

THREE-TAXON TRANSFORMATION IN PHYLOGENETIC INFERENCE: AMBIGUITY AND DISTORTION AS REGARDS EXPLANATORY POWER

Arnold G. Kluge

Museum of Zoology and Department of Biology, University of Michigan,
Ann Arbor, Michigan 48109, U.S.A.

Received for publication 29 March 1993; accepted 29 May 1993

Nelson and Platnick (1991: 351; see also Nelson and Ladiges, 1991a-d, 1992a; and Nelson, in 1993) stated that "[t]ransforming characters into three-taxon statements may increase the sensitivity of parsimony to differences in the fit of data to alternative cladograms", and elsewhere they spoke of that transformation increasing precision (e.g. p. 351). Further, throughout their paper they implied that the transformation does not distort the "data" as *evidence for sister-group relationships*. If there really are such benefits as increased sensitivity and precision, and if the data are not distorted, then the three-taxon transformation of Nelson and Platnick should become a routine operation in cladistics. In order to understand better these preconditions, I re-examine the details of their transformation. Then, in terms of the issues raised from this inspection, I consider the impact their transformation has on maximizing the *explanatory power of cladistic data with parsimony*, that being the largest number of synapomorphies in a data set that can be interpreted as homologues (Farris, 1979, 1980a, b, 1982a, 1983, 1985, 1989: 107; Farris and Kluge, 1985, 1986).

Also, I examine the conceptual relationship between Nelson and Platnick's (1991) transformation and information and parsimony, as well as their claims of increased sensitivity and precision. I do not attempt to judge the use of three-taxon statements in historical biogeography because the application of parsimony in that research program has yet to be precisely defined and convincingly justified (Page, 1990). For example, compared to phylogenetic inference where parsimony's *raison d'être* is to minimize *ad hoc* hypotheses of homoplasy, it is unclear what it means for "items of error" to be "economically explained" in biogeography (Nelson and Platnick, 1981: 417; Nelson and Ladiges, 1991b; see, however, Kluge, 1988: 316).

Synapomorphy, Homology, Parsimony, Explanatory Power

While I subscribe to the taxic concept of homology (Patterson, 1982), I do not equate homology with synapomorphy (*contra* Patterson, 1982; and de Pinna, 1991). Rather, I judge homology and synapomorphy to lie in the realms of ontology and epistemology, respectively. As Sober (1988: 212; see also Rieppel, 1991: 95) pointed out, "[h]omologies are not observed in advance of the genealogies we wish to infer". Thus, homologues are considered contingent facts because they assume a particular phylogenetic hypothesis, and synapomorphies are the empirical basis for those propositions on which homology is retrodicted. Explanatory power is to be understood in this contingent sense, not in terms of the *truth* of synapomorphies as marks of history.

I agree with Platnick (1989: 21) that “[t]he parsimony criterion allows hypotheses about characters and their distribution to impinge on hypotheses about relationships and hence about what *natural groups exist*, and the role of parsimony in that process is essential” (my italics). And as Farris (1970: 92) pointed out, the optimization of a character’s states on a cladogram is “used to increase the parsimony”. In seeking the best-fitting cladogram for a given data set, the principle of parsimony is applied in order to minimize requirements for *ad hoc* hypotheses of homoplasy, and it is on such a cladogram that explanatory power is maximized. As with Hennig’s auxiliary principle (1966: 121; hypotheses of homology are to be preferred in the absence of evidence to the contrary), maximizing explanatory power “amounts to the precept that homoplasy ought not to be postulated beyond necessity, that is to say parsimony” (Farris, 1983: 8). Thus, parsimony provides a basis for postulating a connection between the present (synapomorphies) and the past (homology), and that principle must even be evoked to hypothesize that all or any one of a perfectly congruent set of synapomorphies are not parallelisms (Farris, 1983: 13). “The explanatory power of a genealogy is consequently measured by the degree to which it can avoid postulating homoplasies” (Farris, 1983: 18), and the cladist can do no better than explain all of the relevant available evidence with character congruence (Kluge, 1989; Kluge and Wolf, 1993). Logically, all that can be said about the history of two incongruent synapomorphies in a data matrix is that both characters cannot have had unique and unreversed histories (Kluge, 1976: 28).

Basic Operations

Nelson and Platnick (1991) repeatedly referred to their three-taxon method as involving “transformation”, transforming an “original” matrix of synapomorphies into a three-taxon statement matrix, which is then analyzed with parsimony for sister-group relationships. Ordinarily, transformation is understood to be an *additional*

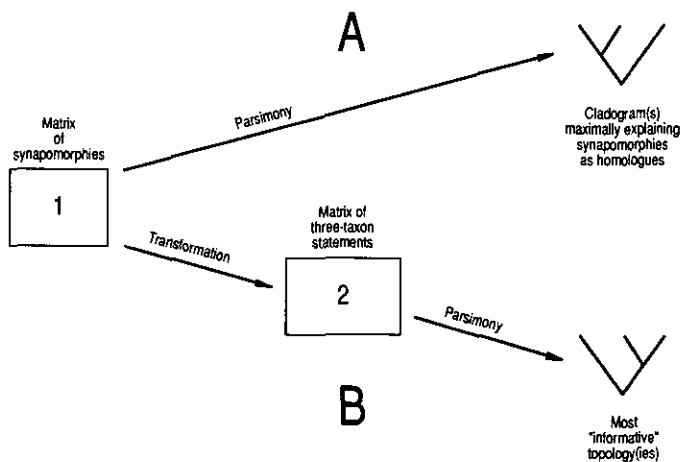


Fig. 1. A comparison of parsimony analyses of untransformed (A) and three-taxon statement transformed (B) synapomorphies. A consequence of the former analysis is a cladogram(s) which maximizes the explanatory power of the synapomorphies as homologues. The added three-taxon transformation can significantly alter that evidence, and the most parsimonious proposition has been termed most “informative” by its advocates (Nelson and Platnick, 1981, 1991).

operation compared to the usual sequence [Fig. 1(B)], which in the present context usually proceeds from a matrix of synapomorphies directly to the phylogenetic hypothesis(es) that maximally explains those shared-derived states as homologues [Fig. 1(A)]. Similarly, the additive binary coding method (Farris et al., 1970: 180), which transforms additive multistate characters into sets of binary characters, is an extra operation, and one that must be employed if *a posteriori* iterative weighting is to be applied correctly to additive data (Carpenter, 1988; Fig. 2). This transformation allows each character's weight, as some function of extra steps, to be measured relative to one synapomorphy, rather than to a variable number of synapomorphies which is typical of multistate characters.

There is no disputing the fact that either of these added operations changes the physical appearance of the original data matrix (the number of characters and their states in the untransformed matrix are not the same in the transformed matrix (Figs 1 and 2; e.g. see Nelson's, 1993, reanalysis of Barrett's et al., 1991, data). However, the more important issue is whether or not the change impacts the data as evidence for sister-group relationships. It has long been known that additive binary coding (Farris et al., 1970) has no effect on synapomorphies potentially interpretable as homologues.

I use a simple example to demonstrate some of the effects the three-taxon transformation has on the original data as evidence for sister-group relationships. I believe this example describes unambiguously some of the consequences of the transformation because the synapomorphies in the original matrix are *perfectly congruent*. Nelson and Platnick (1991) provided many more complex examples, some of which are considered briefly below under the heading of Sensitivity and Precision (see also Harvey, 1992).

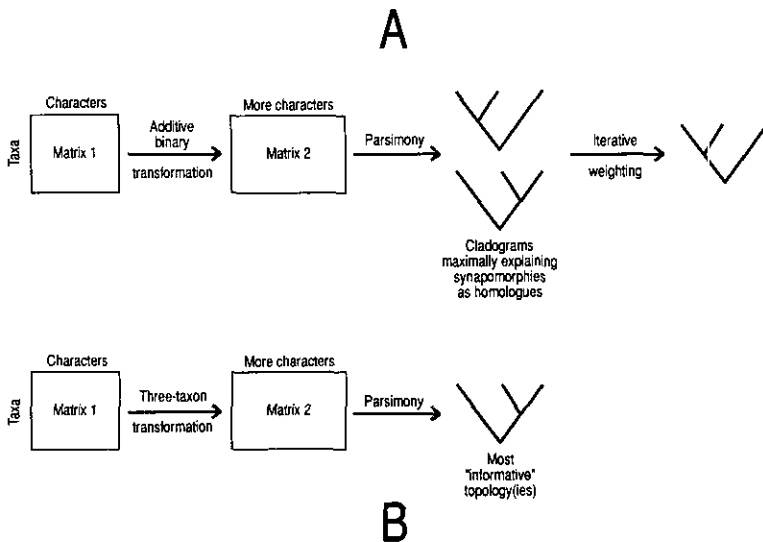


Fig. 2. A comparison of additive binary (A) and three-taxon statement (B) transformations in phylogenetic analysis. The former transformation is applied to additive multistate characters when *a posteriori* iterative weighting is used as the basis for choosing among equally parsimonious secondary cladograms. Additive binary coding (A) does not add ambiguity to the evidence, nor does it distort it, whereas three-taxon statement transformation (B) can significantly do so. See text and Fig. 1 legend for further discussion of the three-taxon transformation, and matrices 1 and 2, in particular.

Consider the following data matrix, which consists of three binary variables (labeled 1–3) *observed on organisms* representative of five terminal taxa (denoted A–E):

Original data matrix 1

	1	2	3
A	0	0	0
B	1	0	0
C	1	1	0
D	1	1	1
E	1	1	1

Original data matrix 1 is transformed into the following three-taxon statement matrix, and Nelson and Ladiges' (1992b) TAX software was used in this operation:

Three-taxon statement matrix 2

	1	2	3
	abcdef	abcdef	abc
A	000000	000???	0??
B	111???	???	0??
C	1??11?	11?	??0
D	?1?1?1	1?11?1	111
E	??1?11	?11?11	111

The transformed matrix consists of 15 three-taxon statements. The numerical label of each original character is retained, to which I have added a lower-case letter for each of the possible corresponding statements.

Basically, the operation involves transforming a character in the original matrix into as many "statements" as there are *possible data* for terminal taxon sister-group relationships. The transformation is constrained by the original data, in particular by the shared-derived states.

It is important to emphasize that, in applying the transformation, the special similarity of all but two apomorphic terminal taxa is replaced by a ? mark(s), thereby constraining each statement to only two *terminal* taxa. For example, character 1 in matrix 1 delimits the group (B, C, D, E). Taxon A is the sister-group to (B, C, D, E) because A is considered plesiomorphic (see polarity operation below). The special similarity evident in this character corresponds to the six possible statements of sister-group relationship shown in matrix 2, 1a (B, C), 1b (B, D), 1c (B, E), 1d (C, D), 1e (C, E) and 1f (D, E), each of which is unique. Character 2 in matrix 1 is less general. It delimits the group (C, D, E), and while it also transforms to six statements, 2a (C, D), 2b (C, E), 2c (D, E), 2d (C, D), 2e (C, E) and 2f (D, E), only three of those are unique. Character 3 in matrix 1 is even less general. It delimits the group (D, E)—all three corresponding three-taxon statements, 3a (D, E), 3b (D, E) and 3c (D, E), describe just that group.

Nelson and Platnick (1991) based many of their claims on the sister-group relationships they obtained from the original and three-taxon statement matrices. Following their recommendation, I add an all-zero ancestor taxon to matrices 1 and 2 in order to polarize the characters, and I use Hennig86 (Farris, 1988) to find the best-fitting hypothesis of relationships for those data sets. A parsimony analysis of each matrix leads to the same best-fitting cladogram, (A(B(C(D, E)))). While the two matrices differ in number of characters, both sets of data are completely congruent on that cladogram (C = 1.0; R = 1.0).

Issues

AMBIGUOUS EXPLANATION

All three synapomorphies in matrix 1 are interpretable unambiguously as homologues on the best-fitting cladogram (A(B(C(D, E))))), whereas statements 1d, 2a, 1e, 2b, 1f, 2c, 3a, 2f and 3b in matrix 2 do not unequivocally delimit a particular sister-group with the FARRIS method of optimization employed by Hennig86 (Farris, 1970). The explanatory power of the best-fitting cladogram has been realized completely relative to matrix 1 (3/3, or 100%), but that power is diminished considerably relative to matrix 2 (6/15, or 40%). If each set of identical three-taxon statements is treated as a singular delimiter, 1d and 2a, 1e and 2b, 1f, 2c and 3a, and 2f and 3b, then more of the characters in matrix 2 can be unambiguously interpreted as homologues (6/10), but the total is still less than 100%. Further improvement in the number of characters unambiguously interpretable as homologues might be realized by applying other types of character optimization (e.g. ACCTRAN, DELTRAN or MINF); however, those alternatives make extra assumptions about evolutionary processes. Thus, not only has the transformation of matrix 1 into 2 increased the number of characters (from three to 15), but it is also clear that the transformation has rendered the original data *less decisive*, at least as regards interpreting synapomorphies as homologues on the best-fitting hypothesis (Farris, 1980a).

MISSING DATA

The matrix resulting from the three-taxon transformation has considerable missing data, *where none existed before* (e.g. compare matrix 2 to 1). According to Platnick (1988: 311), missing data are evidentially ambiguous, and therefore the transformation from original data to three-taxon statements adds substantial equivocation to the transformed matrix (matrix 2 in the above example).

One might argue that the problem lies with the current parsimony programs, such as Hennig86 and PAUP, in terms of how they treat data coded as missing and find spurious clades, i.e. those without support. As Platnick et al. (1991) clearly outlined, there are three possible types of missing data—unknown, inapplicable or variable.¹ For example, in the simple binary case, an unknown datum is *either* 0 or 1, inapplicable datum *neither* 0 or 1 and variable datum *both* 0 and 1. While recognizing that three-taxon statements involve missing data of the first type, I believe the ? marks are not unknown in any ordinary sense (e.g. an observation of an invariant nature is recorded in each of the cells in matrix 1). Nelson's (1992: 358) assertion, that "[w]hatever the appearances, three-item analysis does not transform known data into missing data", does not help to convince me otherwise. In any case, as a consequence of accepting Nelson's position on missing data, (A₂(B₁, C₁)) would then be an allowable three-taxon statement, but I can find no allowance for this possibility in Nelson and Platnick (1991).

Thus, I conclude that whatever arguments might be used to justify Nelson and Platnick's (1991) transformation they must be reconciled with the fact that the

¹ I prefer to restrict the term polymorphism to the special case where different states, like alleles, can be present in a population. Thus, I replace the category "polymorphic" of Platnick et al. (1991) with the more general term "variable".

operation adds ambiguity where none existed before (see also Nelson and Platnick, 1991: 355). Further, as Platnick et al. (1991) underscored, currently available algorithms for discovering best-fitting cladograms do not distinguish between the three different interpretations of missing data, and the three-taxon transformation only compounds that problem.

EVOLUTIONARY REVERSAL

Nelson and Platnick's (1991) three-taxon transformation eliminates all possible ingroup symplesiomorphies—one or more plesiomorphic states in each character is replaced with a ? mark. This amounts to assuming analytical priority (namely, that polarity is more important than matches), that evolutionary reversals are relatively rare, if not absent altogether, or that such a novelty cannot be diagnostic. On the latter count, I believe it is fair to say that the transformation distorts the original data as evidence for genealogical relationships because symplesiomorphies are no longer available to be tested with congruence for phylogenetic informativeness as evolutionary reversals, i.e. when matching 0s delimit sister-groups (Nelson and Platnick's, 1991: 362–363, matrices 17–18, example not withstanding). For instance, the issue arises when two or more reversals delimit a clade, a part of which is supported by *fewer incongruent* shared apomorphies.

Nelson and Platnick (1991: 359–360; e.g. their matrices 11 and 12) provided excellent examples of this distortion. Not only can the number of equally parsimonious cladograms increase by eliminating the symplesiomorphies, but that loss can lead to a different solution (compare their “trees” 1 and 2 derived from their matrices 12 and 11, respectively).

CHARACTER DEPENDENCE AND REDUNDANCY

The three-taxon transformation of Nelson and Platnick (1991) produces non-independent characters. As Nelson and Ladiges (1992a) noted, it necessarily produces *logically dependent* statements. For example, in the simplest case where taxon A is plesiomorphic and taxa B, C and D are apomorphic, any two of the three-taxon statements, (A(B, C)), (A(B, D)) and (A(C, D)), logically imply the third.

The transformation also *creates* redundant evidence. In particular, the generality of the apomorphic state in the original character affects the number of unique three-taxon statements into which the character can be transformed (compare characters in matrices 1 and 2). This redundancy of three-taxon statements is due to the transformation eliminating all symplesiomorphies—the number of instances of apomorphy and plesiomorphy a character exhibits is inversely related, and redundant three-taxon statements are required to express the different combinations of 0 and its substitute ? mark. For example, as noted above, the transformation of character 2 leads to six three-taxon statements (a–f), but only three of which are unique as regards the sister-group identified: 2a and 2d delimit (C, D), 2b and 2e delimit (C, E) and 2c and 2f delimit (D, E). Similarly, character 3 transforms to three three-taxon statements, but only one sister-group is described, (D, E). The taxonomic generality of states 0 and 1 is two (A and B) and three (C, D and E), and three (A, B and C) and two (D and E), in untransformed characters 2 and 3, respectively.

Nelson and Ladiges (1992a) pointed out that additional redundancy is realized when a additive multistate character is transformed. Consider the simple case of a three-state character, where state 0 is exhibited by taxon A, state 1 describes clade (B, C, D), and taxa C and D have state 2. (A(B, C, D)) transforms to (A(B, C)), (A(B, D)) and (A(C, D)), and (A, B(C, D)) transforms to (A(C, D)) and (B(C, D)), with the redundant (A(C, D)) statement resulting from both transformations.

Given that choice among competing cladograms is based on maximizing explanatory power, character independence is important because that assumption is related to how much and which evidence there is to be explained (Kluge, 1989). Thus, whatever the source of non-independent evidence, logical dependence or redundancy, the three-taxon transformation can affect which best-fitting hypothesis is selected (see Discussion below). Nelson and Platnick (1991: 355, 363) and Nelson and Ladiges (1992a) recognized this distortion, and they suggested using *a priori* differential character weighting to remove the effects of non-independence that result from the transformation. However successful such weighting might be, the problem remains as to whether or not each of the dependent and redundant synapomorphies resulting from the transformation is to be explained separately as homologues (or as homoplasies). For example, how are we to understand the evidential nature of characters 3a, 3b and 3c (matrix 2) on the best-fitting cladogram? To put the issue graphically, are they each a part of a homology, and if not, how can they be considered separate evidence for the common ancestry of D and E? I can only conclude that whatever justifies weighting three-taxon statements it does not appear to be explanatory power in any ordinary phylogenetic sense, i.e. synapomorphies interpretable as homologues. Of course, one could avoid this dilemma in the case of redundancy by simply deleting all but one of the non-independent transformed characters, but that strategy would only call into question why the three-taxon transformation was applied in the first place.

OPERATIONAL LIMITATIONS

It is evident that Nelson and Platnick's (1991) three-taxon transformation is limited to statements of sister-group relationships among *terminal taxa*. But it is not clear why a three-taxon statement should be construed as something more specific than merely sister-groups, (X, Y). For example, (A(B(C, D, E))) is a three-taxon statement for character 1 in matrix 1 (the three taxa in the statement being [B], [C, D, E] and [A]), and Nelson and Platnick (1981: 250–257) recognized just such a statement in their solutions to three-taxon problems in minimum and maximum modes (see also Page's 1990: 127–129, biogeographic example). Of course, one could argue that (B(C, D, E)) reduces exhaustively to statements (B, C), (B, D), (B, E), (C, D), (C, E) and (D, E), but the question remains, what is the justification for such a reduction? The issue becomes more complicated if terminal taxa, (A), (B), . . . , are themselves composite wholes of other taxa, because then we are forced to ask what is the justification for *differentially* exhaustively enumerating all possible statements pertaining to higher taxa, such as (B(C, D, E))? Does the three-taxon transformation only apply to completely resolved terminal taxa? What justifies such a constraint, and what arguments can be mounted for its inconsistent application?

There is also the issue of the three-taxon statement itself, as a sufficient basis for providing the *globally* most parsimonious unambiguous hypothesis. Nelson and Platnick's (1991) use of an outgroup is tacit admission that the three-taxon statement (A(B, C)) is in fact not sufficient. Further, it is generally accepted that *multiple outgroups* are required to be able to present the evidence (synapomorphies as homologues) as parsimoniously and unambiguously as possible (Farris, 1982b; Maddison et al., 1984).

Information and Parsimony

Thus far I have judged the transformation of synapomorphies into three-taxon statements by the impact that operation has on explanatory power [Fig. 1(A)]. Maximizing that power with parsimony is but one of several justifications that have been offered for choosing the simplest hypothesis of relationships (Sober, 1988). Other justifications have been based directly, or indirectly, on information (Farris, 1979; Kluge, 1984). While Nelson and Platnick (1991; see, however, Nelson and Platnick, 1981: 39) did not explicitly consider explanatory power, they argued for their methodology in terms of information [Fig. 1(B)]. Therefore, it is appropriate to discuss their notion of information in order to consider whether or not it avoids any of the criticisms I have concerning ambiguity and distortion of explanatory power.

Nelson and Platnick summarized their views on information in systematics and biogeography in their 1981 book. They stated (p. 14) that the primary information content of a classification is "what groups there are, and how they are subdivided"; the secondary content being information about the names and ranks of the groups and subgroups. They characterized cladograms as summarizing *structural elements of knowledge*, which in "[t]he minimum mode ... is a suite of 3-taxon problems that, once solved, result in the informative components of the cladogram that, with respect to a certain sample of information, is the true and final resolution" (p. 254). Thus, they took the informative component, (X, Y), to be "the solution to a particular 3-taxon problem in the minimum mode; and a cladogram may be understood as the combined solutions to a suite of 3-taxon problems" (p. 255). However, they acknowledged that "[m]inimum-mode resolution is sufficiently complex so that it probably can never be consistently achieved in practice" (p. 253). They also recognized term information, which, together with component information, defined total information. Term information is simply one less than the number of taxa in an informative component, which is the number of maximum-mode problems (p. 256). Thus, Nelson and Platnick's (1981: 238–257) definition of information is the set of three-taxon statements allowed by the topology (Page, 1992: 87, 93).

In contrast, Nelson and Platnick (1991: 352–3, 355, 359, 362–365) defined information in terms of characters in their paper on three-taxon statement transformation. For example, they concluded that the more putatively apomorphic entries a character has the more informative it is, and (p. 353) "[p]erhaps the number of three-taxon statements implied by a character is itself an accurate measure of the character's information content". Further (p. 362), "[i]f ... characters are hypotheses of homology and synapomorphy, then they must be relational, and the units of those relations are three-taxon statements (taxa A and B are homologous for this feature, relative to taxon C)", and from which they concluded that the three-taxon

approach is more precise than the untransformed data matrix because it "pays attention to [the] difference in information content". However, it is difficult to see what character-based concept of information would justify their line of reasoning, because the number of question marks (evidential ambiguity; Platnick, 1988: 311) their transformation adds to the data increases as the number of putative apomorphies increases.

In contrast to Nelson and Platnick (1981, 1991), others have related information in systematics directly to parsimony. In fact, these arguments have been considered justifications for using the principle of parsimony generally and in systematics in particular (Kluge, 1984). For example, Sober (1975) concluded that the simplest hypothesis is objectively the most informative. Informativeness in this sense is the measurable extent to which the hypotheses alone answer questions about individuals in their domain. That is, the simplest hypothesis requires less extra information than a more complex one to be able to answer a question derived from the data (E. Sober, pers. comm.). Restated, the greater the information content of the hypothesis the greater the number of things it forbids (see paragraph above for the *opposite conclusion reached by Nelson and Platnick, 1991*).

In addition, Farris (1977, 1979; 1980a, b, 1982a) connected information to the characters responsible for hypotheses and, in turn, to parsimony. As Mickevich and Farris (1981: 362; see also Page, 1992) pointed out, "two bushes [polytomies] are just alike—entirely similar—but they have no information in common, because they have no information at all". "Classifications as such are just suites of groups, and contain no information on characters directly, but do so only indirectly, through the diagnoses of their taxa (Farris, 1977: 841–842; 1979: 489–490). Assessing the information content of classifications then amounts to evaluating their diagnoses" (Farris, 1980a: 389). "The diagnoses are constructed so that the data may be recovered from the classifications together with the diagnoses. This is part of the sense in which the classifications (or the diagnoses) are informative" (Farris, 1980a: 389). And, "[e]fficiency of diagnoses is ... interpretable directly in terms of information content (Farris, 1979: 506). Since the diagnoses together with the classification contain all the information in the data table, the simplest, most succinct, or more efficient suite of diagnoses has the greatest information content per entry. This is the only useful sense in which classificatory information content on characters can be measured" (Farris, 1980a: 391). "It is only efficiency—information content per diagnostic entry—that distinguishes descriptively useful from descriptively worthless classifications" (Farris, 1980a: 391). *Therefore, according to Farris, descriptive efficiency is the same as informativeness and the classification with the greatest information content, being the most parsimonious, is the phylogeneticist's estimate of the single historical pattern.*

Most agree that the information relevant to phylogenetic inference lies in character evolution (synapomorphies interpretable as homologues; see, however, Nelson and Platnick 1981). Further, as Farris (1980a: 392) concluded, "[m]ost parsimonious trees, those that require fewest postulated origins of features (*'origins' include losses and reversals*) ... correspond to classifications with most efficient diagnoses" (my italics). Thus, I claim that Nelson and Platnick's (1991) information concept relates to either descriptive efficiency or explanatory power, and any operation, such as the three-taxon transformation, that detracts from maximizing either of those qualities is open to criticism.

Sensitivity and Precision

According to Nelson and Platnick (1991: 351, 355), the three-taxon transformation may add precision by increasing “the sensitivity of parsimony to differences in the fit of data to alternative cladograms”. This claim is based on different arguments scattered among their several examples, pairs of matrices—original and transformed. The data sets included incongruent synapomorphies, and for which there was usually more than one best-fitting hypothesis. It is clear from Nelson and Platnick’s examples, that any possible outcome can result from the transformation—the operation can leave the number of equally parsimonious cladograms unchanged, reduced (e.g. their matrices 1 and 2) or increased (e.g. their matrices 11 and 12) in number, or it can even lead to a novel topology (e.g. their matrices 15 and 16). As might be imagined given such variable outcomes, generalities regarding the sensitivity–parsimony relation are difficult to identify. In fact, the only obvious generality is the increase in number of characters in the transformed matrix (*cladogram length is longer*). If such is the only basis for increasing sensitivity, as it appears to be (Nelson and Platnick, 1991: 365), it comes at the expense of distorting the original data and increasing its ambiguity. How can the duplication of synapomorphies, which accompanies the transformation, result in “a more precise summary” of hypotheses of homology (p. 362) when that operation has such negative effects on the original data?

Conclusions

Nelson and Platnick’s (1991) three-taxon transformation applied to the data of phylogenetic inference has several consequences: (1) Relatively fewer synapomorphies are interpretable unambiguously as homologues—the transformation has rendered the data “weaker” than it is in fact. (2) Missing data are added, where none existed before. This amounts to the unscientific exercise of throwing away observations. (3) The data are distorted because the phylogenetic informativeness of evolutionary reversals is discarded with the removal of symplesiomorphies. (4) The number of logically dependent and redundant characters increases. Thus I conclude, in general, that explanatory power is negatively affected by Nelson and Platnick’s (1991) three-taxon transformation. Additional criticisms include three-taxon statements being limited to terminal taxa, and local, rather than global, parsimony being emphasized. Nelson and Platnick’s (1991) argument for the three-taxon transformation, because it improves precision by increasing the sensitivity of parsimony to the original data, is flawed given that the operation distorts and adds ambiguity to the original observations.

Transformed and untransformed matrices will always yield the same most parsimonious cladogram when homoplasy is absent, as illustrated by my example above. Therefore, that the three-taxon transformation can affect the amount of homoplasy present, in ways other than removing symplesiomorphies, must be examined more closely. Harvey (1992) has already argued at length that the transformation’s differential character “weighting” (I would say multiplication of homoplasies) produces cladograms different from those obtained from the untransformed matrix. As I discussed above, how much and which evidence there is to be explained is related to the transformation’s affect on character dependence and redundancy,

which is a function of the number of taxa in the data matrix, and the steepness of the weight differentials among characters is correlated with the number of taxa being investigated (Harvey, 1992: fig. 2)—the differential is steeper the more taxa there are in the original matrix. Thus, Harvey (1992: 353) was led to conclude that “the significant results presented by Nelson and Platnick (1991) for the three-taxon method are artifactual consequences of a character weighting scheme that does not reflect any phylogenetically useful criteria” (see, however, Nelson’s 1992, rejoinder). Given that the three-taxon transformation adds homoplasy to the original data, one must also conclude that Nelson and Platnick’s (1991) method is antithetical to the parsimony criterion.

Discussion

A consistent effect of applying the three-taxon transformation is longer cladograms. What other considerations, besides those discussed by Nelson and Platnick (1991), might justify seeking those suboptimal, less parsimonious hypotheses? Can it be considered a hypothetico-deductive operation (e.g. a special device not provided by congruence for exploring the relation of similarity to homology, or of taxon to homology)? Perhaps the relation between synapomorphy and homology is too burdened by theory, and the three-taxon transformation is seen as not appealing to any particular definition of homology (Panchen, 1992: 72–74). Perhaps, applying the transformation dispenses with the formalisms and attendant evolutionary theory of binary and additive multistate characters (Nelson, 1993). Is the three-taxon transformation merely an operation that one hopes will lead to some special insight (Nelson, 1992)? Is the “true method or methods” (Nelson, 1989a: 293) simply to be revealed by its application? Perhaps, it could be argued that if the transformation delivers “empirical findings that conflict with [some] model ... then surely the model must, if possible, adapt to the findings or go extinct” (Nelson, 1989a: 288; Nelson, 1993). One might even claim (Nelson, 1989a: 293) that the discovery process in cladistics may continue without “perfect justification” for the operation of three-taxon transformation.

These possible positions emphasize operationalism, the doctrine that theoretical concepts, such as the natural classification of species, have to be defined in terms of measuring operations (Popper, 1965: 62). Phenetics failed largely, if not entirely, because it was based on that tenet (Farris and Platnick, 1989: 302), and I do not find the arguments listed in the preceding paragraph to be any more salutary for Nelson and Platnick’s (1991) three-taxon transformation.

To my way of thinking, the larger issue associated with Nelson and Platnick’s (1991) three-taxon transformation is whether or not theory-neutral classifications are possible ... in particular whether the natural pattern of species relationships can be discerned without assuming anything about process (Minaka, 1988). I don’t believe sister-group relationships can be so delimited (*contra* Nelson, 1989b) because I understand a cladogram to be more than a hierarchy of character generalities. My position necessarily follows because, like Farris (1983: 18; see also Farris, 1989: 107) and Platnick (1989: 21), I believe parsimony is the criterion for choosing among alternative cladograms *because* its application minimizes *ad hoc* hypotheses of homoplasy, and consequently maximizes explanation of shared similarity in terms of inheritance from a common ancestor. It is this very justification for parsimony

that forces me to reject Brady's (1983, 1985) claim that the *explanandum* must be independent of the *explanans*, i.e. the study of pattern must be prior to that of process. The fact that inheritance is a process, and that it is assumed when cladistic parsimony is applied, means that pattern and process cannot be independent. In my world view, homology is more than a relation. It has a cause, and parsimony is useful, perhaps even necessary, in exploring that explanation.

Of course, there are other justifications for parsimony (see review by Sober, 1988), but they either involve other assumptions of process or, in my opinion, they fall short of Farris' (1979, 1980a, b, 1982a, 1983, 1985, 1989) persuasive arguments having to do with minimizing homoplasy and maximizing explanatory power. At least for the time being, I accept Hennig's (1966: 121 and elsewhere) position that there is more to constructing cladograms than operationalism. Is there any less to transforming original data into three-taxon statements?

Acknowledgments

I thank Jim Carpenter, Doug Eernisse, Mike Donoghue, Steve Farris, Mary McKittrick, Mary Mickevich, Gavin Naylor, Norm Platnick, Elliott Sober, Ed Wiley, and an anonymous reviewer for providing critical comments on an early version of this paper. Gareth Nelson was particularly helpful in pointing out "factual errors" in a preliminary draft of this paper, and in sharing unpublished manuscripts. Elliott Sober was also helpful in explaining his position (1975) on informativeness. The University of Michigan Systematics Discussion Group, where this paper was first read, provided additional critical comments.

REFERENCES

- BARRETT, M., M. J. DONOGHUE AND E. SOBER. 1991. Against consensus. *Syst. Zool.* 40: 486-493.
- BRADY, R. H. 1983. Parsimony, hierarchy and biological implications. *In*: N. I. Platnick and V. A. Funk (eds). *Advances in Cladistics*, Vol. 2. Proceedings of the Second Meeting of the Willi Hennig Society. Columbia Univ. Press, New York, pp. 49-60.
- BRADY, R. H. 1985. On the independence of systematics. *Cladistics* 1: 113-126.
- CARPENTER, J. M. 1988. Choosing among multiple equally parsimonious cladograms. *Cladistics* 4: 291-296.
- DE PINNA, M. C. C. 1991. Concepts and tests of homology in the cladistic paradigm. *Cladistics* 7: 367-394.
- FARRIS, J. S. 1970. Methods for computing Wagner trees. *Syst. Zool.* 19: 83-92.
- FARRIS, J. S. 1977. On the phenetic approach to vertebrate classification. *In*: M. K. Hecht, P. C. Goody and B. M. Hecht (eds). *Major Patterns in Vertebrate Evolution*. NATO Advanced Study Institute Series 14. Plenum Press, New York, pp. 823-850.
- FARRIS, J. S. 1979. The information content of the phylogenetic system. *Syst. Zool.* 28: 483-519.
- FARRIS, J. S. 1980a. The efficient diagnoses of the phylogenetic system. *Syst. Zool.* 29: 386-401.
- FARRIS, J. S. 1980b. Naturalness, information, invariance, and the consequences of phenetic criteria. *Syst. Zool.* 29: 360-381.
- FARRIS, J. S. 1982a. Simplicity and informativeness in systematics and phylogeny. *Syst. Zool.*, 31: 413-444.
- FARRIS, J. S. 1982b. Outgroups and parsimony. *Syst. Zool.*, 31: 328-334.
- FARRIS, J. S. 1983. The logical basis of phylogenetic analysis. *In*: N. I. Platnick and V. A. Funk (eds). *Advances in Cladistics*, Vol. 2. Columbia Univ. Press, New York, pp. 7-36.
- FARRIS, J. S. 1985. The pattern of cladistics. *Cladistics* 1: 190-201.

- FARRIS, J. S. 1988. Hennig86, Version 1.5. Port Jefferson Station, New York.
- FARRIS, J. S. 1989. Entropy and fruit files. *Cladistics* 5: 103–108.
- FARRIS, J. S. AND A. G. KLUGE. 1985. Parsimony, synapomorphy, and explanatory power: A reply to Duncan. *Taxon*, 34: 130–135.
- FARRIS, J. S. AND A. G. KLUGE. 1986. Synapomorphy, parsimony, and evidence. *Taxon* 35: 298–306.
- FARRIS, J. S., A. G. KLUGE. AND M. J. ECKARDT. 1970. A numerical approach to phylogenetic systematics. *Syst. Zool.* 19: 172–189.
- FARRIS, J. S. AND N. I. PLATNICK. 1989. Lord of the flies: The systematist as study animal. *Cladistics* 5: 295–310.
- HARVEY, A. W. 1992. Three-taxon statements: More precisely, an abuse of parsimony? *Cladistics* 8: 345–354.
- HENNIG, W. 1966. *Phylogenetic Systematics*. Univ. of Illinois Press, Chicago.
- KLUGE, A. G. 1976. Phylogenetic relationships in the lizard family Pygopodidae: An evaluation of theory, methods and data. *Misc. Publ. Mus. Zool., Univ. Michigan* (152): 1–72.
- KLUGE, A. G. 1984. The relevance of parsimony to phylogenetic inference. *In*: T. Duncan and T. F. Stuessy (eds). *Cladistics: Perspectives on the Reconstruction of Evolutionary History*. Columbia Univ. Press, New York, pp. 24–38.
- KLUGE, A. G. 1988. Parsimony in vicariance biogeography: A quantitative method and a Greater Antillean example. *Syst. Zool.* 37: 315–328.
- KLUGE, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Syst. Zool.* 38: 7–25.
- KLUGE, A. G. AND A. J. WOLF. 1993. Cladistics: What's in a word? *Cladistics* 9: 183–199.
- MADDISON, W. P., M. J. DONOGHUE AND D. R. MADDISON. 1984. Outgroup analysis and parsimony. *Syst. Zool.* 33: 83–103.
- MICKEVICH, M. F. AND J. S. FARRIS. 1981. The implications of congruence in *Menidia*. *Syst. Zool.* 30: 351–369.
- MINAKA, N. 1988. Order and causality: Logical and biological implications of cladistic structures. *Bull. Biogeog. Soc. Japan* 43: 1–14.
- NELSON, G. 1989a. Cladistics and evolutionary models. *Cladistics* 5: 275–289.
- NELSON, G. 1989b. Reconstructing the past: Parsimony, evolution, and Inference.—Elliot[sic] Sober [book review]. *Syst. Zool.* 38: 293–294.
- NELSON, G. 1992. Reply to Harvey. *Cladistics* 8: 355–360.
- NELSON, G. 1993. Why crusade against consensus? A reply to Barrett, Donoghue, and Sober. *Syst. Biol.* 42: 215–216.
- NELSON, G. AND P. Y. LADIGES. 1991a. Standard assumptions for biogeographic analysis. *Aust. Syst. Bot.* 4: 41–58.
- NELSON, G. AND P. Y. LADIGES. 1991b. Three-area statements: Standard assumptions for biogeographic analysis. *Syst. Zool.* 40: 470–485.
- NELSON, G. AND P. Y. LADIGES. 1991c. TAS: MSDOS program for cladistic systematics. Published by the authors, New York and Melbourne.
- NELSON, G. AND P. Y. LADIGES. 1991d. TAX: MSDOS program for cladistic systematics. Published by the authors, New York and Melbourne.
- NELSON, G. AND P. Y. LADIGES. 1992a. Information content and fractional weight of three-item statements. *Syst. Biol.* 41: 490–494.
- NELSON, G. AND P. Y. LADIGES. 1992b. TAS and TAX: MSDOS programs for cladistics, Version 3.0. Published by the authors, New York and Melbourne.
- NELSON, G. AND N. I. PLATNICK. 1981. *Systematics and Biogeography: Cladistics and Vicariance*. Columbia Univ. Press, New York.
- NELSON, G. AND N. I. PLATNICK. 1991. Three-taxon statements: A more precise use of parsimony? *Cladistics* 7: 351–366.
- PANCHEN, A. L. 1992. *Classification, Evolution, and the Nature of Biology*. Cambridge Univ. Press, New York.
- PAGE, R. D. M. 1990. Component analysis: A valiant failure? *Cladistics* 6: 119–136.
- PAGE, R. D. M. 1992. Comments on the information content of classifications. *Cladistics* 8: 87–95.

- PATTERSON, C. 1982. Morphological characters and homology. *In*: K. A. Joysey and A. E. Friday (eds). *Problems of Phylogenetic Reconstruction*. Academic Press, London, pp. 21-74.
- PLATNICK, N. I. 1988. Systematics, evolution, and biogeography: A Dutch treat. [A book-review of] *Systematics and evolution: A matter of diversity*. P. Hovenkamp (general ed.). 1987. Institute of Systematic Botany, Utrecht Univ. *Cladistics* 4: 308-313.
- PLATNICK, N. I. 1989. Cladistics and phylogenetic analysis today. *In*: B. Fernholm, K. Bremer and H. Jörnvall (eds). *The Hierarchy of Life: Molecules and Morphology in Phylogenetic Analysis*. Proceedings from Nobel Symposium 70 held at Alfred Nobel's Björkborn, Karlskoga, Sweden, 29 August-2 September 1988. Excerpta Medica, Amsterdam, Netherlands, pp. 17-24.
- PLATNICK, N. I., C. E. GRISWOLD AND J. A. CODDINGTON. 1991. On missing entries in cladistic analysis. *Cladistics* 7: 337-343.
- POPPER, K. R. 1965. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Harper Torchbooks, New York.
- RIEPEL, O. 1991. Things, taxa and relationships. *Cladistics* 7: 93-100.
- SOBER, E. 1975. *Simplicity*. Clarendon Press, Oxford.
- SOBER, E. 1988. *Reconstructing the Past: Parsimony, Evolution, and Inference*. Massachusetts Institute of Technology, Cambridge.

