

INTERNATIONAL BULLETIN OF BACTERIOLOGICAL
NOMENCLATURE AND TAXONOMY

Vol. 12, No. 4

October 15, 1962

pp. 171-179

SOME COMMENTS ON THE ADANSONIAN
TAXONOMIC METHOD

C. D. Floodgate¹

Torry Research Station, Department of Scientific
and Industrial Research, Aberdeen, Scotland

SUMMARY. The meaning of the term "feature" as applied to Adansonian taxonomy is reviewed and the difficulties of adequate definition pointed out. Features which are negative in each of two organisms compared require differentiation between negative results that represent absurdities and negative results that are rational. The meaning of absurdity and nonabsurdity in bacteriology is a function of the bacteriologist's knowledge and experience. It is from the sensible, significant and practical features that random selection of features that may be thought necessary will be made.

For the benefit of workers without adequate access to an electronic computer, a relatively simpler approach to the problem of scoring similarities is outlined.

- - - - -

I. On the Nature of Features

Very frequently, and rightly, authors of papers on bacterial taxonomy emphasize the necessity for a standardised procedure in performing the tests which are used for taxonomic purposes. Unfortunately, little emphasis is laid on the equally important process by which the data thus gained are employed in the course of classification. Nevertheless, the need to treat the data by a standard, or at any rate, a well defined method is as important as the efficient gathering of information itself. In short, taxonomic method is an

¹ Present address: Marine Biology Station, Menai Bridge, Anglesey, N. Wales.

important, but largely neglected, part of classification. It has been argued (Floodgate 1962) that the aim of the bacterial taxonomist is to place bacteria into groups or classes so that the resulting arrangement allows the greatest number of propositions and predictions to be made. (The terms "classes" and "groups" are used here in their everyday sense; not in the nomenclatural meaning laid down by the International Code of Bacterial Nomenclature and Taxonomy (Judicial Commission 1958)). Such a taxonomy can properly be called natural using the word in the same sense as it is used in logic (Gilmour 1937). One interesting suggestion that has been put forward recently is that such a natural, multipurpose taxonomy can be achieved by classifying bacteria by using the concept of over-all similarity (Sneath 1957a).

Since over-all similarity is determined by a mathematical calculation on certain entities called features, the nature of these features is of great importance. Indeed, the problem as to what shall be called a feature and what "weight" shall be given to it is one of the most vigorously debated and interesting points in taxonomic methodology at the present time.

It is, in fact, very difficult satisfactorily to define a "feature." One way of looking at the problem is to say that a feature is a question asked by a taxonomist about a certain organism together with the answer his researches provide. Thus the question "does organism X form acid from glucose?", together with the answer "it does" forms a positive feature, while the question "does organism Y form acid from glucose?" together with the answer "it does not" forms a negative feature. One bacteriological event (or series of events) can provide the answer to a number of questions, i.e. supply a number of features. Consider an organism producing acid in glucose-nutrient broth. The following features may, among others, be taken from the event. Feature 1: does the organism produce acid in glucose-nutrient broth? Yes. Feature 2: does the cotton wool plug inhibit acid formation? No. Of these, feature 1 only would be used by a taxonomist. Feature 2 would be dismissed as trivial, that is, given zero weight, at any rate until such time as an organism is found whose acid production is inhibited by cotton wool. Similarly the task of obtaining answers to some questions would be considered to be too complicated, too

BACTERIOLOGICAL NOMENCLATURE
AND TAXONOMY

costly or too time consuming, and so these features would also be given zero weight. Again, since some questions would be judged to be absurd, some features must be absurd also. These too are then weighted zero. Alternatively, absurdities may be scored as negative features by stating that (say) bacteria do not grow feathers or play football. This is satisfactory if, when computing the S values, only those comparisons involving at least one positive feature are used, but if features which are negative in both organisms are to be included in the computation, a differentiation must be made between negative results that represent absurdities and negative results that are rational. This distinction may not always be easy to make. Presumably the reason why it would be thought absurd to take into account for taxonomic purposes the fact that neither of two bacterial strains are able to form the Gram complex, is because there are some bacteria which do form the Gram complex but none are known to form feathers. Furthermore bacteria are not known to do anything similar to or analogous to growing feathers. In other words, the meaning of absurdity and nonabsurdity in this context is a function of the bacteriologist's knowledge and experience. But there are cases where it is more difficult to decide if a feature is absurd or not. For instance, it appears that no strain of bacteria is known which requires any of the vitamin K group of substances as a growth factor; although vitamin K₂ is found in Mycobacterium tuberculosis and M. phlei, and in addition methylnaphthoquinone derivatives may be involved in the respiratory chain and oxidative phosphorylation of both Escherichia coli and M. phlei (Dam and Søndergaard 1960). A feature then concerned with a requirement of vitamin K for growth might be considered absurd on the grounds that no bacteria are known which require this substance, but sensible on the grounds of the analogy that some other coenzymic moieties have to be supplied to some bacteria as essential nutrilites and that bacteria requiring vitamin K may well be found one day. At all events the important point is that what is considered absurd, trivial or impossible to do depends on the taxonomist's judgment and hence has a subjective element.

Moreover, it is unlikely that any taxonomist will have the time, opportunity or inclination to study all the features which are left when absurdities, trivialities and impossible features have been eliminated. Clearly it is from the

sensible, significant and practical features that any random selection that may be thought necessary will be made. Even so, in the past no attempt has been made at random sampling; in fact it has been usual to select those features which are of particular interest to the bacteriologist who is studying the organisms. In this way a bias, which has been called "accidental bias," (Floodgate 1962) has been introduced into the classification. Paradoxically, in spite of these theoretical difficulties, the selection of characteristics that have been used for Adansonian classification have been sufficiently varied for reasonably satisfactory results to be obtained.

It is also worth noting that if a feature is a question and an answer, then it cannot be a gene or an enzyme. Neither genes nor enzymes can be absurd or trivial as features can. Furthermore, Sneath (1957b) has suggested that for any property which is measured quantitatively, two positive features may be scored, "one allowing expression of the characteristic and one determining the quantity." Such a suggestion does not take into account the genes and enzymes involved in the manifestation of the characteristic, but is only concerned with the most advantageous way of expressing it. Of course, there is a connection between features, genes and enzymes, but it is not a direct relationship of the first being identical with either of the other two.

Another objection that can be raised to Adansonian classification techniques is that they lead to a kind of disguised classification of features rather than a true classification of bacteria because each of the terms in the formula of over-all similarity involves features and not bacteria. This difficulty arises from a misunderstanding of the process of classification. It is only possible to classify bacteria by abstracting one or more qualities from them, and putting into one group all those which possess the quality (or qualities) and putting in another group all those which do not. Now the quality abstracted in this case is over-all similarity which is assessed by a mathematical formula, the individual components of which are numbers of features. To classify features it would be necessary to abstract a quality from them so that, for example, some were classed as morphological features, others nonmorphological features, or physiological features and so on.

One unusual feature of the Adansonian method is that the abstracted quality is continuously variable, and not just

BACTERIOLOGICAL NOMENCLATURE
AND TAXONOMY

present or absent as in the case in most taxonomic schemes. This unusual characteristic may be particularly useful in bacterial taxonomy since bacteria are sometimes thought of as a continuously varying "spectrum" or complex (Cowan 1955). It seems probable that the clusters of organisms in a multidimensional space, that result from the application of Adansonian techniques, will represent the relationship between bacteria more usefully than the rigid hierarchical system that bacteriology inherited from zoology and botany, and originally from Aristotle.

Having then decided which features he intends to weight as zero and those which he intends to give at least unit weight, the taxonomist now has to decide if he intends to add any further weighting. Now since feature A cannot make organism X more or less like organism Y than can feature B, all the features must be equally weighted.

Sometimes it is reasonable to suspect that a bias is being introduced because of our ignorance of bacterial processes. For example, the shape of an organism is probably dependent upon a large number of factors. Unfortunately all too little is known as to what they are. Some questions can be asked, but their answers are unknown, so that no features can be scored for them. The result is that morphology may be represented only by a few features concerned with shape and size whereas every enzyme of the glycolytic cycle can be scored. This suggests a weighting in favour of biochemistry to the detriment of morphology. The lack of balance, however, lies in our knowledge; not in the taxonomic method. Faced with the alternatives of guessing the number of positive features involved or adhering strictly to what is known, it is better to take the latter course.

II. A Rapid Method of Classification

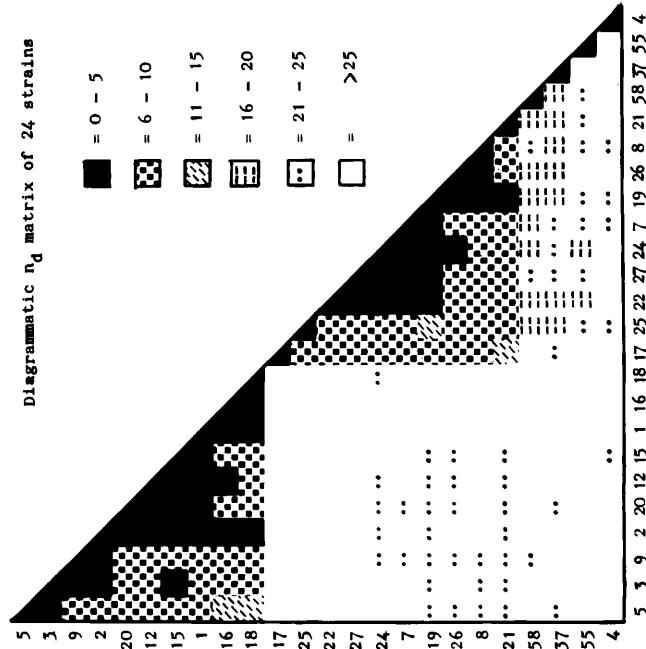
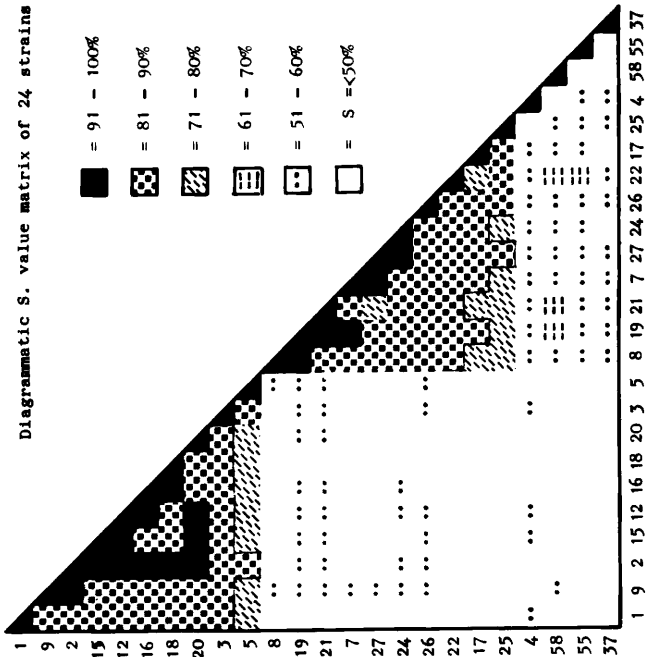
Over-all similarity is an abstract quality and, therefore, cannot be measured directly. When treated mathematically, however, the abstract concept is translated into a form which can be manipulated and used. The problem then arises as to what is the best mathematical expression of the concept. A general discussion of this large and difficult problem will not be attempted in this paper, but an attempt made to solve the practical problem that faces a taxonomist who

does not have immediate access to an electronic computer, ready programmed for taxonomic work; or who has not yet amassed sufficient data to justify a computer run, but who would like to get some idea of the way his data are shaping. The work of calculating and sorting S values of even a moderate number of strains is tedious and time-consuming. A simpler method would be welcome.

Now obviously bacteria which are alike have only a few differences between them, and the greater the number of differences, the more unlike each other the organisms are. It follows that simply counting those features where one organism of a pair is scored positive and the other negative, that is by n_d in Sneath's symbols, will give a measure of the dissimilarity between them. By then arranging the organisms so that those with a low dissimilarity are clustered together, it should be possible to arrive at a classification similar to that obtained by using Sneath's formula of similarity.

Therefore, n_d was calculated for 24 cultures selected at random from a collection of 62 yellow pigmented marine bacteria whose S values to each other were all known. A table of dissimilarity values (n_d) was compiled and then sorted so that those organisms with few differences were close together. The result, together with the corresponding table of similarity values, is shown diagrammatically in Figs. I and II. It will be seen that the bacteria are sorted into the same pleista as before, although the position of each strain relative to the other strains has slightly changed. Unfortunately, the assumption that n_g is always large when n_d is small is not always justified, so that this quick method will not always give the same result as Sneath's formula of similarity. For example, suppose that 4 organisms, A, B, C and D, were examined over 20 tests. A was positive to all the 20 tests, B was positive to tests 1-12 inclusive and negative to the rest, C was positive to tests 1-11 inclusive and negative to the rest while D was positive to tests 1-3 only. The following $S\%$ values are obtained using Sneath's formula:

	A	B	C	D
A	100			
B	60	100		
C	55	92	100	
D	15	25	27	100



FIGURES I and II.

The following are the n_d values:

	A	B	C	D
A	0			
B	8	0		
C	9	1	0	
D	17	9	8	0

In the $S\%$ values A and B are represented as closely similar organisms while C and D are a long way apart. In the n_d table, however, C is as close to D as A is to B.

If one further calculates S values but also includes in n_s , those tests for which both strains in each comparison are negative as well as those that are positive, to get the index S_m (Sneath 1962) the following table can be drawn:

	A	B	C	D
A	100			
B	60	100		
C	55	95	100	
D	15	55	60	100

As might be expected, the resulting table is very similar to that obtained with n_d values, but as pointed out above it is necessary to distinguish between absurd and nonabsurd negative tests. If then there are no blanks in the data, n_d values will be an exact representation of the similarity value which counts negative components (S_m) and is given by the formula

$$n_d = \frac{N_s (1 - S_m)}{S_m}$$

It can be concluded then that classification using n_d values reduces the arithmetic, is worthy of further exploration and can provide a rough guide to the classification when a full computation of S values is difficult to obtain.

- - - - -

The work described in this paper was carried out as part of the programme of the Department of Scientific and Industrial Research.

BACTERIOLOGICAL NOMENCLATURE
AND TAXONOMYReferences

- Cowan, S.T. 1955. Symposium. The principles of microbial classification. Introduction. The philosophy of classification. *Jour. Gen. Microbiol.* 12:314-321.
- Dam, K. and E. S ndergaard. 1960. Vitamin K. *The Enzymes*. Vol. 3. 2nd ed. Edited by Boyer, P.D., H. Lardy, and K. Myrb ck. Academic Press, New York and London.
- Gilmour, J.S.L. 1937. A taxonomic problem. *Nature*, London, 139:1040.
- Floodgate, G.D. 1962. Some remarks on the theoretical aspects of bacterial taxonomy. *Bact. Rev.* (in press).
- Judicial Commission. 1958. International Code of Nomenclature of the Bacteria and Viruses. Edited by R.E. Buchanan, S.T. Cowan, T. Wiken, and W.A. Clark. Iowa State University Press, Ames.
- Sneath, P.H.A. 1957a. Some thoughts on bacterial classification. *Jour. Gen. Microbiol.* 17:184-200.
- _____. 1957b. The application of computers to taxonomy. *Jour. Gen. Microbiol.* 17:201-226.
- _____. 1962. Construction of taxonomic groups Symposium: Microbial classification. Cambridge University Press, Cambridge.

