

MISSING DATA AND RHYNCHOSAUR PHYLOGENY

MARK WILKINSON¹ and MICHAEL J. BENTON²

¹*Department of Zoology, University of Glasgow Glasgow, G12 8QQ, Scotland, U.K.*

²*Department of Geology, University of Bristol, Bristol, BS8 1RJ, England, U.K.*

(Received August 11, 1994; in final form October 3, 1994)

Consensus methods have been widely applied to the problem of summarising relationships that are common to a set of equally parsimonious trees. Unfortunately, the most commonly used consensus methods frequently fail to represent all such relationships. In phylogenetic palaeontology, inclusion of poorly known fossils with abundant missing data can lead to a plethora of equally parsimonious trees and increased loss of resolution/information in consensus trees. In such cases, recently developed consensus methods, and a method of eliminating taxa, termed safe taxonomic reduction, can be used to circumvent these difficulties. These methods are illustrated through their application to cladistic data for the Rhynchosauria. In particular we investigate the use of disqualifier faithful trees as a framework for representing consensus relationships and describe a method of determining the possible positions of taxa that are excluded from a disqualifier faithful tree.

Keywords: Rhynchosauria, consensus trees, missing data, parsimony, phylogeny.

INTRODUCTION

Missing entries in cladistic data have begun to attract considerable attention. The issues that have been raised in the recent literature cover both practical and conceptual problems concerning the use of missing entries and their treatment and effects in computerised parsimony analysis (Nixon and Davis, 1991; Platnick *et al.*, 1991; Nixon and Wheeler, 1992; Novacek, 1992a, 1992b; Wilkinson, 1992a; Maddison, 1993). In this paper, we focus upon a practical problem that frequently arises when poorly known fossil taxa, with abundant missing entries, are included in parsimony analyses. As has been noted previously (e.g. Gauthier, 1986; Novacek, 1992a; Wilkinson, 1992a) the inclusion of such taxa may result in a dramatic increase in the numbers of equally most parsimonious trees (MPTs). This frequently occurs because the poorly known taxa have insufficient data to place them unambiguously within a framework of relationships that can be established for the better known taxa. Such *underdetermined* taxa can be equally parsimoniously interpreted as joining the framework in a number of places resulting in multiple MPTs.

Faced with multiple MPTs, phylogeneticists usually turn to consensus methods to produce a consensus tree that can serve as a summary of those phylogenetic relationships that are common to all or to some proportion (majority-rule consensus) of the MPTs. Unfortunately, the most widely available and widely used consensus methods can fail to provide a comprehensive summary of these relationships (Swofford, 1991; Wilkinson, 1992a; 1994a). In palaeontological studies, this problem is frequently compounded by the increase in MPTs caused by the inclusion of poorly known, underdetermined taxa, leading in turn to a serious loss of resolution in consensus trees. Indeed, the failings

of these consensus methods can be attributed primarily to the problems of accommodating undetermined taxa (Wilkinson, 1994a).

Where multiple MPTs are specifically associated with poorly known taxa, phylogeneticists have sometimes removed these taxa in order to reduce the number of MPTs and thereby increase resolution in any resulting consensus tree (e.g. Gauthier, 1986; Fraser and Benton, 1989; Benton, 1990). However, this strategy carries with it the danger that the excluded taxa might have provided information of significance for the parsimonious interpretation of relationships among the remaining taxa. It is well known that adding or deleting taxa from a parsimony analysis can have profound effects upon the relationships inferred among other terminal taxa (Arnold, 1981; Gauthier *et al.*, 1988). Ideally, we should only remove taxa from an analysis if we can be sure that their removal will have no effect upon the parsimonious interpretation of relationships of the remaining taxa. To eliminate taxa when this condition is not met may be tantamount to ignoring data.

Thus the problem we shall address has two sides. Firstly, the failures of consensus methods and secondly the multiplicatory effect of underdetermined taxa upon the number of MPTs. The latter component further necessitates the use of consensus methods and exacerbates their failings. Similarly, there are two potential solutions to the problem. Firstly, consensus methods have recently been developed that do not suffer from some of the limitations of older methods (Wilkinson, 1994a, in prep.). These can be used to provide more comprehensive summaries of the relationships that are common to a set of MPTs, and their sensitivity is not negatively correlated with numbers of MPTs. Secondly, problematic taxa can be excluded from analyses subject to the constraint that their exclusion can have no effect upon relationships of the remaining taxa. This constraint leads to a strategy of *safe taxonomic reduction* (Wilkinson, 1992a).

THEORY

Consensus Methods

The strict (component) consensus has been the most widely applied consensus method in phylogenetics. It has an unambiguous interpretation, but frequently fails to represent all phylogenetic relationships that are common to a set of MPTs (Swofford, 1991; Wilkinson, 1994a). Thus, the two trees shown in Figures 1a and 1b both agree that A and B are more closely related to each other than they are to C; symbolically (A,B)C. In contrast, the trees do not agree in any such positive statement about the relationships of D. Ideally, a consensus of these two trees would represent the common relationship (A,B)C, but the strict (component) consensus is an unresolved bush (Figure 1c). The less frequently used Adams consensus method (Adams, 1972) produces a consensus tree (Figure 1d) that might appear to succeed where the strict (component) consensus tree fails. However, (A,B)C is not a legitimate interpretation of the Adams consensus. The only legitimate interpretation of the Adams consensus is that (A,B)C and/or (A,B)D is true of each fundamental tree: the Adams consensus tree would have exactly the same topology in each case so that its polytomies are highly ambiguous (Rohlf, 1982; Wilkinson, 1994a).

The recently developed reduced cladistic consensus (RCC) method has been designed to provide a solution to the problems of identifying and graphically representing all positive statements of relationships that are common to a set of MPTs (Wilkinson, 1994a). It avoids problems associated with underdetermined taxa by excluding them from consensus trees if, by this exclusion, additional relationships among the remaining taxa can be specified unambiguously. This method may produce a single RCC tree (e.g. Figure

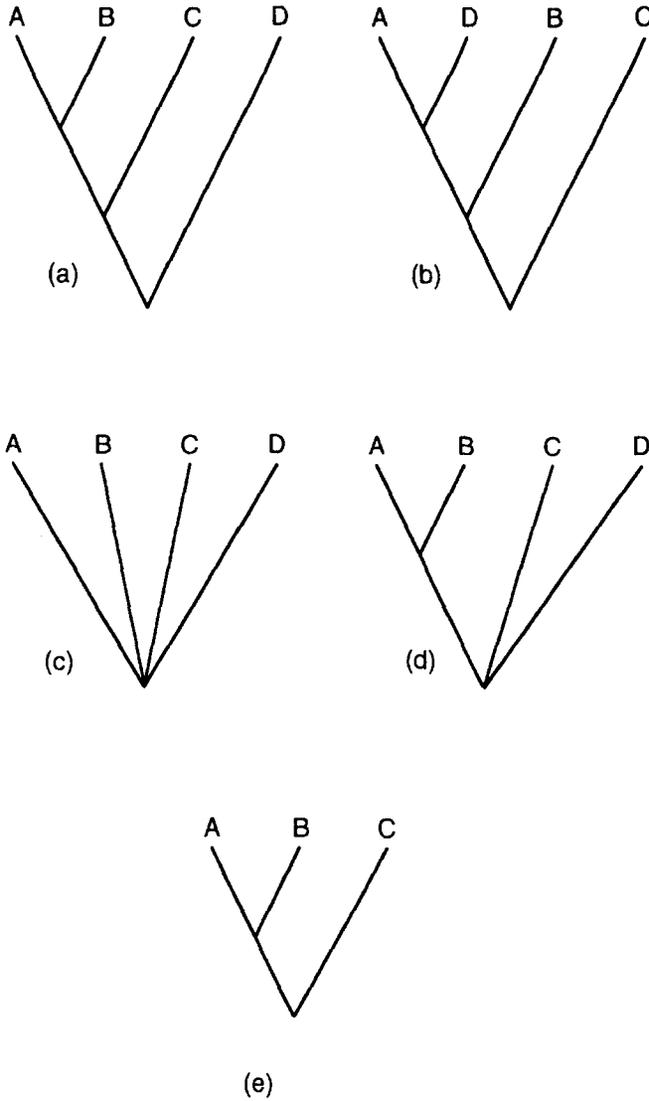


Figure 1 Two trees (a and b) and the corresponding strict (component) consensus tree (c), Adams consensus tree(d), and RCC tree (e).

le) or a number of RCC trees that together comprise an RCC profile. In addition, RCC profiles can be divided into a set of basic RCC trees and derivative forms that can be constructed by combining aspects of different basic RCC trees. The set of basic RCC trees represent an efficient means of representing all common positive statements of relationships, but derivative trees may have the desirable property of being more highly resolved than any single basic RCC tree. If the strict (component) consensus is informative (i.e. not completely unresolved) it will also be included among the basic RCC trees (i.e. it is subsumed within the RCC concept). Other trees in the RCC profile gain additional resolution at the cost of excluding underdetermined taxa.

A *disqualifier* is a negative statement of relationships such as '*C and D are not more closely related to each other than they are to A or B*': symbolically $\text{not}(C,D)A,B$, (Wilkinson, 1994a). This disqualifier is common to the trees in Figures 1a and 1b, but is not represented in any of the consensus trees (Figures 1c, 1d and 1e). Ideally, consensus summaries of MPTs would identify and represent all common disqualifiers as well as all common positive statements of relationships. Disqualifier faithful (DF) trees have the useful properties of (1) including all disqualifiers that are common to the included taxa (adopting the soft interpretation of polytomies) and (2) not implying any disqualifiers that are not common to the MPTs (Wilkinson, 1994a; in prep). Given these properties, DF trees can be used as a framework for the unambiguous representation of additional disqualifiers that describe restrictions on the *possible positions* of taxa that are not included in the framework. Thus of the consensus trees in Figure 1, only the RCC tree (Figure 1e) is disqualifier faithful, and it can be used to provide a description of the possible position of the excluded taxon D. We further explore this use of DF trees and the determination of possible positions in this paper.

Safe Taxonomic Reduction

In order for the inclusion or exclusion of a terminal taxon to have any effect upon inferred relationships of other taxa it must have a combination of character states that are not seen in other included taxa. However, if the character states of one taxon are fully replicated in another, then the former can be safely eliminated from the analysis without any danger of affecting inferences of the relationships among the remaining taxa. The only effect of this strategy, termed safe taxonomic reduction, should be a reduction in the number of MPTs. Note that this implies that the simple proportion of missing entries associated with a terminal taxon does not provide a secure basis for eliminating taxa. Rather, safe taxonomic reduction requires comparisons of the combinations of character states seen in taxa that are candidates for deletion because of their abundant missing entries. If taxa have unique combinations of character states for those characters that are not coded as missing then they should be retained in the analysis (Wilkinson, 1992a).

METHODS

All parsimony analyses were performed using PAUP 3.1.1 (Swofford, 1993). Heuristic searches with TBR branch swapping and 10 random addition sequences were used in the analysis of the complete data set. Other searches used the branch and bound method. Multistate characters were treated as unordered. Full explanations of the character data are given in Benton (1990). The TAXEQ program (Wilkinson, 1992b) was used to identify taxa that could be eliminated in safe taxonomic reduction, and REDCON (Wilkinson, 1994b) was used to identify all basic RCC trees in the RCC profile. The derivative RCC tree was constructed manually by the *fusion* of basic RCC trees (Wilkinson, in prep). The DF tree was constructed by identifying undetermined taxa through manual inspection of pruned MPTs and removing these from consensus trees. Similarly, possible positions of excluded taxa were determined by manual inspection of pruned MPTs (see below for further details).

RHYNCHOSAUR PHYLOGENY

Benton's (1990) cladistic data for the Rhynchosauria comprises 26 characters, 16 ingroup terminal taxa, and an additional hypothetical ancestral taxon based on other early diapsids and used for rooting (Table 1). The extent of missing data in the matrix is high (34% of all data points) and Benton considered that only nine of the terminal taxa were sufficiently well-known to be included in his parsimony analysis. The excluded taxa were those that were coded as missing for more than 33% of the characters. Analysis of this reduced data set yields three MPTs (Figure 2a-c) with a length (L) of 34 steps (originally reported as 35) a consistency index (CI) of 0.853 (originally reported as 0.857) and a retention index (RI) of 0.889 which varied only in the placements of the two included species of *Rhynchosaurus*. Benton represented his phylogenetic conclusions with a core cladogram (Figure 2d, his figure 43), equivalent to a strict (component) consensus of the three MPTs. In addition, the excluded less well known taxa were attached to the core cladogram with dashed lines based on inspection of their character states.

Our PAUP analysis of the entire data set, including all terminal taxa, produced a staggering 21,700 MPTs (L=35, CI=0.829, RI=0.891). There were very likely more MPTs but the search was limited by available memory. The strict (component) consensus of these MPTs (Figure 4a) is almost completely unresolved, and comprises a polytomy grouping all taxa, with the exception of *Mesosuchus* which is recovered as the sister group to all other rhynchosaurians. The lack of resolution in the strict (component) consensus compared with Benton's (1990) tree illustrates the potentially obfuscatory effect of missing entries when they lead to numerous MPTs. Given this effect, Benton's strategy of deleting the most poorly known taxa appears to be a highly effective way of avoiding these problems. However, as we shall see, application of safe taxonomic reduction or use of the RCC method leads to slightly different conclusions. Note also that the trees based on all the taxa require a single additional step. This implies that the poorly known taxa that were eliminated by Benton (1990) are responsible for some additional homoplasy that could potentially alter relationships among the better known taxa.

Table 1 Character data for rhynchosaurian species (from Benton, 1990). Letters in parentheses are abbreviations used in Figure 2d.

	Characters					
	1-5	6-10	11-15	16-20	21-25	26
Hypothetical ancestor	00000	00000	00000	00000	00000	0
<i>Supradapedon</i> (S)	?????	?????	?????	11110	?????	?
<i>Hyperodapedon gordonii</i>	11111	01111	10111	11100	11111	1
<i>H. huxleyi</i>	11111	01111	10111	11110	11111	1
'Nova Scotia rhynchosaur' (N)	?11??	?????	?????	11110	01???	?
'Texas rhynchosaur' (T)	?11??	?????	?????	?????	?????	?
<i>Scaphonyx fischeri</i>	11111	01111	10111	11110	01101	1
<i>S. sanjuanensis</i>	11110	??11?	10011	11110	01101	1
<i>Isalorhynchus</i> (I)	?11??	?????	?????	11100	?????	?
<i>Acrodenta</i> (A)	?????	?????	?????	1???	?????	?
<i>Rhynchosaurus articeps</i>	01110	01110	01011	11201	21001	1
<i>R. brodiei</i>	01111	0111?	?011	11201	21???	?
<i>R. spenceri</i> (R.s.)	11111	?????	00011	11201	31???	?
<i>Mesodapedon</i> (M)	?????	?????	?????	11201	2???	?
<i>Stenaulorhynchus</i>	01110	11110	01001	11201	31001	1
<i>Howesia</i>	0???	01110	?001	10001	30001	0
<i>Mesosuchus</i>	01010	00000	01000	00000	00001	0

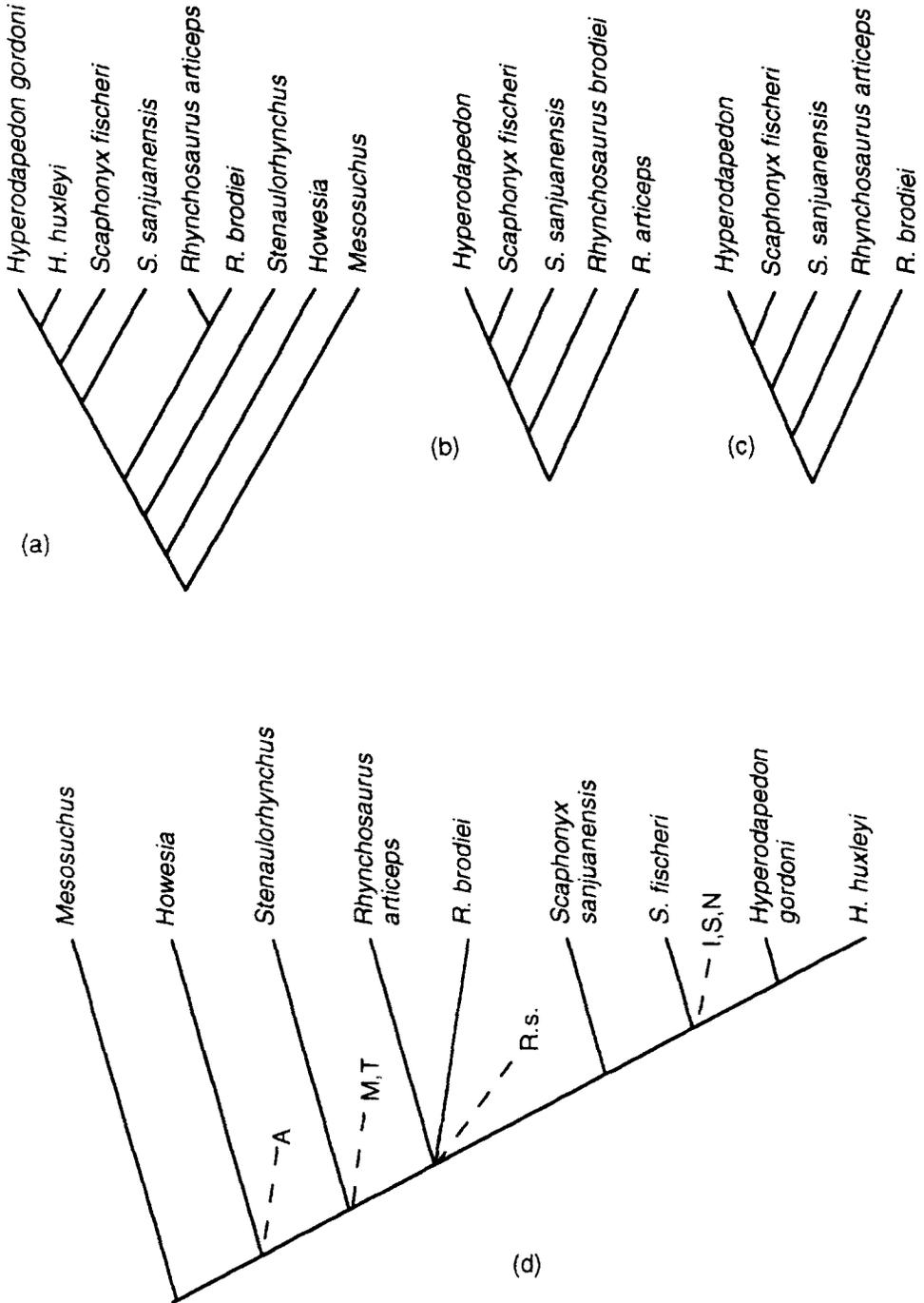


Figure 2 An MPT (a) and two equally parsimonious topological variants (b-c) found in Benton's (1990) study, and his summary diagram (d). Abbreviations as in Table I.

Six terminal taxa (*Acrodonta*, *Isalorhynchus*, *Supradapedon*, *Mesodapedon*, the 'Nova Scotia rhynchosaur' and the 'Texas rhynchosaur') satisfy the criteria for safe taxonomic reduction (Table 2). These taxa were also eliminated by Benton (1990) on the basis of the extent of their missing entries. However, a single taxon that was excluded from Benton's analysis does not satisfy the criteria for safe taxonomic reduction. Despite a high proportion of missing entries (34.6%), *Rhynchosaurus spenceri* has a combination of character states that are not seen in any other taxon and cannot be eliminated from the analysis without some danger of its affecting relationships among the remaining taxa. Analysis of the data set after safe taxonomic reduction yielded two MPTs (Figure 3a and 3b) ($L=35$; $CI=0.829$, $RI=0.891$). The unique combination of character states of *R. spenceri* is sufficient to provide an unambiguous most parsimonious interpretation of its relationships to the other included taxa. Based on the parsimonious interpretation of the character data *R. spenceri* is more closely related to the *Scaphonyx-Hyperodapedon* clade than to the other species of *Rhynchosaurus*, and the genus *Rhynchosaurus* is accordingly paraphyletic. This conclusion rests on a single character (character 1) that describes variation in the overall shape (length and breadth) of the skull. In addition, the introduction of *R. spenceri* decreases the number of MPTs from three to two by placing additional limits on the possible interrelationships of the other two species of *Rhynchosaurus*.

The variation in the positions of *R. brodiei* and *R. articeps* in the two MPTs is due to character incongruence rather than to missing data. Character 5 supports a closer relationship of *R. brodiei* to the *R. spenceri-Scaphonyx-Hyperodapedon* clade than to *R. articeps*, with a reversal in *S. sanjuanensis*. Character 21 supports a sister-group relationship between *R. brodiei* and *R. articeps*. These two trees share a common disclaimer: that *R. articeps* is not more closely related to the *R. spenceri-Scaphonyx-Hyperodapedon* clade than it is to *R. brodiei*. In Benton's original analysis, the possible positions and interrelationships of *R. brodiei* and *R. articeps* are less constrained because the character states of *R. spenceri* are not taken into account when optimising these two characters. As a result character 5 can be equally parsimoniously interpreted as

Table 2 Levels of missing data and scope for safe taxonomic reduction within the rhynchosaurian data of Table I. Taxa that have their character states subsumed within those of other taxa can be safely eliminated.

Taxon	%Missing data	Other taxa that include the same combination of character states
1 <i>Supradapedon</i>	76.9	3,4,6,7
2 <i>Hyperodapedon gordonii</i>	0	-
3 <i>H. huxleyi</i>	0	-
4 'Nova Scotia rhynchosaur'	61.5	6,7
5 'Texas rhynchosaur'	92.3	2-4, 6-8, 10-12, 14
6 <i>Scaphonyx fischeri</i>	0	-
7 <i>S. sanjuanensis</i>	19.2	-
8 <i>Isalorhynchus</i>	69.2	2
9 <i>Acrodonta</i>	92.3	1-4, 6-8, 10-15
10 <i>Rhynchosaurus articeps</i>	0	-
11 <i>R. brodiei</i>	26.9	-
12 <i>R. spenceri</i>	34.6	-
13 <i>Mesodapedon</i>	73.1	10,11
14 <i>Stenaulorhynchus</i>	0	-
15 <i>Howesia</i>	19.2	-
16 <i>Mesosuchus</i>	0	-

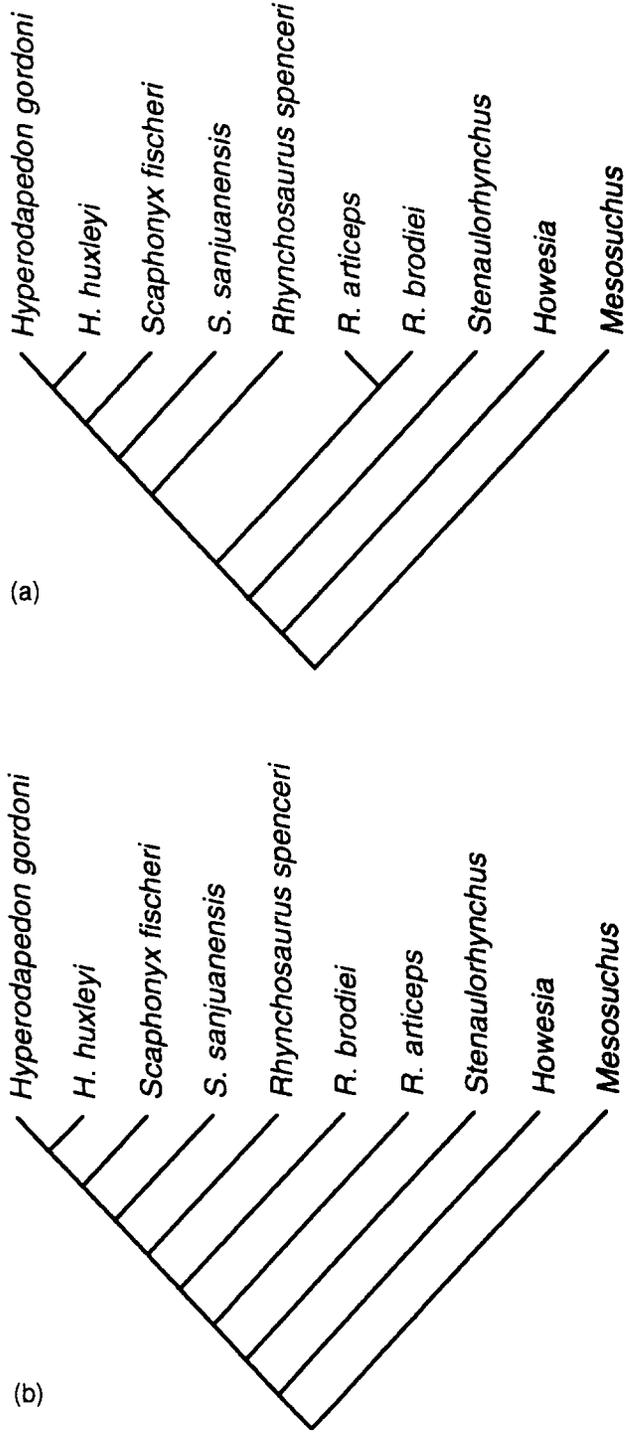


Figure 3 Two MPTs found after safe taxonomic reduction.

independently derived in *R. brodiei* and within the *Scaphonyx-Hyperodapedon* clade, with no reversal in *S. sanjuanensis* (Benton, 1990). Similarly, state 2 of character 21, which unites *R. brodiei* and *R. articeps* when *R. spenceri* is included, can be equally parsimoniously interpreted as plesiomorphic when *R. spenceri* is ignored.

We also constructed the RCC profile directly from the 21,700 MPTs found when all the terminal taxa are included in the analysis. There are five basic RCC trees in the profile (Figure 4), including the strict (component) consensus (Figure 4a). The other

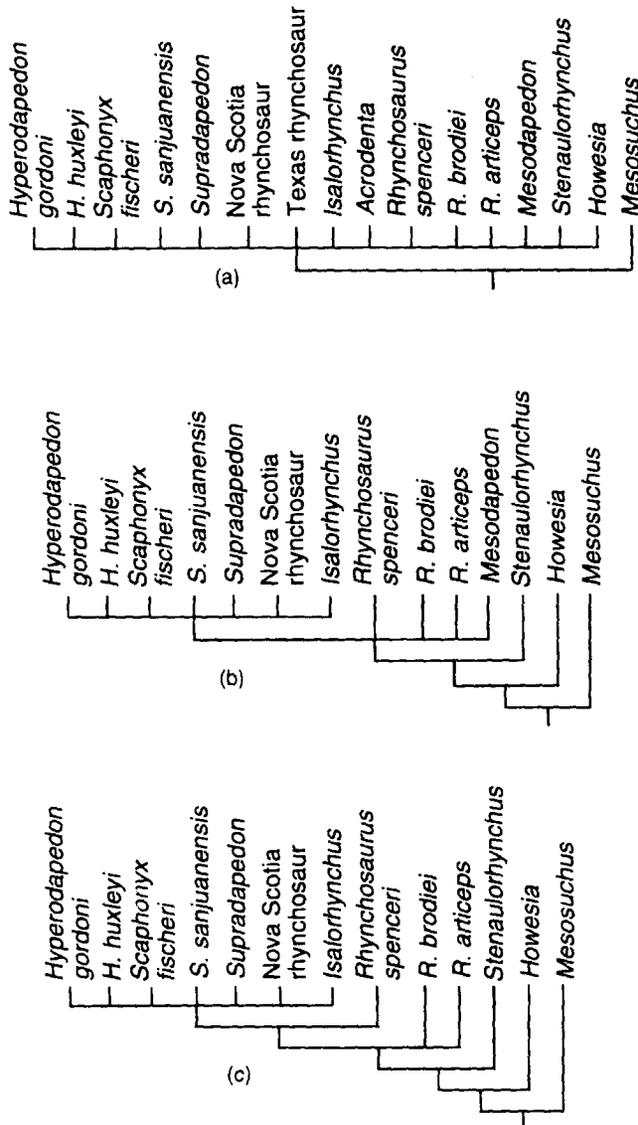
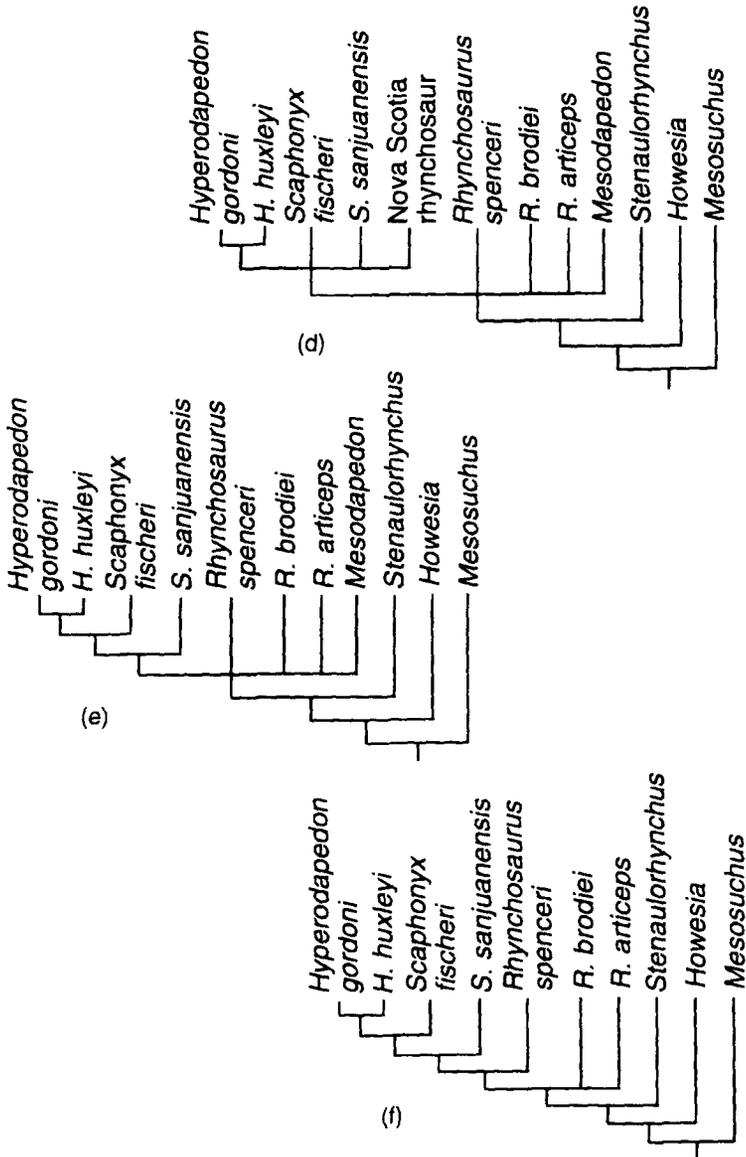


Figure 4 The set of five basic RCC trees for the 21,700 MPTs produced by the analysis of the whole data set (a-e) and a single more fully resolved derivative RCC tree (f). The first basic RCC tree (a) is also the strict (component) consensus.

four basic RCC trees are more fully resolved than the strict (component) consensus and achieve this additional resolution at the expense of excluding underdetermined taxa. In each case the excluded taxa could not be attached to the RCC tree at any position without producing either an erroneous interpretation of their relationships or an ambiguous polytomy that would also be open to incorrect interpretation. The deletion of the 'Texas rhynchosaur' and *Acrodonta* triples the number of nodes in the corresponding RCC tree (Figure 4b). Deletion of *Mesodapedon* provides a single additional node that further resolves the relationships of the *Rhynchosaurus* species (Figure 4c), whereas deletion of *Supradapedon* and *Isalorhynchus* and then the 'Nova Scotia rhynchosaur' add additional nodes within the *Hyperodapedon-Scaphonyx* clade (Figure 4d and 4e).



The RCC trees should not be construed as consensus trees for MPTs produced after the deletion of taxa that are not included in the tree. Rather, they represent summaries of the relationships of these taxa that are common to the 21,700 MPTs produced when all the taxa are included in the analysis. Taken as a whole, the set of five basic RCC trees include all positive statements of relationships that are common to the 21,700 MPTs. Two derivative RCC trees can be produced by the fusion of two pairs of basic RCC trees (c and d, d and e). Fusion involves further deleting taxa that do not feature in *both* the contributing RCC trees and including all nodes that represent relationships among the remaining taxa that occur in *either* of the RCC trees. The derivative RCC formed by the fusion of basic RCC trees c and e (Figure 4f) is more fully resolved than any other RCC tree and provides a good starting point for the construction of a DF tree. Again it achieves this greater resolution at the expense of the exclusion of additional underdetermined taxa.

Note that the derivative RCC tree (Figure 4f) is also the strict (component) consensus produced from the two MPTs after safe taxonomic reduction, and that *R. brodiei* and *R. articeps*, are collapsed to a polytomy. Given a soft interpretation, this polytomy implies no common disqualifiers, which as we have seen is incorrect. Thus the primary RCC tree is not a DF tree. It can be converted into a DF tree by the removal of either *R. articeps* or *R. brodiei*, the choice being arbitrary. Here we have excluded *R. articeps* and provided the nodes of the resulting DF tree with numbers (Figure 5). Thus prepared, the DF tree can be used as a basis for exploring common disqualifiers for taxa that are excluded from the DF tree and how they limit their possible positions relative to the DF tree.

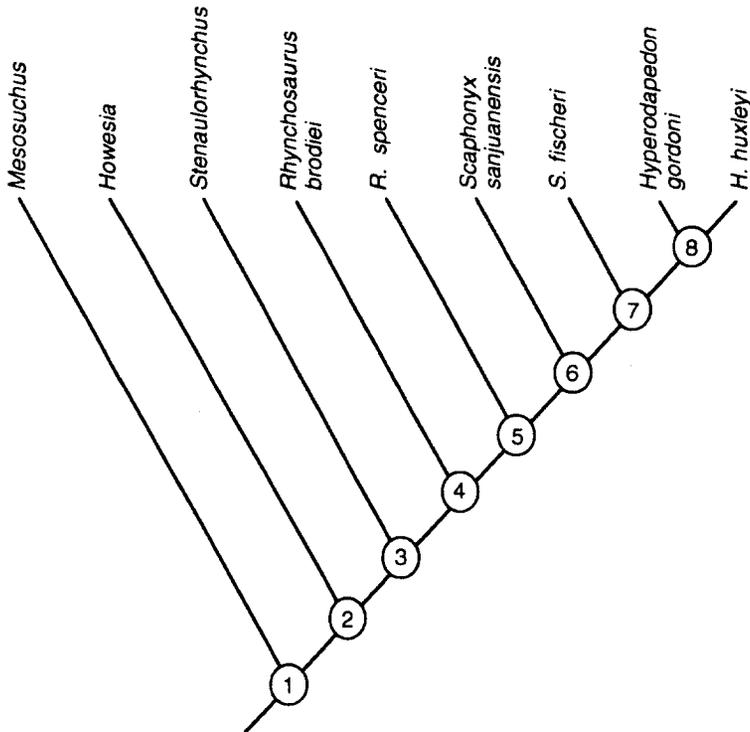


Figure 5 Disqualifier faithful tree used as a framework for summarising relationships (including disqualifiers) that are common to the 21,700 MPTs.

To determine possible positions of excluded taxa relative to the DF tree we exploited facilities in PAUP for pruning taxa from trees in memory. This involved loading the 21,700 MPTs into PAUP, and then instructing PAUP to delete all but one of the underdetermined taxa that are not included in the DF tree (repeating this operation for each of the underdetermined taxa). When taxa are deleted and there are trees in memory, PAUP provides the option of pruning the deleted taxa from the trees and discarding any resulting duplicate trees (ones that have been rendered identical to others by the pruning). In each case, this operation produced a sufficiently dramatic reduction in the number of distinct trees (Table III) to render individual inspection of each practicable. Manual inspection of the pruned MPTs allowed us to determine the possible positions of the single underdetermined taxon relative to the constant framework provided by the DF tree which was replicated across all the pruned trees. The resulting list of possible positions of underdetermined taxa relative to the DF tree is given in Table 3.

Table 3 List of possible positions of underdetermined taxa relative to the numbered nodes of the disqualifier faithful tree of Figure 5 and the numbers of topologically distinct pruned MPTs which were examined to determine these possible positions after the pruning of all other underdetermined taxa. A = above (i.e. on any branch above the node), B = below (i.e. on any branch below the node), - = between (i.e. on any branch on the direct line between the nodes).

<i>Underdetermined taxa</i>	<i>Possible positions</i>	<i>Numbers of pruned MPTs</i>
<i>Supradapedon</i>	5-6, A6+B8	6
'Nova Scotia rhynchosaur'	5-6, A6+B8	5
'Texas rhynchosaur'	1-2, A2	17
<i>Isalorhynchus</i>	5-6, 8-H. <i>gordoni</i>	2
<i>Acrodonta</i>	1-3, A3	16
<i>Mesodapedon</i>	3-6, 5-R. <i>spenceri</i>	5
<i>Rhynchosaurus articeps</i>	3-4, 4-R. <i>brodiei</i>	2

DISCUSSION

Phylogenetic inference advances both through the accumulation and analysis of new data and through the development of new analytical techniques. The problems posed by poorly known taxa with abundant missing entries in the context of parsimony analysis have been widely appreciated. However, attempts to ameliorate this problem by eliminating the most poorly known taxa have been rather *ad hoc* and, given the potential for eliminating/adding taxa to effect the parsimonious interpretation of relationships among other taxa, must be viewed with some caution. Safe taxonomic reduction (Wilkinson, 1992a) is based on the simple fact that taxa that do not have a unique combination of character states add nothing to an analysis that is not already represented by other taxa that have the same combination of character states. Such taxa can be safely eliminated (i.e. without affecting the relationships inferred among the remaining taxa) allowing phylogeneticists to reduce the numbers of MPTs under the constraint of not selectively ignoring pertinent data.

Multiple MPTs may result from either contradictory evidence of relationships (homoplasy) or because of insufficient data (missing data). Safe taxonomic reduction can only be used to reduce the numbers of MPTs that result from a lack of data. If character data is highly homoplastic there may still be large numbers of MPTs. In such

cases, the deficiencies of consensus methods such as the strict (component) consensus and the Adams consensus cannot be avoided by the simple strategy of safe taxonomic reduction. The RCC method provides a solution to the problem of graphically representing all positive statements of relationship that are common to a set of MPTs, and as such it can be used both when multiple MPTs are the result of homoplasy or missing data. Comparing the strict (component) consensus (Figure 4a) for the 21,700 MPTs with complete set of basic RCC trees of which it is a part (Figure 4) demonstrates the degree to which relationships may be obscured by the strict (component) consensus method.

Extending this new consensus method to produce a DF tree and a list of the possible positions of underdetermined taxa allows disqualifiers, an additional class of negative statements of relationship, to be investigated and represented in a consensus of multiple MPTs. In our reanalysis of the rhynchosaur data, the combination of DF tree and table of possible positions of excluded taxa achieves a high degree of sensitivity to relationships that are supported by the data and are common to all the 21,700 MPTs. What is missing from this summary are any disqualifiers that pertain to the interrelationships of underdetermined taxa rather than to relationships of these taxa to the taxa included in the DF tree, and any information provided by additional MPTs that were not examined because of constraints on available memory.

Although the use of novel consensus methods provides a more comprehensive solution to the problems of multiple MPTs, this does not render safe taxonomic reduction redundant. Safe taxonomic reduction may have the practical benefit of reducing the number of taxa sufficiently to allow the application of exact rather than heuristic tree searches (Wilkinson, 1992a). In addition, applying both techniques, phylogeneticists will be better able to assess whether unstable taxa are underdetermined because of missing data, homoplasy or a combination of both.

Acknowledgements

This work was supported in part by SERC grant GR/F 87912, and by a grant from the University of Glasgow New Initiatives Fund. TAXEQ and REDCON, both PC-compatible software, are available from MW upon receipt of a disc.

References

- Adams, E.N. (1972) Consensus techniques and the comparison of taxonomic trees. *Systematic Zoology*, **21**, 390–397.
- Arnold, E.N. (1981) Estimating phylogenies at low taxonomic levels. *Zeitschrift für zoologisches und Systematische Evolutiones-forschung*, **19**, 1–35.
- 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, Series B*, **328**, 213–306.
- Fraser, N.C. and M.J. Benton, M.J. (1989). The Triassic reptiles *Brachyrhinodon* and *Polysphenodon* and the relationships of the sphenodontids. *Zoological Journal of the Linnean Society*, **96**, 413–445.
- Gauthier, J. (1986) Saurischian monophyly and the origin of birds. *Memoires of the California Academy of Sciences*, **8**, 1–47.
- Gauthier, J., Kluge, A.G. and Rowe, T. (1988) Amniote phylogeny and the importance of fossils. *Cladistics*, **4**, 105–209.
- Maddison, W.P. (1993) Missing data versus missing characters in phylogenetic analysis. *Systematic Biology*, **42**, 576–581.

- Nixon, K.C. and Davis, J.I. (1991) Polymorphic taxa, missing values and cladistic analysis. *Cladistics*, **7**, 233–241.
- Nixon, K.C. and Wheeler, Q.D. (1992) Extinction and the origin of species. In *Extinction and Phylogeny*, edited by M.J. Novacek and Q.D. Wheeler, pp. 119–143. New York: Columbia University Press.
- Novacek, M.J. (1992a) Fossils as critical data for phylogeny. Pages 46–88 In *Extinction and phylogeny*, edited by M.J. Novacek, M. J. and Q.D. Wheeler, pp. 46–88. New York: Columbia University Press.
- Novacek, M.J. (1992b) Fossils, topologies, missing data, and the higher level phylogeny of eutherian mammals. *Systematic Biology*, **41**, 58–73.
- Platnick, N.I., Griswold, C.E. and Coddington, J.A. (1991) On missing entries in cladistic analysis. *Cladistics*, **7**, 337–343.
- Rohlf, F.J. (1982). Consensus indeces for comparing classifications. *Mathematical Biosciences*, **59**, 131–144.
- Swofford, D.L. (1991) When are phylogeny estimates from molecular and morphological data incongruent? In *Phylogenetic Analysis of DNA Sequences*, edited by M.M. Miyamoto and J. Cracraft, pp. 295–333. Oxford: Oxford University Press.
- Swofford, D.L. (1993) PAUP: phylogenetic analysis using parsimony, version 3.1.1, Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.
- Wilkinson, M. (1992a) Consensus, compatibility and missing data in phylogenetic inference. Unpublished PhD thesis. University of Bristol.
- Wilkinson, M. (1992b) TAXEQ - safe taxonomic reduction, software and documentation. Department of Geology, University of Bristol
- 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) REDCON reduced consensus program, software and documentation. Department of Zoology, University of Glasgow.
- Wilkinson, M. Reduced consensus trees: additions and corrections. in prep.