Journal of Systematics and Evolution 46 (2): 109–129 (2008) doi: 10.3724/SP.J.1002.2008.07166
(formerly Acta Phytotaxonomica Sinica) http://www.plantsystematics.com

Fifty years of character compatibility concepts at work

George F. ESTABROOK*

(*Department of Ecology and Evolutionary Biology, The University of Michigan,* Ann Arbor, MI 4809-1048, USA)

Abstract In the mid 19th century, systematic biologists realized that observable similarities and differences among a group of related species could be the basis for hypotheses about the evolutionary relationships among the species and their ancestors. Such hypotheses can be expressed as characters. A character is comprised of two or more character states of species considered to be similar with respect to a basis for comparison. The states of a character may also be arranged into a character state tree to hypothesize speciation events associated with changes from one character state to another. In the mid 20th century, some systematists realized that sometimes pairs of characters (or character state trees) could be incompatible as hypotheses, i.e., they could not both be true. Through the 1950s, '60s and '70s, tests for, and ways to resolve, incompatibilities were used to estimate an ancestor relation based on mutually compatible characters. An estimate was often shown as a diagram connecting ancestors to their immediate descendants (not quite correctly) called a phylogenetic tree. More recently, other applications of compatibility concepts have been developed, including: identify characters that appear to be random in the context of their data set; combine estimates of ancestor relations for subsets of taxa in a larger collection into a single estimate (a so-called supertree) for the whole collection; and interpret geographic patterns in an evolutionary context.

Key words biogeography, character compatibility, character evaluation, convex groups, phylogenetic trees, reciprocal illumination, speciation, supertrees

This review begins with a discussion of character state change, which distinguishes between changes that occur within species, often in similar ways within several related species, and changes that occur when a new species evolves from its ancestral species. The former changes are sometimes modeled as a random process; the latter are often used to describe or identify species or higher taxa. A challenge faced by a systematist working near the species level is to distinguish the former from the latter kind of change. Although we prefer the latter for describing species and studying their evolutionary relationships, changes stably associated with the evolution of new species may remain difficult to distinguish. Some changes become more clearly associated with evolutionary events when taxa above the species level, such as species complexes, sections, or genera, are compared. Generally, throughout this review, I will speak of the evolutionary unit, EU, as if it were a species, and of the process by which one EU evolves from its immediate ancestral EU as speciation. However, compatibility concepts apply above the species level, and in the view of many systematists are more appropriately applied above the species level. In any case, to understand correctly the concepts and applications of character compatibility theory, it is important to realize that characters and character state trees should be based on changes stably associated with speciation (or its analog among higher taxa).

How character state trees express hypotheses of speciation associated with character state change and how true character state trees can be combined are the subjects of sections 1, 2, and 3. Character state trees are partial estimates of the ancestor relation, which indicates which species evolved from which ancestors. Character state trees, as hypotheses, are either true or false. If two character state trees are true, then they can be added (combined) to make a single character state tree that refines the partial estimates of each. Of course we do not know which character state trees are true as hypotheses, but we can attempt to apply this addition process to any two character state trees; the process will fail to produce a new character state tree if and only if the two character state trees are incompatible as hypotheses. The explanations and examples of these three sections are important to a basic understanding of character compatibility concepts, but they can be skimmed by readers more interested in history.

Section 4 recognizes authors writing in the 1960s and '70s who became aware that characters could be incompatible. Soon it was realized that incompatibility could result from the particular arrangement of the

Received: 19 December 2007 Accepted: 4 March 2008
* E-mail: gfe@umich.edu; Tel.: 1-734-764- 6219; Fax: 1-734-763-0544.

states into character state trees, so that re-arrangements of the states could resolve incompatibility. However, the membership of EUs in the states themselves could also cause incompatibility. Tests to reveal the various causes of incompatibility were discovered and described. By the early 1980s, these concepts were in place and their properties elucidated.

Section 5 examines how incompatibilities were resolved to produce mutually compatible characters from which an estimate of the ancestor relation could be easily constructed. When Hennig's manuscript was translated from German into English and published in 1966, it provided many early career systematists in the non-German speaking world with some very specific guidelines for resolving incompatibilities. Although Hennig held radically new views of what constitutes a higher taxon, for many early career systematists his Teutonic explicitness was a refreshing alternative to the very subjective approaches of the well established evolutionary systematists, who were making it difficult for other than their favored apprentices to access professional opportunities. Predating DNA sequence data, Hennig's approach advocated the use of almost anything known about natural history of organisms to resolve incompatibilities. Soon a new, powerful and doctrinaire school of thought (Cladism) became established, nominally recognizing Hennig as their guide but largely forgetting his advocacy, adopting instead an automatic criterion (parsimony, proposed by others in the 1960s) for resolving incompatibilities that could be applied by computers to DNA sequence data without the participation (intervention, responsibility) of the systematist. Fortunately, the power of Cladism over young minds in our field has largely dissipated.

The application of probability to compatibility concepts is the subject of section 6. Through the 1970s and '80s, population geneticists used probability concepts applied to changes that they construed as happening at random to study evolution below the species level, using maximum likelihood to estimate "phylogenetic" tree branching patterns. Although systematists using compatibility concepts would seek to avoid random changes as the basis for characters used to estimate the ancestor relation, maximum likelihood estimates were used by some systematists, especially those relying primarily on DNA sequence data, because it too enabled an automatic computer estimate of the ancestor relation. In the 1990s, probability concepts were used with character compatibility to recognize characters whose state composition could not be distinguished from random, in the context of their data set. This is especially important for DNA sequence data in which some sites may represent bases for comparison more appropriate for analysis by random models, such as maximum likelihood. Preliminary evidence suggests that elimination of such sites may clarify estimates based on distance or parsimony methods.

Section 7 discusses other applications of compatibility, which include elucidating biogeographical patterns, suggesting hybridization events in the evolutionary history of a study group, incorporating stratigraphic information into the estimation of the ancestor relation, and constructing supertrees. The relationships among compatibility, monophyly and classification are discussed in the final section 8.

1 Concepts of character state change

For centuries, people have recognized groups of organisms that are similar, and arranged them hierarchically into larger, increasingly less similar, groups. In the late 19th century, the theory of evolution provided systematists with a mechanism to explain similarities and differences among kinds or organisms, but it had very little impact at that time on our view of relationships reflected in the higher taxa traditionally recognized. Students of evolution developed an understanding of speciation and change, well explained by the great writers of the mid 20th century, such as Stebbins (1950), Mayr (1963) and Grant (1963). More recently, the mechanisms and consequences of speciation are discussed by many authors in Otte and Endler (1989). At the individual and population level, genetic changes occur "at random" by mutation, or by chromosomal rearrangement from one generation to the next. Within a breeding population, over generations these changes may be lost or they may spread to many other individuals (by natural selection); or they may be irrelevant to survival and reproduction and so drift at random. When populations become genetically isolated (spatially, temporally, behaviorally, etc.), it is no longer possible for changes in one to be spread to the other. Over time, this may result in the accumulation of sufficient differences that it is no longer genetically possible for members of one population to breed with members of the other; speciation has occurred.

Thus, systematic biologists looked for observable expressions of the changes associated with speciation to recognize and describe distinct species. They are confounded in their task by many natural phenomena

that produce observable differences between individual organisms that are NOT changes associated with speciation, such as juvenile and adult forms, sexual dimorphism, developmental anomalies caused by damage or disease, facultative response to environment, clinal variation over space, and of course the within population genetic variation described above. In the early 20th century, common gardens, experimental breeding, and larger numbers of specimens to study have helped systematists recognize more accurately changes associated with speciation (Briggs & Walters, 1969). Although changes associated with speciation provide potentially relevant data with which to estimate the history of speciation, it was not until the middle of the 20th century that some systematists began to carefully consider concepts that would enable them to use changes associated with speciation to estimate evolutionary relationships among species, and use those estimates to recognize higher taxa.

One concept construes changes associated with speciation to differ from changes that occur "at random" in individuals and that are sometimes spread over generations through breeding populations; species changes have the same origin, but after speciation they usually can no longer be spread by breeding between members of different species. Changes associated with speciation interrupt phyletic continuity over time (Estabrook, 1972). This concept gives rise to a historical-biological species concept in which a species evolves at a time in the past, usually in a somewhat restricted geographic area, persists through time possibly dispersing to other geographic areas, and ultimately goes extinct (a few still extant species have not gone extinct yet). During the life of a species, a population may become isolated and independently evolve enough genetic difference so that its members can no longer breed with members of the species from which it was isolated, as described above. In this way, one species becomes the immediate ancestor of another that evolved from it. This process may happen repeatedly, so that one species may come to be the immediate ancestor of several distinct species. In this way, the study of evolutionary relationships among extant species implicates ancestral species that existed over past time, and has for one of it principal objectives an estimate of the ancestor relation among related species, past and present. This view of speciation and systematics has come to be called evolutionary systematics and is represented by Simpson (1961) and Mayr (1969) among many others over the past 40 years, including more recently Skelton (1993).

2 Characters and character state trees

We rarely know for sure the branching pattern of phyletic lines leading up through time to the extant species under study, but to illustrate the concepts, we consider a hypothetical case in which we do. Suppose that we are studying a group S of six extant species (a, b, c, d, e, f) whose phyletic lines branch upward through time as shown on the left in Fig. 1. Whenever a phyletic line branches, one branch represents a new species and the other represents the continuation of the ancestral species. Arrows identify speciation events and point to the branch created by changes that produced a new species. The diagram on the right of Fig. 1, shows the ancestor relation that results from these speciation events; a line is drawn upward from an ancestral species to any immediate descendant species. Thus, each line in the diagram represents a speciation event on the phylogenetic tree. We say that species x is an ancestor of species y if there is a series of one or more upward lines leading from x to y.

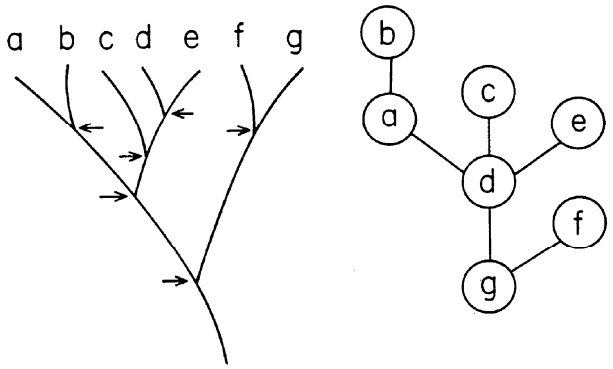

Fig. 1. Hypothetical branching of phyletic lines leading to 7 species (a, b, c, d, e, f, g) on which speciation events are indicated by arrows. The diagram of the corresponding ancestor relation is shown on the right.

On the same hypothetically true branching pattern of phyletic lines, the speciation events could have occurred at different times and places. Figure 2 shows an example. You can see from the diagram of the resulting ancestor relation that it is quite different from the ancestor relation of Fig. 1. This example should make it clear that there is not a one-to-one relationship between phyletic branching patterns (phylogenetic trees) and ancestor relations, because the latter are a result of the historical speciation events that created the species we study. Although it is possible to have extinct species represented in S, some ancestral species may not be represented in S because of extinction. Figure 3 shows the same speciation

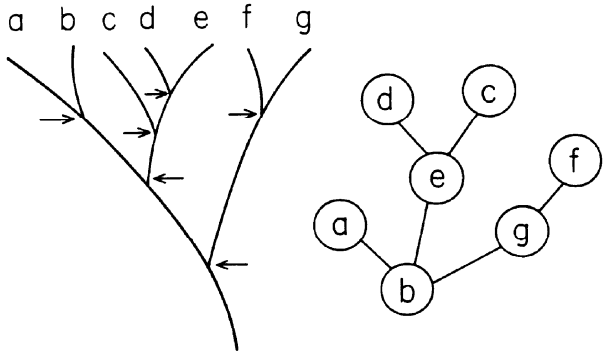

Fig. 2. The same phyletic lines as in Fig. 1, but with speciation events in different places. The diagram of the corresponding ancestor relation is shown on the right. This illustrates how different ancestor relations can be even when the branching of the phyletic lines is the same.

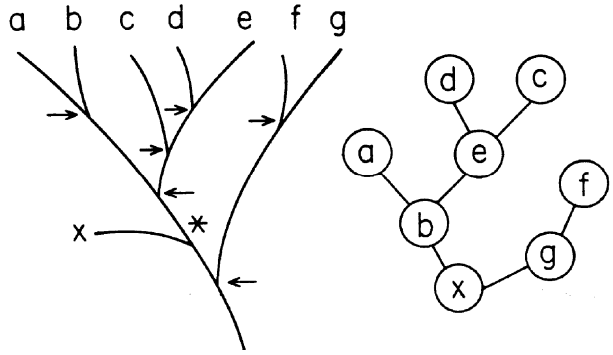

Fig. 3. The same phyletic lines as in Fig. 1 with the addition of an extinct species, x. The diagram of the corresponding ancestor relation is shown on the right.

events as Fig. 2 together with an additional speciation, indicated with *, after which the ancestral species, x, goes extinct and so is not found among the species in S; the resulting ancestor diagram is shown on the right. You can imagine the variety of speciations and extinctions that could occur on a phylogenetic tree and the resulting variety of ancestor relations.

One of the tasks of systematics is to use the similarities and differences that can be observed among the species in a group under study to estimate their ancestor relation. To examine more explicitly how these concepts relate to this task, it is appropriate to define some terms. If one can observe a given structure for each species in a collection S of species under study, and recognize distinct variations, then the species can be placed into groups so that those in the same group look the same with respect to that basis for comparison, but those in different groups look different. Such a basis for comparison is called a character, and the groups that result are called its character states. For a character to be relevant to the ancestor relation, its states should be based on changes associated with some of the speciation events that created the species in S, as discussed above. Of course, not all speciation events need be associated with a change in the structure defining a given character, but when that structure did change, it should have been in association with a speciation event. Such a character can be used as the basis for a hypothesis about the ancestor relation. This hypothesis is expressed as an ancestor relation diagram in which the character states play the role of individual species, and as before, a line leading up from one state to another represents a speciation event on the phylogenetic tree, as shown in Fig. 4.

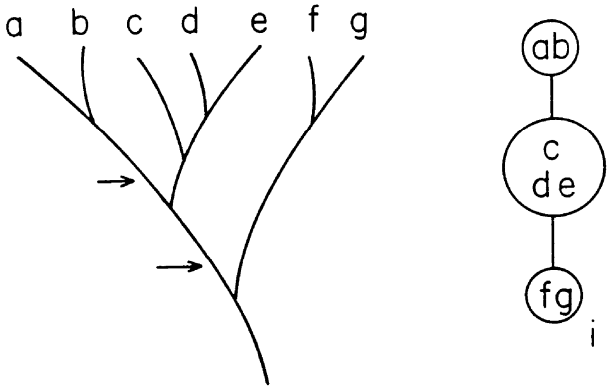

Fig. 4. The same phyletic lines as in Fig. 1, but with only two speciation events indicated. The corresponding character state tree is shown on the right.

The character state tree (CST) shown on the right represents the two speciation events shown by arrows near the phylogenetic tree on the left. The speciation indicated by the lower arrow changed a structure from the form exhibited by species f and g to the form exhibited by species c, d and e; the speciation indicated by the upper arrow changed that form to the one exhibited by species a and b. If the phylogenetic tree and the speciation events shown on the left of Fig. 4 are historically correct, then we would say that the hypothesis of the CST is true (or simply that the CST is true) because it corresponds to speciations on the true phylogenetic tree (Estabrook et al., 1975).

If two CSTs are true, then by considering all the speciation events that correspond to one or the other or both of them, another true CST (called the sum of the first two) is determined, as shown in Fig. 5. The "sum" CST is a refinement of either of the two CSTs that were added, because it represents all the

speciation events of either. In the same way, another true CST could be added to this sum to create an even more refined CST, as shown in Fig. 6. The two changes distinguishing state (d, e) represented by CSTs iii and iv may have been associated with the same speciation event, or with different speciation events suggesting the possibility of an extinct ancestral species represented by the dotted circle. However, the most ancestral state contains only extinct ancestors because of speciation events on both branches of the phyletic lines leading up to the extant taxa. A moment's consideration should make it clear that if enough true CSTs are added, then the sum CST becomes the diagram of the ancestor relation itself, in which will be shown, in their historical place, ancestors not represented in S. Thus, an ancestor relation is a CST sufficiently refined so that each state has at most one species.

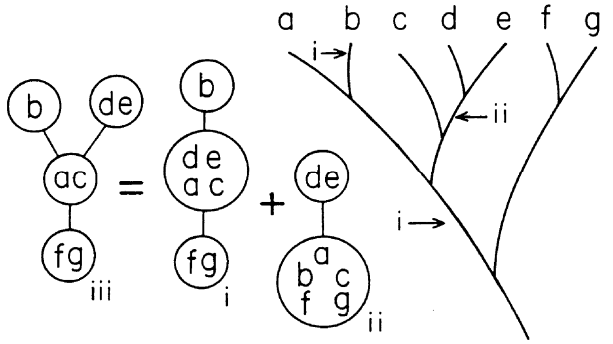

Fig. 5. The sum of two character state trees, with corresponding speciation events indicated on the phyletic lines of Fig. 1 at the right.

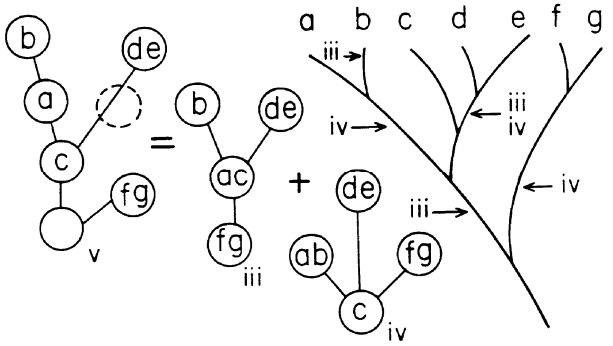

Fig. 6. A third character state tree is added to the sum shown in Fig. 5.

3 Compatible character state trees

Of course, not all CSTs are true; there are three basic ways in which they can be false, as shown in Fig. 7. On the right of Fig. 7 we again see our hypothetically true phylogeny, and on the left three false CSTs. They are false because there is no possible way that speciations could have occurred on this true phylogenetic tree so that they would correspond to the lines in the CSTs. The leftmost misrepresents the direction (trend) of the changes, because changing the direction of change so that state (f, g) is primitive does make it possible to put speciation events on the phylogenetic tree so that this CST would be true. The middle CST connects pairs of states that cannot all be next to each other and still represent speciation events on the phylogenetic tree; redirecting these proximity relations cannot make a CST whose speciation events can be placed on the true phylogenetic tree. However, attaching state (c) to state (d, e) instead of to state (a, b) does make this possible. For the rightmost CST, there is no way to place a speciation on the phylogenetic tree so that even its states would result.

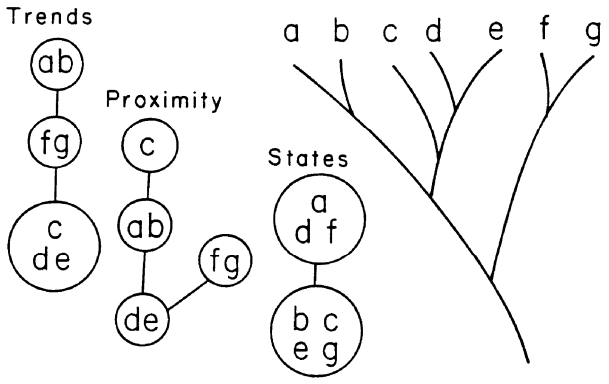

Fig. 7. Three character state trees each false for a different reason.

Two CSTs do not have to be both true for it to be possible to add them; they can be added so long as there is some phylogenetic tree (true or not) on which all their speciations can be simultaneously placed. Then from the placement of these speciations on this phylogenetic tree the sum CST can be constructed. But how can we find such a phylogenetic tree? Estabrook and McMorris (1980) demonstrated that we do not need to. They showed that there is a one-to-one correspondence between CSTs and trees of subsets of S. A collection of subsets of S is called a tree of subsets of S if it satisfies two properties: S itself is one of the subsets, and any two subsets either have no species in common or one contains all the species that are in the other. Each character state in a CST has a subset in its tree of subsets consisting of all the species in that state plus all the species in any descendant

state. Thus, the most primitive state has for its subset S, the entire study collection of species. States with no descendant states have for their subset only the species that they contain themselves. The correspondence is shown in Fig. 8, where below each CST is shown its tree of subsets, arranged so that derived states are above their ancestors. The sum is determined by combining the subsets from both trees of subsets; if the result is itself a tree of subsets, then the sum is the corresponding CST. In Fig. 8, the trees of subsets of the two CSTs to be added are combined to make the tree of subsets in the lower right; finally the CST above is constructed using the principles described above. This CST is the sum of the first two.

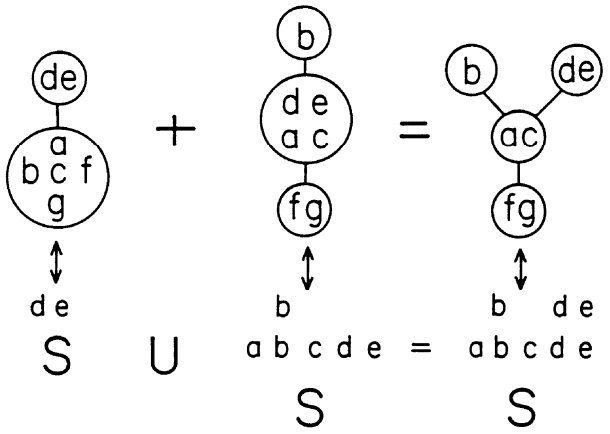

Fig. 8. The sum of two character state trees determined by their respective trees of subsets.

Two CSTs do not have to be true for one to be a refinement of the other. We can readily see from the sum of two CSTs that if the tree of subsets of one CST contains all the subsets in the tree of subsets of another CST, then the first is a refinement of the second. The diagram of the relation "is a refinement of" among CSTs makes a semi-lattice. A diagram of this semi-lattice for the simple case in which S contains only 3 species is shown in Fig. 9. The refinement relation has been studied theoretically by Estabrook and McMorris (1980), McMorris and Zaslavsky (1981), and Janowitz (1984).

Not every pair of CSTs can be added. When the union of their two trees of subsets is NOT a tree of subsets, as shown in Fig. 10, then there is no phylogenetic tree on which to place speciations that correspond to the lines in both CSTs. Typically we do not know the true phylogenetic tree so we can not test a CST to discover whether it is true or not. However, if two CSTs cannot be added, then they cannot both be true; they are incompatible as hypotheses about the ancestor relation among the species and their ancestors under study (Estabrook, 1984). Two CSTs that can be added are compatible as hypotheses about the ancestor relation. This concept of character compatibility, and others related to it, form the basis of a variety of compatibility methods developed and used over the past 50 years.

4 Early history of compatibility concepts

In the mid 1960s several workers independently became aware that it may be the case that a feature of some (but not all) species in a group under study cannot be associated with a single speciation event on a phylogenetic tree if another feature is. If the historical branching pattern of phyletic lines leading upward to extant species does in fact form a tree, then the two hypotheses that they each can be so associated are incompatible. Wilson (1965) pointed out that if the group of species with one feature were either distinct from, or entirely contained in, the group with the other, then the hypotheses were compatible, but if the groups partially overlapped then they were not. You can see that, in the case of two-state CSTs, this test is identical to addition of CSTs by trees of subsets. Hennig (1966, page 121 and Fig. 36) describes, if more prolix, the same basic test. He (or his translators) calls incompatible features "incongruous", and points out that one of two incongruous features must have been wrongly interpreted, i.e., false. Of course they both may have been "wrongly interpreted". Camin and Sokal (1965) consider several possible exclusive states of the same homologous structure arranged in a sequence to hypothesize the order in which they evolved to produce a linear, multi-state CST. They recognized that such CSTs could be contradictory under the assumption that the branching pattern of the phyletic lines leading to the extant species under study forms a tree.

Once aware of these contradictory hypotheses, the question of how to resolve them needed to be addressed. Some of the contradictions may be a consequence of errors in direction of evolution, i.e., a feature may have been lost instead of gained, in which case the group of species not exhibiting the feature should be the one participating in the tests. One approach to this problem is to hypothesize different directions in an attempt to resolve contradictions. Another approach is to eliminate direction from the hypothesis by considering CSTs undirected. LeQuesne (1969) proposed a compatibility test for

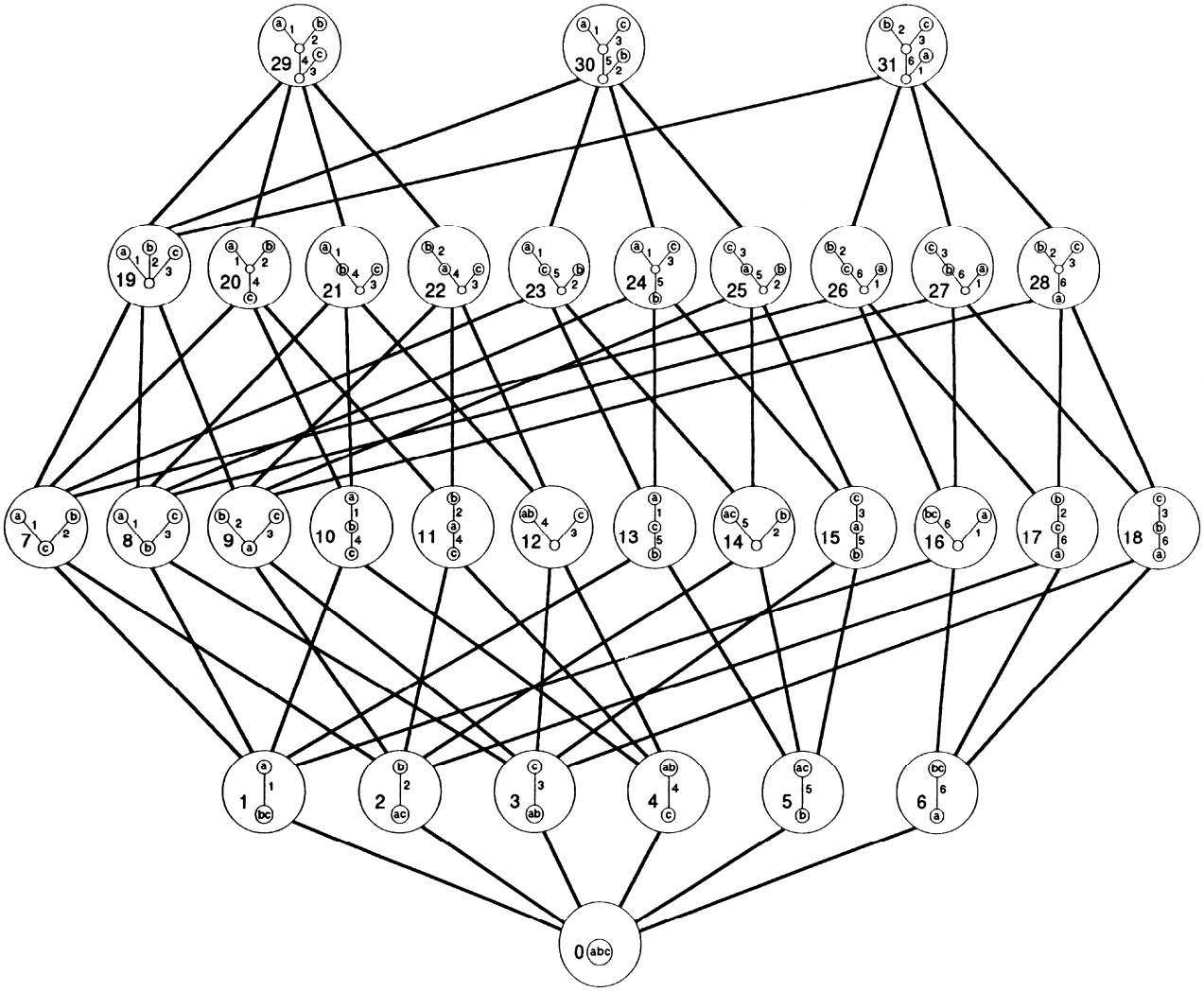

Fig. 9. The lower semi-lattice of the refinement relation among all possible ancestor relations for a study collection of three species. The original figure was drawn by C. A. Meacham in 1989 as class notes for his students at the University of California at Berkeley, CA, USA.

undirected two-state CSTs; if all four possible combinations of present and absent for two features are represented among the species in a study collection, then the two-state CSTs associated with those features are incompatible. Estabrook and Meacham (1979) presented a test for undirected multi-state CSTs. They proved that in a CST there is always a state that can be designated as the most primitive so that the number of species in any state x immediately derived from it, plus the sum of the number of species in all the states derived from state x, is never more than half the number of species in S. A CST directed with such a state most primitive is said to be directed "common equal primitive". They then proved that if two common equal primitive CSTs are incompatible, then they will remain so directed in any other way. Thus if CSTs are directed common equal primitive a compatibility test by trees of subsets will also test them as undirected hypotheses. Estabrook (1977) suggested that systematists might be tempted to believe that common equal primitive because directing change that way eliminates conflicts due to direction alone. Meacham (1984a) made the then controversial suggestion that the role of hypothesizing direction of evolutionary change before estimating the ancestor relation among species under study could be reduced, or even eliminated, by reasoning with undirected CSTs, especially in cases with little or no a priori evidence to identify a primitive condition. Donoghue and Maddison (1986) objected on philosophical

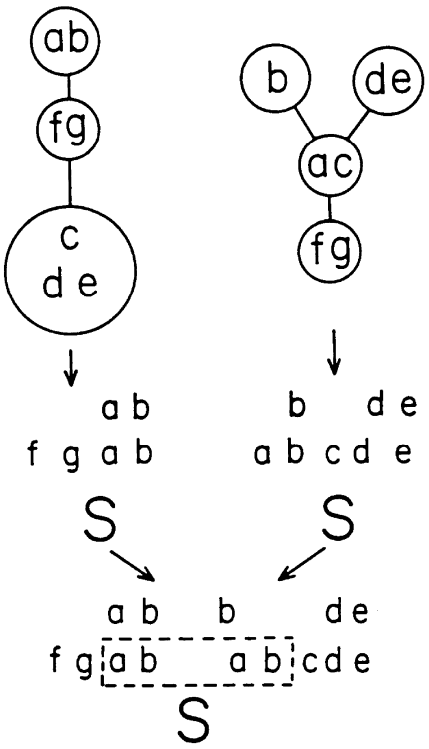

Fig. 10. The incompatibility of two character state trees, discovered using their respective trees of subsets.

grounds that were later shown to be irrelevant to modern methods.

With two-state CSTs, there are no incompatibilities that can be resolved by changing proximities of states in CSTs because there is only one proximity. However, with more general CSTs, we can ask for two incompatible CSTs whether there are any other CSTs with the same respective states that are compatible. This is especially relevant when the data source provides predominantly multi-state characters. In the 1970s protein sequencing became more common. Species could be compared based on which amino acid appeared in any given position of a sequenced protein. The resulting characters often had more than two states, and could have as many as 20 states. Multi-state characters continued to be relevant through the 1980s as DNA sequencing became more available. Here nucleotide bases represent character states.

A character consisting of just its character states, with no hypothesized state proximities or direction of evolutionary change, is called a qualitative taxonomic character (QTC). Without a CST for its states, a QTC does not represent an explicit hypothesis about speciation events on the true phylogenetic tree. However, we can hypothesize that there are speciation events on the true phylogenetic tree that produce a CST with the same states as a QTC. In this way, a QTC represents a hypothesis about speciation events; it is weaker than the hypothesis associated with a CST and asserts nothing about direction of change. Two QTCs are potentially compatible if there are two compatible CSTs with the same states respectively. Those CSTs realize that potential. A character state is convex if the unique path of (ancestor, immediate descendant) pairs connecting any two species in that state contains only species in that state. This mathematical term generalizes a convex lens inside of which any two points determine a line that lies entirely within the lens. A true QTC will have states that are "convex" on the diagram of the true ancestor relation. If two QTCs are potentially compatible, then there exists some ancestor relation (not necessarily true) for which all the states of both are convex. Thus, all true QTCs are necessarily potentially compatible with each other. Often, when working within a context of only QTCs, potentially compatible QTCs are called simply compatible.

In the mid 1970s, Fitch (1975), Sneath et al. (1975) and Estabrook and Landrum (1975) independently suggested essentially the same way to test whether two QTCs were potentially compatible, i.e., there existed compatible CSTs with their respective states. Estabrook and McMorris (1977) mathematically proved the validity of the test of Estabrook and Landrum (1975). To make the test, EUs are entered in the cells of a matrix: states of the first character label the rows, and states of the second character label the columns; each EU is placed in the cell whose row and column labels are the states to which it belongs, as shown in Fig. 11. Moving only from one occupied cell to another in a straight line horizontally or vertically but never retracing a path already taken, if you can return to an occupied cell you have already visited then the two characters are incompatible, as for I and III, otherwise the two characters are potentially compatible, as for I and II. In Fig. 11, possible moves are shown with dashed lines. To discover two CSTs that realize potential compatibility, first connect with horizontal or vertical lines any cell or cell group not yet connected to other cells in any way that does not close a loop, and then designate as primitive any cell in the connected network. Thus, for QTCs I and II in Fig. 11, you could connect cell (IG IIT) containing (d, f) to the empty cell (IC IIT) on the path between cell (IC IIA) containing (b) and cell (IC IIC) containing (e, g); then designate cell (IC IIT) primitive. The resulting CST for I has state C primitive with states A and

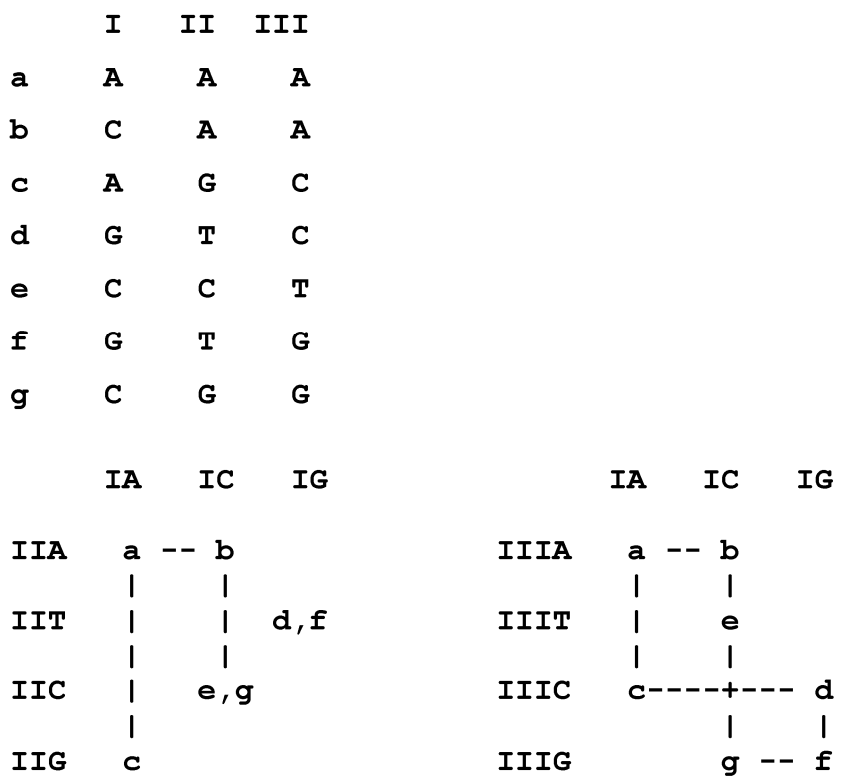

Fig. 11. Illustration of a test for the potential compatibility of two qualitative taxonomic characters, using seven species and three sites. A closed loop in the coincidence matrix, shown lower right, indicates that the pair of sites is not potentially compatible.

G separately derived from it. The resulting CST for II has state T primitive with states C and A separately derived from it and state G derived from state A.

Some characters are inappropriate no matter what their compatibility relations are with other characters, for example, random changes within populations not ultimately associated with speciation events, as discussed at the beginning of this essay. Characters based on such changes are likely to be incompatible with more appropriate characters. Another cause of inappropriate characters is mistaken homology, i.e., comparing non-comparable structures. A structure in one species is homologous to a structure in another species if each structure evolved from the same structure in the most recent common ancestor of those two species. This concept of homology is idealized because virtually never can we make observations to tell for sure whether these conditions have been met. We have to guess based on things we can observe. This concept, together with some ideas about how to guess whether structures are homologous, is very old (Owen, 1848; Lankester, 1870). Operational approaches to estimating homologies were discussed by Inglis (1966), Jardine (1967, 1969), and more recently Estabrook (1997, Figs. 21–24). Characters inappropriate for these reasons are likely to be incompatible with true characters. Incompatibilities with such inappropriate characters are best resolved by simply eliminating such inappropriate characters from further consideration, if more plausible homologies cannot be estimated.

Estimating homology is essential before protein amino acid, or DNA base, sequence data can be used as characters for any kind of analysis (including and especially by computers). An estimate of homology takes the form of sequence alignment. Sometimes short sub-sequences that are identical in all taxa occur in about the same place in the full sequence. These are useful to aid in estimating homology of positions (alignment), but once sequences are aligned with their help, these positions become useless for estimating relationships because they do not vary. Variable sites can be used to estimate relationships provided that variation accurately reflects relationship. But, especially if evolution has created gaps in aligned sequences, variable sites are more difficult to align.

Needleman and Wunsch (1970) were one of the first to suggest an operational procedure for alignment, and increasingly sophisticated criteria and algorithms to estimate homology among DNA base sequences have been devised and discussed over the past few decades by Waterman (1984), Thompson et al. (1994), Day and McMorris (1994) and Kumar et al. (2004), to name a few. The possibility of errors in homology still remains for characters based on aligned DNA base sequences. Discovering and eliminating sites whose alignment is questionable should improve estimates of relationship.

By the late 1970s, all the basic concepts for testing the compatibility of hypotheses of speciation events on a phylogenetic tree, based on comparative observations of species in a group under study, had been developed. Contemporary reviews are available from McMorris (1975), Estabrook (1978, 1984), Cartmill (1981), LeQuesne (1982), Meacham and Estabrook (1985), and more recently Xu (1994) described for the Chinese speaking world compatibility for the special case of characters with unbranched CSTs.

5 Resolution of incompatibility to estimate an ancestor relation

Construed as they are here, the more useful and reliable characters are based on observable changes associated with speciation events, in which a descendant species becomes different in some respects from its ancestral species. Characters based on random and often reversible changes that occur within species, especially if they occur within several less closely related species, are less likely to compatibly reflect much about the history of the speciation process. Thus, as we attempt to reason with the patterns of compatibility among groups of more reliable characters, what we can estimate is an ancestor relation, commonly expressed as a directed diagram with an arrow from any ancestor to each of its immediate descendants, or as an undirected diagram that can represent an estimate of the ancestor relation after its most primitive place (root) is estimated. In particular, we cannot directly estimate with characters, construed as we have, the branching pattern of phylogenetic lines. As we have seen from Figs. 2, 3, the same branching pattern of phyletic lines can give rise to strikingly different ancestor relations, depending where the speciation events take place. With this in mind, we examine some of the early approaches to using compatibility of pairs of characters to estimate the ancestor relation.

Early workers proposed three basic approaches. (1) Make considerations of the development, adaptive functions, parasites, diseases, biogeography, natural history, etc. of the species under study to modify (if possible) one or both CSTs in any incompatible pair to resolve their incompatibility, until enough CSTs could be added together to produce an estimate of the ancestor relation for the species under study. This process has been called "reciprocal illumination" because when some natural, biological factors suggest ways to resolve some incompatibilities, the relationships suggested by the sum of now more compatible characters invite consideration of other natural, biological factors with which to resolve other incompatibilities. CSTs whose incompatibilities could not be resolved by this process were set aside, as less reliable or problematic. (2) Leave characters as originally constructed, but apply some operational (often quantitative) criterion to choose one character (or a compatible group) to make an initial partial estimate. Then, within the context of subsets of S that are convex on this partial estimate, apply the criterion again, until the ancestor relation is resolved to the satisfaction of the investigator. This approach is operationally possible because of the mathematical fact that if two CSTs or two QTCs are compatible in the context of S, then they will remain compatible in the context of any subset of S; but some pairs of CSTs or QTCs incompatible in the context of S may become compatible in the context of only a subset of S. (3) Sub-divide character states without consideration of the development, form, adaptive functions, parasites, diseases, biogeography, natural history, etc. of the group under study, to create new characters with more, smaller states, that are all mutually compatible. This approach is operationally possible because of the mathematical fact that if two CSTs or two QTCs are compatible, there is always a way to subdivide any state (with two or more species) of either to produce CSTs or QTCs that are still compatible; and for any two incompatible CSTs or QTCs, it is always possible to sub-divide states enough to create two new compatible characters. In fact, there are in general very many ways to do this, especially if S is large. For this reason, this approach imposes the additional criterion that as few as possible new character states should be created to resolve all the incompatibilities among the characters.

Among practitioners of the first approach, one of the most influential in his time was Hennig (1966). This book is a translation into English by D.D. Davis and R. Zangerl of an unpublished MS composed by

Hennig shortly before that date as a major revision of a less well known book he had published 15 years earlier. In his preface, Zangerl himself warns us of the linguistic difficulties of making such a translation, especially while the original German revision remains unavailable to most scholars. Indeed, many of the terms in Hennig (1966) are taken from evolutionary biologists writing in English, where their meanings have been well understood for decades, but given different meanings, either by Hennig or by his translator in an attempt to translate Hennig's German. Unfortunately, this resulted in some serious misunderstandings during the 1970s and 1980s, which are only now beginning to be resolved. Mayr (1974) and Sokal (1975) discuss some of these issues in more detail.

Making allowances for the flagrant misuse of established terms by Hennig (1966), pages 120 and 121 of this book clearly describe a test for the compatibility of two 2-state CSTs; incompatible CSTs are there called "incongruent". Much of the rest of the book is devoted to techniques of "reciprocal illumination" to consider the development, form, adaptive functions, parasites, diseases, biogeography, natural history, etc. of the group under study to modify one or both of the pairs of incompatible CSTs to resolve incompatibilities. Examples of applications of these techniques are shown as branching patterns of phylogenetic lines, with speciation events associated with character state changes marked on them. Although for its time, this concept was enlightened, I suspect that Hennig (1966), similar to almost everyone else at that time, did not realize that branching patterns of phyletic lines could not be explicitly estimated with character state trees, only ancestor relations. This went on to confuse Hennig's followers (of which there were many) over the next three decades, and ramifications of this confusion are still with us today.

Many others who did not consider themselves followers of Hennig also considered natural, biological criteria to restructure CSTs to provide a more consistent estimates, with authors typically publishing only the completely compatible CSTs and the resulting ancestor relation (or inappropriately, phylogenetic tree). Such results appear internally very consistent, but often, specifically how they were achieved remained unspecified. Good examples of natural criteria for estimating CSTs are given by Marx and Rabb (1972), who discuss criteria for structuring and modifying CSTs in this spirit, applying them in explicit detail to structure 50 CSTs for the morphological characters of snakes. Other examples are given by DeMarco et al. (1985) presented in Table 2 of Estabrook (2001), Stein et al. (1984), Gardner and LaDuke (1979), and more recently Strasser and Delson (1987) and Chen (1994). Few investigators use this approach today, in part because molecular data have come to dominate as the basis for estimating relationships, and it is not yet clear how to apply considerations from the natural world explicitly to restructure QTCs arising from molecular data. This approach may become more useful again as macromolecular data, e.g., chromosomal rearrangements or other large scale genetic changes, become more widely implicated in the estimation of ancestor relations.

The second approach uses an operational criterion to select some of the CSTs or QTCs to use compatibly to make a (possibly only partial) estimate of the ancestor relation. LeQuesne (1969) considered 2-state CSTs, and described a test for their compatibility. He called true CSTs "uniquely derived"; an apt term because it reminds us that if a character is true then the observable quality shared by all the species in any character state evolved only once when the most recent common ancestor of all the species in that state evolved. LeQuesne (1974) evaluates all possible 2-state undirected CSTs with criteria related to the number of other CSTs with which they are compatible, chooses one of these as a first division in an hierarchical classification, and then iterates the process to resolve a classification. It is not clear how this classification relates to an estimate of the ancestor relation. In the case of CSTs, directed or undirected, Estabrook et al. (1976b) proved that if all pairs of CSTs in a collection of CSTs are compatible, then the entire collection is compatible, i.e., there are ancestor relations (namely all refinements of their sum) that are refinements of every CST in the collection. Thus, a maximal collection of pairwise compatible CSTs could be chosen as the basis for a first (usually only partial) estimate of the ancestor relation. Discovering such a collection is equivalent to discovering the maximal cliques in an undirected graph, a computationally difficult (NP complete) problem as S becomes large. Bron and Lerbosch (1973) published an algorithm to discover the largest collections (cliques) of pairwise compatible CSTs. In 1976, Kent Fiala used this algorithm, among others, to write the computer program CLINCH (CLadistic INference by Compatibility of Hypothesised CSTs) used by Estabrook et al. (1977) to estimate an ancestor relation based on the largest collection of pairwise compatible characters. LaDuke and Crawford (1979), Duncan (1980), Varadarajan and Gilmartin (1983) and Crins (1990) have

also used CLINCH. Warnow (1993, 1994) discussed efficient ways to implement this criterion. There may be two or more largest (in numbers of CSTs) collections of pairwise compatible characters, often with many CSTs in common. Voss and Voss (1983) used the intersection of the two largest cliques to estimate the ancestor relation. Fitch (1984) suggested choosing the maximal clique whose CSTs were compatible with the most other CSTs. Estabrook and Anderson (1978) chose the single CST compatible with the most other CSTs as the first partially resolved estimate of the ancestor relation; subsequent analyses of two subsets of S, convex on this partially resolved estimate, (so-called secondary analyses) resolved the estimate. Strauch (1978) makes extensive use of secondary analyses, and Strauch (1984) explains in more detail some of the tactics of secondary analysis. As CLINCH became more widely used, subsequent versions incorporated these and other criteria; Fiala (1984) documents its sixth version.

Many investigators were not comfortable hypothesizing CSTs. The observable states of a QTC seemed to have more objective reality than a CST, which includes a hypothesis about how those states evolved from one another. A collection of QTCs for a study collection S of species (or other evolutionary units) can be tested for potential compatibility and maximal groups of pairwise potentially compatible QTCs discovered, but as we have pointed out above, there may be no possible estimate of the ancestor relation on which all the states of the QTCs in such a group are convex, i.e., these QTCs may not be group-wise potentially compatible. This poses an interesting problem whose mathematical analog has been studied by Gavril (1974) and McMorris et al. (1994). Related to this, Benham et al. (1995) generalized the concept of characters and their compatibility.

Of course, for a group of pairwise compatible QTCs there may be an ancestor relation on which all or most of them have convex states. To look for it, start with two QTCs, make their matrix of intersecting states and connect them as illustrated in Fig. 11. Then choose any occupied cell as primitive, and direct evolutionary change from it along the lines connecting the occupied cells, to make a CST of their sum. Use other QTCs in the group to refine (if possible) this CST. Boulter et al. (1979) in their study of amino acid sequences of plastocyanin from flowering plants were among the first to apply compatibility of QTCs to a major study. Estabrook (1991) and Camacho et al. (1997) provide later examples.

Especially with the advent of molecular data sets, it became unclear how to hypothesize CSTs, or how to take the first approach to resolve incompatibilities in consideration of other natural data. Especially in data sets with a large number of more distantly related EUs, the second approach applied to CSTs or to the potential compatibility of QTCs often resulted in typically only a small fraction of the data participating in estimates of the ancestor relation. In molecular data sets, high levels of incompatibility may result from a larger fraction of molecular data evolved as random changes not associated with particular speciation events. With other forms of data, high levels of incompatibility may result from adaptive syndromes being selected repeatedly on different phyletic lines as populations of less related species were subject to similar selection pressures. In either case, especially with larger data sets, the first two approaches were difficult for many investigators to apply; an automatic incompatibility resolving criterion was desired.

The third approach is to use an automatic incompatibility resolving procedure: subdivide some of the states of the CSTs a minimum number of times so that all CSTs become mutually compatible. This criterion, suggested by Camin and Sokal (1965), has been called "parsimony" because it minimizes the number of times that an ad hoc character state change had to be hypothesized to eliminate logical incompatibility among the CSTs. The parsimony criterion can be easily modified to apply to QTCs (subdivide states a minimum number of times to make them all convex on an ancestor relation). This made parsimony especially attractive to workers who wanted to avoid hypothesizing CSTs. Discovering which ancestor relations were parsimonious in the context of a collection of CSTs is a mathematically difficult problem. Early workers took one of two basic approaches: (1) Study this problem mathematically in an effort to create algorithms, or (2) devise heuristics that might make plausible guesses. Estabrook (1968) was one of the first to address this problem mathematically. Nastansky et al. (1974) continued his mathematical approach to derive more powerful results. However, this mathematical approach proved computationally impractical except for small data sets; heuristic approaches, typically based on swapping branches to look for more parsimonious trees, proved to be more practical. Computer programs to implement these parsimony heuristics were among a collection of several, called PHYLIP, written by Felsenstein in the 1970s. Easy to use and readily accessible, PHYLIP was revised periodically for the next two decades. Its version 3.5 was published by Felsenstein (1993). Soon

parsimony algorithms themselves became the subject of mathematical study; Hendy and Penny (1982) established branch and bound criteria, and this mathematical study has continued (Argawala et al., 1995, Ganapathy et al., 2003). Swofford (1991) produced a powerful set of programs to implement parsimony (and other criteria), which has been intensely maintained, widely distributed, and remains one of the principal technologies in use (Swofford, 2003).

When parsimony divides character states to resolve incompatibilities, the original CSTs are converted to new CSTs, typically with more states. Commonly in larger data sets, there are many different ways to subdivide a minimum number of character states to resolve incompatibilities. These ways result in different collections of new CSTs. The same ancestor relation may refine all new CSTs in each of these different collections, or some may be refined by one ancestor relation, others by another, etc. Sometimes parsimony produces a very large number of collections of new CSTs that resolve incompatibilities equally parsimoniously, and suggest many, often quite disparate ancestor relations.

Typically, parsimony heuristics evaluate only "fully resolved" ancestor relations, in which none of the species (or other EUs) in S have ever served as an ancestor of any other, and no ancestral species has a representative in S. The true ancestor relation is never "fully resolved" whenever S contains EUs that have served as ancestors for others. This is commonly the case in paleontological studies; for examples see Smith (1994). True ancestor relations that are not fully resolved may also be the case with many studies of extant taxa; for example Gates (1982) describes *Banisteriopsis campestris*, a weedy shrub spread over the aluminum-rich, laterite soils of the central savannah of Brazil, and several other species of that genus that occur uniquely on several geographically isolated quartzite outcrops each not more than a few tens of square kilometers in size. These species seem to have evolved directly from *B. campestris* when populations became isolated on these outcrops. If this were the case, then the true ancestor relation would be unresolved; its diagram would be fan-shaped with *B. campestris* primitive and the other species descended directly from it.

Some variations of parsimony do not weight all subdivisions of characters states equally. In one variation, after a character state has been subdivided once, further subdivisions are weighted less or not at all. In another variation, after any state of a character has been subdivided once, subdivisions of that state or other states in that character are weighted less, or not weighted at all, in which case (if the heuristic process gets the right answer) the heuristic will discover an ancestor relation for which the fewest number of characters need to have any of their states subdivided; such an ancestor relation is a refinement of every character in a largest maximal clique of compatible characters. In this way we can see that there is a continuum of criteria from strict parsimony to estimating an ancestor relation that refines all the characters in a largest clique of compatible characters. Because the popular PHYLIP package contains a program CLIQUE that finds an ancestor relation that requires subdivision of states in a minimum number of characters, many workers have used PHYLIP (and not CLINCH) to discover an ancestor relation that is a refinement of a largest clique of compatible characters. A recent example is given by Gupta and Sneath (2007) who used PHYLIP to discover largest maximal cliques of compatible 2-state characters in a very large data set of DNA sequences comprising thousands of sites for 24 species representing a wide diversity of proteobacteria. The five major groups identified by this compatibility method were the same as those identified by more computationally intense and mathematically sophisticated recent methods, and the ancestor relations among these groups were all very similar. The authors concluded from their results that compatibility analysis is a useful new tool. This conclusion is clearly wrong in one respect; compatibility analysis used in this way is an old tool.

Through the 1960s and '70s, several other methods for estimating the ancestor relation from QTCs, but not directly related to compatibility concepts, were proposed. They are mostly outside the scope of this review, but Felsenstein (2004) provides an excellent review of some of them.

6 Probability of compatibility

Some mid 20th century investigators of evolutionary history, especially those working below the species level in human genetics, such as Edwards and Cavelli-Sforza (1964), construed changes in gene frequencies, DNA base pairs, or other indications of evolution as random processes, not stably associated with speciation events. Because populations within the same species (especially human) are generally less genetically isolated than distinct species, tree branching patterns may be inappropriate representations of the historical evolutionary relationships among

populations within a single species (Baum & Estabrook, 1978; Ward et al., 1991). Nevertheless, population geneticists used probability to model these random processes as occurring along tree branching patterns connecting populations. Once a particular random process is hypothesized, maximum likelihood methods can be used to estimate a (usually) undirected tree branching pattern. Random models of character state change and maximum likelihood estimation of tree branching patterns of phyletic lines became more widely used to estimate evolutionary relationships among species as well. Felsenstein (1983) pointed out that in some modeled cases, especially those including a distantly related evolutionary unit, maximum likelihood estimates differed from those made by compatibility or parsimony, and would continue to do so even if more data were generated. From this he concluded that in such cases the estimates made by compatibility or parsimony were misleading. However, if many changes are random, but some changes are stably associated with speciation events, maximum likelihood may get the "wrong" answer and compatibility or parsimony do better, especially if random changes are somehow recognized and removed before a compatibility or parsimony estimate is made. Qiu and Estabrook (2008) observed that when the apparently more random sites were removed parsimony estimates became more stable with higher branch support while maximum likelihood estimates produced a variety of branching patterns with weaker branch support. Recently, Kolaczkowski and Thornton (2004) have clearly demonstrated this effect with simulation.

Probability concepts have been applied to compatibility to try to recognize, so to eliminate or underweight, characters that seemed to be more random. LeQuesne (1972), presented a formula to calculate the probability that two 2-state characters would be potentially compatible at random; he used this concept to rank characters in order of "merit" to be considered uniquely derived (true), or to select characters for further consideration. Meacham (1981, 1984b) defined clearly the concept of "at random" implicit in LeQuesne (1972) and described mathematically how to calculate the probability that two (or a group) of CSTs (with any number of states) would be compatible at random. It became clear that some kinds of CSTs were more likely than others to be compatible at random with others; CSTs with many large advanced states were less likely to be compatible at random. This suggested another criterion for choosing a collection of compatible characters for the initial estimate (or subsequent secondary refinements) of the ancestor relation; instead of choosing the largest clique, which might have many CSTs likely to be compatible at random, choose the clique of CSTs least likely to be compatible at random. Meacham's computer program COMPROB enabled investigators to implement this criterion accurately to identify maximal cliques least likely to be cliques at random.

Especially in data sets with a large number of taxa, the number of characters (CSTs or QTCs) in the largest (or least likely) clique was often a small fraction of the total number of characters. An immediate consequence of this is that in many cases most characters are false, because true characters always belong to the same clique (but perhaps not always the largest one). Flagrantly false characters, whose states would have to be subdivided many times to become convex on the true ancestor relation, might be as likely to be compatible with other characters as a character to whose states EUs were assigned at random. For a given QTC in a collection describing the EUs in S, the probabilities with which it would be compatible at random with each other QTC in the collection could be summed to give the number of other QTCs with which it would be expected to be compatible at random. This could be compared with the number with which it was actually compatible; if this number were not substantially more than would be expected at random, then the character could be set aside as indistinguishable from random. Remaining characters could be dealt with in any of the three approaches described above.

Meacham (1994) construed the number of other characters with which a given character is compatible as a random variable, and undertook to estimate its distribution under the hypothesis that the given character was random. Because exact algorithms in the style of Meacham (1981) become complicated and computationally intractable for QTCs, or for CSTs for data sets with a large number of EUs, Meacham (1994) used simulation to make close approximations to compatibility related probabilities. To estimate the probability that a given character (QTC or CST) would be compatible with at least as many other characters as it actually was, under the hypothesis that it was a random character, he chose with equal probability another character from among those with the same number of EUs in each state as the given character, and counted the number of other characters in the data set with which it was compatible, noting whether this was at least as many as the given character. This was repeated 1000 (or more) times. The fraction of these random characters that were

compatible with at least as many other characters as the given character estimates this probability. Meacham (1994) called this probability *Cf*, the Frequency of Compatibility Attainment. He applied this to evaluate the 53 morphological characters of angiosperms published by Donoghue and Doyle (1989), ranking them by their *Cf*, the probability that a random character would be compatible with at least as many other characters as observed. By this criterion, 25 characters were significantly non-random at the $Cf < 0.05$ level. He then used parsimony to reconcile incompatibilities among only the top ranked characters, which produced an estimate similar to, but less ambiguous than, that of Donoghue and Doyle (1989). Camacho et al. (1997) used the character evaluation method of Meacham (1994) in conjunction with potential compatibility of QTCs to estimate relationships among species of a subterranean genus of Crustacea. Qiu and Estabrook (2008) apply the character evaluation method of Meacham (1994) to a large molecular data set to choose the less apparently random sites for further analysis with bootstrap parsimony and maximum likelihood using PAUP. Pisani (2004) applied the criterion of Meacham (1994) to evaluate 866 DNA sites for 47 species chosen to represent the diversity of Arthropoda. There were 172 sites with $Cf > 0.5$, i.e., can not reject at $p = 0.05$ the hypothesis that the character is random based on the number of other characters with which it is observed to be compatible. These 172 sites were removed and those remaining were subject to maximum likelihood and neighbor joining analyses to estimate branching pattern of phyletic lines. Pisani (2004) suggests that removing characters with $Cf > 0.5$ may reduce the effects of long branch attraction; he also observed that when characters with Cf much lower than 0.5 were also removed, results began to deteriorate. However Qiu and Estabrook (2008) observed increased clarity of parsimony estimates of relationships among key groups of angiosperms when all characters with $Cf > 0.2$ were removed. Day et al. (1998) used the number of compatible pairs of characters in a whole data set as a random variable under the hypothesis that all the characters in the data set were random in the sense of Meacham (1994). They analyzed 102 published data sets, of which 12 had fewer compatible pairs than would be expected at random. In general, they observed that inclusion of outgroups increased the probability that compatibilities levels are random, sometimes substantially so.

Salisbury (1999) suggested another way to use probability to evaluate cliques of compatible characters, which he termed "strongest evidence"; where Meacham (1994) calculates the probability that a clique of compatible characters would be mutually compatible if they were all random characters, i.e., that there is some ancestor relation with which they are all compatible, Salisbury (1999) calculates the probability that the characters in a clique of compatible characters would be independently each compatible at random with the ancestor relation that they jointly determine, which is a stricter criterion with a lower probability. Salisbury implemented strongest evidence compatibility in a computer program, called SECANT, which also contains most of the functions of CLINCH.

7 Other applications

Compatibility concepts have been applied to several other areas. Ringe et al. (2002) recognize the parallel between the historical development of related languages and the evolution of species to structure features of related languages as characters, and find maximal cliques of compatible characters as a basis for estimating their historical relationships.

O'Keefe and Wagner (2001) used character compatibility as a basis for testing hypotheses of character independence. An advantage of their approach is that it does not require estimating an ancestor relation, which may itself assume character independence. Characters with similar patterns of compatibility with the other characters in the data set are candidates for dependence; simulation under the hypothesis of independence estimates significance levels. This is a second example, along with Meacham (1994), of using compatibility concepts to evaluate characters in a phylogenetic context without needing to estimate an ancestor relation.

One way to resolve incompatibilities among characters is to hypothesize explicitly reticulate events (such as hybridization) in the evolution of the EUs under study. Corti et al. (1986) hypothesized hybridization events, severely constrained by the meiotic failure of certain chromosomal rearrangements, to resolve incompatibilities. von Haeseler and Churchill (1993), Bandelt (1994) and Jakobsen and Easteal (1996) used compatibility concepts to describe phylogenetic networks more generally.

Nelson and Platnik (1981) explored the possibility, suggested earlier by Hennig (1966) and many others, that there may be a link between the phylogenetic relationships among species, and the history of the processes that have resulted in their occupancy of

particular geographic areas. Estabrook (1985) used compatibility concepts to demonstrate the potential tenuousness of this relationship. However, character states convex on an undirected estimate of the ancestor relation may suggest a historical relationship among geographic areas if species evolved either by rare dispersal into geographically disjunct areas, or as a consequence of those areas becoming disjunct through the establishment of barriers to dispersal. Estabrook (2001) gives an example of this, and considers available natural history data to judge which of these two processes seems most likely. Craw (1988) has interpreted compatibility relations among "characters", construed as the presence or absence of species in distinct geographic areas, to suggest historical patterns. The simultaneous evolution of species with the "evolution" of the isolated areas they occupy is a fascinating phenomenon, especially because conceptually it is analogous to parasitic species evolving simultaneously with their hosts, or gene duplications (like speciation for gene lineages) evolving simultaneously with the species of whose genomes they are a part. It is possible that an estimate of ancestor relation for the areas (hosts, species) is incompatible with the ancestor relation implied for the area (hosts, species) by an estimate of the ancestor relation for the species (parasites, gene lineages). How to resolve such incompatibilities is a difficult problem, which was effectively addressed by Page (1994, 1996). Much progress to understand it more thoroughly has been made by a number of workers since, but its discussion is beyond the scope of this review.

Rock strata containing fossils have been used to estimate the interval of time from the evolution to the extinction of species (or other higher taxon). These estimates could place so-called stratigraphic constraints on estimates of the ancestor relation. For the past two decades or more, the question of whether such stratigraphic constraints should be imposed on estimates made with comparative characters has been hotly debated. Cladists, who believe that ancestral species have ceased to exist, generally oppose stratigraphic constraints, as do others who believe that such stratigraphic estimates are generally too inaccurate to impose constraints. Huelsenbeck (1991) presents some of these issues in more detail. Estabrook and McMorris (2006) examined the consequences of stratigraphic constraints on the mathematical foundations of character compatibility analysis established by Estabrook et al. (1976a, b) and discovered that these consequences can be quite severe.

Methods to estimate ancestor relations, such as parsimony, often produce many different estimates, so the question arises how to combine them into a single estimate. Investigators may also want to combine estimates based on different data sources, e.g., different genes, or different members of the same gene family, in which case the several estimates may not involve exactly the same species. Many ways to combine estimates have been suggested; the description of all is beyond the scope of this review, but one way uses compatibility. An estimate is represented by several 2-state CSTs; to each ancestor corresponds a 2-state CST with its advanced state containing that ancestor and all its descendants. These CSTs are all compatible and their sum reconstructs that estimate of ancestor relation. All the CSTs of all the estimates to be combined are analyzed for compatibility and a maximal clique of compatible CSTs are added to produce the combined estimate. Rodrigo (1996) argues convincingly for the appropriateness of this method, and suggests a way to combine estimates that do not each involve every one of the species in S. His suggestion becomes computationally unwieldy when many estimates each missing many species are to be combined. This problem is avoided by using a parsimony hill climbing heuristic with the optimality criterion that the number of characters for which *any* state needs to be subdivided is minimized. Ross and Rodrigo (2004) make a thorough assessment of this compatibility method, and Wilkinson et al. (2005) test this method, together with 13 others, to demonstrate its consistency and stability.

Some forms of data distinguish a group of some of the species under study; such a group might be a plausible candidate for a character state convex on the ancestor relation, but there may be no compelling evidence to imagine that the remaining species will also make a convex state. An example of this is the presence of an endonuclease binding site which may have evolved only once but may have been subsequently lost in descendant lineages; in this way the group of species that do not posses the site would not make a group convex on the ancestor relation. McMorris (1977), Meacham (1983) and Templeton et al. (1992) have grappled with such so-called partial binary characters.

Instead of estimating an entire ancestor relation, investigators may be primarily concerned to test the credibility of specific hypotheses of monophyletic in the context of a group of species under study for which character data are available, especially when those hypotheses are incompatible with each other. Archie (1989) was among the first to suggest an

approach to this, and a variety of other approaches, whose description is beyond the scope of this review, have since been proposed by several investigators. One of those approaches uses compatibility concepts. For each competing hypothesis, the probability that it would be compatible at random with each character with which it is compatible is transformed to its negative logarithm and summed. This provides a test statistic whose realized significance is estimated by simulating the hypothesis that the monophyletic group in question could have been any collection of EUs the same size with equal probability. Competing hypotheses of monophyly can be compared using the significance with which each rejects the hypothesis that they seem random in the context of the data. A computer program, MEAWILK, (Frohlich & Estabrook, 2000) performs this analysis. Qiu et al. (2006) used MEAWILK to test competing hypotheses about the deepest divergences in land plants.

8 Compatibility, monophyly and higher taxa

Not all systematists of the mid 20th century, in attempting to formulate well defined concepts with which to study evolutionary relationships among species, embraced the concept that distinct species evolve from their ancestral species, which may continue to have an independent existence through time until their extinction. Cladists did not recognize an ancestor relation among species, but considered an ancestral species to be "identical with all the species that have arisen from it" (Hennig, 1966, Fig. 18). Hennig (1966) clearly describes incompatible CSTs and discusses a wide variety of ways to consider additional data to resolve incompatibilities, but contrary to the claims of his followers, this work does not advocate automatic criteria, such as parsimony, for resolving incompatibilities. Fundamental to Hennig's explanations of the methods he describes are the two concepts: synapomorphy (shared derived character states), and monophyly (group contains all the descendants of its most recent common ancestor). Both of these concepts depend on the direction of evolutionary change. The relevance of these concepts to estimating branching pattern of phylogenetic lines was defended at the time by some very respected systematists, such as Donoghue and Maddison (1986). As I have shown above, Meacham (1984a) pointed out that branching patterns can be estimated independent of estimates of direction of evolutionary change, and because this involves fewer a priori hypotheses, many argue that they should be. Virtually all modern "computer" methods to estimate phylogenetic relationships, such as parsimony, maximum likelihood, neighbor joining, etc., make their estimates with no consideration of the direction of evolutionary change. Direction is estimated after branching pattern has been estimated, often by choosing a primitive place in the undirected branching pattern. The dangers of including a distantly related species as a surrogate for this place were pointed out by Baum and Estabrook (1996) and evidenced by the results of Day et al. (1998). If direction of evolutionary change plays no role in the estimation of branching pattern of phyletic lines, then neither do synapomorphies nor monophyly. However convex characters states, a fundamental concept of compatibility, remain convex no matter how an undirected branching pattern may be directed.

Many cladists became strong advocates of monophyly (defined above) as the only legitimate basis for the recognition of higher taxa, partly, or perhaps even mostly, as a consequence of denying the continued existence of ancestral species. Monophyly can be lost with the evolution of new species, even when there is no change in the ancestral taxon, e.g., reptiles may have been a monophyletic taxon until birds evolved from them, but then (according to cladists) the existence of birds invalidated reptiles as a natural taxon, even though reptiles themselves did not change. The requirement to constrain the recognition of higher taxa by requiring strict monophyly with respect to an estimate of the ancestor relation is much more stringent than the constraint long recognized by evolutionary systematists, who require only that a system of higher taxa be compatible (as a CST) with a generally accepted estimate of the ancestor relation. An early but very clear example is provided by Hall (1928) Fig. 27, which is the diagram of an ancestor relation of the species in a (convex but not monophyletic) section of the genus *Haplopappus* (Poaceae). Some more recent, but equally clear, examples are given by Michener (1977) and Varadarajan and Gilmartin (1983). Estabrook (1986) presented a computer program CONPHEN to evaluate possible classifications into higher taxa constrained to be compatible with a given ancestor relation. Meacham and Duncan (1987) point out several problems with monophyly as an overly strict criterion for higher taxa. When this monophyly criterion results in the destruction of a convex, phenotypically distinct and long recognized genus or family, it makes scientific nomenclature unstable and interferes with access to older name-based publications. Another methodological problem is becoming apparent; when

workers are trying to clarify relationships near the root of a phylogenetic tree, they are often reluctant to remove members of well established monophyletic groups far from the root, even though the inclusion of these less related taxa are likely to confound their analysis. Sometimes this reluctance stems from a belief that it is illegitimate to analyze a convex, but not monophyletic, group. More progress will be made when we learn to focus taxon sampling near the area in question and avoid the inclusion of distantly related taxa, even if they are descendants of some of the ancestors in question.

References

Archie JW. 1989. A randomization test for phylogenetic information in systematic data. Systematic Zoology 38: 239–252.

Argawala R, Fernandez-Baca D, Slutzki G. 1995. Fast algorithms for inferring evolutionary trees. Journal of Computational Biology 2: 397–407.

Bandelt H-J. 1994. Phylogenetic networks. Verhandlungen des Naturwissenschaftlichen Vereins in Hamburg 34: 51–71.

Baum BR, Estabrook GF. 1978. Application of compatibility analysis in numerical cladistics at the infraspecific level. Canadian Journal of Botany 56: 1130–1135.

Baum BR, Estabrook GF. 1996. Impact of outgroup inclusion on estimates by parsimony of undirected branching pattern of ingroup phyletic lines. Taxon 45: 243–257.

Benham C, Kannan S, Paterson M, Warnow TJ. 1995. Hen's teeth, and whale's feet: generalized characters and their compatibility. Journal of Computational Biology 2: 515–525.

Boulter D, Peacock D, Guise A, Gleaves TJ, Estabrook GF. 1979. Relationships between the partial amino acid sequences of plastocyanin from members of ten families of flowering plants. Phytochemistry 18: 603–608.

Briggs D, Walters SM. 1969. Plant variation and evolution. New York: McGraw-Hill.

Bron C, Lerbosch J. 1973. Algorithm 457: Finding all cliques of an undirected graph. Communications of the Association for Computing Machines 16: 575–577.

Camacho AI, Bello E, Estabrook GF. 1997. A statistical approach to the evaluation of characters to estimate evolutionary history among the species of the aquatic subterranean genus, *Iberobathynella* (Crustacea, Snycardia). Biological Journal of the Linnean Society 60: 221–241.

Camin JH, Sokal RR. 1965. A method for deducing branching sequences in phylogeny. Evolution 19: 311–326.

Cartmill M. 1981. Hypothesis testing and phylogenetic reconstruction. Zeitschrift fur Zoologische Systematik Evolutionsforschung 19: 73–96.

Chen Z-D. 1994. Phylogeny and phytogeography of the Betulaceae. Acta Phytotaxonomica Sinica 32: 1–31.

Corti M, Capanna E, Estabrook GF. 1986. Microevolutionary sequences in house mouse chromosomal speciation. Systematic Zoology 38: 163–173.

Craw R. 1988. Continuing the synthesis between panbiogeography, phylogenetic systematics and geology as illustrated by empirical studies on the biogeography of New Zealand and the Chatham Islands. Systematic Zoology 37: 291–310.

Crins WJ. 1990. Phylogenetic considerations below the sectional level in *Carex*. Canadian Journal of Botany 68: 1433–1440.

Day WHE, Estabrook GF, McMorris FR. 1998. Measuring the phylogenetic randomness of biological data sets. Systematic Biology 47: 604–616.

Day WHE, McMorris FR. 1994. Alignment, comparison, and consensus of molecular sequences. In: Diday E, Lechevallier Y, Schader M, Bertrand P, Burtschy B eds. New approaches in classification and data analysis. Berlin: Sprenger Verlag. 327–346.

DeMarco GA, Altieri A, Estabrook GF. 1985. Relazioni evolutivi e biogeografiche dei popolamenti ad areale disguinte di *Genista ephedroides* DC. Biogeographia 11: 115–131.

Donoghue MJ, Doyle JA. 1989. Phylogenetic analysis of angiosperms and the relationships of the Hamamelidae. In: Crane PR, Blackmore S eds. Evolution, systematics and fossil history of the Hamamelidae. Systematics Association Special Volume 1. Oxford: Clarendon Press. 17–45.

Donoghue MJ, Maddison WP. 1986. Polarity assessment in phylogenetic systematics: A response to Meacham. Taxon 35: 534–545.

Duncan TO. 1980. A cladistic analysis of the *Ranunculus hispidus* complex. Taxon 29: 441–454.

Edwards AWF, Cavelli-Sforza LL. 1964. Reconstruction of evolutionary trees. In: Heywood VH, McNeil J eds. Phenetic and phylogenetic classification. London: Systematics Association Publication #6. 67–76.

Estabrook GF. 1968. A general solution in partial orders for the Camin-Sokal model in phylogeny. Journal of Theoretical Biology 21: 421–438.

Estabrook GF. 1972. Cladistic methodology: a discussion of the theoretical basis for the induction of evolutionary history. Annual Reviews of Ecology and Systematics 3: 427–456.

Estabrook GF. 1977. Does common equal primitive? Systematic Botany 2: 36–42.

Estabrook GF. 1978. Some concepts for the estimation of evolutionary relationships in systematic botany. Systematic Botany 3: 146–158.

Estabrook GF. 1984. Phylogenetic trees and character state trees. In: Duncan T, Stuessy TF eds. Cladistics. perspectives on the reconstruction of evolutionary history. New York: Columbia University Press. 135–151.

Estabrook GF. 1985. Historical biogeography: the relationship between phylogeny and geography. Biogeographia 11: 37–47.

Estabrook GF. 1986. Evolutionary classification using convex phenetics. Systematic Zoology 35: 560–570.

Estabrook GF. 1991. The use of distance measures, such as from DNA-DNA hybridization, in phylogenetic reconstruction and in convex phenetics. Bollettino di Zoologia 58: 299–305.

Estabrook GF. 1997. Ancestor-descendant relations and incompatible data: Motivation for research in discrete

mathematics. In: Mirkin B, McMorris FR, Roberts FS, Rzhetsky A eds. Mathematical hierarchies and biology. DIMACS Applied Mathematics Series. Vol. 37. Providence, RI, USA: American Mathematical Society. 1–28.

Estabrook GF. 2001. Vicariance or dispersal: the use of natural history data to test competing hypotheses of disjunction on the Tyrrhenian coast. Journal of Biogeography 28: 95–103.

Estabrook GF, Anderson WR. 1978. An estimate of phylogenetic relationships within the genus *Crusea* (Rubiaceae) using character compatibility analysis. Systematic Botany 3: 179–196.

Estabrook GF, Johnson CS, McMorris FR. 1975. An idealized concept of the true cladistic character. Mathematical Biosciences 23: 263–272.

Estabrook GF, Johnson CS, McMorris FR. 1976a. An algebraic analysis of cladistic characters. Discrete Mathematics 16: 141–147.

Estabrook GF, Johnson CS, McMorris FR. 1976b. A mathematical foundation for the analysis of cladistic character compatibility. Mathematical Biosciences 29: 181–187.

Estabrook GF, Landrum LR. 1975. A simple test for the possible simultaneous evolutionary divergence of two amino acid positions. Taxon 24: 609–613.

Estabrook GF, McMorris FR. 1977. When are two qualitative taxonomic characters compatible? Journal of Mathematical Biology 4: 195–200.

Estabrook GF, McMorris FR. 1980. When is one estimate of evolutionary relationships a refinement of another? Journal of Mathematical Biology 10: 367–373.

Estabrook GF, McMorris FR. 2006. The compatibility of stratigraphic and comparative constraints on estimates of ancestor-descendant relations. Systematics and Biodiversity 4: 9–17.

Estabrook GF, Meacham CA. 1979. How to determine the compatibility of undirected character state trees. Mathematical BioSciences 46: 251–256.

Estabrook GF, Strauch JG Jr, Fiala KL. 1977. An application of compatibility analysis to the Blackiths' data on Orthopteroid insects. Systematic Zoology 26: 269–276.

Felsenstein J. 1983. Parsimony in systematics: biological and statistical issues. Annual Review of Ecology Systematics 14: 313–333.

Felsenstein J. 1993. PHYLIP, version 3.5c. Seatle, WA: University of Washington.

Felsenstein J. 2004. Inferring phylogenies. Sunderland, MA: Sinauer.

Fiala KL. 1984. CLINCH: Cladistic inference by compatibility of hypothesized character state trees. Version 6.1. FORTRAN program and User's Document. Ann Arbor, MI: Museum of Zoology, University of Michigan.

Fitch WM. 1975. Towards finding the tree of maximum parsimony. In: Estabrook GF ed. Proceedings of the Eighth International Conference on Numerical Taxonomy. San Francisco: W. H. Freeman. 189–230.

Fitch WM. 1984. Cladistic and other methods: problems, pitfalls and potentials. In: Duncan T, Stuessy TF eds. Cladistics: perspectives on the reconstruction of evolutionary history. New York: Columbia University Press. 221–232.

Frohlich MW, Estabrook GF. 2000. Wilkinson support calculated with exact probabilities: An example using Floricaula/LEAFY amino acid sequences that compares three hypotheses involving gene gain/loss in seed plants. Molecular Biology and Evolution 17: 1914–1925.

Ganapathy G, Ramachandran V, Warnow TJ. 2003. Better hill-climbing searches for parsimony. In: Benson G, Page R eds. Algorithms in Bioinformatics, Proceedings of the Third International Workshop, WABI 2003, Budapest, Hungary, September, 2003. Berlin: Springer. 245–258.

Gardener RC, LaDuke JC. 1979. Phyletic and cladistic relationships in *Lipochaeta* (Compositae). Systematic Botany 3: 197–207.

Gates B. 1982. *Banisteriopsis*, *Diplopterys* (Malpighiaceae). Flora Neotropica, monograph 30. New York: The New York Botanical Garden.

Gavril F. 1974. The intersection graphs of subtrees in trees are exactly the chordal graphs. Journal of Combinatorial Theory, series B 16: 47–56.

Grant V. 1963. The origins of adaptations. New York: Columbia University Press.

Gupta RS, Sneath PHA. 2007. Application of the character compatibility approach to generalized molecular sequence data: Branching order of the proteobacterial subdivisions. Journal of Molecular Evolution 64: 90–100.

Hall HM. 1928. The genus *Haplopappus*. Washington, DC.: Carnegie Institution.

Hendy MD, Penny D. 1982. Branch and bound algorithms to determine evolutionary trees. Mathematical Biosciences 60: 133–142.

Hennig W. 1966. Phylogenetic systematics (Translated from German manuscript by Davis DD and Zangerl R). Chicago: University of Illinois Press. 263.

Huelsenbeck JP. 1991. When are fossils better than extant taxa in phylogenetic analysis? Systematic Zoology 40: 458–469.

Inglis WG. 1966. The observational basis of homology. Systematic Zoology 15: 219–228.

Jakobsen IB, Easteal S. 1996. A program for calculating and displaying compatibility matrices as an aid in determining reticulate evolution in molecular sequences. Computer Applications in the Biosciences 12: 291–295.

Janowitz MF. 1984. On the semi-lattice of the weak orders of a set. Mathematical Social Sciences 8: 229–239.

Jardine N. 1967. The concept of homology in biology. British Journal of the Philosophy of Science 18: 135–139.

Jardine N. 1969. The observational and theoretical components of homology. Biological Journal of the Linnean Society 1: 327–361.

Kolaczkowski B, Thornton JW. 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. Nature 431: 980–984.

Kumar S, Tamura K, Nei M, 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5: 150–163.

LaDuke J, Crawford DJ. 1979. Character compatibility and phyletic relationships in several closely related species of *Chenopodium* of the western United States. Taxon 28: 307–314.

Lankester ER. 1870. On the use of the term homology in

modern zoology. Annals of the Magazine of Natural History 6: 34–43.
LeQuesne WJ. 1969. A method of selection of characters in numerical taxonomy. Systematic Zoology 18: 201–205.
LeQuesne WJ. 1972. Further studies based on the uniquely derived character concept. Systematic Zoology. 21: 281–288.
LeQuesne WJ. 1974. Uniquely derived character concept and its cladistic application. Systematic Zoology 23: 513–517.
LeQuesne WJ. 1982. Compatibility analysis and its applications. Zoological Journal of the Linnean Society 74: 267–275.
Marx H, Rabb GB. 1972. Phyletic analysis of fifty characters of advanced snakes. Fieldiana, Zoology 63: 1–321.
Mayr E. 1963. Animal species and evolution. Cambridge, MA: Belknap Press.
Mayr E. 1969. Principles of systematic zoology. New York: McGraw Hill.
Mayr E. 1974. Cladistic analysis or cladistic classification? Zeitschrift fur Zoologische Systematik und Evolutionsforschung 12: 94–128
McMorris FR. 1975. Compatibility criteria for cladistic and qualitative taxonomic characters. In: Estabrook GF ed. Proceedings of the Eighth International Conference on Numerical Taxonomy. San Francisco: Freeman. 189–230.
McMorris FR. 1977. On the compatibility of binary qualitative taxonomic characters. Bulletin of Mathematical Biology 39: 133–138.
McMorris FR, Warnow TJ, Wimmer T. 1994. Triangulating vertex colored graphs. SIAM Journal of Discrete Mathematics 7: 296–306.
McMorris FR, Zaslavsky T. 1981. The number of cladistic characters. Mathematical BioSciences 54: 3–10.
Meacham CA. 1981. A probability measure for character compatibility. Mathematical Biosciences 57: 1–18.
Meacham CA. 1983. Theoretical and computational considerations of the compatibility of qualitative taxonomic characters. In: Felsenstein J ed. Numerical taxonomy. NATO Advanced Study Institute. Series G1. Berlin: Springer-Verlag. 304–314.
Meacham CA. 1984a. The role of hypothesized direction of characters in the estimation of evolutionary history. Taxon 33: 26–38.
Meacham CA. 1984b. Evaluating characters by character compatibility analysis. In: Duncan TO, Stuessy TF eds. Cladistics: perspectives on the reconstruction of evolutionary history. New York: Columbia University Press. 152–165.
Meacham CA. 1994. Phylogenetic relationships at the basal radiation of Angiosperms: further study by probability of compatibility. Systematic Botany 19: 506–522.
Meacham CA, Duncan TO. 1987. The necessity of convex groups in biological classification. Systematic Botany 12: 78–90.
Meacham CA, Estabrook GF. 1985. Compatibility methods in systematics. Annual Review of Ecology and Systematics 16: 431–446.
Michener CD. 1977. Discordant evolution and the classification of Allodapine bees. Systematic Zoology 26: 32–56.
Nastansky L, Selkow SM, Stewart NF. 1974. An improved solution to the generalized Camin-Sokal model for numerical cladistics. Journal of Theoretical Biology 48: 413–424.
Needleman SB, Wunsch CD. 1970. A general method applicable to the search for similarities in the amino acids sequences of two proteins. Journal of Molecular Biology 48: 443–453.
Nelson G, Platnik N. 1981. Systematics and biogeography: Cladistics and vicariance. New York: Columbia University Press.
O'Keefe FR, Wagner PJ. 2001. Inferring and testing cladistic character dependence using character compatibility. Systematic Biology 50: 657–675.
Otte D, Endler JA eds. 1989. Speciation and its consequences. Sunderland, MA: Sinauer.
Owen R. 1848. On the archetype and homologies of the vertebrate skeleton. London: John van Voorst.
Page RDM. 1994. Maps between trees and cladistic analyses of historical associations of genes, organisms, and areas. Systematic Biology 43: 58–77.
Page RDM. 1996. Temporal congruence revisited: comparison of mitochondrial DNA sequence divergence in cospeciating pocket gophers and their chewing lice. Systematic Biology 45: 151–167.
Pisani D. 2004. Identifying and removing fast-evolving sites using compatibility analysis: An example from the Arthropoda. Systematic Biology 53: 978–989.
Qiu Y-L, Estabrook GF. 2008. Inference of phylogenetic relationships among key angiosperm lineages using a compatibility method on a molecular data set. Journal of Systematics and Evolution 46: 130–141.
Qiu Y-L, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J, Estabrook GF, Hendry TA, Taylor DW, Testa CM, Ambros M, Crandall-Stotler B, Duff RJ, Stech M, Frey W, Quandt D, Davis CC. 2006. The deepest divergences in land plants inferred from phylogenomic evidence. Proceedings of the National Academy of Sciences USA 103: 15511–15516 and supporting text.
Ringe D, Warnow TJ, Taylor A. 2002. Indo-European and computational cladistics. Transactions of the Philological Society 100: 59–129.
Rodrigo AG. 1996. On combining cladograms. Taxon 45: 267–274.
Ross HA, Rodrigo AG. 2004. Assessment of matrix representation with compatibility in supertree construction. In: Bininda-Emonds ORP ed. Phylogenetic supertrees: Combining information to reveal the tree of life. Berlin: Springer. 35–64.
Salisbury BA. 1999. Strongest evidence in compatibility: Clique and tree evaluations using apparent phylogenetic signal. Taxon 48: 755–766.
Simpson GG. 1961. Principles of animal taxonomy. New York: Columbia University Press.
Skelton PW. 1993. Adaptive radiation: definition and diagnostic test. In: Lees RD, Edwards D eds. Evolutionary patterns and processes. Linnean Society Symposium Series 14. London: Academic Press. 46–58.
Smith AB. 1994. Systematics and the fossil record. Oxford: Blackwell.
Sneath PHA, Sackin MJ, Ambler RP. 1975. Detecting evolutionary incompatibilities from protein sequences.

Systematic Zoology 24: 311–332.
Sokal RR. 1975. Mayr on cladism — and his critics. Systematic Zoology 24: 257–262.
Stebbins GL. 1950. Variation and evolution in plants. New York: Columbia University Press.
Stein WE Jr, Wight DC, Beck CB. 1984. Possible alternatives for the origin of Sphenopsida. Systematic Botany 9: 102–118.
Strasser E, Delson E. 1987. Cladistic analysis of Cercopithecid relationships. Journal of Human Evolution 16: 81–99.
Strauch JG Jr. 1978. The phylogeny of the Charadriiformes (Aves): a new estimate using the method of character compatibility analysis. Transactions of the Zoological Society of London 34: 263–345.
Strauch JG Jr. 1984. The use of homoplastic characters in compatibility analysis. Systematic Zoology 33: 167–177.
Swofford DL. 1991. PAUP: Phylogenetic analysis using parsimony. version 3.1. Champaign, IL: Illinois Natural History Survey.
Swofford DL. 2003. PAUP*4.0: Phylogenetic analysis using parsimony. Sunderland, MA: Sinauer.
Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132: 619–633.
Thompson JD, Higgins DJ, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties, and weight matrix choice. Nucleic Acid Research 22: 4673–4680.
Varadarajan GS, Gilmartin AJ. 1983. Phenetic and cladistic analyses of North American *Chloris* (Poaceae). Taxon 32: 380–386.
von Haeseler A, Churchill GA. 1993. Network models for sequence evolution. Journal of Molecular Evolution 37: 77–85.
Voss NA, Voss RS. 1983. Phylogenetic relationships in the cephalopod family Cranchiidae (Oegopsida). Malacologia 23: 397–426.
Ward RH, Frazer BL, Dew-Jager K, Paabo S. 1991. Extensive mitochondrial diversity within a single Amerindian tribe. Proceedings of the National Academy of Sciences USA 88: 8720–8724.
Warnow TJ. 1993. Constructing phylogenetic trees efficiently using compatibility criteria. New Zealand Journal of Botany 31: 239–247.
Warnow TJ. 1994. Tree compatibility and inferring evolutionary history. Journal of Algorithms 16: 388–407.
Waterman MS. 1984. Efficient sequence alignment algorithms. Journal of Theoretical Biology 1108: 333–337.
Wilkinson M, Cotton JA, Creevey C, Eulenstein O, Harris SR, LaPointe F-J, Levasseur C, McInery JO, Pisani D, Thorley JL. 2005. The shape of supertrees to come: tree shape related properties of fourteen supertree methods. Systematic Biology 54: 419–431.
Wilson EO. 1965. A consistency test for phylogenies based on contemporaneous species. Systematic Zoology 14: 214–220.
Xu K-X. 1994. The compatibility concept and compatibility analysis methods in Cladistics. Acta Phytotaxonomica Sinica 32: 380–388.

Availability of computer programs

Many computer programs related to the concepts discussed here are available from the authors of the publications cited. In addition, many are available to be freely downloaded from my own web site: www-Personal.umich.edu/~gfe/. These include:

* SECANT by Ben Salisbury. Accepts character data, implements strongest evidence compatibility, and also incorporates many of the functions of CLINCH by Kent Fiala.
* MEACHAM accepts sequence data and applies the character evaluation criterion of Meacham (1994), then selects sites with *Cf* less than a specified threshold, and writes them to a file for subsequent analysis.
* CONPHEN accepts an ancestor relation and distance matrix, discovers close groups convex on that ancestor relation.
* MEAWILK accepts sequence data and hypotheses of monophyly, evaluates hypotheses using compatibility criteria, as well as some others.
* ADQUARC4 accepts area x taxon presence/absence data, evaluates compatibility of “taxa” , finds cliques and reveals the “ancestor” relation they determine.
* STRATCOM accepts character data and a stratigraphic range for each taxon, evaluates character compatibility in a stratigraphic context.