CHAPTER 3

New Methods in Taxonomy

In this chapter we shall briefly review a number of methods recently introduced into taxonomy. This will include a short section on early developments related to numerical taxonomy.

Some techniques, such as comparative serology, may be regarded as quantitative techniques which yield measures of overall taxonomic similarity. Other techniques which yield many characters in a single technical procedure are here called *polyphenic* methods. These do not themselves yield measures of affinity directly, but they can be made to do so by appropriate mathematical methods such as those described in Chapters 5 and 6. Finally, brief mention is made of newer or less usual characters which may be employed in systematics.

We believe that the results obtained by these methods should be incorporated into the body of taxonomic knowledge; they should, whenever feasible, be included with other more usual characters, such as morphological ones, in analyses for the estimation of affinity by numerical methods. Indeed, their advent makes the use of modern computing techniques a necessity for taxonomy in the future. Just as the problem in Linnaeus' time was the flood of new organisms which were being discovered, so one problem of today is the flood of new characters now available which must be handled in some fashion (Cain, 1959c).

3.1. THE DEVELOPMENT OF NUMERICAL METHODS IN TAXONOMY

The earliest attempts to apply numerical methods to taxonomy date from the rise of biometrics in the last century. As early as 1898 Heincke

used a measure of phenetic distance to distinguish between races of the herring. It was early realized that biometrics could be applied to systematics, but the only important eventual development was that of discriminant functions (Fisher, 1936), which is useful in only one specialized problem of taxonomy.

One of the earliest statistics of interest to systematists was the "Coefficient of Racial Likeness" (Pearson, 1926). It was extensively applied in physical anthropology but does not seem to have been taken up by taxonomists. The C.R.L. was close to being a coefficient of taxonomic similarity, and was subsequently developed by Mahalanobis in the form of the "Generalized Distance" statistic, which is also formally a coefficient of this kind (see Rao, 1948). Anderson and Whitaker (1934) and Anderson and Abbe (1934) employed a similar statistic, which was also equivalent to a diagonal in a multidimensional Euclidean space, and which they called the "General Index." These statistics, though mathematically adequate, did not lead to notable advances for two reasons: (1) they were developed as discriminant functions to aid in the allocation of individuals to existing taxa and not as methods for creating taxa; and (2) as a consequence of (1) these workers selected principally those characters which gave the best discrimination between the taxa, and—since it is usually only necessary to employ a small number of such characters, once they have been found—these methods were in practice based on few rather than many characters. Some of the characters were selected on a priori grounds, and their small number led to instability on repeating the work with other characters. These techniques, with others developed later, have been very widely and successfully used for the study of certain limited taxonomic problems. We may, for example, cite the elegant work of Blackith (1957) on sexual and phase variation in locusts.

Some studies with aims similar to those of numerical taxonomy today should be mentioned. Smirnov (1925) established types on a quantitative basis. His work has been discussed and evaluated from different points of view by Hennig (1950) and Sokal (1962b). Haltenorth (1937) in a study of similarities among eight species of large cats developed a coefficient similar to that of Cain and Harrison (1958). At about the same time Zarapkin (1939) developed a rather elaborate technique called *Divergenzanalyse*, arriving at a quantity analogous to taxonomic distance. The *Affinitätsrechnung* by Schilder and Schilder (1951) is also a computation of taxonomic distance. We believe that these methods did not succeed at the time they were developed because of the entrenched

nature of phylogenetic systematics and since for any substantial number of characters or taxa the methods advocated by these authors presented computational difficulties insurmountable at the time.

Other early methods, but ones specifically intended for taxonomy, are those of Forbes (1933), Anderson and Owenbey (1939), Sturtevant (1939, 1942), Boeke (1942), James (1953), Stallings and Turner (1957), Hudson, Lanzillotti, and Edwards (1959), and Chillcot (1960), based on variations of matching coefficients. These authors did not develop their methods sufficiently to meet the main needs of numerical taxonomy and hesitated to give equal weight to every feature or to employ large numbers of characters. Similar trends can also be seen in the history of bacterial classification, where the earlier reliance on a few morphological or physiological characters has given place to attempts at classification in the Adansonian tradition (see Sneath, 1962).

We believe that one of the main conceptual difficulties which retarded progress in numerical taxonomy was the problem of weighting, even if this was to some extent an illusory one, and the liberating effect of accepting equal weighting can scarcely be overemphasized. The use of many characters is also a prerequisite. In addition, the use of methods of cluster analysis in building the taxonomic hierarchy has been a major advance. It is in these three points that numerical taxonomy chiefly differs from the earlier ideas and methods. No comprehensive review of numerical taxonomy has yet appeared, but we list here some of the literature which discusses or reviews specific topics and methods in this field.

General articles are those of Sokal (1960), Rogers (1961), Sneath (1961, 1962), Sneath and Sokal (1962), Sokal (1963), and a number of reports of meetings—for example, Gilmour (1961a), Williams and Lambert (1961c), and papers in a recent issue of *Systematic Zoology* (Ehrlich, 1961b, c; Russell, 1961; Daly, 1961; Jahn, 1961). The following reviews are restricted to bacteriology: Sneath (1958), Cowan (1959), Silvestri (1960), Brisbane and Rovira (1961), and Floodgate (1962); these have recently been summarized by Sneath (1962).

Below are listed some papers discussing certain topics in numerical taxonomy. Despite a good deal of overlap in subject matter, they can be roughly divided into theoretical and methodological papers. Mainly theoretical discussions are those of Michener and Sokal (1957), Sneath (1957a), Michener (1957), Cain and Harrison (1958), Ehrlich (1958), Sneath (1961), and Sokal (1962b). Papers discussing mainly methods are those of Sneath (1957b), Sokal and Michener (1958), Sokal (1958),

Rogers and Tanimoto (1960), Sokal (1961), Rohlf and Sokal (1962), Sneath (1962), and Sokal and Rohlf (1962).

A comprehensive listing of applications of numerical taxonomy to various systematic groups since 1957 is given in Section 10.1.1.

3.2. COMPARATIVE SEROLOGY

Comparative serology yields measures of taxonomic relations just as numerical taxonomy does. It grew out of medical immunology, from the pioneering work of Nuttall (1901, 1904). It has been greatly expanded by Boyden and his colleagues, and the reviews of Erhardt (1931) and Boyden (1942, 1953, 1958) and on comparative serology of plants by Chester (1937) are valuable. The symposium on Taxonomic Biochemistry, Physiology, and Serology (Leone, 1963) contains much recent material on this field and on other methods discussed in this chapter.

The basic principle, that the proteins of one organism will react strongly with antibodies to the proteins of a very similar organism, but less so in the case of a dissimilar organism, has been shown to have a very wide application. It is being used, despite certain pitfalls, as a court of appeal in doubtful taxonomic cases. The precise meaning of the resemblance shown by this immunological technique will inevitably be somewhat obscure until a great deal more is known about the nature of antigen-antibody reactions. However, it is likely that serological resemblance can best be expressed as an overall similarity of the structure of the relevant proteins, in which a very large number of small differences and resemblances (due to amino acid sequences and perhaps also to folding of the polypeptide chain) are reflected in the reactions with the antisera. This seems more than plausible, since those relatively rare examples in which serology indicates a relationship clearly discordant with accepted taxonomy are commonly due to the close similarity of those widely distributed cell components which give the cross-reaction. Usually carbohydrates and not proteins are responsible for the unexpected serological findings. It is not surprising that two creatures which both produce large amounts of, say, a polymer of glucuronic acid, could show a strong serological cross-reaction because of this.

The many small differences and resemblances in the proteins may perhaps be validly thought of as a large sample of the features of the organism. Modern genetic theory suggests that this is so, since the protein structure is considered to be determined by the fine structure of the

Rogers and Tanimoto (1960), Sokal (1961), Rohlf and Sokal (1962), Sneath (1962), and Sokal and Rohlf (1962).

A comprehensive listing of applications of numerical taxonomy to various systematic groups since 1957 is given in Section 10.1.1.

3.2. COMPARATIVE SEROLOGY

Comparative serology yields measures of taxonomic relations just as numerical taxonomy does. It grew out of medical immunology, from the pioneering work of Nuttall (1901, 1904). It has been greatly expanded by Boyden and his colleagues, and the reviews of Erhardt (1931) and Boyden (1942, 1953, 1958) and on comparative serology of plants by Chester (1937) are valuable. The symposium on Taxonomic Biochemistry, Physiology, and Serology (Leone, 1963) contains much recent material on this field and on other methods discussed in this chapter.

The basic principle, that the proteins of one organism will react strongly with antibodies to the proteins of a very similar organism, but less so in the case of a dissimilar organism, has been shown to have a very wide application. It is being used, despite certain pitfalls, as a court of appeal in doubtful taxonomic cases. The precise meaning of the resemblance shown by this immunological technique will inevitably be somewhat obscure until a great deal more is known about the nature of antigen-antibody reactions. However, it is likely that serological resemblance can best be expressed as an overall similarity of the structure of the relevant proteins, in which a very large number of small differences and resemblances (due to amino acid sequences and perhaps also to folding of the polypeptide chain) are reflected in the reactions with the antisera. This seems more than plausible, since those relatively rare examples in which serology indicates a relationship clearly discordant with accepted taxonomy are commonly due to the close similarity of those widely distributed cell components which give the cross-reaction. Usually carbohydrates and not proteins are responsible for the unexpected serological findings. It is not surprising that two creatures which both produce large amounts of, say, a polymer of glucuronic acid, could show a strong serological cross-reaction because of this.

The many small differences and resemblances in the proteins may perhaps be validly thought of as a large sample of the features of the organism. Modern genetic theory suggests that this is so, since the protein structure is considered to be determined by the fine structure of the

genes. In this context we may cite the comments of Crick (1958) on the application to taxonomy of the details of protein structure.

> Biologists should realize that before long we shall have a subject which might be called "protein taxonomy"—the study of the amino acid sequences of the proteins of an organism and the comparison of them between different species. It can be argued that these sequences are the most delicate expression possible of the phenotype of an organism and that vast amounts of evolutionary information may be hidden away within them.

We may add to this the comment that such studies will allow an evaluation of immunological relationships and that in addition the use of techniques of numerical taxonomy will be essential to the interpretation and analysis of the amino acid sequences. Information on the sequences in the small protein molecules of different species is already becoming available (see Anfinsen, 1959, pp. 142–162), though the data are insufficient to warrant numerical analysis at present. That different proteins of the same animal usually do not cross-react serologically (see Cinader, 1957) suggests that they do not possess extensive common amino acid sequences, and therefore they are presumably the products of separate genes.

When we turn to comparison between organisms, it is probably only a small part of the protein molecule whose structure is fixed because of the functional requirements of the protein; the remainder of the molecule, perhaps the major part, could then vary extensively without greatly affecting the function of the molecule and could therefore reflect intraorganismic diversity as well as the genetic differences between different kinds of organisms. In this connection the relation between serological and genetic characters shown by the work of Irwin and his colleagues (Irwin, 1959) is of interest. Whether these differences in protein structure can be considered a reasonably random sample of the genetic differences of the whole organism is a problem to which no clear answer can yet be given. There is the serious danger that if we study serologically only a single protein (which is desirable for other reasons) we may in effect be studying the fine structure of only one gene, and we have as yet no knowledge of whether the fine structure will be representative of the differences in structure of the other genes. For example, if there had been no evolution of the fine structure of such a gene during the radiating evolution of a group of organisms, the serological studies

would show a disproportionate similarity—even identity—between the evolved taxa. If all birds had retained the same serum proteins that their reptilian ancestor possessed, then avian serum serology would be uniformly uninteresting. And conversely, rapid evolution in such a gene would yield a disproportionate dissimilarity between the descendant forms. Therefore, the value of serology (and of other techniques which study fine chemical structure) does not depend on the "conservatism" of the proteins; extreme conservatism and precociousness are alike fatal. Most usefully, the evolution in protein structure should be generally proportionate to that in other characters. We have no strong reasons to expect that individual genes controlling protein structure will not accumulate changes proportionate to the changes in other genes; if one did find a protein which was aberrant in this respect, it seems unlikely that other proteins would also be aberrant. The use of several proteins would greatly add confidence to the conclusions of comparative serology, so that this might well repay the extra labor involved.

The decoding of fine protein and fine genetic structure in the next few decades will help us to answer whether evolution in protein structure is proportionate to that in other characters, and there are many signs that differences in protein structure do in general parallel other differences. The well-known correspondence between orthodox taxonomy and serology is one piece of evidence. Internal evidence points in the same direction: Boyden, DeFalco, and Gemeroy (1951) showed that there is a very close parallel between the serological resemblances based on serum albumins and those based on serum globulins; recent work on other proteins such as those of egg white, hemoglobins, and red cell antigens shows close agreement with earlier comparative serology and among themselves (Mainardi, 1958a, b; Sibley, 1960); there is high serological similarity between proteins in different stages of the life cycle (Wilhelmi, 1940; Spiegel, 1960).

If the assumption is true—that in general the fine structure of one or two proteins is an adequate and random sample of features—we would expect that each feature of fine structure would in practice contribute equal weight to the serological results. Though our knowledge is scanty on this point, it seems likely that single amino acid differences will generally have approximately equal effects on the serology. Therefore it is not implausible that comparative serology is a quantitative taxonomic method which will generally yield the same conclusions as numerical taxonomy. One might speak of it as a method for estimating affinity in

which the immunized animal acts as the computing machine when it produces the specific antibodies.

There are pitfalls in the use of serology, which Boyden (1942, 1958) has discussed in detail. In brief, one may obtain misleading results because of: (1) the apparently fortuitous occurrence in distant taxa of very similar substances, usually carbohydrates, which dominate the serological reactions; (2) the poor antigenicity of some substances and the variation in discrimination and antibody response given by different animals used for immunization; and (3) distortions due to the use of uncontrolled mixtures of antigens or unsuitable techniques. Despite these, the method is very valuable and can profitably be used not only in deciding difficult cases but also in making exploratory studies. It yields measures of affinity which are formally quite independent from phylogeny, though not all its practitioners have made this distinction in their published work. There is also the need, not appreciated by systematic serologists, for cluster analysis of the serological affinities themselves in order to yield taxonomic groupings, as discussed in Chapter 7.

Several measures of serological similarity have been devised. In the early work of Nuttall (1904) a rough indication of the similarity between two taxa was given by listing the percentage of cross-reactions of a certain strength obtained by serological comparison of a number of species of the two taxa. Boyden (1932) expressed the similarity between two samples as the arithmetic mean of the two reciprocal heterologous titrations after the titers had been first expressed as percentages of the homologous titers; he converted this index M into a measure of serological distance by subtracting it from 100. Chu, Andrewes, and Gledhill (1950) and Mainardi (1958a, 1959b) employed the geometric mean instead of the arithmetic mean, and Mainardi calls this the immunological distance, *I.D.*:

$$I.D. = \sqrt{\frac{Ho_a}{He_a} \times \frac{Ho_b}{He_b}},$$

where *Ho* and *He* stand for the homologous and heterologous titers respectively with the antisera a and b (note that the titers are properly the dilutions of serum giving the end point, not the concentrations). This is an improvement on Boyden's 100 − M index, but it suffers from the defect that immunological identity corresponds to an immunological distance of 1; it is also unduly sensitive to fluctuations in the end point

of the titrations, which are seldom accurate to within one doubling dilution. *I.D.* is better expressed in a logarithmic scale, and with doubling dilutions the appropriate statistic would be $\log_2 I.D.$

This statistic is still inefficient in that each reaction is obtained from the end point only, and a scoring system which makes use of the results with each dilution is better (see Dömök, Szafir, and Farkas, 1954; Race and Sanger, 1950, p. 166). If a score of 1 is given for each maximal reaction and a fractional value for a lesser reaction, the sum of the scores for a titration is equivalent to a logarithmic measure, and with doubling dilutions, $\log_2 I.D.$ will be equivalent to $\frac{1}{2}$ (sum of homologous scores) $-$ $\frac{1}{2}$ (sum of heterologous scores).

The quantitative precipitation reaction has been used extensively since the introduction of nephelometry (measuring turbidity of suspensions). The importance of studying the reaction at successive dilutions is stressed by Bolton, Leone, and Boyden (1948) and Boyden (1942). The usual way of expressing the serological similarity seems to be reasonably adequate. This is to plot the curve of the amount of precipitate obtained at the various dilutions, to calculate the area under the curve, and then to express this area as a percentage of the area obtained from the homologous titration.

Since serology measures so many features at once, it might be argued that it should be given great weight in a taxonomic study. One cannot convert a serological cross-reaction into a single character to be incorporated, with other observed features, into a numerical analysis. The serological results are already a matrix of similarity coefficients. This matrix is not symmetrical, for the reaction—for example, of anti-horse with cow serum, and anti-cow with horse serum—may not be the same in degree. It is possible to break down serological data to give antigenic formulas for the different reacting antigenic factors. This is not simple, but where it can be done these factors can be included like other characters in numerical taxonomic analyses.

Since, for the reasons given above, we have no way yet of knowing what weight should logically be given to a coefficient of serological cross-reaction compared with the weight of an affinity value obtained by numerical analysis, it would be difficult to combine the two coefficients. Comparisons between the two methods are thus of some importance and should be encouraged.

Some less well-known forms of comparative serology may be mentioned. The technique can be applied equally well to many sorts of protein: serum proteins, red corpuscle antigens, egg-white proteins,

proteins of seeds, insect proteins, and so on (see Mainardi, 1958a, b, 1959a, b; Dujarric de la Rivière, Saint-Paul, and Eyquem, 1953; Leone, 1947). Even allergic reactions in man may reflect the taxonomic relationships between the organisms which produce the offending substances (Perlman, 1961). One of the technical difficulties in the past has been the use of mixed antigens such as crude serum; another has been that in microbiology the cell-surface antigens may entirely dominate the serology. Both of these difficulties may be overcome by developments in gel precipitation methods, which should in any event be a valuable adjunct to the usual methods (for example, Gell, Hawkes, and Wright, 1960; and see Ouchterlony, 1958, 1962, and Crowle, 1960). Particularly valuable would be a critical re-examination of the serological relations between plants described by Mez and his colleagues and summarized in the *Königsberger Stammbaum* (Mez and Ziegenspeck, 1926), for, as Chester (1937) points out, their findings have been generally ignored despite the fact that no convincing criticism has been made of much of their work.

3.3. POLYPHENIC METHODS

3.3.1. Chromatography

The use of chromatography, especially two-way paper chromatography, is a relatively new technique in taxonomy, pioneered by Proom and Woiwod (1949) and Micks and Ellis (1952). Like serology, it can be applied to a wide range of tissue fluids and tissue extracts, and a variety of classes of chemical substances may be detected and estimated in a semiquantitative fashion. Wright (1959) and Buzzati-Traverso (1960) have reviewed these methods in zoology, mentioning for example work on insects, fish, molluscs, and echinoderms, and the use of tissue fluids, muscle squashes, and mucus, which were examined for amino acids, pigments, and fluorescent substances. Mainardi (1958a) has studied tissue extracts of birds, and there have been many examples in botany, such as the works by Turner and Alston (1959) and by Pecket (1959); other work has been reviewed by Thompson et al. (1959). In bacteriology it has been used by Cummins and Harris (1956, 1958) to study the relation of the cell-wall composition to taxonomy; Mattick, Cheeseman, Berridge, and Bottazzi (1956) have employed extracts of bacteria, and Proom and Woiwod (1949) have used changes in the culture media for taxonomic purposes.

Unlike serology, the result of these examinations is not a similarity index but is instead a set of data on the occurrence of individual chemical substances, which are perfectly good characters for taxonomic use. The data should obviously be handled by numerical taxonomic methods, and Cheeseman and Berridge (1959) have given an example using the bacterial genus *Lactobacillus*.

It is probably seldom that the number of compounds in the chromatograms will be numerous enough and of a sufficiently wide genetic origin to give an adequate sample of the characters of the organism; therefore, these methods should be used as an adjunct to others. They may be particularly useful for the identification of taxa (as contrasted with their classification).

3.3.2. Electrophoresis

A similar technique for separating and identifying chemical constituents of organisms is electrophoresis on paper or in gels. The principles and difficulties found in chromatography prevail also in this technique. Electrophoresis has been applied to the proteins of insect hemolymph (Brezner and Enns, 1958; van Sande and Karcher, 1960) and to hemoglobin in birds (Mainardi, 1958a; Conterio and Mainardi, 1959). The elegant work of Sibley (1960) on the proteins of egg white of birds is a notable example of this method. Sibley's findings have been of the greatest interest and generally correlate well with other estimates of taxonomic affinity. Electrophoretic patterns should be subjected to numerical analysis; at present the interpretations of the electrophoretic curves have been largely made by eye, and the estimates of similarity may be highly subjective.

3.3.3. Infrared spectroscopy

Another new technique is that of infrared spectroscopy. The pattern of absorption of infrared light by tissues or biological products depends on their chemical composition and can therefore yield many features useful in taxonomy. This technique seems to have been applied mainly to microorganisms, starting with the work of Randall et al. (1951) and Stevenson and Bolduan (1952), although Micks and Benedict (1953) applied it to mosquitoes.

The subject has been excellently reviewed by Norris (1959). While mainly used for identification of bacteria, there is no reason why it should not also be employed for the creation of taxonomic groups and

the assessment of affinity. The development of automatic analyses and comparisons between spectra (Rogoff, 1957) could yield similarity coefficients based on many chemical attributes. As in other methods discussed above, care must be taken that an excess of some single chemical compound does not dominate the spectrum to such an extent as to give patently false estimates of affinity.

3.4. OTHER METHODS

Almost every new technique in biology gives new characters which can be employed in systematics. A few of the more outstanding recent examples are mentioned here. These new characters must be incorporated into the existing body of taxonomic knowledge, and it is our belief that only numerical taxonomy can adequately do this.

Much information on chromosomes is now available from the intensive cytotaxonomic work of the past few decades. It has been the disposition of cytotaxonomists to give this information a very heavy weight (Löve and Löve, 1961), and the equal disposition of others to give it very little—at least when it does not fit previous systematic schemes, as is often the case (see Frahm-Leliveld, 1958). There seems to us to be no warrant for either practice. Clearly, a large number of features can be obtained from cytology, and these can be legitimately included in numerical analyses with the same weight as any other features.

Chemistry is now giving the systematists many new characters, both in plants and animals (see Bate-Smith, 1959, and Florkin, 1949) and even in paleontological material (Abelson, 1957). Newer cytological methods, especially the use of the electron microscope, behavioral studies, ecology, histology (as in Andrew, 1959), and parasitology can all yield a wealth of new material, which we should use as it becomes available. There seems to be no likelihood that any of these newer methods will prove to be an adequate sole basis for taxonomy; to qualify as such, a method would have to reflect accurately the entire genotype. A step in this direction has been taken by estimating the degree to which samples of single-strand DNA from two organisms can form hybrid double-strand DNA (Doty, Marmur, Eigner, and Schildkraut, 1960; Schildkraut, Marmur, and Doty, 1961). This depends on the degree to which the two forms of DNA are similar chemically and are homologous in a genetic sense. Such homology itself depends on the base composition, which can also be a guide to genetic similarity (Lee, Wahl, and Barbu, 1956; Sueoka, 1961).