A Method for Deducing Branching Sequences in Phylogeny
Author(s): Joseph H. Camin and Robert R. Sokal
Source: *Evolution,* Vol. 19, No. 3 (Sep., 1965), pp. 311-326
Published by: Society for the Study of Evolution
Stable URL: http://www.jstor.org/stable/2406441
Accessed: 26/01/2015 14:37

A METHOD FOR DEDUCING BRANCHING SEQUENCES IN PHYLOGENY[1, 2]

Joseph H. Camin and Robert R. Sokal[3]

Department of Entomology, The University of Kansas, Lawrence

With the advent of relatively objective classifications, such as the phenetic classifications produced by the operational techniques of numerical taxonomy (Sokal and Sneath, 1963), it was inevitable that biologists would wonder what phylogenetic conclusions could be drawn from them and with what reliability. If these phenetic taxonomies did not reflect all of the elements of phyletics (Sokal and Camin, 1965), could techniques be devised for deducing the latter? For example, could operational methods be devised for deducing the cladistic relationships among taxa, so that, given the same initial information, different investigators would obtain the same results? By cladistic relationships we mean the evolutionary branching sequences among taxonomic units without regard to phenetic similarities among them or to an absolute time scale.

There is no question that phylogenies could probably be reconstructed without error for any taxonomic group if complete fossil sequences for that group were available. However, can cladistic reconstructions be carried out with any degree of reliability if only characters of recent forms are considered? Several recent studies have also considered this question from different points of view (Doolittle and Blombäck, 1964; Edwards and Cavalli-Sforza, 1964; Simpson, 1963; Throckmorton, 1965; Wilson, 1965).

Since 1962 a group at the Entomology Department of The University of Kansas has been examining the principles by which phylogenies are constructed conventionally, as well as the relation between the principles and practices of phylogeny and those of taxonomy, both orthodox and numerical. In addition to the authors, the group includes G. W. Byers and C. D. Michener and several graduate students. The study was based on a group of imaginary animals possessing a number of morphological characteristics generated by one of us (JHC) according to rules known so far only to him, but which are believed to be consistent with what is generally known of transspecific evolution. Genetic continuity was accomplished by tracing the drawings of the animals from sheet to sheet, permitting the preservation of all characters except for such modifications as were desired. Although the study is still in progress, it has already led to an empirical method which we believe capable of deducing probable cladistics from the characters of existing organisms.

Detailed studies of subsets of the assemblage of hypothetical animals by orthodox phylogenetic methodology resulted in differing, but internally consistent, cladistic schemes, the choice among which was not apparent to those uninitiated in the true phylogeny. Comparison by Camin of these various schemes with the "truth" led him to the observation that those trees which most closely resembled the true cladistics

invariably required for their construction the least number of postulated evolutionary steps for the characters studied. Subsequently we examined the possibility of reconstructing cladistics by the principle of evolutionary parsimony. The following technique seems capable of doing this.

TECHNIQUE

The technique requires a conventional data matrix as used for numerical taxonomy (Table 1). The columns of this matrix represent the *operational taxonomic units* (OTU's), which can stand for any taxonomic unit from individual through species up to higher categories. The rows of the matrix represent characters scored into different *character states* as qualitative or quantitative subdivisions of each character, differing among the OTU's.

The basic assumptions underlying character coding are fundamental to the entire technique and must be carefully examined.

1. We assume that characters can be expressed in discrete states differing among at least some of the OTU's of the study.

2. The character states can be arrayed in some logical order. If the characters are quantitative (*e.g.*, counts of bristles, segments, or leaves, or increments of size) linear order for the states is easily accomplished. Qualitative characters (shapes, colors, etc.) may require some ingenuity as well as some arbitrariness in coding states. If a logical order cannot be found for a qualitative character, the states may have to be recoded as several two-state characters. For details of this procedure see Sokal and Sneath (1963, p. 74ff).

From our knowledge of evolutionary processes the following three asumptions are not valid for all cases, although they are probably true for the majority of characters and taxa. Thus, they are only working assumptions which, as we shall show, can themselves be tested by the technique and may be relaxed in certain instances.

3. It is assumed that we have knowledge of the direction of the evolutionary trends within characters, and therefore the character states can be arrayed in a presumed evolutionary sequence from primitive (ancestral) to derived. In the linear sequence of character states, 1, 2, 3, 4, 5, 6, the presumed evolutionary sequence may be from 1 to 6, or from 6 to 1, or from 3 in two directions (toward 1 as well as toward 6). The primitive character state in these three examples would be 1, 6, or 3, respectively. For convenience, the primitive state is coded zero, derived states

positively or negatively, as required (see Table 1).

4. The ancestral state arose only once in the taxa at hand. Wilson (1965) has called such character states *unique*. Derived character states may, however, have arisen repeatedly in different branches of the group studied.

5. Evolution is irreversible, *i.e.*, a line, having attained a derived character state, cannot return to a character state ancestral to the derived one.

Under these assumptions the minimum number of evolutionary steps necessary to evolve c states of a character is $c - 1$. The number of character states and the minimum number of evolutionary steps necessary for each character are shown in Table 1. The most parsimonious cladistic dendrograms for characters 5 and 6 are shown at the bottoms of Tables 2A and 2B. We suggest the term *cladogram* to distinguish a cladistic dendrogram from a phenetic one which might be called a *phenogram*.[4] The principles of our approach are best illustrated during a preliminary step in the computations, the calculation of the so-called compatibility matrix.

All characters in the study are fitted to the pattern of the cladogram of each character. By this we mean that we compute the number of evolutionary steps required to arrive at the correct character states for all the OTU's in the study via the various cladistic patterns provided. The pathways of a pattern cladogram are unidirectional from the base of the dendrogram to the tips. Since changes in character states are irreversible (assumption 5) an evolutionary step in a character affects all pathways beyond that step. Provisionally any time several branches come off a stem at the same place a single evolutionary step suffices to produce the same change in any or all branches. Evolutionary steps increasing character state codes are shown graphically as short lines crossing the stems, while those decreasing the character state codes are shown as X-marks across the stems. They are marked with the num-

---

[4] Ernst Mayr (1965) has independently suggested the same terms with identical meanings. Arnett (1963) has used *phenogram* for a profile-type summary of characters.

TABLE 1. *Data matrix.*

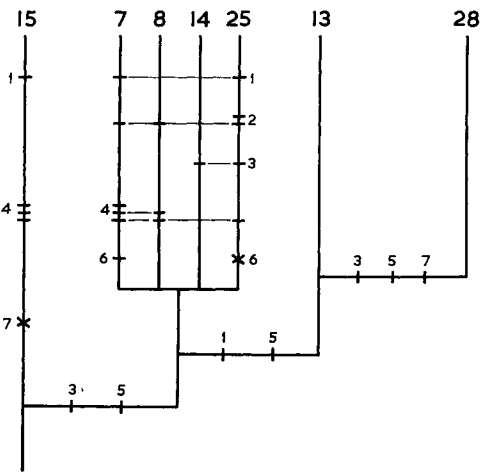|					t OTU's				Character states (c)	Minimum steps ($c-1$)
		7	8	13	14	15	25	28		
	1.	1	0	1	0	1	1	1	2	1
	2.	1	1	0	0	0	2	0	3	2
	3.	1	1	1	2	0	2	2	3	2
n Characters	4.	3	2	0	0	3	1	0	4	3
	5.	1	1	2	1	0	1	3	4	3
	6.	1	0	0	0	0	-1	0	3	2
	7.	0	0	0	0	-1	0	1	3	2
										15

FIG. 1. Pattern cladogram of character 5. This cladogram follows the pattern shown in Table 2A. Numbers at the tips of the branches are code numbers of OTU's. The evolutionary steps for each character are marked on the branches. Evolutionary steps increasing character state codes are shown as lines across branches, those decreasing character state codes as X-marks. The number of the character represented is next to each mark. Single steps affecting several members of a cluster are diagrammed as heavy cross lines which indicate changes for the stems concerned, connected by thin lines to indicate that these comprise a single step.

ber of the character which they represent (Fig. 1). Each mark represents a character state increment or decrement of 1 only. To indicate a change from state 3 to 5, for example, two marks must be shown on a stem. The state code at the base of the dendrogram is assumed to be 0 (primitive) for all characters. It is convenient to set up the data in a *pattern table* as shown in Table 2. This arranges the OTU's by the state codes of the characters that form the bases for the patterns (*e.g.*, characters 5 and 6). The character state codes inside the table are the columns of the data matrix in Table 1 rearranged in the new order. The cladogram at the bottom of the table may be helpful, although we have not generally found it necessary for our computations.

Fitting character 1 to the pattern of character 5 (Table 2A) we find that it possesses state 1 in OTU 15. A mark cannot be placed at the base of the cladogram because OTU's 8 and 14 are coded 0 and would be changed by such an evolutionary step. A positive mark is therefore placed on the branch to OTU 15. A single step suffices to change OTU's 7 and 25 to character state code 1. The mark for this step cannot be placed on the stem leading to the cluster (7–25) as it would raise all members of the cluster to state 1. We therefore place the marks over the cluster as shown in Fig. 1, heavy cross lines indicating changes for the stems concerned and thin lines connecting the heavy ones to indicate that these comprise a single step. Such clusters may then be resolved later in the

procedure. Finally a mark before the point where OTU's 13 and 28 separate will raise these to code 1. These marks and those for fitting the other characters to pattern 5 are shown graphically in Fig. 1.

Character 2 needs a single positive mark to raise OTU's 7, 8, and 25 to state code 1 and a second for OTU 25 only, to raise it to code 2. Character 3 has a step prior to cluster (7–25) raising all OTU's other than 15 to code 1. A single positive mark raises OTU's 14 and 25 to code 2 and another raises OTU 28 to code 2. In character 4, because OTU's 14, 13, and 28 are at state 0, three separate positive steps are required to change OTU 15 to code 3. One step will raise OTU's 7, 8, and 25 to state 1, a second will raise OTU's 7 and 8 to state 2, and a third will raise OTU 7 to state 3.

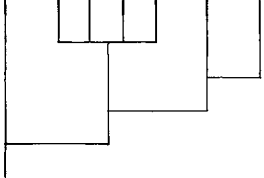Character 5, being the pattern character,
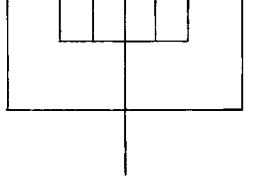
TABLE 2. *Sample pattern tables.*

To demonstrate the method of computing the number of necessary evolutionary steps and the measure of compatibility between the various characters and patterns constructed according to other characters. The cladograms at the bottoms of the tables are the most parsimonious pathways for evolution of characters 5 and 6, respectively, whose states are underlined in the tables.

A. Pattern table of character no. 5.

	OTU's							Total steps	Extra steps
	15	7	8	14	25	13	28		
1.	1	1	0	0	1	1	1	3	2
2.	0	1	1	0	2	0	0	2	0
3.	0	1	1	2	2	1	2	3	1
4.	3	3	2	0	1	0	0	6	3
5.	0	1	1	1	1	2	3	3	X
6.	0	1	0	0	-1	0	0	2	0
7.	-1	0	0	0	0	0	1	2	0
								21	6

(Characters label on left side)

B. Pattern table of character no. 6.

	OTU's							Total steps	Extra steps
	25	8	13	14	15	28	7		
1.	1	0	1	0	1	1	1	2	1
2.	2	1	0	0	0	0	1	3	1
3.	2	1	1	2	0	2	1	4	2
4.	1	2	0	0	3	0	3	6	3
5.	1	1	2	1	0	3	1	4	1
6.	-1	0	0	0	0	0	1	2	X
7.	0	0	0	0	-1	1	0	2	0
								23	8

(Characters label on left side)

requires only the minimum number of steps as indicated. In character 6 a single positive and a single negative step change OTU's 7 and 25 to codes 1 and −1, respectively. Finally, character 7 needs a positive and a negative step to change OTU's 28 and 15 similarly. The total number of steps required for any one character to fit a given pattern is listed to the right of the pattern table (Table 2A). After a little practice the number of steps required can be written down simply by inspection of the pattern table.

Some additional problems in evaluating the number of evolutionary steps are illustrated by Table 2B which is a pattern table based on character 6, showing a V-shaped evolutionary trend. Evolutionary steps now have to be calculated in both directions from the pivotal stem. Parsimony may result by following the provisional rule on branches arising from one place on the stem. For example, a single step may turn the left and right arms or the pivotal stem and one of the arms in the same direction. For character 1 in Table 2B a single step turns OTU's 25 and 7 to character state 1 and a second step turns OTU's 13, 15, and 28 to state 1, leaving OTU's 8 and 14 at state 0. Thus two evolutionary steps are required to fit character 1 to the pattern of character 6.

Subtracting the minimum number of steps $(c - 1)$ for each of the characters (see Table 1) from the total number of steps yields the extra number of steps necessary to fit a given character to a pattern. These values are shown in the last column of each pattern table (Table 2). Whenever this value is 0, the character is compatible with the pattern provided. The number of extra steps is a measure of the incompatibility of the character to any given pattern. A check on the computation is provided since the sum of the total necessary steps minus the sum of the extra steps must give the sum of the minimum number of steps, $\Sigma(c - 1)$. When all (n) characters have been fitted to the n patterns (one for each character) the numbers of extra steps for each pattern are assembled in a *compatibility matrix* (Table 3). The diagonal elements of the compatibility matrix are zeros since, obviously, every character is compatible with

its own pattern. They are indicated by X's and excluded from the computation. Rows and columns of this matrix are summed in two ways. The number of zeros is counted (and recorded as "compatibilities") and the numbers of extra steps are summed.

The compatibility matrix provides information of two kinds. It shows which characters provide "good" patterns and thus are relatively close to the presumed correct cladogram. Such characters would have high column compatibilities, *i.e.*, have a large number of characters compatible with their pattern and consequently few extra steps in their column. The compatibility matrix also supplies us with information about "poor" characters which are those whose patterns have relatively few characters compatible with them and call for a large number of extra steps, and which also fit poorly to most other patterns, showing few compatibilities and requiring many extra steps. Thus in Table 3 the pattern of character 4 has only 2 compatibilities and requires 11 extra steps by the other characters in order to conform to it. In addition, character 4 is incompatible with all other patterns and requires 18 extra steps to fit it to the patterns of the other characters. Such characters may be poor because of miscoding of their character states. The latter can arise from errors in transcription of data (such a case occurred in the analysis of the horses discussed below), or by an incorrect interpretation of the evolutionary trends in character states. Coding a character 0, 1, 2, 3 implies that evolution has proceeded in steps from 0 to 3. If, in fact, evolution proceeded from state 3 to 0, the compatibility matrix would show the miscoded character to be poor as a pattern and in fitting other patterns. The assumption of irreversibility of evolutionary steps may not be true in a specific case. When state 2 of a character arose from state 1 as well as by reversion from state 3, the character will show up as miscoded, if we consider all OTU's exhibiting the operationally homol-

TABLE 3. *Compatibility matrix.*

		Patterns							Compatibilities	Extra steps
		1	2	3	4	5	6	7		
Characters	1.	X	2	2	2	2	1	1	0	10
	2.	1	X	1	2	0	1	0	2	5
	3.	2	2	X	4	1	2	1	0	12
	4.	2	3	4	X	3	3	3	0	18
	5.	1	1	1	3	X	1	1	0	8
	6.	0	0	0	0	0	X	0	6	0
	7.	0	0	0	0	0	0	X	6	0
Compatibilities:		2	2	2	2	3	1	2	14	–
Extra steps:		6	8	8	11	6	8	6	–	53

ogous character state 2 as identical. Thus the method provides a check on some of its assumptions. If we knew which instances of apparent state 2 were really reversions from state 3, we could recode these as state 4 to preserve the linear sequence, which would improve the pattern and fit of this character in the compatibility matrix. It has therefore been our practice to exclude characters which show few compatibilities and large numbers of extra steps in their columns and rows on the assumption that these characters are miscoded. The reconstruction of the cladogram is then carried out without considering these characters which are later fitted separately to the reconstructed cladogram. Frequently reexamination of poor characters and fitting them to the final reconstruction in their own most parsimonious sequence will reveal the source of miscoding and permit their use in subsequent studies.

A number of different approaches to the reconstruction of the cladogram have been developed. None work perfectly so that they directly provide the most parsimonious solution. All methods provide a *procladogram*, which represents a state of considerable parsimony but must be adjusted by inspection or preferably a systematic program of trial and error to change it to the final most parsimonious arrangement. Our first approach fitted all characters to the cladogram of a good pattern as defined above, making adjustments as necessary. The steps necessary to fit

TABLE 4. *The monothetic method for reconstructing cladograms.*

A. Data matrix for group A (Table 1) with characters 6 and 7 recoded and character 4 omitted.

		7	8	13	14	15	25	28	Cycle 1, Step 1
	1	1	0	1	0	1	1	1	OTU's 14 and 15 have
	2	1	1	0	0	0	2	0	6 zeros each. Removal
Characters	3	1	1	1	2	0	2	2	of OTU 14 leaves no
	5	1	1	2	1	0	1	3	"non-zero" rows. Re-
	6+	1	0	0	0	0	0	0	moval of OTU 15 leaves
	6−	0	0	0	0	0	1	0	rows 3 and 5 non-zero
	7+	0	0	0	0	0	0	1	and 7− all-zero. There-
	7−	0	0	0	0	1	0	0	fore, remove OTU 15.
Number of zeros		3	5	5	6	6	3	4	

B. Data matrix A with OTU 15 removed.

		7	8	13	14	25	28	Cycle 1, Step 2
	1	1	0	1	0	1	1	
	2	1	1	0	0	2	0	
Characters	3	1	1	1	2	2	2	Subtract unity
	5	1	1	2	1	1	3	from rows 3
	6+	1	0	0	0	0	0	and 5; delete
	6−	0	0	0	0	1	0	row 7−.
	7+	0	0	0	0	0	1	
	7−	0	0	0	0	0	0	

C. Data matrix B with unity subtracted from rows 3 and 5 and row 7− deleted.

		7	8	13	14	25	28	Cycle 2, Step 1
	1	1	0	1	0	1	1	Recompute number of ze-
	2	1	1	0	0	2	0	ros for remaining OTU's.
Characters	3	0	0	0	1	1	1	OTU's 8 and 14 have 6 ze-
	5	0	0	1	0	0	2	ros each. Removal of either
	6+	1	0	0	0	0	0	8 or 14 leaves no non-zero
	6−	0	0	0	0	1	0	rows. Removal of both 8
	7+	0	0	0	0	0	1	and 14 leaves row 1 non-
Number of zeros		4	6	5	6	3	3	zero. Therefore, remove OTU's 8 and 14 together.

D. Data matrix C with OTU's 8 and 14 removed.

		7	13	25	28	Cycle 2, Step 2
	1	1	1	1	1	
	2	1	0	2	0	Subtract unity from
Characters	3	0	0	1	1	row 1. Row 1 be-
	5	0	1	0	2	comes all-zero, so de-
	6+	1	0	0	0	lete.
	6−	0	0	1	0	
	7+	0	0	0	1	

E. Data matrix D with row 1 deleted.

		7	13	25	28	Cycle 3, Step 1
	2	1	0	2	0	Recompute number of zeros for
Characters	3	0	0	1	1	remaining OTU's. OTU 13 has
	5	0	1	0	2	5 zeros. Removal of OTU 13
	6+	1	0	0	0	leaves no non-zero rows. Simi-
	6−	0	0	1	0	larly for OTU 7 with 4 zeros.
	7+	0	0	0	1	Removal of OTU's 7 and 13
Number of zeros		4	5	3	3	leaves row 3 non-zero. There- fore, remove OTU's 7 and 13.

F. Data matrix E with OTU's 7 and 13 removed

		25	28	Cycle 3, Step 2
	2	2	0	
Characters	3	1	1	Subtract unity from row 3.
	5	0	2	Row 3 becomes all-zero, so
	6+	0	0	delete. OTU's 25 and 28 are
	6−	1	0	a terminal bifurcation.
	7+	0	1	

all characters are marked on the chosen pattern cladogram (Fig. 1). Branches of this basic cladogram are rearranged if this achieves greater parsimony of evolutionary steps. The basic outline of the tree is retained but the provisionally shared steps in cluster (7–25) must be resolved. In small studies, such as this one, it is simplest to make a frequency distribution of steps shared by OTU pairs, triplets, etc., which are subsets of a cluster. Not counting character 4, as explained in the results for Group A, we find that OTU's 8 or 14 have only one shared step with OTU 25 which, however, has 2 steps in common with OTU 7. Since OTU 8 shares a step with both OTU's 7 and 25, OTU 14 is least related to the cluster and becomes its basal branch. The origin of the stem for OTU 14 is at the same point as that for the stem leading to OTU's 13 and 28 because there are no steps common to all of cluster (7–25) that

do not also affect OTU's 13 and 28. The joint step of OTU's 14 and 25 (character 3) cannot be made compatible with such an arrangement and, therefore, must be an example of parallelism. Next to arise from the stem bearing the rest of the cluster is OTU 8 because it shares one step with OTU's 7 and 25. The latter share two steps and represent the terminal branches of the former cluster.

The completed reconstruction can be inspected in Fig. 2. In this cladogram evolutionary steps for characters 8 through 12 are also added. These were not included in the data matrix or the computation, because they occurred in the derived state in only one OTU each. The steps for these characters are compatible with the general cladogram and emphasize that OTU 25 is very specialized.

In larger clusters, frequency distributions of shared steps are more tedious without a digital computer. Another approach consists of a single-linkage cluster analysis (Sokal and Sneath, 1963) of the OTU's based on the number of common evolutionary steps.

The third approach is the monothetic method which works directly from the original data matrix. However, all characters having both positive and negative states must be recoded as two characters as shown in Table 4A for the data matrix of Table 1. Although the method is monothetic in operation, its results are polythetic (Sokal and Sneath, 1963, p. 13). The steps are as follows.

1. Zero states are counted for each OTU (column). This provides some measure of "primitiveness," as OTU's with greater numbers of characters in state "0" should branch off the main trunk of the cladogram near its base.

2. The OTU with the greatest number of zeros is removed and the remaining data matrix is checked for rows (characters) without zeros. (If there are ties for largest number of zeros, the first of the tied OTU's is removed.)

3. If no "non-zero" rows appear, the OTU with the next largest number of zeros is removed (or the second of the tied OTU's is removed) and the previously removed OTU column is placed back into the matrix.
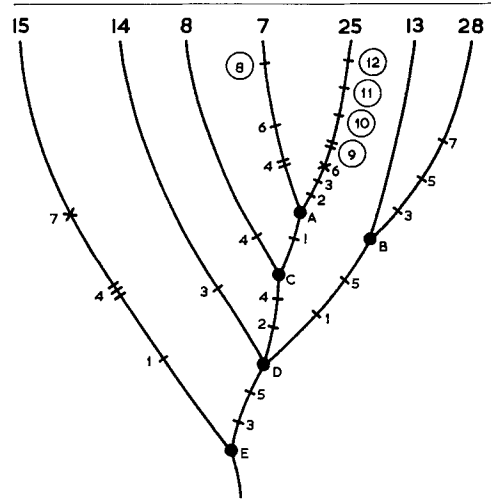
4. If there are still no non-zero rows, then

FIG. 2. Reconstruction of cladogram of Group A of the hypothetical animals, based on the pattern cladogram of Fig. 1. The OTU's and evolutionary steps are indicated as in Fig. 1. Characters 8 through 12 (circled) were not included in the data matrix (Table 1) or Fig. 1 because they occurred in the derived state in only one OTU each. The total number of evolutionary steps, not counting characters 8 through 12, is 23 compared with the minimum number of 15 steps from Table 1. The ancestors of the OTU's at the tips are indicated by black circles and identified by capital letters. Their probable character states can be easily obtained from the cladogram by going from the base to the ancestor. Thus for ancestor C, characters 1 through 7 will be in states 0, 1, 1, 1, 1, 0, 0, respectively.

the first and second OTU's are removed simultaneously. If this still leaves no non-zero rows, the OTU columns are placed back into the matrix and the OTU with the next largest number of zeros is removed. If no non-zero rows appear, the third and first, third and second, or all three OTU's are removed in that order. This process is systematically continued until at least one non-zero row appears (Table 4A).

5. When a non-zero row appears after the removal of one or more OTU's, the OTU or OTU's whose removal from the matrix produced a non-zero row is drawn as a branch from the base of the procladogram (Fig. 3A).

6. Unity (1) is then subtracted from each character state code in each non-zero row (Table 4B). This is repeated, if the row remains non-zero after the subtraction of one. Another branch, the main trunk of the procladogram, is drawn adjacent to the branch bearing the removed OTU or OTU's and evolutionary steps
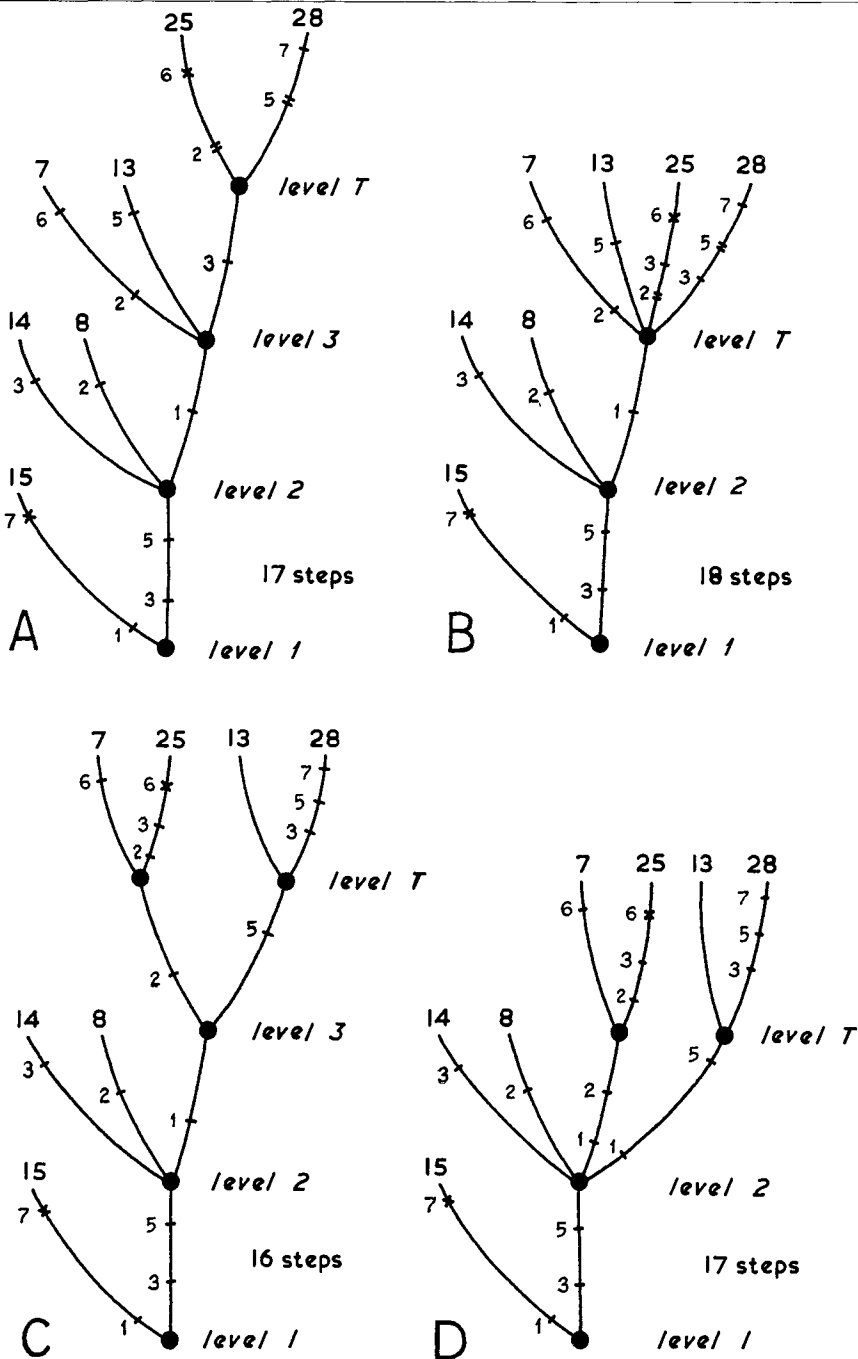
FIG. 3. Steps in the reconstruction of the cladogram of Group A (Fig. 2) by the monothetic method. Symbolism as in earlier figures. Level numbers and T refer to levels of furcation from 1 to terminal level. A. Procladogram resulting from monothetic technique illustrated in Table 4. Total number of evolutionary steps is 17. B. OTU 25 moved down one branching point; 18 evolutionary steps result. C. OTU's 25 and 7 grouped as are OTU's 13 and 28. Achieved parsimony of 16 steps. This is equally

representing the subtraction of one are drawn on that trunk with a number beside it for identification representing the character affected (Fig. 3A). At this time, any rows which are left all-zero may be removed from the matrix and dropped from subsequent consideration. Steps for such characters are placed on the branch leading to the removed OTU's (Fig. 3A).

7. The number of column zeros for the remaining OTU's is recomputed and the process is repeated through a second cycle (Tables 4C and 4D) and continued until the reconstruction of the procladogram is completed (Tables 4E and 4F).

8. When the procladogram (an initial approximation of the most parsimonious solution) is completed, all character states (evolutionary steps) are added to the cladogram and final readjustments for greater parsimony are made (Fig. 3A).

Because all of the methods devised so far yield only good approximations of the most parsimonious solution, it is necessary to test for parsimony and to make adjustments when these are indicated. This can best be practiced with a systematic procedure. First remove all internodes, *i.e.*, segments of stems between furcations, which do not bear any evolutionary steps because there is no reason to assume separate branching points in the absence of intervening evolutionary steps. Next move all common evolutionary steps found on adjacent branches to the base stem of these branches. This practice is parsimonious by making one evolutionary step do the work of two. Finally, and of most consequence, is the trial and error moving of branches which we shall illustrate in Fig. 3. From the procladogram resulting from the monothetic method (Fig. 3A) we try moving one of the terminal branches, the branch carrying OTU 25 or OTU 28, down one furcation to the furcation level 3 indicated in Fig. 3A. This now becomes furcation level T (the terminal level) and OTU's 25, 28, 7, and 13 all emerge from this point

(Fig. 3B). While our original procladogram required 17 evolutionary steps, the new adjustment requires 18 steps, and is thus moving away from the intended direction. However, we can now check members of the cluster 25, 28, 7, and 13, for common steps which can be removed to base stems. We find in Fig. 3C that OTU's 25 and 7 can be placed together with character step 2 made common, and OTU's 28 and 13 can be joined with character step 5 in common. This necessitates parallel steps in OTU's 25 and 28 for character 3, which previously was a common step, but we have now reduced the number of steps for the cladogram to 16. This is the same number required by the true cladogram but Fig. 3C is not the correct solution. Moving the branch which bears OTU's 7 and 25 in Fig. 3C from furcation level 3 to level 2 (Fig. 3D) will result in an equally parsimonious cladogram which is the same as Fig. 2, the correct solution. This illustrates an important point. Different, but equally parsimonious solutions may occur and in order to distinguish between them one must have added information from further characters. When character 4 was laid on the cladogram of Fig. 3C, without correcting it for apparent miscoding, this took an additional eight steps. Further rearrangement following the principles outlined above reduced the added steps for character 4 to only seven and yielded the correct solution. However, when character 4 is fitted and recoded in its most parsimonious sequence to either the cladogram of Fig. 3C or of Fig. 2, it results in only four additional steps and the two solutions remain equally parsimonious. Therefore, in such cases, additional characters must be sought in order to find the most probable solution to

←

parsimonious but not identical to cladogram in Fig. 2. Addition of character 4 to this cladogram and further adjustments for parsimony result in a cladogram identical to Fig. 2. D. Branch bearing OTU's 7 and 25 move down one branching point; 17 evolutionary steps result. If OTU's 8, 7, and 25 are now rearranged so that they share their common step for character 2, the cladogram of Fig. 2 is obtained.

the cladogram and to the recoding of character 4.

We should point out that the small studies reported in this paper are based on very few characters and decisions on alternative cladograms are frequently taken on the saving of a single step. When more characters are studied, such decisions generally are more soundly based. However, even in larger studies equally parsimonious solutions may occur in certain portions of the tree, where the determination of structure depends on very few evolutionary steps.

Three computer programs have been developed by Ronald Bartcher for carrying out the above methods of numerical cladistics. The first program calculates a compatibility matrix, finds the optimal pattern, and then fits the characters to this pattern. A second program carries out the monothetic method of finding a procladogram. The third program parctices parsimony on a cladogram for any given data matrix. The output is shown as an actual cladogram with the evolutionary steps marked in. These programs, called CLADON I, II, and III, respectively, were prepared in FORTRAN IV, for the IBM 7040 with 16K memory, at The University of Kansas Computation Center. As currently written they can handle 30 OTU's and 50 characters. Persons interested in obtaining copies of CLADON for adaptation to their computational equipment are invited to write to the authors. To provide some idea of running time, four of the reconstructions reported in this paper (including one with 14 characters) took approximately one-half minute at a cost of $1.50 by CLADON I, or two minutes ($6.00) by CLADON II including drawing of cladograms. Attempts to improve parsimony by CLADON III (unsuccessful—apparently correct solutions were obtained by either CLADON I or II) took five minutes.

RESULTS

The technique was applied experimentally to data from several hypothetical cases, including Group A, seven OTU's from Camin's imaginary animals, which furnished the illustrative example of the technique section. Data from several groups of real organisms were also analyzed.

Group A

The cladogram of Fig. 2, which was obtained directly from the compatibility matrix (Table 3), proved to be entirely correct. Besides leading to the reconstruction, the compatibility matrix provides additional information of interest. Characters which provide poor patterns usually fit well to many other patterns and are, therefore, usually compatible with the optimal pattern (e.g., characters 2, 6, and 7). Some characters which provide poor patterns may fit only moderately well to other patterns. This usually indicates some lack of parsimony, i.e., parallel evolution for that character (e.g., characters 1 and 3 in Table 3 and Fig. 2). We have already noted character 4 which provides a poor pattern and also fits quite poorly to most other patterns. In view of these considerations, we excluded character 4 from procedures for finding the cladogram. However, as we have seen in the monothetic method, it was necessary to employ character 4 in order to obtain the correct solution. That it is not unduly discordant can be seen from the moderate number of extra steps in rows and columns. Nevertheless, when equally parsimonious solutions occur, it is probably preferable to seek additional new characters in order to resolve such solutions. Examination of the true phylogeny of the OTU's in Group A revealed an error in tracing the forms, which unintentionally produced reversibility in character 4. When character 4 is recoded to fit the cladogram of Fig. 2 in its most parsimonious sequence, reversibility plus parallelism is revealed.

Another use of a reconstructed cladogram is to predict the character states of the ancestral forms at the branching points (see Fig. 2). Because no evolutionary changes have taken place since OTU 28 branched off their common ancestral stem, OTU 13 is identical with the ancestral form B, for the characters under consideration. This relationship presents a method for introducing fossil forms into a study along with recent forms. All OTU's which show no evolutionary steps subsequent to their last point of branching can be considered ancestral to all OTU's derived subsequent to the branch. An analysis of

Group A, including fossils *A* through *E*, again resulted in the correct cladogram.

Two other studies of simulated phylogenies showed that cladograms with little or moderate amounts of parallelism were reconstructed without error. Further work with hypothetical phylogenies led to the following tentative conclusions.

1. If reversible characters greatly outnumber irreversible ones, the most parsimonious tree will probably be incorrect. However, if the reversible characters are not highly correlated with each other, the compatibility matrix will provide criteria for removing them prior to the reconstruction and, if the remaining irreversible characters are numerous enough to give the correct cladogram, the most parsimonious tree will be correct.

2. Parallel or miscoded convergent characters (Sokal and Camin, 1965) will not be detected as such when they outnumber divergent characters and will show as recent divergences in the most parsimonious tree. If divergent characters are more numerous than others, the most parsimonious tree will show all evolutionary steps correctly.

Fossil Horses

W. A. Clemens of the Zoology Department, The University of Kansas, kindly provided us with data on lineages of fossil horses. These lineages are reputedly among the best known in the animal kingdom. He chose species within those genera believed to represent some of the major lines of horse evolution (see Fig. 4), although the actual species are not necessarily in the direct cladistic lines. Characters chosen (Table 5) were among those considered significant by authorities in the field and for which data on all species used in the analysis were available. Table 6 shows the data matrix and the compatibility matrix from which the reconstruction shown in Fig. 4 can be obtained. Presumed ancestral forms are shown circled and by dashed stems.

The reconstruction of equid cladistics is correct according to the studies of Stirton
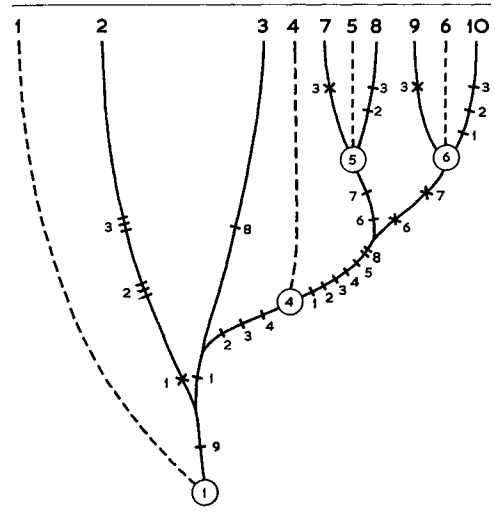

FIG. 4. Reconstructed cladogram of the fossil horses based on data in Table 6. The OTU code numbers and evolutionary steps are indicated as in Figs. 1 and 2. Character 3 has been recoded as explained in the text. Dashed branches represent OTU's ancestral to others in the study. Their code numbers are also shown in circles at points of branching. There are 31 evolutionary steps in this cladogram, compared with a minimum number of 20 (Table 6). OTU numbers represent the following fossil horse species: 1. *Mesohippus barbouri* Schlaikjer, Oligocene, data from Schlaikjer (1932, 1935). 2. *Hypohippus osborni* Gidley, Miocene, data from Gidley (1907) and Osborn (1918). 3. *Archaeohippus blackbergi* (Hay), Miocene; following White (1942), *A. nanus* is regarded as a synonymous species. Generic reference follows Stirton (1940). Data from Simpson (1932), White (1942), and Bader (1956). 4. *Parahippus pristinus* Osborn, Miocene, data from Osborn (1918). 5. *Merychippus* (*Merychippus*) *seversus* (Cope), Miocene, data from Downs (1956, 1961). 6. *Merychippus* (*Protohippus*) *secundus* Osborn, Miocene, data from Osborn (1918). 7. *Nannipus* cf. *minor* (Sellards), Pliocene, data from Lance (1950). 8. *Neohipparion occidentale* (Leidy), Pliocene, data from Gregory (1942). 9. *Calippus placidus* (Leidy), Pliocene, data from Gidley (1906, 1907) and Osborn (1918). 10. *Pliohippus mexicanus* Lance, Pliocene, data from Lance (1950).

(1940) and of Simpson (1951). During the analysis two characters appeared to be miscoded. One of these (character 7) was discovered to have been erroneously transcribed and is shown corrected in Table

TABLE 5. *Characters and character states of fossil horses.*

	Character State Codes*			Character State Codes*	
	A	B		A	B
Character 1. FEET:			**4. CROCHET:**		
3 toes, all supporting, phalanges broad	−1	0	Absent or present in only a few individuals	0	0
3 toes, lateral digits reduced in size and slightly shorter than middle digit, phalanges broad	0	−	Present, but prefossette closure not complete	1	−
			Prefossette closure complete	2	1
3 toes, lateral digits reduced in size and slightly shorter than middle digit, phalanges narrow	1	1	**5. CEMENT:**		
			Absent or irregularly developed	0	0
3 toes, lateral digits distinctly shorter	2	2	Present	1	1
1 toe, laterals reduced to splints	3	3	**6. FOSSETTES:**		
			With simple borders	−1	−1
2. LENGTH OF METATARSAL III:			Fossette closure not complete	0	0
Less than 140 mm	0	0	With complex borders	1	1
140 to 169 mm	1	−			
170 to 200 mm	2	1	**7. PROTOCONE:**		
Greater than 200 mm	3	2	Linked to protoselene	−1	−1
			Fossette closure not complete	0	0
3. AVERAGE CROWN LENGTHS (anteroposterior) of 4th upper premolar and 2nd upper molar:			Separated from protoselene	1	1
			8. CROWN HEIGHT:		
Less than 14.5 mm	0	0	Brachydont	0	0
14.5 to 16.9 mm	1	1	Hypso-brachydont	1	1
17.0 to 20.9 mm	2	−	Hypsodont	2	2
Greater than 20.9 mm	3	2	**9. METALOPH AND ECTOLOPH:**		
			Not connected	0	−
			Connected	1	0

* Column A. Character states for entire study. Column B. Character states for restricted study of species from genera represented in late Miocene or early Pliocene faunas (species 2, 3, 7, 8, 9, and 10), recoded when necessary.

6A. The other was character 3, average crown length of two cheek teeth, which Clemens, unknown to us, deliberately had coded from the point of view of operational homology, *i.e.*, from smallest to largest, in order to test the method. It is generally assumed from fossil evidence that each of the subgenera of *Merychippus*, here represented by OTU's 5 and 6, gave rise to at least one lineage of horses of larger size and another lineage in which there was a reduction in size. Interestingly this interpretation is sustained by character 3, reflecting changes in dimensions of the dentition, but not character 2, length of metatarsal III. This might be an artifact reflecting the choice of OTU's.

If the characters are operationally coded, reversals of evolutionary trends tend to confound the true cladogenesis of the data. When the compatibility matrix indicated character 3 to be miscoded, it was laid aside until after the cladogram had been reconstructed from the other characters. However, within clusters (5, 7, 8) and (6, 9, 10) we were unable to differentiate species 5 and 7 or 6 and 9. The generally accepted phylogenetic interpretation of the evolutionary changes for character 3 emerged automatically when the assumption of irreversibility was relaxed and the character was fitted most parsimoniously to the cladogram. Thus by a consistent application of the technique the originally

miscoded character was recognized and contributed to the final reconstruction.

For a subsequent analysis of "recent" forms we selected six species from genera represented in late Miocene or early Pliocene faunas (species 2, 3, 7, 8, 9, and 10). Although some of the ancestral genera (species 1, 4, 5, and 6) survived into this period, their characteristics were assumed to be unknown. However, the predicted characteristics of the ancestral genera are correct and the cladogram emerges as before. Character 9 was not used since it is invariant in this analysis. Thus, no prediction could be made about the nature of character 9 for OTU 1. Also, since species 5 and 6 were not included, it was not possible to infer the evolutionary changes of character 3 correctly.

Other Organisms

The Fusulinidae, a group of paleozoic protozoa were analyzed using data obtained from Dunbar (1963), Dunbar and Henbest (1942), and Dunbar and Skinner (1937) by Roger Kaesler of the Geology Department at The University of Kansas. Cladograms of genera as well as species corresponded well with ideas on cladistic relationships expressed by Dunbar. Studies of 25 species of bees of the *Hoplitis* complex and of 24 genera of Mecoptera (scorpionflies) were also carried out using recent material. In both instances the cladograms obtained by our method corresponded well with ideas on cladistic relationships expressed by authorities in the field. Separate publications on all of these studies are in preparation. Such studies are continuing and are suggesting methods for analyzing OTU's of supraspecific rank.

DISCUSSION

General Considerations

From the findings reported above it would appear to be possible to deduce cladistic sequences from the characteristics of recent organisms. It may be argued that the cladistic solutions obtained from recent organisms merely reflect the thinking of

TABLE 6. *Fossil horses.*
A. Data matrix.

		OTU's										Character states (c)	Minimum steps (c − 1)
		1	2	3	4	5	6	7	8	9	10		
Characters	1.	0	-1	1	1	2	2	2	2	2	3	5	4
	2.	0	3	0	1	2	2	2	3	2	3	4	3
	3.	0	3	0	1	2	2	1	3	1	3	4	3
	4.	0	0	0	1	2	2	2	2	2	2	3	2
	5.	0	0	0	0	1	1	1	1	1	1	2	1
	6.	0	0	0	0	1	-1	1	1	-1	-1	3	2
	7.	0	0	0	0	1	-1	1	1	-1	-1	3	2
	8.	0	0	1	0	2	2	2	2	2	2	3	2
	9.	0	1	1	1	1	1	1	1	1	1	2	1
													20

B. Compatibility matrix.

		Patterns									Compatibilities	Extra steps
		1	2	3	4	5	6	7	8	9		
Characters	1.	X	4	5	1	1	1	1	1	0	1	14
	2.	5	X	1	3	3	4	4	3	0	1	23
	3.	6	1	X	3	3	5	5	3	0	1	26
	4.	1	3	4	X	1	1	1	1	0	1	12
	5.	0	1	2	0	X	0	0	0	0	6	3
	6.	1	2	4	0	0	X	0	0	0	5	7
	7.	1	2	4	0	0	0	X	0	0	5	7
	8.	1	3	5	1	1	1	1	X	0	1	13
	9.	0	1	1	1	1	1	1	1	X	1	7
Compatibilities:		2	0	0	3	2	2	2	3	8	22	−
Extra steps:		15	17	26	9	10	13	13	9	0	−	112

the taxonomists who furnished us with the data. This is true in a general sense. However, the method proposed here tests the assumptions behind character coding. Repeatedly, through our methodology, we have been able to point out to our colleagues errors in their reasoning about evolutionary trends in characters.

While evolutionists probably have a relatively thorough understanding of modes of evolutionary change, assumptions about the relative frequencies of these phenomena may be in error. We therefore do not know how frequently assumptions 3, 4, and 5 about character coding will be valid in any given study.

The correctness of our approach depends

on the assumption that nature is indeed parsimonious. Alternative, equally parsimonious solutions may appear and the choice between them may not be evident from the data at hand. While the addition of a single new character may permit a decision, we should be on our guard against relying too firmly on the cladograms so obtained. Far-reaching decisions about stems are sometimes taken on the weight of a single evolutionary step. Partly this has been due to our choice of few characters for the analyses which initially were carried out by hand. While it is remarkable that even with few characters we obtained results consistent with the known facts, it is obvious that more characters would make decisions on junctions less likely to be dependent on the presence or absence of single evolutionary steps. The larger studies with more extensive suites of characters currently being processed by computer should lead to firmer cladograms. The probability of correctness of any portion of the tree varies with the relative reliability of our interpretation of any one character and with the number of characters and character states on which it is based.

The method as described above assumes equal probability of all evolutionary steps[5] after the characters have been coded. The method of coding characters and our initial assumptions about the evolutionary trends do reflect judgments based on biological knowledge of the material. The criteria by which this may be done have been outlined by several authors (Hennig, 1957; Maslin, 1952). Thus in actuality all evolutionary steps are not assumed to be equally probable.

The method proposed here is not substantially different from the conventional cladistic approaches of phylogenists. It simply quantifies and systematizes these procedures, making them objective in the process and permitting them to be put on

a computer. Thus, they have the same relation to conventional (cladistic) phylogeny that numerical taxonomy has to conventional phenetic taxonomy. Just as the study of numerous characters and the preparation of dendrograms in numerical taxonomy enhance knowledge and understanding of systematic relationships, so an analysis of cladogenesis along the lines proposed here leads systematists to critical tests of their ideas and assumptions about a phylogeny.

The proposed method does not weight characters equally in the construction of the cladogram, since compatible characters are preferred over those that are incompatible. Characters with few states tend to be more compatible than those with many. Since evolutionary steps are equally weighted, those with more states will be more heavily weighted. However, the weighting procedure agrees with the principles of numerical taxonomy (Sokal and Sneath, 1963); it is automatic and *a posteriori*, based on the entire available evidence rather than on *a priori* or character-by-character weighting as employed in conventional phylogenetic procedures.

Technical Points

The method illustrated here and several variations currently being investigated are empirical approaches to finding the most parsimonious cladogram. A possible pattern cladogram might be a phenogram, if phenetics is closely related to cladistics. However, cladograms and phenograms will be similar only when similarities are due to recent divergence. It is to be expected that the two types of dendrograms will not be entirely alike because they measure different aspects of phyletic relationship. Locating the cladogram requiring the minimum number of evolutionary steps by trial and error is a stupendous computational task, but might be made manageable by a Monte Carlo method. We have therefore attempted to reach a near parsimonious solution by one of the methods reported above, before applying trial and

---

[5] This point has been called to our attention by E. C. Minkoff, Harvard University.

error improvements. An analytical mathe- matical solution which would give the single most parsimonious cladogram is a difficult mathematical problem.

We do not yet know how to evaluate all of the information in the compatibility matrix. It indicates when some of our assumptions, such as irreversibility in evo- lution, are wrong. The matrix may also point out cases in which the basic assump- tion of evolutionary parsimony is invalid.

A feature for changing the primitive state of a character can be built into the computer program to try a variety of as- sumptions adopted by the operator to re- veal the most parsimonious, internally consistent evolutionary pattern. In this way we could investigate hypotheses about evolutionary trends in the character. There seems to be no fundamental obstacle to assuming more complicated evolutionary trends than the V-shaped ones discussed above. Thus a character coded

$$-2 \leftarrow -1 \leftarrow 0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 4$$
$$\downarrow \qquad \downarrow$$
$$-2' \qquad 3'$$
$$\downarrow$$
$$4'$$
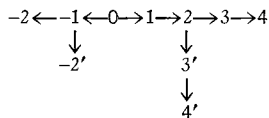
could be included in the computations.

We have as yet no technique for dealing with cases of hybridization. Similarly, we have not yet explored the consequences of missing data for character states of given OTU's. These could be handled by adding the characters to the cladogram after the construction is complete, in the same way in which we now add the characters which differ in a single state for one OTU only. The cladogram might then provide a means of predicting the missing character states as it does for the character states of an- cestral forms.

Implications for Systematics

The development of a technique for de- ducing cladistic relationships among orga- nisms appears to furnish a base for classi- fication alternative to the phenetic system espoused by numerical taxonomy. Since

the present method appears to be the "phyletic" one professed by orthodox taxonomists, it might appear that classifi- cations should be established on the basis of it or a similar method. However, we have pointed out elsewhere (Sokal and Camin, 1965) that phyletic relationships are always a composite of phenetic, cladis- tic, and chronistic relationships not always clearly separated in the minds or writings of systematists and that systematics as a whole must be based on all of these con- siderations. The degree to which phenetics and cladistics coincide is not yet known, although we may assume it to be consider- able. Since no operational system for com- bining phenetic and cladistic relationships is available we must choose between a phenetic or cladistic basis for classification. We distinguish here between "systematics" and "classification," the former including not only the study of the order of living things but also the causes and processes bringing this about, while classification is simply the arrangement of organized nature into categories for the convenience of biologists. For a variety of reasons, de- tailed elsewhere (Sokal and Camin, 1965), it would seem that a phenetic basis is pre- ferable for classification in this narrow sense until an operational system, combin- ing cladistics and phenetics can be estab- lished.

Comparison of phenograms with clado- grams may lead to the resolution of phenetic resemblance into its components discussed by us in detail in Sokal and Camin (1965). If by a comparative cladistic and phenetic study of a group of organisms it has been shown that an apparently similar char- acter in two organisms could not possibly be due to primitive patristic similarity, it must therefore be a case of parallelism (derived patristic similarity) or of classical convergence. This will stimulate biologists to a study of the underlying structural and physiological phenomena which lead to an apparently identical result. Such studies may permit the separation of parallelisms from divergence and may facilitate the

recognition of characters which have been miscoded as convergent. The joint consideration of chronistics, cladistics, and phenetics will also enable objective measurement of evolutionary rates.

SUMMARY

A method is described for reconstructing presumed cladistic evolutionary sequences of recent organisms and its implications are discussed. Characters of the organisms to be studied are presented in a data matrix of the type employed in numerical taxonomy with the character states arrayed according to a presumed evolutionary sequence. The reconstruction proceeds on the hypothesis that the minimum number of evolutionary steps yields the correct cladogram. The method has been programmed for computer processing.

LITERATURE CITED

ARNETT, R. H., JR. 1963. The phenogram, a method of description for studies on *Oxacis* (Coleoptera, Oedemeridae). Coleopt. Bull., **17**: 6–18.

BADER, R. S. 1956. A quantitative study of the Equidae of the Thomas Farm Miocene. Bull. Mus. Comp. Zool., **115**: 47–78.

DOOLITTLE, R. F., AND B. BLOMBÄCK. 1964. Amino-acid sequence investigations of fibrinopeptides from various mammals: evolutionary implications. Nature, **202**: 147–152.

DOWNS, THEODORE. 1956. The Mascall fauna from the Miocene of Oregon. Univ. California Publ. Geol. Sci., **31**: 199–354.

——. 1961. A study of variation and evolution in Miocene *Merychippus*. Contrib. Sci., Los Angeles County Mus., **45**: 1–75.

DUNBAR, C. O. 1963. Trends of evolution in American fusulines. *In* Von Koenigswald, G. H. R., *et al.* (eds.), Evolutionary trends in Foraminifera. Elsevier Publ. Co., New York, pp. 25–34.

DUNBAR, C. O., AND L. G. HENBEST. 1942. Pennsylvanian Fusulinidae of Illinois. Bull. Illinois State Geol. Surv., **67**: 1–218.

DUNBAR, C. O., AND J. W. SKINNER. 1937. Upper Paleozoic ammonites and fusulinids. 2. Pernian Fusulinidae of Texas. Univ. Texas Bull., **3701**: 517–825.

EDWARDS, A. W. F., AND L. L. CAVALLI-SFORZA. 1964. Reconstruction of evolutionary trees. *In* Phenetic and phylogenetic classification. Systematics Assoc. Publ., **6**: 67–76.

GIDLEY, J. W. 1906. New or little known mammals from the Miocene of South Dakota. Bull. Amer. Mus. Nat. Hist., **22**: 135–153.

——. 1907. Revision of the Miocene and Pliocene Equidae of North America. Bull. Amer. Mus. Nat. Hist., **23**: 865–934.

GREGORY, J. T. 1942. Pliocene vertebrates from Big Spring Canyon, South Dakota. Univ. California Publ. Geol. Sci., **26**: 307–446.

HENNIG, W. 1957. Systematik und Phylogenese. Ber. Hundertjahrfeier Deutsch. Entomol. Ges., pp. 50–70, Berlin.

LANCE, J. F. 1950. Paleontología y estratigrafía del Plioceno de Yepomera, Estado de Chihuahua. Iᵃ parte: Equidos, excepto Neohipparion. Bol. Inst. Geol. Mex., **54**: 1–83.

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

MAYR, E. 1965. Numerical phenetics and taxonomic theory. Syst. Zool., **14**: 73–97.

OSBORN, H. F. 1918. Equidae of the Oligocene, Miocene, and Pliocene of North America; iconographic type revision. Mem. Amer. Mus. Nat. Hist. (n.s.), **2**: 1–330.

SCHLAIKJER, E. M. 1932. The osteology of *Mesohippus barbouri*. Bull. Mus. Comp. Zool., **72**: 391–410.

——. 1935. Contributions to the stratigraphy and palaeontology of the Goshen Hole area, Wyoming. IV. New vertebrates and the stratigraphy of the Oligocene and early Miocene. Bull. Mus. Comp. Zool., **76**: 97–189.

SIMPSON, G. G. 1932. Miocene land mammals from Florida. Bull. Florida Geol. Surv., **10**: 7–10.

——. 1951. Horses. The story of the horse family in the modern world and through sixty million years of history. Oxford Univ. Press, New York, 247 pp.

——. 1963. The meaning of taxonomic statements. *In* S. L. Washburn [ed.], Classification and human evolution. Aldine, Chicago, pp. 1–31.

SOKAL, R. R., AND J. H. CAMIN. 1965. The two taxonomies: areas of agreement and of conflict. Syst. Zool., in press.

SOKAL, R. R., AND P. H. A. SNEATH. 1963. Principles of numerical taxonomy. W. H. Freeman and Co., San Francisco and London, 359 pp.

STIRTON, R. A. 1940. Phylogeny of North American Equidae. Univ. California Publ. Geol. Sci., **25**: 165–198.

THROCKMORTON, L. H. 1965. Similarity versus relationship in *Drosophila*. Syst. Zool., in press.

WHITE, T. E. 1942. The lower Miocene mammal fauna of Florida. Bull. Mus. Comp. Zool., **92**: 1–49.

WILSON, E. O. 1965. A consistency test for phylogenies based on contemporaneous species. Syst. Zool., in press.