



OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

Fragmentary Taxa, Missing Data, and Ambiguity: Mistaken Assumptions and Conclusions

Author(s): Maureen Kearney

Source: *Systematic Biology*, Apr., 2002, Vol. 51, No. 2 (Apr., 2002), pp. 369-381

Published by: Oxford University Press for the Society of Systematic Biologists

Stable URL: <https://www.jstor.org/stable/3070797>

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/3070797?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



and Oxford University Press are collaborating with JSTOR to digitize, preserve and extend access to *Systematic Biology*

JSTOR

- K. Bremer, and H. Jornvall, eds.). Elsevier Science Publishers, Biomedical Division, New York.
- GAUTHIER, J., A. G. KLUGE, AND T. ROWE. 1988. The early evolution of the Amniota. Pages 103–155 in *The phylogeny and classification of the tetrapods, volume 1: amphibians, reptiles, birds* (M. J. Benton, ed.). Clarendon Press, Oxford, U.K.
- HAY, J. M., I. RUVINSKY, S. B. HEDGES, AND L. R. MAXSON. 1995. Phylogenetic relationships of amphibian families inferred from DNA sequences of mitochondrial 12S and 16S ribosomal RNA genes. *Mol. Biol. Evol.* 12:928–937.
- HEDGES, S. B., AND L. R. MAXSON. 1993. A molecular perspective on lissamphibian phylogeny. *Herpetol. Monogr.* 7:27–42.
- HEDGES, S. B., K. D. MOBERG, AND L. R. MAXSON. 1990. Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Mol. Biol. Evol.* 7:607–633.
- LAURIN, M. 1998. The importance of global parsimony and historical bias in understanding tetrapod evolution. Part I—Systematics, middle ear evolution, and jaw suspension. *Ann. Sci. Nat. Zool.* (13 Ser.) 19:1–42.
- LAURIN, M., AND R. R. REISZ. 1997. A new perspective on tetrapod phylogeny. Pages 9–59 in *Amniote origins—Completing the transition to land* (S. Sumida and K. Martin, eds.). Academic Press, London.
- LAURIN, M., AND R. R. REISZ. 1999. A new study of *Solenodonsaurus janenschii*, and a reconsideration of amniote origins and stegocephalian evolution. *Can. J. Earth Sci.* 36:1239–1255.
- LOMBARD, R. E., AND S. S. SUMIDA. 1992. Recent progress in understanding early tetrapods. *Am. Zool.* 32:609–622.
- MILNER, A. R. 1993. The Paleozoic relatives of lissamphibians. *Herpetol. Monogr.* 7:8–27.
- PANCHEN, A. L., AND T. R. SMITHSON. 1988. The relationships of the earliest tetrapods. Pages 1–32 in *The phylogeny and classification of the tetrapods, volume 1: Amphibians, reptiles, birds* (M. J. Benton, ed.). Clarendon Press, Oxford, U.K.
- PATON, R. L., T. R. SMITHSON, AND J. A. CLACK. 1999. An amniote-like skeleton from the Early Carboniferous of Scotland. *Nature* 398:508–513.
- TRUEB, L., AND R. CLOUTIER. 1991. A phylogenetic investigation of the inter- and intrarelationships of the Lissamphibia (Amphibia: Temnospondyli). Pages 223–313 in *Origins of the higher groups of tetrapods—Controversy and consensus* (H.-P. Schultze and L. Trueb, eds.). Cornell Univ. Press, Ithaca, NY.
- ZARDOYA, R., AND A. MEYER. 2001. On the origin of and phylogenetic relationships among living amphibians. *Proc. Natl. Acad. Sci. USA* 98:7380–7383.

Received 27 June 2001; accepted 11 November 2001

Associate Editor: Jack Sites

Syst. Biol. 51(2):369–381, 2002

Fragmentary Taxa, Missing Data, and Ambiguity: Mistaken Assumptions and Conclusions

MAUREEN KEARNEY

Field Museum of Natural History, Department of Zoology, Division of Amphibians and Reptiles, Chicago, Illinois 60605, USA; E-mail: mkearney@fmnh.org

It is unclear under what circumstances incomplete taxa will be problematic in phylogenetic analyses, and whether such potential problems are outweighed by the benefits of their inclusion. In Gauthier et al. (1988), the potential importance and uniqueness of fossil taxa despite their fragmentary nature was emphasized, and a total evidence approach (Kluge, 1989; Eernisse and Kluge, 1993) was advocated (see also Doyle and Donoghue, 1987). Later papers explicated analytical problems associated with missing data from incomplete fossil taxa in combined analyses (e.g., Crepet and Nixon, 1989a,b; Nixon and Wheeler, 1992; Novacek, 1992; Wilkinson, 1995a; Wiens, 1998). Many systematists are now reluctant to include fragmentary taxa because of perceived

problems associated with missing data, especially the relationship between missing data and ambiguity of results. Recently, some have suggested that, because of missing data associated with fossils, the positive effects of total evidence versus the potentially negative effects of missing data must be weighed (e.g., Nixon, 1996; Grande and Bemis, 1998; O'Leary, 2000).

Despite all the papers mentioned above, and despite the now popular buzzwords "missing data," it remains to be clarified whether commonly made generalizations about missing data hold true. Certainly, the assertion that adding incomplete taxa *necessarily* increases the number of primary trees or overall ambiguity is simplistic because many studies belie this assumption (e.g.,

TABLE 1. Comparison of levels of ambiguity in some recent studies with various degrees of missing data and fossil taxa. Studies are listed in increasing order of percentage of missing data.

Study	Percentage fossil taxa	Percentage missing data	Number of primary trees	Number of resolved nodes/ number of terminal taxa
Novacek (1992)	26	13	>6,800 ^a	16/28 = 57%
Messenger and McGuire (1998)	2	14	>45,000 ^a	33/56 = 59%
Wu et al. (1996)	25	19	2	25/28 = 89%
Grande and Bemis (1998)	98	20	>10,000 ^a	19/44 = 43%
Fraser and Benton (1989)	100	21	82	9/15 = 60%
Gao and Norell (1998)	58	34	>32,000 ^a	13/33 = 39%
Norell and Gao (1997)	80	49	395	8/20 = 40%
Gatesy et al. (1999)	0	57	6	72/79 = 91%
O'Leary (1999)	75	73	30	22/37 = 59%

^aSearch stopped at set limit of equivalent trees, so more primary trees actually exist.

Norell and de Queiroz, 1991). Table 1 suggests there may not be a strict correlation between percentage of fossil taxa and numbers of primary trees or between percentage of missing data and numbers of primary trees. Instead, as suggested by Novacek (1992), the effects of incomplete taxa and concomitant missing character data are not general, but matrix-specific, and depend on the precise distribution of question marks and character states across taxa.

Some researchers have focused on developing or advocating alternative consensus methods as a solution to the “missing data problem” (e.g., Swofford, 1991; Wilkinson, 1994, 1995b; Wilkinson and Benton, 1995). These methods attempt to preserve consensus resolution by identifying and removing ambiguously resolved taxa. Such methods are somewhat compelling, in that exclusion of incomplete taxa a priori is not required; instead, attempts to resolve ambiguity are made subsequent to cladogram construction. However, these methods still neglect an important consideration, that is, that instability of taxa may be the result of missing data, conflicting data, or both.

I argue here that the core of the problem lies in operationally treating all ambiguity equally, a practice that can misrepresent the character evidence. Also, one need not necessarily exclude incomplete taxa or characters to solve the missing data problem. Instead, all relevant characters and taxa (including their associated question marks) can be combined at the outset and, if results are ambiguous, appropriate methods can be used after analysis, such as constructing summary trees based on character support. In an attempt to assess the generality of the problem, I have reexamined several recent anal-

yses that combine living and fossil taxa and that have broached the missing data problem in some manner.

AMBIGUOUS AMBIGUITY

Nixon and Wheeler (1992) coined the term wildcard to describe an incompletely known taxon that floats into many different positions on a cladogram as the result of alternative optimizations of question marks by computer algorithms. In the case of a taxon with a large amount of missing data, many primary trees may be produced (corresponding to all the possible placements of the wildcard), and a strict consensus tree may be poorly resolved (Fig. 1). Such wildcard behavior is not necessarily restricted to fossils but may occur whenever missing data are concentrated in a single taxon for any reason, such as when data sets are combined and one subset of character data is unknown for a taxon (Carpenter, 1987).

Some researchers have cited this “wildcard problem” as a reason to exclude incomplete taxa or characters that add missing data to a matrix, often using exclusion strategies for taxa based on a cutoff percentage of missing character data (e.g., Rowe, 1988; Benton, 1990; Wu et al., 1996; Grande and Bemis, 1998; Anderson, 2001). But fragmentary taxa do not always behave as wildcards, and those that do may do so for different reasons (Fig. 2). Taxa may be unstable because of missing data, character conflict, or both (as in Nixon and Wheeler’s example), but many recent studies assume a simple causal connection between missing data associated with fragmentary taxa and ambiguity in results.

Wilkinson (1994, 1995a) categorized both sources of ambiguity as underdetermination

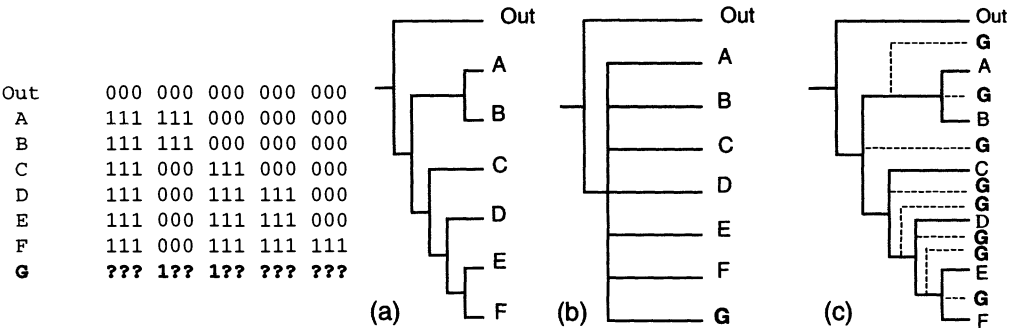


FIGURE 1. Effect of inclusion or exclusion of a wildcard taxon (redrawn from Nixon and Wheeler, 1992). (a) Single tree resulting from analysis of the data set excluding the wildcard taxon G. (b) Strict consensus of eight trees resulting from analysis of the data set including the wildcard taxon G. (c) The eight possible positions of G are indicated with dashed lines, a result of analysis of the data set with the program HENNIG86 (Farris, 1988) or PAUP (Swofford, 1993). If analyzed in NONA (Goloboff, 1993), four of those eight trees are found due to the algorithm's different approach to ambiguous character optimizations. The strict consensus of those four is still completely unresolved, however, and the basic wildcard problem remains.

and introduced a new method, reduced consensus, for dealing with both: "... we could develop alternative consensus methods that are not subject to the general difficulties of those presently in widespread use.

Such a solution would apply equally well to the limitations of these consensus methods in cases when numerous MPTs result from positive rather than negative underdetermination, homoplasy or missing entries" (Wilkinson, 1995a:504). Anderson (2001) devised a similar method that treats all ambiguity equally. In contrast, I suggest here that treating all ambiguity equally may be empirically misleading because the two sources of ambiguity reflect observational differences: In one case, the uncertainty of a taxon's relationship reflects a lack of character evidence and, in the other case, it is the result of conflicting character evidence.

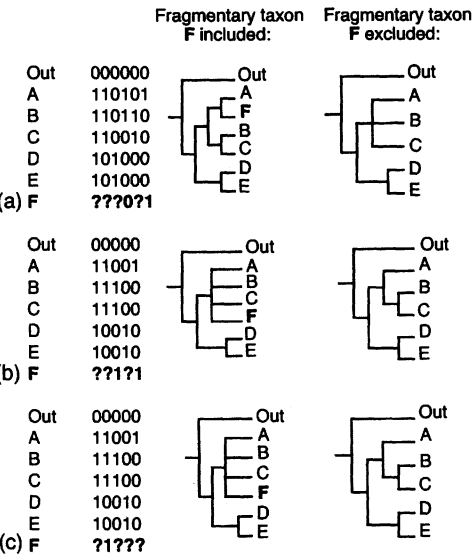


FIGURE 2. Whether incomplete taxa will increase ambiguity or not is unpredictable based on the amount of missing data before analysis. (a) Inclusion of fragmentary taxon F increases resolution. (b) Inclusion of fragmentary taxon F decreases resolution. Ambiguity is caused by character conflict among taxa A, B, C, and F. (c) Inclusion of fragmentary taxon F decreases resolution. F contains sufficient information to place it in the group A(BC), but missing data cause it to behave as a wildcard within that clade. The supported (BC) clade is obscured by the wildcard effect of taxon F.

PREVIOUSLY PROPOSED SOLUTIONS
AND THEIR LIMITATIONS
Safe Taxonomic Reduction

Wilkinson (1992, 1995a) recognized the drawbacks of removing fragmentary taxa from analyses based on percentage of missing data and proposed instead a strategy of Safe Taxonomic Reduction (STR). STR, which is implemented in the TAXEQ3 computer program (Wilkinson, 2001), identifies taxonomic equivalents in a data matrix and deletes them. Taxonomic equivalents are taxa that contribute missing data and also completely overlap in character states with other taxa in the matrix, thus contributing no unique information (Fig. 3). As Wilkinson (1992, 1995a) suggested, removal of such taxa will often drastically reduce the number of equivalent trees. When such taxa are deleted,

Data matrix:

Out	000	000	000	000
A	100	000	000	100
B	111	000	000	000
C	111	110	000	010
D	111	111	100	000
E	111	111	100	001
F	000	000	111	000
G	000	000	111	000
H	000	000	011	000
I	1??	???	???	111
J	0??	???	?11	???

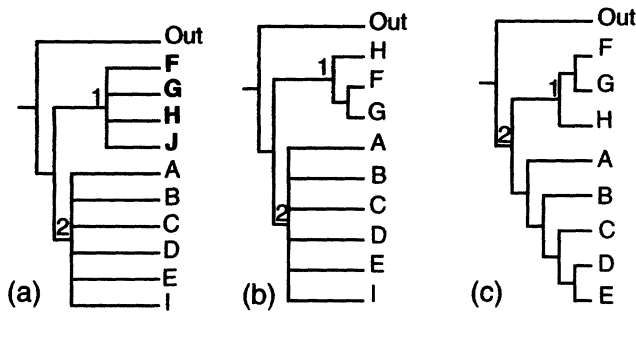


FIGURE 3. Effects of deleting taxonomic equivalents. (a) Analysis of the complete data matrix results in a strict consensus tree containing the clades 1 and 2. Safe Taxonomic Reduction identifies J as a taxonomic equivalent of F, G, and H, and thus J can be "safely removed" from the matrix. (b) Reanalysis yields the same two supported clades, and clade 1 now has more apparent resolution. Clade 2 remains poorly resolved because taxon I, although not a taxonomic equivalent, is unstable. (c) Resolved cladogram if taxon I were also removed from the analysis.

more apparent resolution is obtained because redundant taxa are pruned from polytomies shared with their equivalents. However, contrast this with the example given by Nixon and Wheeler (1992) (Fig. 1 here), in which the taxon G is unstable because of a mixture of homoplasy and missing data. If that type of wildcard problem exists in real data sets, it cannot be solved by STR because the fragmentary taxon in this case is not a taxonomic equivalent.

Less Strict Consensus Methods

Swofford (1991) reviewed various consensus methods and argued for the use of methods other than strict consensus (Schuh and Farris, 1981; Schuh and Polhemus, 1981) that could preserve more of the structure of the primary trees. Wilkinson (1994, 1995a) also argued that strict consensus methods are too strict. In contrast, Nixon and Carpenter (1996b) argued that all methods other than strict consensus are compromise methods, methods to be avoided because they do not necessarily reflect the agreement in grouping among all primary trees (but see Wilkinson and Thorley, 2001). I consider here methods that have been advocated specifically for dealing with the effects of missing data.

The Adams consensus method (Adams, 1972) identifies unstable taxa and collapses the nodes corresponding to different positions for those taxa in the alternative cladograms to the first node that includes those alternative placements (Fig. 4a). Gordon's (1980) common pruned trees (Fig. 4b) and

Wilkinson's (1994) reduced consensus methods (Fig. 4c) are variants of a general taxon-pruning approach (see also Anderson, 2001). These methods yield consensus trees that may contain fewer taxa than the primary cladograms. Increased resolution is obtained by pruning one or more taxa until a resolved topology is acquired (the possible placements for pruned taxa may be annotated in some manner). All of these methods preserve structure by removing or collapsing unstable taxa, but they also represent the compromise trees of Nixon and Carpenter (1996b): They may contain groups contradicted in some of the primary trees or may lack supported groups present in the primary trees. Further, they do not distinguish between no-data-ambiguity and conflicting-data-ambiguity. This is unsurprising, given that consensus methods are blind to causes of ambiguity.

USING SUMMARY TREES INSTEAD

An accurate summary of primary trees would reveal all groups supported in the primary trees, contain no groups contradicted in the primary trees, and not allow missing data to obscure supported groups. No existing consensus method meets these goals. Problems with consensus methods occur because the relationship between the consensus tree and the data matrix is ignored when only topological agreement among primary trees is focused upon. A more accurate method requires a consideration of character support for groups—both conflicting and

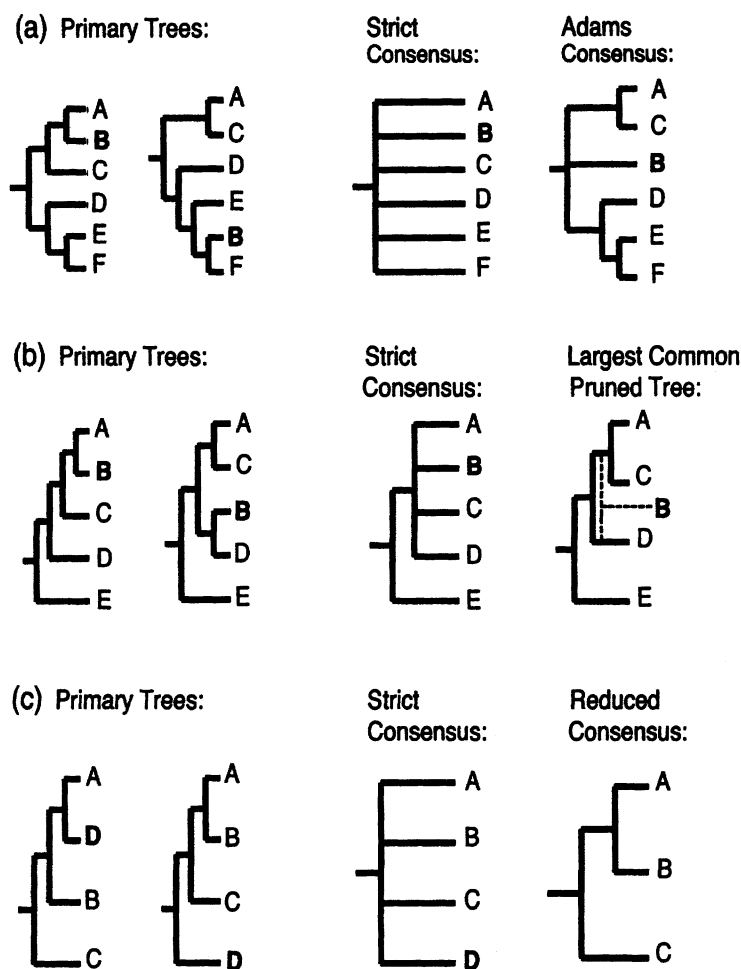


FIGURE 4. Consensus methods proposed as alternatives to the strict method for decreasing ambiguity resulting from unstable taxa. (a) The strict consensus of the two primary trees is completely unresolved because of the alternative placements of taxon B. The Adams consensus of these two trees identifies the unstable taxon B and places it unresolved at the base of this clade. The alternative evidence that exists for groups (AB) and (FB) among the primary trees is concealed by the Adams consensus. (b) The largest common pruned tree method identifies and prunes the unstable taxon B. The same comments from (a) above apply here except that the regrafted tree indicates the possible positions of taxon B. (c) The reduced cladistic consensus method identifies and excludes unstable taxa. Possible placements for excluded taxa may be annotated in some manner. The same comments apply.

nonconflicting support. This can be accomplished by assessing supported conflict for nodes and by constructing summary trees on that basis.

To build a summary tree, the unstable taxon or taxa first must be identified and, second, the cause of the instability must be identified. For small numbers of primary trees, this will not present a problem. Ambiguously resolved taxa can be identified by comparing all trees, and the cause of the instability can then be identified by simply examining character optimizations on those trees by eye.

However, the problem with many analyses is that tens of thousands of trees may make direct examination of all topologies too difficult. An a posteriori solution that can assess the cause of instability for groups will then be necessary.

Consideration of supported intercladogram conflict was first addressed by Nixon and Carpenter (1996b) with the development of the Clade Concordance (CC) index. The CC index measures the degree of supported conflict among primary trees. Nixon and Carpenter (1996b) suggested the use of the

CC index for assessing whether nodes are unambiguously supported under all possible character optimizations. CC measures the degree of character conflict that occurs between (not within) alternative cladograms, essentially by making use of the information provided by consensus tree length relative to primary tree length. CC will approach 1 when differences among primary trees are not due to any character conflict and 0 when differences among cladograms are entirely the result of character conflict. In some simple cases, one may be able to utilize the CC index to identify the contributed character conflict of particular unstable taxa and thereby identify a wildcard effect in a large data matrix. For example, if a single fragmentary taxon is contributing no character conflict but is behaving as a wildcard, the CC index should be greater if the taxon is included than if it is excluded. In Figure 2c, the CC index is 0 when F is excluded from the analysis and 1 when F is included.

An obvious brute-force strategy, in cases where direct examination of all trees is intractable, is to use an Adams or reduced consensus tree as a heuristic method for flagging unstable taxa, and to then identify the cause of instability (missing data, or homoplasy, or both) by tracing characters on this consensus. (This exercise is intended only as a method for identifying supported conflict for groups, not for examining character evolution per se, which would be undesirable on a consensus tree.) Figure 5 illustrates this method for two simple data sets.

An algorithmic solution to the wildcard problem that would produce a summary tree based on character support is desirable. Such an algorithm might be based on an assessment of character support for each node in the set of primary trees before collapsing only those groups with supported conflict into a summary tree. Using certain metrics to determine the amount of extra homoplasy, if any, that results from the inclusion of suspected wildcard taxa might also be a way to develop shortcuts for such algorithms. For example, in the case of a taxon that is unstable because of both missing data and character conflict, the length of the strict consensus tree is equal to the length of the Adams consensus tree and both of these exceed the length of the primary trees (see Fig. 5b). For a taxon unstable as a result of character conflict only, the length of the strict consensus tree exceeds

that of the Adams consensus tree (see Fig. 5c). The difference is because, in the former case, the characters reach their maximum length on the Adams consensus.

DIFFERENT TYPES OF WILDCARDS—EXPLORING RECENT ANALYSES

Nixon and Wheeler (1992) characterized a wildcard as a taxon that is unstable because of alternative optimizations of question marks. But others have construed the meaning of wildcards less conservatively, to refer to any highly unstable taxon. This has caused some confusion in the literature. The three potential types of wildcards are as follows (see also Fig. 5):

- 1) Missing data wildcard—A taxon that contributes a mixture of completely congruent characters as well as question marks (i.e., Wilkinson's [1992, 1995a] taxonomic equivalent). This type of wildcard's instability is due entirely to missing data.
- 2) Mixed wildcard—A taxon that contributes both question marks and character incongruence; its instability is due to a combination of missing data and character conflict (i.e., Nixon and Wheeler's [1992] original wildcard example).
- 3) Conflict wildcard—A taxon that carries some character incongruence and is unstable entirely because of character conflict. Such a taxon may or may not also contribute missing data to the matrix but in any case, the missing data are not the cause of its instability. This type of wildcard is the one that has figured most prominently in discussions of alternative consensus methods such as those of Swofford (1991) and Wilkinson (1994, 1995b).

In exploring the types of wildcard problems commonly encountered in combined analyses, I reexamined the published combined analyses of Gauthier (1986), Novacek (1992), Grande and Bemis (1998), Norell and Gao (1997), and Gao and Norell (1998), all of whom have discussed problems related to missing data, fossil taxa, and ambiguity.

Gauthier.—Gauthier (1986) explored the results of two analyses, one that included all taxa and yielded thousands of most-parsimonious trees (Fig. 6a), and one that excluded the most incomplete taxa and yielded

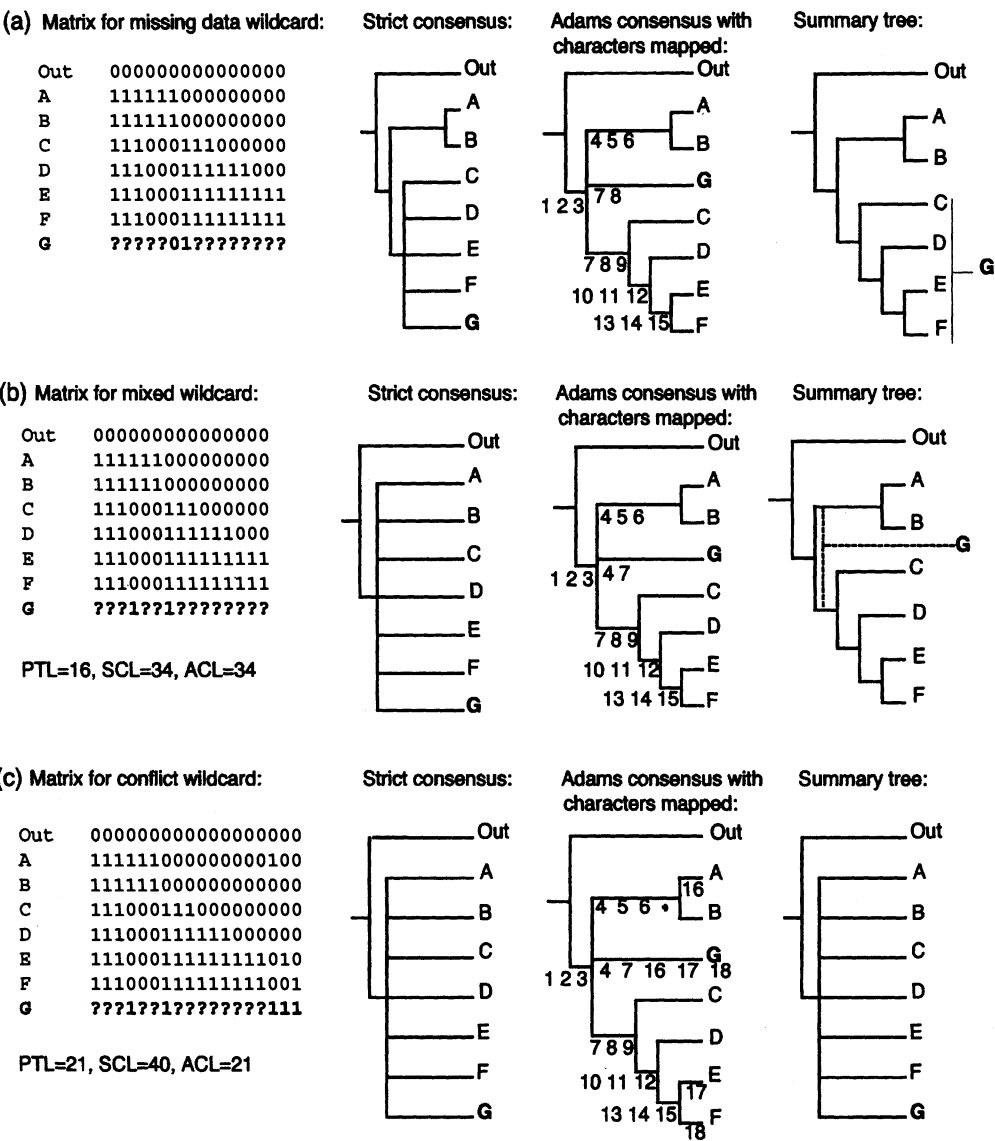


FIGURE 5. (a) When taxon G is included in the analysis, the resulting strict consensus is poorly resolved. An Adams consensus identifies G as unstable. Mapping characters onto that tree shows the cause of taxon G's instability to be mainly missing data. Because G is a taxonomic equivalent of the group (C(D(EF))), the summary tree shown here is more appropriate than the unresolved strict consensus tree, which obscures the supported group (C(D(EF))). (b) When taxon G is included in the analysis, the resulting strict consensus is completely unresolved. An Adams consensus identifies G as unstable. In this case, mapping characters onto that tree indicates that taxon G contains two conflicting characters (4 and 7) and is missing data for all the other characters, causing the structure of the entire tree to collapse due to both character conflict and missing data. The summary tree is more appropriate than the strict consensus because it reveals the supported groups (AB) and (C(D(EF))). (c) The data matrix without taxon G supports a fully resolved cladogram for taxa A–F. Including taxon G in the analysis results in a completely unresolved strict consensus. An Adams consensus identifies G as unstable. In this case, mapping characters onto that tree reveals the cause of G's instability to be character conflict. Accordingly, an unresolved strict consensus is the most appropriate summary of the primary cladograms and the data matrix. ACL = length of Adams consensus, PTL = length of primary trees, SCL = length of strict consensus.

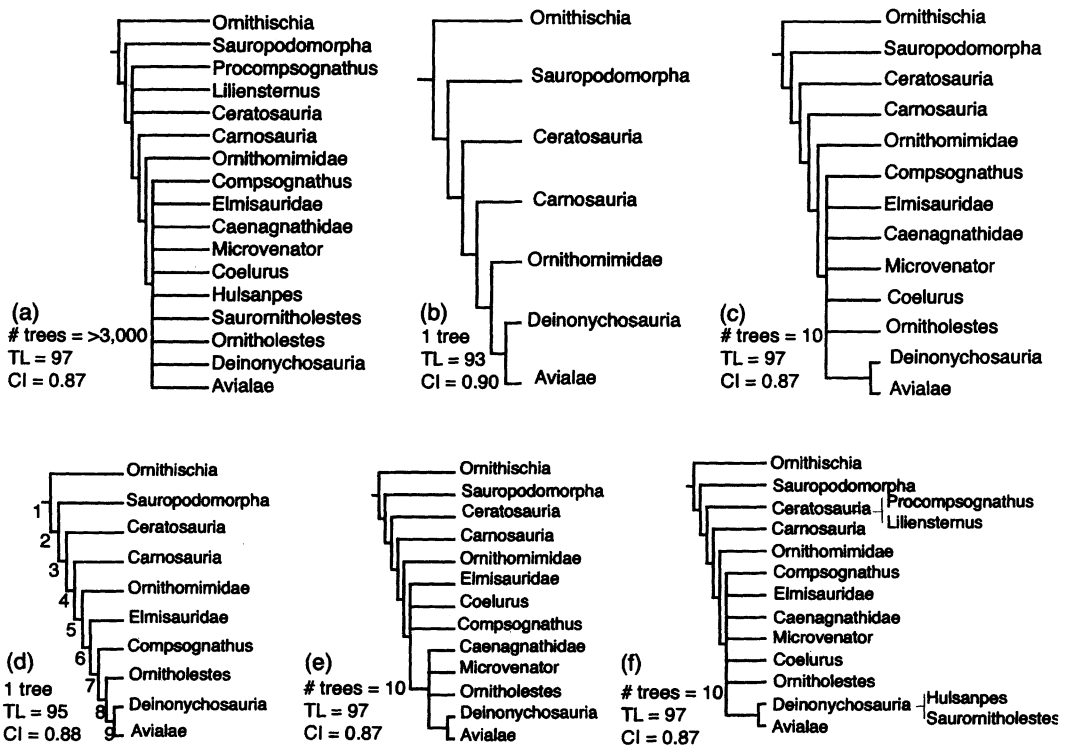


FIGURE 6. (a) Strict consensus of >3,000 shortest trees from Gauthier's (1986) saurischian analysis. (b) Single tree obtained by Gauthier (1986) by removing the most incomplete taxa. (c) Strict consensus of 10 trees after removal of taxonomic equivalents. (d) Reduced consensus [from Wilkinson, 1995a] of those same 10 trees. (e) Adams consensus of those same 10 trees. (f) Summary tree of those same 10 trees.

one most-parsimonious tree (Fig. 6b). Gauthier proposed that the numerous alternative trees produced from analysis of the complete data set resulted from missing data in the most incomplete taxa. Wilkinson (1995a), reanalyzing Gauthier's data by implementing his STR method, identified *Procompsognathus* and *Liliensternus* as taxonomic equivalents of the more complete taxon *Ceratosauria*, and *Saurornitholestes* and *Hulsanpes* as taxonomic equivalents of the more complete taxon *Deinonychosauria*. Removal of those four redundant taxa from the data set and subsequent reanalysis resulted in 10 most parsimonious trees, the strict consensus of which is shown in Figure 6c. Wilkinson's final procedure was the construction of the reduced consensus tree shown in Figure 6d, which also pruned the unstable taxa *Coelurus*, *Microvenator*, and *Caenagnathidae* and resulted in a fully resolved tree. Using numbered nodes allows a description of the possible positions for excluded taxa.

This latter step (reduced consensus) is one I do not necessarily recommend as an actual summary of results. The cause of the instability for these three taxa is observed character conflict, not missing data, and removing them may be misleading by obscuring the existence of evidence for alternative relationships. The usual justification for a method such as reduced consensus or common pruned trees is that stable relationships exist among some taxa in the primary trees when unstable taxa are ignored (Swofford, 1991; Wilkinson, 1994, 1995b). Although this may be true, alternative relationships suggested by conflicting data are just as much a part of cladistic results as are "stable relationships." Such methods may have heuristic value—they identify and remove taxa that exhibit a high degree of instability, but they should not be viewed as actual representations of the results of cladistic analysis.

The Adams consensus (Fig. 6e) differs from the strict consensus in that *Microvenator*, *Caenagnathidae*, and *Ornitholestes* are placed

in a group with the Deinonychosauria–Avialae clade, but character optimization does not indicate any remaining ambiguity due to missing data. Figure 6f illustrates a summary tree of these results. Further resolution could potentially be obtained by character reanalysis, addition of new data, or both. In summary, some of the ambiguity in Gauthier's original tree was the result of missing data, but some resulted from conflicting evidence. The summary tree (Fig. 6f) preserves the supported Deinonychosauria–Avialae clade that is obscured by taxonomic equivalents in the all-taxa tree (Fig. 6a); provides at least some information on the affinities of taxa such as *Compsognathus* and *Caenagnathidae* that is absent in the extant-only tree (Fig. 6b); contains only unambiguously supported groups from the primary trees, unlike the reduced consensus (Fig. 6d); and illustrates the taxonomic equivalence of the pruned taxa.

Novacek.—Novacek (1992) discussed the effects of incomplete fossil taxa and missing data in his analysis. When 20 extant terminal taxa alone were analyzed, eight shortest trees were obtained (strict consensus shown in Fig. 7a). Inclusion of seven incomplete fossil taxa increased the number of trees to more than 10,000, and Novacek noted a less resolved strict consensus tree (Fig. 7b). Screening the Novacek (1992) data set with the TAXEQ3 program (Wilkinson, 2001) identified four taxonomic equivalents (*Asio-ryctes*, equivalent of Carnivora; *Desmostylia*, equivalent of Sirenia; *Microsyopidae*, equivalent of Carnivora; and *Kennalestes*, equivalent of Carnivora) that could be removed. Reanalysis of the TAXEQ3 data set resulted in eight shortest trees and the strict consensus shown in Figure 7c. The Adams consensus (Fig. 7d) differs from the strict consensus in two respects: The position of Carnivora is resolved and the Scadentia–Primates and Dermoptera–Chiroptera clades are united rather than unresolved, but examination of character optimizations shows the strict consensus ambiguity to be the result of character conflict. Figure 6e illustrates a summary tree of these results. In this data set, therefore, most of the ambiguity was apparent ambiguity, caused by the clustering of taxonomic equivalents with more complete terminals. The summary tree provides more information than the extant-only tree (Fig. 7a), showing that Leptictida is the sister-group to

Insectivora, and Tubulidentata is the sister-group to that clade. In comparison with the all-taxa tree (Fig. 7b), the summary tree reveals the sister-group relationship between Sirenia and Proboscidea, which had previously been obscured by the taxonomic equivalent *Desmostylia*. An Adams consensus or reduced consensus of these trees would imply more resolution than is supported by the data.

Gao and Norell.—Gao and Norell (1998) analyzed the relationships of anguimorph lizards on the basis of a data set containing 58% fossil taxa and 34% overall missing data. Results of their analysis appeared quite ambiguous, having more than 32,700 (the user-specified maximum) shortest trees (Fig. 8a). After removal of the nine most incomplete taxa, the authors obtained a more highly resolved strict consensus tree (Fig. 8b). Screening the Gao and Norell (1998) data set with the TAXEQ3 program (Wilkinson, 2001) identified three taxonomic equivalents that could be removed from the analysis. Reanalysis of the TAXEQ3 data set resulted in 111 shortest trees (Fig. 8c). Use of an Adams tree (Fig. 8d) and character mapping indicated no remaining ambiguity that could be attributed to missing data. The Gao and Norell (1998) data set contained three missing data wildcards. Removal of those three taxa and reanalysis with the use of a summary tree (Fig. 8e) provides an accurate summary of results.

Other reanalyses.—Similar results were obtained with reanalysis of two other recent combined analyses that encountered “missing data problems” (Norell and Gao, 1997; Grande and Bemis, 1998). Grande and Bemis (1998), in particular, discussed problems associated with missing data extensively and presented a serious critique on the inclusion of fragmentary fossils. They concluded that highly fragmentary taxa should be excluded at the researcher's discretion because such taxa can collapse supported nodes in consensus trees (wildcard effect). However, in each of these two studies, reanalysis showed that most ambiguity was attributable to taxonomic equivalency, and well-resolved summary trees were produced. In the case of Grande and Bemis (1998), only five primary trees were found after the removal of taxonomic equivalents.

Thus, despite extensive discussion about missing data problems in the literature and

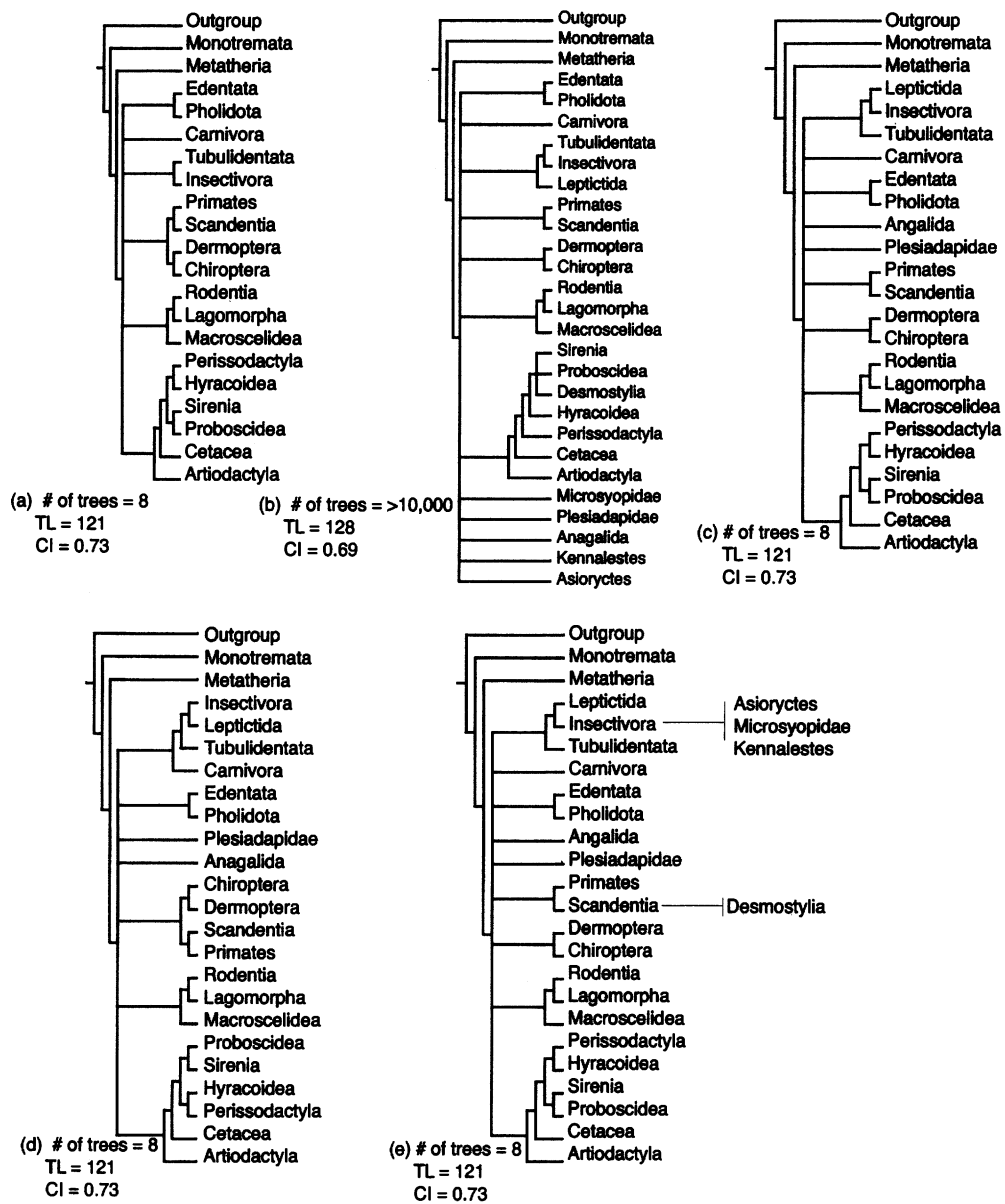


FIGURE 7. (a) Strict consensus of eight shortest trees from Novacek's (1992) analysis of the extant mammalian taxa only. (b) Strict consensus of >6,800 shortest trees from Novacek's (1992) analysis of the entire data set, including seven incompletely known taxa. (c) Strict consensus of eight shortest trees after removal of four taxonomic equivalents. (d) Adams consensus of those same eight trees. (e) Summary tree for Novacek's (1992) data set.

the widespread assumption of a connection between fragmentary taxa (especially fossils), missing data, and ambiguity, recent combined analyses apparently (1) sometimes achieve increased resolution by including incomplete fragmentary taxa, (2) often encounter problems where most of the apparent ambiguity is a result of redundancy in incomplete taxa and can be easily resolved with the STR method (Wilkinson, 1992, 1995a, 2001), or (3) encounter ambiguity that has nothing to do with missing data. The common assumption of a missing data ambiguity problem and strategies of simply excluding fragmentary taxa for such a perceived problem are ill-advised. Instead, it is prudent to

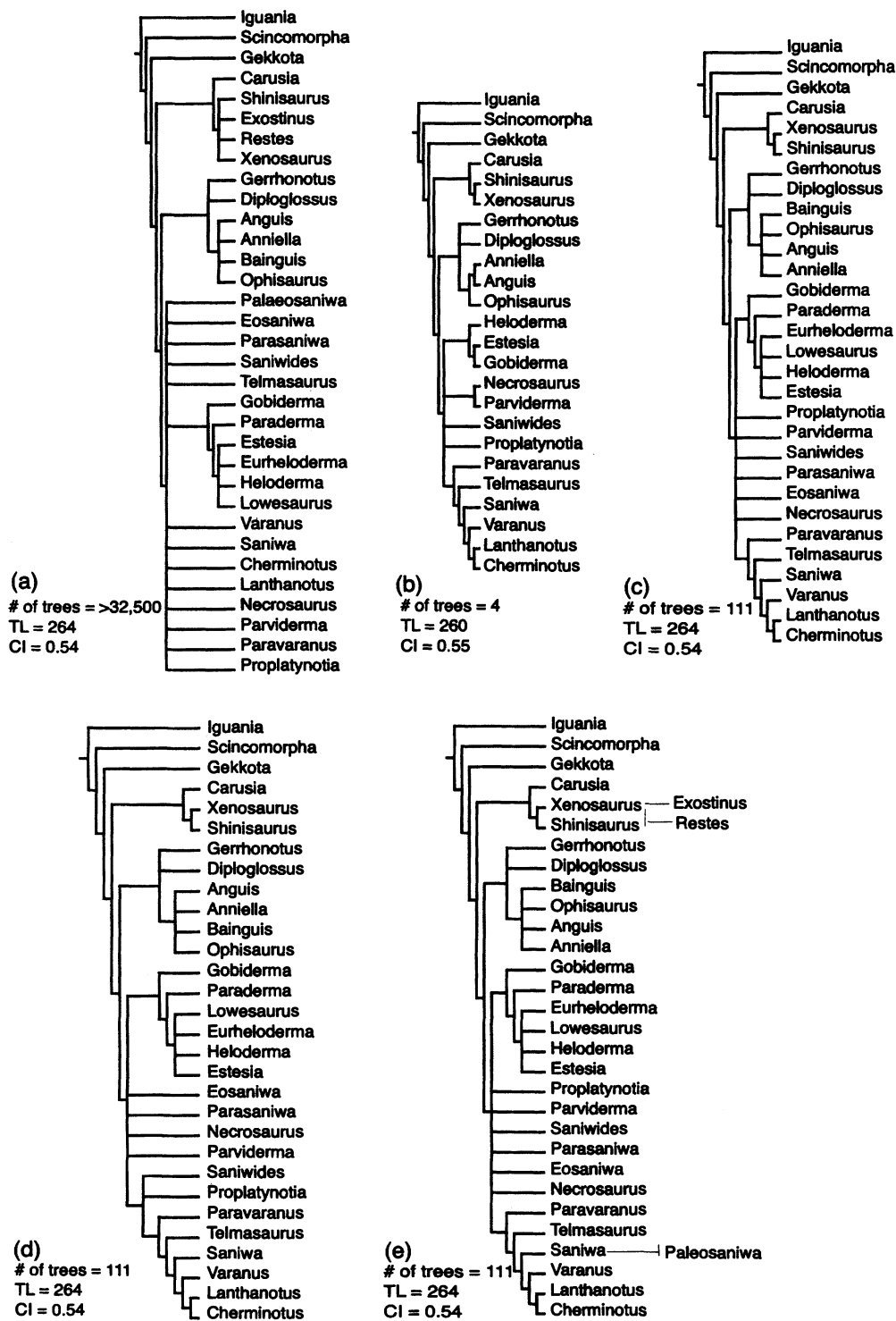


FIGURE 8. (a) Strict consensus of >32,000 primary trees obtained by Gao and Norell (1998) from analysis of their entire data set. (b) Strict consensus of four trees obtained by Gao and Norell (1998) from analysis of a reduced data set, excluding the nine most incomplete taxa. (c) Strict consensus of 111 primary trees obtained by analysis of the Gao and Norell (1998) data set after removal of three taxonomic equivalents, *Exostinus*, *Restes*, and *Palaeosaniwa*. (d) Adams consensus of those 111 trees. (e) Summary tree for the Gao and Norell (1998) data set.

assess the effects of incomplete taxa empirically on a case-by-case basis.

RECOMMENDED STRATEGY FOR COMBINED ANALYSES INCLUDING FRAGMENTARY TAXA

All relevant taxa and characters can be included at the outset, whether complete or incomplete. Because the degree of missing data may not be strictly correlated with the degree of ambiguity in results, a wildcard problem must be deduced after the cladogram is built rather than anticipated in advance by excluding fragmentary taxa or characters a priori. That satisfactory resolution may result despite extensive missing data is entirely possible, depending on the distribution of congruent characters, homoplasy, and missing data in a specific matrix.

A strict consensus can be used to initially assess agreement among alternative cladograms. If a strict consensus appears highly ambiguous, the next step would be to screen for and delete taxonomic equivalents (Wilkinson, 1992, 1995a); this will remove one type of wildcard from the analysis. If results are still ambiguous, and if other taxa are identified as possible wildcards, the cause of the ambiguity for those taxa can be determined either with direct methods or with potential algorithmic shortcuts. Finally, results can be depicted in a summary tree that is consistent with different causes of ambiguity.

CONCLUSION

In spite of potential computational problems, fragmentary taxa (including fossils) can provide data capable of testing phylogenetic hypotheses; indeed, some investigators argue that fossil taxa, in particular, provide a unique type of data that compels their inclusion. All taxa, no matter how incomplete, have the potential to carry some evidence for resolving relationships or to carry some character incongruence (or both) and thus to either decrease or increase ambiguity. Three mistaken assumptions have been made in several recent studies: (1) Fragmentary taxa will always behave as wildcards; (2) analyses that include fragmentary taxa and yield many primary trees demonstrate a causal relationship between missing data in those taxa and ambiguity of results; and (3) ambiguity contributed by fragmentary

taxa is solely the result of missing data. Actually, however, large, combined data sets that contain incomplete taxa are likely to be complicated mixtures of signal, homoplasy, and missing data. Thus, consensus ambiguity cannot be assumed to be exclusively attributable to missing data or the wildcard effect. In contrast to some recent suggestions, missing data in incomplete fossils are not an impediment to a total evidence (Kluge, 1989), or simultaneous analysis, (Nixon and Carpenter, 1996a) approach.

Alternative consensus methods proposed as a solution to the problem of ambiguity caused by highly unstable taxa have emphasized the topological consequences of missing data, with relatively little regard for the exact causes of these changes in topology. As such, they fail to distinguish between lack of character evidence for resolving relationships and conflicting character evidence for alternative relationships. Because of this distinction, it is preferable to address this problem in terms of character support for groups in consensus trees.

A common theme in many phylogenetic studies is the desire to obtain better resolution. Although resolution of relationships is obviously a goal of phylogenetic analysis, it should not be obtained at the expense of ignoring data. Ambiguity of results calls for reexamination of data and addition of new data, rather than use of methods that may imply more resolution than the data support.

ACKNOWLEDGMENTS

For discussions regarding missing data, total evidence, wildcards, and computer programs, I thank J. Carpenter, J. Clark, M. Daly, K. de Queiroz, A. Kluge, D. Lipscomb, K. Nixon, M. O'Leary, E. Strong, and J. J. Wiens. I thank D. Sikes for constructive criticism and M. Wilkinson for extensive discussion and debate.

REFERENCES

- ADAMS, E. N. 1972. Consensus techniques and the comparison of taxonomic trees. *Syst. Zool.* 21:390–397.
- ANDERSON, J. S. 2001. The phylogenetic trunk: Maximal inclusion of taxa with missing data in an analysis of the Lepospondyli (Vertebrata, Tetrapoda). *Syst. Biol.* 50:170–193.
- BENTON, M. J. 1990. The species of *Rhynchosaurus*, a rhynchosaur (Reptilia, Diapsida) from the middle Triassic of England. *Philos. Trans. R Soc. London* 328:213–306.
- CARPENTER, J. M. 1987. Phylogenetic relationships and classification of the Vespinae (Hymenoptera: Vespidae). *Syst. Entomol.* 12:413–431.

- CREPET, W. L., AND K. C. NIXON. 1989a. Earliest megafossil evidence of Fagaceae: Phylogenetic and biogeographic implications. *Am. J. Bot.* 76:842–855.
- CREPET, W. L., AND K. C. NIXON. 1989b. Extinct transitional Fagaceae from the Oligocene and their phylogenetic implications. *Am. J. Bot.* 76:1493–1505.
- DOYLE, J. A., AND M. J. DONOGHUE. 1987. The importance of fossils in elucidating seed plant phylogeny and macroevolution. *Rev. Paleobot. Palynol.* 50:63–95.
- EERNISSE, D. J., AND A. G. KLUGE. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules and morphology. *Mol. Biol. Evol.* 10:1170–1195.
- FARRIS, J. S. 1988. HENNIG86, version 1.5. Program and documentation. Port Jefferson Station, New York.
- FRASER, N. C., AND M. BENTON. 1989. The Triassic reptiles *Brachyrhinodon* and *Polysphenodon* and the relationships of the sphenodontids. *Zool. J. Linn. Soc.* 96:413–445.
- GAO, K., AND M. A. NORELL. 1998. Taxonomic review of *Carusia* (Reptilia: Squamata) from the late Cretaceous of the Gobi Desert and phylogenetic relationships of anguimorph lizard. *Am. Mus. Novit.* 3230:1–51.
- GAUTHIER, J. 1986. Saurischian monophyly and the origin of birds. *Mem. Calif. Acad. Sci.* 8:1–56.
- GAUTHIER, J., A. G. KLUGE, AND T. ROWE. 1988. Amniote phylogeny and the importance of fossils. *Cladistics* 4:105–208.
- GOLOBOFF, P. A. 1993. NONA, version 1.6. Program and documentation. American Museum Natural History, New York.
- GORDON, A. D. 1980. On the assessment and comparison of classifications. Pages 149–160 in *Analyse de donnees et informatique* (R. Tomassone, ed.). INRIA, LeChesnay.
- GRANDE, L., AND W. E. BEMIS. 1998. A comprehensive phylogenetic study of amiid fishes (Amiidae) based on comparative skeletal anatomy. An empirical search for interconnected patterns of natural history. *Soc. Vert. Paleontol. Memoir* 4, suppl. to volume 18.
- KLUGE, A. G. 1989. A concern for evidence and a phylogenetic hypothesis for relationships among *Epicrates* Boidae, Serpentes). *Syst. Zool.* 38:1–25.
- MESSINGER, S. L., AND J. A. MCGUIRE. 1998. Morphology, molecules and the phylogenetics of cetaceans. *Syst. Biol.* 47:90–124.
- NIXON, K. C. 1996. Paleobotany in cladistics and cladistics in paleobotany: enlightenment and uncertainty. *Rev. Paleobot. Palynol.* 90:361–373.
- NIXON, K. C., AND J. M. CARPENTER. 1996a. On simultaneous analysis. *Cladistics* 12:221–242.
- NIXON, K. C., AND J. M. CARPENTER. 1996b. On consensus, collapsibility, and clade concordance. *Cladistics* 12:305–322.
- NIXON, K. C., AND WHEELER Q. D. 1992. Extinction and the origin of species. Pages 119–143 in *Extinction and phylogeny* (M. J. Novacek and Q. D. Wheeler, eds.). Columbia Univ. Press, New York.
- NORELL, M. A., AND K. DE QUEIROZ. 1991. The earliest iguanine lizard (Reptilia: Squamata) and its bearing on iguanine phylogeny. *Am. Mus. Novit.* 2997:1–16.
- NORELL, M. A., AND K. GAO. 1997. Braincase and phylogenetic relationships of *Estesia mongoliensis* from the Upper Cretaceous, Gobi Desert and the recognition of a new clade of lizards. *Am. Mus. Novit.* 3211:25 pp.
- NOVACEK, M. J. 1992. Fossils, topologies, missing data, and the higher level phylogeny of eutherian mammals. *Syst. Biol.* 41:58–73.
- O'LEARY, M. A. 1999. Parsimony analysis of total evidence from extinct and extant taxa, and the cetacean-artiodactyl question. *Cladistics* 15:315–330.
- O'LEARY, M. A. 2000. Operational obstacles to total evidence analyses considering that 99% of life is extinct. *J. Vert. Paleontol.*, volume 20, Supplement:61A. Abstracts of papers.
- ROWE, T. 1988. Definition, diagnosis and origin of Mammalia. *J. Vert. Paleontol.* 8:241–264.
- SCHUH, R. T., AND J. S. FARRIS. 1981. Methods for investigating taxonomic congruence and their application to the Leptopodomorpha. *Syst. Zool.* 30:331–351.
- SCHUH, R. T., AND J. T. POLHEMUS. 1981. Analysis of taxonomic congruence among morphological, ecological, and biogeographic data sets for the Leptopodomorpha (Hemiptera). *Syst. Zool.* 29:1–26.
- SWOFFORD, D. L. 1991. When are phylogeny estimates from molecular and morphological data incongruent? Pages 295–333 in *Phylogenetic analyses of DNA sequences* (M. M. Miyamoto and J. Cracraft, eds.). Oxford Univ. Press, New York.
- SWOFFORD, D. L. 1993. PAUP, phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign.
- WIENS, J. J. 1998. Does adding characters with missing data increase or decrease phylogenetic accuracy? *Syst. Biol.* 47:625–640.
- WILKINSON, M. 1992. TAXEQ2, software and documentation. Univ. Bristol, Bristol, England.
- WILKINSON, M. 1994. Common cladistic information and its consensus representation: Reduced Adams and reduced cladistic consensus trees and profiles. *Syst. Biol.* 43:343–368.
- WILKINSON, M. 1995a. Coping with abundant missing entries in phylogenetic inference using parsimony. *Syst. Biol.* 44:501–514.
- WILKINSON, M. 1995b. More on reduced consensus methods. *Syst. Biol.* 44:436–440.
- WILKINSON, M. 2001. TAXEQ3. Software and documentation. Natural History Museum, London.
- WILKINSON, M., AND M. J. BENTON. 1995. Missing data and rhynchosaur phylogeny. *Hist. Biol.* 10:137–150.
- WILKINSON, M., AND J. L. THORLEY. 2001. No compromise on consensus. *Taxon* 50:181–184.
- WU, X.-C., D. B. BRINKMAN, AND A. P. RUSSELL. 1996. *Sineoamphisbaena hexatabularis*, an amphisbaenian (Diapsida: Squamata) from the Upper Cretaceous redbeds at Bayan Mandahu (Inner Mongolia, People's Republic of China), and comments on the phylogenetic relationships of the Amphisbaenia. *Can. J. Earth Sci.* 33:541–577.

Received 16 March 2001; accepted 7 November 2001
Associate Editor: T. Collins