REVIEW ARTICLE

Strain, clone and species: comments on three basic concepts of bacteriology

L. DIJKSHOORN, B. M. URSING* and J. B. URSING†

Department of Infectious Diseases, Leiden University Medical Centre, PO Box 9600 2300 RC Leiden, The Netherlands. *Department of Biochemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands and †Department of Medical Microbiology, Malmö University Hospital, Lund University, S-205 02 Malmö, Sweden

Different aspects of the terms strain, clone and species are discussed. The term strain is commonly used to denote a pure culture – here called ‘the strain in the taxonomic sense’ – but does also refer to a natural concept closely related to the clone. The term clone on the other hand is used both in a general and in a more restricted sense, the latter indicating a low degree of genetic exchange. The important distinction between the definition of a species and the criteria for a species is emphasised and the main kinds of criteria are considered.

Introduction

Some of the bacteriological terms in daily use are actually far from well defined. This is particularly the case with laboratory jargon, but in scientific papers bacteriological terms may also be used with different meanings, depending on the context and the subjective preferences of the author. The reason for this may be that unambiguous definitions are difficult to formulate, but in some cases there is disagreement about the proper use of a term. This review attempts to analyse the concepts behind the terms strain, clone and species. Although the examples are taken from clinical bacteriology, the discussion will also be relevant to other fields of bacteriology.

The strain

According to the first edition of Bergey's Manual of Systematic Bacteriology 'A strain is made up of the descendants of a single isolation in pure culture and usually is made up of a succession of cultures ultimately derived from an initial single colony' [1]. This is a rather standard definition of the strain as the basic operational unit in bacteriology and it will be referred to below as the strain in the taxonomic sense.

The definition has several interesting implications.

Perhaps the most important is that the strain in this sense is not a natural concept, as the selection of the 'initial single colony' is made by decision and its descendants are kept in artificial culture. Even if a natural environment such as a sterile body site in a laboratory animal is used for cultivation, the inoculation and recovery of a strain are controlled procedures. A merit of the definition is that there can be no doubt as to the identity of the strain, provided that it has been correctly labelled and protected from contamination.

The starting point of the strain is another important issue. 'A single isolation in pure culture' indicates that the strain has been isolated from a particular site at a particular time. This conflicts with another of the meanings of the term strain in clinical microbiology.

Meningococci isolated at the same time from the nasopharynx, blood and cerebrospinal fluid of one and the same patient are almost certainly derived 'from an initial single colony'. The same applies to repeated isolates of a particular streptococcus from the blood of a patient with infective endocarditis. In both these cases the starting point in space and time and also the further development of bacterial spread and growth are unknown.

Thus, it seems necessary to assume that there is a counterpart in nature to the strain in the taxonomic sense. We will simply refer to it here as the strain in nature. The relationship between the two concepts becomes evident if we say that a strain in the taxonomic sense is a sample from a strain in nature. However, if we try to formulate a definition of the

Received 23 Feb 1999; revised version accepted 1 Oct 1999.
Corresponding author: Dr J. B. Ursing (e-mail: jan.ursing@ mikrobiol.mas.lu.se).
Present address: Torparebron, S-277 55 Brösarp, Sweden.
strain in nature analogous to that of the strain in the taxonomic sense, we run into the difficulties which are always encountered when we look for boundaries in nature.

The strain in nature seems to have been in the minds of Tenover et al. when they defined a strain as ‘... an isolate or group of isolates that can be distinguished from other isolates of the same genus and species by phenotypic characteristics or genotypic characteristics or both’ [2]. This definition was, in all essentials, adopted by a European study group on epidemiological markers [3]. Neither of these groups concerned themselves with what we have called here the strain in the taxonomic sense.

A natural strain is rarely ‘pure’. Rather, it has to compete with a number of other strains for its existence. There are of course exceptions, such as the clinical examples given above which concern normally sterile sites of the body.

Bacterial strains – both in the taxonomic sense and in nature – change over time. They undergo mutations and they may lose plasmids. Strains in the taxonomic sense retain the identity given to them even if the phenotype is changed, which happens occasionally to old strains. The strain in nature may also acquire genetic material from other strains in the environment. The decision as to whether two samples represent the same strain in nature becomes merely a matter of opinion. In the two clinical examples above, the common origin of the isolates is inferred from similarity and clinical data. We will return to the problem when we have discussed the term clone.

The clone

This term denotes the progeny of one individual through asexual reproduction [4] and was originally used in botany [5]. It seems to have been introduced rather late in bacteriology; the progenitor here is a bacterial cell. As bacteria reproduce by fission and lack meiosis they are by definition clonal in this original sense. In evolutionary terms, the clone is assumed to be monophyletic, which means not only that all cells have the same ancestor (opposite: polyphyletic), but also that the clone includes all descendants of the progenitor (opposite: paraphyletic) (Fig. 1). The latter prerequisite is merely of theoretical interest, as in bacteriological practice we are always dealing with samples of the clone. And the samples – again – are strains in the taxonomic sense.

The term clone has become useful in epidemiology, particularly in the study of the relationships between isolates representing widely separate geographical areas. A good working definition was offered by Ørskov and Ørskov: ‘... the word clone will be used to denote bacterial cultures isolated independently from different sources, in different locations, and perhaps at different times, but showing so many identical phenotypic and genotypic traits that the most likely explanation for this identity is a common origin’ [6]. Tenover et al. [2] are less specific: ‘Genetically related isolates (clones) are isolates that are indistinguishable from each other by a variety of genetic tests . . . or that are so similar that they are presumed to be derived from a common parent.’

Most of what was said above about the strain in nature is also valid for the clone. The similarity between the two concepts is also evident from the definitions cited above. However, for similar isolates recovered over wide geographical areas, clone rather than strain is used. Even if the clone is basically a natural concept, the strain in the taxonomic sense may be regarded as an artificial clone.

The clonal relationship may be obscured by horizontal transfer of genetic material. This has been studied by methods of population biology, notably multilocus enzyme electrophoresis [7] and lately by multilocus sequence typing [8]. Diversification of clones has been revealed and clonality has then been regarded as a relative property of a bacterial species implying a low frequency of horizontal transfer [9]. The opposite behaviour is panmictic, a term also borrowed from eukaryote biology, which denotes free interbreeding [4]. Intermediate forms also occur. The distinction between clonal – in this restricted sense – and panmictic behaviour is useful, as an epidemiological relationship may not be evident if horizontal gene transfer has significantly changed the genotype and phenotype of a clone.

Thus, no rules can be formulated for the differences
allowed between samples of a clone or the strain in nature. For each situation, several parameters have to be considered: the discriminatory power of the typing systems, connections in space–time, and degree of ‘clonality’ – if known. The latter may differ considerably even between closely related species [9].

The species

The species is the basic category in biological classification. While the definition of the eukaryotic species has focused on the biology of reproduction, the bacterial species has mainly been discussed in terms of similarity between strains (from now on: in the taxonomic sense if not specified).

Despite the many different opinions that have been expressed about the bacterial species, the general concept of species held by most bacteriologists could probably be formulated as follows: A species consists of strains of common origin which are more similar to each other than they are to any other strain. The difference between the two prerequisites of this definition is smaller than it appears to be at first sight, as common origin in bacteria is largely inferred – at least at present – by similarity of semantides, information-rich macromolecules like proteins and nucleic acids. It should further be noted that this is a working definition of an artificial entity. It may be monophyletic or paraphyletic but the definition excludes polyphyly.

If the maximum possible information is used for the comparison of strains, this species concept will imply a maximum achievable predictivity. From a philosophical point of view, this principle has been regarded as the signum of a natural or multi-purpose classification; natural referring here to what may be considered natural to the human mind [10]. Overall similarity is of course the basis of numerical taxonomy [11], and also of the ‘polyphasic’ approach, where the aim is to use as many phenotypic and genotypic properties as possible [12].

Although the species definition – what we mean by a species or want it to be – may not be a major problem, difficulties arrive when we have to delineate species; we then have to look for species criteria – features that can be used for differentiation. The species criteria become important when it has to be decided whether or not a bacterial group should be given a specific epithet, i.e., should formally be regarded as species.

Given consensus about the definition, the criteria may be regarded as a theory that has to be tested against the definition. In our definition, it is taken for granted that in the bacterial world there are discontinuities – more or less clear-cut. In other words, we suppose that there exists in nature a counterpart to the species that we create and name for our convenience. If not, boundaries would be arbitrary, like the wavelengths chosen to define a specific colour, e.g., bluish green. The species in nature would consist of a finite number of clones and ideally would be monophyletic.

The theoretical approach to the criterion problem has developed in an interesting way over time. Up to the middle of this century bacterial taxa were mainly differentiated by morphology and selected sets of biochemical tests. A species should be characterised by features invariably present, a view rooted in Plato’s world of ideas and put into practice in the Aristotelian logic. The term monothetic was coined by Sneth for groups created in this way [13]. As the species criteria were selected by decision, identification became easy and reliable. A recent example is the Kauffmann-White scheme for classification of salmonellae, based on the assumption that the surface antigens could be used as species criteria [14].

The monothetic approach was succeeded by the polythetic, used in numerical taxonomy where members of a group have a maximum of properties in common [13]. However, it is inherent in the polythetic view that no single property is essential for membership of the group. There may be situations where it is difficult to find useful differential characteristics and identification then becomes a matter of probability.

The progress of molecular biology has changed the situation in a profound way. New kinds of bacterial data have become available – cell wall composition, whole cell protein analysis, whole cell fatty acid analysis and gene sequences, to mention a few. The ability to study and compare genomes has promised an insight into evolutionary relationships and a classification based on natural affinity – natural in a different sense from above. This has led to the view that genotypic data are the most important – clearly a retreat from the polythetic approach. A side-effect of this development could be added; taxonomic work has been increasingly undertaken by biochemists rather than microbiologists – chemical compounds are easier to classify than organisms.

The species criterion dominating the last 30 years has been provided by DNA–DNA pairing data. Studies of enterobacteria had shown the existence of discontinuities between groups of data, relatedness values in the range 50–70% being comparatively rare. Moreover, DNA from strains of existing species were >70% related [15]. (It should be noted that current methods do not record re-association between DNA strands unless they are c. 80% similar, meaning that the range of 80–100% similarity corresponds to the range of 0–100% re-association. About 90% similar DNAs show only c. 50% re-association.)

A recommendation that implied that strains with a
DNA–DNA pairing value <70% are members of different species [16] triggered an increasing number of proposals of new species. Recently, a comparison of DNA–DNA pairing data and 16S rRNA similarity data showed that strains with rRNA similarity less than c. 97% generally showed no significant DNA–DNA reassociation and thus belong to different species [17]. Similarity >97% may or may not indicate close relationship. The use of this percentage as a rule of thumb has in many cases made rRNA sequencing replace the more cumbersome DNA–DNA pairing technique for the creation of new species. At present, either the 70% or the 97% rule is used to underpin most proposals for new species.

DNA–DNA pairing values as species criteria have the merit of the monothetic approach: they are easily applied and identification is reliable – although not easily accomplished, as few routine laboratories have access to the technique. However, when a species has been described, species-specific methods based on genomic data may present alternative identification methods. A major shortcoming is that the recommended values were a generalisation of results for a limited spectrum of organisms; they make no provision for bacterial groups related at different DNA–DNA pairing values [18]. Groups that are phenotypically and ecologically different may not merit the status of different species by the DNA–DNA pairing criteria; there are also situations in which genomically distinct groups cannot be differentiated phenotypically [19].

This means that DNA–DNA pairing has failed to provide generally applicable criteria for the bacterial species as defined above. The recommendations cited [16] include the statement that a DNA–DNA pairing group should not be named unless it can be 'differeniated from some property'. The use of the singular form implies a rather modest contribution to a polyphasic approach [16]. Indeed, adherence to the 70% and 97% rules with few or no phenotypical or ecological requirements means that the species definition above is narrowed to suit only those interested in phylogenetic relationships.

In the early stages of our knowledge of horizontal genetic transfer, there was a hope that capability of gene exchange could be used as a species criterion [20]. The idea that the bacterial species – like the species of higher organisms – could be defined as a gene pool has recently been revived [21]. Protein-coding genes may better recognise ecological populations than DNA–DNA pairing data, which should make them better suited as species criteria [22]. Much work lies ahead in this area.

Concluding remarks

We have been looking into the meaning of words. The question of true or false is of course not relevant here. To cite Humpty Dumpty: ‘When I use a word, . . . it means just what I choose it to mean – neither more nor less’ [23]. A definition cannot be wrong but it can be unsuitable if it is contrary to general usage or ambiguous.

The definition of strain in the taxonomic sense is probably endorsed by most bacteriologists. In any event, they have to accept it as inherent in the type strain concept, which is the backbone of current nomenclature praxis [24]. Nevertheless, we also have to live with the term strain as it refers to the strain in nature. Of course, clone could be used here, although this could be confusing – e.g., if it referred to gonococci isolated from the blood and from the genital region of one and the same person, gonococci being the standard example of a bacterium with a 'non-clonal' population structure [9].

For groups of bacteria capable of horizontal genetic transfer, Raven coined the term genospecies [20]. He also suggested taxospecies for a phenotypically circumscribed group and nomenspecies for any named species. Genospecies has since been used in different senses, among others to denote a DNA–DNA pairing group. For this purpose, the terms genomic species and genomospecies are also in use. This slum of species has given rise to much confusion and would be better discarded. There should be no room for more than one kind of species in one and the same classification [19].

The revived monothetic approach to species criteria, leaving phenotypic properties behind, is a source of frustration for bacteriologists in the field [25]. Most of them are interested in bacteria as they are now and do not care much about how they evolved. Their feeling that they have been abandoned by taxonomists is aggravated by the growing number of proposals of new species based on a single strain [26]. Knowledge of intraspecies variation is necessary for all identification work. Nevertheless, there is hope for the future. There is growing support for the polyphasic species concept [27, 28]. It is to be hoped that the prevailing, rather rigid and dogmatic approach to bacterial systematics will be replaced by a more pragmatic view.

References