*J. gen. Microbiol.* (1965), **40**, 143–148
*Printed in Great Britain*



# Microbial Identification*

BY THE LATE K. J. STEEL
*National Collection of Type Cultures, Central Public Health Laboratory, Colindale Avenue, London, N.W. 9*

(*Received* 26 *February* 1965)

SUMMARY

Identification is the practical application of taxonomic knowledge. Dichotomous keys and diagnostic tables form the backbone of everyday identification, but computers will be used in the future. Standardization of methods for characterizing tests, the development of multiple inoculation apparatus, and the use of mass cultures will enable more reliable tests to be carried out and more strains to be tested.

## *Introduction*

In an earlier paper (Steel, 1962) I considered classification to be an art and identification a science; recent developments in numerical taxonomy have lessened this distinction between them. Taxonomy, here equated with Simpson's (1961) definition of systematics, consists of (i) classification, (ii) nomenclature, (iii) identification; and the components should be taken in this order so that communication of results is made possible. Cowan (1965) regards the practice of identification as the utilitarian aspect of systematics or taxonomy and I propose to concentrate on this, and I shall use the terms 'identification' and 'diagnosis' as synonyms.

## *Aims of diagnosis*

The identifier or diagnostician aims to identify a micro-organism accurately in the shortest practical time but when a pathogenic organism is to be identified he is often under pressure for a quick report; however, speed must always be secondary to accuracy. Nungester (1963) stated five objectives in identifying micro-organisms: (1) to determine quickly the susceptibility to antimicrobial drugs, (2) to gain information which may have prognostic value for physicians, (3) to identify pathogens in terms of their potential danger to people in contact with patients, (4) to aid epidemiologists in tracing sources of infections, (5) to accumulate data of interest to those studying infectious diseases. These objectives are primarily for the clinical microbiologist but can be adapted and amended for all concerned with microbial identification, irrespective of their field of study.

Ideally, every specimen for identification should be treated as a research problem, but time and facilities generally preclude this. Consequently the diagnostician

* Based on a paper 'Microbial identification: theory and practice', read at Quebec in August 1964 a few weeks before Dr Steel died, and edited by his colleagues in the National Collection of Type Cultures. As presented to the meeting organized by the Canadian Committee on Culture Collections the paper contained several extracts from Cowan & Steel (1965) and Cowan (1965); these were removed as that material can be read in the original; cross headings have been added.

has developed presumptive tests for those micro-organisms believed to be important, and relegates those he considers unimportant to ill-defined dump-heaps with labels such as 'achromobacter', 'paracolon', or 'non-pathogen'. The magnitude of the diagnostician's task can be gauged by the routine work of the Public Health Laboratory Service of England and Wales which, in 1963, examined two-and-a-half million specimens. Complete identification may not always be warranted and the extent to which identification will be carried will depend on the nature of the specimen and the purpose for which it was submitted.

### *Practice of identification*

The identification of a micro-organism involves comparison of an unknown and a known unit, and eventually giving a name to the former. Both processes depend on adequate information for characterizing the known unit and many ways have been proposed for making such information readily available. Some are based on the use of dichotomous keys and others on diagnostic keys and tables; one recent scheme which uses a computer as an automatic library facility may soon become a practical proposition (Payne, 1963).

*Dichotomous keys.* Skerman (1959) devised a comprehensive key which enables the searcher to place an unknown in its genus, but for further differentiation and diagnosis of species the supplementary keys in *Bergey's Manual* (1957) are required. Difficulties in interpretation of keys arise where strains behave inconsistently in some respect, and to make allowances for the variable reactions given by such strains Manclark & Pickett (1961) developed flow charts in which a strain may then appear in more than one place at the extremities of the chart.

Some characters are almost invariably positive or negative, but characters of such constancy are usually shared by similar organisms and, although they are important in characterizing an organism, have little value in distinguishing it from its neighbours. On the results of a limited number (up to ten) of selected cytological and physiological tests, Cowan & Steel (1961, 1965) found that most organisms encountered in clinical bacteriology could be placed in a genus or group of genera; this constituted the first stage of identification with their diagnostic tables.

*Diagnostic keys and tables.* These have one important limitation, the specimen under examination must belong *a priori* to the group of organisms for which the scheme was devised; otherwise, mis-identification or failure to identify will result. If the diagnostician has ready access to an electronic computer he could use an almost infinite number of characters, but with tables he is restricted by memory or is limited by his ability to recognize similarities and differences when making comparisons simultaneously. These limitations led to the construction of the Determinator and the compilation of tables suitable for use with it (Cowan & Steel, 1960, 1961, 1965).

Another way of comparing the characters of known and unknown micro-organisms is to use punched cards; the characters of the unknown are punched on a card which is then compared, mechanically or by hand, with a series of cards containing the characters of known organisms.

*Diagnosis by computer.* The use of an electronic computer to assist in the identification of micro-organisms is a new venture, but its potential usefulness in bacteriology has already been proved. Fundamentally the computer is provided with a

'microbial memory'. This consists of a table of microbial characters but with the usual plus and minus signs replaced by numerical data which express the constancy of the reactions quantitatively; the names of the organisms and of the reactions are stored in numerical form.

The results of tests on an unknown organism are fed into the computer, which then compares the pattern of these results with the pattern of results for the same tests already held in its 'microbial memory'. The computer selects from its memory those organisms whose behaviour most closely resembles that of the unknown. Finally the computer is programmed to consider the remaining tests for their potential value in differentiating between the suspected organisms.

### *Diagnostic tests*

Of the three approaches to identification (Cowan, 1965), the third or progressive method aims to determine a few fundamental characters so that an isolate can be placed in one genus or small group of genera; other appropriate tests are then carried out so that specific identification can be made. Additional tests for the better identification of a species, variety or biotype may be required.

The diagnostician and the taxonomist are dealing with similar material and both aim to decrease subjective bias and increase objectivity; phylogenetic speculation is out of order for both workers, but only the diagnostician may properly use character weighting. The variable weighting attached to these characters is based largely on experience, but it is likely that the assimilation of data from a wide range of organisms and subsequent computer analysis will, in the future, enable the value of a character to be expressed objectively in a quantitative manner.

The tests used should be those which give the most reproducible results; many reactions are influenced by factors that are difficult to control and, in the absence of agreed standard methods, test methods should be those recommended in published manuals (Cowan & Steel, 1965; Skerman, 1959; Society of American Bacteriologists Committee on Bacteriological Technic, 1957). The need for this is well illustrated among the enteric bacteria where, under the conditions that provide adequate -SH compounds few fail to produce $H_2S$; when tested under other conditions only strong $H_2S$ producers are recorded as positive and the test then has good diagnostic value. Acetoin production is another example of a test whose sensitivity is easily altered; many 'soft rot coliforms' appear to be V-P positive when grown in O'Meara's fumarate broth and Barritt's method is used to detect acetoin.

The characters used should be independent of subjective criteria such as rate of growth, odour and recognition of the finer shades of pigments; when too great a reliance is placed on pigment production the occurrence of non-pigmented organisms (whether natural or artificially induced) poses a dilemma. Colonial morphology will vary with the conditions under which the organism is grown and is seldom of diagnostic value; for example, non-rhizoid variants of the typically rhizoidal *Bacillus mycoides* are common.

*Special tests.* Mention must be made of special tests, usually described for the distinction of similar micro-organisms or for use only within a particular group or genus. The diagnostic and taxonomic value of characters revealed by such tests is diminished when they are known only for a restricted number of micro-organisms. In the National Collection of Type Cultures we try to avoid using special tests for

particular organisms and, in evaluating newly described tests, apply them to a wide range of bacteria; this frequently produces unexpected and interesting results and permits the inclusion of the test in routine identification work.

Ideally perhaps, diagnostic characters should be expressions of the occurrence of a particular gene in a micro-organism; current genetical knowledge does not permit this and we are fortunate indeed to have a few tests which are detecting the presence of a single enzyme (for example the $\beta$-galactosidase test). Similarly, it is not known how many or which genes influence a particular enzyme and it may well be that some of the tests used are each reflecting the same gene. Certainly it does not seem defensible to record the production of both acid and gas from a range of carbohydrates when gas production in all of them revolves around the presence of the same enzyme. Again, it is not unreasonable to assume that when an organism produces acid from maltose it will usually produce it from glucose (*Pseudomonas maltophilia* is a notable exception); the converse is however untrue, since a glucose-attacking organism may or may not attack maltose. Similarly, flagellated organisms may be expected to be motile, but even when non-flagellated some organisms may show gliding motion.

The diagnostician must be fully aware of the importance of adaptation and ecology, and consider a microbial culture as a population rather than a collection of cells; the host range, which is of epidemiological importance, is also an ecological problem, but such factors cannot be satisfactorily tabulated or weighted.

### *Construction of diagnostic schemes*

For many years diagnosticians have distinguished three species within the genus *Brucella* by their dye sensitivity, $H_2S$ production, and agglutination with mono-specific sera; the specialist worker recognizes the same named species but on their oxidative metabolic pattern and bacteriophage sensitivity. Thus what is *B. melitensis* to the diagnostician may be a biotype of *B. abortus* to the specialist in this genus. This example raises the question, need a diagnostic scheme be taxonomically correct? Some schemes are entirely artificial and do not bear any resemblance to accepted classifications; examples are those used by water bacteriologists, based upon IMViC reactions, and some of those intended for identification of the rhizosphere flora. They are useful but often static and may fail to take account of recent developments. Much present-day identification is based upon schemes and keys that are monothetic; such keys are analytical tools rather than classifications but are constructed in the manner of an artificial classification, being based upon a few discriminating criteria that happen to provide a ready means of subdivision. We cannot identify micro-organisms at a glance as can often be done with higher organisms. If we are to identify speedily, which is one of the aims of the diagnostician, we must rely on the determination of selected characters (differentiae or key characters) fewer in number than would be needed for classification. Although identification and the construction of diagnostic schemes logically follows classification, the importance of prominent single characters to the diagnostician has reflected back and such characters have been assumed to be important in constructing taxa.

Does the possibility of producing natural classifications affect identification? A monothetic classification makes the preparation of diagnostic keys easy, but there

is always a risk of mis-identification when an organism which is aberrant in one of the key characters selected is encountered; this difficulty is less likely to arise with tables which are essentially polythetic. Tables have the further advantage that it is possible to obtain some idea of the 'nearest fit' for an aberrant strain.

Polythetic taxa resulting from taxonomic methods based upon overall similarity must be augmented by identification schemes that also reflect the current taxonomic trend. In the creation of natural taxa on phenetic evidence it is possible that a single common character may not be found; this is disturbing for the diagnostician but the dilemma may be resolved by the adoption of a polythetic diagnostic scheme in which the possession of any one character is not essential for identification.

The basic data upon which such schemes are made could be fed into a computer; by suitable programming this information could be retrieved in a form suitable for the preparation of keys. This might best be done by using the discriminant analysis of Fisher (1936) in which each character is given a weighting such that there is the least probability of mis-identifying a specimen taken at random. Another approach is based on the method of Brisbane & Rovira (1961) who calculated the association coefficients for all character-pair combinations and then arranged characters with the highest association coefficients in a dichotomy; this arrangement was unfortunate as dichotomies may lead to error, and the application of cluster analysis to reveal clusters of characters with a high degree of mutual association would be better.

The theory and practice of the preparation of diagnostic keys has been discussed in some detail by Ainsworth (1941), Metcalf (1954), and mathematically by Maccacaro (1958). Hill & Silvestri (1962) and Möller (1962) introduced the concept of probability into the construction of keys; their method permits the ascription of a specimen to a taxon on a probabilistic basis. The choice of tests used for the key was decided on their information content, and for each taxon the mean probability of identification was calculated.

### *The future?*

What does the future hold for microbial identification? About 1880, Koch was working with crude fluid media and gelatin-solidified media but his bacteria grew about as quickly as ours do today. Without forced aeration of liquid cultures, which involves technical problems, we are unlikely greatly to accelerate microbial growth. How then are we to answer the challenge for speedier identification? The use of computers to aid identification has been mentioned. Spot tests and micro-reactions may well play an important role; latex-fixation reagents provide the pathologist with rapid screening tests for agammaglobulinaemia, hypofibrinogenaemia, systemic lupus erythematosus, and other conditions; similar rapid tests would be of value to microbiologists.

The number of organisms submitted for identification increases annually and the examination of larger numbers of strains envisaged in numerical taxonomy puts a considerable burden on the laboratory. If this challenge is to be met, attention must be paid to techniques for the mass handling of cultures as well as to the handling of data resulting from such action. Replica plating, plate inoculators, and apparatus such as that used in colicine and phage typing will play a more important part in the microbiologist's work. However, these methods may be insufficient and thought must be given to ways of relieving the tedium of inocu-

lating scores of cultures into perhaps hundreds of tubes. A method that comes to mind is the application of the fraction collector used in chromatography. Automation has become an accepted tool in clinical biochemical determinations; when the problems of possible cross-contamination and of sterilization of the apparatus have been overcome we may expect automated microbial identification to become a practical proposition. With the possible exception of fluorescent antibody techniques, all diagnostic methods require that the unknown micro-organisms be isolated in pure culture.

These consideratious of future developments do not assume that the diagnostician will be unnecessary; rather they suggest that he should be in a better position to identify more efficiently and accurately, to devise new test methods, and to aid the study of ecological and epidemiological patterns.

## REFERENCES

AINSWORTH, G. C. (1941). A method for characterizing smut fungi exemplified by some British species. *Trans. Brit. mycol. Soc.* **25**, 141.

*Bergey's Manual of Determinative Bacteriology.* (1957). 7th ed. Edited by R. S. Breed, E. G. D. Murray, & N. R. Smith. Baltimore: The Williams and Wilkins Co.

BRISBANE, P. G. & ROVIRA, A. D. (1961). A comparison of methods for classifying rhizosphere bacteria. *J. gen. Microbiol.* **26**, 379.

COWAN, S. T. (1965). Principles and practice of bacterial taxonomy—a forward look. *J. gen. Microbiol.* **39**, 143.

COWAN, S. T. & STEEL, K. J. (1960). A device for the identification of microorganisms. *Lancet*, i, 1172.

COWAN, S. T. & STEEL, K. J. (1961). Diagnostic tables for the common medical bacteria. *J. Hyg., Camb.* **59**, 357.

COWAN, S. T. & STEEL, K. J. (1965). *Manual for the Identification of Medical Bacteria.* London: Cambridge University Press.

FISHER, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugen., Lond.* **7**, 179.

HILL, L. R. & SILVESTRI, L. (1962). Quantitative methods in the systematics of Actinomycetales. III. The taxonomic significance of physiological-biochemical characters and the construction of a diagnostic key. *Giorn. Microbiol.* **10**, 1.

MACCACARO, G. A. (1958). La misura delle informazione contenuta nei criteri di classificazione. *Ann. Microbiol. Enzimol.* **8**, 231.

MANCLARK, C. R. & PICKETT, M. J. (1961). Diagnostic bacteriological screening procedures. *Lab. World*, **12**, 446.

METCALF, P. Z. (1954). The construction of keys. *Systematic Zool.* **3**, 38.

MÖLLER, F. (1962). Quantitative methods in the systematics of Actinomycetales. IV. The theory and application of a probabilistic identification key. *Giorn. Microbiol.* **10**, 29.

NUNGESTER, W. F. (1963). Contributions of microbiology and immunology to medicine and some unfinished business. *Texas Rep. Biol. Med.* **21**, 315.

PAYNE, L. C. (1963). Towards medical automation. *World med. Electron.* **2**, 6.

SIMPSON, G. C. (1961). *Principles of Animal Taxonomy.* New York: Columbia University Press.

SKERMAN, V. B. D. (1959). *A Guide to the Identification of the Genera of Bacteria.* Baltimore: The Williams and Wilkins Co.

Society of American Bacteriologists Committee on Bacteriological Technic. (1957). *Manual of Microbiological Methods.* New York: McGraw-Hill Book Co., Inc.

STEEL, K. J. (1962). The practice of bacterial identification. *Symp. Soc. gen. Microbiol.* **12**, 405.