

TAXONOMIC NOTE

Is characterization of a single isolate sufficient for valid publication of a new genus or species? Proposal to modify Recommendation 30b of the *Bacteriological Code* (1990 Revision)

¹ Department of Veterinary Microbiology, Stigbøjlen 4, The Royal Veterinary and Agricultural University, DK-1870 Frederiksberg, Denmark

² Rødding Tværvej 3, 6630 Rødding, Denmark

³ Institute of Medical Microbiology and Hospital Hygiene, Philipps University, Pilgrimstein 2, D-35037 Marburg, Germany

⁴ Institute of Veterinary Bacteriology, University of Bern, Laenggass-Strasse 122, CH-3012 Bern, Switzerland

Henrik Christensen,¹ Magne Bisgaard,¹ Wilhelm Frederiksen,² Reinier Mutters,³ Peter Kuhnert⁴ and John Elmerdahl Olsen¹

Author for correspondence: Henrik Christensen. Tel: +45 3528 2783. Fax: +45 3528 2757. e-mail: hech@kvl.dk

From 1990 to 2000, the number of published named taxa based upon new isolates at species and genus levels in *International Journal of Systematic and Evolutionary Microbiology*, formerly *International Journal of Systematic Bacteriology*, have increased by approximately four- and sevenfold, respectively. New taxa based upon characterization of only a single isolate remained at around 40% for both categories. The *Bacteriological Code* (1990 Revision) has no recommendations on the number of strains required for definition of new taxa. For a few groups, a minimum number of 5–10 strains has been suggested in minimal standards. Since an exponential increase in new taxa can be expected in the future, the authors discuss problems related to naming new species and genera based upon descriptions of a single isolate and suggest that this practice is re-evaluated. It is proposed that the following should be added to Recommendation 30b of the *Bacteriological Code*: ‘Descriptions should be based on as many strains as possible (minimum five), representing different sources with respect to geography and ecology in order to be well characterized both phenotypically and genotypically, to establish the centre (from which the type strain could be chosen) and the extent of the cluster to be named. In addition, comparative studies should be performed, including reference strains that represent neighbouring species and/or genera, in order to give descriptions that are sufficiently detailed to allow differentiation from these neighbours.’

Keywords: single-isolate description, minimal standards, bacterial nomenclature, *Bacteriological Code*

Present recommendations on the selection of isolates

The rules of bacterial (domains *Archaea* and *Bacteria*) nomenclature are stated in the 1990 Revision of the *International Code of Nomenclature of Bacteria* (in short *Bacteriological Code*) (Lapage *et al.*, 1992); however, the *Bacteriological Code* does not deal with the classification of bacteria, including the criteria for selection of isolates. Recommendation 30b of the *Bacteriological Code* (1990 Revision) states that ‘Before publication of the name and description of a new species, the examination and description should conform at least to the minimal standards (if available) required for the relevant taxon of bacteria’. Minimal

standards were included in the *Bacteriological Code* to ensure the optimal circumscription of new taxa to be named within specific groups of bacteria (Lessel, 1971). Fewer than ten minimal standards have been published by the Subcommittees on Taxonomy under the International Committee on Systematic Bacteriology (ICSB). In these cases, recommendations as to the selection and the number of isolates have been given. It has been recommended to base descriptions of new species on the characterization of not fewer than five, or ten, and preferably as many isolates as possible. These isolates should be independently isolated and obtained from different locations (Dewhirst *et al.*, 2000; Graham *et al.*, 1991; Oren *et al.*, 1997; Ursing *et al.*

al., 1994). It is, in our opinion, necessary that certain principles are followed for all new taxa to be named, including recommendations as to the number of strains studied in order to ensure the quality of taxonomy. Formally, this could be executed through the Editors of the *International Journal of Systematic and Evolutionary Microbiology* (IJSEM) [formerly *International Journal of Systematic Bacteriology* (IJSB)] in the form of instructions to authors and reviewers or through new recommendations included in the Code.

The diversity of isolates used for characterization of a new taxon and for selection of the type strain is of importance to other disciplines which are dependent on the reliability of bacterial taxonomy, such as human and veterinary medicine, agricultural and environmental sciences, as studies of the type strain are often used to draw general conclusions (Young, 2000). Rule 15 of the *Bacteriological Code* requires that a nomenclatural type strain is designated for each named taxon but it is not required that the nomenclatural type strain is the most representative element. With only one strain per taxon (the nomenclatural type), information about character divergence, ecological or geographical distribution of the new named taxon would be unavailable (Sneath, 1977). For example, in medical and veterinary microbiology, description of a new species based upon a single isolate leaves the users with major uncertainty as to diagnosis, clinical significance, subsequent treatment and prophylaxis. Such information is formally irrelevant for naming a new taxon according to the *Bacteriological Code*, but it is convenient that information about ecology and distribution is provided. Such information might easily be published, e.g. in IJSEM, which aims at the publication of all types of systematics.

An increasing number of new genera and species are named and published in IJSEM based on the characterization of a single isolate (Frederiksen *et al.*, 1999). To analyse to what extent this has happened during the past decade and to examine whether this has been a problem for the scientific quality of bacterial taxonomy, we decided to evaluate the publications in IJSB/IJSEM and, based on that, to discuss recommendations as to diversity of collections of strains used for taxonomy.

Numbers of isolates used to characterize new genera and species during 1990–2000

Publications of new, named bacterial genera and species during 1990–2000, but not taxa or new combinations or reclassifications based on strains previously named, were included. The number of strains per new taxon was the maximum number of strains listed in the paper. Only publications where some or all isolates were characterized *de novo* were included. The subspecies category was not included in the analysis, because the number of new named subspecies has been at a steady low level (data not shown). In addition, the

majority of new named subspecies has been suggested on the basis of two or more isolates. Nomenclatural changes effectively published outside IJSB/IJSEM were not included.

Since 1990 the number of bacterial names for new taxa at genus and species levels has increased almost every year and in recent years the numbers have increased in an exponential-like manner (Fig. 1). From 1990 to 2000, the increase has been approximately four- and sevenfold for species and genera, respectively. The proportion of new named taxa based upon characterization of a single isolate has, however, remained steady at around 40% for both genera and species (Fig. 1, insets).

Bacterial population structure and strategy of classification

The publication of a new genus or species based on the characterization of only a single isolate would be justified if all isolates of a new taxon studied had turned out to be pheno- and genotypically identical. The evolution of such a taxon would have required an endemic population with complete genetic isolation (a population is here defined as a genotypic and functional subset of a bacterial taxon). In addition, new variants arising within the population should have been completely suppressed by a strong selection pressure. It cannot be excluded that such genotypically homogeneous taxa with endemic distribution exist. However, the probability is low and this type of population structure is not a likely *a priori* assumption for new groups with unknown population structure. One reason is that a global or disjunctive, rather than an endemic, distribution seems to be the general pattern at the level of genera and species representing both the *Bacteria* and the *Archaea* (Achtman & Pluschke, 1986; Beltran *et al.*, 1988; Enright & Spratt, 1998; Williams *et al.*, 1996). Another property of bacterial populations is that genetic rearrangements frequently occur, often related to environmental changes (Madigan *et al.*, 2000; Wright, 2000). Genetic recombination processes are less studied in the *Archaea* than in the *Bacteria* (Tumbula & Whitman, 1999), but the high phenotypic and evolutionary diversity of representatives within the *Archaea* makes it likely that genetic recombination processes are no less frequent within this group compared to the *Bacteria*. As to the view of population structure, it is unknown how often very narrow and rarely isolated bacterial species exist in nature. However, the *a priori* assumption would be a predominance of species separated in two or more populations (Palys *et al.*, 1997).

Populations and not individuals became the units of zoological systematics and the level for classification already in the 1930s and 1940s (Simpson, 1961). Information about population structure is not available for most new bacterial taxa and an increased research effort in the field of population genetics would be beneficial for taxonomic studies.

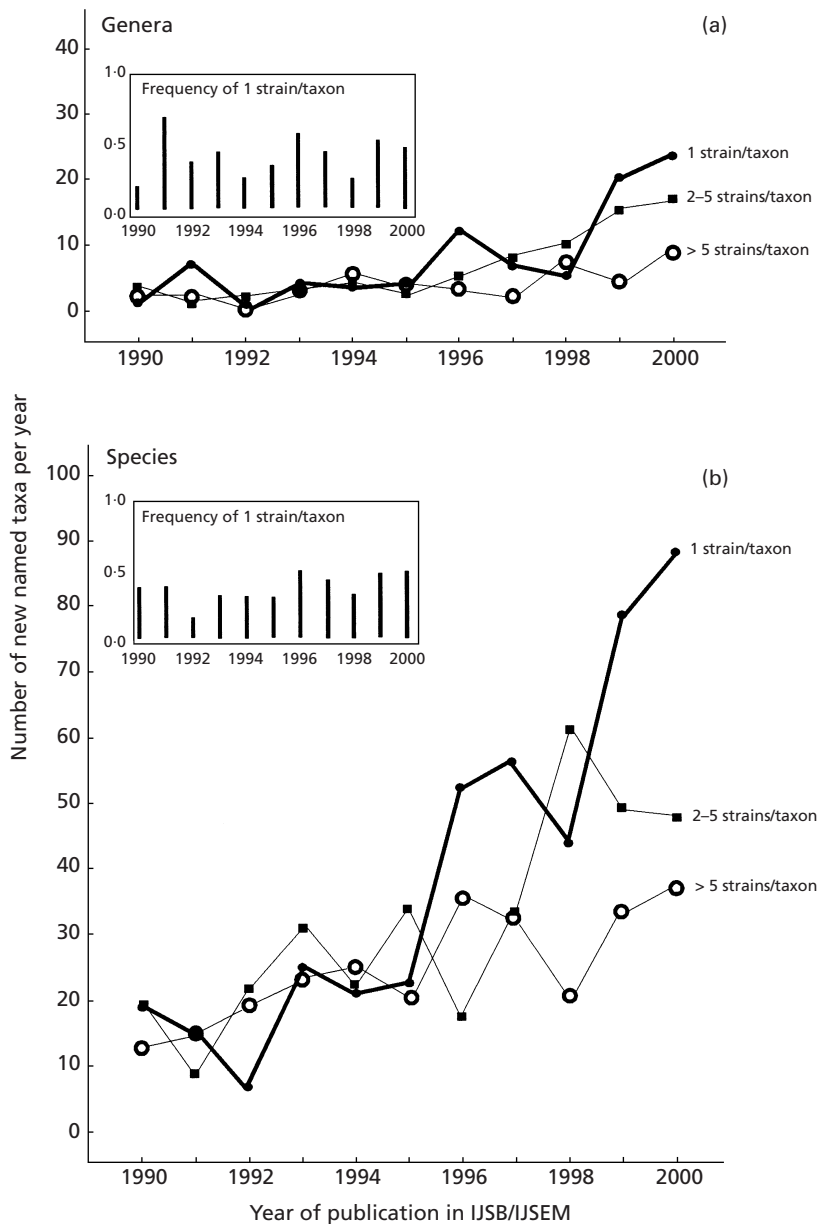


Fig. 1. Distribution of numbers of new named bacterial taxa (domains *Archaea* and *Bacteria*) based upon new isolates and published in IJSB/IJSEM during 1990–2000 in relation to the number of isolates used at the level of genus (a) and species (b). Insets: frequency of genera (a) and species (b) named on the basis of single strains.

Selection of isolates versus diversity in methods used for their characterization

The stringency of criteria used for selection of isolates weighted against the work put into their characterization was discussed by Trüper & Schleifer (1992), who documented good theoretical reasons to include many isolates. Nevertheless, historical facts, according to them, have shown descriptions of most bacterial species during the last century based on the characterization of only a few isolates.

It has been recommended recently 'that new species containing a single strain or new genera containing a single species be validly named only if their genotype and phenotype have been thoroughly and adequately characterized' (Vandamme *et al.*, 1996). While we agree that adequate tests for characterization should

be obligatory in taxonomic investigations, we argue that this, in itself, will not protect against subsequent problems of identification caused by lack of diversity in the assembly of the strains that formed the basis for description and naming of the taxon.

The unusually strong weight given to phylogenetic 16S rRNA sequence analysis compared to other geno- and phenotypic characters in classification is the most obvious reason for the publication of new names at genus and species levels based on characterization of a single isolate. The genotypic definition of a new genus is often based on monophyly and a low sequence similarity as derived by 16S rRNA sequence comparison. The occasionally high 16S rRNA similarity at the species level has been stressed (Fox *et al.*, 1992; Stackebrandt & Goebel, 1994), and thus a new species

is justified from a genotypic point of view if the 16S rRNA sequence analysis indicates the genus level and if DNA–DNA hybridization shows DNA binding of less than 70% between the new taxon and all other known strains (Ursing *et al.*, 1995).

The strong weight that is put on phylogenetic 16S rRNA sequence comparison and DNA–DNA hybridization in classification is in accordance with the recommendations by ICSB that nomenclature should reflect genomic relationships to the greatest extent possible (Wayne *et al.*, 1987). This concept has improved the scientific quality of taxonomic studies, but, in light of the data presented, it seems obvious to discuss inclusion of further recommendations with respect to the selection of isolates for comparative investigations. The problem was recognized for new taxa when the accuracy in relation to identification was found to be low based upon a single isolate representative 16S rRNA sequence deposited in GenBank (Clayton *et al.*, 1995). As an example, for members of *Pasteurellaceae*, 6% of published 16S rRNA sequences are associated with annotation errors such as wrong or old genus or species names (data not shown).

Nomenclatural problems?

A particular problem associated with establishing new genera or species on the basis of characterization of a single or few isolates is the subsequent identification problems due to differences among isolates in diagnostic characters in relation to the type strain, because the type strain does not necessarily represent the median values (Frederiksen *et al.*, 1999; Sneath, 1977; Young, 2000). This again might subsequently necessitate reclassification of new genera or species with resulting nomenclatural changes. Such frequent changes in nomenclature should be avoided (Lapage *et al.*, 1992), especially for microbiologists who depend on an unambiguous name for a pathogen.

A good taxonomic practice is to select the type strain from the most representative population of the habitat of the new taxon (Mayr, 1944). Whether this practice of defining type strains is useful or whether it would be better to assign a 'nominifer' to name-bearing strains was the topic of a recent taxonomic note (Young, 2000). In light of the difficulties already present in handling the type strain concept, it would probably lead to more confusion if both nomenclatural and diagnostic type strains were used. For new taxa to be named, a pragmatic solution would be to select the type strain to be the most representative with respect to characters defining the new taxon among the strains studied. This would not conflict with Rule 15 of the *Bacteriological Code* (1990 Revision), but requires detailed information on diversity within the taxon.

To our knowledge, frequent nomenclatural changes due to initial characterization of too few isolates have not yet caused problems with bacteria of veterinary importance. This statement, however, is based only on

a search through nomenclatural changes since 1984 (Christensen *et al.*, 1999). A 'spot test' in the List of Bacterial Names with Standing in Nomenclature (Euzéby, 1997; updated with each new issue of IJSEM), however, showed that around 1.5% of new species names published based upon characterization of only a single isolate during 1990–2000 were later reclassified. The problem of reclassification shows that microbiologists have to aim for a robust taxonomy, which respects on the one hand our endeavour to describe and investigate microbial life and biodiversity, and on the other hand the necessity to enable a reliable diagnosis.

Proposal

We propose a modification of Recommendation 30b of Rule 30 of the *Bacteriological Code* (1990 Revision) to ensure that strains used for classification are geographically and epidemiologically independent and that proposals for new genera, species and subspecies are not accepted if they are based on only single strains, since the characterization of a single strain is insufficient to represent variability without *a priori* knowledge of population structure. This Recommendation will also serve as an editorial guideline for systematics journals.

In conclusion, it is proposed that the following should be added to Recommendation 30b of the *Bacteriological Code*: 'Descriptions should be based on as many strains as possible (minimum five), representing different sources with respect to geography and ecology in order to be well characterized both phenotypically and genotypically, to establish the centre (from which the type strain could be chosen) and the extent of the cluster to be named. In addition, comparative studies should be performed, including reference strains that represent neighbouring species and/or genera, in order to give descriptions that are sufficiently detailed to allow differentiation from these neighbours.'

Acknowledgements

This work was supported by the Danish Agricultural and Veterinary Research Council through grant 9702797.

References

- Achtman, M. & Pluschke, G. (1986). Clonal analysis of descent and virulence among selected *Escherichia coli*. *Annu Rev Microbiol* **40**, 185–210.
- Beltran, P., Musser, J. M., Helmuth, R. & 8 other authors (1988). Toward a population genetic analysis of *Salmonella*: genetic diversity and relationships among strains of serotypes *S. choleraesuis*, *S. derby*, *S. dublin*, *S. enteritidis*, *S. heidelberg*, *S. infantis*, *S. newport*, and *S. typhimurium*. *Proc Natl Acad Sci USA* **85**, 7753–7757.
- Christensen, H., Aabo, S., Aalbæk, B., Bisgaard, M. & Olsen, J. E. (1999). Nomenclatural changes and new bacterial species of

- veterinary importance. *Dan Vet J* **82**, 108–111 (in Danish with English summary).
- Clayton, R. A., Sutton, G., Hinkle, P. S., Jr, Bult, C. & Fields, C. (1995).** Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. *Int J Syst Bacteriol* **45**, 595–599.
- Dewhurst, F. E., Fox, J. G. & On, S. L. W. (2000).** Recommended minimal standards for describing new species of the genus *Helicobacter*. *Int J Syst Evol Microbiol* **50**, 2231–2237.
- Enright, M. C. & Spratt, B. G. (1998).** A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144**, 3049–3060.
- Euzéby, J. P. (1997).** List of bacterial names with standing in nomenclature: a folder available on the Internet (URL <http://www.bacterio.cict.fr/>). *Int J Syst Bacteriol* **47**, 590–592.
- Fox, G. E., Wisotzkey, J. D. & Jurtschuk, P., Jr (1992).** How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int J Syst Bacteriol* **42**, 166–170.
- Frederiksen, W., Magee, J. T. & Ursing, J. (1999).** Proposed new bacterial taxa and proposed changes of bacterial names published during 1997 and considered to be of interest to medical or veterinary bacteriology. *J Med Microbiol* **48**, 113–116.
- Graham, P. H., Sadowsky, M. J., Keyser, H. H. & 8 other authors (1991).** Proposed minimal standards for the description of new genera and species of root- and stem-nodulating bacteria. *Int J Syst Bacteriol* **41**, 582–587.
- Lapage, S. P., Sneath, P. H. A., Lessel, E. F., Skerman, V. B. D., Seeliger, H. P. R. & Clark, W. A. (editors) (1992).** *International Code of Nomenclature of Bacteria (1990 Revision)*. *Bacteriological Code*. Washington, DC: American Society for Microbiology.
- Lessel, E. F. (1971).** Judicial Commission of the International Committee on Nomenclature of Bacteria. Minutes of meeting, 5 August 1970. *Int J Syst Bacteriol* **21**, 100–103.
- Madigan, M. T., Martinko, J. M. & Parker, J. (2000).** *Biology of Microorganisms*, 9th edn. London: Prentice-Hall.
- Mayr, E. (1944).** *Systematics and the Origin of Species*. New York: Columbia University Press.
- Oren, A., Ventosa, A. & Grant, W. D. (1997).** Proposed minimal standards for description of new taxa in the order *Halo-bacteriales*. *Int J Syst Bacteriol* **47**, 233–238.
- Palys, T., Nakamura, L. K. & Cohan, F. M. (1997).** Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int J Syst Bacteriol* **47**, 1145–1156.
- Simpson, G. G. (1961).** *Principles of Animal Taxonomy*. New York: Columbia University Press.
- Sneath, P. H. A. (1977).** The maintenance of large numbers of strains of microorganisms, and the implications for culture collections. *FEMS Microbiol Lett* **1**, 333–334.
- Stackebrandt, E. & Goebel, B. M. (1994).** Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* **44**, 846–849.
- Trüper, H. G. & Schleifer, K.-H. (1992).** Prokaryote characterization and identification. In *The Prokaryotes*, 2nd edn, vol. 1, pp. 126–148. Edited by A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K.-H. Schleifer. New York: Springer.
- Tumbula, D. L. & Whitman, W. B. (1999).** Genetics of *Methanococcus*: possibilities for functional genomics in *Archaea*. *Mol Microbiol* **33**, 1–7.
- Ursing, J. B., Lior, H. & Owen, R. J. (1994).** Proposal of minimal standards for describing new species of the family *Campylobacteraceae*. *Int J Syst Bacteriol* **44**, 842–845.
- Ursing, J. B., Rosselló-Mora, R. A., García-Valdés, E. & Lalucat, J. (1995).** Taxonomic note: a pragmatic approach to the nomenclature of phenotypically similar genomic groups. *Int J Syst Bacteriol* **45**, 604.
- Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K. & Swings, J. (1996).** Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* **60**, 407–438.
- Wayne, L. G., Brenner, D. J., Colwell, R. R. & 9 other authors (1987).** International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* **37**, 463–464.
- Williams, R. A. D., Smith, K. E., Welch, S. G. & Micallef, J. (1996).** *Thermus oshimai* sp. nov., isolated from hot springs in Portugal, Iceland, and the Azores, and comment on the concept of a limited geographical distribution of *Thermus* species. *Int J Syst Bacteriol* **46**, 403–408.
- Wright, B. E. (2000).** A biochemical mechanism for nonrandom mutations and evolution. *J Bacteriol* **182**, 2993–3001.
- Young, J. M. (2000).** Suggestions for avoiding on-going confusion from the Bacteriological Code. *Int J Syst Evol Microbiol* **50**, 1687–1689.