FORMATION OF MONOTHETIC GROUPS IN QUANTITATIVE BACTERIAL TAXONOMY[1]

W. R. LOCKHART AND P. A. HARTMAN

Department of Bacteriology, Iowa State University, Ames, Iowa

Received for publication 25 July 1962

ABSTRACT

LOCKHART, W. R. (Department of Bacteriology, Iowa State University, Ames) AND P. A. HARTMAN. Formation of monothetic groups in quantitative bacterial taxonomy. J. Bacteriol. **85**:68–77. 1963.—Previous applications of quantitative methods to bacterial classification have resulted in polythetic groupings in which no organism necessarily has all the features characteristic of its group, and no single property is necessarily possessed by every member of a group. Such classifications are not altogether analogous to present taxonomies, and (since the possibile groups are not mutually exclusive) hence cannot be evaluated completely by computer methods. A technique is presented for formation of hierarchical monothetic groups in which the criterion for addition of each new member is the possession of an array of properties common to all organisms already in the group. These monothetic relationships among individuals are expressed in terms of cumulative difference, which is linearly related to the taxonomic distance (a function of similarity ratio) of polythetic classifications. Monothetic groupings obtained for 50 representative microorganisms were essentially similar to the polythetic groups evaluated by earlier methods. The necessary computation is suitable for analyzing relationships among large numbers of organisms.

Recent applications of quantitative methods to taxonomy have been based on the principles that classification is a measure of the over-all similarity among organisms and that all properties of organisms are potentially of equal value in creating taxa, so that no *a priori* assumptions need be made of the relative importance of particular features. The investigator gathers data concerning a large number of properties of each organism studied, records these data in suitable form, and calculates (with the aid of an electronic computer) the similarity of each organism to all the others. This quantitative estimate usually is expressed as the numerical ratio or proportion of the properties in which two organisms are alike to the total properties compared. The resulting matrix of similarity values is then arranged or "ordered" into groups of mutually similar organisms (Sneath and Sokal, 1962).

This approach to bacterial systematics has yielded encouraging results in the hands of many investigators (reviewed by Sneath, 1962), although all problems associated with use of the technique have not been resolved. It is not known, for example, whether negative matches (e.g., the fact that neither of two organisms produces indole) should be counted as similarities. Additional problems arise in scoring properties wherein neither result may be considered negative (e.g., colony type), when more than two alternatives exist (e.g., cell morphology), or where a quantitative response is involved (e.g., drug resistance). Although a number of proposed solutions to such problems have been tested (Hill et al., 1961; Beers and Lockhart, 1962; Beers et al., 1962), no procedure has completely satisfied all workers.

A more serious problem is presented by the available techniques for group formation. The several alternative means of calculating similarities and arranging organisms into groups result in groupings which (since they are derived from measures of similarities between pairs of organisms) are polythetic (Sneath, 1962); that is, no property is necessarily possessed by all individuals in a group, and no organism necessarily has all the properties generally characteristic of members of its group. Consequently, any given organism may score mathematically as being equally similar (though in different respects) to two or more other individuals which, in turn, are quite dissimilar to each other. Since polythetic groups

[1] Journal Paper No. J-4401 of the Iowa Agricultural and Home Economics Experiment Station, Ames, Iowa. Project No. 1509.

are thus not mutually exclusive, it becomes virtually impossible to compose a computer program which infallibly will sort the organisms studied into optimal groupings. The resulting necessity for supplementing machine programs for group formation with visual examination of data eventually imposes severe limitations, because the number of possible pairs (hence the number of similarity values obtained) increases in proportion to the square of the number of organisms studied. Further, the groupings obtained by this means are not theoretically analogous to the hierarchical taxa of present classifications. The latter groupings are monothetic (Sneath, 1962); all individuals in such a group are presumed similar to one another in the same respects, with a core of common properties shared by every member of the group.

The purpose of the work reported in this paper was to devise and test a procedure for formation of groups on a monothetic basis, to learn whether such a technique might be useful in handling larger quantities of data, and to compare the groups formed according to monothetic vs. polythetic criteria.

Materials and Methods

The organisms used in this study are listed in Table 1. The strain numbers only are shown in subsequent figures. Names enclosed in parentheses in Table 1 indicate that the isolate in question did not agree with any species described in *Bergey's Manual of Determinative Bacteriology*, but was judged to resemble most closely the organism designated. These strains were used by Beers et al. (1962) in an earlier comparison of various scoring methods, and the reader is referred to that paper for details as to the sources of the various strains and the diagnostic tests used in their characterization. Sixty properties of each organism were determined, and the results scored as indicated by Beers et al. (1962, Method III). Briefly, the properties were numbered, and for each organism the symbol A, B, C, or D was recorded for every property. [The scoring convention proposed by Beers and Lockhart (1962) and used in earlier experiments with these strains (Beers et al., 1962) was identical in effect with the present method, but differed in detail. The previous technique made use of three symbols (+, −, 0) and assigned a separate feature to each alternative expression of a property. Since the binary number system of computer "language" permits use of only two symbols (1 or a cipher) in any position, two positions were needed to accomodate the three symbols used. Because at least two features (alternatives) were assigned to any property, four or more machine positions were required for each property. Substitution of the four symbols A, B, C, D permits any property to be recorded in only two positions and makes it possible for the computer "memory" to accomodate substantially larger amounts of data. We have found four alternatives sufficient for recording the properties used in these studies; if more were needed, three positions could be used and would permit coding of up to eight alternative expressions of each property.] Examples of the scoring of typical properties are shown in Table 2. The computer was instructed to score a similarity if the same symbol appeared

TABLE 1. *Test organisms*

1 *Mima polymorpha*	26 *A. aerogenes*
2 *M. polymorpha*	27 *A. aerogenes*
3 *M. polymorpha*	28 (*Aerobacter* sp.)
4 *M. polymorpha*	29 (*Aerobacter* sp.)
5 *Herellea* sp.	30 *Proteus vulgaris*
6 *Alcaligenes faecalis*	31 *Serratia marcescens*
7 *A. faecalis*	32 *S. marcescens*
8 (*Vibrio* sp.)	33 *S. marcescens*
9 (*Achromobacter* sp.)	34 *S. marcescens*
10 (*Achromobacter* sp.)	35 *S. marcescens*
11 *Pseudomonas reptilivora*	36 *S. marcescens*
12 *P. reptilivora*	37 *S. marcescens*
13 *P. reptilivora*	38 *S. marcescens*
14 *P. boreopolis*	39 *S. marcescens*
15 *P. boreopolis*	40 *S. marcescens*
16 *P. boreopolis*	41 *S. marcescens*
17 *P. boreopolis*	42 *Streptococcus faecalis*
18 *Escherichia coli*	43 *S. faecalis*
19 *E. coli*	44 *S. faecalis*
20 *E. coli*	45 *S. faecalis* var. *liquefaciens*
21 *E. coli*	46 *S. faecalis* var. *liquefaciens*
22 *E. coli*	47 *S. faecalis* var. *liquefaciens*
23 *Aerobacter cloacae*	48 *S. faecalis* var. *liquefaciens*
24 *A. cloacae*	49 *S. faecalis* var. *liquefaciens*
25 *A. cloacae*	50 *S. mitis*

TABLE 2. *Examples of the scoring of properties of the test organisms for computer analysis*

Property	Alternatives	Code symbols
Production of indole	Indole produced	A
	Indole not produced	B
Colony morphology	Punctiform	A
	Circular	B
	Irregular	C
Tolerance to NaCl	Sensitive	A
	Weakly resistant	B
	Moderately resistant	C
	Strongly resistant	D

under a given property for two organisms, and to score a difference whenever different symbols appeared. The result was that two organisms scored a single similarity if they were alike in a given property (either positive or negative), and a single difference if they were unlike. The advantages of such a scoring system, which avoids possible bias due to unequal numbers of comparisons between various strains (Silvestri et al., 1962), were outlined by Beers and Lockhart (1962).

Computer programs for calculating similarities and for group formation, described in Results, were performed on the CYCLONE digital computer of Iowa State University.

RESULTS

Relationships among the test organisms were first determined in terms of polythetic groupings, making use of the similarity ratios between pairs of organisms:

$$S = \frac{n_s}{n_s + n_d}$$

where n_s is the number of similarities between two organisms and n_d is the number of differences. The value of S would be 1.00 for two identical strains and 0.00 for two totally dissimilar strains. The value of n_s used includes both positive and negative similarities. The polythetic groupings obtainable with these particular strains through use of various scoring methods were determined previously; part of the diagram of groupings shown in Fig. 1 appeared elsewhere in slightly different form (Fig. 3 in Beers et al., 1962).

The grouping shown in Fig. 1 is reasonably satisfactory. Group I consists of a tight cluster of *Mima polymorpha* (DeBord, 1942) loosely surrounded by strains of *Alcaligenes faecalis*, "*Vibrio* sp.," and *Herellea* sp. (DeBord, 1942). Group II contains a tight cluster of strains of *Pseudomonas boreopolis* allied with three strains of *P. reptilovora* and two of *Achromobacter* spp. The enterobacteria are related at a higher level of similarity, with separate groups containing *Escherichia coli* (III), *Aerobacter* (IV), with three strains of *A. cloacae* forming a tight cluster, and *Serratia marcescens* (V). The single strain of *Proteus vulgaris* (30) appears to constitute a separate subgroup attached to group II. Group V consists of several subgroups, one of which (A) seems to be an eco-group containing two unidentified isolates and a single *S. marcescens*, all of which are insect pathogens (Beers et al., 1962). Group VI, the streptococci, also contains two subgroups.

The groupings shown are satisfactory and seem to confirm the identifications of the various strains, but were obtained only after some manual rearrangement of the matrix of S values. It was particularly difficult to assign certain strains to subgroups in groups V and VI, and to decide the proper position of those strains shown as joining groups I and II at low levels of similarity. Obviously it would be advantageous to devise criteria for group formation which are simple enough to be incorporated fully into computer programs. The first step toward inducing the computer to reproduce from original data a dendrogram like that in Fig. 1 is relatively easy. The high points of the diagram (e.g., strains 3 and 4, or 28 and 29, etc.) may be recovered simply by instructing the machine to locate the most similar pair (i.e., the two strains with the fewest differences) among all the organisms studied. To these (the "nucleus pair" of a group) could then be added the strain most similar to the nucleus pair, then that which is most similar to the first three, *ad infinitum*. Unfortunately, additional strains would not be found which were equally similar to all present members of a group. Even if mutually exclusive groups existed, methods for recovering polythetic groups (Sneath, 1962) require the calculation and comparison of all the possible pair-similarities involved, a

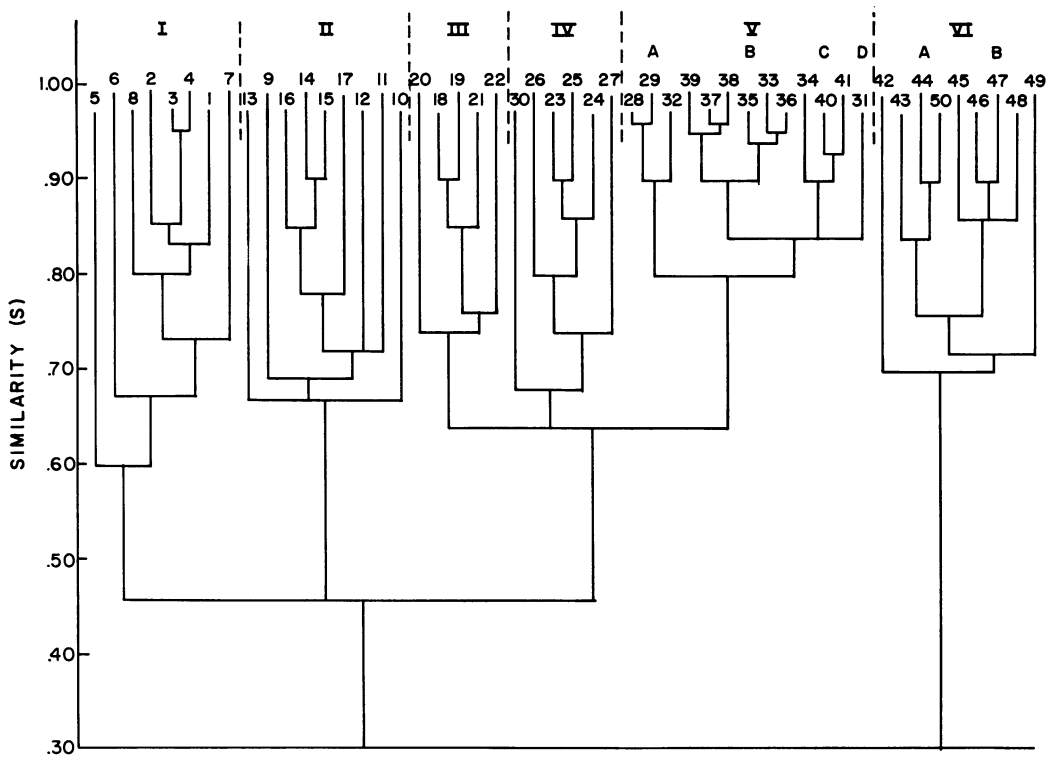

FIG. 1. *Polythetic grouping of the test organisms. Numbers correspond to the strain designations listed in Table 1; roman numerals indicate major groups, separated by dotted lines in the diagram.*

time-consuming process which would become impracticable if very large numbers of strains were involved. To simplify this process, then, a monothetic criterion was adopted for addition of organisms to the group.

If we assume that any properties in which the nucleus pair differ are nonessential (in the sense that they do not contribute to similarity), we may eliminate these features from further consideration. The computer is then instructed to locate the third strain having the fewest differences from the nucleus pair in those properties that remain. Any properties in which all three organisms are not identical are again eliminated from consideration, and the computer then locates the strain most similar to the group on the basis of the remaining features. This process is continued until all strains have been added, or until no property remains that is common to all. At each step, this process is equivalent to comparing only a single strain with all remaining strains, since all members of the present group are identical with respect to those properties remaining under consideration. The computer is directed to print only the number of each new strain added and the cumulative difference, d_c (the total differences which have accumulated in the group being formed). This criterion for group formation gradually eliminates from consideration any similarities involving properties not shared by all members of a group. The groups formed are strictly monothetic; all individuals are similar in the same respects at any given level of d_c. The computation is greatly shortened by the smaller number of comparisons necessary, and it remains only to see whether the groups thus formed are meaningful ones.

When the data regarding the test strains were submitted to the computer program just outlined, the results shown in Fig. 2 were obtained. This "stair step" diagram was found to be the most convenient means of examining such data. The machine will be seen to have neatly located subgroup V-A. Furthermore, it is evident that this

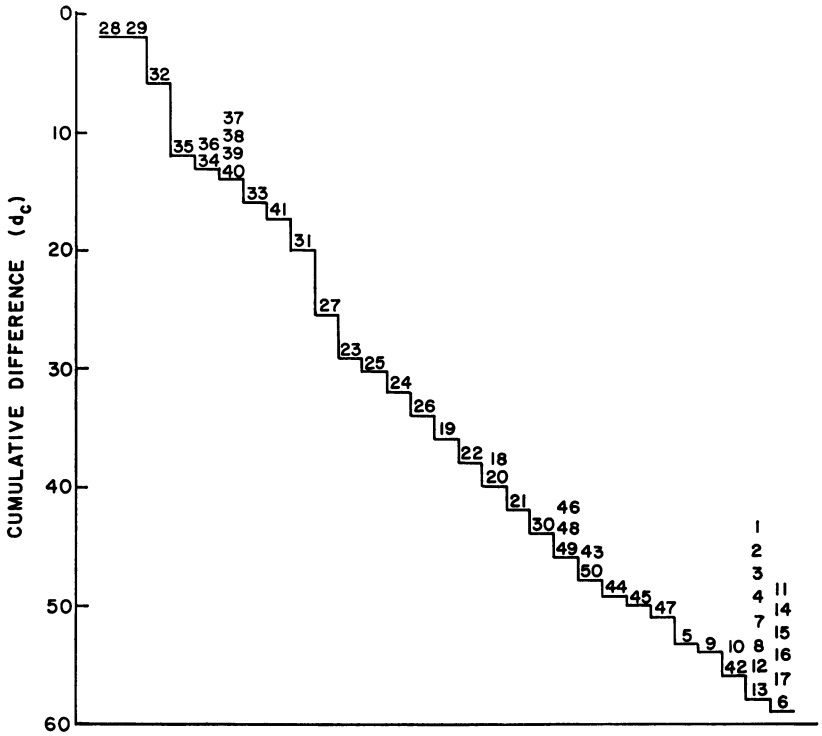

FIG. 2. *Monothetic relationships of the test organisms to a nucleus pair consisting of strains 28 and 29.*

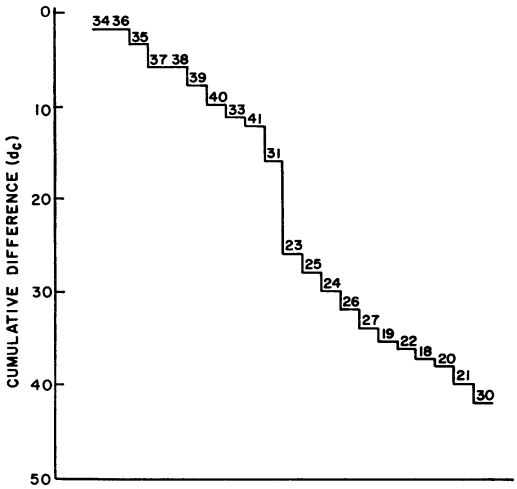

FIG. 3. *Monothetic relationships of the test organisms to a nucleus pair consisting of strains 34 and 36. Strains 28, 29, and 32 were omitted from the computation.*

subgroup is adjacent to the rest of group V, and that the most closely associated groups are III and IV. The rest of the strains are added in a diffuse, haphazard manner at high values of d_c, which is not surprising since members of groups I, II, and VI would hardly be expected to fit the monothetic criteria for enterobacteria.

The machine was then directed to form another group, omitting strains 28, 29, and 32 from the computation. The result is shown in Fig. 3. The nucleus pair this time was 34 and 36 at one difference. (The machine selected the equally similar pair 28 and 29 the first time simply because it encountered them first.) The remainder of group V is now definitely established, and its relationship to the adjacent groups III and IV is shown more clearly. The rest of the data were again diffuse, and are not included in Fig. 3.

An apparent pattern has emerged in these data. First, a group is built around the nucleus pair, with additional strains being added that possess relatively few additional differences. Then, there is an abrupt and appreciable increase in cumulative difference, and the next adjacent group appears in a cluster with little additional increase in cumulative difference. If two adjacent groups are equally near the one formed first, they will appear as a mixture. There is then another relatively abrupt increase in d_c, and the data beyond this point generally are not helpful.

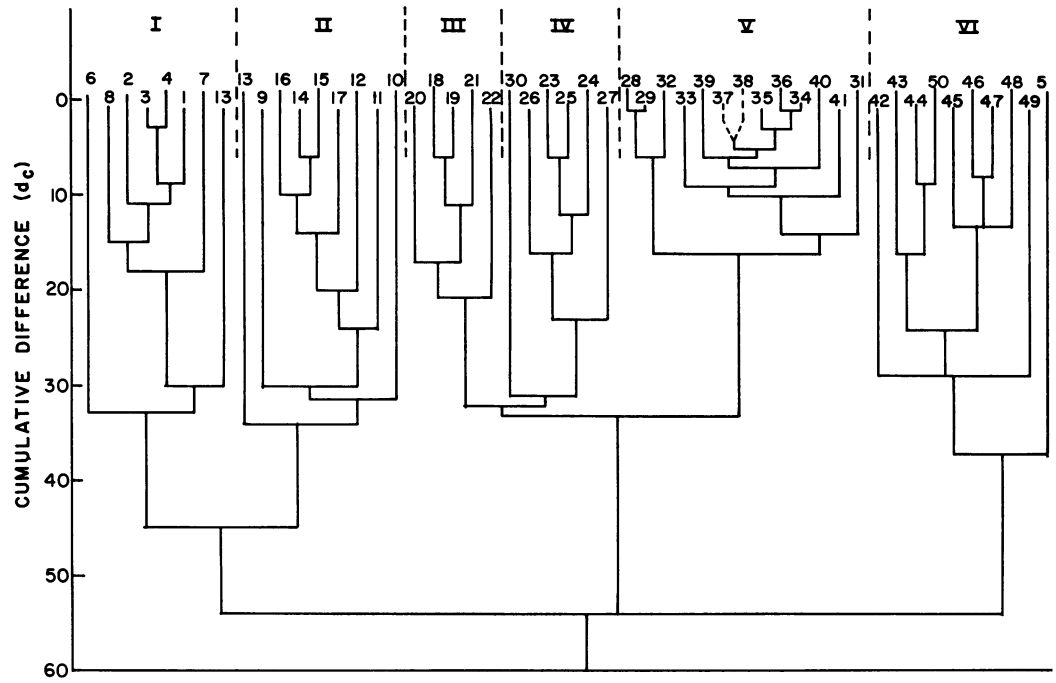

FIG. 4. *Monothetic grouping of the test organisms. Strain numbers and major groups are the same as those shown in Fig. 1.*

Further computations verified this pattern. The next, with all of group V eliminated, began in group I with the nucleus pair 3 and 4 at three differences, and quickly added strains 1, 2, 8, and 7, then 13 (d_c = 27) and 6 (d_c = 33). After another increase in d_c, strains 9 to 17 (group II) were added together over the interval from d_c = 40 to d_c = 49. Strain 5 was not added until later, joining in at d_c = 55, along with a mixture of the remaining organisms. Strains 6 and 13, since they appeared to join group I at high values of d_c, were not eliminated from succeeding computations. Strain 6 never joined any other group, but strain 13 was later added also to group II. In further experiments, group II was formed around the nucleus pair 14 and 15, group III around 18 and 19, group IV around 23 and 25, and groups VI-A and VI-B around the pairs 44 and 50, and 46 and 47.

For arranging these data into a diagram of the relationships among groups, it is necessary only to decide the basis on which the point of confluence between adjacent groups will be designated. Figures 2 and 3 suggest three possible choices: the greatest cumulative difference at which all members of a second group join the first (43 for the confluence of groups III and IV with V), the smallest cumulative difference at which any member of the second group joins the first (25 in this example), or the mean difference at which members of the second group join the first (about 35 in this case). The first of these alternatives was considered most compatible with the monothetic criteria by which the groups had been formed, and was the basis for the diagram shown in Fig. 4.

Comparison of the polythetic grouping in Fig. 1 with the monothetic one in Fig. 4 reveals that the two are surprisingly, and encouragingly, similar. The differences are rather minor—strain 5 no longer appears in group I, and strain 13 is found to fit almost equally well into either group I or II. Re-examination of the *S*-matrix from which Fig. 1 was derived showed that such arrangements would also have been justifiable in the polythetic diagram; strains 5 and 13 had been rather arbitrarily assigned to the groups shown, although they might have fit equally well (or badly) elsewhere. We have pointed out already that such arbitrary decisions are almost inevitable in the construction of polythetic groups. Subgroups V-B and V-C of Fig. 1 were also combined in the monothetic grouping of Fig. 4. The order in which such highly similar strains are added to

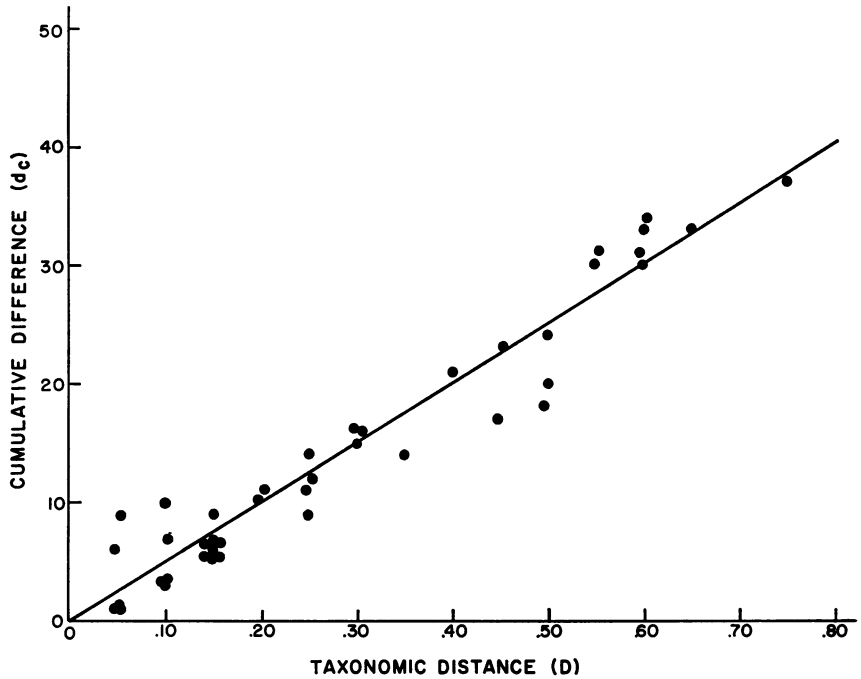

FIG. 5. *Relationship between taxonomic distance* $\left(D = log_2 \frac{1}{S}\right)$ *in the polythetic arrangement of the test organisms from Fig. 1, and cumulative difference in the monothetic arrangement from Fig. 4.*

a monothetic group depends largely on the order in which the strains are encountered during the computation (i.e., the order in which they are listed in the computer "memory"). It is doubtful, in any event, whether subgroupings which differ from one another by only one or two differences have much significance.

The proportions of Fig. 4 are different from those of Fig. 1; the distances between strains within a group, and between the groups themselves, are not the same in the two diagrams. However, it has been suggested on theoretical grounds (Beers and Lockhart, 1962) that distance

$$\left(D = \log_2 \frac{1}{S}\right)$$

rather than similarity might be the most appropriate parameter for expressing quantitative relationships among organisms. The similarity ratios shown in Fig. 1 were converted to their corresponding D values, and the D levels at which strains joined each group in Fig. 1 were plotted against the values of d_c at which the corresponding strains (and groups) were joined together in Fig. 4. The result (Fig. 5) suggests a nearly linear relationship between taxonomic distance in a polythetic arrangement and cumulative difference in the corresponding monothetic arrangement. This function will deviate from linearity at high values of d_c, since the distance between organisms approaches infinity as their similarity approaches zero.

These findings appear to justify the use of monothetic groupings, but the program used does not fully realize the purpose for which it was intended. The process of finding nucleus pairs upon which to build groups still requires comparisons of all possible pairs, and would be prohibitively time-consuming when dealing with large samples. Further, since it is necessary each time to eliminate all strains which have already been assigned to groups as a result of previous computations, an error would be introduced if the investigator incorrectly assigned an organism (such as strain 13 in Fig. 4) to the first group it appeared to join. The computer program therefore was modified further to minimize such problems when dealing with larger samples.

A specific strain, chosen at random, is designated; the computer is instructed to identify the strain most similar to it, and then to construct a group around this artificial nucleus pair, which may be assumed to lie somewhere in one of the groups. This process greatly reduces the computer time necessary to find a nucleus pair. It is not necessary to eliminate any strains from future computations, since the investigator need only specify that each succeeding group be constructed around one of the strains not yet included in a group. A table of random numbers might be used to select the strain to be specified, but it is convenient and probably equally justifiable simply to begin with strain 1 at the first computation and to start succeeding ones with the lowest-numbered strain not yet included in a group. Since no further useful information is obtained once a group has been found and the nature and position of the next adjacent group ascertained, an instruction is also included in the program to stop the computation after a specified number of strains has been added. On the basis of previous knowledge of our groups, we selected 20 strains as the "cut-off" point. An investigator dealing with material not yet analyzed could either predict the size of groups expected or decide this empirically after examining his early results.

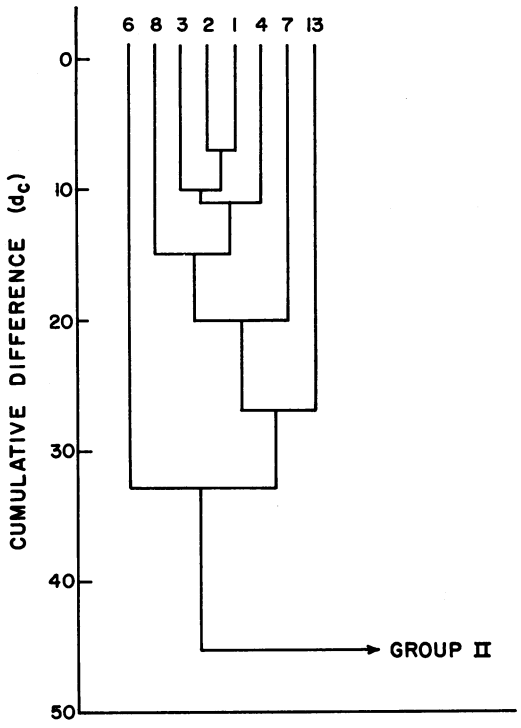

FIG. 6. *Monothetic arrangement of group I of the test organisms, as derived from a computation specifying strain 1 as a member of the nucleus pair.*

When this program was applied, specifying strain 1 as the starting point, the nucleus pair 1 and 2 was located at $d_c = 7$, and strains 3 and 4 were added at $d_c = 10$ and 11, respectively. Strains 8, 7, 13, and 6 were then added exactly as in the previous computation; there was an abrupt increase in d_c, and strains 9 to 17 (group II) were added at between 40 and 49 differences, just as before. A diagram of group I, as it would be constructed from these data (Fig. 6) is identical with group I (Fig. 4), once the level is reached at which the initially specified strain joins the group.

Logically, strain 5 should be specified as the nucleus for the next group. Since strain 5 is known to be a cryptic individual which does not really seem to belong to any group, the result should indicate whether an investigator working with unknown material could be misled if he happened to specify an unusual strain as the nucleus for a group. When this was done, the nucleus pair 5 and 48 was recovered at 16 differences. Strains 42 to 50 (group VI) were added immediately at one or two additional differences each, followed by an abrupt increase to $d_c = 54$, at which point a mixture of groups III and IV began to be added. An investigator would have no trouble concluding that strain 5 is atypical, though perhaps remotely related to a group consisting of strains 42 to 50. A later computation, specifying one of the latter strains as the nucleus, would confirm the existence of group VI and would place strain 5 in its proper relationship to the group.

Specifying strain 9 as the nucleus for the next group produced a similar (though less extreme) result. The pair 9 and 10 was designated at $d_c = 9$; there was a drop to $d_c = 19$, and the rest of group II was added at one or two additional differences each. After an abrupt increase to $d_c = 38$, a mixture of all the other strains began to be added. The presence of group II, and its relationship to other groups, thus was evident. A cautious investigator might begin his next group with strain 11 to assure himself that strains 9 and 10 did not constitute a separate

subgroup, but the existence of group II is already established.

The rest of the experiment was routine. Successive computations beginning, respectively, with strains 18, 23, 31, and 42 revealed the existence of each of the major groups and showed their relationships to each other. Although we were guided by results of earlier experiments, it seems likely that any investigator using this modified approach to monothetic grouping would be led inescapably to these same groups. If confusion arose during analysis of a large number of strains, an additional computation beginning with a doubtful strain should easily clarify the relationships involved.

It is not necessary to resort to this modified program unless the number of strains under study exceeds 100 to 200. The "best pair" technique consumes relatively little time for modest numbers of strains (e.g., less than 1 hr to form the first group from 200 strains in the CYCLONE computer). A larger array of organisms could be divided by the modified method into major groups which, in many cases (Fig. 4 and 6), would not be greatly different from those obtainable by other means and which, for many purposes, would be satisfactory without further analysis. However, studies restricted to the individual groups could then provide good monothetic (or even polythetic) arrangements within the groups. Small samples of representatives from various groups also could be used to confirm intergroup relationships on either a monothetic or polythetic basis.

DISCUSSION

This monothetic technique seems to yield meaningful groupings, and, with the suggested shortcut device of specifying one member of each nucleus pair, it should permit economical computer analysis of quite large numbers of organisms. Our program for the CYCLONE computer, for example, could handle as many as 1,500 strains scored for 200 properties each. Whether the more complex interrelationships among larger numbers of organisms will prove considerably more difficult to interpret in monothetic terms remains to be seen. We are at present undertaking the analysis of a collection of approximately 1,200 strains of streptococci by this technique; the result should provide an adequate test of the applicability of the method to large samples.

Rather curiously, our monothetic groupings were nearly identical to those obtained on a polythetic basis. It is conceivable, though it seems quite unlikely, that this apparent identity is coincidental and unique to the material used in this study. We hope that other investigators, who have at their disposal collections of strains already analyzed on a polythetic basis, will also test monothetic sorting methods and compare the results. Our findings suggest that the polythetic groups of organisms occurring in nature have a monothetic core of common properties. This is not surprising when one considers that previous computer analyses of a wide variety of organisms have produced polythetic groupings which nearly always were quite similar to the hierarchical, essentially monothetic, taxa of existing classifications. The concept of a monothetic core of properties within a group is anticipated by the observations of Liston, Wiebe, and Colwell (1962; *personal communication*), who found through computer analysis that features characteristic of a species are normally distributed among individual strains. At one extreme of this distribution (the "maximum calculated organism") are those strains possessing all the properties ever found among members of the group; at the other extreme (the "minimum calculated organism") are those strains having only the essential properties which are found in all members of the group. Presumably this "minimum" organism is the base line for a group and will be detected as such since its properties constitute the monothetic core. In practice, it is not likely, even with rather large samples, that a strain often will be found so atypical that it possesses only these minimal properties, and the array of mutual properties which serve to define each group may well be nearer the mode than the extreme of the distribution. The nucleus pairs of our groups presumably are those nearest the "maximum" organism; those strains nearest the "minimum" will be included at a greater value of d_c. Thus, an atypical member will not be excluded; the minimal group is operationally redefined as each organism is added, and any property is eliminated from this definition as soon as a strain is located which otherwise fits the group but does not possess that particular property. The key to group formation is the apparent fact that the monothetic core of minimal properties for each group is sufficiently unique to produce a perceptible increase in d_c before the

group fuses with another to form a more inclusive group (or "higher taxon") with an equally characteristic but more limited core of properties. The cumulative difference at each level seems to be proportional to taxonomic distance; this will prove convenient if verified in future work, for such distances have a mathematically defined relationship to the similarity levels, or phenons (Sneath and Sokal, 1962), used in earlier studies.

If future work confirms that monothetic sorting methods are workable and that the groupings thus obtained are essentially the same as those found on a polythetic basis, a number of controversial points in the theoretical aspects of quantitative taxonomy may be found simply to have disappeared. Fears would be allayed that it might one day be necessary to abolish present classifications and to replace them with taxonomies and diagnostic schemes which somehow would take into account the multidimensional polythetic arrangements of organisms which actually seemed to occur in nature. The Adansonian principle, accepted by most quantitative taxonomists, that equal weight must be given all properties of organisms becomes less heretical when one considers that, in monothetic schemes, the "nonessential" properties have been eliminated during group formation, thus in effect weighting those which remain. The essential axiom that no *a priori* weight be assigned to any feature is maintained, for the weighting of key characters occurs only as these are identified and made the criteria for group formation. Finally, the question of whether negative matches should be counted as similarities becomes rather academic; groups are formed by enumerating differences rather than by computing similarities. There may be endless argument as to whether two organisms are similar if neither ferments lactose, but everyone agrees that they are different if one does so and the other does not. Likewise, we need no longer attempt to avoid an upward bias in similarity values by examining data and discarding any properties found to be the same for all the organisms studied. The presence of a number of such mutual properties during monothetic sorting does no particular harm and may even prove useful in establishing the point at which groups fuse together.

Although these results are encouraging, the proposed methods doubtless will require more exhaustive testing before it is known whether their apparent promise will be realized in practice.

ACKNOWLEDGEMENT

We are indebted to June Fisher Smith, who prepared our computer programs and furnished much helpful advice about computer capabilities.

LITERATURE CITED

BEERS, R. J., J. FISHER, S. MEGRAW, AND W. R. LOCKHART. 1962. A comparison of methods for computer taxonomy. J. Gen. Microbiol. **28**:641–652.

BEERS, R. J., AND W. R. LOCKHART. 1962. Experimental methods in computer taxonomy. J. Gen. Microbiol. **28**:633–640.

DEBORD, G. G. 1942. Descriptions of *Mimeae* trib. nov. with three genera and three species and two new species of Neisseria from conjunctivitis and vaginitis. Iowa State Coll. J. Sci. **16**:471–480.

HILL, L. R., M. TURRI, E. GILARDI, AND L. G. SILVESTRI. 1961. Quantitative methods in the systematics of Actinomycetales. II. Giorn. Microbiol. **9**:56–72.

LISTON, J., W. WIEBE, AND R. R. COLWELL. 1962. Preliminary studies of variation within species. Bacteriol. Proc., p. 50.

SILVESTRI, L., M. TURRI, L. R. HILL, AND E. GILARDI. 1962. A quantitative approach to the systematics of actinomycetes based on overall similarity. Symp. Soc. Gen. Microbiol. **12**:333–360.

SNEATH, P. H. A. 1962. The construction of taxonomic groups. Symp. Soc. Gen. Microbiol. **12**:289–332.

SNEATH, P. H. A., AND R. R. SOKAL. 1962. Numerical taxonomy. Nature **193**:855–860.