The Out-Group Comparison Method of Character Analysis

Larry E. Watrous; Quentin D. Wheeler

Systematic Zoology, Vol. 30, No. 1 (Mar., 1981), 1-11.

Syst. Zool., 30(1), 1981, pp. 1–11

THE OUT-GROUP COMPARISON METHOD OF CHARACTER ANALYSIS

LARRY E. WATROUS AND QUENTIN D. WHEELER

Abstract

Watrous, Larry E., and Quentin D. Wheeler (Division of Insects, Field Museum of Natural History, Chicago, Illinois 60605, and Department of Entomology, Cornell University, Ithaca, New York 14853) 1981. The out-group comparison method of character analysis. Syst. Zool., 30:1–11.—An operational rule for analyzing character polarity with out-group comparison is presented and a series of observations, including potential problems in applying the rule, are discussed. The "commonality principle" (="frequency of occurrence," "common equals primitive") for determining character polarity is reviewed and dismissed as a reliable alternative to out-group comparison. Based on the rule and observations, a general method for character analysis is synthesized. [Cladistics; character analysis; out-group comparison; polarity.]

Although many authors have discussed criteria for distinguishing plesiomorphic (primitive) and apomorphic (derived) character states (see Maslin, 1952; Hennig, 1965, 1966; Kluge and Farris, 1969; Marx and Rabb, 1970, 1972; Lundberg, 1972; Ross, 1974; Munroe, 1974; Ekis, 1977; Crisci and Stuessy, 1980; and references therein), there has been little discussion of the deficiencies of some of these criteria, and in some instances there is considerable misunderstanding of what we consider to be the most reliable criterion, that of out-group comparison.

Our goals here are to discuss some problems with one criterion in particular, the so-called "commonality principle" (Eldredge, 1979), and to develop a general method of character analysis based on out-group comparison. Many authors have fully understood out-group comparisons (e.g., Kluge and Farris, 1969; Lundberg, 1972; Ross, 1974), but others have confused the concept with that of commonality. Because out-group comparison can often be invoked where commality is arbitrarily used, and because the resultant hypotheses are subject to rigorous testing, its central position in character analysis is assured. It is because of the importance of out-group comparison and because of its confusion with commonality (e.g., Crisci and Stuessy, 1980), that we attempt to formalize the criterion and synthesize the method.

PROBLEMS WITH COMMONALITY

From the following quotations, abstracted from a cursory review of some recent literature, it is apparent that the "commonality principle" for determining in-group character polarity receives wide support in both botanical and zoological circles. At least some authors (Crisci and Stuessy, 1980; Eldredge, 1979) consider this to be the most widely accepted and applied of all criteria.

> "The character state that consistently appears in many members of each of the taxa being studied is deemed most likely to be primitive for the group of taxa."—H. B. Boudreaux (1979:8).
>
> "Criterion of Frequency of Occurrence . . . . The objective is to determine how extensively a character state is distributed within the taxon under study . . ." —G. Ekis (1977:117).
>
> "A primitive state is more likely to be widespread within a group than is any one advanced state."—A. G. Kluge and J. S. Farris (1969:5).
>
> ". . . the relatively more primitive states are likely to be distributed more generally throughout the group under study . . ."—G. Estabrook (1972: 439).
>
> "The criterion of common is primitive is not only the oldest, but also the most widely used of all the criteria of primitiveness . . . the concept is a good one to employ in most groups."—J. V. Crisci and T. F. Stuessy (1980:119, 130).

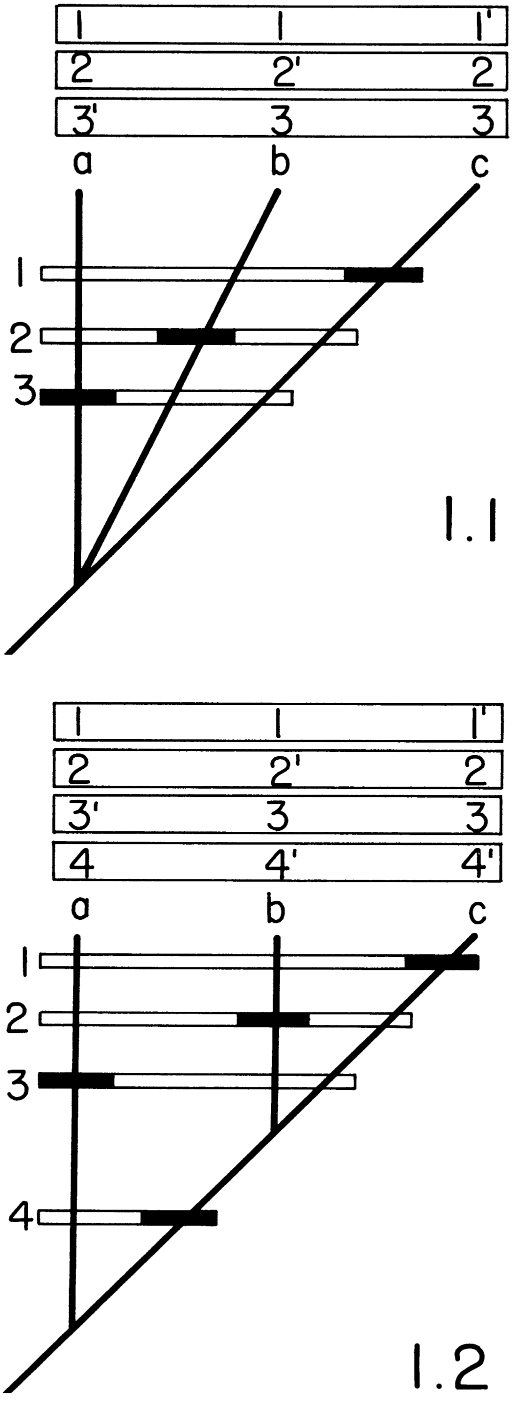

FIG. 1.—A three-taxon problem: fig. 1.1, where common equals primitive for all characters; fig. 1.2, where common equals derived for one character (4).

The first problem in dealing with the notion of "common equals primitive" is in recognizing when it is invoked. Occasionally it is unclear whether the author is using "common equals primitive" explicitly, out-group comparison in the strict sense, or is failing to differentiate the two. Crisci and Stuessy (1980), for example, divide their "Common is primitive" criterion into two categories: "in-group comparisons" and "out-group comparisons." They fail to distinguish between authors who have used the out-group method correctly (e.g., Ross, 1974; Lundberg, 1972) and those who have not. The following points about the "commonality principle" apply to instances where the frequency of occurrence of character states is used solely to infer polarity. This includes "in-group" criteria of most authors, and "out-group" criteria of some (e.g., Crisci and Stuessy, 1980).

1. In-group "common = primitive" will not resolve relationships within a 3-taxon group (Fig. 1.1). Because the construction of cladograms rests on the use of synapomorphies, a 3-taxon group can only be resolved if there is at least one state distribution wherein "common = derived," that is, if two of the three taxa share one or more derived (apomorphic) states (Fig. 1.2).

2. In-group "common = primitive" analysis of characters tends to produce "balanced" cladograms (i.e., branching sequences rooted near the middle) regardless of the real evolutionary relationships. For example, if Figures 2.1 and 2.2 represent the real evolutionary relationships of the taxa shown, in-group "common = primitive" results in cladograms in Figure 2.2 and 3.2. Further, in-group "common = primitive" alone will not fully resolve a cladogram, but will yield unresolved trichotomies or quadritomies (Figs. 2.2, 3.2, 4.2).

3. The incorrect hypotheses noted in point 2 will be continually supported by additional characters when polarities are determined by use of "common = primitive." Additions of such characters, pre-

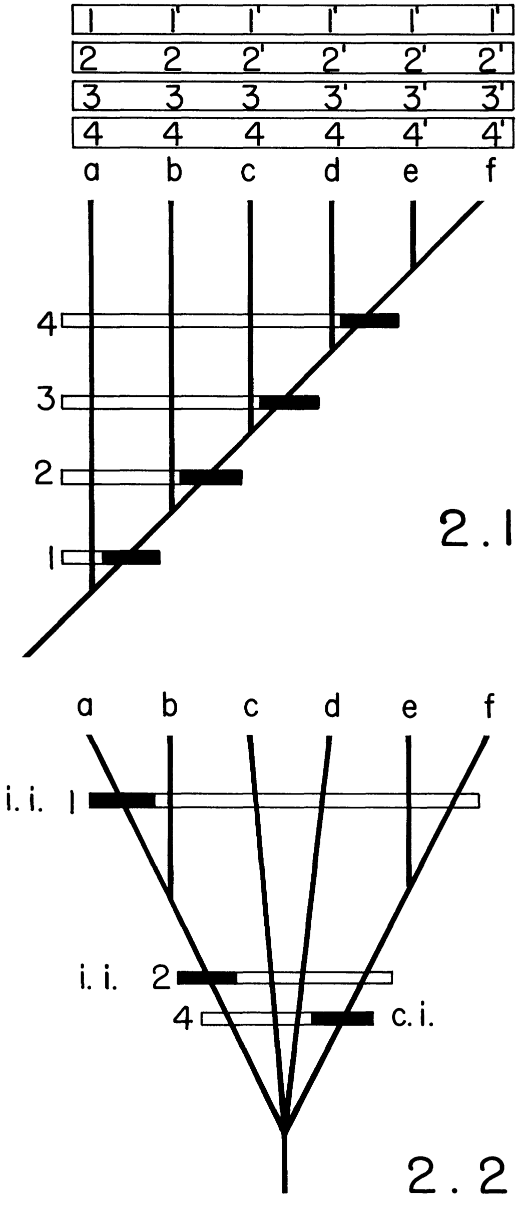

FIG. 2.—fig. 2.1, hypothetical cladogram showing actual (evolutionary) relationships; fig. 2.2, cladogram of same taxa as 2.1, based on common = primitive for all characters (characters 1 and 2 are incorrect inferences in this case).

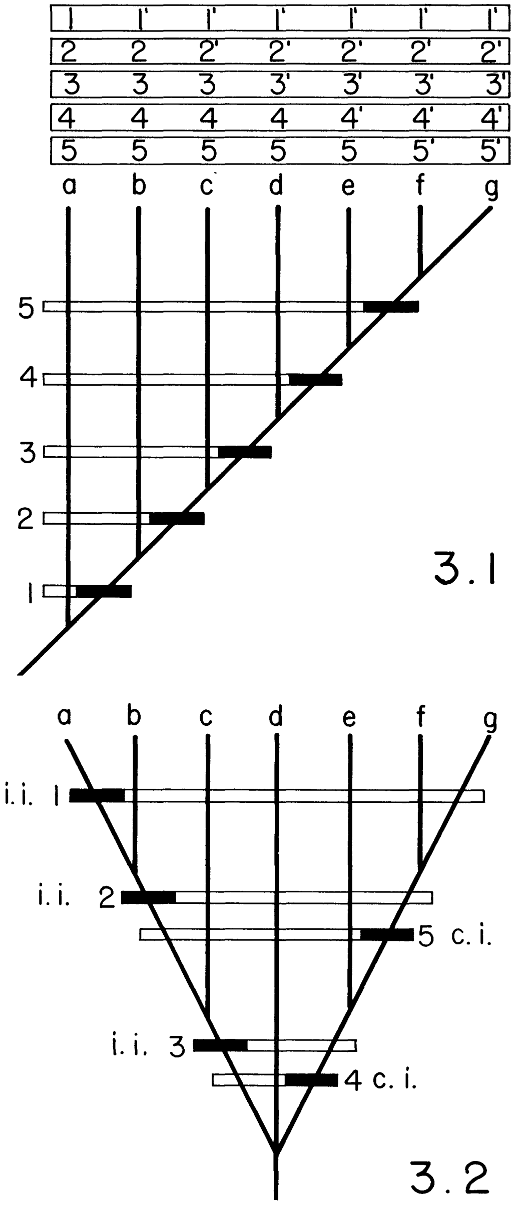

FIG. 3.—fig. 3.1, hypothetical cladogram showing actual (evolutionary) relationships; fig. 3.2, cladogram of same taxa as 3.1, based on common = primitive for all characters.

sumably to "test" earlier hypotheses, will not improve estimates of relationships even with the use of parsimony.

Estabrook (1977:37) raised the question "Does common equal primitive?", and responded, "all agree that the answer is not necessarily." We concur with this answer, and hasten to add that it is impossible without out-group comparison

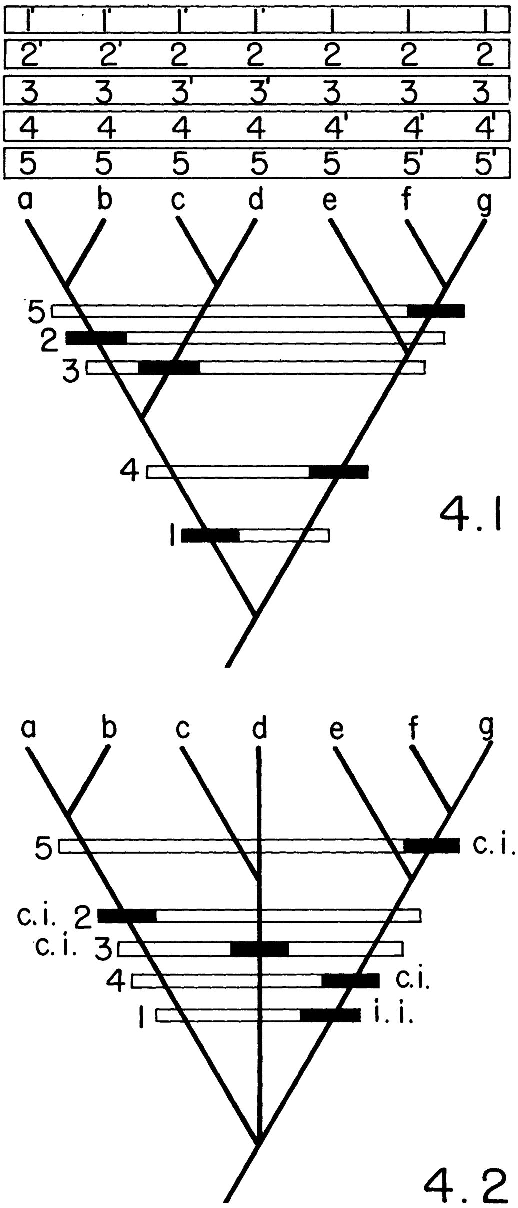

FIG. 4.—fig. 4.1, hypothetical cladogram showing actual (evolutionary) relationships; fig. 4.2, cladogram of same taxa as 4.1, based on common = primitive for all characters.

to differentiate those cases when common does equal primitive from those when it does not. We suggest that in-group commonality should not be used to infer polarities. As stated by Eldredge (1979:170), "the commonality principle . . . is actually an *a priori*, explicit statement . . . ," and in our view, one that may or may not reflect reality. As an alternative, polarity decisions might be based on out-group comparisons.

CHARACTERS AND CHARACTER STATES

Our use of the terms character and character state essentially follows that of Platnick (1979). Because of the importance of these concepts in our discussion of character analysis, and because of potential confusion, we restate and develop some of the ideas introduced by Platnick (1979).

A state, according to Platnick (1979:542) is equivalent to an attribute. An attribute is generally regarded as either an empirical observation (i.e., condition seen in an individual organism) or an hypothesis (i.e., an attribute of a group of organisms, such as a population, species, or any higher taxon). We interpret a state as an attribute of a species or supraspecific group, and as an hypothesis. This hypothesis is tested simply by observing attributes in additional individuals belonging to the group. An example (borrowed from Popper, 1968) would be observing several individual swans, and forming an hypothesis that "all swans are white." Discovery of a black one would falsify this hypothesis.

A character, on the other hand, is interpreted by Platnick as a "unit of sameness," that is, it consists of a group of states that are considered to be modifications or alternate forms of the same thing (i.e., states that are homologous). A character is also an hypothesis, and is tested by examining its congruence or incongruence with other characters (Platnick, 1979:544).

Formally defined, a character is an original form plus all of its subsequent modifications. The original form, similarly, is a modification of some pre-existing form. As an example, consider the following sequence of character states:

$$\begin{tikzcd} & & b \\ x \arrow[r] & a \arrow[ur] \arrow[dr] & & & \\ & & c \arrow[r] & d \arrow[r] & e \end{tikzcd}$$

character 1

In this transformation series, character 1 includes five states (*a–e*). The original form is state *a*, its subsequent modifications are states *b–e*, and state *a* itself is a modification of pre-existing form *x*. Within this complex of character states, in addition to character 1, there are several more restricted characters. This set of nested characters is as follows:

Character	States				
1	*a*	*b*	*c*	*d*	*e*
2		*b*			
3			*c*	*d*	*e*
4				*d*	*e*
5					*e*

On the other hand, some of the states or combinations of states in this example do not form characters (e.g., states *a, c, d, c+d,* and other combinations). Also consistent with this interpretation is that a character is identical with an apomorphy as defined by Hennig (1966). Those states or combinations of states that do not form characters are plesiomorphies or convergencies. So, the goal of character analysis is to distinguish those states that form sets of nested characters from those that do not. Put another way, character analysis attempts to determine the relative generality or restrictedness of characters (=apomorphies).

OUT-GROUP COMPARISONS

We have not attempted to discover the originator of the concept of out-group comparison and contemporary advocates are truly too numerous to cite. However, two lucid discussions of out-group comparison were given by Lundberg (1972: "related group"), and Ross (1974: "exgroup"). The operational rule below was modified from these sources, and a series of observations on the application of out-group comparison is provided to develop the concept beyond this rule.

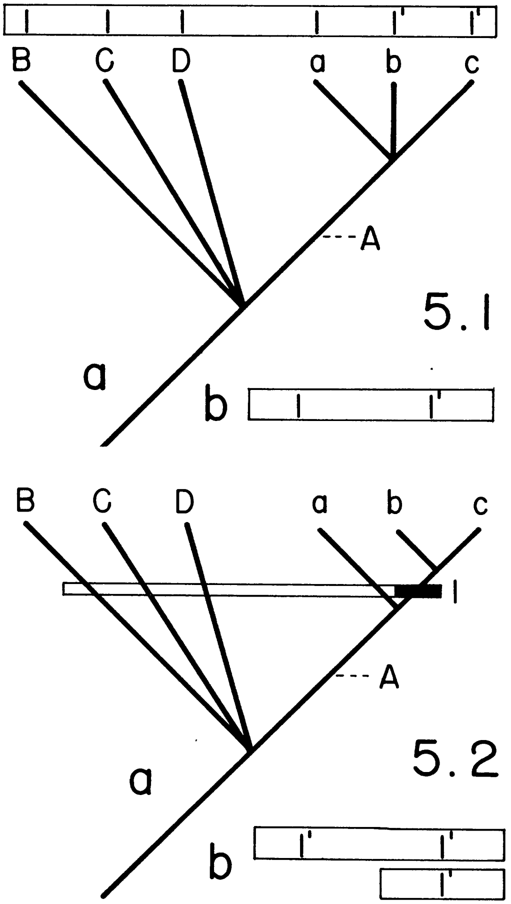

FIG. 5.—Simple example of out-group comparison; see discussion in text.

Operational rule.—For a given character with 2 or more states within a group, the state occurring in related groups is assumed to be the plesiomorphic state. If the character contains only 2 states, the alternative state is assumed to be apomorphic, thereby forming a more restricted character.

Figures 5.1–5.2 illustrate a simple example of out-group comparison. A hypothetical genus A contains species a, b, and c, and has related genera B, C, and D. Given a 2-state character (1, 1′) within genus A, the state found also in related groups (state 1 in genera B, C, and D) is plesiomorphic. Therefore, the alternative

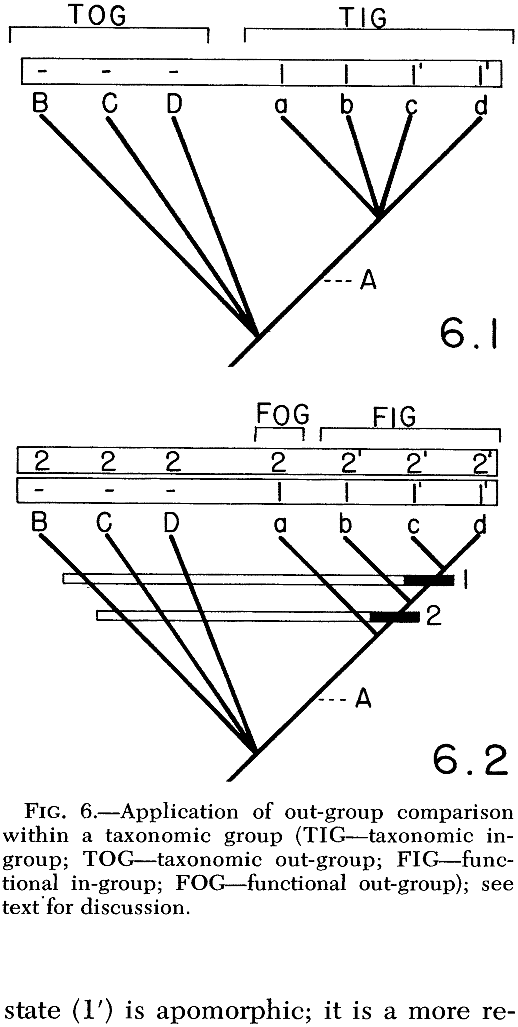

FIG. 6.—Application of out-group comparison within a taxonomic group (TIG—taxonomic in-group; TOG—taxonomic out-group; FIG—functional in-group; FOG—functional out-group); see text for discussion.

state (1′) is apomorphic; it is a more restricted character grouping b–c (Fig. 5.2).

Observations.—(1) *The number of times each character state occurs (i.e., its "frequency of occurrence") has a bearing on the polarity decision only in that the plesiomorphic state occurs in both the in-group and out-group (twice) and the apomorphic state only in the in-group (once).*

(2) *Out-group comparisons are not constrained by nomenclatural or taxonomic barriers, that is, they are applicable at all levels of a cladogram or to all monophyletic groups.*

Because any group may be used in out-group comparison, we refer to them as functional in-groups (FIG's) and functional out-groups (FOG's). Application of out-group comparison in the past has been limited by unnecessary coupling of the groups to formal classifications (i.e., with nomenclatural units, or what we call taxonomic in-groups [TIG's] and taxonomic out-groups [TOG's]). The concepts of FIG/FOG versus TIG/TOG will help to illustrate the applicability of out-group comparison to all levels (i.e., for all monophyletic groups).

(3) *All taxonomic groups (TIG/TOG) are also functional groups (FIG/FOG), but not all functional groups are taxonomic ones.*

This observation reenforces the idea that out-group comparisons need not be constrained by nomenclatural rank or Linnean hierarchical structure. Consider an example (Fig. 6.1) where 2 states of a character (1 and 1′) occur in a TIG (Genus A, with species a, b, c, and d), but neither state of which is present in related taxa (the TOG: genera B, C, and D). Given only this much information it is impossible to resolve the polarity of character states 1 and 1′ using the out-group criterion. If a second character is found, however, with 2 states (2 and 2′) in the TIG, one state of which (2) also occurs in the TOG, then a FIG (spp. b, c, and d) and a FOG (sp. a) can be formed within the initial TIG (Fig. 6.2). In the context of this smaller FIG/FOG system, character state 1 is plesiomorphic and 1′ is a synapomorphy grouping c–d.

This idea of reducing the problem to a smaller functional-group level has clearly been used by authors in the past, although usually expressed in other terms (e.g., as a form of in-group comparison, within a single TIG). For example, Ball (1975) stated that for characters confined to the taxonomic group that "one proceeds by determining for an unclassified character which of its states is associated with that member taxon of the complex under consideration judged to be 'most plesiotypic.'" In Ball's study of *Phloeoxena* carabid beetles some characters only occurred in the *picta*-group and therefore could not be analyzed with out-group comparisons. Using other, independent characters (body proportions), Ball determined the *nigricollis-limbicollis* lineage to be "most plesiotypic" (i.e., to be sister-group to remaining species of the *picta*-group). Thus, Ball used the *nigricollis-limbicollis* lineage as a FOG to analyze characters in the remaining *picta*-group species (the FIG).

At the same time, however, so-called "in-group criteria" have been misinterpreted and incorrectly applied by others. A common misuse involves the assumption that common = primitive. This is clearly to be avoided when functional groups can be used.

(4) *If both states of a two-state character occur in the TIG and also the TOG (and if relative polarities of the two states in the TOG are unknown), then the problem cannot be immediately resolved by out-group comparison.*

One of the two states in the TOG may be removed from consideration either by decreasing the size of the TOG or by indicating nonhomology between the in-group and out-group occurrences of either of the two states.

(5) *Characters with more than 2 states can be completely resolved by out-group comparison only if the most plesiomorphic state in each more-restricted character occurs in 2 groups, one more inclusive than the other.*

The cladogram in Figure 7.1 shows a 4-state character (character 1 with states, 1′, 1″, 1‴) distributed among 7 taxa, a–g. Within character 1 are three less inclusive characters (Fig. 7.2, characters 2–4). The most plesimorphic state in character 2, for example, is state 1′, which occurs in two groups, d–g and c–g; the latter, of course, is more inclusive. The plesiomorphic state in each of the other characters also occurs in two groups, one more inclusive than the other. Put another way, out-group comparisons cannot be applied to resolve a complex character (i.e., one with more than 2 states) unless FIG's and FOG's can be established (using independent characters) and one, and only one, of the states found in the FIG also occurs in the FOG.

Unfortunately, many characters with more than two states do not conform to the general pattern outlined above. For such characters there may be a complex array of equally plausible character state sequences. Consider the cladogram in Figure 7.3 (which is the same as Fig. 7.1, with taxa c and e missing). Assuming no homoplasies, the possible sequences for states 1, 1′, 1″, and 1‴ are shown in Figure 7.4. All of these sequences are theoretically equally probable, though morphological evidence (e.g., morphoclines based on Dollo's principle) might indicate that some are unlikely.

(6) *Polarity decisions are hypotheses, subject to testing and falsification.*

Determination of character state polarity amounts to forming an hypothesis, which is subject to testing and potential falsification. A more thorough examination of the out-group taxa is one such test. Perhaps initial hypotheses of polarity were based on a small sample of the out-group taxa. Study of other taxa might reveal that both states within a FIG also occur in the FOG, thus the out-group rule no longer is applicable (see observation 4).

(7) *Original groupings do not have to be monophyletic.*

The FIG + FOG does not have to be monophyletic (*sensu* Hennig, 1966). This observation is directed to an apparent problem implicit in the question "Where do out-group comparisons begin?". Monophyletic groups are used to decide polarities of character states, and apomorphies (characters) are in turn used to define monophyletic groups. However, we agree with Eldgredge (1979:171, footnote) that some progress has been made during the last few centuries, and a practical starting point is to follow the existing classification. Given a preliminary TIG/TOG system, some polarities are likely to be determinable. These will

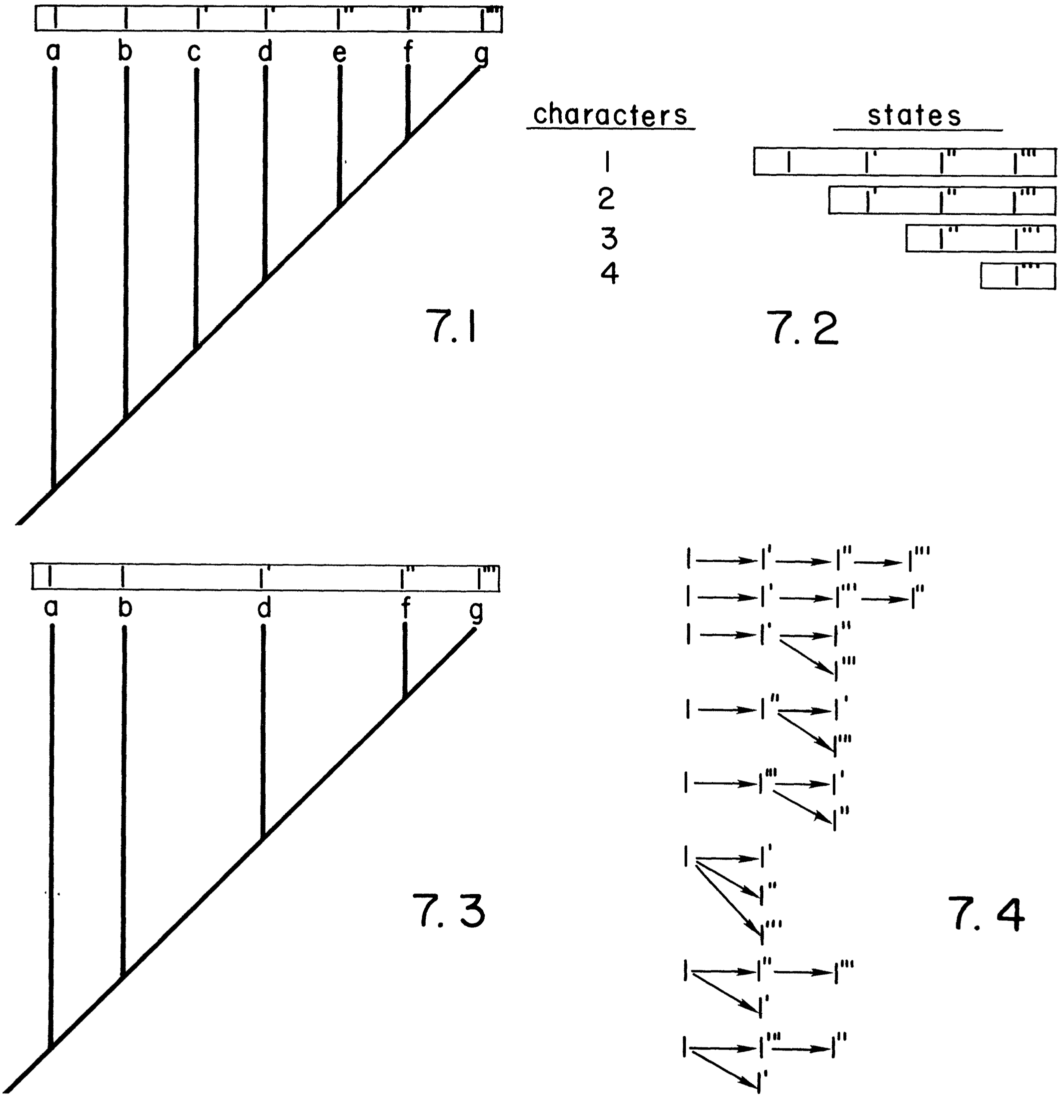

FIG. 7.—A 4-state character: fig. 7.1, distribution of states among seven taxa, a–g; fig. 7.2, nested characters within character 1; fig. 7.3, cladogram in fig. 7.1, with taxa c and e missing; fig. 7.4, possible sequences for 4 states, based on cladistic relationships in fig. 7.3, and assuming no homoplasies occur.

then support more refined FIG/FOG arrangements to provide further resolution of polarities. When mistakes occur they can be detected (see observation 8).

(8) *Misinterpretations of character state polarities based on out-group comparisons (due to reversals or independent occurrences) or, incorrect initial TIG/TOG hypotheses (see observation 7), will be revealed through conflicts with other characters.*

The original set of characters analyzed will form the basis for an hypothesis about relationships among a set of taxa, expressed as a cladogram. Study of additional characters (polarized by out-group comparisons) will test these relationship hypotheses, and narrow the gap between our best estimates of relationships and reality (i.e., we are specifying the single, historical pattern of nature more precisely). The same is not true for the "commonality principle" where additional characters continually support incorrect hypotheses (Figs. 3.1–3.2, 4.1–4.2, and point 3 under the discussion of commonality above).

(9) *A technique of cladistic analysis which requires ranking all character states as apomorphic or plesiomorphic prior to formation of groups often cannot be based entirely on out-group comparisons.*

Because the polarity of some character states cannot be determined until FIG/FOG refinements have been established within a particular taxon, *a priori* ranking of all character states as plesiomorphic or apomorphic may not be possible based on out-group comparisons alone. This observation implies that a cladistic analysis is often a step-by-step process, with some levels of resolution being dependent upon preceding polarity decisions and FIG/FOG groupings (Figs. 6.1– 6.2 again illustrate this concept).

(10) *Neontological and paleontological data have equal applicability in out-group comparisons.*

Fossils can reveal something about the minimum age of a taxon (or character state), and perhaps something about the paleoecological surroundings of the organism and/or its past geographic distribution. However, they contain no special historical information about phylogeny. A framework of monophyletic groups is hypothesized from characters, and cladistics does not distinguish between characters based on fossils or those based on extant organisms. In fact, extant organisms allow us to observe features that are invisible in or absent from fossil specimens (e.g., fine details, internal organs, cytological features). Extant organisms are also much easier to evaluate in terms of variation, simply because more material is generally available. Cracraft (1979) provides a lucid discussion of fossils and cladistics.

SYNTHESIS OF THE OUT-GROUP COMPARISON METHOD

The out-group comparison criterion was formalized in our "operational rule," and several important implications and potential problems of this criterion were discussed in our observations above. While these observations do confront the important issues crucial to out-group comparison, they do not alone constitute a clearly formulated method of character analysis. Our goal beyond a definition and discussion of out-group comparisons, is to interpret these observations and synthesize a general method of character analysis consistent with the out-group criterion.

The simplest character analysis based on out-group comparison proceeds as follows. Two states are observed in a TIG, only one of which also occurs in the TOG. The restricted state (occuring only in the TIG) is inferred to be apomorphic; the general state (occuring in the TIG + TOG) is inferred to be plesiomorphic. This polarized character is then compared with others for congruence, and with no conflicts is accepted as is. This scheme of our out-group comparison method, and the major types of complications (discussed below) are summarized in Figure 8.

If neither state occurs in the TOG, we have to use a less inclusive FIG/FOG system, based on independent characters as discussed under observation 3. If both states occur in the TOG, we would either demonstrate that the TIG and TOG states are not homologous or narrow the TOG such that only one of the states is present (see observation 4).

If the character is incongruent with other characters, we have three alternatives. We can select the most parsimonious arrangement, and assume that the majority of congruent characters are correctly polarized. This implies a problem with homologies for the conflicting character states (i.e., homoplasies). We can reinterpret the homologies of characters from a morphological (or other) point of

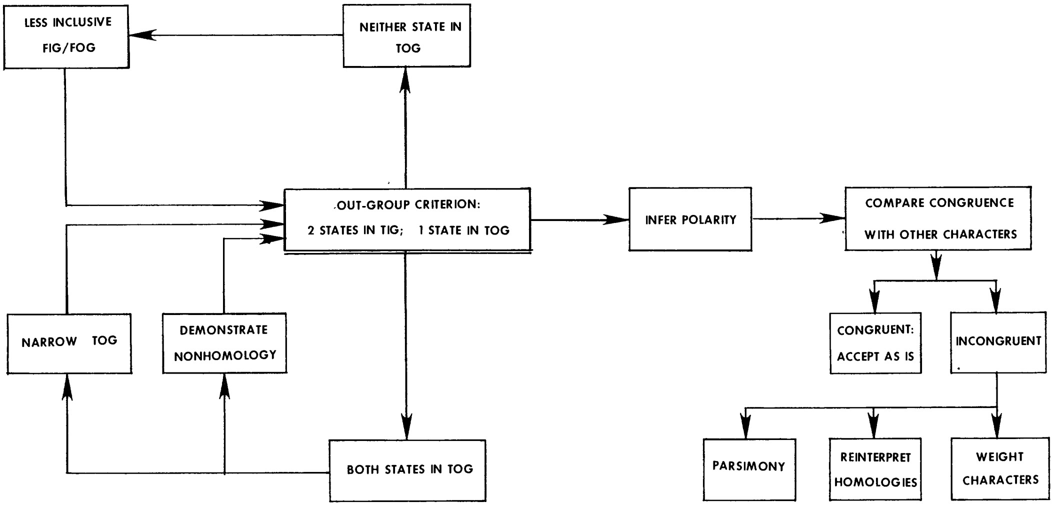

FIG. 8.—Flow chart showing method of character analysis based on out-group comparison (TIG—taxonomic in-group; TOG—taxonomic out-group; FIG—functional in-group; FOG—functional out-group); see text for discussion.

view. Or, we can weight characters to resolve conflicts.

DISCUSSION

Misunderstandings in the correct application of the out-group criterion in resolving character polarity questions indicated the need for a more formal presentation of this criterion, which we have done in our operational rule. A series of observations are made on this "rule," which reenforce its definition, and explain some important implications and problems with application of the criterion. All of these combine to add stability to the meaning of this longstanding criterion, and to clarify its proper use.

The topic of character polarity cannot be effectively dealt with in the absence of some reference to the commonality principle, which has and continues to receive support in systematics. Although there certainly are instances wherein common does equal primitive, we have explained that it is not possible to sort these out *a priori*, and that false hypotheses of relationship based on this explicit assumption can be incorrectly, and undetectably, supported by additional (and also false) characters. Based on these, and other considerations we conclude that "common = primitive" is not a reliable indicator of polarity.

To elucidate the full implications of the out-group operational rule and our observations to analysis of characters, we synthesize a general method of character analysis which rests on out-group comparisons. We believe this to be the most reliable method for determination of polarities based on comparative studies. While the basis for our method is not new (and, in fact, is among the oldest and most wide-spread criteria), persistent misunderstanding in the correct application of the criterion emphasizes the need for standardization. Our goal has been to outline a method of character analysis which, when rigorously applied, can avoid the errors of the past, and add stability to one of the fundamental parts of cladistics.

ACKNOWLEDGMENTS

Preparation of the final form of this paper was made possible by a Presidential Fellowship of The Ohio State University, held by each of us. This paper originated in a discussion group on cladistics at the North-Central Branch meetings of the Entomological Society of America in Indianapolis in 1979, which we chaired jointly. Among the persons who have discussed these ideas with us we especially thank George Ball, Dan Brooks, Jim Carpenter, Ginter Ekis, Vicki Funk, Don Johnston, Norman Platnick, Tod Stuessy, and the participants in our systematic discussion group at Ohio State University.

REFERENCES

Ball, G. E. 1975. Pericaline Lebiini: notes on classification, a synopsis of the New World genera, and a revision of the genus *Phloeoxena* Chaudoir (Coleoptera: Carabidae). Quaest. Entomol., 11:143–242.

Boudreaux, H. B. 1979. Arthropod phylogeny with special reference to insects. J. Wiley and Sons, New York.

Cracraft, J. 1979. Phylogenetic analysis, evolutionary models, and paleontology. Pp. 7–39 *in* Phylogenetic analysis and paleontology (J. Cracraft and N. Eldredge, eds.). Columbia University Press, New York, 233 pp.

Crisci, J. V., and T. F. Stuessy. 1980. Determining primitive character states for phylogenetic reconstruction. Systematic Botany, 5:112–135.

Ekis, G. 1977. Classification, phylogeny, and zoogeography of the genus *Perilypus* (Coleoptera: Cleridae). Smithson. Contrib. Zool., no. 227.

Eldredge, N. 1979. Cladism and common sense. Pp. 165–198 *in* Phylogenetic analysis and paleontology (J. Cracraft and N. Eldredge, eds.). Columbia University Press, New York, 233 pp.

Estabrook, G. F. 1972. Cladistic methodology: a discussion of the theoretical basis for the induction of evolutionary history. Ann. Rev. Ecol. Syst., 3:427–456.

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

Hennig, W. 1965. Phylogenetic systematics. Ann. Rev. Entomol., 10:97–116.

Hennig, W. 1966. Phylogenetic systematics. University of Illinois Press, Urbana (Reprinted, 1979).

Kluge, A. G., and J. S. Farris. 1969. Quantitative phyletics and the evolution of Anurans. Syst. Zool., 18:1–32.

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

Marx, H., and G. B. Rabb. 1970. Character analysis: an empirical approach applied to advanced snakes. J. Zool. (London), 161:525–548.

Marx, H., and G. B. Rabb. 1972. Phyletic analysis of fifty characters of advanced snakes. Fieldiana, Zool., 63:1–321.

Maslin, P. P. 1952. Morphological criteria of phylogenetic relationships. Syst. Zool., 1:49–70.

Munroe, D. D. 1974. The systematics, phylogeny, and zoogeography of *Symmerus* Walker and *Australosymmerus* Freeman (Diptera: Mycetophilidae: Ditomyiinae). Mem. Entomol. Soc. Canada, 93:1–183.

Platnick, N. I. 1979. Philosophy and the transformation of cladistics. Syst. Zool., 28:537–546.

Popper, K. R. 1968. The logic of scientific discovery. Harper and Row, New York.

Ross, H. H. 1974. Biological systematics. Addison-Wesley, Reading, Massachusetts.

Manuscript received June 1980
Revised October 1980