The Role of Hypothesized Direction of Characters in the Estimation of Evolutionary History
Author(s): Christopher A. Meacham
Source: *Taxon*, Vol. 33, No. 1 (Feb., 1984), pp. 26-38
Published by: International Association for Plant Taxonomy (IAPT)
Stable URL: http://www.jstor.org/stable/1222026
Accessed: 18/06/2014 22:28

TAXON 33(1): 26–38. FEBRUARY 1984

THE ROLE OF HYPOTHESIZED DIRECTION OF CHARACTERS IN THE ESTIMATION OF EVOLUTIONARY HISTORY

Christopher A. Meacham[1]

Summary

The effect of directing characters on cladistic analysis is discussed. Characters can conflict because of the character state membership of the EUs, the ordering of character states, or direction. Conflicts of characters were analyzed in twenty-three data sets containing a total of 1024 characters. At least 93 percent of character conflicts in these data sets were not due to incorrectly hypothesized direction. Most character conflicts are caused by similarities among EUs that are due to parallelisms or reversals that were not recognized by the systematist. Arguments against a priori directing of characters are presented. It is recommended that an undirected analysis be performed first and that the undirected tree be directed subsequently.

To document and explain the diversity of organisms is one of the fundamental goals of biology. The theory of evolution provides a general explanation of the mechanism that created the multitude of life forms that inhabit the earth today. Many systematists are currently interested in the possibility of reconstructing the evolutionary history of organisms from the attributes they possess. The development of an objective basis for estimating the branching pattern of evolutionary history is the concern of cladistics, a subdiscipline of quantitative systematics.

Several recent authors (Crisci and Stuessy, 1980; Stevens, 1980; Watrous and Wheeler, 1981; see Stevens, 1980 for earlier references) have discussed the problem of determining the evolutionary direction of character trends. These authors consider the designation of an ancestral state a "critical first step" in the estimation of evolutionary history. Many systematists presume that cladistics requires the use of directed characters and that any method that does not require hypothesized evolutionary direction of characters is phenetic. An evaluation of the contribution that character direction makes to the estimation of evolutionary history is therefore in order.

The Definition of a Character

We begin with a group of organisms called the *study collection.* Each kind of organism present in the study collection is called an *evolutionary unit* (EU) and the study collection as a whole is assumed to have some evolutionary history that can be represented by a tree diagram. A qualitative character is an operationally defined basis for comparing the members of the study collection, the EUs, that allows us to decide which EUs are alike and which EUs are different. EUs that are alike with respect to the character are said to share the same *character state.* In this way, a character allows us to recognize discrete classes of EUs. Because possession of a particular character state defines a particular class, at times it will be convenient for our purposes to call the class itself the character state. Two EUs that are alike with respect to the character are members of the same character state. In short, a qualitative character is a set of character states, which are mutually exclusive,

---

[1] Botany Department, University of Georgia, Athens, GA 30602, U.S.A.

exhaustive subsets of the study collection. If, for example, counting chromosomes were the operational basis for comparing EUs, and the organisms in the study collection had chromosome numbers of either $n = 6$ or $n = 7$, then this character would have two character states; one consisting of all those EUs in the study collection with $n = 6$, and the other would be the set of EUs with $n = 7$. It may seem peculiar at first to view characters as sets of subsets, but cladistic analysis is essentially the analysis of the relationships among characters (cf. Hull, 1979). The relationships of qualitative characters can be completely described as relationships among sets. A set theoretical view of characters enables one to gain more direct access to the properties of characters, and it becomes simpler to visualize character relationships.

The creation of a qualitative character is a complex operation that requires a great deal of biological interpretation and intuition. Because synthesis and interpretation are so predominant in character construction, this process resembles an act of invention more than discovery. In the course of the following discussion, I hope to show that the invention of the operational basis for comparison is, in fact, the critical first step in the creation of characters. It is hoped that this recognition will focus attention on this important aspect of cladistic analysis.

Character Types

Although inventing a qualitative character is the first step in developing a character for use in cladistic analysis, few current methods can deal directly with this type of character. Most methods require an hypothesis of the historical relationships among the character states. Some methods of analyzing characters require hypothesizing only the *order* in which the character states evolved without specifying the direction. For example, if we had chromosome numbers $n = 6$, $n = 7$, and $n = 8$, we might feel confident that 7 should go between 6 and 8 even though we would be uncertain about which of these states is ancestral. Other methods require, besides hypothesized order, the additional hypothesis of which character state is ancestral, an hypothesis of *direction.* Thus there are three levels of hypotheses one can make starting with a qualitative character as summarized below.

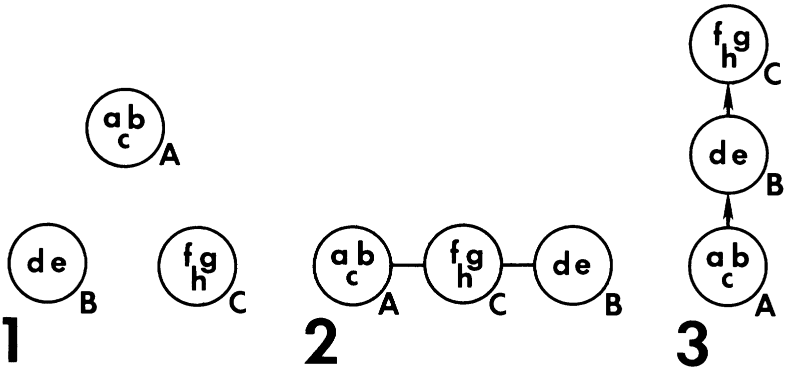
Fig. 1 is a diagram of a three-state qualitative character. The basis for comparison allows us to recognize three kinds of EUs. EUs *a, b,* and *c* are alike with respect to this basis in possessing the feature that distinguishes state *A* and are therefore in the same state; *d* and *e* both possess state *B,* and *f, g,* and *h* all possess state *C.* Characters without hypothesized relationships among their states are unordered and undirected. Qualitative characters without hypotheses of ordering or direction will be called *unordered.*

Fig. 2 shows a three-state character that is ordered but undirected. Character state *C* is hypothesized to be evolutionarily between *A* and *B,* but the historical direction of trends is unspecified. An undirected, ordered character will hereafter be called simply *undirected.*

Fig. 3 shows a three-state character that is both directed and ordered. In this character, *A* is the ancestral state. Conventionally, direction is indicated by orientation of the diagram, but here the hypothesized direction is explicitly indicated by arrows. Characters like this will be called *directed.*

Note that the characters in Figs. 2 and 3 differ in ordering. In Fig. 3, *B,* is between *A* and *C,* although in Fig. 2, *C* is between *A* and *B.* The third possible ordering for a three-state character would have *A* in the middle. There are three possible orderings for an undirected three-state character. If direction is hypothesized for any of these undirected three-state characters, it can, in each case, be done three ways depending on which state is designated ancestral. Thus, a three-state qualitative character can give rise to any of three undirected characters, each of which, in turn, can give rise to any of three directed characters, a total of nine possibilities.

To some extent, the distribution of EUs among the character states is more certain than the hypothesized relationships among the states, because the assignment of EUs to states



Figs. 1–3. Examples of the types of characters. The lower-case letters are labels of EUs. EUs that belong to the same character state are enclosed in a circle. The upper-case letters are labels of the character states. 1. A three-state unordered character. 2. A three-state undirected character. 3. A three-state directed character.

was accomplished by direct observation. The relationships hypothesized to exist among the states, however, do not have such a firm observational basis.

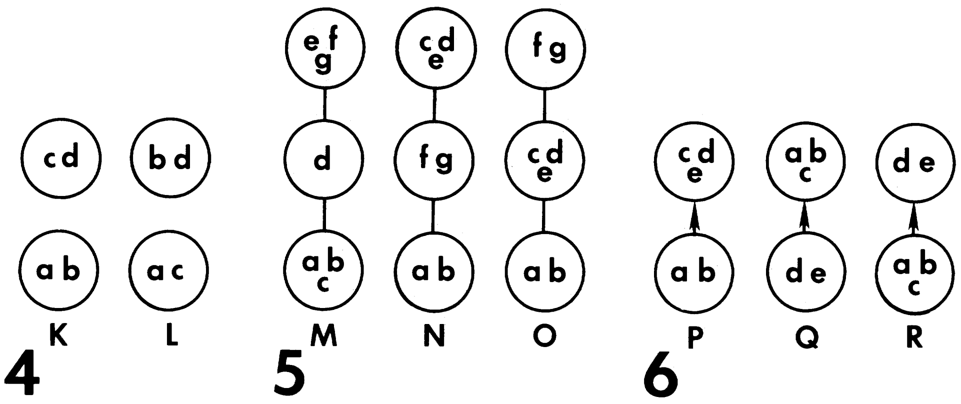
How Characters are Used

Cladistic methods share the point of view that because the features of descendant organisms are inherited from their ancestors, the features of extant organisms preserve, in some sense, a record of the evolutionary history of the organisms to which they belong. Cladistic methods seek to reconstruct the evolutionary history of organisms from the information contained in their characters. All methods for analyzing qualitative characters use a principle that could justifiably be called the fundamental rule of cladistic inference: the possession of the same state by two EUs 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. That character information is used to infer the character states of organisms not included in the study collection is one of the identifying features of cladistic methods. This fundamental rule that the possession of the same state by two EUs is evidence of retention of that state through descent from their most recent common ancestor is the "auxiliary principle" of Hennig (1966, p. 121) and axiom AIII of Farris et al. (1970) and is basic to the concept of the true cladistic character of Estabrook et al. (1975). If characters actually reflected evolutionary history in this way, there would be no conflicts among characters (Estabrook et al., 1975) and the estimation of evolutionary history would be straightforward. However, characters often conflict.

The Nature of Character Conflicts

Basically, two characters conflict when the evolutionary relationships that they suggest in the sense of the fundamental rule as given above cannot both be historically true. That is, if the evolutionary relationships as inferred from one character are correct, then the evolutionary relationships as inferred from the other are not and vice versa. In this case, it is not possible to find an evolutionary tree that would allow the inferences made from both characters to be correct. On any evolutionary tree, at least one of the characters must show parallelism or reversal and hence, there must be some pair of EUs that share the same state of one of the characters and yet did not inherit that state from their most recent common ancestor.

Several authors have, apparently independently, discovered conditions under which



Figs. 4–6. The types of character conflicts. 4. Two unordered binary characters that conflict because of the character-state membership of the EUs. 5. Three ordered characters; *M* and *N* conflict because of ordering, *M* and *O* do not. 6. Three directed characters; *P* and *Q* conflict because of direction, *P* and *R* do not.

characters conflict (Wilson, 1965; Hennig, 1966, pp. 120–121; Le Quesne, 1969; Platnick, 1977). (There is a misprint in Hennig's discussion: on page 120, third line from the bottom, "A" should be "B."). The mechanics of how one determines whether characters conflict has been discussed elsewhere (Estabrook and Landrum, 1975; Fitch, 1975; Estabrook and Meacham, 1979; Meacham, 1980, 1981a, 1983; Estabrook, in press). Only the kinds of conflicts will be covered here.

The most severe type of conflict is that of *membership.* As Le Quesne (1969) discovered, two characters can conflict because of the way EUs are distributed among the character states. If there exist two states in each of two characters such that all four possible combinations of character states are present in the study collection, then no tree can be found that would allow valid inferences by the fundamental rule from both characters. For the two binary characters, *K* and *L* in Fig. 4, EU *a* is in the bottom state of both characters, EU *d* is in the top state of both, *b* is in the bottom state of *K* and the top state of *L,* and *c* is in the bottom state of *L* and the top of *K.* These two characters conflict because of the character-state membership of the EUs (see Meacham, 1980). No matter what direction or ordering is hypothesized, the evolutionary relationships of these EUs as inferred (by the fundamental rule) from one of these characters conflicts in a basic way with the relationships as inferred from the other character. All three types of characters, unordered, undirected, and directed, can conflict because of membership.

A less severe type of conflict is that of *ordering.* Sometimes two characters conflict if their states are ordered in one way but do not conflict if ordered in another. Fig. 5 shows three undirected character state trees. Characters *M* and *N* conflict, but characters *M* and *O* do not. Note that *N* and *O* have the same states; only the ordering of the states has been changed. All three types of characters can conflict because of ordering. However, because binary (two-state) characters have only one possible ordering, they cannot conflict because of ordering. Binary characters can always be considered ordered.

The least severe type of conflict is that of *direction.* Fig. 6 shows three binary characters. Characters *P* and *Q* conflict. However, characters *P* and *R,* which is *Q* with the direction reversed, do not conflict. Only directed characters can conflict because of direction.

Estabrook et al. (1976) proved that if each pair of characters in a set of directed characters does not conflict, the set as a whole does not conflict. McMorris (1977) proved a similar result for undirected characters. However, Fitch (1975, 1977) presents an example of three unordered characters that, taken a pair at a time, do not conflict. Yet, as a set, they do conflict. For two of these characters to be nonconflicting, they must have a certain ordering imposed on them. And it turns out that the ordering that must be imposed on character 1 so that it does not conflict with character 2 is not the same ordering that must be imposed on 1 if it is not to conflict with 3. Pairwise absence of conflicts does not imply setwise absence of conflicts for unordered characters.

The Definition of a Method of Cladistic Analysis

Because all current methods of cladistic analysis for qualitative characters are based on the fundamental rule that the possession of the same state by two EUs is evidence of descent from a common ancestor that possessed the same state that they retain, all methods will produce identical results if presented with characters that do not conflict. When characters do conflict, the conflicts must be resolved. The factor that distinguishes different methods is the rule that is used to resolve conflicts (Meacham, 1980; Felsenstein, 1982). One reasonable rule is to resolve the conflicts so that the smallest number of character state transitions is required to explain the observed distribution of character states. This criterion is the basis of the Wagner parsimony technique (Kluge and Farris, 1969; Farris et al., 1970). The method of character compatibility analysis is directed more toward explicitly identifying conflicts than resolving them. Characters that conflict are incompatible; characters that do not conflict are compatible. Because characters that allow correct influences by means of the fundamental rule cannot conflict, it may be reasonable to base estimates of evolutionary history on some, perhaps the largest, set of nonconflicting characters, resolving conflicts in favor of these at the expense of others. Because both these techniques are based on the fundamental rule, in data sets with fewer conflicts, the results obtained from parsimony and suggested by character compatibility are largely concordant. Hennig (1966, p. 121) was certainly aware of character conflicts and suggests that they be resolved by further study of the organisms. But because he was not explicit about how to proceed when further study does not resolve the conflict, his method lacks a crucial element. Felsenstein (1982) has called this "Hennig's dilemma."

In the discussion that follows, the effect of hypothesized direction on the methods of compatibility analysis and Wagner parsimony will be examined. Because many computer programs that currently perform parsimony analysis (e.g., Kluge and Farris, 1969) are not guaranteed to discover the truly most parsimonious tree, but only find an approximation, these programs may not behave exactly as described. Any algorithm that did always discover the most parsimonious tree would be influenced by hypothesized direction as described below. The algorithm developed by Hendy et al. (1978) is one of this class, but its use is limited to data sets that contain about 12 or fewer EUs.

Hypothesized Direction is a Restriction

The fundamental rule does not itself embody any statement about supposed direction of trends between character states, only the relationship of EUs within the same state. For this reason, both parsimony and character compatibility can be performed in an undirected sense (Farris, 1970; Estabrook and Meacham, 1979). Any conflicts among characters that are due solely to direction can be eliminated by making the characters undirected. Conflicts among undirected characters are due to either membership or ordering and cannot be eliminated by directing the characters in some other way. When we do an undirected analysis, we discover the best result with respect to our criterion for resolving conflicts. For the same data set, a most parsimonious tree obtained from directed characters can never be shorter than an undirected tree from the same characters without direction, because any directed tree can be changed to an undirected one simply by ignoring the direction. It is evident from these considerations that hypothesizing the direction of characters in cladistic analysis is essentially a restriction (cf. Sneath and Sokal, 1973, p. 325). Whether we are doing parsimony or character compatibility, directing characters can only increase conflicts, never remove them.

Hypothesizing Direction is Identical to Hypothesizing an Ancestor

Directing a character consists of identifying one state as ancestral. By hypothesizing a state of each character as ancestral, we postulate the combination of character states possessed by the most recent common ancestor of the study collection. A directed tree is constrained to have the ancestor at the root. Consider what would happen if we carried out an undirected parsimony analysis on the study collection with the hypothetical ancestor added. The most parsimonious undirected tree that includes the ancestor cannot be longer than the directed tree. Because if it were longer, we could ignore the direction on the directed result to get a shorter undirected tree that includes the ancestor by placing the ancestor at the root. Also, the most parsimonious undirected tree including the ancestor cannot be shorter than the directed result, because we could root the undirected result at the ancestral node to obtain a shorter directed tree. For any most parsimonious directed tree, there is (with regard to branching structure) an identical most parsimonious undirected tree that includes the hypothetical ancestor (Farris, 1970). The similar principle for character compatibility is that two directed characters are incompatible if and only if the corresponding undirected characters are incompatible for the study collection plus hypothetical ancestor.

The Effect of Adding a Hypothetical Ancestor

Adding a hypothetical ancestor to an undirected analysis is identical to directing characters with regard to the determination of branching structure. Therefore, the role of direction in the estimation of evolutionary history can be evaluated by considering the effect of adding a hypothetical ancestor.

If all the characters allow correct inferences by the fundamental rule and, hence, do not conflict, then adding a correctly assessed ancestor will make no difference at all. No extra steps must be necessary if the ancestor is added, because an extra step must be a parallelism or a reversal, contrary to the assumption that all characters allow valid inferences by the rule. In the case of perfect characters like these, the true ancestor must fall directly on the tree obtained by undirected analysis. (If the ancestor has a state not present in the study collection, an extra step will be required; but this transition will not already occur on the tree and so will not be a parallelism or reversal.) An incorrectly hypothesized ancestor may distort the results if characters are not misleading. Thus, if our data are perfect, a hypothetical ancestor will hurt or make no difference at all in reconstructing the branching pattern of evolution.

Most data sets contain many conflicts and are hence imperfect. The effect of adding a hypothetical ancestor to a data set like this depends, of course, on the data set. An absolute limit of the effect can be deduced, however. Assume we have found a most parsimonious undirected tree excluding the hypothetical ancestor. It would be possible to obtain a directed tree by attaching the ancestor to this tree. Consider the EU in the study collection that is most similar to the hypothesized ancestor. Assume that they are identical in all characters but seven. The ancestor could be attached to this tree by seven character state transitions between the ancestor and its most similar EU. This may not be the shortest directed tree, but the actual shortest tree can be no more than seven steps longer than the shortest undirected tree. Even though this longer directed tree may differ substantially from the undirected tree, the number of characters that distinguish the hypothetical ancestor from the most similar extant EU is an upper limit on the number of steps by which the most parsimonious directed and undirected trees can differ. A similar relationship holds for character compatibility. In the same way that adding the hypothetical ancestor to an undirected tree may introduce extra steps in at most the number of characters that distinguish the hypothesized ancestor from the most similar observed EU, changing an undirected character compatibility analysis to a directed analysis by adding an ancestor will only influence the compatibility relationships of those same characters, increasing their incompatibilities (conflicts) with other characters or leaving them unchanged.

These observations show that the effect of hypothesized direction on cladistic analysis is limited by the extent to which the hypothesized ancestor differs from the most similar observed EU. In practice, the hypothetical ancestor is usually quite similar to some EU in the study collection, and consequently, relatively few additional conflicts are introduced into the data sets. The effect of hypothesizing direction is often minimal. To determine which conflicts among characters are due solely to direction, one can perform both a directed and an undirected character compatibility analysis to see which incompatibilities are removed by making the characters undirected.

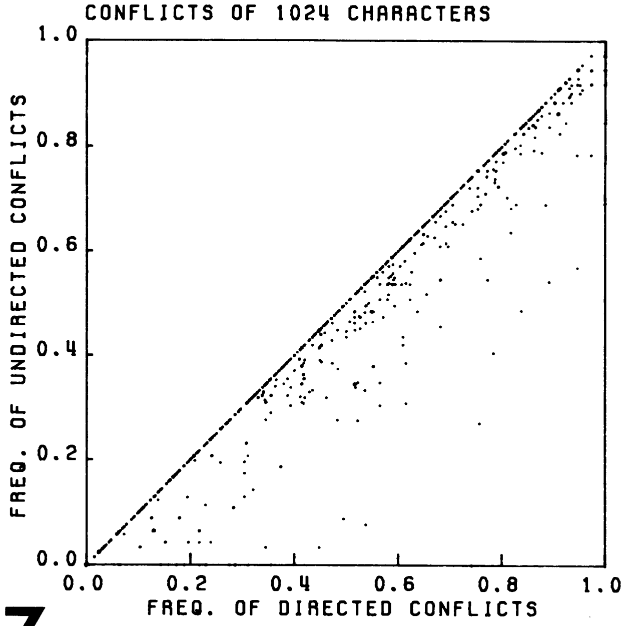
These considerations also shed light on another interesting relation between directed and undirected analyses. If the hypothesized ancestor is identical with any EU in the study collection, no conflicts of direction will be added to the data (Farris, pers. comm.). In the same way that adding an ancestor to an undirected analysis will produce a directed result, performing a directed analysis with a hypothetical ancestor that is identical to any observed EU will produce results that are identical with undirected analyses, ignoring direction.

Undirected Methods are Cladistic

Direction, we see, is an attribute of the data, not of the method being used. An undirected technique can be used to produce a directed result simply by including the hypothetical ancestor in the data set. Using an observed EU as an ancestor in a directed analysis will produce a result with the same branching pattern as an undirected analysis. The only factor that distinguishes the directed method from the corresponding undirected one is that the directed method *requires* the inclusion of the hypothetical ancestor by the specification of the ancestral states. The logic of inferring the tree from the data is the same for both directed and undirected techniques. I suggest that it is the *logic* of character analysis that distinguishes cladistic techniques from phenetic ones. Because the logic of character analysis in cladistic methods remains the same whether a hypothetical ancestor is introduced into the data set or not, we must conclude that the use of directed characters does not in itself distinguish cladistic techniques from phenetic ones.

Analysis of Data Sets

Table 1 lists 23 cladistic data sets, which include a total of 1024 characters. A sample of this size and variety seems adequate to form a basis for general conclusions. Using the program CLINCH, written by Kent L. Fiala, the conflicts of characters in these data sets were examined when the characters were directed and when they were undirected. Fig. 7 is a scatter plot of the conflicts of the 1024 characters. Each dot represents one character. The horizontal coordinate of the dot corresponding to each character is the frequency of conflicts with other characters in its own data set when all characters are directed. The vertical coordinate is the frequency of conflicts when all characters are undirected. (The frequency of conflicts for a given character is the number of observed conflicts divided by one less than the number of characters in its own data set.) The frequency of conflicts for most characters is the same whether directed characters are used or not. Of the 27,447 pairwise comparisons of characters made in these 23 data sets, there are 11,602 pairwise conflicts when the characters are directed and 10,892 when they are undirected. There are thus 710 conflicts of direction among the directed characters. Even if *all* these conflicts were due to an error in the hypothesized direction (hypothetical ancestor), this would account for only 6.12 percent of the conflicts among the directed characters. Most conflicts in these 23 data sets are not caused by incorrectly directed characters.



7

Fig. 7. Scatter plot of frequency of directed conflicts vs. frequency of undirected conflicts for the 1024 characters in the 23 data sets in Table 1.

Why Hypothesize an Ancestor?

The problem with identifying the ancestral character states is, of course, that the ancestor is not available for direct examination. The ancestral states are often inferred from the distribution of character states found in taxa that are a priori considered related to the study collection. The character states for the extant EUs were obtained by direct observation. Because in a directed analysis the ancestor is treated the same as any EU, the inclusion of a hypothetical ancestor in a cladistic analysis on an equal footing with the observed EUs implies that we are as certain of the character states of the ancestor as we are of those of the study collection. This is surely not the case. It seems that characters are directed in the hope that this will remove conflicts from the data. But this is never true. We could hypothesize any other unobserved descendant of the most recent common ancestor of the study collection that we might imagine to have existed and include its hypothesized character states in the analysis. This would certainly be called an ad hoc hypothesis. Yet, its influence on discovering the branching pattern of evolution would be the same as adding a hypothetical ancestor. And, under reasonable criteria for guessing what states it might have had, would be based on no less firm a foundation.

As a concrete example, assume we have performed a directed analysis on a genus. As we have seen, this is logically equivalent to introducing a hypothetical ancestor into the data set. At some later time we carry out a directed analysis on the family that includes this genus. Would we be justified in including the hypothetical ancestor for the genus in the family analysis? I am sure most systematists would object on the grounds that there is no end to the number of such intermediate ancestors that could be hypothesized. Yet if it is not proper to include the hypothetical ancestor of the genus in the family analysis, why

Table 1. Twenty-three data sets containing 1024 characters used in analysis.

Taxon	Rank	Major group	Author(s)	Characters	EUs
Avena sativa	Species	Plant	Baum and Estabrook (1978)	23	16
Balistoidea	Superfamily	Fish	Matsuura (1979)	32	33
Berberidaceae	Family	Plant	Meacham (1980)	30	15
Beryciformes	Order	Fish	Zehren (1979)	70	26
Oneirodidae	Family	Fish	Pietsch (1974)	30	13
Charadriiformes	Order	Bird	Strauch (1978)	70	227
Cichlidae	Family	Fish	Cichocki (1976)	62	47
Cranchiidae	Family	Mollusk	Voss and Voss (1982)	14	14
Crusea	Genus	Plant	Estabrook and Anderson (1979)	58	17
Fundulus & *Profundulus*	Genera	Fish	Farris (1969)	35	34
Gymnophiona	Order	Amphibian	Nussbaum (1979)	43	13
Leptopodomorpha	Suborder(?)	Insect	Schuh and Polhemus (1980)	47	8
Lipochaeta diploids	Genus (part)	Plant	Gardner and La Duke (1979)	32	18
Lipochaeta tetraploids	Genus (part)	Plant	Gardner and La Duke (1979)	32	9
Macrouridae	Family	Fish	Okamura (1970)	57	15
Myrceugenia	Genus	Plant	Landrum (1981)	38	45
Orthoptera	Order	Insect	Blackith and Blackith (1968)	92	12
Pygopodidae	Family	Lizard	Kluge (1976)	86	21
Ranunculus hispidus	Spp. complex	Plant	Duncan (1980)	14	87
Tetraodontiformes	Order	Fish	Winterbottom (1974)	66	11
—	—	Fish	Unpublished	30	30
—	—	Fish	Unpublished	36	15
—	—	Fish	Unpublished	27	30

is it proper to include this same ancestor when analyzing the genus? Our only justification for including it is that we wish to obtain a directed result. The point I wish to emphasize is that a priori directing of characters is not the only means of producing a directed estimate.

Another argument against a priori directing of characters is that after the cladistic analysis has been carried out, some characters show many transitions on the tree. If we believe that a character has undergone many parallelisms or reversals during the evolution of the study collection, it would seem rash to think we can determine its primitive condition by examining its distribution in related taxa. It seems that hypothesis of direction should be reserved for characters that, according to the available evidence, appear to be conservative.

How to Produce a Directed Estimate

Crisci and Stuessy (1980), Stevens (1980), and Watrous and Wheeler (1981) and earlier authors (e.g., Colless, 1969; Kluge and Farris, 1969; Lundberg, 1973) have discussed in detail many criteria for hypothesizing the direction of character trends. These criteria can also be used to suggest the direction of character state transitions on undirected trees. This approach has several advantages. It frees the analysis from a priori restrictions, removing conflicts that must be resolved. It allows us to be open to possibilities we might not have thought of. Directions of character state transitions we are more sure of can suggest the direction of those we are less sure of. It lets the characters speak for themselves. And as argued above, a perfect data set will not be improved by the addition of a hypothetical ancestor. I recommend that an undirected tree analysis be performed first, and then that the undirected tree be directed by explicit arguments (see Lundberg, 1973).

Common-Equals-Primitive

Estabrook (1977) discussed this criterion for determining direction and pointed out that fewer characters will conflict if directed in this way. Estabrook went on to suggest that this fact might explain why this criterion was so widely employed. He suggested that in the effort to achieve maximum agreement among directed characters, workers had designated the most common state of each character as the primitive state without explicitly realizing that this criterion will always produce maximum agreement and so, does not necessarily indicate the actual primitive state. Based on an earlier proof by McMorris (1977), Estabrook and Meacham (1979) proved that characters whose binary factors are directed common-equals-primitive are compatible if and only if they are compatible as undirected characters. Common-equals-primitive characters cannot conflict in direction. Hence if this criterion is used in compatibility analysis, the result should be considered undirected and the root ignored. CLINCH, a program to perform character compatibility analysis, uses the fact that common-equals-primitive characters cannot conflict in direction to perform undirected analysis. Common-equals-primitive should not be considered a criterion to discover the actual primitive state. It is instead a logical device to eliminate conflicts due solely to direction.

Perspective

The preceding arguments are not meant to diminish the importance of direction to a complete description of evolutionary history. Evolution, because it occurs through time, is fundamentally a directed process. A priori directing of characters is only one way of achieving a directed estimate of evolutionary history. These arguments suggest that a priori directing of characters is not crucial to the logic of Wagner parsimony or character compatibility. The logic of deriving the branching structure of the tree is the same for these two techniques whether the characters are directed or not. Consequently, undirected Wagner parsimony and undirected character compatibility analysis have as much right to be called cladistic as the corresponding directed techniques. A more flexible approach to developing a directed estimate of evolutionary history should be considered.

Conclusion

In the absence of character conflicts, the fundamental rule directly gives us the most reasonable evolutionary estimate. No other criteria are necessary. The task of cladistics is essentially the analysis and resolution of character conflicts. Different rules for conflict resolution distinguish the different cladistic methods. It is recommended that an undirected cladistic analysis be performed first. A directed estimate can then be produced by evaluating the evidence for direction of characters in light of the undirected estimate.

In the 23 data sets examined, at least 93 percent of character conflicts were not due to incorrect direction. The incorrect hypothesis of ancestral character states is not the major source of conflicts among characters. This leaves ordering and membership as sources of the remaining conflicts. Because most of the characters examined are binary and cannot conflict because of ordering, the character-state membership of EUs is the predominant source of character conflicts. For any two characters that conflict because of membership, at least one of the states of at least one of the characters arose at least two distinct times in the evolution of the study collection. Thus, the conflict is caused by similarities among EUs that are due to parallelisms or reversals that were not recognized as such by the systematist. I suggest that the development of the basis for comparison by which EUs are assigned to character states is the critical first step in a cladistic analysis and that a careful evaluation of the ways in which similarities and differences among organisms are translated into characters for cladistic analysis will prove worthwhile.

Acknowledgments

I thank William R. Anderson, George F. Estabrook, Wayne P. Maddison, Gerald R. Smith, Ann Stoneburner, and Warren H. Wagner, Jr. for their helpful comments on an earlier draft. I thank S. G. Poss for providing many fish data sets. The views expressed are my own. I am grateful to the National Science Foundation for the support of a Graduate Fellowship and to the Faculty of Science of the Memorial University of Newfoundland for the support of a Postdoctoral Fellowship, which was financed in part by the Natural Sciences and Engineering Research Council of Canada under Grant A-4142 to Dr. William H. E. Day. This paper is based on a portion of a doctoral dissertation (Meacham, 1981b) submitted to the Horace H. Rackham School of Graduate Studies of The University of Michigan.

Literature Cited

Baum, B. R. and G. F. Estabrook. 1978. Application of compatibility analysis in numerical cladistics at the infraspecific level. *Canadian J. Bot.* 56: 1130–1135.

Blackith, R. E. and R. M. Blackith. 1968. A numerical taxonomy of Orthopteroid insects. *Australian J. Zool.* 16: 111–131.

Cichocki, F. G. 1976. Cladistic history of cichlid fishes and reproductive strategies of the American genera *Acarichthys, Biotodoma,* and *Geophagus,* vol. I. Ph.D. Dissertation, Univ. of Michigan, Ann Arbor, Michigan.

Colless, D. H. 1969. The phylogenetic fallacy revisited. *Syst. Zool.* 18: 115–126.

Crisci, J. F. and T. F. Stuessy. 1980. Determining primitive character states for phylogenetic reconstruction. *Syst. Bot.* 5: 112–135.

Duncan, T. 1980. A taxonomic study of the *Ranunculus hispidus* Michaux complex in the Western Hemisphere. *Univ. California Publ. Bot.* 77: 1–125.

Estabrook, G. F. 1977. Does common equal primitive? *Syst. Bot.* 2: 36–42.

———. In press. Phylogenetic trees and character state trees. *In:* T. Duncan and T. F. Stuessy (eds.), *Cladistics: Perspectives on the reconstruction of evolutionary history.* Columbia Univ. Press, New York.

——— and W. R. Anderson. 1979. An estimate of phylogenetic relationships within the genus *Crusea* (Rubiaceae) using character compatibility analysis. *Syst. Bot.* 3: 179–196.

———, C. S. Johnson, Jr. and F. R. McMorris. 1975. An idealized concept of the true cladistic character. *Math. Biosci.* 23: 263–272.

———, ——— and ———. 1976. An algebraic analysis of cladistic characters. *Discrete Math.* 16: 141–147.

——— and L. R. Landrum. 1975. A simple test for the possible simultaneous evolutionary divergence of two amino acid positions. *Taxon* 24: 609–613.

——— and C. A. Meacham. 1979. How to determine the compatibility of undirected character state trees. *Math. Biosci.* 46: 251–256.

Farris, J. S. 1969. The evolutionary relationships between the species of the killifish genera *Fundulus* and *Profundulus* (Teleostei: Cyprinodontidae). Ph.D. Dissertation, Univ. of Michigan, Ann Arbor, Michigan.

———. 1970. Methods for computing Wagner trees. *Syst. Zool.* 19: 83–92.

———, A. G. Kluge and M. J. Eckardt. 1970. A numerical approach to phylogenetic systematics. *Syst. Zool.* 19: 172–189.

Felsenstein, J. 1982. Numerical methods for inferring evolutionary trees. *Quart. Rev. Biol.* 57: 379–404.

Fitch, W. M. 1975. Toward finding the tree of maximum parsimony. *In:* G. F. Estabrook (ed.), *Proceedings of the Eight International Conference on Numerical Taxonomy,* pp. 189–230. W. H. Freeman, San Francisco.

———. 1977. On the problem of discovering the most parsimonious tree. *Amer. Natur.* 111: 223–257.

Gardner, R. C. and J. C. La Duke. 1979. Phyletic and cladistic relationships in *Lipochaeta* (Compositae). *Syst. Bot.* 3: 197–207.

Hendy, M. D., D. Penny and L. R. Foulds. 1978. Identification of phylogenetic trees of minimal length. *J. Theor. Biol.* 71: 441–452.

Hennig, W. 1966. *Phylogenetic systematics.* Transl. by D. D. Davis and R. Zangerl. Univ. of Illinois Press, Urbana.

Hull, D. L. 1979. The limits of cladism. *Syst. Zool.* 28: 416–440.

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.

——— and J. S. Farris. 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.* 18: 1–32.

Landrum, L. R. 1981. The phylogeny and geography of *Myrceugenia* (Myrtaceae). *Brittonia* 33: 105–129.

Le Quesne, W. J. 1969. A method of selection of characters in numerical taxonomy. *Syst. Zool.* 18: 201–205.

Lundberg, J. G. 1973. Wagner networks and ancestors. *Syst. Zool.* 21: 398–413.

Matsuura, K. 1979. Phylogeny of the superfamily Balistoidea (Pisces: Tetraodontiformes). *Mem. Faculty Fisheries Hokkaido Univ.* 26: 49–169.

McMorris, F. R. 1977. On the compatibility of binary qualitative taxonomic characters. *Bull. Math. Biol.* 39: 133–139.

Meacham, C. A. 1980. Phylogeny of the Berberidaceae with an evaluation of classifications. *Syst. Bot.* 5: 149–172.

———. 1981a. A manual method for character compatibility analysis. *Taxon* 30: 591–600.

———. 1981b. The estimation of evolutionary history with reference to the Berberidaceae. Ph.D. Dissertation, Univ. of Michigan, Ann Arbor, Michigan.

———. 1983. Theoretical and computational considerations of the compatibility of qualitative taxonomic characters. *In:* J. Felsenstein (ed.), *Numerical taxonomy: Proceedings of a NATO Advanced Study Institute,* pp. 304–314. *NATO Advanced Study Institute Series G (Ecological Sciences)*, No. 1. Springer-Verlag, Berlin, Heidelberg, and New York.

Nussbaum, R. A. 1979. The taxonomic status of the caecilian genus *Uraeotyphlus* Peters. *Occ. Pap. Mus. Zool. Univ. Michigan* 687: 1–20.

Okamura, O. 1970. Studies on the macrourid fishes of Japan. Morphology, ecology, and phylogeny. *Repts. Usa Marine Biol. Sta.* 17: 1–179.

Pietsch, T. W. 1974. Osteology and relationships of ceratioid anglerfishes of the family Oneirodidae, with a review of the genus *Oneirodes* Lütken. *Sci. Bull. Nat. Hist. Mus. Los Angeles Co.* 18: 1–113.

Platnick, N. I. 1977. Parallelism in phylogeny reconstruction. *Syst. Zool.* 26: 93–96.

Schuh, R. T. and J. T. Polhemus. 1980. Analysis of taxonomic congruence among morphological, ecological, and biogeographic data sets for the Leptopodomorpha (Hemiptera). *Syst. Zool.* 29: 1–26.

Sneath, P. H. A. and R. R. Sokal. 1973. *Numerical taxonomy.* W. H. Freeman, San Francisco.

Stevens, P. F. 1980. Evolutionary polarity of character states. *Ann. Rev. Ecol. Syst.* 11: 333–358.

Strauch, J. G., Jr. 1978. The phylogeny of the Charadriiformes (Aves): A new estimate using the method of character compatibility analysis. *Trans. Zool. Soc. London* 34: 263–354.

Voss, N. A. and R. S. Voss. 1982. Phylogenetic relationships in the cephalopod family Cranchiidae (Oegopsida). *Malacologia* 23: 397–426.

Watrous, L. E. and Q. D. Wheeler. 1981. The out-group comparison method of character analysis. *Syst. Zool.* 30: 1–11.

Wilson, E. O. 1965. A consistency test for phylogenies based on contemporaneous species. *Syst. Zool.* 14: 214–220.

Winterbottom, R. 1974. The familial phylogeny of the Tetraodontiformes (Acanthopterygii: Pisces) as evidenced by their comparative myology. *Smithsonian Contrib. Zool.* 155: 1–201.

Zehren, S. J. 1979. The comparative osteology and phylogeny of the Beryciformes (Pisces: Teleostei). *Evol. Monog.* 1: 1–389.