

- Saether, O. A. 1983. The canalized evolutionary potential. Inconsistencies in phylogenetic reasoning. *Syst. Zool.* 32(4): 343–359.
- Sober, E. 1975. *Simplicity*. Clarendon Press, Oxford.
- Thomas, W. W. 1984. The systematics of *Rhynchospora* section *Dichromena*. *Mem. N.Y. Bot. Gard.* 37: 1–116.
- Wagner, W. H., Jr. 1961. Problems in the classification of ferns. In: *Recent advances in botany*, vol. 1, pp. 841–844. Univ. Toronto Press, Toronto.
- . 1980. Origin and philosophy of the groundplan-divergence method of cladistics. *Syst. Bot.* 5: 173–193.
- Wiley, E. O. 1981. *Phylogenetics. The theory and practice of phylogenetic systematics*. John Wiley & Sons, New York.
- Steven P. Churchill, *New York Botanical Garden, Bronx, NY 10458, U.S.A.*;
E. O. Wiley, *Museum of Natural History and Department of Systematics and Ecology,*
University of Kansas, Lawrence, KS 66045, U.S.A.; and
Larry A. Hauser, *Department of Botany,*
University of Illinois, Urbana, IL 61801, U.S.A.

PARSIMONY, SYNAPOMORPHY, AND EXPLANATORY POWER: A REPLY TO DUNCAN

According to Duncan (1984), “directed character compatibility [clique] analysis is equivalent to Hennig’s method,” because Hennig restricts the term synapomorphy “to shared apomorphies that are postulated to be uniquely derived.” The connection between Hennig’s method and parsimony analysis, Duncan maintains, is a misconception that arose because “Farris, [Kluge and Eckardt] (1970) redefine synapomorphy to include shared non-uniquely derived characters, abandon Hennig’s criterion that monophyly be based on shared uniquely derived character states, and arbitrarily impose the criterion of parsimony on Hennig’s method.” The crux of Duncan’s argument, then, is his claim concerning Hennig’s definition of synapomorphy.

In the usual conception, if an apomorphic trait shared by two taxa is inherited from a common ancestor, then this constitutes synapomorphy. If the “same” trait originated separately in the two lines, then the similarity between the two is not synapomorphy, but homoplasy (convergence, parallelism, homoiology, etc). Duncan quotes Hennig (1966: 120) to this effect, and apparently takes this idea to imply that only unique derivations can constitute synapomorphies. In fact there is no such implication. Consider as an example endothermy in vertebrates. It is not uniquely derived, having arisen independently in birds, in mammals, and in some fishes. That birds and mammals both have that trait does not comprise synapomorphy, as the most recent common ancestor of the two groups lacked the property. That birds share that same property among themselves, however, does constitute a synapomorphy of the members of that group, and the same applies to mammals, for within each group the trait is inherited from a common ancestor.

Duncan writes that Hennig would not consider similarity in a multiply-derived trait as synapomorphy, but his grounds for that claim are somewhat vague. He offers a number of quotations from Hennig in ostensible support of his interpretation, but none of those passages states that synapomorphy is to be found exclusively in unique derivations, and Duncan provides no argument relating the actual wording to his notion. He refers to examples given by Hennig (1966: 90–92) in which the synapomorphies are coincidentally unique derivations, but those examples include no multiply derived traits and so say nothing of how Hennig would treat such characters. Indeed, Hennig himself (1966: 93) notes that this part (pp. 90–92) of his discussion disregards such practical problems as reversal, parallelism, and convergence. (In citing the examples from those pages Duncan neglects to mention this.) But if the parts of Hennig’s writings that Duncan does cite are not pertinent to this claim, other parts are more decisive. Hennig (1983: 152) lists endothermy as an “abgeleitete Grundplanmerkmal” (i.e. synapomorphy; cf. Hennig, 1983: 9) of birds, and also (1983: 169) among the “apomorphen Grundplanmerkmale” of mammals.

Such examples can easily be multiplied: Hennig (1983: 57) lists paired appendages as a synapomorphy of gnathostomes, while the secondary loss of those structures is (1983: 105) a synapomorphy of

apodans—and also, of course, of snakes. But there is neither space nor need to enumerate these cases further. It is already clear that Duncan's description of Hennig's views was not motivated by careful study of Hennig's writings. It would appear instead that Duncan has proceeded simply by attributing to Hennig whatever ideas seem necessary to support Duncan's own opinions, avoiding mention of any statement by Hennig that cannot be made to appear to conform to Duncan's position.

Farris, Kluge, and Eckardt suggested a connection between parsimony and Hennig's (1966: 121) point that (as Duncan quotes it) "the more . . . apomorphic [character states] . . . [that] are present in a number of different species, the better founded is the assumption that these species form a monophyletic group." According to Duncan this passage "asserts that monophyletic groups are more compelling if a congruent set of transformation series define these groups" and implies that phylogenetic inferences should be drawn from cliques rather than parsimony analyses. He advances two lines of argument for his position. The first is a citation of Felsenstein's (1983) contention that choice of the parsimony criterion is arbitrary. The second involves a hypothetical data set devised by Duncan.

Felsenstein's (1983: 316) discussion of this point consists in its entirety of the assertion that the argument of Farris, Kluge and Eckardt "contains one step that leads to the arbitrary selection of parsimony." He does not identify that step, nor explain why it is arbitrary. He presents, in short, no argument whatever. Duncan's unqualified citation of Felsenstein's opinion is little less misleading than his portrayal of Hennig's usage of synapomorphy.

Duncan uses his hypothetical data set in an attempt to show by example that parsimonious cladograms may "contain fewer monophyletic groups defined by synapomorphies in Hennig's sense" than do clique diagrams. (In the process he passes off as a "Hennigian Cladogram" a string-of-beads diagram, his fig. 2D, that resembles nothing ever proposed by Hennig.) Now in fact the parsimonious cladogram in Duncan's example has more groups than the clique tree for the same data. Duncan arrives at the opposite count by deleting from the parsimonious diagram any groups that are not set off by unique derivations. To rationalize that step he again falls back on the claim that Hennig would allow only unique deviations as synapomorphies, and it has already been seen that that claim is false. There is thus no legitimate reason for deleting groups from the parsimonious tree, and Duncan's conclusion is unsupported.

Of course the quotation from Hennig concerns the number of characters applied as synapomorphies of a group, not the number of groups, and so Duncan's argument is irrelevant anyway. If Duncan's example shows anything about numbers of characters, it is only that clique trees may indicate a larger number of putatively unique derivations than do parsimonious reconstructions. Duncan, no doubt, feels that clique methods therefore provide a greater number of synapomorphies in "Hennig's" sense than does the parsimony criterion, so yielding groups that are better founded according to Hennig's quoted principle. But that idea rests on the supposition that Hennig would count only unique derivations in applying this rule, and the support for that premise would seem to reside entirely in Duncan's imagination. The passage from Hennig (as distinguished from Duncan's rewording of it) makes no mention whatever either of unique derivation or of congruence, and it is already clear that Hennig would indeed apply non-unique derivations as apomorphies of groups.

Which method, then, better conforms to Hennig's principle? Clique procedures apply only putatively unique derivations as synapomorphies. Parsimony analysis aims to minimize cases of homoplasy, which is to say that it applies as many apomorphic similarities as can be managed as synapomorphies. Plainly the former method will apply fewer synapomorphies than the latter, and in practice it is often found that the difference in the number of characters used is substantial. By Hennig's criterion parsimonious arrangements are thus better founded than those produced by clique techniques.

Duncan et al. (1980) had applied several methods of analysis to an artificial data set, the dendrogrammaceae, and had claimed to show from those results that Wagner's groundplan-divergence method yielded the same tree as clique analysis, while parsimony analysis produced a different arrangement. Churchill et al. (1984) analyzed those data again and found that Hennigian, groundplan-divergence, and parsimony analyses produced the same results, while clique analysis yielded a disparate arrangement, so disputing both the position of Duncan et al. and Duncan's claim that the clique technique is Hennigian. Duncan denies that the Hennigian analysis of Churchill et al. is truly Hennigian and maintains that their groundplan-divergence diagram violates Wagner's principles. His basis for rejecting their Hennigian analysis is only that it applies non-unique derivations as synapomorphies, and it has already been seen that that reasoning is specious. His discussion of groundplan-divergence, however, is rather more elaborate.

The groundplan-divergence diagram of Churchill et al. is more parsimonious than that proposed

by Duncan et al., and Duncan defends his preference by trying to explain why groundplan-divergence analysis should not yield a parsimonious result. He presents as an “appealing” feature of his arrangement that all its lineages are “defined by synapomorphies (sensu Hennig)” — that is, of course, by putatively unique derivations. But, again, there is no basis for the exclusive use of unique derivations to be found in Hennig’s views. The quality is “appealing” only because Duncan wishes to bolster cliques. He cites Sober (1983: 356) to support his preference for a less parsimonious arrangement, but in this he merely quotes Sober out of context. Sober (1983: 356) in fact concluded: “Cladists [advocate parsimony] on the grounds that it correctly identifies which genealogical hypotheses are best supported by, or provide the best explanations of, observed character distributions The likelihood argument presented shows that parsimony has the virtues cladists have claimed for it.” Indeed, Sober (1983: 350) explicitly notes that the main argument of his paper both justifies parsimony and criticizes clique methods.

Duncan also maintains that “Although parsimonious explanations of data sets is [sic] important in Groundplan-Divergence, strict application of parsimony is not part of the method.” He does not explain, however, why it is desirable not to apply an important principle, nor does he provide any satisfactory account of how parsimony is to be applied laxly, or of what other principle is to be applied in its place. His preferred diagram, he writes, was prepared so that it “has all lineages defined by uniquely derived apomorphies. Then, in the lineages defined by synapomorphies (sensu Hennig) the number of homoplasious changes was minimized.” But this is nonsense. So long as the restriction that each group be distinguished by a putatively unique derivation is maintained, minimizing homoplasy — or doing anything else — within those preselected groups can have no effect on the grouping. A “groundplan-divergence” diagram prepared according to that prescription could not be anything other than a clique tree, and Duncan’s reason — “Hennigian” synapomorphies — for preferring his arrangement amounts to no more than that tree is the one produced by clique analysis.

Duncan criticizes Churchill et al. for comparing results in order to establish equivalence of methods. They had supposed, he writes, that the groundplan-divergence method comprised parsimony and had calculated accordingly. That their groundplan-divergence and parsimony trees agreed, he argues, is not independent evidence of the agreement of the methods, but instead merely reflects the consequences of their supposition. That reasoning is correct as far as it goes, but it also reveals an inconsistency in Duncan’s own position. Duncan et al. had calculated their own “groundplan-divergence” tree so that it could not be anything but a clique tree, but they were quite content to present comparisons of trees as evidence on the agreement of cliques with “groundplan-divergence.” Plainly, if the comparisons of Churchill et al. offer no evidence, then neither do those of Duncan et al.

This only shows that evidence on the nature of the proper groundplan-divergence method must be sought elsewhere. Duncan gives the impression that it is Wagner’s ideas that lead to the requirement that each group be set off by putatively unique derivations — that is, to selection of the clique tree — but there is a curious lack of detail in his defense of that position. He offers no quotation from Wagner in support of the exclusive use of putatively unique derivations. The only evidence for the claim seems to be Duncan’s “groundplan-divergence” diagram itself. That, according to Duncan et al. (1980: 281), had been “developed by Wagner,” but Duncan now admits that he had “assisted in the construction” of it himself. Wagner without Duncan’s “assistance” (for example 1969: fig. 5) has published groundplan-divergence diagrams with groups distinguished only by non-unique derivations. Duncan’s version of “Wagner’s” views, it would seem, has much in common with his portrayal of “Hennig’s.”

Churchill et al. also observed that secondary clique procedures seem not to have been fully described in the literature. Duncan denies this, and cites as evidence some comments by Estabrook et al. (1977). The method appears well-defined in Duncan’s treatment because he cites just that one source, but consideration of other literature suggests a different picture. As Farris and Kluge (1979) have discussed in detail, the secondary clique method of Estabrook and Anderson (1979) differs from that of Estabrook et al. in the selection of both the primary and secondary character sets, and the two methods give different results. Clearly there is more than one way to proceed in secondary clique analyses, but neither paper provides any indication of the rules governing selection among these alternatives. Duncan offers no argument against any of this; as with all other literature that does not fit into his position, he simply does not mention these 1979 papers. It seems, then, that the criticism by Churchill et al. is entirely well-taken.

Duncan evidently feels that secondary clique analyses are consistent with (Duncan’s version of) Hennig’s method, but this is nothing but a self-contradiction in his position. Despite the ambiguities in their formulation, it is certainly true that secondary clique trees may have groups distinguished

only by non-unique derivations. Duncan does not deny this, but when parsimonious trees have that same property he decries them as "contrary to Hennig's original method." Evidently what is "Hennigian," and what is not, is for Duncan a matter of the convenience of the moment. That is by itself no criticism of secondary cliques, to be sure, inasmuch as Duncan's version of "Hennigian" bears no relationship whatever to Hennig's views. Rather the main defect of that "method"—at least insofar as one can tell from publication—is that it is not a method at all, but a slogan used to provide the appearance of a basis for conclusions in fact derived on undisclosed grounds.

Duncan closes his paper by suggesting—astoundingly, in view of the nature of his argument to this point—that methods ought not be chosen for their conformity to Hennig's views, but according to "the appropriate evolutionary model for groups under study." According to Duncan a variety of methods may reasonably be employed, and the benefit of using many methods is that "meaningful statements about relationships can be drawn in relation to recency of common ancestry, patterns of cladogenesis, degrees of anagenesis, overall similarity, homoplasy and geographical distribution."

The idea of selecting a method according to an "appropriate evolutionary model" contains a crucial pitfall: How was the model itself selected? It seems obvious that one would not wish to use a false model. But in that case some assurance seems required that the model be correct, and establishing the truth of a model would involve determining the facts of evolution. It seems that there is a vicious circle: The purpose of the model is to select a method, and the aim of the method is to make inferences about evolution. If the course of evolution were already known, choosing a model would be simple, but then no method would be needed. Evolution is not yet known, of course, and so methods are needed, but for that same reason choice of model is problematical.

If it were required to establish the proper model in order to select a method for evolutionary study, then nothing could ever have been learned about evolution. This does not mean that evolutionary study is impossible, only that a different way of arriving at a method is needed.

Hypotheses of phylogenetic relationship are scientific theories, and, like all such theories, are to be judged by their ability to account for (or explain) relevant observation. The observations comprise shared characters—points of similarity—among organisms, and hypotheses of genealogical relationship are potentially able to explain such similarities as the result of inheritance. Usually no one possible genealogy can explain all observed similarities in that way, so that some homoplasy must be postulated. Homoplasies by definition are similarities that cannot be accounted for as resulting from inheritance; such similarities are not explained by a postulated genealogy, but are instead coincidental from the standpoint of the genealogical hypothesis.

If a putative genealogy is consistent with a single origin of some trait, then all observed similarities in that feature can be accounted for in terms of inheritance. If a second postulated phylogeny instead divides the species showing a trait into two widely separated groups, so requiring two origins of the feature, then the similarity between the groups cannot be accounted for as inheritance, though the similarity within the groups still can be so explained. If one of those two groups were split further into two according to a third genealogical hypothesis, still fewer of the observed similarities would be explicable in terms of common descent. With each additional requirement for a separate origin of the trait, the explanatory power of the genealogical theory is reduced still further.

These simple observations provide a means of establishing preference among alternative hypotheses of relationship. The postulated genealogy that requires the fewest independent origins of features is the one best able to explain observed similarities. That genealogy is most consistent with available data in the sense that it minimizes cases in which apparent correspondences between organisms need be dismissed as coincidental. This criterion, of course, is parsimony, and so it is seen that the fundamental scientific criteria for evaluation of theories—explanatory power and conformity to observation—lead to preference for most parsimonious trees.

These words are not Hennig's, but the result is the same. Hennig's (1966: 121) auxiliary principle requires that shared apomorphies be taken as indicating kinship whenever possible, and they do indeed indicate kinship unless they originated independently. This means that interpretations of independent origin are to be avoided: postulated homoplasies ought not to be multiplied beyond necessity. And that is parsimony again.

This derivation has presumed little about evolution beyond viewing that process as descent with modification, and so it provides a way of studying phylogeny without presupposing some unsubstantiated model. It nonetheless has implications for models and their use. A worthwhile model would have to postulate specific characteristics of evolution, and, since such characteristics could not be known a priori, would therefore comprise a scientific theory in itself. Like any theory, again, a model would

then have to be judged on its explanatory power. To see the implications of this, consider a common example.

It is not uncommon (among nonphylogeneticists) to try to study evolution by phenetic cluster analysis, and this is typically defended by recourse to a model under which rates of evolution are invariable, or nearly so. If this is done with "overall" (raw) similarities calculated from character data, then the underlying observations are points of similarity, just as before. The tree resulting from clustering might be largely the same as one arrived at by parsimony. If so, then the phenetically-derived clustering, interpreted as a genealogy, is well able to account for the similarities in the data. But if the phenogram interpreted as a genealogy requires many more homoplasies than the parsimonious arrangement, then it is much less able to account evolutionarily for the observed similarities. Though ostensibly based on those similarities, the phenogram effectively dismisses them as coincidental.

If a model (a theory) has inferior ability to explain observations, those observations comprise evidence against the model. When a model of rate constancy is used to justify phenetic clustering as an evolutionary method, and when that method yields an unparsimonious arrangement, consequently, the data call the model itself into question. If it develops that the phenetically-derived clustering is parsimonious, on the other hand, then the data are in accord with the model. This means that models—evolutionary theories—can be studied. They can be corroborated, when they are in accord with a parsimonious—explanatory—arrangement, and they can be refuted, when found to have inferior explanatory power. But it is also seen that no theoretical model can justify preference for an unparsimonious arrangement, for whenever a model seems to dictate such a conclusion in practice, that is evidence against the model.

The ideas just presented have been summarized in only the briefest form; they are discussed much more fully by Farris (1983; see also Farris, Kluge and Mickevich, 1982). Duncan offers no rebuttal to those papers—like much of the pertinent literature, he does not mention them at all. It is worth noting in particular that Sober (1983) attributes the same significance to the auxiliary principle as that described above, and Farris, Kluge and Eckardt discussed that principle as well. While Duncan cites both those papers, he makes no mention of this idea. Apparently he was unable to devise a way to reconstrue Hennig's auxiliary principle so that it would appear to favor clique analysis.

Indeed, the only advocate of clique methods who has mentioned Hennig's auxiliary principle is Meacham (1984: 28), who describes it as "the fundamental rule of cladistic inference: the possession of the same state by two [taxa] is evidence of descent from a common ancestor that possessed the same state that they retain. In other words, no unnecessary character state transitions are hypothesized." Meacham seems to regard this idea as consistent with clique procedures, though how that could be he does not explain. Avoiding unnecessary transitions is plainly parsimony, and it is just as plain that clique techniques in general do postulate unnecessary transitions, so violating Hennig's auxiliary principle. It seems generally true of clique advocates that they pay only lip service to Hennig's ideas, having no intention of following the implications of his thought.

It is readily seen even from this cursory treatment that these results reveal profound defects in Duncan's position. The aspects of evolution concerning which he wishes to make statements—common ancestry, cladogenesis, anagenesis, and homoplasy—are all phylogenetic conclusions. None of them can be studied independently of an inferred phylogeny, and no finding on these subjects can sensibly be regarded as better founded than the relevant aspects of the phylogeny on which it is based. Parsimonious phylogenetic inferences are chosen to conform to, and account for, available information as well as possible. Any alternative phylogeny would be weaker as a theory precisely because it would conform less well with data: the data would offer grounds for discarding the tree. The same would apply to conclusions on any facet of evolution drawn from a putative phylogeny. If theories on those aspects are to be chosen to conform to observation—if the study of evolution is to be empirical—then they must be based on parsimonious inferences. Recourse to other methods has no other effect than discarding of evidence and consequent loss of explanatory power.

Duncan contends that methods may be chosen according to models and that the use of methods other than parsimony analysis is "reasonable." He does not explain why it is reasonable to base conclusions on models—that is, on presuppositions—rather than on evidence, nor why it is reasonable to discard evidence, rather than seek conformity to it. But that is hardly his only omission, for in fact he does not explain his position at all. One would think from his claim that clique methods are justified by their dependence on some evolutionary model, but even if such a justification were possible it would require justification of the model itself, and Duncan offers no hint of what that model might be—let alone how it might be justified.

But if Duncan's talk of models is then merely a pretext, lacking any underlying substance, it is scarcely different from the rest of his defense of clique methods. To the extent that his treatment is not simply self-contradictory, it consists of nothing beyond misrepresentation. Nor is Duncan unique among clique advocates in this respect. Estabrook (1979) likewise claimed Hennigian status for clique methods, but—as Farris and Kluge (1979) have documented—produced not a single quotation from Hennig in support of his assertion. Estabrook and Anderson (1979: 189) declared that a highly satisfactory tree at which they had arrived “will attest to the utility of the method of compatibility analysis.” They did not bother to inform their readers that, as Farris and Kluge discovered, that tree was in fact a parsimonious arrangement. Gardner and La Duke (1979: 206) maintained that the assumptions of clique methods are “more acceptable” than those of other procedures. Much like Duncan, they did not discuss the assumptions of any method, let alone disclose those of clique analysis.

There is nothing mysterious about the basis of Hennigian phylogenetic analysis, that is, of the parsimony criterion. It relates preference among hypotheses of kinship to their ability to account for observed character distributions. One can only wonder what motivates clique methods, and one must wonder as well why their proponents go to such lengths to avoid disclosing the premises of those procedures.

Literature Cited

- Churchill, S. P., E. O. Wiley and L. A. Hauser. 1984. A critique of Wagner groundplan-divergence studies and a comparison with other methods of phylogenetic analysis. *Taxon* 33: 212–232.
- Duncan, T. 1984. Willi Hennig, character compatibility, and the “dendrogrammaceae” revisited. *Taxon* 33: 698–704.
- , R. B. Phillips and W. H. Wagner, Jr. 1980. A comparison of branching diagrams derived by various phenetic and cladistic methods. *Syst. Bot.* 5: 264–293.
- Estabrook, G. F. 1979. Some concepts for the estimation of evolutionary relationships in systematic botany. *Syst. Bot.* 3: 146–158.
- and W. R. Anderson. 1979. An estimate of phylogenetic relationships within the genus *Crusea* (Rubiaceae) using character compatibility analysis. *Syst. Bot.* 3: 179–196.
- , J. G. Strauch and K. L. Fiala. 1977. An application of compatibility analysis to the Blackiths' data on orthopteroid insects. *Syst. Zool.* 26: 269–276.
- Farris, J. S. 1983. The logical basis of phylogenetic analysis. In: N. I. Platnick and V. A. Funk (eds.), *Advances in cladistics*, Volume 2. Columbia Univ. Press, New York.
- and A. G. Kluge. 1979. A botanical clique. *Syst. Zool.* 28: 400–411.
- , ——— and M. J. Eckardt. 1970. A numerical approach to phylogenetic systematics. *Syst. Zool.* 19: 172–191.
- , ——— and M. F. Mickevich. 1982. Phylogenetic analysis, the monothetic group method, and myobatrachid frogs. *Syst. Zool.* 31: 317–327.
- Felsenstein, J. 1983. Parsimony in systematics: Biological and statistical issues. *Ann. Rev. Ecol. Syst.* 14: 313–333.
- Gardner, R. C. and J. C. La Duke. 1979. Phyletic and cladistic relationships in *Lipochaeta* (Compositae). *Syst. Bot.* 3: 197–207.
- Hennig, W. 1966. *Phylogenetic systematics*. Univ. Illinois Press, Urbana.
- . 1983. *Stammesgeschichte der Chordaten*. Verlag Paul Parey, Hamburg.
- Meacham, C. A. 1984. The role of hypothesized direction of characters in the estimation of evolutionary history. *Taxon* 33: 26–38.
- Sober, E. 1983. Parsimony in systematics: Philosophical issues. *Ann. Rev. Ecol. Syst.* 14: 335–358.
- Wagner, W. H., Jr. 1969. The construction of a classification. In: C. G. Sibley (Chairman), *Systematic biology*. Pub. 1692, U.S. Nat. Acad. Sci., Washington.

*James S. Farris, Department of Ecology and Evolution,
State University of New York, Stony Brook, NY 11794, U.S.A.; and
Arnold G. Kluge, Museum of Zoology, The University of Michigan,
Ann Arbor, MI 48104, U.S.A.*