

A New Method for Representing Trees

CHRISTIAN CHARLES EMIG

Station Marine d'Endoume, Rue de la Batterie-des-Lions,
13007 Marseille, France

When constructing a tree according to Wagner's groundplan divergence analysis (Wiley, 1981), the single available value for plotting each taxon on the graph is the Divergence Index (*DI*), which just allows the investigator to place the taxon somewhere on a concentric semicircle that equals its *DI* (Fig. 1a). However, the position on the semicircle cannot be determined for actual taxa or for hypothetical common ancestors. Therefore, it is desirable for a second parameter to be introduced on the graph to avoid such subjectivity.

In this paper, I develop a technique for representing on such a graph the number of characters that have been used in the analysis as this second parameter. This technique can be used in conjunction with any method of phylogenetic character analysis to construct a tree, showing in summary form more of the information in the data matrix.

METHODS

The number of characters used in the analysis is indicated on the graph by centripetal lines crossing the semicircles as shown on Figure 1b; these lines are numbered from 0, a horizontal line, that equals no apomorphy (or only plesiomorphies) up to the vertical line which equals the total number of characters used (or all apomorphies). Thus, these centripetal lines involve the number of apomorphies (synapomorphies and autapomorphies) determined in the character analysis and of course are directly related to the *DI*, which is the total number of apomorphies of each taxon. The place of each taxon on the graph is plotted at the intersection of the concentric semicircle and the centripetal line that together indicate the *DI*.

To determine the *DI* of each taxon, Wag-

ner methods (Wiley, 1981) are followed: a plesiomorphic character receives a score of 0 and each apomorphic character a score of 1; if a transformation series contains several "intermediate apomorphies," the latter might be scaled between 0 and 1, as recommended by Wiley (1981). The *DI* totals the values for each taxon for all characters. A data table of coded characters is constructed, the taxa being ranked according to the increasing *DI*. This has been done in Tables 1 and 2, but for a large data set only the apomorphic characters can be indicated (as shown in Table 3), thus making the table easier to comprehend.

When the taxa are plotted on the graph, applying the technique described above, they are linked to one another through hypothetical common ancestors. These hypothetical ancestors are determined from the data table by taking the median (*M*), which is defined as the number of shared apomorphies (=synapomorphies) between two or several taxa, and *M* is plotted on the graph according to its value by the same manner as for the taxa. The method proceeds by taking the *M* of taxon 1 (generally the first taxon in the data set) which shows the smallest number of apomorphies shared with all other taxa (taxon 2 to taxon *n*): this median is *M*₁ (taxon 1, taxon 2 to taxon *n*) which indicates the hypothetical common ancestor of all the taxa. If taxon 1 shows no other synapomorphy with another taxon, it is connected to *M*₁ on the graph, and one begins again on the data set by taking *M*₂ (taxon 2, taxon 3 to taxon *n*), which is connected on the graph to *M*₁ and to taxon 2, and so on (see examples). When a taxon shares one or several other synapomorphies with one or several other taxa, a sister group is determined (see example 3). The tree of the relationships is constructed when all taxa are linked, and such a graphic may be

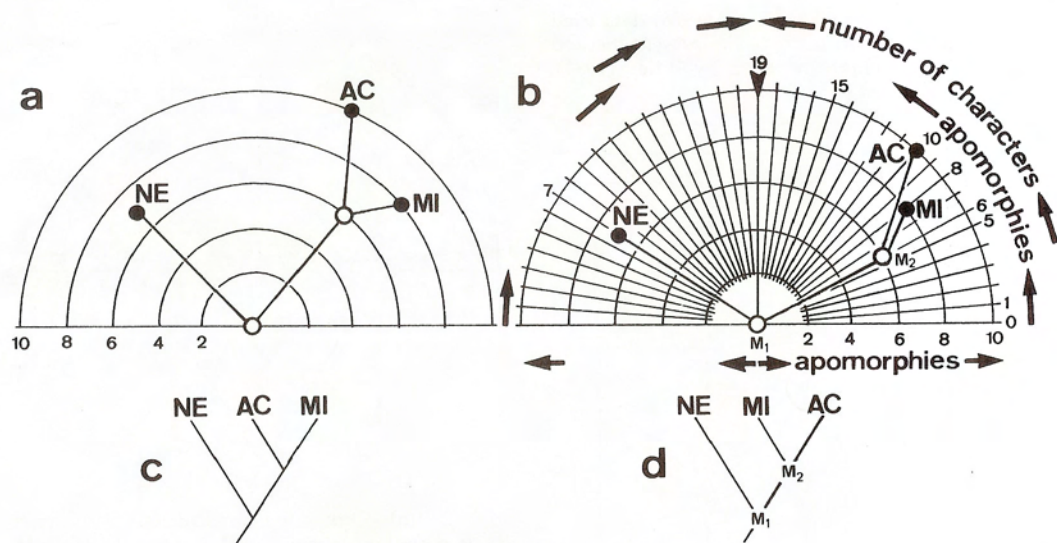


FIG. 1. (a) Results of phylogenetic analysis using the Wagner groundplan divergence method (from Wiley, 1981). (b) Results of a phylogenetic analysis using the new method. (c) Relationships among the three selected taxa constructed from Figure 1a (adapted from Platnick and Shadab [1976] by Wiley [1981]). (d) Relationships constructed from Figure 1b.

TABLE 1. Data adapted (see text) from data (of Platnick and Shadab, 1976) used by Wiley (1981) to demonstrate a Wagner groundplan divergence analysis. Characters are coded as 0 (plesiomorphic) and 1 (apomorphic). The medians are $M_1(NE, MI \text{ to } AC) = 0$ and $M_2(MI, AC) = 6$.

Character number	Taxa		
	NE	MI	AC
1	1	0	0
2	1	0	0
3	1	0	0
4	1	0	0
5	1	0	0
6	1	0	0
7	1	0	0
8	0	1	1
9	0	1	1
10	0	1	1
11	0	1	1
12	0	1	1
13	0	0	1
14	0	0	1
15	0	1	1
16	0	0	1
17	0	0	1
18	0	1	0
19	0	1	0
DI	7	8	10

represented by a standard cladogram, as shown in Figures 1d, 2b, and 3b.

EXAMPLES

Example 1.—To illustrate the above theoretical argument, I have chosen the example given by Wiley (1981), from data of Platnick and Shadab (1976), to compare my results with those proposed by Wiley using Wagner's basic method. Figure 1b and Table 1 show the results of my method applied to data from three taxa. Figure 1a represents the tree constructed by Wiley (1981), whose hypothesis of relationships led him to the more justifiable cladogram, shown in Figure 1c. According to Table 1 and Figure 1b, MI is less distant than AC from the hypothetical common ancestor $M_2(MI, AC)$ that produces the tree in Figure 1d. Thus, Figure 1b is more informative, applying the direct measure of the distance between taxa, and the same line of proceeding will be used in the following examples to construct a standard cladogram (i.e., when taxa have the same

TABLE 2. Platnick and Shadab's (1976) data used to compare the tree according to the present method with the Wagner algorithm used by Wiley (1981). Characters are coded as in Table 1.

Character number	Taxa			
	MA	NE	MI	AC
1	0	1	0	0
2	0	1	0	0
3	0	1	0	0
4	0	1	0	0
5	0	1	0	0
6	0	1	0	0
7	0	1	0	0
8	0	0	1	1
9	0	0	1	1
10	0	0	1	1
11	0	0	1	1
12	0	0	1	1
13	0	0	0	1
14	0	0	0	1
15	0	0	1	1
16	0	0	0	1
17	0	0	0	1
18	0	0	1	0
19	0	0	1	0
20	0	1	1	1
21	1	0	0	0
DI	1	8	9	11

M , the taxon with a smaller DI is connected first to M , and so on).

Because $M_1 = 0$, NE has been placed on the left side of the median line in Figure 1b, and the two other taxa, which share together a common ancestor M_2 , are on the other side. Even when NE is plotted on the right side (Fig. 2a) the graph analysis leads to the same result, but by a more complex diagram. My method, when compared to the Wagner groundplan divergence method, produces a more precise and more justifiable tree.

Example 2.—My method has similarities to the Wagner-algorithm method proposed by Kluge and Farris (1969) and Farris (1970). In order to compare the results of my method with those of the previous one, I have used the first example (Table 1) expanded to include taxon MA as the outgroup, according to the data computed by Wiley (1981).

The matrix of characters is arranged as previously done in Table 2. The four taxa are plotted in Figure 2a according to the

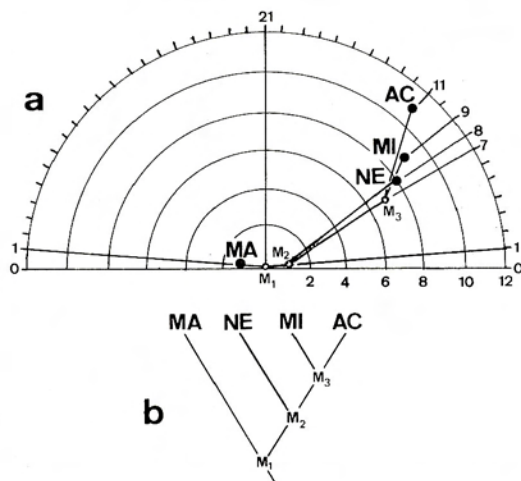


FIG. 2. (a) Wagner tree constructed according to the new method from Table 2. Centripetal lines are omitted except as they lead to the margin from a hypothetical or actual taxon. (b) Phylogenetic relationships constructed from Figure 2a.

DI. MA was selected as the outgroup and placed on the left side of the graph, and the other taxa were placed on the right side. The distances between the taxa can be seen directly on the graph. To link MA to the hypothetical common ancestor, the median $M_1(MA, NE \text{ to } AC) = 0$ is determined from Table 2 and plotted in Figure 2a. MA is connected to M_1 , then $M_2(NE, MI \text{ to } AC) = 1$ is determined, placed on the graph, and connected to M_1 and to NE (which shares no other apomorphy with the following taxa). Finally, $M_3(MI, AC) = 7$ is connected to M_2 and to MI and AC . The tree has now been built, and it may be represented by a standard cladogram (Fig. 2b).

The present results can be obtained rapidly by hand without a computer program and produce the same tree as that found with the Wagner algorithm given by Wiley (1981). Even when there is a data matrix with suspected homoplasy, my method will disclose the impossibility of establishing and connecting the hypothetical common ancestor between two or several taxa, thus indicating a possible error or misinterpretation of polarity in the character analysis. Obviously, the present

TABLE 3. Data of Lynch (1982)^a used to construct trees according to the present method. Characters are coded 1 (apomorphic) and X (homoplasious); pleiomorphies coded 0 are omitted.

Character number	Taxa						
	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1
3		1	1	1	1	1	1
4		1	1	1	1	1	1
5		1	1	1	1	1	1
6	1						
7	1						
8		1	1	1			
9					1	1	1
10					1	1	1
11					1	1	1
12			1	1			
13				X			X
14							1
15							1
16							1
17							1
18							1
19							1
20		X		X		1	1
DI	4	6	7	7	8	9	15

^a Correspondence with Lynch's (1982) taxa: my A = Lynch's A, B = E, C = F, D = B, E = G, F = C, G = D.

method can be applied on rather small data sets. With larger data matrices a computer program would be helpful.

Example 3.—I used the table from Lynch (1982) to test my method on real data. The DI was determined for each of the seven taxa, coded by Lynch (1982), and the taxa were ranked according to the increasing DI in Table 3 and plotted in Figure 3a. To represent the relationships among taxa, various medians were determined from Table 3 and plotted on graphs (Fig. 3). The steps followed were:

$$M_1(A, B \text{ to } G) = 2. \quad (1)$$

Since A shows no synapomorphy with other taxa, it was connected to M_1 .

$$M_2(B, C \text{ to } G) = 5. \quad (2)$$

M_2 is connected to M_1 but, as B shares another synapomorphy (character 8) with C and D, these three taxa are a sister group and have together a median:

$$M_3(B, C \text{ to } D) = 6. \quad (3)$$

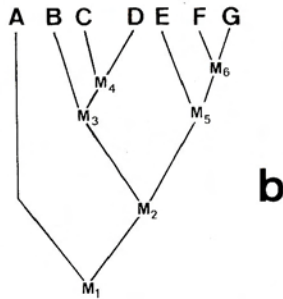
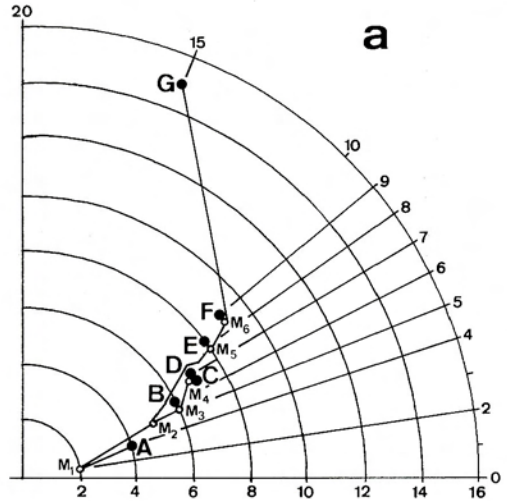


FIG. 3. (a) Wagner tree constructed according to the new method from Table 3 (adapted from data of Lynch, 1982). (b) Phylogenetic relationships constructed from Figure 3a.

M_3 is connected to M_2 . Here the positions of M_3 and B are the same on the graph (Fig. 3a) but, given that no taxon will occupy an ancestral position when using Wagner methods (Wiley, 1981), this position indicates only a link of B to M_3 (see Fig. 3b). Taxa C and D share character 12,

$$M_4(C, D) = 7. \quad (4)$$

M_4 is connected to M_3 . The remarks about M_3 also apply here for C and D (for linking C and D in Fig. 3b, see text above). As taxa B, C, and D are a sister group, one goes back to M_2 and (in Table 3) to the next taxon which is E,

$$M_5(E, F \text{ to } G) = 8. \quad (5)$$

M_5 is connected to M_2 , and the same remarks apply here for M_5 and E . Lastly,

$$M_6(F, G) = 9. \quad (6)$$

M_6 is connected to M_5 ; then F and G are joined to M_6 . All the taxa are linked, with the tree shown in Figure 3b. This is the same as the hypothesis proposed by Lynch (1982). His tree can be constructed directly without going through the graphic represented in Figure 3a. The distance between taxa is in this case estimated on the data table. However, the tree in Figure 3a summarizes more of the information that can be visually appreciated. The statement of Lynch (1982), that the autapomorphic attributes 14-19 are uninformative for purposes of determining relationships, does not appear to be correct because all apomorphies have to be coded to construct a tree of relationships using the *DI* (i.e., all selected characters are informative, even homoplasious ones). Such information operates to determine the position of a taxon on the tree. In the present example, when characters 14-19 are disregarded, the linking of F and G to M_6 (Fig. 3b) may be F and G or G and F , both taxa then having a *DI* equal to 9. With respect to homoplasy that has been coded under X in Table 3, I suggest such a condition be used to connect taxa which display the same M and the same *DI*. Consequently, as taxon D possesses two homoplasious attributes (Table 3; characters 13 and 20) and taxon C none, C is connected first to M_4 , then D , as shown in Figure 3b. In the same way, if characters 14-19 are disregarded, the homoplasious character 13 could be used to link F and G to M_6 .

CONCLUSIONS

The theoretical argument and the examples presented demonstrate that the

present method, derived from Wagner methods, is a technique that improves the information content of Wagner trees, packing more information to be visually appreciated and providing methods for summarizing more information in the data matrix. This technique works on a phylogenetic analysis, applying the principle of parsimony. Still, the present method demonstrates that it is possible to take into account not only the number of apomorphies by the *DI*, but also the number of plesiomorphies. The latter are indicated on the graph by the distances of various taxa to the vertical line, which indicates no plesiomorphy. The graphic technique (Figs. 1b, 2a, 3a) shows visual images that make sense from an evolutionary point of view. Homoplasy is introduced and coded in the data matrix, and such information may be used to determine the position of two taxa linked to the same M and showing the same *DI*.

REFERENCES

- FARRIS, J. S. 1970. Methods for computing Wagner trees. *Syst. Zool.*, 19:83-92.
- KLUGE, A. G., AND J. S. FARRIS. 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.*, 18:1-32.
- LYNCH, J. D. 1982. Relationships of the frogs of the genus *Ceratophrys* (Leptodactylidae) and their bearing on hypotheses of Pleistocene forest refugia in South America and punctuated equilibria. *Syst. Zool.*, 31:166-179.
- PLATNICK, N. I., AND M. U. SHADAB. 1976. A revision of the mygalomorph spider genus *Neocteniza* (Araneae, Actinopodidae). *Am. Mus. Novitat.*, 2603:1-19.
- WILEY, E. O. 1981. *Phylogenetics: The theory and practice of phylogenetic systematics*. Wiley-Interscience, New York.

Received 5 November 1984; accepted 31 January 1985.