The Logic and Limitations of the Outgroup Substitution Approach to Cladistic Analysis

Michael J. Donoghue; Philip D. Cantino

The Logic and Limitations of the Outgroup Substitution Approach to Cladistic Analysis

MICHAEL J. DONOGHUE

Department of Botany, San Diego State University,
San Diego, California 92182

PHILIP D. CANTINO

Department of Botany, Ohio University, Athens, Ohio 45701

ABSTRACT. The outgroup substitution approach can be used when a well-corroborated hypothesis of the more inclusive cladistic relationships of a study group is unavailable. It is particularly appropriate when the set of outgroups, each of which could plausibly be the sister group, includes some that may be only distantly related to others. All plausible sister groups are used as outgroups, alone and in various arrangements, to assess character polarities. In each case an ingroup cladogram is constructed and these are searched for areas of congruence. Through this approach, robust phylogenetic hypotheses can be generated in spite of uncertain outgroup relationships. Caution must be exercised because polarity assessment is sensitive to the exact arrangement of outgroups. Even if completely congruent cladograms are obtained, it may not be possible to specify character support for particular clades. The approach may be most useful when one's primary objective is to establish an outgroup hypothesis for a subsequent cladistic analysis at a lower taxonomic level.

Outgroup comparison is now widely regarded as one of the most reliable, if not the only logically justified, method of assessing character polarity for cladistic analysis (e.g., Arnold 1981; Eldredge and Cracraft 1980; Farris 1982; Nelson and Platnick 1981; Stevens 1980; Watrous and Wheeler 1981; Wheeler 1981; Wiley 1981). However, the outgroup method is often difficult to use in studies of angiosperm groups because there are few well-corroborated hypotheses of cladistic relationships among families and orders, and there is frequently lack of agreement among phylogenists about such relationships. If a suite of related groups is used collectively as a single outgroup without resolution of their relationships to each other (e.g., Baum 1983; Bolick 1981; Kress and Stone 1983; Seaman and Funk 1983), it is necessary either to apply the "commonality principle" (Eldredge 1979) within the outgroup or to restrict the analysis to only those characters that are uniform throughout this composite outgroup but vary in the ingroup. Barabé (1982) suggested that, given the extensive parallelism in the angiosperms, only a rather restricted outgroup is invariant in enough characters to be of much use in constructing an ingroup cladogram. Because our present ignorance of angiosperm phylogeny precludes confident restriction of the outgroup in most cases, Barabé (1982) re-

jected outgroup comparison in favor of other criteria such as ontogeny, fossil evidence, and character correlation. Likewise, Guédès (1983, p. 277) stated: "I doubt that the outgroup method can be used to establish phylogeny because it requires knowledge of phylogeny."

We maintain that even in the absence of a "known" (i.e., well-corroborated) sister group, it is possible to employ outgroup comparison in a rigorous and explicit manner. If the problem is simply one of non-resolution of relationships within a set of closely related outgroups, it may be best to attempt simultaneous resolution of the outgroups plus the ingroup (the "global parsimony" approach; Maddison et al., in press). However, this method is most appropriate when one can assume that the plausible outgroups plus the ingroup form a monophyletic group. In plants, at least, this assumption is often likely to be incorrect. Angiosperm phylogenists sometimes express radically different and logically incompatible opinions about the affinities of a group. One author will consider taxon "A" to be closely related to taxon "B," another will consider "A" to be closely related to "C," and yet both will agree that "B" and "C" could not possibly be closely related to one another. The resolution of such a controversy would, in some cases, require a cladistic analysis of a large segment of the angio-

sperms. This is not usually possible as a prelude to the phylogenetic analysis of a family or genus. The global parsimony approach also assumes that extensive data are available on all of the outgroups considered, even about characters that are invariant in the ingroup. Most specialists on an ingroup will not have such a broad data base for the outgroups.

One approach to the problem of uncertain, and possibly only distantly related, outgroups was suggested by Wiley (1980, p. 200) and was more explicitly formulated by Donoghue (Coombs et al. 1981, p. 364; Donoghue 1983b). This method, which we call the outgroup substitution approach, has since been used by both of us (Cantino 1982; Donoghue 1983b), but not in a manner as rigorous as we now realize is necessary. Our purpose here is to clarify the logic of this procedure and to explore its limitations.

The idea behind the outgroup substitution approach is straightforward: If relationships of outgroups to the study group are uncertain and it is not feasible to perform a higher level cladistic analysis to resolve them, one can employ as outgroups all plausible sister groups, singly and in all plausible arrangements, construct a cladogram for each using a parsimony method, and search for areas of congruence in the resulting cladograms. In this context, a "plausible" sister group or outgroup arrangement is any for which there is evidence based on previous cladistic analysis or that has been suggested by phylogenists. There are as many cladograms obtained by this method as there are plausible outgroup arrangements. Congruent portions of the resulting cladograms represent robust hypotheses of cladistic relationship within the study group, because they remain the same even when assumptions about the broader cladistic relationships of the group are changed. Acceptance of these robust hypotheses assumes, of course, that one of the plausible sister groups is the real sister group, i.e., that previous generations of systematists have narrowed down the possible close relatives of the study group to a set that includes the real sister group (see Eldredge 1979, p. 171, and Donoghue 1983b for discussion of this assumption). This work of previous systematists is what Donoghue (1983b) referred to as the "phenetic groping" phase, a point that was unfortunately misinterpreted by Platnick and

Funk (1983, p. 199). The examination of a variety of possible outgroups is certainly not "phenetic groping." Indeed, it is not phenetic and can be carried out in a rigorous and explicit fashion.

The outgroup substitution approach is a practical way to draw tentative conclusions about cladistic relationships when outgroups are uncertain. It need not be used if a well-corroborated hypothesis of more inclusive relationships is already available or if it is feasible to try to establish such a hypothesis before beginning an analysis of the study group. However, if a hypothesis is available but is not very well supported, it may be appropriate to consider how changes among the outgroups affect cladistic analysis of the study group.

Although the outgroup substitution approach is, in our experience, useful, it should be employed with consideration of: 1) the influence of secondary outgroups (i.e., those beyond the immediate sister group) on polarity assessments, 2) the assessment of congruence and treatment of incongruence among resulting cladograms, and 3) the nature of character support for cladogram structure when the outgroup substitution approach is used.

INFLUENCE OF SECONDARY OUTGROUPS ON POLARITY ASSESSMENT

The following observations are based on an analysis of the logic of outgroup comparison carried out by Maddison et al. (in press; see also Donoghue and Maddison 1982). They considered the case of a monophyletic study group (the "ingroup") within which cladistic relationships are unresolved. In figure 1 we show such a group with variation in a binary character with states "a" and "b." This situation will henceforth be represented by a triangle as in figure 2. The problem is to determine which state is plesiomorphic (ancestral) and which apomorphic (derived). One or more outgroups are considered in order to arrive at an assessment of the most parsimonious state assignment to the "outgroup node" (ON in fig. 2). This state is considered the plesiomorphic condition for the ingroup. In the following discussion, we will consider hypothetical examples in which the cladistic relationship of more than one outgroup to the ingroup is "known" (figs. 2–8). We will refer to the sister group as
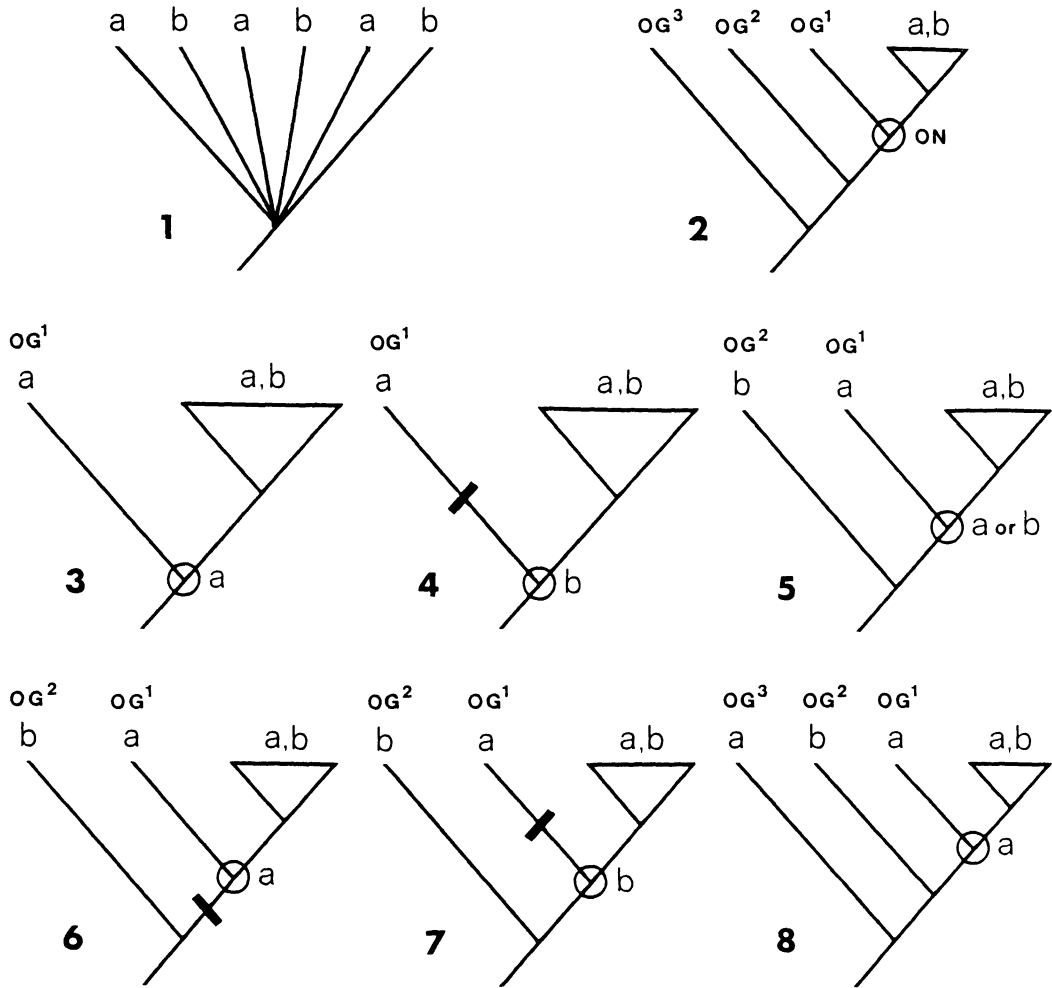
FIGS. 1–8. Outgroup analysis. 1. Unresolved study group (ingroup) of six taxa with variation in a binary character. 2. Unresolved ingroup (shown as triangle) and three outgroups (OG); ON = outgroup node. 3–4. One outgroup; "a" most parsimonious at ON (shown as circle); a character state change (shown as bar) is required in the outgroup portion of the cladogram if "b" is assigned to ON. 5–7. Two outgroups; polarity assessment equivocal; "a" or "b" at ON both require one character state change (shown as bar). 8. Three outgroups; "a" most parsimonious at ON.

"outgroup 1" and will number the secondary outgroups consecutively outward.

Maddison et al. (in press) demonstrated that outgroup 1 is the most important determiner of the outcome of outgroup polarity assessment. Either the most parsimonious assignment to ON will be the state possessed by outgroup 1 or it will not be possible to decide on the basis of parsimony which state should be assigned to ON (an "equivocal assessment"). Whether it is possible to reach a decision or not depends on whether more distant outgroups are "known" and, if so, on the distribution of states among these groups. If only outgroup 1 is "known" and it has state "a," then the assignment to ON will be "a" (fig. 3). Like any outgroup assessment, this is a parsimony decision, i.e., the hypothesis assigning state "a" to ON requires one fewer character state change in the outgroup portion of the cladogram than does the alternative hypothesis (compare figs. 3 and 4). In either case, one or

more character state changes are required within the ingroup, but the number and position of these changes does not affect the most parsimonious state assignment at ON supplied by the outgroups. Maddison et al. (in press) have demonstrated that the state assignments to the outgroup node that are provided by the outgroups will allow one to find a globally most parsimonious ingroup cladogram, i.e., an ingroup cladogram that is most parsimonious in the context of the larger cladogram that includes the outgroups.

If a second outgroup is "known" and has state "b" (fig. 5), the assignment to ON is equivocal, because regardless of which state, "a" or "b," is assigned to ON, a minimum of one character state change is required in the outgroup portion of the cladogram (figs. 6–7). If a third outgroup is "known" and has state "a," it is again possible to assess polarity (fig. 8). At least one character state change is required in the outgroup portion of the cladogram if state "a" is assigned to ON, whereas at least two are required with "b" assigned to ON.

Whether a particular character state assignment at ON is decisive or equivocal may affect the outcome of cladistic analysis in the ingroup, because, at least initially, an equivocal character cannot be used to infer ingroup cladogram structure (but see p. 200). Because character state assignment at ON is influenced by secondary outgroups, it is important to consider how these groups might affect the analysis. We will illustrate this with examples, mostly drawn from our own applications of the outgroup substitution approach.

Cantino (1982) analyzed cladistic relationships in an ingroup consisting of three terminal taxa (Lamiales, Scrophulariales, and Boraginaceae), in relation to five plausible sister groups (Gentianales, Polemoniales, Loasales, Campanulales, and the "Cornalean core-complex"). He concluded that more apomorphies link Lamiales with Scrophulariales than with Boraginaceae, regardless of which outgroup was used as outgroup 1. In this analysis each outgroup was used singly. As demonstrated below, if secondary outgroups had been considered, several polarity decisions based on a single outgroup would have changed from decisive to equivocal or vice versa.

The leaves are opposite in nearly all members of Lamiales and in most of the large families of Scrophulariales; they are alternate in Boraginaceae. Among the outgroups, opposite leaves predominate in Gentianales, and that arrangement is considered ancestral in the order, while alternate leaves are considered to be ancestral in Campanulales. [The reader will note that the commonality principle was applied within the outgroup here. This is not an essential, or even a recommended, step in the outgroup substitution approach but happened to have been used in this example. The problem of character variation within outgroups is discussed on p. 198 and by Cantino (1982).] Both character states are common in the other three outgroups, and at present it is not possible to decide which state is ancestral within these groups. Using each outgroup singly, Cantino concluded that opposite leaves is an apomorphy linking Lamiales and Scrophulariales if Campanulales is the outgroup, but not if any of the other four outgroups is used (fig. 9).

However, if Campanulales is outgroup 1 and Gentianales is outgroup 2, the polarity assessment becomes equivocal (fig. 10) by the argument illustrated in figures 5–7. If any of the other plausible sister groups is used as outgroup 2 (provided Gentianales is not also a secondary outgroup), it is most parsimonious to assign alternate leaves to ON (compare figs. 11 and 12). However, if Gentianales is admitted as one of the secondary outgroups, then no matter where it occurs in relation to Campanulales and the ingroup it will render the assessment equivocal (fig. 13).

Floral symmetry provides an example of the opposite possibility: Consideration of secondary outgroups permits an unequivocal assessment whereas one outgroup used singly does not. If Campanulales is used as the sole outgroup, no polarity decision is possible because both actinomorphy and zygomorphy are common in Campanulales (fig. 14). If, however, Gentianales is outgroup 2, then actinomorphy is the most parsimonious assignment to ON (Fig. 15). This is also true if any of the other plausible sister groups is used as outgroup 2, because in each of them actinomorphy is either the universal or heavily predominant state and hence was considered by Cantino (1982) to be the ancestral condition.

Discovery of the limitations of using only single outgroups has necessitated a re-examination of the affinities of Lamiales. In the new
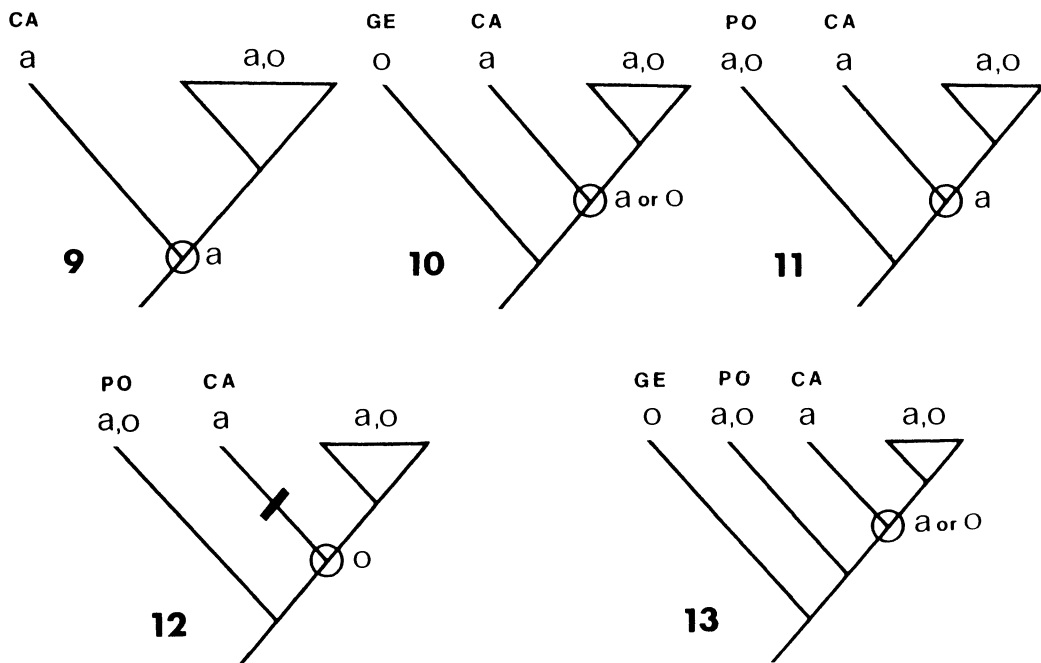
FIGS. 9–13. Polarity assessments for leaf arrangement in an ingroup (triangle) comprising Lamiales, Scrophulariales, and Boraginaceae, with Campanulales (CA) as OG[1]. 9. Alternate leaves (a) more parsimonious than opposite leaves (o) at ON (circle). 10. Gentianales (GE) as OG[2]; polarity assessment equivocal. 11–12. Polemoniales (PO) as OG[2]; "a" more parsimonious than "o" at ON. 13. PO as OG[2] and GE as OG[3]; polarity assessment equivocal.

analysis, all possible combinations of the five plausible sister groups were considered. Regardless of the outgroup arrangement, it is most parsimonious to hypothesize that Scrophulariales, rather than Boraginaceae, is the sister group of Lamiales. Thus the original conclusion of Cantino (1982) still holds.

Donoghue (1983b) investigated the phylogenetic relationships of *Viburnum*. In one of his analyses he considered an ingroup (Caprifoliaceae s.l.) consisting of three terminal taxa: 1) the ten genera of Caprifoliaceae s.str., 2) *Sambucus* and *Adoxa*, and 3) *Viburnum*. Using Valerianaceae or Dipsacaceae or both (i.e., the re-
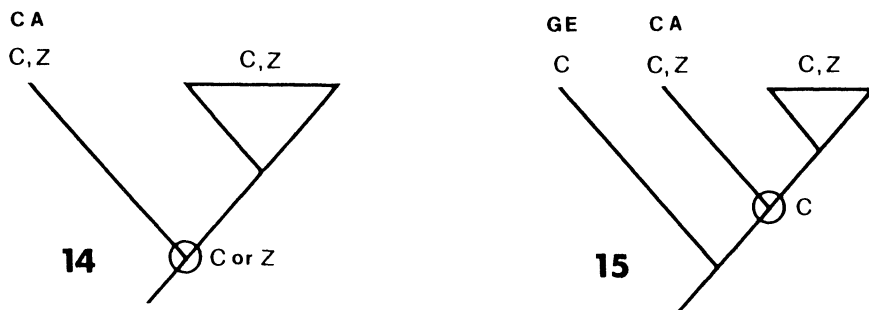
FIGS. 14–15. Polarity assessments for floral symmetry with ingroup as in figures 9–13 and Campanulales (CA) as OG[1]. 14. Actinomorphic flowers (c) and zygomorphic flowers (z) equally parsimonious at ON (circle). 15. Gentianales (GE) as OG[2]; "c" more parsimonious than "z" at ON.
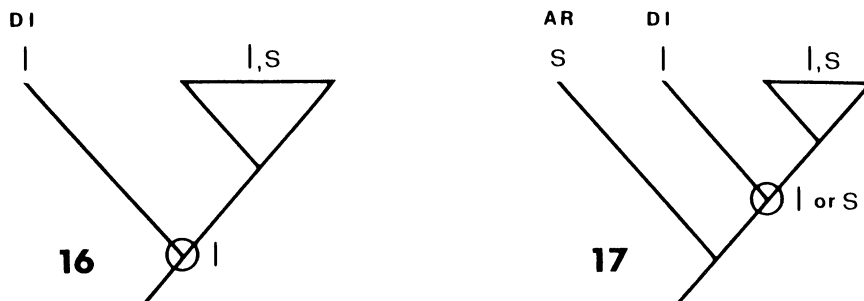
FIGS. 16–17. Polarity assessments for style length in an ingroup (triangle) comprising Caprifoliaceae s.l. with Dipsacales (DI) as OG[1]. 16. Long style (l) more parsimonious than short style (s) at ON (circle). 17. Araliales (AR) as OG[2]; polarity assessment equivocal.

maining Dipsacales) as outgroup 1, he concluded that more apomorphies link *Viburnum* with *Sambucus* and *Adoxa* than with Caprifoliaceae s.str. If he had employed other plausible sister groups as secondary outgroups, the polarity assessment for some characters would have changed. For example, using Valerianaceae and Dipsacaceae together as outgroup 1, short style length was scored as a synapomorphy of *Viburnum*, *Sambucus*, and *Adoxa* (fig. 16). If Rubiaceae were outgroup 2, this assessment would not be altered, because many Rubiaceae have long styles. If Araliales (Araliaceae and Apiaceae) were outgroup 2, however, a polarity decision for this character would not have been possible, because long styles and short styles would be equally parsimonious at ON (fig. 17). Hence, under these circumstances, style length could not have been used to unite *Viburnum* with *Sambucus* and *Adoxa*. However, it is most parsimonious to conclude that *Viburnum*, *Sambucus*, and *Adoxa* form a monophyletic group regardless of which combination of outgroups is used.

In Bolick's (1981) cladistic analysis of *Salmea* (including *Chrysosalmea*, pers. comm.), polarities were assessed by comparison with three other genera (*Acmella*, *Otopappus*, and *Spilanthes*) whose cladistic relationships to the ingroup are unresolved. For most characters an unambiguous polarity assessment was possible because all three outgroups shared the same state. A few characters, however, varied within or among the outgroups. For character 9 in her table II ("apex of pale flaring"), *Acmella* and *Otopappus* lack the condition and *Spilanthes* has

the condition. In spite of this variation, presence was scored as apomorphic. If, however, *Spilanthes* were outgroup 1 (and her higher level analysis did not rule out this possibility), then the assessment would be equivocal (fig. 18) and presence of a flaring apex on the pale could not be treated as apomorphic within *Salmea* on the basis of this outgroup comparison. Bolick's character 23 ("hairs on upper leaf surface") illustrates the opposite possibility. Although she concluded that it was not possible to determine polarity owing to variation in *Acmella* and *Spilanthes*, a decision is possible; it is most parsimonious to conclude that the presence of hairs is plesiomorphic in *Salmea* (figs. 19–20). This is true whether *Otopappus* is outgroup 1, 2, or 3, but only so long as relationships among the species within *Acmella* and *Spilanthes* remain unresolved.

These examples demonstrate that consideration of alternative outgroup arrangements can result in changes in polarity assessment from decisive to equivocal or vice versa. In some cases this will alter conclusions drawn about cladistic relationships in the study group. Awareness of this problem allows one to specify more precisely the conditions under which particular cladistic conclusions will hold.

Confidence is greatest in those polarity assessments for which all plausible sister groups are invariant, because in such cases any arrangement of outgroups will yield the same result. In other cases, polarity assessments will depend upon particular character state distributions among the outgroups being used and their assumed relationships to one another and
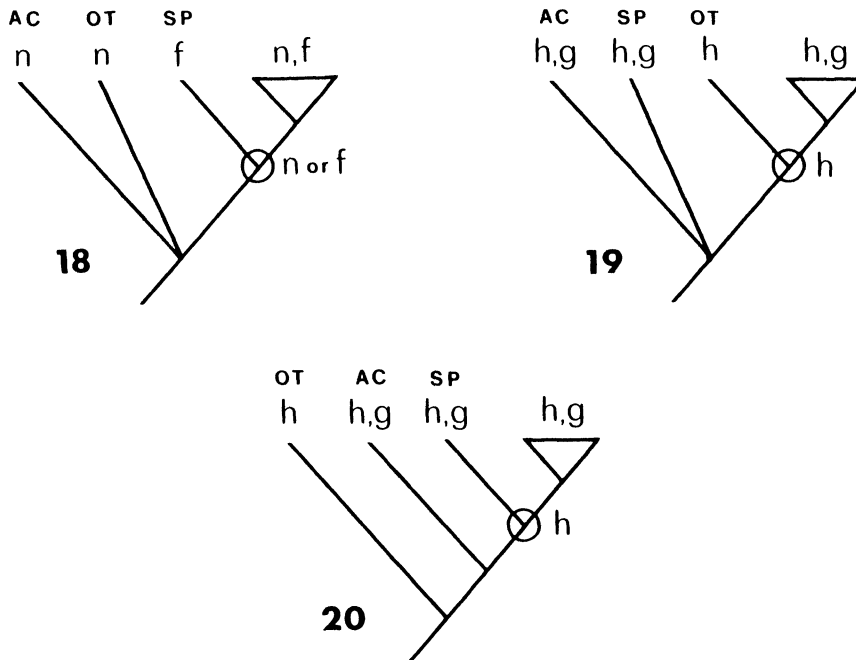
FIGS. 18–20. Polarity assessments for an ingroup (triangle) comprising *Salmea*. 18. *Spilanthes* (SP) as OG[1] and *Otopappus* (OT) and *Acmella* (AC) as secondary outgroups; f = apex of pale flaring, n = pale not flaring; polarity assessment equivocal. 19. OT as OG[1] and SP and AC as secondary outgroups; h = pubescent on upper leaf surface, g = glabrous; "h" most parsimonious at ON (circle). 20. SP as OG[1] and AC and OT as secondary outgroups; "h" most parsimonious at ON.

to the ingroup. If outgroup 1 has one state and all secondary outgroups have the opposite state, then the polarity assessment is equivocal. If outgroup 1 is variable but all secondary outgroups are invariant for a single state, then a polarity decision is possible and the character can be used (at least until relationships are resolved within outgroup 1). In other cases, some arrangements of outgroups will yield different polarity decisions than others (Maddison et al. in press, provide a general algorithm for outgroup polarity assessment).

Characters that vary *within* all plausible sister groups should generally not be used, because polarity cannot be assessed with confidence until cladistic relationships within one or more of them can be resolved. If however, the distribution of character states within an outgroup is extremely lopsided, it is common practice to proceed by invoking the commonality principle within that outgroup. This may sometimes be justifiable on pragmatic grounds, but it must be recognized that an additional assumption has been made that may introduce an error in the analysis (see Cantino 1982 for further discussion of this problem). If subsequent study of relationships within that outgroup indicates that the rare state is in fact ancestral in the group, then the initial analysis in which the commonality principle was applied must be revised.

When polarity assessments depend on precise knowledge of outgroup relationships, the best way to proceed is to perform a higher level cladistic analysis to resolve these relationships. In some instances this may be feasible but in others it might require a decade or lifetime of work. If the latter is true, it may still be possible to make some headway. Certain combinations of outgroups may be less reasonable than others. For example, judging from published "phylogenies" of the angiosperms, it is far less likely that Rubiaceae and Araliales are outgroups 1 and 2 of Dipsacales than that Cornales and Araliales are outgroups 1 and 2 (i.e., phylogenists have suggested that Dipsacales are

most closely related *either* to Rubiaceae *or* to Cornales/Araliales, while maintaining that Rubiaceae and Cornales/Araliales are only distantly related to each other). Use of this kind of argument extends the concept of "plausibility" to combinations of outgroups, thereby eliminating the need to consider every possible arrangement of plausible sister groups. To do so would indeed be tedious, e.g., for five plausible sister groups there are 120 ladderlike cladograms involving all five, 205 more using only four, three, two, or one of them, and many others that are not ladderlike.

CONGRUENCE AND INCONGRUENCE OF CLADOGRAMS

After character polarities are assessed using every plausible arrangement of outgroups, an ingroup cladogram is constructed for each. For complex data sets involving many ingroup terminal taxa and many characters, such cladograms are best obtained using a computer-assisted parsimony algorithm such as a Wagner program (Farris 1970). The cladograms are then compared to locate areas of congruence. Whole cladograms need not, and often will not, be entirely congruent. Discovery of even a few congruent clades may be helpful and may make it possible to resolve the ingroup further by use of the functional ingroup/functional outgroup approach described in the next section (Watrous and Wheeler 1981).

With ingroup cladograms that have relatively few terminal taxa, such as those used in our examples, areas of congruence will be easy to identify. With more complicated cladograms this may become increasingly difficult, and a rigorous method for finding such groups could be employed. For our purposes, constructing an Adams consensus tree (Adams 1972) is inappropriate because a clade in such a tree need not be present in either of two cladograms that are being compared, i.e., the tree may contain compromise branching points (Sokal and Rohlf 1981; Seaman and Funk 1983). We are interested in finding components shared by all cladograms, and for this purpose a Nelson consensus tree (Nelson 1979; Nelson and Platnick 1981) seems appropriate.

Because a number of equally parsimonious but topologically different cladograms can often be obtained from a given data set, one may ask how such suites of cladograms should be compared both within and among different data sets (each data set represents one outgroup arrangement). We recommend a conservative approach in which congruent clades are sought among all of the most parsimonious cladograms obtained from all of the outgroup arrangements considered. However, as Schuh and Farris (1981) pointed out, this conservative approach is probably inappropriate if one's objective is a quantitative comparison of congruence for purposes of assessing the stability obtained by different classificatory methods.

What conclusions can be drawn from incongruent cladograms or portions thereof? In our opinion, only complete congruence is an acceptable basis for phylogenetic conclusions. If one wishes to discuss relationships involving incongruent areas in the cladograms, the different ingroup cladograms derived from different outgroup hypotheses should be clearly presented. This is helpful in that it identifies a problem that can perhaps be solved by future work on the more inclusive cladistic relationships of the ingroup.

It might be tempting to invoke a probability argument to draw a conclusion despite incongruence of cladograms. For example, if three out of four cladograms are identical, one might contend that there is a 75% chance that this cladogram is correct. In our view this form of argument is inappropriate. A basic assumption of the outgroup substitution approach is that one and only one of the outgroup arrangements being considered is the correct one. However, because the relationships of the plausible sister groups to the study group are not known, it is assumed for the sake of the analysis that each outgroup arrangement is equally likely to be correct. We can only hope that the first assumption is true; the second is obviously not true but is necessary because we are aware of no rigorous procedure for assessing and quantifying probable relationships of the outgroups to the study group. To employ probability in choosing among incongruent ingroup cladograms would place too much weight on the operational assumptions, which, after all, are based on ignorance. In so doing, one might effectively rule out an outgroup arrangement that may be the true one. Moreover, assignment of a numerical probability to an ingroup cladogram may convey an unwarranted

sense of confidence in a particular outcome. The strength of the approach outlined here, with its insistence on complete congruence among cladograms (or portions thereof), is that it can generate robust hypotheses in spite of initial uncertainties. Anything less than complete congruence maintains or may even compound the initial uncertainties and, in our opinion, renders any conclusions about relationships far less compelling. Ultimately, the best way to rule out a particular incongruent cladogram is to demonstrate that the outgroup assumptions that underlie it are less parsimonious than other outgroup assumptions, which would require a higher level cladistic analysis.

The likelihood of obtaining incongruent cladograms is a function both of the number of terminal taxa in the ingroup and the number of plausible sister groups employed. The fewer the ingroup taxa and the fewer the plausible sister groups, the greater the likelihood that completely congruent cladograms will be obtained. However, even with numerous ingroup taxa and many plausible sister groups, it may still be possible to identify some robust clades.

CHARACTER SUPPORT FOR CLADOGRAM STRUCTURE

In this section we consider the causes of congruence among cladograms obtained with the outgroup substitution approach and implications regarding the study of character evolution within the ingroup. The most obvious, and perhaps most important, cause of congruence is the use of characters for which all plausible sister groups have the same state, and hence identical polarity assessments are obtained for all outgroup arrangements. If every character were of this sort, the resulting cladograms would be fully congruent, and it would be possible to specify which character states support each clade. This ideal situation is probably rare.

Even if most or all characters used in the analysis vary among the outgroups, and/or particular characters are used with some outgroup arrangements and not others, some clades or even the entire ingroup cladogram may be congruent. This is so because particular clades may be supported by different sets of synapomorphies depending on which arrangement of outgroups is used.

There are two main reasons why one might use a character with some outgroup arrangements but not with others. First, if an equivocal polarity assessment is obtained with a particular outgroup arrangement, then for that arrangement the character cannot be used to infer ingroup cladogram structure at this stage in the analysis. The character may still be useful, however, with other outgroup arrangements and in subsequent analyses of unresolved portions of the ingroup cladogram. If some cladistic structure within the ingroup can be established, certain clades within the ingroup can be used as functional outgroups for the unresolved clades (Watrous and Wheeler 1981). In doing so, one should consider the functional outgroup along with the plausible sister groups of the ingroup in determining polarity. This could be very helpful. For example, if the two basal clades within the ingroup have the same state, this will be the most parsimonious plesiomorphic state for the rest of the ingroup. This assessment will hold regardless of the distribution of states among outgroups further removed from the ingroup (the "first doublet rule" of Maddison et al. in press).

Secondly, a character that can be used with one outgroup may not be applicable with another because of uncertainty regarding the homologous condition in the latter. For example, if all members of the ingroup have simple leaves, some with entire margins and some with lobed, and a plausible outgroup has compound leaves with entire-margined leaflets, is the entire margin of the simple leaf in the ingroup homologous with the entire margin of the leaflets in that outgroup? Quite possibly not. The lobed leaves of some ingroup members might represent an intermediate state in a transformation series from simple, entire leaves to compound leaves or vice versa. Because of the uncertainty about homology, leaf margin would have to be omitted from the analysis when this outgroup is used.

An important implication of the above paragraphs is that it may be possible to arrive at a robust hypothesis of ingroup relationships using the outgroup substitution approach but not to detail which character states support each clade. This is a serious limitation of our approach, inasmuch as one purpose of cladistic analysis is to formulate hypotheses concerning character evolution within a study group and to detail character support for particular clades.

However, the omission of a character in relation to some outgroup arrangements does not necessarily mean that it will be impossible to make positive statements about the evolution of the character in the ingroup. There may be a single most parsimonious arrangement of its states on the robust cladogram obtained using other characters. This a posteriori fitting of characters to cladograms is essential if one is interested in tracing evolution of particular characters (Farris 1970; Mickevich 1982; Swofford and Maddison, in prep.).

It may be most appropriate to use the outgroup substitution approach when a hypothesis of cladistic relationships is desired but when it is not critical to know where on the cladogram particular character state changes occur. For example, the method is especially useful if one's interest in cladistic relationships in the ingroup is primarily to establish an outgroup hypothesis for a subsequent lower level cladistic analysis. It has been successfully employed to establish Scrophulariales as the best-supported outgroup for subsequent cladistic analysis of Lamiales (Cantino 1982) and to determine that the clade comprising *Sambucus* and *Adoxa* is the sister group of *Viburnum* (Donoghue 1983b) in order to assess polarities and determine cladistic relationships within *Viburnum* (Donoghue 1983a).

Although we have observed that congruent cladograms can be obtained for different reasons, we have not addressed an interesting question: Under what circumstances will an incorrect outgroup arrangement result in an ingroup cladogram that is congruent with that obtained using the true outgroup arrangement? We cannot answer this question. There are so many different ways to obtain congruent cladograms that it may not be possible to make meaningful generalizations. Fortunately, the logic of the outgroup substitution approach does not rest on the answer. It is not essential to know why congruent cladograms are obtained with incorrect outgroup arrangements, but only to know that one of the outgroup arrangements used is the correct one. The outgroup substitution approach is very conservative; only congruent cladograms, or portions thereof, are accepted. Thus, although this approach may yield a highly unresolved cladogram, one can be certain that whatever robust clades are obtained would also be obtained if a cladogram were constructed using only the true outgroup arrangement.

CONCLUSIONS

Although there is no completely satisfactory substitute for a well-corroborated hypothesis of outgroup relationships, we recognize the need to develop methods to cope with the common problem of uncertain outgroup relationships. The approach advocated here is particularly helpful when some of the plausible sister groups may be only distantly related to others, a common situation given our current state of uncertainty about angiosperm phylogeny.

The outgroup substitution approach is not a rigid set of procedures. It does not necessarily require an exhaustive testing of every possible arrangement of conceivable outgroups. Rather, it is a flexible approach to the problem of uncertain outgroups that permits the generation of robust phylogenetic hypotheses in an explicit manner. Some subjectivity is involved inasmuch as the investigator must exercise judgment in selecting "plausible" sister groups and outgroup arrangements. Because of this, it is imperative that it be stated which ingroup cladograms were generated on the basis of which outgroup arrangements. Some workers may disapprove of an approach that permits such subjectivity. But our procedure is at least explicit and is thus an improvement on the rather vague applications of multiple outgroups that have appeared in recent botanical cladistic literature. We look forward to the day when angiosperm phylogeny is well enough understood that approaches such as ours are no longer necessary. In the meanwhile, however, we recommend it as a practical method of generating tentative phylogenetic hypotheses on the basis of clearly stated outgroup assumptions.

LITERATURE CITED

ADAMS, E. N. 1972. Consensus techniques and the comparison of taxonomic trees. Syst. Zool. 21: 390–397.

ARNOLD, E. N. 1981. Estimating phylogenies at low

taxonomic levels. Z. Zool. Syst. Evol.-Forsch. 19: 1–35.

BARABÉ, D. 1982. Le critère de comparaison hors-groupe et son application à la systématique des angiospermes: Cas de Hamamélidales. Compt. Rend. Hebd. Séances Acad. Sci., Série III 295: 755–758.

BAUM, B. R. 1983. A phylogenetic analysis of the tribe Triticeae (Poaceae) based on morphological characters of the genera. Canad. J. Bot. 61:518–535.

BOLICK, M. R. 1981. A cladistic analysis of *Salmea* DC. (Compositae-Heliantheae). Pp. 115–125 in *Advances in cladistics* [Proceedings of the first meeting of the Willi Hennig Society], eds. V. A. Funk and D. R. Brooks. New York: New York Bot. Gard.

CANTINO, P. D. 1982. Affinities of the Lamiales: A cladistic analysis. Syst. Bot. 7:237–248.

COOMBS, E. A. K., M. J. DONOGHUE, and R. J. MC-GINLEY. 1981. Characters, computers, and cladograms: A review of the Berkeley cladistics workshop. Syst. Bot. 6:359–372.

DONOGHUE, M. J. 1983a. A preliminary analysis of phylogenetic relationships in *Viburnum* (Caprifoliaceae s.l.). Syst. Bot. 8:45–58.

———. 1983b. The phylogenetic relationships of *Viburnum*. Pp. 143–166 in *Advances in cladistics*, vol. 2 [Proceedings of the second meeting of the Willi Hennig Society], eds. N. I. Platnick and V. A. Funk. New York: Columbia Univ. Press.

——— and W. P. MADDISON. 1982. Outgroups and ontogeny: Parsimony, information, and polarity assessment (Abstract). Bot. Soc. Amer. Misc. Publ. 162:78.

ELDREDGE, N. 1979. Cladism and common sense. Pp. 165–198 in *Phylogenetic analysis and paleontology*, eds. J. Cracraft and N. Eldredge. New York: Columbia Univ. Press.

——— and J. CRACRAFT. 1980. *Phylogenetic patterns and the evolutionary process*. New York: Columbia Univ. Press.

FARRIS, J. S. 1970. Methods for computing Wagner trees. Syst. Zool. 19:83–92.

———. 1982. Outgroups and parsimony. Syst. Zool. 31:328–334.

GUÉDÈS, M. 1983. A response from Guédès. Taxon 32:277–278.

KRESS, W. J. and D. E. STONE. 1983. Morphology and phylogenetic significance of exine-less pollen of *Heliconia* (Heliconiaceae). Syst. Bot. 8:149–167.

MADDISON, W. P., M. J. DONOGHUE, and D. R. MADDISON. In press. Outgroup analysis and parsimony. Syst. Zool.

MICKEVICH, M. F. 1982. Transformation series analysis. Syst. Zool. 31:461–478.

NELSON, G. 1979. Cladistic analysis and synthesis: Principles and definitions, with a historical note on Adanson's *Familles des Plantes* (1763–1764). Syst. Zool. 28:1–21.

——— and N. PLATNICK. 1981. *Systematics and biogeography: Cladistics and vicariance*. New York: Columbia Univ. Press.

PLATNICK, N. I. and V. A. FUNK. 1983. Summary. Pp. 199–200 in *Advances in cladistics*, vol. 2 [Proceedings of the second meeting of the Willi Hennig Society], eds. N. I. Platnick and V. A. Funk. New York: Columbia Univ. Press.

SCHUH, R. T. and J. S. FARRIS. 1981. Methods for investigating taxonomic congruence and their application to the Leptopodomorpha. Syst. Zool. 30:331–351.

SEAMAN, F. C. and V. A. FUNK. 1983. Cladistics analysis of complex natural products: Developing transformation series from sesquiterpene lactone data. Taxon 32:1–27.

SOKAL, R. R. and F. J. ROHLF. 1981. Taxonomic congruence in the Leptopodomorpha re-examined. Syst. Zool. 30:309–325.

STEVENS, P. F. 1980. Evolutionary polarity of character states. Annual Rev. Ecol. Syst. 11:333–358.

SWOFFORD, D. L. and W. P. MADDISON. In prep. Fitting character state distributions to Wagner Trees. Submitted to Syst. Zool.

WATROUS, L. E. and Q. D. WHEELER. 1981. The outgroup comparison method of character analysis. Syst. Zool. 30:1–11.

WHEELER, Q. D. 1981. The ins and outs of character analysis: A response to Crisci and Stuessy. Syst. Bot. 6:297–306.

WILEY, E. O. 1980. Phylogenetic systematics and vicariance biogeography. Syst. Bot. 5:194–220.

———. 1981. *Phylogenetics: The theory and practice of phylogenetic systematics*. New York: John Wiley and Sons.