PROBLEMS DUE TO MISSING DATA IN PHYLOGENETIC ANALYSES INCLUDING FOSSILS: A CRITICAL REVIEW

MAUREEN KEARNEY[1] and JAMES M. CLARK[2]

[1]Department of Zoology, Field Museum of Natural History, Chicago, Illinois 60605, U.S.A., mkearney@fieldmuseum.org;
[2]Department of Biological Sciences, George Washington University, Washington, DC 20052, U.S.A.

ABSTRACT—We review the widespread notion that the inclusion of taxa scored for relatively few characters is problematic in phylogenetic analyses. Taxa scored for few characters may lead to lack of resolution, but need not. Lack of resolution may be unrelated to missing data when characters conflict. Missing data cannot produce groupings for which there is no evidence. A common approach to avoid the "missing data problem" is to exclude incomplete taxa, but excluding such taxa is inadvisable because the information content of taxa is not necessarily correlated with degree of completeness. Another prevalent strategy—excluding characters with a high proportion of missing data—may actually contribute to the low resolution problem rather than ameliorate it because removing any character data removes potentially informative synapomorphies. Other approaches, including the use of less-than-strict consensus techniques, have the potential to obscure evidence for alternative relationships or, at best, provide incomplete summaries of the primary trees. Missing data simply represent the unknown and should not be viewed as an impediment to considering all available evidence in phylogenetic analyses, nor used as justification for excluding specific taxa or characters.

INTRODUCTION

Homology statements are the basis of phylogenetic analysis, but homologies cannot be assessed for parts of organisms that are unknown. These "missing data" are perceived to be a problem in paleontology because the relationships of taxa for which few homology statements can be made are sometimes difficult to resolve. When an incompletely known taxon is incorporated into a data matrix with other taxa, large numbers of equivalent cladograms sometimes result, the strict consensus of which may be poorly resolved. It is often assumed that missing data create this problem, and therefore that missing data should be avoided via the exclusion of incomplete taxa, but the nature of the problem has been little explored. Missing data are associated with fossil taxa, but they are neither limited to fossils nor necessarily more problematic with them. Missing data are introduced to an analysis during the first stage of homology assessment ("primary homology" of de Pinna, 1991) when there is no topological equivalent available to form a homology statement. This commonly occurs with fossils which, for example, typically lack all soft tissues, but this also may occur with complete organisms due to transformation and resulting character inapplicability. Thus, digital characters cannot be coded in snakes, which lack digits, and genetic characters are susceptible to the same problem when, for example, an entire gene is absent.

Analyzing living and fossil taxa simultaneously is a powerful phylogenetic approach because it results in the most strongly corroborated global hypothesis of relationships (Kluge, 1989; Nixon and Carpenter, 1996a). Some have questioned whether such a total evidence or simultaneous analysis approach might be problematic due to the inflation of missing data in data matrices that combine fossil and living taxa, but the evolutionary study of fossils must involve phylogenetic analysis, and missing data will always be a factor in phylogenetic studies. Furthermore, the benefits and/or drawbacks of simultaneous analysis vs. increased missing data associated with combined analyses have not been thoroughly explored.

Potential problems thought to be associated with missing data notwithstanding, fossils inarguably provide data capable of testing phylogenetic hypotheses, and some paleontologists have argued further that they provide a unique type of data that compels their inclusion (e.g., Gauthier et al., 1988; Donoghue et al., 1989). In the end, all evidence bearing on a hypothesis should be considered, and the exclusion of evidence that could either refute or support a hypothesis is difficult to justify. However, the introduction of poorly known taxa is thought to lead to the operational problem of multiple equally parsimonious cladograms and concomitant ambiguity in results. Phylogeneticists therefore find themselves balancing the desire to unambiguously resolve relationships with the desire to accommodate as much evidence as possible. Whether this dichotomy is more real or imaginary remains relatively unknown.

EVOLUTION OF THE "MISSING DATA PROBLEM"

While there have been many general discussions regarding the role that fossils play compared to extant taxa in systematics (e.g., Patterson, 1981; Ax, 1987; Gauthier et al., 1988; Donoghue et al., 1989), the first specific mention that we can find of "missing data" pertaining to quantitative phylogenetic analyses that include incomplete fossils comes from Gauthier (1986). With the advent of computerized algorithms for phylogenetic analysis, computational problems associated with missing character state entries came into focus and largely replaced previous theoretical issues surrounding the information provided by fossil taxa. In discussing the results of two analyses, one that included only well-known taxa, and another that included less well-known taxa, Gauthier (1986:8) pointed out that, "The multiple trees obtained in the second run resulted from the missing data in the 10 less well-known taxa." From that point until today, the issue of missing data has evolved into a topic significant enough for an SVP symposium and this volume. Yet, the issue remains more of a catchphrase than a well-understood problem, and a review of the evolution of this topic below reveals the perpetuation of several misconceptions that unfortunately have begun to influence systematists' inclusion criteria for fossil taxa in phylogenetic analyses.

Rowe (1988) measured the completeness of taxa as the percentage of the total number of characters that could be scored for each taxon, and included only those taxa that fell within the range of completeness for the extant taxa in his analysis (a range of 88–96% completeness). This is the first paper we have
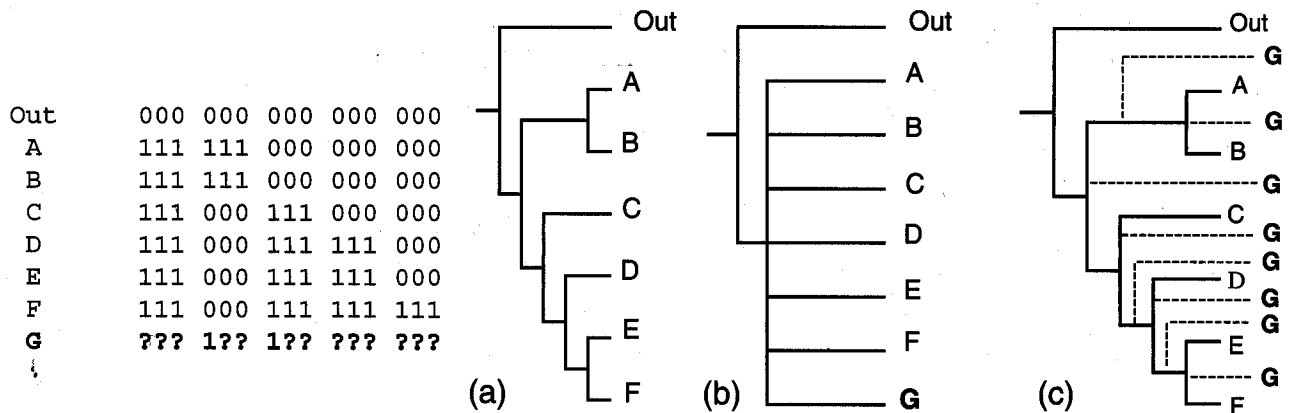
Out	000	000	000	000	000
A	111	111	000	000	000
B	111	111	000	000	000
C	111	000	111	000	000
D	111	000	111	111	000
E	111	000	111	111	000
F	111	000	111	111	111
G	???	1??	1??	???	???

(a)   (b)   (c)

FIGURE 1. The wildcard problem (redrawn from Nixon and Wheeler, 1992). **a.** Single tree resulting from parsimony analysis of the data set excluding the wildcard taxon G. **b.** Strict consensus of eight trees resulting from parsimony analysis of the data set including the wildcard taxon G. **c.** The eight possible positions of G are indicated with dashed lines. Note that these eight trees were obtained by analysis with the program HENNIG86 (Farris, 1988) or PAUP (Swofford, 1993). When the data set is analyzed with NONA or PAUP* with the "amb-" option, 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.

found that utilized a specific threshold of completeness as an inclusion criterion for fossil taxa, a criterion that has now become widespread among paleontologists. For example, Fraser and Benton (1989) successively deleted taxa from their analysis of sphenodontids based on percentage of missing data and showed results for each of these analyses. Gao and Norell (1998) pursued a similar strategy, and it is repeated in many subsequent studies. This approach involves two important assumptions—one, that ambiguity of results is solely attributable to missing data in fragmentary fossil taxa and, two, that the proportion of missing data in a taxon is correlated with degree of ambiguity.

Nixon and Wheeler (1992) elucidated the computational problem that may occur when data matrices that contain missing entries are subjected to phylogenetic computer programs. Certain incompletely known taxa may "float" into many different positions during tree searches due to alternative optimizations of question mark entries by computer algorithms. Such taxa were termed "wildcards" by Nixon and Wheeler (1992). In some cases, large numbers of 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 problems are not necessarily restricted to fossils, but may occur whenever missing data are concentrated within a single taxon, such as when data sets are combined and one subset of character data is unknown for a taxon (Carpenter, 1987). Subsequent to the Nixon and Wheeler (1992) paper, some researchers began to cite the "wildcard problem" as a reason to exclude incomplete taxa that increase the percentage of missing data in a matrix, again typically using exclusion strategies for taxa based upon a cutoff percentage of missing character data. However, what constitutes a "wildcard" has become another assumption-laden problem and it seems that the catchword "wildcard" simply replaced the catchphrase "missing data" at about this time as a vague justification for excluding certain taxa in some studies.

Wilkinson (1995b) more strongly characterized missing data as a significant problem to phylogenetic analyses combining fossil and living taxa, and discussed possible solutions to this problem. He suggested that missing data obfuscate relationships by causing the generation of so many equivalent trees that the identification of non-ambiguous relationships becomes impossible. One solution offered by Wilkinson (1995b) was "Safe Taxonomic Reduction," a method that identifies redundant,

fragmentary taxa that cannot have an effect on the relationships of other taxa and then deletes those taxa from the data matrix (Fig. 2). This method is a useful tool for combined analyses as evidenced, for example, by its effectiveness in the study of Norell and Gao (1997). Norell and Gao (1997) analyzed varanoid relationships based on a data set that included 49% missing data and several highly fragmentary fossil taxa. Analysis of the complete data set resulted in 395 shortest trees and a poorly resolved strict consensus tree. Norell and Gao (1997) implemented Wilkinson's (1995b) Safe Taxonomic Reduction and identified one taxon as a taxonomic equivalent. Removal of that single taxon and reanalysis resulted in 9 trees and a much more highly resolved strict consensus.

Wilkinson (1994, 1995c) also pursued other strategies, including the use of alternative consensus methods in cases where strict consensus trees are poorly resolved, presumably due to missing data problems. Such alternative consensus methods aim to target and remove unstable taxa and thereby improve resolution of relationships in consensus trees. These methods seem not to have been widely adopted among paleontologists, but the papers bolstered the burgeoning viewpoint that missing data pose a serious problem to combined phylogenetic analyses and that new methods and solutions are required to solve such problems.

Recently, Grande and Bemis (1998) discussed problems associated with missing data extensively and gave perhaps the most damning critique of the issue. They suggested that "These programs in theory choose the character state that provides the most parsimonious distribution of known characters and in doing so this method of analysis logically decreases the empirical quality of a data matrix (i.e., each question mark is effectively assigned a character state) . . . Combining largely incomplete taxa with well-preserved taxa in computerized phylogenetic analyses increases methodological circularity" (Grande and Bemis, 1998:569). Also: "There is, no doubt, information present even in those taxa that have question mark entries. The challenge is to weigh the positive effect of that information against potentially misleading effects of question marks on the phylogenetic program being used for data analysis. This suggests that some factor of completeness should be considered when choosing taxa for computer analysis. . . ." (Grande and Bemis, 1998: 570).

Beyond the papers focusing mainly on resolution issues listed above, the "missing data problem" has also been addressed

```
A      0001120100111221
B      ?????2??0???1???      redundant fragmentary taxon, can be safely deleted from matrix
C      ?????0??1???0???      non-redundant fragmentary taxon, cannot be safely deleted from matrix
D      0011201000121112
```

FIGURE 2. Fragmentary Taxon B is a redundant, taxonomic equivalent of Taxon A and can be safely removed from the data matrix according to the rules of STR without affecting the placement of other taxa in the data matrix. Fragmentary Taxon C contains an equivalent amount of missing data, but is not a taxonomic equivalent and cannot be safely removed from the data matrix since it has the potential to affect topological relationships among other taxa.

with simulation studies that have assessed the effects of missing data in relation to phylogenetic accuracy (Huelsenbeck, 1991; Wiens and Reeder, 1995; Wiens, 1998; Matthee et al., 2001). Conclusions reached in these papers vary—that completeness of taxa improves phylogenetic resolution and that inclusion of incomplete taxa may sometimes decrease the probability of finding the correct tree relative to the inclusion of complete taxa (Huelsenbeck, 1991); that incomplete taxa should be included because they do not significantly detract from phylogenetic accuracy (Wiens and Reeder, 1995); and that inclusion of incomplete characters can increase phylogenetic accuracy up to a point, but can decrease accuracy in some cases and thus should not always be uncritically included (Wiens, 1998). All of these studies rely on the supposition that question marks can be positively misleading.

It seems clear from the above history that the "missing data problem" has transformed significantly over the last 15 years from the realization that missing entries may pose operational problems with phylogenetic computer algorithms, to the idea that incomplete taxa cause inordinate ambiguity and should be avoided, and even to the idea that missing entries are positively misleading. On the other side of the coin, two seminal papers (Doyle and Donoghue, 1987; Gauthier et al., 1988) advocated the inclusion of fossil taxa in phylogenetic analyses in spite of their less complete nature, partly because of their potentially unique capacity to retain plesiomorphic character states. It remains to be seen whether the criticisms leveled against missing data in fossil taxa have overshadowed the importance attributed to fossil taxa in the latter papers.

HOW COMPUTER ALGORITHMS FOR PHYLOGENETIC ANALYSIS WORK

Much of the discussion of the missing data problem concerns the results of computer-assisted searches for most parsimonious cladograms, so we will briefly describe some general aspects of these searches. For a more detailed discussion of quantitative phylogenetic analysis see Farris (1970), Fitch (1971), Swofford and Maddison (1987), Goloboff (1993, 1996), and Nixon (1999a). The programs, such as Hennig86 (Farris, 1988), PAUP* (Swofford, 2001), NONA (Goloboff, 2000), and TNT (Goloboff et al., 2002) are "black boxes" to the majority of people using them, in part because the precise algorithms are not published for the most commonly used programs. However, the primary goal of all of these programs is the same, so that the same optimality criterion is used during tree searches.

The primary goal of parsimony analysis is to obtain the set of relationships that most economically explains the distribution of characters among taxa (Farris, 1983). In quantitative analysis, the most parsimonious explanation is the one that minimizes the number of "steps" (changes in character states) implied by the topology of the cladogram for characters distributed among taxa. For a binary (two state) character a single step implies no homoplasy—from the plesiomorphic (more general) condition to the apomorphic (less general) condition; any steps beyond this imply non-homology (homoplasy). The total number of steps for a given cladogram is quantified as its Length. Thus, a cladogram requiring no homoplasy for a set of binary characters will have a length equal to the number of characters.

The concept of parsimony was first explicitly applied to the construction of evolutionary trees by Camin and Sokal (1965), in discussing minimum spanning trees. These trees simply minimize the number of steps between terminal taxa. In 1970, Farris et al. presented an algorithm for calculating most parsimonious cladograms, called Wagner Trees, in which the length of the tree is minimized not between taxa, but between the nodes connecting them. When constructing Wagner trees, nodes are treated like terminal taxa, so that a state for each character can be inferred for each node. The characters inferred for a node are interpretable as the ancestral characters of the terminal taxa grouped within it (Farris et al., 1970).

Cladograms with the same nodes can be generalized as simpler topologic structures termed unrooted networks. Thus, different cladograms identifying alternative groups of taxa but in which the same nodes occur between the same taxa can be generalized in a network diagram specifying the nodes but not the direction, or "root," of the cladogram (Fig. 3). Because the lengths of the cladograms derived from a single network are identical, so that all of the cladograms of the most parsimonious network will also be most parsimonious cladograms, it is simpler to search for a network and then "root" the network than it is to search for all of the individual cladograms. This is feasible because cladogram length is independent of the location of the root.

All computer algorithms that search for most parsimonious cladograms therefore operate using the same criterion to identify these cladograms: the length of unrooted networks. Length is calculated as the minimum number of steps between all nodes and terminal taxa over the network. The calculation of minimum network length requires the specification of at least a range of possible character states at the nodes, but does not require a specific state attribution (Goloboff, 1993). Calculation of cladogram length also does not require the specification of states for missing entries, and indeed, searches would be much slower if they were specified.

Cladogram searches ideally lead to the discovery of all topologies of equal length, but an additional consideration is whether character support exists for all of the groups indicated by each topology (Platnick et al., 1991; Coddington and Scharff, 1994; Wilkinson, 1995a; Nixon and Carpenter, 1996b). A group supported by a zero length branch, along which no character state changes, lacks any evidential support for the taxon and is considered by most systematists to be an artifact. Groups with zero length branches are found by some computer searches when there are mutually exclusive character optimizations at the node, which may involve missing data but need not (Wilkinson, 1995a) (Fig. 4). The problem of zero length branches is due to ambiguity in optimization, not to missing data per se.

Swofford and Begle (1993) identified three possible rules for treating zero length branches after a cladogram search. The first
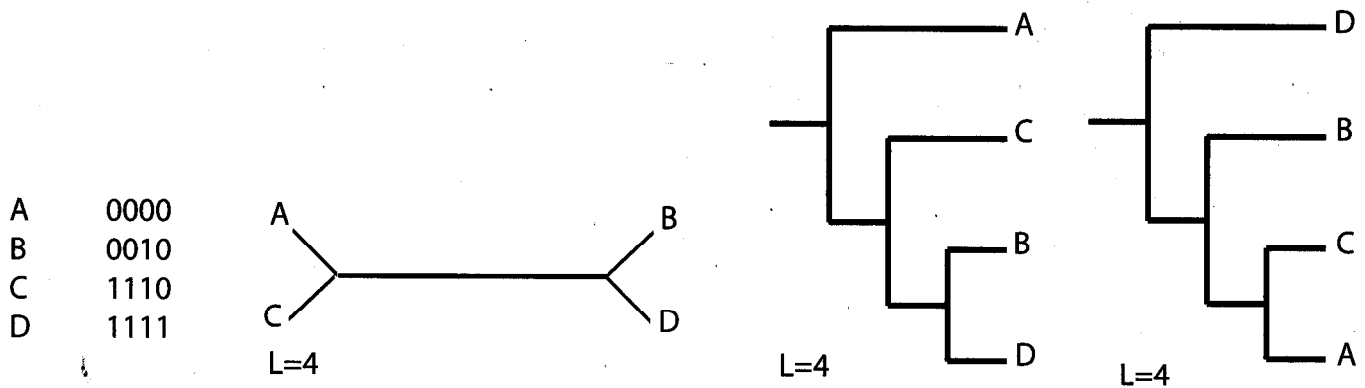
FIGURE 3.   An unrooted network for four taxa based on a simple data set and the two cladograms resulting from rooting the network either on Taxon A or Taxon D. Both cladograms have a length of four, equivalent to the length of the unrooted network.

would collapse an interior branch if the minimum possible length is zero (i.e., it is zero under any one of the possible optimizations). The second is to choose a specific optimization procedure (e.g., Acctran) and collapse those zero length branches found under this optimization criterion only. The third is to collapse an interior branch if the maximum possible length of the branch is zero (i.e., if it is never more than zero under any of the possible optimizations). Coddington and Scharff (1994: 420) suggested a fourth rule, "discard all trees that must contain a zero-length branch," because simply collapsing a zero length branch may not necessarily result in a cladogram with all groups fully supported. The third option is clearly justified, but the second requires justification for the particular optimization procedure and the first excludes all branches with character support only under some optimizations. Nixon and Carpenter (1996b) argued that groups supported under only some optimizations should not be considered, and that only strictly supported branches should be reported.

Some computer programs output all of the possible resolutions regardless of zero length branches (e.g., Hennig86), whereas others (e.g., NONA, PAUP*) allow cladograms with

zero length branches to be ignored (under the "amb-" option). Zero length branches can also be identified (and filtered out, if so desired) using programs for optimizing characters onto cladograms, such as MacClade (Maddison and Maddison, 1992) and Winclada (Nixon, 1999b), although this may be labor intensive.

Zero length branches are problematic for computer algorithms searching for most parsimonious networks because they may be optimization-dependent. To maximize efficiency, and therefore speed, it is faster to use the length of networks as the sole selection criterion during initial tree searches, regardless of how characters optimize on the rooted cladograms. To do otherwise would severely limit the speed of the programs, and therefore the size of the data sets that can be analyzed.

Thus, one potential problem that missing data can introduce to the most widely used computerized phylogenetic algorithms centers around the fact that "?" entries may increase the number of potentially ambiguous character optimizations. If there is a "real" missing data problem then it is this—their contribution to the zero-length branch problem. If zero-length branches are not suppressed, every possible optimization for a "?"
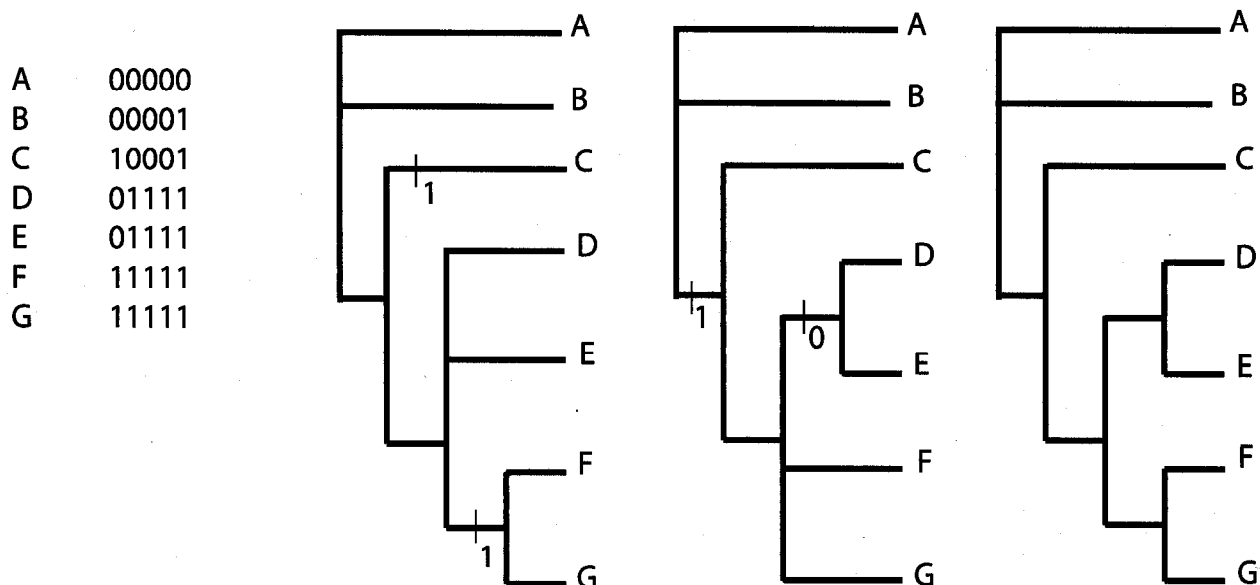

FIGURE 4.   The problem of zero-length branches is a result of alternative, mutually exclusive optimizations rather than missing data. In this example with no missing data, the groups (DE) and (FG) are supported under different optimizations of character 1, but not simultaneously in the same cladogram as is shown in cladogram 3. Either character state 0 or character state 1 is plesiomorphic for (DEFG), but not both simultaneously. Modified from Coddington and Scharff (1994).

TABLE 1. Some recent strategies used for solving the "missing data problem."

A priori
1. Do not combine fossil and living taxa.
2. Exclude characters that cannot be scored in both fossil and extant taxa.
3. Delete fossil taxa according to percentage of missing character data.
4. Use Safe Taxonomic Reduction to delete taxonomic equivalents.
5. Construct composite taxa.

A posteriori
7. Identify and delete wildcard taxa and re-run analysis.
8. Use a less-than-strict consensus method.

entry may be considered with certain algorithms. The Nixon and Wheeler (1992) data set (Fig. 1) is a good example of this problem. One potential solution to missing data or wildcard problems, then, is to filter out all but strictly supported clado- grams, which are those cladograms with no zero-length branch- es (Nixon and Carpenter, 1996b). However, this is complicated by the fact that strict consensus trees including wildcard taxa may obscure supported groups (Wilkinson, 1995b; Kearney, 2002).

PREVIOUS APPROACHES TO THE "MISSING DATA PROBLEM" AND THEIR DRAWBACKS

The paleontological community has responded to the wide- spread discussion of "missing data problems" described above by adopting new approaches to increase resolution of taxonom- ic relationships in data sets that include incomplete fossil taxa (see Wilkinson, 1995b; Kitching et al., 1998; and Grande and Bemis, 1998 for recent discussions). These approaches to the missing data problem can be conceptually divided into a priori and a posteriori approaches (Table 1). All of these approaches have emphasized the topological consequences of missing data and the criterion of resolution, with relatively little regard for the exact causes of these changes in topology or resolution, and many of these strategies can only resolve ambiguity at the ex- pense of concealing or ignoring relevant data. Ultimately, this is because these approaches do not distinguish between ambi- guity caused by lack of data and ambiguity caused by character conflict (Kearney, 2002).

Not Combining

The first approach is to simply not combine fossil and living taxa at all in order to avoid question marks in the data matrix. There are several obvious problems with this approach. First, fossils can be critical to the phylogenetic analysis of living taxa (Doyle and Donoghue, 1987; Gauthier et al., 1988) and vice versa. Second, although increased percentages of missing data and ambiguity of results are the reasons most often cited for not combining, in practice this approach does not necessarily decrease ambiguity, so that information may be sacrificed for no reason. For example, Norell and de Queiroz (1991) found that the inclusion of two very fragmentary fossil taxa (scored for less than 35% of the total characters in their data matrix) in their analysis of iguanine lizards provided increased resolu- tion despite causing a significant increase in the percentage of question marks in the data matrix. Other examples where a fragmentary taxon increased, rather than decreased, resolution in relationships among more complete taxa can be readily found (e.g., Wilkinson and Benton, 1995).

Thus, extent of missing data in a data matrix is not a general predictor of degree of resolution, and a few compatible char- acters may provide greater resolution than many conflicting

characters. Resolution of relationships depends on the exact dis- tribution of question marks, character congruence, and homo- plasy in the data matrix and is therefore always matrix-specific (Novacek, 1992; Kearney, 2002). In any case, certain evolu- tionary questions involving groups with living and fossil mem- bers can only be resolved with a combined analysis so that missing data simply cannot be avoided in this manner. We rec- ommend that, whenever possible, both extant and fossil mem- bers of a group be considered in phylogenetic analyses.

Excluding Unscorable Characters

Another common strategy to avoid missing data is to com- bine living and fossil taxa but exclude those characters that cannot be scored for both (e.g., soft tissue characters, behavioral characters, non-preserved osteological characters). An example of this approach can be found in the study of Wu et al. (1996), in which the authors were attempting to ascertain the systematic position of a newly discovered fossil. They utilized a previous data set (that of Estes et al., 1988) but excluded all the char- acters from that data set that could not be scored for the new fossil taxon because of concern about the effects of missing data. For their reduced data set, they obtained 28 most parsi- monious trees and the strict consensus tree shown in Figure 5a. However, inclusion of the soft tissue characters in this analysis results in a decrease in the number of most parsimonious trees (to 6) and a more resolved strict consensus tree (Fig. 5b) despite an overall increase in the percentage of question marks in the data matrix.

The decreased resolution found when characters with missing data are excluded should not be surprising because data resolve relationships. Excluding character data cannot increase resolu- tion unless it lessens character conflicts, which is generally un- likely. This approach is not only misguided, it is counter-pro- ductive, as it will generally yield less resolution. It should be noted that this strategy is becoming commonplace in paleon- tological studies to the point where certain classes of data (usu- ally non-osteological) are often excluded with little justification other than the implicit assumption that missing data should be avoided. Ironically, this approach limits many phylogenetic studies that include fossils to the overutilized and limited set of osteological characters that has been unable to resolve the re- lationships of many groups satisfactorily, especially those groups in which there appears to be a high degree of homoplasy among these characters. We recommend that, whenever possi- ble, all relevant character data, regardless of degree of appli- cability to all taxa, be included in phylogenetic analyses.

Deleting Taxa According to Percentage of Question Marks

Several researchers have explored the approach of deleting fossil taxa according to the percentage of question marks they contribute to the data matrix (Rowe, 1988; Fraser and Benton, 1989; Benton, 1990; Gao and Norell, 1998; Grande and Bemis, 1998). With this approach it is typically demonstrated that when fragmentary taxa are included in analyses increased ambiguity may result, and this is used as a basis for setting minimum threshold levels of completeness for the inclusion of fossil taxa.

Fraser and Benton (1989) successively deleted taxa based on percentage of missing data in their analysis of sphenodontids and illustrated results for each of these analyses (Fig. 6). The unpredictability of this strategy is apparent in that numbers of trees did not decrease linearly as fragmentary taxa were deleted; more trees were produced in Analysis 3 than in Analysis 2, with an increase in ambiguity despite a decrease in missing data at this particular step. Additionally, the last analysis, while re- sulting in a single, completely resolved cladogram, was limited to only 5 of the original 15 taxa, so that very little could be concluded regarding the original question asked. Fraser and
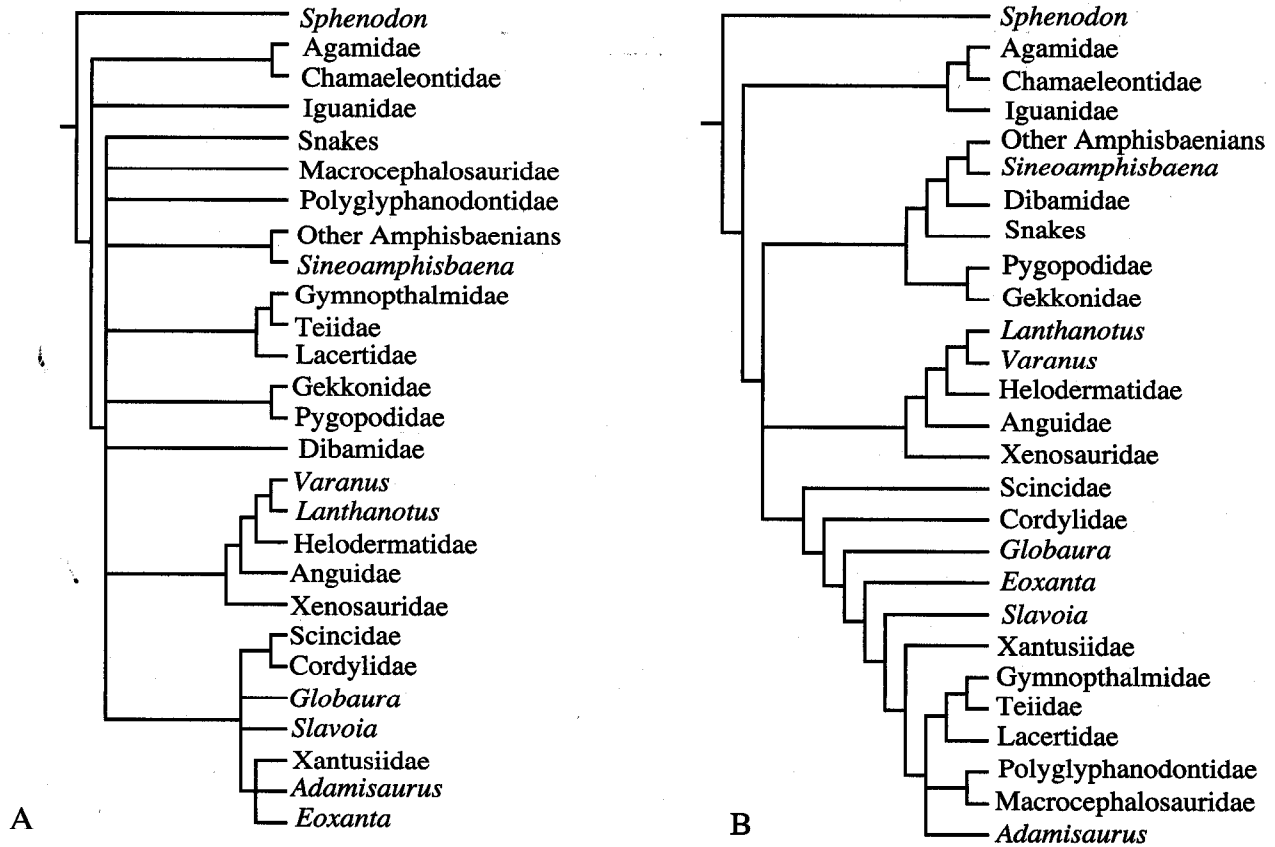
FIGURE 5. **A**, strict consensus of 28 most parsimonious cladograms from Wu et al. (1996), resulting from analysis of a reduced data set that excluded characters that could not be scored for the fossil taxon of interest and contained 5.6% missing data. **B**, strict consensus of 6 most parsimonious cladograms from Wu et al. (1996), resulting from analysis of the complete data set containing 18.9% missing data.

Benton's (1989) phylogenetic conclusions were equivocal because of their concerns regarding missing data and lack of resolution in some trees; however, as in the other studies listed above, all ambiguity was assumed to be attributable to missing data in the fragmentary fossils.

Using quantity of preserved character data in a taxon as the criterion for its inclusion is ill-advised for several reasons. First, there is no objective way to designate a cutoff threshold for missing data. Indeed, the spectrum of cutoff levels for missing data from just a few recent studies indicates a high degree of subjectivity: 12% (Rowe, 1988); 33% (Benton, 1990); 36% (Grande and Bemis, 1998); 45% (Ebach and Ahyong, 2001). Second, quantity of missing data is not directly associated with the information content of a taxon. Consider the Norell and de
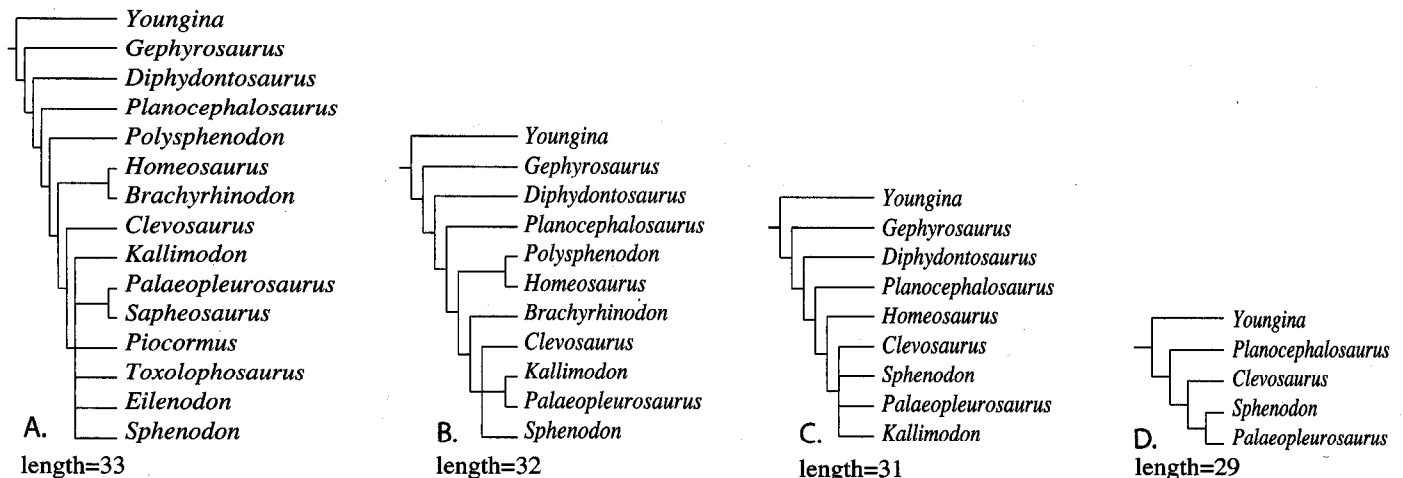

FIGURE 6. Analysis of Fraser and Benton (1989). **A**, strict consensus of 82 most parsimonious cladograms based on analysis including all taxa. **B**, strict consensus of 2 most parsimonious cladograms based on analysis excluding four taxa with greater than 50% missing data. **C**, strict consensus of 4 most parsimonious cladograms based on analysis excluding two additional taxa (all taxa with greater than 80% missing data excluded). **D**, single most parsimonious cladogram based on analysis including only completely scored taxa in data matrix.

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 ???
```
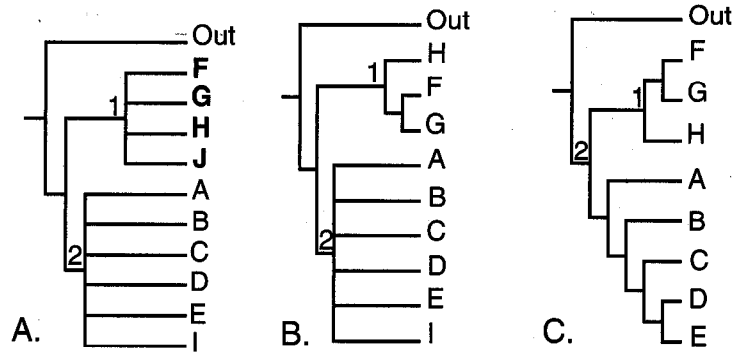
FIGURE 7. Illustration of Safe Taxonomic Reduction (Wilkinson, 1995b) (taken from Kearney, 2002). **A,** analysis of the complete data matrix results in a strict consensus tree containing the clades 1 and 2. STR identifies Taxon J as a taxonomic equivalent of taxa F, G and H, and thus J can be "safely removed" from the matrix. **B,** reanalysis excluding Taxon J 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, still causes ambiguity. **C,** completely resolved cladogram if Taxon I were also removed from the analysis.

Queiroz (1991) study mentioned previously—if they had used a similar cutoff threshold based on percentage of missing data, they would have lost the increased resolution provided by two highly incomplete fossil taxa (70% and 68% missing data). Third, even if the inclusion of fragmentary taxa does result in decreased resolution, this may be due to character conflict and not necessarily to missing data, and deleting such taxa may therefore conceal alternative evidence for groups. In sum, the percentage approach assumes that fragmentary taxa will always increase ambiguity and, further, that all ambiguity contributed by fragmentary taxa is due solely to missing data. We recommend that taxa not be excluded a priori from phylogenetic analyses based on the criterion of number of preserved characters.

Deleting Taxonomic Equivalents

Wilkinson (1992, 1995a) pointed out the problems involved in simply removing fragmentary taxa from analyses based on quantity of missing data and proposed instead a strategy of Safe Taxonomic Reduction (STR). STR, which is currently implemented in the TAXEQ3 computer program (Wilkinson, 2001), identifies taxonomic equivalents in a data matrix and deletes them. Taxonomic equivalents contribute large amounts of missing data and also completely overlap in character states with other taxa in the matrix, thus contributing no unique information. Deletion of such taxa may drastically reduce the number of equivalent trees (Fig. 7).

STR targets fragmentary taxa that are completely congruent with other taxa and that are also missing data that could place them more specifically. When such taxa are deleted, more apparent resolution obtains because redundant taxa are pruned from nodes shared with their equivalents. This is in contrast with the example given above by Nixon and Wheeler (1992) (Fig. 1), in which a taxon is unstable not simply due to redundancy and missing data, but to a mixture of homoplasy and missing data. If this latter type of wildcard problem exists in real data sets, it cannot be solved by STR because these types of wildcard taxa are not taxonomic equivalents. However, it is thus far unknown which types of wildcard problems occur predominantly in combined analyses. Kearney (2002) reviewed some recent combined analyses and found that most encountered problems that were related to unrecognized redundant terminals that could be deleted from the analyses as described by Wilkinson's (1995b) method. This method, therefore, may de-

serve greater attention as a tool for phylogenetic studies combining living and fossil taxa.

Using Composite Taxa

A common practice in vertebrate paleontology is to combine poorly known lower taxa into a single higher taxon with a greater proportion of the characters scored (e.g., Sereno, 1999). Thus, a species in which only the skull is known might be combined with another from the same genus in which only the post-cranial skeleton is known, and the genus then treated as a single taxon for which the entire skeleton is known. This certainly lessens the proportion of missing data and any problems potentially caused by it, but it obviously circumvents testing the monophyly of the composite group. For example, in extreme cases such as the skull and post-cranial taxa mentioned above, there may be evidence placing the two specimens or species in different taxa. Furthermore, unless characters are coded polymorphically, it requires a priori decisions to be made when there is variation within the ingroup in a character, rather than searching for the most globally parsimonious solution. This is a less than ideal solution that hides potential problems rather than exposing them.

Deleting Wildcards

Several authors have suggested that wildcard taxa may require removal from analyses when huge numbers of trees are produced in order to preserve resolution (Nixon and Wheeler, 1992; Wilkinson, 1995b). With this strategy, wildcard taxa are identified subsequent to cladistic analysis and then pruned from the tree in order to glean more resolution. But fragmentary taxa do not always behave as wildcards and wildcard behavior is not always caused by missing data (Fig. 8). In Figure 8a, inclusion of the fragmentary taxon F increases resolution despite its incompleteness. In Figure 8b,c, inclusion of F causes a wildcard scenario, but for different reasons in each case: in b it is due to character conflict, whereas in c it is due to missing data. Thus, taxa may be unstable (behave as wildcards) due to missing data, due to character conflict, or due to both (as in Nixon and Wheeler's example). The deletion of wildcard taxa, like most of the other strategies pursued to date, is an oversimplified response to the missing data problem because it focuses on resolution of results and does not consider different causes of ambiguity.
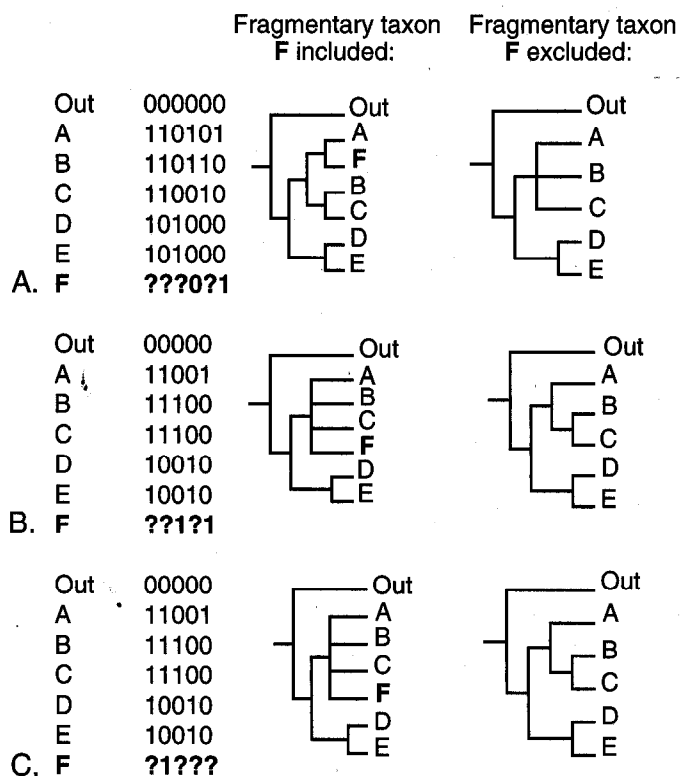
FIGURE 8. Fragmentary taxa and ambiguity (taken from Kearney, 2002). **A**, inclusion of fragmentary taxon F increases resolution among other taxa. **B**, inclusion of fragmentary taxon F increases the number of most parsimonious trees from 1 to 4. Ambiguity in the strict consensus tree is due to character conflict among taxa A, B, C and F. C, the same strict consensus is obtained when fragmentary taxon F is included, but for a different reason—F contains sufficient information to place it in the group A(BC), but question marks cause it to behave as a wildcard within that clade, thus the ambiguity is caused by missing data. The supported (BC) clade is obscured by the wildcard effect of taxon F.

Using A Less-Than-Strict Consensus Method

Whether strict consensus methods (Schuh and Farris, 1981; Schuh and Polhemus, 1981) are too sensitive in regard to conflict among primary trees has been the topic of considerable debate and some have advocated alternative methods such as majority rule consensus trees (Margush and McMorris, 1981), combinable component or semi-strict consensus trees (Bremer, 1990), reduced consensus trees (Wilkinson, 1994), common pruned trees (Gordon, 1980), and Adams consensus trees (Adams, 1972). Swofford (1991) reviewed these methods and argued against the use of strict consensus trees because they do not preserve enough structure found among the primary trees. In contrast, Nixon and Carpenter's (1996b) review of consensus methods led to the opposite conclusion: that all methods other than strict consensus are "compromise methods," methods which should be eschewed because they may not accurately reflect the agreement and disagreement in grouping among all the primary trees.

Certain alternative consensus methods have been advocated specifically for dealing with the missing data effects of incomplete taxa. The clear advantage these methods have over omitting taxa is that they include all taxa that may affect the topology of the tree. The Adams consensus method identifies unstable taxa and collapses nodes corresponding to different positions for those taxa in the primary cladograms to the first node that includes those alternative placements (Fig. 9a). In an Ad-

ams consensus, wildcards will not obscure supported groups. However, Adams trees may contain resolved groups that are contradicted in some of the primary trees, and do not distinguish between taxa that are unstable due to conflicting data vs. lack of data.

Gordon's (1980) "common pruned trees" (Fig. 9b) and Wilkinson's (1994, 1995c) "reduced consensus" methods (Fig. 9c) are variants of a general "taxon pruning" approach. (Anderson's [2001] "phylogenetic trunk" method seems to be fundamentally indistinguishable from these earlier methods in that it relies on degree of taxon instability to prune taxa from consensus trees.) Unlike the Adams consensus, these consensus trees contain fewer taxa than the primary cladograms (although the possible placements for pruned taxa may be annotated in some manner or the pruned taxa may be regrafted). Increased resolution is obtained by pruning one or more taxa from the consensus tree until a resolved topology is acquired, rather than collapsing unstable taxa to the most basal inclusive node as in an Adams consensus. This increased resolution, however, comes at the expense of losing information from the primary trees (and by extension, therefore, from the data set). These
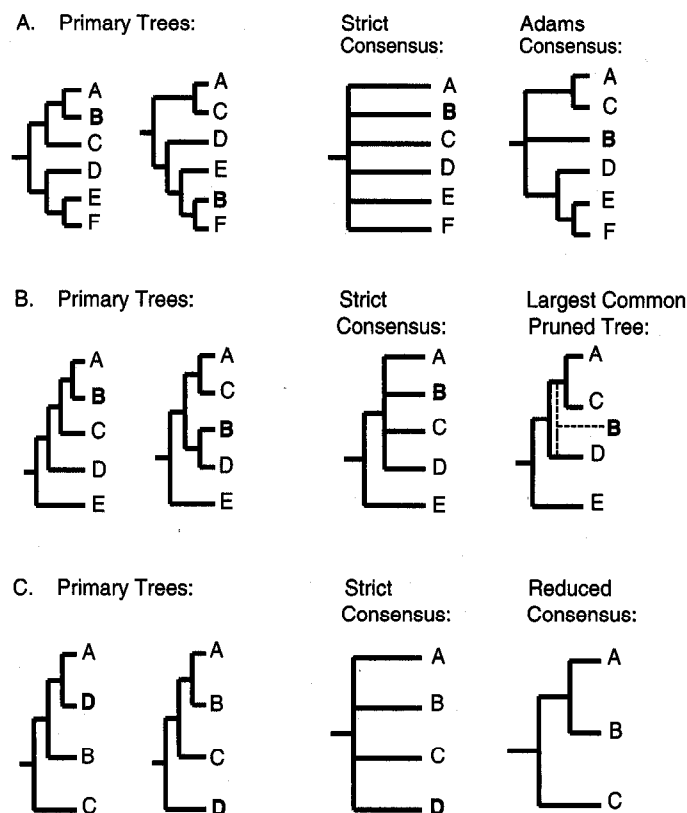

FIGURE 9. Alternative consensus methods proposed for decreasing ambiguity due to unstable taxa (taken from Kearney, 2002). **A**, the strict consensus of the two primary trees is completely unresolved due to the different possible positions of taxon B. An Adams consensus of these two trees identifies the unstable taxon B and places it unresolved at the base of the clade. However, alternative evidence for the groups (AB) and (FB) exists among the primary trees and this is ignored, or concealed, by the Adams consensus. Also, the cause of Bs instability is unknown. **B**, the largest common pruned tree method searches for taxa to prune, leaving as much common resolution as possible. Taxon B is identified as unstable and pruned. The same criticisms from A above apply here except that the regrafted tree indicates more specifically the possible positions of B. **C**, a reduced cladistic consensus (RCC) method reveals those groups that are supported in every tree and excludes unstable taxa that may reduce resolution. Possible placements for excluded taxa may be annotated in some manner. The same criticisms apply.

TABLE 2. Comparison of levels of ambiguity in some recent studies with varying degrees of missing data and fossil taxa. Studies are listed in increasing order of percentage of missing data (from Kearney, 2002: table 1).

Study	Percentage of fossil taxa	Percentage of missing data	Number of primary trees
Novacek (1992)	26%	13%	>6,800*
Messenger and McGuire (1998)	2%	14%	>45,000*
Wu et al. (1996)	25%	19%	2
Grande and Bemis (1998)	98%	20%	>10,000*
Fraser and Benton (1989)	100%	21%	82
Gao and Norell (1998)	58%	34%	>32,000*
Norell and Gao (1997)	80%	49%	395
Gatesy et al. (1999)	0%	57%	6
O'Leary (1999)	75%	73%	30

*Search stopped at set limit of equivalent trees, so more primary trees actually exist.

methods may preserve supported structure that could be lost by the inclusion of wildcard taxa but they represent the "compromise trees" of Nixon and Carpenter (1996b)—they may either contain groups contradicted in some of the primary trees or they may not contain some supported groups present in the primary trees. And, like Adams trees or strict consensus trees, they do not distinguish between no-data-ambiguity vs. conflicting-data-ambiguity. We do not recommend using less-than-strict consensus methods as a strategy to obtain unambiguous results. Although these methods are useful in flagging wildcard taxa and for other heuristic purposes, they are incomplete summaries of the results of phylogenetic analysis.

We conclude that recent approaches for ameliorating the "missing data problem," being focused exclusively on resolution of results, have made some unsound assumptions. Resolution of results cannot in and of itself be a criterion for including or excluding data, else any evidence can be excluded until some arbitrary level of resolution is reached. Since the cause of ambiguity in results may itself be ambiguous, treating all ambiguity equally will most likely entail problems.

THE "MISSING DATA PROBLEM"—MISCONCEPTIONS, REAL PROBLEMS, AND POSSIBLE SOLUTIONS

The veracity of commonly made generalizations about incomplete fossil taxa and missing data in phylogenetic analysis remains unclear. For example, a simple and common notion is that adding incomplete taxa will increase ambiguity, but many studies that combine complete and incomplete taxa repudiate this assumption (Table 2). These examples show that, contrary to the negative connotations attached to missing data, combining incomplete fossils with living taxa is just as likely to yield a highly resolved cladogram than will a separate analysis with less missing data.

Another common misconception is that missing data produce artifactual resolution, and this is partially based on a semantic problem. In some discussions of the missing data problem, the "over-resolution" that can occur due to semi-strictly supported branches has been termed "spurious" resolution. For example, Kitching et al. (1998:82), when referring to the problem of ambiguous optimizations of missing entry cells, state "In other words, all of the branches resolving groups DEFG, DFG, DG and DF are spurious; none has unambiguous support in the data." The use of the term "spurious" here seems meant to convey that certain nodes may be supported only under some optimizations, but not others—in other words, that they are not strictly supported nodes. It is true that a semi-strictly supported tree may include taxa that are placed on the basis of how homoplastic characters or "?" entries can be optimized. This is far different, however, from the use of the term "spurious" that is sometimes used by other authors who seem to fear that missing data are actually building erroneous cladograms. For ex-

ample, recently Ebach and Ahyong (2001:5) interpreted the discussion of Kitching et al. (1998) as support for their exclusion of fragmentary taxa and avoidance of missing data: "The deletion of taxa with more than 45% missing data (denoted by "?") yielded fewer equally parsimonious trees than Analysis A. Similar results, observed by Kitching et al. (1998) using hypothetical data, led to the conclusion that missing data may not only increase the number of equally parsimonious trees, but also cause some cladistic computer programs to yield spurious cladograms. The exclusion of terminals with a high proportion of missing data as outlined in Kitching et al. (1998) is herein justified." The Kitching et al. (1998) discussion and figures therein did illustrate the problem of semi-strictly supported nodes. However, they did not mention that such over-resolution can be suppressed by choosing the "amb-" setting in PAUP* or by the default setting in NONA, and they did not advocate excluding incomplete taxa on that basis.

Another common misconception is that, during searches for most parsimonious trees, computer algorithms assume a particular state for each missing entry in each taxon. In reality, tree search procedures such as the Wagner algorithm (Farris, 1970) simply search for the trees with the shortest length computed on the basis of the minimum number of steps (i.e., changes in character states) a tree implies for all of the characters in a data set. Since missing data cannot add any changes, missing data add no length to the tree and are not considered in length calculations. Thus, no additional assumptions are required to interpret missing data, other than that the data missing from a matrix do not contradict the known data in the matrix (i.e., an assumption that all missing character data would be completely congruent with the known character data). But this assumption is no different than assuming that any data not presently known are consistent (do not conflict) with the known data—an assumption that can hardly be avoided.

Some Real Problems and their Solutions

Different Computer Programs Treat Zero Length Branches Differently—The various computer programs that are in popular use for phylogenetic analysis do not all treat zero length branches in the same manner. HENNIG86 (Farris, 1988) optimizes question marks as one of the possible states for a given character, using parsimony as a criterion for assigning those states, and reports all trees including semi-strictly supported branches. PAUP* (Swofford, 2001) by default treats question marks in the same manner as HENNIG86, but also contains an option (amb-) which can be implemented in order to suppress ambiguous resolutions of question marks. In NONA (Goloboff, 2000), the default option (amb-) will not yield semi-strictly supported trees. This is a stricter interpretation than HENNIG86's algorithm or PAUP*'s default, which will both resolve branches if they are supported only under some opti-

mizations. A "wildcard" taxon that floats among various positions in a cladogram will result in alternative resolutions involving branches that have a minimum length of zero. In other words, these are semistrictly supported branches, analogous to those produced by the ambiguous optimizations that can occur due to homoplasy. Using NONA or using PAUP* with the "amb-" option will suppress these arbitrary resolutions. The solution to this particular problem, therefore, is to allow only strictly supported branches and to use only strict consensus trees.

Fragmentary Taxa May Sometimes Behave as Wildcards and Result in Large Numbers of Equivalent Trees and Unresolved Consenses of Those Trees—In some analyses containing fragmentary fossil taxa, such taxa may behave as wildcards and very large numbers of most parsimonious trees may be produced, along with a poorly resolved strict consensus of those trees. When such an analysis yields so many primary trees that examination of each tree to determine the cause of the ambiguity is intractable, some other solution will be required. This solution cannot be the simple use of a less-than-strict consensus method because such methods obscure evidence for alternative groups (Nixon and Carpenter, 1996b) and do not discriminate between different causes of ambiguity (Kearney, 2002). A better solution is to use a combination of Wilkinson's (1992, 1995b) STR method to identify redundant taxa and reduce the number of equivalent trees, as well as methods to assess character support or lack thereof in strict consensus trees (Kearney, 2002).

Support Measures—The effect of missing data on character support for groups has been of less concern. Nevertheless, character support for groups is considered by many to be an integral component of phylogenetic analysis and little attention has been paid in that context to the effect that missing data can have on various support measures such as bootstrap and Bremer support. Missing data may indeed affect statistical support values and some researchers (e.g., Norell and Wheeler, 2003) are attempting to "correct for" this effect. The point of these methods seems to be to correct for the implicit assumption of complete congruence of missing data cells with known data. Unfortunately, such corrections rely on random replacement of missing data cells with unobserved homoplastic character state values, so that these methods simply correct for one assumption (complete congruence of unobserved data) by using another (randomly incongruent unobserved data). While it is true that the unknown data (if they were known) would most likely contain some homoplasy, the pattern of homoplasy that they would contain is something we cannot surmise.

A different approach to correcting for the effect of missing data on support measures can be found in the double decay analysis introduced by Wilkinson et al. (2000). Double decay analysis is a technique that provides an alternative support measure to the traditional decay index of Bremer (1988), and is based upon reduced consensus methodology (Wilkinson, 1994). This technique addresses the problem of artificially altered support measures caused by taxa with many missing entries, but is naturally subject to the same criticism leveled at all other methods which focus on the removal of unstable taxa—that is, it does not discriminate between taxon instability due to missing data vs. conflicting data. Double decay analysis is a logical extension of Wilkinson's (1994) reduced consensus methods, and is therefore also contingent upon the logic of those methods.

Beyond the Computational—Highly Incomplete Taxa May Make Homology Hypotheses Less Rigorous—Beyond the purely operational considerations discussed above, there may be important theoretical issues related to missing data in incomplete fossil taxa, and such issues have been almost completely ignored in deference to computational problems. Nixon

(1996) argued that primary homology assessments are weaker in fossil taxa because we usually do not have the same level of detail in structure, topology and variation as we do for living taxa. It is true that the classical criteria used for assessing homology (e.g., Remane, 1952) may not be as readily available in fragmentary taxa as in more complete taxa and that other, more subsidiary criteria may have to be invoked in such cases. If the less complete nature of fossil taxa simply means that fewer homology statements (characters) can be made for them, then the only danger is that incomplete taxa may be more difficult to place phylogenetically due to limited character data. If, on the other hand, it means that homology statements made for fragmentary taxa are more ambiguous than those made for more complete taxa, then this is a more serious concern because the quality of homology statements (as assessed by such factors as independence and non-arbitrariness, for example), and not just the number of such statements, is an important aspect of testability (Rieppel and Kearney, 2002). However, this same high level of scrutiny for homology statements might also be leveled in the other direction—extant taxa that are highly transformed may pose the same problem, i.e., homology hypotheses may be less rigorous due to transformation and consequent difficulties in identifying topological equivalence (Gauthier et al., 1988). For example, this would include molecular sequence regions that are difficult to align and positional homologies of supernumary digits. In both extant in extinct taxa, lack of context can make primary homology hypotheses more ambiguous.

SUMMARY

The phrase "missing data" remains more of a buzzword than a well-understood problem and taxa with missing data have become stigmatized. The phrases "missing data problem" and "wildcard problem" have become part of the systematic lexicon and have led to misguided attempts to solve a problem which may be somewhat of a red herring. Thus, the phrase "adding missing data" is an oxymoron; one cannot "add" missing data to an analysis because one cannot "add" something that does not exist. If one adds a taxon to a data matrix which is scored for one character and contains question marks for all other characters, one has added one more observation to the analysis. Of course, one could say that one has added question marks to the data matrix as well, but question marks are simply indicators of the unknown and they do not build trees. To be sure, operational problems associated with the treatment of missing entries by phylogenetic computer programs exist, but these do not outweigh the rigor of bringing as much evidence as possible to bear on a phylogenetic question. We recommend an approach where all relevant characters and taxa are included in a simultaneous analysis, degree of missing data is not used as an inclusion or exclusion criterion for characters or taxa, and specific tools such as the identification of zero length branches and redundant taxonomic equivalents, and the assessment of character support for nodes in consensus trees, are used for assessing causes of ambiguity and dealing with them.

ACKNOWLEDGMENTS

LITERATURE CITED

Adams, E. N. 1972. Consensus techniques and the comparison of taxonomic trees. Systematic Zoology 1:390–397.

Anderson, J. S. 2001. The phylogenetic trunk: maximal inclusion of taxa with missing data in an analysis of the Lepospondyli (Vertebrata, Tetrapoda). Systematic Biology 50:170–193.

Ax, P. 1987. The phylogenetic system: the systematization of organisms on the basis of their phylogenesis. Wiley, New York, 340 pp.

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.

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

——— 1990. Combinable component consensus. Cladistics 6:369–372.

Camin, J. H., and R. R. Sokal. 1965. A method for deducing branching sequences in phylogeny. Evolution 19:311–327.

Carpenter, J. M. 1987. Phylogenetic relationships and classification of the Vespinae (Hymenoptera: Vespidae). Systematic Entomology 12:413–431.

Coddington, J., and N. Scharff. 1994. Problems with zero-length branches. Cladistics 10:415–424.

de Pinna, M. C. C. 1991. Concepts and tests of homology in the cladistic paradigm. Cladistics 7:367–394.

Donoghue, M. J., J. A. Doyle, J. Gauthier, A. G. Kluge, and T. Rowe. 1989. The importance of fossils in phylogeny reconstruction. Annual Review of Ecology and Systematics 20:431–460.

Doyle, J. A., and M. J. Donoghue. 1987. The importance of fossils in elucidating seed plant phylogeny and macroevolution. Review of Paleobotany and Palynology 50:63–95.

Ebach, M. C., and S. T. Ahyong. 2001. Phylogeny of the trilobite subgenus *Acanthopyge* (*Lobopyge*). Cladistics 17:1–10.

Estes, R., K. de Queiroz, and J. Gauthier. 1988. Phylogenetic relationships within Squamata; pp. 119–282 *in* R. Estes and G. K. Pregill (eds.), Phylogenetic Relationships of the Lizard Families. Essays Commemorating Charles L. Camp. Stanford University Press, Stanford, California.

Farris, J. S. 1970. Methods for computing Wagner Trees. Systematic Biology 19:83–92.

——— 1983. The logical basis of phylogenetic analysis; pp. 7–36 *in* N. I. Platnick and V. A. Funk (eds.), Advances in Cladistics, Vol. 2. Columbia University Press, New York.

——— 1988. HENNIG86, version 1.5. Program and documentation. Port Jefferson Station, New York.

———, A. Kluge, and M. J. Eckhardt. 1970. A numerical approach to phylogenetic systematics. Systematic Zoology 19:172–191.

Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Systematic Zoology 20:406–416.

Fraser, N. C., and M. Benton. 1989. The Triassic reptiles *Brachyrhinodon* and *Polysphenodon* and the relationships of the sphenodontids. Zoological Journal of the Linnean Society 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 anguimorphan lizards. American Museum Novitates 3230:1–51.

Gatesy, J., P. O'Grady, and R. H. Baker. 1999. Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics 15:271–314.

Gauthier, J. 1986. Saurischian monophyly and the origin of birds. Memoirs of the California Academy of Sciences 8:1–56.

———, A. G. Kluge, and T. Rowe. 1988. Amniote phylogeny and the importance of fossils. Cladistics 4:105–208.

Goloboff, P. A. 1993. Character optimization and calculation of tree lengths. Cladistics 9:433–436.

——— 1996. Methods for faster parsimony analysis. Cladistics 12:199–220.

——— 1999. Analyzing large data sets in reasonable times: solutions for composite optima. Cladistics 15:415–428.

——— 2000. NONA, version 1.6. Program and documentation. American Museum of Natural History, Department of Entomology, New York.

———, J. S. Farris, and K. Nixon. 2002. T. N. T.: Tree Analysis Using New Technology. Program available at www.cladistics.com.

Gordon, A. D. 1980. On the assessment and comparison of classifications; pp. 149–160 *in* R. Tomassone, (ed.), Analyse de donnees et informatique. 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. Society of Vertebrate Paleontology Memoir 4.

Huelsenbeck, J. 1991. When are fossils better than extinct taxa in phylogenetic analysis? Systematic Zoology 40:458–469.

Kearney, M. 2002. Fragmentary taxa, missing data, and ambiguity: mistaken assumptions and conclusions. Systematic Biology 51:369–381.

Kitching, I., P. Forey, C. Humphries, and D. Williams. 1998. Cladistics, The Theory and Practice of Parsimony Analysis. Systematics Association Special Publication 11, Oxford University Press, 228 pp.

Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis for relationships among *Epicrates* Boidae, Serpentes. Systematic Zoology 38:1–25.

Maddison, W. P., and D. R. Maddison. 1992. MacClade: Analysis of Phylogeny and Character Evolution, Version 3.04. Sinauer, Sunderland, Massachusetts.

Margush, T., and F. R. McMorris. 1981. Consensus n-trees. Bulletin of Mathematical Biology 43:239–244.

Matthee, C. A., J. D. Burzlaff, and J. P. Bollback. 2001. Mining the mammalian genome for artiodactyl systematics. Systematic Biology 50:367–390.

Messenger, S. L., and J. A. McGuire. 1998. Morphology, molecules and the phylogenetics of cetaceans. Systematic Biology 47:90–124.

Nixon, K. C. 1996. Paleobotany in cladistics and cladistics in paleobotany: enlightenment and uncertainty. Review of Paleobotany and Palynology 90:361–373.

——— 1999a. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15:407–414.

——— 1999b. Winclada (BETA) Ver. 0.9.9, published by the author, Ithaca, New York.

———, and J. M. Carpenter. 1996a. On simultaneous analysis. Cladistics 12:221–242.

———, and ——— 1996b. On consensus, collapsibility, and clade concordance. Cladistics 12:305–322.

———, and Q. D. Wheeler. 1992. Extinction and the origin of species; pp. 119–143 *in* M. J. Novacek and Q. D. Wheeler (eds.), Extinction and Phylogeny. Columbia University Press, New York.

Norell, M. A., and K. de Queiroz. 1991. The earliest iguanine lizard (Reptilia: Squamata) and its bearing on iguanine phylogeny. American Museum Novitates 2997:1–16.

———, 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. American Museum Novitates 3211:1–25.

———, and W. Wheeler. 2003. Missing entry replacement data analysis: a replacement approach to dealing with missing data in paleontological and total evidence data sets. Journal of Vertebrate Paleontology 23:275–283.

Novacek, M. J. 1992. Fossils, topologies, missing data, and the higher level phylogeny of eutherian mammals. Systematic Biology 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.

Patterson, C. 1981. Significance of fossils in determining evolutionary relationships. Annual Review of Ecology and Systematics 12:195–223.

Platnick, N. I., C. E. Griswold, and J. A. Coddington. 1991. On missing entries in cladistic analysis. Cladistics 7:337–343.

Remane, A. 1952. Die Grundlagen des natürlichen Systems, der vergleichenden Anatomie und der Phylogenetic. Akademische Verlagsgesellschaft, Leipzig.

Rieppel, O., and M. Kearney. 2002. Similarity. Biological Journal of the Linnean Society 75:59–82.

Rowe, T. 1988. Definition, diagnosis and origin of Mammalia. Journal of Vertebrate Paleontology 8:241–264.

Schuh, R. T., and J. S. Farris. 1981. Methods for investigating taxonomic congruence and their application to the Leptopodomorpha. Systematic Zoology 30:331–351.

———, and J. T. Polhemus. 1981. Analysis of taxonomic congruence among morphological, ecological, and biogeographic data sets for the Leptopodomorpha (Hemiptera). Systematic Zoology 29:1–26.

Sereno, P. 1999. The evolution of dinosaurs. Science 284:2,137–2,147.

Swofford, D. L. 1991. When are phylogeny estimates from molecular

and morphological data incongruent; pp. 295–333 *in* M. M. Miyamoto and J. Cracraft (eds.), Phylogenetic Analyses of DNA Sequences. Oxford University Press, New York.

——— 2001. PAUP*: Phylogenetic Analysis Using Parsimony*, Version 4.0b8. Sinauer, Sunderland, Massachusetts.

———, and D. Begle. 1993. PAUP, Phylogenetic Analysis Using Parsimony, Version 3.1.1. Illinois Natural History Survey, Champaign, Illinois.

———, and D. Maddison. 1987. Reconstructing ancestral character states under Wagner parsimony. Mathematical Biosciences 87:199–299.

Wiens, J. J. 1998. Does adding character with missing data increase or decrease phylogenetic accuracy? Systematic Biology 47:625–640.

———, and T. Reeder. 1995. Combining data sets with different numbers of taxa for phylogenetic analysis. Systematic Biology 44:548–558.

Wilkinson, M. 1992. TAXEQ2, software and documentation. University of Bristol, Bristol, England.

——— 1994. Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. Systematic Biology 43:343–368.

——— 1995a. Arbitrary resolutions, missing entries, and the problem of zero-length branches in parsimony analysis. Systematic Biology 44:109–111.

——— 1995b. Coping with abundant missing entries in phylogenetic inference using parsimony. Systematic Biology 44:501–514.

——— 1995c. More on reduced consensus methods. Systematic Biology 44:436–440.

——— 2001. TAXEQ3, software and documentation. Department of Zoology, Natural History Museum, London.

———, and M. J. Benton. 1995. Missing data and rhynchosaur phylogeny. Historical Biology 10:137–150.

———, J. L. Thorley, and P. Upchurch. 2000. A chain is no stronger than its weakest link: double decay analysis of phylogenetic hypotheses. Systematic Biology 49:754–776.

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. Canadian Journal of Earth Sciences 33:541–577.