# Taxonomic Note

# Notes on the characterization of prokaryote strains for taxonomic purposes

B. J. Tindall,[1] R. Rosselló-Móra,[2] H.-J. Busse,[3] W. Ludwig[4] and P. Kämpfer[5]

Correspondence
Brian J. Tindall
bti@dsmz.de

[1]DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Inhoffenstraße 7B, D-38124 Braunschweig, Germany

[2]Grup de Microbiologia Marina, Departament d'Ecologia I Recursos Marins, IMEDEA (CSIC-UIB), C/Miquel Marqués 21, E-07190, Esporles, Spain

[3]Institut für Bakteriologie, Mykologie und Hygiene, University of Veterinary Medicine Vienna, Veterinärplatz 1, A-1210 Vienna, Austria

[4]Lehrstuhl für Mikrobiologie, Technische Universität München, Am Hochanger 4, D-85354 Freising-Weihenstephan, Germany

[5]Institut für Angewandte Mikrobiologie, Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 26-32 (IFZ), D-35392 Giessen, Germany

Taxonomy relies on three key elements: characterization, classification and nomenclature. All three elements are dynamic fields, but each step depends on the one which precedes it. Thus, the nomenclature of a group of organisms depends on the way they are classified, and the classification (among other elements) depends on the information gathered as a result of characterization. While nomenclature is governed by the Bacteriological Code, the classification and characterization of prokaryotes is an area that is not formally regulated and one in which numerous changes have taken place in the last 50 years. The purpose of the present article is to outline the key elements in the way that prokaryotes are characterized, with a view to providing an overview of some of the pitfalls commonly encountered in taxonomic papers.

## INTRODUCTION

The characterization of a strain is a key element in prokaryote systematics. Although various new methodologies have been developed over the past 100 years, both the newer methodologies and those considered to be 'traditional' remain key elements in determining whether a strain belongs to a known taxon or constitutes a novel one. In the case of a known taxon, a selected set of tests may be used to determine whether a strain has been identified as a member of an existing taxon. **However, in the case of a strain or set of strains shown to be novel taxa, they should be characterized as comprehensively as possible. The goal of this characterization is to place them within the hierarchical framework laid down by the Bacteriological Code (1990 revision) (Lapage *et al.*, 1992), as well as to provide a description of the taxa.** Strains should be allocated to species (and/or subspecies), but the nature of the 'species name' (a binomial or combination) dictates that it must also be assigned to a genus. The genus may be either an existing or a novel genus. The Bacteriological Code also recommends that the placement of a genus in a family should be mentioned, and this can be extended to the other hierarchical levels as these become defined. Although this approach may appear novel, with much emphasis currently being placed on the species, the advent of 16S rRNA gene sequencing forces us to choose between primers that are specific for members of the *Archaea* or for members of the *Bacteria*, so the first step in that direction is already routine in many laboratories. However, such a classification system is only possible if strains are comprehensively and properly characterized. **A further key element is the way in which datasets are compared and it is here too that some degree of guidance and a discussion of the potential problems needs to be provided.** In the case of species, various recommendations have been made with respect to the ways in which species may be delineated and it is important to consider these aspects when considering how new strains are to be placed in novel species. However, far too little attention has been paid to the way in which taxa above the rank of species should be characterized and modern prokaryote taxonomy would benefit from paying greater attention to the ranks above those of species.

The purpose of the present article is to deal with current methodologies and to outline how these methods should be best used and implemented. It does not set out to guide

Abbreviations: AFLP, amplified fragment-length polymorphism; DDH, DNA–DNA hybridization; LPS, lipopolysaccharides; MLSA, multilocus sequence analysis; MLST, multilocus sequence typing.

the reader as to how the results should be interpreted, although there are some aspects that are not widely appreciated, where an indication of the value of the dataset may be helpful. The article is divided into genetic (including protein sequence-based methods) and phenotypic methods. The latter include aspects such as cell shape, size, physiological and biochemical tests, as well as methods of chemical analysis of the cell.

## Information on the strains being studied

**In a publication, the information presented on the strains being studied should be as complete as possible. It should include the location from which the strains were isolated (geographical references may also include GPS or latitude/longitude data where possible), the strain designations (including any culture collection accession numbers) and any information that relates to the environment from which the strains were isolated (e.g. pH, salinity, temperature, chemical composition of the environment, depth of the water column or soil profile).** However, care should be taken in using such data to formulate conclusions regarding the ecological significance of the novel strain in the environment without further studies (role of the organism in the environment, cell counts, etc.). Note that it is unacceptable to cite a culture collection number if that strain has not been obtained directly from a collection (e.g. DSM 1234 vs derived from strain DSM 1234 and supplied by x). This information may also be required by culture collections or databases such as GenBank, DDJB or EBI/EMBL and it is the responsibility of the author(s) to make sure that all the information is consistent in order to avoid the accumulation of errors. **Where reference is made to a strain as a type strain, this should be clearly indicated in the publication, database entry (GenBank/DDBJ /strain= '<strain_id>' /note='<type strain of> <Genus> <species>' or EMBL /strain='type strain: <strain_id>'), or the culture collection accession form.** The current policy of GenBank, DDBJ and EBI/EMBL is not to accept new taxonomic names until they have appeared in print. In order for the database staff to update the names, it is important that the strain can be accurately identified by using the associated information submitted to the database. This information will also enable the entry to be linked to the relevant publication.

## Source of the data in the publication

The data presented in a manuscript may be derived from a number of different sources. Original data should be based on results collected using the methods described in the text. So that others may repeat the experimental work described in a paper, the methods used should be given clearly or a reference should made to another publication in which the methods are fully described. When data are extracted from the literature, clear and unambiguous references should be made to the publication(s) in which the data were first published. Data may be supplied by colleagues who do not wish to be co-authors, but have given their consent for the results to be published, or they may be supplied as part of a scientific service. In both cases, the source of the data should be given clearly and unambiguously in order to make it obvious that the data were not collected by the authors. Methods must be given.

## The importance of types

**It should be remembered that in prokaryote taxonomy the inclusion of types is of central importance.** Although not laid down by the Code, since the Code deals with matters of nomenclature and not matters of taxonomy, it is taxonomic common sense to include all type strains that are relevant to a study. Where members of different genera are included and not all species belonging to those genera can be taken into consideration, the inclusion of the type species of the genera under study is of utmost importance. Naturally, the type strain of that type species must be used. It must be emphasized that the type species of the genus is the most important reference organism to which a novel species has to be compared if it is considered to be a member of the same genus. Other species of the same genus may be misclassified and may be subject to reclassification in the future. Similarly, a species placed in a new genus must be compared with species of closely related taxa, which must include the type species of the genera under study. In the case of comparisons across families, the type genera must be included, and by extension, the type strains of the type species of the type genera. Many papers published in the IJSEM seriously violate this elementary taxonomic principle. It should be borne in mind that the types of certain taxa may be organisms that are pathogenic to humans, animals or plants. Not all researchers have access to the facilities or permits for handling such organisms. This should be taken into consideration before research is begun and also by the editors of manuscripts that would normally require comparative studies with such organisms.

## GENETIC-BASED METHODS

Modern prokaryote taxonomy has been strongly influenced by developments in genetic methods. The elucidation of the structure of DNA and the deciphering of the genetic code were major steps forward in biology. The long-term goal of researchers in the 1950s was to be able to gain rapid access to the genome sequence data, a goal that was realized by the mid 1990s. Even prior to the elucidation of whole genome sequences, the widely differing DNA G+C values within the prokaryotes were recognized as having a direct link to codon usage (De Ley, 1968), a topic that is becoming more transparent as more genomes are sequenced. The development of nucleic acid hybridization methods (DNA–DNA and DNA–RNA) has allowed the indirect comparison of gene sequences. The introduction

of the analysis of the 16S rRNA gene by cataloguing (Fox *et al.*, 1977), reverse transcriptase-sequencing (Sanger *et al.*, 1977; Lane *et al.*, 1988) and finally PCR-based gene sequencing (Saiki *et al.*, 1988) has provided a useful working hypothesis on which other elements may be compared when investigating the taxonomy and evolution of prokaryotes. It is realistic to assume that the recognition of novel taxa often centres on the use of 16S rRNA gene-based techniques. Despite the widespread use of 16S rRNA gene sequencing, there are a number of points that need to be considered when evaluating the data. 16S rRNA gene sequences are one of the most widely used datasets. Although there is justification for using 23S rRNA gene sequences, this dataset is currently much smaller and the 16S rRNA gene sequence presently remains the gene sequence of choice. As more whole genome sequences become available, a greater selection of genes with different degrees of resolution will become available.

## Recommendations for sequence analysis of the 16S rRNA gene

### Primary data quality

- **Use (almost) complete sequences.**
- **Check the quality of new and reference sequences: ambiguities, primary and secondary structure consensus violations, overlay of potential cistron heterogeneities (direct PCR fragment sequencing). The quality of the sequences should also be checked against a set of properly aligned sequences (see below). The quality should be checked before sequences are deposited in databases, used in publications or sent to culture collections along with type strain deposits.**
- **Remember that the sequence databases are full of incorrectly labelled and poor quality sequences. There is NO justification for using a sequence of poor quality/dubious origin simply because it is in the database. When characterizing new taxa, a taxonomist should use the best quality data available, including resequencing if appropriate.**

### Multiple alignment

- **The use of expert-maintained seed alignments comprising only high quality sequence data is highly recommended, e.g. ARB (www.arb-home.de), RDP (http://rdp.cme.msu.edu/), SILVA (www.arb-silva.de) and LTP (www.arb-silva.de/projects/living-tree/). The European rRNA database at the Department of Plant Systems Biology, University of Gent, is no longer updated (as of February 2007) but a link is provided to the SILVA website. A limitation is that seed alignments may not be universally compatible with some of the programs used by authors of articles in the IJSEM. Alternatively, high quality sequences that were not previously aligned can be obtained from public databases and aligned using robust multiple alignment programs (e.g.**
  CLUSTAL_X, CLUSTAL W, CLUSTAL X2, CLUSTAL W2, MEGA, T-COFFEE, MUSCLE)**, followed by manual editing.**
- Evaluate the alignment of new and reference sequences with respect to primary and secondary structure consensus. The alignment MUST be made available to the editors and reviewers at the time of manuscript submission – this should be extended to the deposition of the alignments as supplementary data as outlined below. Examples of programs that can display secondary structure for sequence editing are ARB and PHYDIT (http://plaza.snu.ac.kr/~jchun/phydit/), which has been discontinued in favour of jPHYDIT (http://plaza.snu.ac.kr/~jchun/jphydit/index.php; Jeon *et al.*, 2005). RnaViz (De Rijk *et al.*, 2003; http://rnaviz.sourceforge.net/) was also developed specifically to display RNA secondary structure. The Gutell Lab. at the University of Austin Texas maintains the Comparative RNA website and project (http://www.rna.ccbb.utexas.edu/) and this resource is an excellent source of reference secondary RNA structures.
- Pairwise nucleotide sequence similarity values should be calculated with caution. The following should be considered:

  (i) the way in which the similarity was calculated should be given in detail. The following programs are recommended for similarity calculations: ARB, PHYDIT and jPHYDIT. Programs such as CLUSTAL or PHYLIP give the dissimilarity values. EzTaxon (www.eztaxon.org) provides a web-based tool.

  (ii) pairwise similarity values obtained from local alignment programs, such as BLAST and FASTA, should not be used. These programs are primarily useful for database searches.

  (iii) corrected evolutionary distance (e.g. Jukes and Cantor model) should not be used for pairwise similarity calculations.

  (iv) full-length sequences should always be used.

### Assignment to defined taxa

- **Overall sequence similarity values (ranges) might be sufficient if comprehensive high quality reference datasets are available.** There is extensive documented evidence that two strains sharing less than 97 % 16S rRNA gene sequence similarity are not members of the same species (Amann *et al.*, 1992; Collins *et al.*, 1991; Fox *et al.*, 1992; Martínez-Murcia & Collins, 1990; Martínez-Murcia *et al.*, 1992). There have been suggestions that this value should be set higher; however, it should be remembered that the values should be based on almost full-length, high quality sequences. **16S rRNA gene sequences alone do not describe a species, but may provide the first indication that a novel species has been isolated (less than 97 % gene sequence similarity). Where 16S rRNA gene sequence similarity values are more than 97 % (over full pairwise comparisons), other methods such as DNA–DNA hybridization or**

analysis of gene sequences with a greater resolution must be used. **These methods must also be correlated with the characterization based on phenotypic tests.** While the resolving power of the 16S rRNA gene with respect to the delineation of novel species has been extensively debated, less attention has been paid to other taxonomic ranks, such as the genus. 16S rRNA gene sequence similarities between strains belonging to the same genus may vary dramatically from ~97 % in members of the family *Enterobacteriaceae* and some members of the *Actinomycetes*, etc., to much lower levels, for example, in the genera *Deinococcus* and *Hymenobacter*. This variation may be in part due to either an overinflation of the number of genera (high similarity values) or a failure to recognize the presence of additional taxa (low similarity values).

- **At values above ~95 % 16S rRNA gene sequence similarity (over full pairwise comparisons), taxa should be tested by other methods to establish whether separate genera are present. In BOTH cases, the establishment of novel species or new genera (irrespective of the degree of sequence similarity) should be clearly and unambiguously documented.** In the case of species, authors should document differences between the novel species and existing species within the genus. If the genus is too heterogeneous (e.g. members of the genera *Bacillus* or *Clostridium*), then the authors should provide a reasonable scientific justification for using only a limited set of species from the genus. However, the type species of the genus in question must be included. In the case of genera, reasonable attempts should be made to differentiate the new genus from other genera that are 'closely related'. **Many genus descriptions published in recent years do not distinguish the new genus from existing genera, violating an elementary principle in taxonomy.** A 95 % gene sequence similarity 'lower cut-off window' for genera seems reasonable, but in practice it may not be easy to implement. A key issue is often the interpretation of chemotaxonomic data and few scientists now have extensive experience of this procedure. A recent article (Yarza *et al.*, 2008) has also indicated that a 'lower cut-off window' of 95 % 16S rRNA gene sequence similarity may be reasonable, but it must be emphasized that this was based on the evaluation of taxa that are delineated by a broad spectrum of methods and not on 16S rRNA gene sequences alone.

## Graphical representation – phylogenetic trees

- Use (almost) complete high quality sequences.
- Use high quality alignments.
- Do not mix full and partial sequences.
- Never truncate an alignment to the region covered by partial sequences (due to a reduction in the information content) – instead the truncated

sequences should be removed or the organism in question should be resequenced.

- **Establish sequence conservation profiles for the group of interest and higher ranks (Ludwig & Klenk, 2001; Peplies *et al.*, 2008).**
- Use these profiles as filters to recognize branch attraction effects probably resulting from plesiomorphic sites. Alignment columns should be successively removed according to positional variability and the resulting tree topologies compared. Applying no and 50 % filters has proved to be a reasonable compromise. The filters MUST be clearly and unambiguously identified otherwise the work is NOT reproducible. When taxonomic interpretation is based on filters/masks, these MUST be identified in the alignments that accompany the paper.
- **Apply alternative treeing methods (distance matrix, maximum-parsimony, maximum-likelihood methods). The latter two are to be preferred; distance matrix methods should be used for raw screening only (Ludwig & Klenk, 2001; Peplies *et al.*, 2008). The use of different methods of evaluation using the same dataset does not identify effects such as gene transfer or convergent or parallel evolution. Credibility should be given to other datasets (alternative genes and non-sequence based methods) that have been shown to reflect evolution.**
- As many high quality sequence data as possible should be included and balanced datasets (with respect to the phylogenetic spectrum) should be preferred to reduce bias resulting from branch attraction effects (Ludwig & Klenk, 2001; Peplies *et al.*, 2008). New maximum-likelihood software tools are available (PHYML, RAXML), which allow the inclusion of a reasonable number of sequences.
- Remember that all methods are based on underlying assumptions. If the dataset violates those assumptions (which is often difficult to test), then the tree produced may well be the most parsimonious or most likely, but it is not necessarily the 'correct' answer in terms of the true evolution.
- Never use sequences from single distantly related organisms as (an) outgroup(s) (branch attraction!). The group of interest is the ingroup, the remaining taxa included in the tree serve as outgroups for that ingroup.
- Always include sequences of the appropriate types – type strains for species, the type strain of the type species for genera, etc. (this applies to ALL datasets!).

## Presentation of trees

- If properly described (via the methods section and in the legends to the figures), it is acceptable to visualize only the part(s) of trees based on comprehensive datasets (as recommended) that is (are) of importance in the context of the taxonomic assignment or description.
- If properly described (via the methods section and in the legends to the figures), it is acceptable to show

consensus trees visualizing local support (resolved topology) or discrepancies (multifurcation, shading, circles) that have been constructed by applying different treeing methods and parameters.

- Alignments must be made available as supplementary material.
- Trees could be included as supplementary material.
- Details of the settings used for some methods, such a maximum-parsimony or maximum-likelihood analysis, are rarely given – one can only assume that the 'default settings' have been used. The settings used in all tree-construction methods should be clearly indicated.
- Only bootstrap proportions of 70 or higher should be included in the dendrogram.

## Other methods of representing the results of sequence analyses

- The graphical representation (trees) of sequence data is currently a two-dimensional representation.
- Heat maps have also been used to some effect in the analysis of large datasets and are a further development of the shaded similarity matrices that have been commonly used in numerical taxonomy in the past (Lilburn & Garrity, 2004; Sneath & Sokal, 1973).
- Three or multi-dimensional representations of the data may be preferable to two-dimensional representations.

## Other genes

- As indicated above, there is growing interest in the use of other genes with a greater degree of resolution (protein-encoding genes) to resolve issues that are not solved by 16S rRNA gene sequencing.
- **Typically multiple genes are used and the sequences are either compared as individual datasets or combined in concatenated sequences. In such cases, the criteria laid down for the compilation of reliable datasets and expertly maintained alignments are important.**

Sequences of other conserved protein-encoding genes, typically housekeeping genes, can provide a higher resolution than 16S rRNA gene sequences and can complement DNA–DNA relatedness or 16S rRNA gene sequence data for taxonomic analysis at the species level. It can be expected that multilocus sequence analysis (MLSA) and multilocus sequence typing (MLST) schemes will become available through publicly accessible databases. However, it is essential that the species delineations based on MLSA schemes are corroborated with DNA–DNA hybridization data, thereby also validating the MLSA scheme itself.

## Nucleic acid hybridization methods

The technique of DNA–DNA or DNA–RNA hybridization was introduced into prokaryote systematics from the 1960s onwards (Brenner et al., 1967; De Ley et al., 1966; Johnson & Ordal, 1968; McCarthy & Bolton, 1963; Pace & Campbell, 1971; Palleroni et al., 1973). Although there is an extensive collection of literature on DNA–RNA hybridization, this method was eventually replaced by 16S rRNA gene sequencing. Those consulting the older literature should remember that there is generally a good correlation between the results obtained in DNA–RNA experiments and 16S rRNA gene sequencing.

Recommendations for DNA–DNA hybridization (reassociation) experiments. **DNA–DNA hybridization (DDH) is to be performed in cases where the new taxon contains more than a single strain, in order to show that all members of the taxon have a high degree of hybridization among each other. DDH is necessary when strains share more than 97 % 16S rRNA gene sequence similarity.** If the new taxon shows this high degree of similarity to more than one species, DDH should be performed with all relevant type strains to ensure that there is sufficient dissimilarity to support the classification of the strain(s) as a new taxon. DDH can be performed using a number of techniques. Most of them have been validated and show comparable results (Grimont et al., 1980; Rosselló-Móra, 2006).

- **The method (with all modifications) must be clearly cited. If a novel method is used, the authors must provide evidence that the new method produces comparable results to the established methods.**
- **DDH data must be provided for the type strain of the proposed novel species with all other strains of the proposed novel species, and for the new type strain with the type strain(s) of the closest related species with validly published names. At least one reciprocal value for DDH of the closest related species to the new type strain must be provided. Note that when applying spectrophotometric methods carried out in solution, no reciprocal values can be obtained as none of the DNAs are labelled (De Ley et al., 1970). Standard deviations of at least three analyses must be given.**
- DDH is optional for new taxa encompassing a single strain that shares less than 97 % 16S rRNA gene sequence similarity to the closest neighbour.
- In prokaryote systematics, DDH methods have largely concentrated on the use of relative binding ratio (RBR) methods. The results should be given as percentage binding or percentage hybridization. It has been recommended that only DNA from strains sharing a difference in the melting temperature ($\Delta T_m$) of 5 °C or less should be included. Methods using the difference in melting point of the heterologous hybrid compared with the homologous hybrid ($\Delta T_m$) have rarely been used in studies with prokaryotes, but have been widely used in zoology.
- Hybridizations are commonly performed at $T_m - 30$ °C and/or $T_m - 15$ °C (stringent conditions). $T_m$ is calculated on the basis of the known DNA G + C content of the strains to be hybridized. The temperature at which the hybridization is carried out must be clearly stated.

- A DDH value equal to or higher than 70 % has been recommended as a suitable threshold for the definition of members of a species (Brenner, 1973; Johnson, 1973; 1984; Wayne *et al.*, 1987), but this value should not be used as a strict species boundary (Ursing *et al.*, 1995). A single species must embrace all the strains that cannot be clearly discriminated by a stable phenotypic property. A single species may embrace groups of strains with DNA–DNA hybridization values of less than 50 % that are indistinguishable by means of other properties tested at the time.
- A single species may contain several genomic groups or genomovars (Ursing *et al.*, 1995) that may be reclassified as novel species once clear and stable discriminative phenotypic properties are found.

## DNA G+C content

DNA G+C content is still a useful parameter and its relationship to codon usage is clearly illustrated in genome analysis. It is also an important prerequisite for determining the conditions used in DNA–DNA hybridizations. While DNA G+C content may be fairly constant in a group, there are notable exceptions, particularly for obligate intracellular parasites. Deviations from the values obtained for other members of the group may also be an indication of problems with the strain being studied. The methods of choice are now HPLC-based (Ko *et al.*, 1977; Tamaoka & Komagata 1984; Mesbah *et al.*, 1989). Although thermal stability of the native DNA and caesium chloride density-gradient centrifugation are alternative methods, these are now largely of historical interest (see Johnson & Whitman, 2007; De Ley, 1970).

## Use of whole genome sequences

The drop in the price of sequencing whole genomes, together with the technical advances that have been made suggest that routine sequencing of prokaryote genomes will be realistic in the near future. Further advances need to be made in the annotation of the genome. The pilot phase study of the Genomic Encyclopedia of Bacteria and Archaea project (GEBA; http://www.jgi.doe.gov/programs/GEBA/index.html) is currently examining the feasibility of sequencing all available type strains. The current plan is for the information to be published as short genome reports in the open access journal *Standards in Genomics Sciences* (http://standardsingenomics.org). A key issue that remains is the reliable annotation of all genes in a genome since identifying gene homologies (preferably orthologues) is of central importance in taxonomy. In principle, there are three basic approaches:

## Genome indexes

- There are several indexes that are obtained by comparing pairwise genomes that could be used in taxonomy. Noteworthy are the Average Nucleotide Identity (ANI; Konstantinidis *et al.*, 2006) and Maximal Unique Matches (MUM; Deloger *et al.*, 2009) indexes as they have been hypothesized to be able to substitute for DDH. ANI has been demonstrated to correlate with DDH, where the range of ~95–96 % similarity may reflect the current boundary of 70 % DDH similarity (Goris *et al.*, 2007). ANI may substitute for DDH analyses in the near future.

## Gene content

- The rate at which this area can develop is dependent on the number of sequenced genomes that are available. Such studies rely on the distribution of particular genes and also on good annotation of the sequenced genomes (Snel *et al.*, 1999; Huson & Steel, 2004).
- Where representatives of only a small number of taxa have been studied, the coverage may be insufficient to delineate taxa adequately. In the case of taxa where a larger number of strains have been sequenced, more detailed analyses are already possible.
- In the apparent absence of ways of defining or differentiating higher taxa, it is often not clear whether a particular taxon may be over differentiated (e.g. members of the family *Enterobacteriaceae*) or under differentiated (e.g. members of the genus *Deinococcus*). Pre-conceived concepts and nomenclatures can easily influence the interpretation of data.

## Multiple (gene) aligned sequence datasets

- The comparison of large datasets of aligned sequences is usually based on the concatenation of the dataset. In such cases, there appears to be some preference for the selection of genes that strengthen a particular viewpoint. When comparing genomes across all members of the *Bacteria* or between the *Bacteria* and the *Archaea*, the major problem appears to be the limited dataset since there are only a small number of genes that appear to be universally distributed (Dagan & Martin, 2006).
- Different algorithms or the selection of a smaller subset of genes may give different results (Wolf *et al.*, 2001).
- As in the case of 16S rRNA gene sequencing, particular attention must be given to accurately reporting the quality of the alignment, to the annotation of the genome and to any form of secondary restrictions (use of masking/filters) used in the dataset.

There is increasing interest in the use of such multilocus sequencing methods to define species. However, it is clear that estimating the influence of recombination within a gene pool is not an easy task. In some cases strains appear to have clonal origin, whereas others appear to be almost freely combining. The development of the concept of the ecotype indirectly puts emphasis on the process of ecological selection on the organism, where it is the expressed phenotype (at the level of the organism) rather than the gene that is exposed to direct selection (Cohan, 2002).

## Nucleic acid fingerprinting

**In general, these methods provide information at the subspecies and/or strain level.** Examples for these

techniques are: amplified fragment-length polymorphism PCR (AFLP), macrorestriction analysis after pulsed field gel electrophoresis (PFGE), random amplified polymorphic DNA (RAPD) analysis, rep-PCR (repetitive element primed PCR, directed to naturally occurring, highly conserved, repetitive DNA sequences, present in multiple copies in the genomes) including REP-PCR (repetitive extragenic palindromic-PCR), ERIC-PCR (enterobacterial repetitive intergenic consensus sequences-PCR), BOX-PCR (derived from the boxA element), $(GTG)_5$-PCR and ribotyping. A major disadvantage of some of these fingerprint-based methods is that the results are often very difficult to compare when they have been obtained in different laboratories due to the lack of standardization. Exceptions include AFLP and ribotyping since these approaches have been adequately standardized. **DNA fingerprinting methods are of limited value for species descriptions, but when used properly they can be valuable for identification at the species and subspecies levels. These typing techniques can however be very useful to demonstrate whether or not isolates of a novel taxon are members of a clone.**

## PHENOTYPIC CHARACTERIZATION

The phenotype of an organism is typically taken to be restricted to parameters such as cell shape, colony morphology, biochemical tests, pH and temperature optima, etc. However, such a definition is too limited or simplistic and does not reflect the true scope of phenotypic characteristics that can now be easily examined. Typically the chemical composition of the cell (fatty acid, polar lipid and respiratory lipoquinone composition, amino acid composition of the peptidoglycan of the cell wall of Gram-positive bacteria, presence and size of mycolic acids, polyamine pattern, etc.) is included under a separate heading, chemotaxonomy, but it is in essence a part of the phenotypic characterization of an organism. A recently published collection of phenotypic methods has incorporated references to standard chemotaxonomic methods under the category of phenotypic characterization (Tindall *et al.*, 2008). This publication includes, and is an extension of, the work by Smibert & Krieg (1994). This view is adopted here and expands on principles outlined by two ad hoc committees of the ICSB (Wayne *et al.*, 1987; Murray *et al.*, 1990). Other references to methods useful for the phenotypic characterization of prokaryotes include Bascomb (1987), Blazevic & Ederer (1975), Holdeman *et al.* (1977) and Barrow & Feltham (2004). This list is not complete and reference to the appropriate published minimal standards (a list of which is given later) should be made where they are available.

### Morphology, physiology and biochemistry

Examinations of morphological, physiological and bio-chemical properties are the oldest tools for the characterization and classification of prokaryotes. All relevant traits should be listed in the protologue of each taxon being described. At the rank of species and subspecies, more than one representative strain should be included in order to determine which characteristics are stable and which are variable. The inclusion of suitable positive and negative controls should be emphasized, particularly where test conditions differ from those given in standard reference works. The issue of the use of more than a single strain is raised at regular intervals (Christensen *et al.*, 2001; Felis & Dellaglio, 2007). The number of strains to be used is not (and cannot be) laid down by the Bacteriological Code (since it deals with nomenclature and not taxonomy or the characterization of strains/taxa), but this does not diminish the importance of using more than a single strain. Morphological, physiological and biochemical traits should be carefully evaluated to determine those that are common (or even unique) to the taxon in question. Properties that are variable may be included in the protologue since they may indicate subgroupings within the taxon in question.

### Morphological criteria

- **Cell shape and size. The diversity in cell form and its underlying structural basis are now becoming apparent, largely due to a combination of improved methods of visualizing prokaryote cells, structural studies at the subcellular and molecular levels, and genome sequencing. Although light microscopy may be adequate for describing coarse features of the cell, electron microscopic images have a higher resolution.** Depending on the organisms concerned, scanning electron micrographs may be used to show the morphology of whole cells, while transmission electron micrographs may be used to determine the infrastructure of the cell envelope or the presence of cytoplasmic inclusions, internal membrane structures, etc.

- Cell shape and form should be adequately described and supported by appropriate photographs (including electron micrographs). Photographs must be of good quality and features discussed in the text must be evident in the photograph(s). In some cases, cells are simply rods with a uniform size and evenly rounded ends or cocci with a regular spherical shape. In other cases, the ends of rods may have characteristic forms or cocci that have just divided may have relatively flat adjoining surfaces. Spirilla and spirochaetes have different forms, with different amplitudes and wavelength. In extreme cases, cells may form spirals and their diameter may be characteristic. Vibrios may also be of different lengths and have different amplitudes.

- Light microscope photographs of cells are best prepared by covering microscope slides with a thin layer of agar. Such slides may be used either while still wet or air-dried and stored for later analysis. Wet mounts should be examined for areas where cells are well contrasted and form a single layer between the agar surface and the coverslip (Pfennig & Wagener, 1986).

- **Special attention should be given to characteristic features of the cell, such as apolar growth, the presence of stalks, or the fact that they undergo a life cycle (*Caulobacter* spp., *Pirellula* spp., etc.).** The formation of prosthecae, budding or branching are also important characteristics, as is the formation of cell aggregates (e.g. diplococci, sarcinas, chains of cells).

- **Spore formation, endospores or exospores. In the case of endospore-forming organisms, the location of the spores and their size in relation to the size of the vegetative cell should be noted.** In the case of exospores, features such as the formation of sporangia and the shape of the spores should be noted.

- The location of flagella should be noted with care: polar, subpolar or laterally inserted. Their occurrence in one location (tufts of flagella) or distributed at different locations should be noted. Special care should be taken to check whether strains are motile by standard tests or under the light microscope before the strain is examined in an electron microscope since flagella may be lost during fixation.

- **Motility should be described accurately since both the form and speed may be of significance.** Some organisms may be rapidly motile, others only slowly. A distinction should be made between motility caused by flagella and gliding or other forms of motility.

- Intracellular structures visible under the light microscope such as gas vacuoles, sulphur granules or polyhydroxybutyrate granules should be investigated using appropriate methods.

- When grown on solid surfaces, the shape and size of the colony, together with other features of the colony form, should be accurately noted. Standard textbooks cover the diversity of colony shape and size, including features such as the nature of the outer edge, whether the colony is opaque or translucent and whether the colony has any visible additional structures/features. Some strains form rapidly spreading colonies, while others are fairly compact.

- In cases where colonies are not formed, it is still important to record whether cell material is pigmented. Cellular pigments may be soluble either in water or organic solvents. Typical pigments include carotenoids (which turn blue in the presence of concentrated sulphuric acid), flexirubins (which change colour reversibly under acid and alkaline conditions), (bacterio)chlorophyll [which is soluble in organic solvents, but becomes water soluble following saponification; treatment with acid results in the formation of (bacterio)phaeophytin], melanin and pyocyanin (fluorescence at 360 nm, reversible colour change at acid/alkaline pH); this list is not exhaustive.

## Staining behaviour of the cell

- **The Gram stain (Gram, 1884) is one of the oldest forms of staining the cells of prokaryotes and** distinguishes between cell-wall structures that allow a dye complex to be washed out of the cell and those that retain it. The structural basis of this reaction has been described using electron microscopy studies (Beveridge & Davies, 1983; Davies *et al.*, 1983). It should be noted that certain organisms with a defective Gram-positive cell-wall structure may also stain Gram-negative (Wiegel, 1981) and that the reaction to the Gram stain may alter as the cells age. However, such properties are not randomly distributed, but are found in restricted groups of taxa.

- In certain groups, such as members of the class *Halobacteria*, Gram staining may only work after the fixation of the cells. The stability of the cells at low salt concentrations or in the absence of salt may be a more important feature than the Gram stain.

- Acid-fast staining has been used in the past for organisms that typically contain mycolic acids. A detailed examination of the nature of the mycolic acids should be undertaken in such cases.

- Lipophilic cellular inclusions (e.g. polyhydroxybutyric acid) may be stained with Sudan Black.

- Extracellular polymers may be visualized by suspending cells in a solution of India Ink and observing wet mounts under the microscope. Bright haloes around the cells indicate the presence of extracellular material, such as polysaccharides.

## Recommendations for phenotypic analyses

- **The methods (with all modifications) must be clearly given. If novel methods are used, the authors must provide evidence that the new method produces comparable results to established methods.**

- **Physiological and biochemical tests should be carried out in test media and under conditions that are identical or at least comparable. When different methods are used, authors must provide evidence that they give comparable results.**

- Scientists are encouraged to use tests that may not be covered by API (bioMérieux) tests or Biolog (Biolog Inc.) test plates. Such tests may be documented in the original literature or in minimal standards. For certain groups of organisms, the degradation of polymers is an important characteristic and should be tested by appropriate methods. The use of multi-point inoculators and multi-well plates can be helpful for handling large numbers of strains.

- It is important to differentiate between methods that require the demonstration of enzyme activity (Bascomb, 1987; Blazevic & Ederer, 1975), growth on the substrate or metabolic activity without associated growth (Bochner, 1989).

- **As a general recommendation, authors should include strains of the most closely related taxa for comparison in their phenotypic analysis rather than using the data from previously published work. The**

comparisons must include the type strain of the type species of the appropriate genera. Exceptions may be made where laboratories have collected the data used in previous publications themselves using identical methods.

- Accurate citation of the source of the methods is important, particularly where standard reference works may give more than one method. Authors must clearly attribute the tests to the relevant reference work in which the test is described in detail. In cases where more than one method for performing a particular test is given in a reference work, the authors must indicate which of the methods was used, e.g. lipase activity was determined using method 2 listed in Tindall *et al.* (2007).

- All references to methods used should give the source in which the method is described in detail, allowing others to verify the method and the results obtained. In many instances, reference is made to a work that itself also refers to another paper where the method may, or may not, be described in detail. This can make it difficult to locate the original publication and description of the method.

## Chemical characterization

Chemical characterization of the cell (traditionally referred to as chemotaxonomy) deals with various structural elements of the cell including the outer cell layers (peptidoglycan, techoic acids, mycolic acids, etc.), the cell membrane(s) (fatty acids, polar lipids, respiratory lipoquinones, pigments, etc.) or constituents of the cytoplasm (polyamines).

As a general recommendation, authors have to study the chemotaxonomic features of the most closely related taxa for comparison, especially where novel genera are being proposed. The discriminatory power of the different methods may vary (see specific comments below).

**Peptidoglycan.** The discriminatory power of the structure of peptidoglycan is apparently restricted to Gram-positive bacteria, whereas no variation has been reported among members of the phylum *Proteobacteria* and phylum *Bacteroidetes* (Schleifer & Kandler, 1972). Analyses of the peptidoglycan structure can be performed at different levels. The simplest analysis is the determination of the characteristic diamino acid in the cross-linking peptide. Analysis of the peptidoglycan type (A type: cross-linkage of the two peptide side chains via amino acid three of one peptide subunit to amino acid four of the other peptide subunit; B type: cross-linkage of the two peptide side chains via amino acid two of the one peptide subunit to amino acid four of the other peptide subunit), mode of cross-linkage (direct or interpeptide bridge and amino acids in the bridge) and complete amino acid composition provide more detailed information. B type peptidoglycan is characteristic of all genera of the family *Microbacteriaceae*

and members of the *Erysipelothrix/Holdemania* group. All other murein-containing bacteria so far analysed exhibit the A-type peptidoglycan. The amino acid composition of the peptide side chain, including the characteristic diamino acid is usually common to all species of a genus. However, a higher degree of variability has been detected in the mode of cross-linkage between the peptide side chains, which often differ among species of certain genera (e.g. members of the genus *Microbacterium*), but may also differ between strains of a single species as reported for *Micrococcus luteus*. Analysis of the peptidoglycan structure is a requirement for all members of novel Gram-positive genera when they are described and at least the amino acid composition should be provided for every novel Gram-positive species described. In the majority of cases, the amino acid composition of the peptide side chain of the type species of a novel genus may be shared by future species assigned to the genus and hence it should be listed in the genus description as a characteristic trait. The complete composition of the peptidoglycan of a novel species of a recognized genus should be in agreement with the characteristics of the genus and may provide differences in the interpeptide bridge that are useful for differentiation of the novel species from other species. A list of peptidoglycan variations can be found at http://www.dsmz.de/microorganisms/main.php?content_id=35. This system is slightly different to that developed by Schleifer & Kandler (1972).

Pseudomurein, a characteristic peptidoglycan, has been detected in some Gram-positive staining members of the *Archaea*, in which *N*-acetyl muramic acid is replaced by *N*-acetyl talosuronic acid (König *et al.*, 1982; 1983).

**Respiratory lipoquinones.** Respiratory lipoquinones are widely distributed in both anaerobic and aerobic organisms within the *Bacteria* and *Archaea*. These may be divided into two basic structural classes, naphthoquinones and benzoquinones. A third class includes the benzothiophene derivatives, such as *Sulfolobus* quinone and *Caldariella* quinone, but data available to date indicate that this class is restricted to members of the order *Sulfolobales* (Tindall, 2005).

Benzoquinones
- Benzoquinones include ubiquinones, rhodoquinones and plastoquinones.
- **Ubiquinones appear to be restricted to members of the classes *Alphaproteobacteria*, *Gammaproteobacteria* and *Betaproteobacteria* (Tindall, 2005).** Reports to the contrary should be treated with caution. Within members of the class *Alphaproteobacteria*, the predominant ubiquinone is Q-10, although some taxa are known with Q-9 or Q-11. Members of the class *Betaproteobacteria* typically have Q-8, while there is a greater degree of variation within the members of the class *Gammaproteobacteria*. Quinones Q-7 to Q-14 may be found in different taxa within the members of the class

*Gammaproteobacteria*, although there appears to be no evidence of organisms producing only Q-10 as the major ubiquinone in members of this class. Members of the genus *Legionella* typically synthesize more than one ubiquinone, with chain lengths between Q-10 and Q-14 (Collins and Gilbart, 1983; Karr *et al.*, 1982).

- Some members of the classes *Alphaproteobacteria* and *Gammaproteobacteria* may also synthesize menaquinones and rhodoquinones. Their relative concentrations are usually dependent on the substrates utilized and whether a strain is using oxygen as a terminal electron acceptor. Based on the report of Hiraishi *et al.* (1984), there are betaproteobacterial taxa such as *Rhodocyclus tenuis* and *Rubrivivax gelatinosus* that, in addition to the synthesis of ubiquinones, also synthesize menaquinone 8 (MK-8).

- There is evidence for side chain modifications of the ubiquinones in some methanotrophs (Collins, 1994).

## Naphthoquinones

- **The vast majority of evolutionary lineages within the *Bacteria* and *Archaea* that are known to produce respiratory lipoquinones synthesize naphthoquinone (menaquinone) derivatives.** These include menaquinones, demethylmenaquinones, monomethylmenaquinones, dimethylmenaquinones and menathioquinones (Collins, 1985; Collins, 1994; Tindall, 2005).

- The side chain lengths recorded to date of known respiratory lipoquinones range from 5 to 15 isoprenoid units.

- The isoprenoid side chains are often fully unsaturated, but certain groups have characteristic, stereospecific patterns of hydrogenation (saturation) in the side chain. Typically certain members of the class *Deltaproteobacteria* and members of the class *Halobacteria* synthesize menaquinones in which one point of saturation occurs at the aliphatic end of the side chain. The isoprenoid side chains in members of the class *Halobacteria* are octaprenyl; in members of the class *Deltaproteobacteria* they are penta- to heptaprenyl (Collins, 1994).

- In the high G + C Gram-positives, the isoprenoid side chains show different patterns of hydrogenation. In some cases, the first position of saturation is at the second isoprenoid double bond (from the naphthoquinone ring), followed by position three. In other cases, the second position of saturation is at the aliphatic end of the chain. There is currently a single report of a tetra-hydrogenated menaquinone in which saturation is only present at the aliphatic end (Collins, 1994).

- Terminal ring structures have also been reported (Collins, 1994).

- Fully saturated side chains have been reported in certain taxa within the *Archaea*.

- The chain length, degree of saturation (if any) and position of saturation closely follows groupings as determined by 16S rRNA gene sequence analysis (Tindall, 2005).

- It would appear that there are no differences in the qualitative variation between strains within a species. Demethyl-, monomethyl-, dimethylmenaquinones and menathioquinones also have characteristic distribution patterns.

## Methodological aspects

- Respiratory quinones should be extracted from cell material taking care to avoid extremes of pH, exposure to strong light or highly oxidizing conditions.

- Pre-screening using TLC will identify the various respiratory lipoquinone classes: ubiquinones, rhodoquinones, menaquinones, *Sulfolobus-/Caldariella*-quinones, menathioquinones and demethyl-, monomethyl- and dimethylmenaquinones.

- Separation of simple quinone mixtures may be undertaken using reverse phase TLC, but a properly calibrated reverse phase HPLC column provides greater reproducibility and allows more accurate quantification (Collins, 1994; Tindall, 2005).

- If quinones are detected that cannot be identified, information such as UV-visible spectroscopy results and accurate reports of the behaviour of the quinones in TLC and HPLC systems would be helpful. Full structural identification can usually be achieved by a combination of MS and NMR (Collins, 1994).

**Hydrophobic side chains of lipids.** Hydrophobic side chains of intact lipids are typically obtained by alkaline or acid hydrolysis of the intact lipids. A variety of side chains has been recorded: members of the *Bacteria* contain a diverse range of hydrophobic side chains in their lipids, whereas members of the *Archaea* have only isoprenoid-based side chains.

### Isoprenoid-based ether-linked lipids

- **Isoprenoid ether-linked side chains are currently only known to occur in members of the *Archaea*. The stereochemistry of the glycerol-isoprenoid side chain is opposite to that found in the glycerol linked to the hydrophobic side chains in members of the *Bacteria*.**

- In members of the *Archaea*, the intact glycerol isoprenoid ethers are normally analysed and are typically divided up into diethers, hydroxylated diethers, macrocyclic diethers, tetraethers, and polyol derivatives of the tetraether. The presence of unsaturation or penta- and hexa-cyclic rings is particularly important in the further modification of the isoprenoid side chain.

- Despite reports to the contrary, there is no reliable evidence that members of the *Archaea* have significant amounts of ester-linked fatty acids in their lipids.

- Care should be taken in choosing the methods of hydrolysis since some compounds (in particular

hydroxylated diethers) produce artefacts when hydrolysed under strongly acid conditions (Ekiel & Sprott, 1992; Sprott *et al.*, 1990). On the other hand, some ether-linked lipids are notoriously stable and specially developed methods are required to cleave all head groups (Kumar *et al.*, 1983; Nishihara *et al.*, 2002). The methods used must be fully documented.

- Ether-linked lipids can be rapidly detected using TLC methods (Tornabene & Langworthy, 1979). They may also be detected by HPLC (Demizu *et al.*, 1992; Ohtsubo *et al.*, 1993; Koga *et al.*, 1998), high temperature GC (Nichols *et al.*, 1993) or MS (de Souza *et al.*, 2009; Macalady *et al.*, 2004). There appear to be no universal methods applicable to all compounds.

### Fatty acids and their derivatives

- **Fatty acids, ester-linked to the glycerol are typical constituents of almost all members of the *Bacteria*. The Sherlock MIS system (MIDI Inc.) provides a comprehensive database, but this is certainly not complete and there are some discrepancies that need to be clarified or compounds that are currently not included in the database.**
- **When determining the fatty acid patterns of strains, the cultivation conditions of the strains should be identical prior to fatty acid extraction. There may be exceptions in cases where it is impossible to grow the organisms under the same conditions. These must be carefully documented.**
- **When reporting fatty acids, all components that constitute 1 % or more of the fatty acids must be reported. In cases where major peaks are not identified, they will not be included in the peak-naming table. In such instances, their presence must be reported and the equivalent chain-length (ECL) given. This will allow any future work on the elucidation of the structure to link to the ECL given in publications.**
- When comparing patterns of fatty acids, it should be remembered that they generally do not fluctuate significantly within a taxonomic group. Thus care should be taken of reports of branched chain fatty acids in a group that otherwise synthesizes straight chain and unsaturated fatty acids. Similarly the presence or absence of hydroxylated fatty acids is generally characteristic and their unexpected absence or presence should be treated with caution.
- When comparing fatty acid patterns, it is important to make appropriate comparisons. The Sherlock MIS (MIDI) system extracts fatty acids from intact cells (Miller, 1982) and comparisons with results that have been obtained from the lipid fraction(s) extracted from cells or on fatty acid methyl ester mixtures that have previously been separated into different classes by TLC are not sensible and can falsify any interpretations.
- In additional to simple fatty acids, there is growing evidence for the importance of modified fatty acids in

a variety of polar lipids found in specific groups of bacteria. These compounds appear to be formed by the condensation of a fatty acid with a molecule with a small molecular mass, such as an amino acid. Known examples include sphingolipids (Naka *et al.*, 2003), capnines (Godchaux & Leadbetter, 1984) and alkyl-amines (Anderson, 1983). Long chain diols (Pond *et al.*, 1986) have also been reported to be a significant component in the lipids of *Thermomicrobium roseum*. Where such compounds are known to occur, appropriate methods of analysis should be undertaken to ascertain whether they are actually present. The methods used must be documented.

### Non-isoprenoid based-ether-linked lipids

- **In addition to fatty acids, there is a diverse range of members of the *Bacteria* that synthesize ether-linked lipids. Typically they are either straight chain or simple branched (not isoprenoid) side chains or mono-unsaturated derivatives, with the point of unsaturation adjacent to the ether bond (i.e. plasmalogens or vinyl ethers).** Plasmalogens/vinyl ethers are hydrolysed under acid conditions to give dimethyl acetals that elute with the fatty acid methyl esters prepared by acid catalysis during GC. Where such compounds are known or suspected to occur in a group, efforts should be made to record their presence/absence. There is evidence that dimethyl acetals have been reported to occur in some organisms where hydroxylated fatty acids elute with the same retention time and may have been mistakenly identified as dimethyl acetals instead of the hydroxylated fatty acid (Moore *et al.*, 1994; Helander & Haikara, 1995). For example, 1,1-dimethoxy-12-methyl-pentadecane ($C_{15:0}$ anteiso-DMA) exhibits an ECL of $-0.0008$ in GC analyses compared with $C_{14:0}$ 2-OH (Kämpfer *et al.*, 2000).
- Non-plasmalogen ethers in members of the *Bacteria* include both mono- and diethers (Langworthy *et al.*, 1983, Rütters *et al.*, 2001). They do not co-chromatograph with mono- or diethers from members of the *Archaea*, but methods applicable to the analysis of such compounds are similar to those used for members of the *Archaea*.

**Polar lipids** There is a vast diversity of polar lipids now known to be present in prokaryotes and in many cases their structures have yet to be fully elucidated. There is currently no collective work that adequately covers all aspects of prokaryote lipids, although the work of Ratledge & Wilkinson (1988) is a good starting point. Their biosynthesis is also not fully understood. It should be emphasized that the polar lipid diversity is associated with the cell membrane(s) and is not limited to just phospholipids. Given the large diversity, it is important to document the lipids present by providing an image of the thin layer plate stained with a reagent that will allow all lipids to be visualized.

- **Run TLCs under standardized conditions. It is often wise for beginners or those with a limited amount of experience to include strains in the study for which the polar lipid patterns are already known.**
- Where unknown polar lipids are reported, the use of terms such as 'unidentified phospholipids' or 'glycolipid' must be supported by an image of the TLC plate. **The inclusion of images of TLC plates is a prerequisite for verifying the work in the future and all work reporting polar lipid patterns should include such images.** Good quality 8 bit, grey scale images, measuring $7 \times 7$ cm, scanned with a good quality scanner at 300 d.p.i. are adequate for publication (see Martens *et al.*, 2006 and Biebl *et al.*, 2007 for examples).
- The range of polar lipids known to occur in members of the *Archaea* is currently restricted to phospholipids, aminophospholipids, glycolipids and phosphoglycolipids.
- The range of polar lipids known to occur in members of the *Bacteria* is currently known to include phospholipids, glycolipids, phosphoglycolipids, aminophospholipids, amino acid derived lipids, capnines, sphingolipids (glyco- or phosphosphingolipids) and also hopanoids. Appropriate staining methods should be used to characterize the functional groups in the polar lipids.
- The nomenclatural system of Lechevalier *et al.* (1977) restricts itself to making reference to only phospholipid patterns in members of the actinomycetes. Additional information can be gained by also reporting the presence of other polar lipids, such as glycolipids.
- In some members of the *Bacteria*, the presence of hydroxylated fatty acids in the polar lipids will alter their $R_f$ values compared with the equivalent structures where hydroxylated fatty acids are not present (Cox & Wilkinson, 1989; Kroppenstedt *et al.*, 1990; Kroppenstedt & Goodfellow, 1991). Such compounds are characteristic of specific taxa.
- Polar lipids are often referred to as if they are single compounds, e.g. phosphatidylglycerol. However, they are more appropriately treated as being a class of compounds, since each novel combination of fatty acids attached to the glycerol is a novel compound. In any one organism, the spot labelled phosphatidylglycerol, for example, may contain several different compounds with different fatty acid compositions. This is evident when the lipid is subject to MS analysis.
- **The results of the analysis of the hydrophobic side chains and the polar lipids should be correlated.**

## Other extracellular constituents
### Lipopolysaccharides (LPS)
- Few routine studies are now performed, but it is evident from studies over the past six decades that both the nature of the sugar(s) present in Lipid A as well as the nature of the fatty acids and the way they are linked (ester and/or amide linked) to the sugar are of significance (Hase & Rietschel, 1976; Weckesser &

Mayer, 1988; Mayer *et al.*, 1989). The chain length of the fatty acids in the LPS may also differ significantly from those found in the polar lipids. When carrying out whole-cell fatty acids analysis, the fatty acids from the LPS will also be extracted and this should be borne in mind when interpreting the data.

### Mycolic acids
- Mycolic acids are known to occur in certain members of the high G+C Gram-positives (Brennan, 1988). Recent work on *Mycobacterium bovis*, *Mycobacterium smegmatis* and *Corynebacterium glutamincum* has shown conclusively that mycolic acids are involved in the formation of an outer membrane with a typical bilayer structure (Hoffmann *et al.*, 2008; Zuber *et al.*, 2008).
- The length of the mycolate side chain is closely correlated to the 16S rRNA gene sequence grouping reported for the majority of taxa that produce them. These compounds are obviously additional taxonomic markers and their detection and analysis should be documented for all taxa where they are known to occur. Mycolic acids may be analysed by TLC (Dobson *et al.*, 1985), HPLC (Willumsen *et al.*, 2001), GC (Rainey *et al.*, 1995; Müller *et al.*, 1998; Torkko *et al.*, 2003) or MS (Fujita *et al.*, 2005a; 2005b).
- In cases where mycolic acids are known to occur, but taxa are shown not to synthesize these compounds, this information is also taxonomically relevant.

## Non-lipophilic constituents of the cell
### Polyamines
Polyamines are small molecular mass compounds that are usually found in the cytoplasm and appear to have a diverse range of functions such as providing stability to the DNA and intracellular compensation of extracellular changes in osmotic conditions. Certain polyamines are also known to be covalently linked to the peptidoglycan of members of the *Sporomusa–Pectinatus–Selenomonas* evolutionary group (Hirao *et al.*, 2000; Kamio & Nakamura, 1987; Kamio *et al.*, 1981a, b). They have been detected in the majority of prokaryotes, but their cellular concentrations may be below the level of detection in moderately or extremely halophilic bacteria. Polyamine patterns can be very helpful to differentiate and define taxonomic groupings (Busse & Auling, 1988; Hamana & Matsuzaki, 1992; Altenburger *et al.*, 1997; Busse & Schumann, 1999).

- Cells subjected to polyamine analysis should be harvested in the late exponential growth phase (Busse & Auling, 1988).
- Extraction can be easily carried out after acidic hydrolysis and derivatization of polyamines using dansylchloride (Scherer & Kneifel, 1983).
- Analysis is carried out using a HPLC running a solvent gradient and equipped with a fluorescence detector.

## OTHER FEATURES OF TAXONOMIC VALUE

It should be remembered that prokaryotes are chemically diverse and the presence of compounds such as teichoic and teichuronic acids (Fischer, 1988; Neuhaus & Baddiley, 2003; Hancock, 1994), arabino-galactans (Brennan, 1988), other heteropolysaccharides (Hancock, 1994; König, 1994; Kandler & Hippe, 1977; Schleifer *et al.*, 1982), or the substitution of sphingoglycolipids for lipopolysaccharides in members of the family *Sphingomonadaceae* can all contribute to the differentiation of various taxonomic groups (see for example Lechevalier & Lechevalier, 1970) within the prokaryotes (members of the *Bacteria* and *Archaea*). This information is encoded somewhere on the genome. An overview of the distribution of teichoic and lipoteichoic acids in actinomycetes has been published by Rahman *et al.* (2009a, b) which links this information to the underlying genetic information that encodes critical steps in their biosynthesis.

## MINIMAL STANDARDS

Minimal standards are useful documents compiled by experts within the framework of subcommittees set up within the ICSB/ICSP (Lapage *et al.*, 1992) with a view to providing detailed information on the way specific groups of organisms are characterized. Their role is covered by Rule 30, recommendation 30 (formerly recommendation 30b) of the Bacteriological Code (Lapage *et al.*, 1992), as modified at the 1999 meetings of the ICSB and its Judicial Commission (De Vos & Trüper, 2000; Labeda, 2000). Minimal standards are not available for all organisms covered by the subcommittees and those available may not necessarily be current. Reference should be made to these publications, since they may be more detailed and cover additional aspects not mentioned here. At the same time, the outline of methods listed here may also be used to expand on the methods given in these guidelines, as well as helping to highlight problems. **Ideally these guidelines and the minimal standards should complement each other. The minimal standards that have been published to date can be found here:**

Aerobic, endospore-forming bacteria (Logan *et al.*, 2009)
Anoxygenic phototrophic bacteria (Imhoff & Caumette, 2004)
Genus *Brucella* (Corbel & Brinley Morgan, 1975a, b)
Family *Campylobacteraceae* (Ursing *et al.*, 1994)
Family *Flavobacteriaceae* (Bernardet *et al.*, 2002)
Order *Halobacteriales* (Oren *et al.*, 1997)
Family *Halomonadaceae* (Arahal *et al.*, 2007; Arahal *et al.*, 2008)
Genus *Helicobacter* (Dewhirst *et al.*, 2000)
Methanogenic bacteria (*Archaea*) (Boone & Whitman, 1988)
Suborder *Micrococcineae* (Schumann *et al.*, 2009)
Class *Mollicutes* (Division *Tenericutes*, Order *Mycoplasmatales*) (International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of *Mollicutes*, 1979; International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of *Mycoplasmatales*, 1972; Brown *et al.*, 2007)
Genera *Moraxella* and *Acinetobacter* (Bøvre & Henriksen, 1976)

Genus *Mycobacterium* (Lévy-Frébault & Portaels, 1992)
Family *Pasteurellaceae* (Christensen *et al.*, 2007)
Root and stem nodulating bacteria (Graham *et al.*, 1991)
Staphylococci (Freney *et al.*, 1999)
Genus *Streptomyces* (not a minimal standard, but a standard reference work, Shirling & Gottlieb, 1966)

## MATTERS RELATING TO THE CODE

**Despite the fact that the Bacteriological Code (Lapage *et al.*, 1992) is rarely cited in the pages of the IJSB/IJSEM, it remains the instrument that controls the way names that have standing in prokaryote nomenclature are brought into use. A number of articles have appeared in the past decade highlighting important aspects of the existing Code or further developing the principles on which the Code is based. Readers of this article are directed towards the publications by Tindall (1999) and Tindall *et al.* (2006) that deal with important aspects of the workings of the Code. The articles by Tindall (2008) and Tindall & Garrity (2008) draw attention to issues relating to the deposit and availability of type material in culture collections.** Recent events have shown that despite the fact that the wording of the Bacteriological Code has now been modified to require that type strains be deposited under the following conditions:

> Rule 30(3a): 'As of 1 January 2001 the description of a new species, or new combinations previously represented by viable cultures must include the designation of a type strain, and a viable culture of that strain must be deposited in at least two publicly accessible service collections in different countries from which subcultures must be available. The designations allotted to the strain by the culture collections should be quoted in the published description. Evidence must be presented that the cultures are present, viable, and available at the time of publication'

**it is clear that strains have and are being deposited that do not conform to the strain that the authors would like to designate as the type strain. It is important to note that the wording refers to the 'type strain' and not any strain. It remains the responsibility of the author(s) to make sure that type strains are deposited according to the wording of the Code.** The importance of the enforcement of the wording of the Code has been emphasized by Tindall (2008) and is a topic that has been discussed repeatedly at meetings of the ICSB/ISCP and its Judicial Commission in 1999 (De Vos & Trüper, 2000), 2002 (De Vos *et al.*, 2005), 2005 (Tindall *et al.*, 2008) and 2008.

Where collections check that the **type strains** are present, viable and will be available at the time of publication, they must be deposited well in advance of the submission of a manuscript. **Where such work is not carried out, it is difficult to know how an author or the editors of the IJSEM can establish that these conditions, required by the current wording of the Bacteriological Code, have been met.**

In summary, the task of describing a novel taxon is one that requires careful selection and use of a wide variety of methodologies. It is also a process that should not be divorced from the workings of the Bacteriological Code (Lapage *et al.*, 1992). Experience gained over the past six decades has continued to demonstrate the value of comparing different datasets and also of basing the description and delineation of taxa on as wide a dataset as possible. The availability of an increasing number of sequenced genomes from a diverse range of prokaryotes is providing an interesting addition to the methods that are now considered to be 'traditional'. At the same time, some of the methods listed above provide an insight into the structure of cells that indicates where information is currently scant or lacking at the genomic level. Experience has shown that the interplay between genetic and phenotypic datasets provides a sound basis for appreciating and describing the diversity of prokaryotes and has the potential to become the foundation of a more stable, in depth taxonomy of the prokaryotes. The publication of a set of guidelines is long overdue and is something that has been alluded to in the past (Lessel, 1970; Murray & Schleifer, 1994). It is appropriate that this task is now completed and it is hoped that the publication of these guidelines will also have implications for other areas of microbial research.

## ACKNOWLEDGEMENTS

## REFERENCES

**Altenburger, P., Kämpfer, P., Akimov, V. N., Lubitz, W. & Busse, H.-J. (1997).** Polyamine distribution in actinomycetes with Group B peptidoglycan and species of the genera *Brevibacterium*, *Corynebacterium*, and *Tsukamurella*. *Int J Syst Bacteriol* **47**, 270–277.

**Amann, R. I., Lin, C., Key, R., Montgomery, L. & Stahl, D. A. (1992).** Diversity among *Fibrobacter* isolates: towards a phylogenetic classification. *Syst Appl Microbiol* **15**, 23–31.

**Anderson, R. (1983).** Alkylamines: novel lipid constituents in *Deinococcus radiodurans*. *Biochim Biophys Acta* **753**, 266–268.

**Arahal, D. R., Vreeland, R. H., Litchfield, C. D., Mormile, M. R., Tindall, B. J., Oren, A., Bejar, V., Quesada, E. & Ventosa, A. (2007).**

Recommended minimal standards for describing new taxa of the family *Halomonadaceae*. *Int J Syst Evol Microbiol* **57**, 2436–2446.

**Arahal, D. R., Vreeland, R. H., Litchfield, C. D., Mormile, M. R., Tindall, B. J., Oren, A., Bejar, V., Quesada, E. & Ventosa, A. (2008).** [Erratum] Recommended minimal standards for describing new taxa of the family *Halomonadaceae*. *Int J Syst Evol Microbiol* **58**, 2673.

**Barrow, G. I. & Feltham, R. K. A. (eds) (2004).** *Cowan and Steel's Manual for the Identification of Medical Bacteria*, 3rd edn. Cambridge University Press.

**Bascomb, S. (1987).** Enzyme tests in bacterial identification. In *Methods in Microbiology*, vol. 19, pp. 105–160. Edited by R. R. Colwell & R. Grigorova. New York: Academic Press Inc.

**Bernardet, J. F., Nakagawa, Y. & Holmes, B. (2002).** Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. *Int J Syst Evol Microbiol* **52**, 1049–1070.

**Beveridge, T. J. & Davies, J. A. (1983).** Cellular responses of *Bacillus subtilis* and *Escherichia coli* to the Gram stain. *J Bacteriol* **156**, 846–858.

**Biebl, H., Pukall, R., Lünsdorf, H., Schulz, S., Allgaier, M., Tindall, B. J. & Wagner-Döbler, I. (2007).** Description of *Labrenzia alexandrii* gen. nov., sp. nov., a novel alphaproteobacterium containing bacterio-chlorophyll a, and a proposal for reclassification of *Stappia aggregata* as *Labrenzia aggregata* comb. nov., of *Stappia marina* as *Labrenzia marina* comb. nov. and of *Stappia alba* as *Labrenzia alba* comb. nov., and emended descriptions of the genera *Pannonibacter*, *Stappia* and *Roseibium*, and of the species *Roseibium denhamense* and *Roseibium hamelinense*. *Int J Syst Evol Microbiol* **57**, 1095–1107.

**Blazevic, D. J. & Ederer, G. M. (1975).** *Principles of Biochemical Tests in Diagnostic Microbiology*. New York: John Wiley & Sons.

**Bochner, B. R. (1989).** 'Breathprints' at the microbial level. *ASM News* **55**, 536–539.

**Boone, D. R. & Whitman, W. B. (1988).** Proposal of minimal standards for describing new taxa of methanogenic bacteria. *Int J Syst Evol Microbiol* **38**, 212–219.

**Bøvre, K. & Henriksen, S. D. (1976).** Minimal standards for description of new taxa within the genera *Moraxella* and *Acinetobacter*: Proposal by the Subcommittee on *Moraxella* and Allied Bacteria. *Int J Syst Bacteriol* **26**, 92–96.

**Brennan, P. J. (1988).** *Mycobacterium* and other actinomycetes. In *Microbial Lipids*, vol. 1. pp. 203–298. Edited by C. Ratledge & S. G. Wilkinson. London: Academic Press.

**Brenner, D. J. (1973).** Deoxyribonucleic acid reassociation in the taxonomy of enteric bacteria. *Int J Syst Bacteriol* **23**, 298–307.

**Brenner, D. J., Martin, M. A. & Hoyer, B. H. (1967).** Deoxyribonucleic acid homologies among some bacteria. *J Bacteriol* **94**, 486–487.

**Brown, D. R., Whitcomb, R. F. & Bradbury, J. M. (2007).** Revised minimal standards for description of new species of the class *Mollicutes* (division *Tenericutes*). *Int J Syst Evol Microbiol* **57**, 2703–2719.

**Busse, H.-J. & Auling, G. A. (1988).** Polyamine patterns as a chemotaxonomic marker within the Proteobacteria. *Syst Appl Microbiol* **11**, 1–8.

**Busse, H.-J. & Schumann, P. (1999).** Polyamine profiles within genera of the class *Actinobacteria* with LL-diaminopimelic acid in the peptidoglycan. *Int J Syst Bacteriol* **49**, 179–184.

**Christensen, H., Bisgaard, M., Frederiksen, W., Mutters, R., Kuhnert, P. & Olsen, J. E. (2001).** 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). *Int J Syst Evol Microbiol* **51**, 2221–2225.

Christensen, H., Kuhnert, P., Busse, H.-J., Frederiksen, W. C. & Bisgaard, M. (2007). Proposed minimal standards for the description of genera, species and subspecies of the *Pasteurellaceae*. *Int J Syst Evol Microbiol* 57, 166–178.

Cohan, F. M. (2002). What are bacterial species? *Annu Rev Microbiol* 56, 457–487.

Collins, M. D. (1985). Isoprenoid quinone analyses in bacterial classification and identification. In *Chemical Methods in Bacterial Systematics* (Society for Applied Bacteriology Technical Series no. 20), pp. 267–287. Edited by M. Goodfellow & D. E. Minnikin. London: Academic Press.

Collins, M. D. (1994). Isoprenoid quinones. In *Chemical Methods in Prokaryotic Systematics*, pp. 345–401. Edited by M. Goodfellow & A. G. O'Donnell. Chichester: John Wiley & Sons.

Collins, M. D. & Gilbart, J. (1983). New members of the Coenzyme Q series from the *Legionellaceae*. *FEMS Microbiol Lett* 16, 251–255.

Collins, M. D., Rodrigues, U., Ash, C., Aguirre, M., Farrow, J. A. E., Martinez-Murcia, A., Phillips, B. A., Williams, A. M. & Wallbanks, S. (1991). Phylogenetic analysis of the genus *Lactobacillus* and related lactic acid bacteria as determined by reverse transcriptase sequencing of 16S rRNA. *FEMS Microbiol Lett* 77, 5–12.

Corbel, M. J. & Brinley Morgan, W. J. (1975a). Proposal for minimal standards for descriptions of new species and biotypes of the genus *Brucella*. *Int J Syst Bacteriol* 25, 83–89.

Corbel, M. J. & Brinley Morgan, W. J. (1975b). [Erratum] Proposal for minimal standards for descriptions of new species and biotypes of the genus *Brucella*. *Int J Syst Bacteriol* 25, 243.

Cox, A. D. & Wilkinson, S. G. (1989). Polar lipids and fatty acids of *Pseudomonas cepacia*. *Biochim Biophys Acta* 1001, 60–67.

Dagan, T. & Martin, W. (2006). The tree of one percent. *Genome Biol* 7, 118.

Davies, J. A., Anderson, G. K., Beveridge, T. J. & Clark, H. C. (1983). Chemical mechanism of the Gram stain and synthesis of a new electron-opaque marker for electron microscopy which replaces the iodine mordant. *J Bacteriol* 156, 837–845.

De Ley, J. (1968). Molecular biology and bacterial phylogeny. In *Evolutionary Biology*, vol. 2, pp. 103–156. Edited by T. Dobzhansky, M. K. Hects & W. C. Steare. Amsterdam, The Netherlands: North Holland Publishing Co.

De Ley, J. (1970). Reexamination of the association between melting point, buoyant density, and chemical base composition of deoxyribonucleic acid. *J Bacteriol* 101, 738–754.

De Ley, J., Cattoir, H. & Reynaerts, A. (1970). The quantitative measurement of DNA-DNA hybridization from renaturation rates. *Eur J Biochem* 12, 133–142.

De Ley, J., Park, I. W., Tijtgat, R. & van Ermengem, J. (1966). DNA homology and taxonomy of *Pseudomonas* and *Xanthomonas*. *J Gen Microbiol* 42, 43–56.

Deloger, M., El Karoui, M. & Petit, M.-A. (2009). A genomic distance based on MUM indicates discontinuity between most bacterial species and genera. *J Bacteriol* 191, 91–99.

Demizu, K., Ohtsubo, S., Kohno, S., Miura, I., Nishihara, M. & Koga, Y. (1992). Quantitative determination of methanogenic cells based on analysis of ether-linked glycerolipids by high-performance liquid chromatography. *J Ferment Bioeng* 73, 135–139.

De Rijk, P., Wuyts, J. & De Wachter, R. (2003). RnaViz2: an improved representation of RNA secondary structure. *Bioinformatics* 19, 299–300.

de Souza, L. M., Muller-Santos, M., Lacomini, M., Gorin, P. A. J. & Sassaki, G. L. (2009). Positive and negative tandem mass spectro-metric fingerprints of lipids from the halophilic Archaea *Haloarcula marismortui*. *J Lipid Res* 50, 1363–1373.

De Vos, P. & Trüper, H. G. (2000). Judicial Commission of the International Committee on Systematic Bacteriology; IXth International (IUMS) Congress of Bacteriology and Applied Microbiology. *Int J Syst Evol Microbiol* 50, 2239–2244.

De Vos, P., Trüper, H. G. & Tindall, B. J. (2005). Judicial Commission of the International Committee on Systematics of Prokaryotes; Xth International (IUMS) Congress of Bacteriology and Applied Microbiology; Minutes of the meetings, 28, 29 and 31 July and 1 August 2002, Paris, France. *Int J Syst Evol Microbiol* 55, 525–532.

Dewhirst, 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.

Dobson, G., Minnikin, D. E., Minnikin, S. M., Parlett, J. H. & Goodfellow, M. (1985). Systematic analysis of complex mycobacterial lipids. In *Chemical Methods in Bacterial Systematics* (Society for Applied Bacteriology Technical Series no. 20), pp. 237–266. Edited by M. Goodfellow & D. E. Minnikin. London: Academic Press.

Ekiel, I. & Sprott, G. D. (1992). Identification of degradation artifacts formed upon treatment of hydroxydiether lipids from methanogens with methanolic HCl. *Can J Microbiol* 38, 764–768.

Felis, G. E. & Dellaglio, F. (2007). On species descriptions based on a single strain: proposal to introduce the status *species proponenda* (sp. pr.). *Int J Syst Evol Microbiol* 57, 2185–2187.

Fischer, W. (1988). Physiology of lipoteichoic acids in bacteria. *Adv Microb Physiol* 29, 233–302.

Fox, G. E., Pechman, K. R. & Woese, C. R. (1977). Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to procaryotic systematics. *Int J Syst Bacteriol* 27, 44–57.

Fox, G. E., Wisotzkey, J. D. & Jurtshuk, 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.

Freney, J., Kloos, W. E., Hajek, V., Webster, J. A., Bes, M., Brun, Y. & Vernozy-Rozand, C. (1999). Recommended minimal standards for description of new staphylococcal species. *Int J Syst Bacteriol* 49, 489–502.

Fujita, Y., Naka, N., Doi, T. & Yano, I. (2005a). Direct molecular mass determination of trehalose monomycolate from 11 species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiology* 151, 1443–1452.

Fujita, Y., Naka, N., McNeil, M. R. & Yano, I. (2005b). Intact molecular characterization of cord factor (trehalose 6,6′-dimycolate) from nine species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiology* 151, 3403–3416.

Godchaux, W. & Leadbetter, E. R. (1984). Sulfonolipids of gliding bacteria. Structure of the *N*-acylaminosulfonates. *J Biol Chem* 259, 2982–2990.

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P. & Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57, 81–91.

Graham, P. H., Sadowsky, M. J., Keyser, H. H., Barnet, Y. M., Bradley, R. S., Cooper, J. E., De Ley, D. J., Jarvis, B. D. W., Roslycky, E. B. & 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.

Gram, C. (1884). Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpärparaten. *Fortschr Medicin* 2, 185–189 (in German).

**Grimont, P. A. D., Popoff, M. Y., Grimont, F., Coynault, C. & Lemelin, M. (1980).** Reproducibility and correlation study of three deoxyribonucleic acid hybridization procedures. *Curr Microbiol* **4**, 325–330.

**Hamana, K. & Matsuzaki, S. (1992).** Polyamines as a chemotaxonomic marker in bacterial systematics. *Crit Rev Microbiol* **18**, 261–283.

**Hancock, I. C. (1994).** Analysis of cell wall constituents of Gram-positive bacteria. In *Chemical Methods in Prokaryotic Systematics*, pp. 63–84. Edited by M. Goodfellow & A. G. O'Donnell. Chichester: John Wiley & Sons.

**Hase, S. & Rietschel, E. T. (1976).** Isolation and analysis of the lipid A backbone. Lipid A structure of lipopolysaccharides from various bacterial groups. *Eur J Biochem* **63**, 101–107.

**Helander, I. M. & Haikara, A. (1995).** Cellular fatty acyl and alkenyl residues in *Megasphaera* and *Pectinatus* species: contrasting profiles and detection of beer spoilage. *Microbiology* **141**, 1131–1137.

**Hiraishi, A., Hoshino, Y. & Kitamura, H. (1984).** Isoprenoid quinone composition in the classification of *Rhodospirillaceae*. *J Gen Appl Microbiol* **30**, 197–210.

**Hirao, T., Sato, M., Shirahata, A. & Kamio, Y. (2000).** Covalent linkage of polyamines to peptidoglycan in *Anaerovibrio lipolytica*. *J Bacteriol* **182**, 1154–1157.

**Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J. M. & Engelhardt, H. (2008).** Disclosure of the mycobacterial outer membrane: Cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci U S A* **105**, 3963–3967.

**Holdeman, L. V., Cato, E. P. & Moore, W. E. C. (editors) (1977).** *Anaerobe Laboratory Manual*, 4th edn. Virginia Polytechnic Institute and State University, Blacksburg.

**Huson, D. H. & Steel, M. (2004).** Phylogenetic trees based on gene content. *Bioinformatics* **20**, 2044–2049.

**Imhoff, J. F. & Caumette, P. (2004).** Recommended standards for description of new species of anoxygenic phototrophic bacteria. *Int J Syst Evol Microbiol* **54**, 1415–1421.

**International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of *Mollicutes* (1979).** Proposal of minimal standards for descriptions of new species of the class *Mollicutes*. *Int J Syst Bacteriol* **29**, 172–180.

**International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of *Mycoplasmatales* (1972).** Proposal for minimal standards for descriptions of new species of the order *Mycoplasmatales*. *Int J Syst Bacteriol* **22**, 184–188.

**Jeon, Y.-S., Chung, H., Park, S., Hur, I., Lee, J.-H. & Chun, J. (2005).** jPHYDIT: a JAVA-based integrated environment for molecular phylogeny of ribosomal RNA sequences. *Bioinformatics* **12**, 3171–3173.

**Johnson, J. L. (1973).** Use of nucleic-acid homologies in the taxonomy of anaerobic bacteria. *Int J Syst Bacteriol* **23**, 308–315.

**Johnson, J. L. (1984).** Nucleic acids in bacterial classification. In *Bergey's Manual of Systematic Bacteriology*, vol. 1, pp. 8–11. Edited by N. R. Krieg & J. G. Holt. Baltimore: Williams & Wilkins.

**Johnson, J. L. & Ordal, E. J. (1968).** Deoxyribonucleic acid homology in bacterial taxonomy: effect of incubation temperature on reaction specificity. *J Bacteriol* **95**, 893–900.

**Johnson, J. L. & Whitman, W. B. (2007).** Similarity analysis of DNAs. In *Methods for General and Molecular Microbiology*, pp. 624–652. Edited by C. A. Reddy, T. J. Beveridge, J. A. Breznak, G. Marzluf, T. M. Schmidt & L. R. Snyder. Washington DC: American Society for Microbiology.

**Kamio, Y. & Nakamura, K. (1987).** Putrescine and cadaverine are constituents of peptidoglycan in *Veillonella alcalescens* and *Veillonella parvula*. *J Bacteriol* **169**, 2881–2884.

**Kamio, Y., Itoh, Y. & Terawaki, Y. (1981a).** Chemical structure of peptidoglycan in *Selenomonas ruminantium*: cadaverine links covalently to the D-glutamic acid residue of peptidoglycan. *J Bacteriol* **146**, 49–53.

**Kamio, Y., Itoh, Y., Terawaki, Y. & Kusano, T. (1981b).** Cadaverine is covalently linked to peptidoglycan in *Selenomonas ruminantium*. *J Bacteriol* **145**, 122–128.

**Kämpfer, P., Rainey, F. A., Andersson, M. A., Lassila, E.-L. N., Ulrych, U., Busse, H.-J., Mikkola, R. & Salkinoja-Salonen, M. (2000).** *Frigoribacterium faeni* gen. nov., sp. nov., a novel psychrophilic genus of the family *Microbacteriaceae*. *Int J Syst Evol Microbiol* **50**, 355–363.

**Kandler, O. & Hippe, H. H. (1977).** Lack of peptidoglycan in the cell walls of *Methanosarcina barkeri*. *Arch Microbiol* **113**, 57–60.

**Karr, D. E., Bibb, W. F. & Moss, C. W. (1982).** Isoprenoid quinones of the genus *Legionella*. *J Clin Microbiol* **15**, 1044–1048.

**Ko, C. Y., Johnson, J. L., Barnett, L. B., McNair, H. M. & Vercellotti, J. R. (1977).** A sensitive estimation of the percentage of guanine plus cytosine in deoxyribonucleic acid by high performance liquid chromatography. *Anal Biochem* **80**, 183–192.

**Koga, Y., Morii, H., Akagawa-Matsushita, M. & Ohga, M. (1998).** Correlation of polar lipid composition with 16S rRNA phylogeny in methanogens. Further analysis of lipid component parts. *Biosci Biotechnol Biochem* **62**, 230–236.

**König, H. (1994).** Analysis of Archaeal cell envelopes. In *Chemical Methods in Prokaryotic Systematics*, pp. 85–119. Edited by M. Goodfellow & A. G. O'Donnell. Chichester: John Wiley & Sons.

**König, H., Kralik, R. & Kandler, O. (1982).** Structure and modifications of the pseudomurein in *Methanobacteriales*. *Zbl Bakt Hyg I Abt Orig C* **3**, 179–191.

**König, H., Kandler, O., Jensen, M. & Rietschel, E. Th. (1983).** The primary structure of the glycan moiety of the pseudomurein from *Methanobacterium thermoautotrophicum*. *Hoppe Seylers Z Physiol Chem* **364**, 627–636.

**Konstantinidis, K. T., Ramette, A. & Tiedje, J. M. (2006).** The bacterial species definition in the genomic era. *Philos Trans R Soc Lond B Biol Sci* **361**, 1929–1940.

**Kroppenstedt, R. M. & Goodfellow, M. (1991).** The family *Thermomonosporaceae*. In *The Prokaryotes*, 2nd edn, pp. 1085–1114. Edited by A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K. H. Schleifer. New York: Springer.

**Kroppenstedt, R. M., Stackebrandt, E. & Goodfellow, M. (1990).** Taxonomic revision of the actinomycete genera *Actinomadura* and *Microtetraspora*. *Syst Appl Microbiol* **13**, 148–160.

**Kumar, R., Weintraub, S. T. & Hanaham, D. J. (1983).** Differential susceptibility of mono- and di-*O*-alkyl ether phosphoglycerides to acetolysis. *J Lipid Res* **24**, 930–937.

**Labeda, D. P. (2000).** International Committee on Systematic Bacteriology. IXth International (IUMS) Congress of Bacteriology and Applied Microbiology. Minutes of the meetings, 14 and 17 August 1999, Sydney, Australia. *Int J Syst Evol Microbiol* **50**, 2245–2247.

**Lane, D. J., Field, K. G., Olsen, G. J. & Pace, N. R. (1988).** Reverse transcriptase sequencing of ribosomal RNA for phylogenetic analysis. *Methods Enzymol* **167**, 138–144.

**Langworthy, T. A., Holzer, G., Zeikus, J. G. & Tornabene, T. G. (1983).** Iso- and anteiso-branched glycerol diethers of the thermophilic anaerobes *Thermodesulfobacterium commune*. *Syst Appl Microbiol* **4**, 1–17.

**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.

**Lechevalier, M. P. & Lechevalier, H. (1970).** Chemical composition as a criterion in the classification of aerobic actinomycetes. *Int J Syst Bacteriol* **20**, 435–443.

**Lechevalier, M. P., De Bièvre, C. & Lechevalier, H. A. (1977).** Chemotaxonomy of aerobic actinomycetes: phospholipid composition. *Biochem Syst Ecol* **5**, 249–260.

**Lessel, E. F. (1970).** Progress and problems in bacterial systematics. *Int J Syst Bacteriol* **20**, 339–344.

**Lévy-Frébault, V. V. & Portaels, F. (1992).** Proposed minimal standards for the genus *Mycobacterium* and for description of new slowly growing *Mycobacterium* species. *Int J Syst Bacteriol* **42**, 315–323.

**Lilburn, T. G. & Garrity, G. M. (2004).** Exploring prokaryotic taxonomy. *Int J Syst Evol Microbiol* **54**, 7–13.

**Logan, N. A., Berge, O., Bishop, A. H., Busse, H.-J., De Vos, P., Fritze, D., Heyndrickx, M., Kämpfer, P., Rabinovitch, L. & other authors (2009).** Proposed minimal standards for describing new taxa of aerobic, endospore-forming bacteria. *Int J Syst Evol Microbiol* **59**, 2114–2121.

**Ludwig, W. & Klenk, H.-P. (2001).** Overview: A phylogenetic backbone and taxonomic framework for prokaryotic systematics. In *Bergey's Manual of Systematic Bacteriology, The Archaea and the Deeply Branching Phototrophic Bacteria*. pp. 49–65. Edited by D. R. Boone & R. W. Castenholz. New York: Springer-Verlag.

**Macalady, J. L., Vestling, M. M., Baumler, D., Boekelheide, N., Kaspar, C. W. & Banfield, J. F. (2004).** Tetraether-linked membrane monolayers in *Ferroplasma* spp: a key to survival in acid. *Extremophiles* **8**, 411–419.

**Martens, T., Heidorn, T., Pukall, R., Simon, M., Tindall, B. J. & Brinkhoff, T. (2006).** Reclassification of *Roseobacter gallaeciensis* Ruiz-Ponte *et al.* 1998 as *Phaeobacter gallaeciensis* gen. nov., comb. nov., description of *Phaeobacter inhibens* sp. nov., reclassification of *Ruegeria algicola* (Lafay *et al.* 1995) Uchino *et al.* 1999 as *Marinovum algicola* gen. nov., comb. nov., and emended descriptions of the genera *Roseobacter*, *Ruegeria* and *Leisingera*. *Int J Syst Evol Microbiol* **56**, 1293–1304.

**Martínez-Murcia, A. J. & Collins, M. D. (1990).** A phylogenetic analysis of the genus *Leuconostoc* based on reverse transcriptase sequencing of 16 S rRNA. *FEMS Microbiol Lett* **70**, 73–83.

**Martínez-Murcia, A. J., Benlloch, S. & Collins, M. D. (1992).** Phylogenetic interrelationships of members of the genera *Aeromonas* and *Pleisiomonas* as determined by 16S ribosomal DNA sequencing: lack of congruence with results of DNA-DNA hybridizations. *Int J Syst Bacteriol* **42**, 412–421.

**Mayer, H., Masoud, H., Urbanik-Sypniewska, T. & Weckesser, J. (1989).** Lipid A composition and phylogeny of Gram-negative bacteria. *Bull Jpn Fed Cult Collect* **5**, 19–25.

**McCarthy, B. J. & Bolton, E. T. (1963).** An approach to the measurement of genetic relatedness among organisms. *Proc Natl Acad Sci U S A* **50**, 156–164.

**Mesbah, M., Premachandran, U. & Whitman, W. B. (1989).** Precise measurement of the G + C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int J Syst Bacteriol* **39**, 159–167.

**Miller, L. T. (1982).** Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy fatty acids. *J Clin Microbiol* **16**, 584–586.

**Moore, L. V. H., Bourne, D. M. & Moore, W. E. C. (1994).** Comparative distribution and taxonomic value of cellular fatty acids in thirty-three genera of anaerobic Gram-negative bacilli. *Int J Syst Bacteriol* **44**, 338–347.

**Müller, K.-D., Schmid, E. N. & Kroppenstedt, R. M. (1998).** Improved identification of mycobacteria by using the microbial identification system in combination with additional trimethylsulfonium hydroxide pyrolysis. *J Clin Microbiol* **36**, 2477–2480.

**Murray, R. G. E. & Schleifer, K. H. (1994).** Taxonomic notes: a proposal for recording the properties of putative taxa of prokaryotes. *Int J Syst Bacteriol* **44**, 174–176.

**Murray, R. G. E., Brenner, D. J., Colwell, R. R., De Vos, P., Goodfellow, M., Grimont, P. A. D., Pfennig, N., Stackebrandt, E. & Zavarzin, G. A. (1990).** Report of the ad hoc committee on approaches to taxonomy within the Proteobacteria. *Int J Syst Bacteriol* **40**, 213–215.

**Naka, T., Fujiwaraa, N., Yanoc, I., Maedaa, S., Doed, M., Minaminob, M., Ikedab, N., Katob, Y., Watabee, K. & other authors (2003).** Structural analysis of sphingophospholipids derived from *Sphingobacterium spiritivorum*, the type species of genus *Sphingobacterium*. *Biochim Biophys Acta* **1635**, 83–92.

**Neuhaus, F. C. & Baddiley, J. (2003).** A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria. *Microbiol Mol Biol Rev* **67**, 686–723.

**Nichols, P. D., Shaw, P. M., Mancuso, C. A. & Franzmann, P. D. (1993).** Analysis of archaeol phospholipid-derived di- and tetraether lipids by high temperature capillary gas chromatography. *J Microbiol Methods* **18**, 1–9.

**Nishihara, M., Morii, H., Matsuno, K., Ohga, M., Stetter, K. & Koga, Y. (2002).** Structural analysis by reductive cleavage with LiAlH$_4$ of an allyl ether choline-phospholipid, archaetidylcholine, from the thermophilic methanoarchaeon *Methanopyrus kandleri*. *Archaea* **1**, 123–131.

**Ohtsubo, S., Kanno, M., Miyahara, H., Kohno, S., Koga, Y. & Miura, I. (1993).** A sensitive method for quantification of aceticlastic methanogens and estimation of total methanogenic cells in natural environments based on an analysis of ether-linked glycerolipids. *FEMS Microbiol Ecol* **12**, 39–50.

**Oren, A., Ventosa, A. & Grant, W. D. (1997).** Proposed minimal standards for description of new taxa in the order *Halobacteriales*. *Int J Syst Bacteriol* **47**, 233–238.

**Pace, B. & Campbell, L. L. (1971).** Homology of ribosomal ribonucleic acid of diverse bacterial species with *Escherichia coli* and *Bacillus stearothermophilus*. *J Bacteriol* **107**, 543–547.

**Palleroni, N. J., Kunisawa, R., Contopoulou, R. & Doudoroff, M. (1973).** Nucleic acid homologies in the genus *Pseudomonas*. *Int J Syst Bacteriol* **23**, 333–339.

**Peplies, J., Kottmann, R., Ludwig, W. & Glöckner, F.-O. (2008).** A standard operating procedure for phylogenetic inference (SOPPI) using (rRNA) marker genes. *Syst Appl Microbiol* **31**, 251–257.

**Pfennig, N. & Wagener, S. (1986).** An improved method of preparing wet mounts for photomicrographs of microorganisms. *J Microbiol Methods* **4**, 303–306.

**Pond, J. L., Langworthy, T. A. & Holzer, G. (1986).** Long-chain diols: a new class of membrane lipids from a thermophilic bacterium. *Science* **231**, 1134–1136.

**Rahman, O., Dover, L. G. & Sutcliffe, I. C. (2009a).** Lipoteichoic acid biosynthesis: two steps forwards, one step sideways? *Trends Microbiol* **17**, 219–225.

**Rahman, O., Pfitzenmaier, M., Pester, O., Morath, S., Cummings, S. P., Hartung, T. & Sutcliffe, I. C. (2009b).** Macroamphiphilic components of thermophilic actinomycetes: identification of lipoteichoic acid in *Thermobifida fusca*. *J Bacteriol* **191**, 152–160.

**Rainey, F. A., Klatte, S., Kroppenstedt, R. M. & Stackebrandt, E. (1995).** *Dietzia*, a new genus including *Dietzia maris* comb. nov., formerly *Rhodococcus maris*. *Int J Syst Bacteriol* **45**, 32–36.

**Ratledge, C. & Wilkinson, S. G. (1988).** *Microbial lipids*, vol. 1. London: Academic Press.

**Rosselló-Móra, R. (2006).** DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In *Molecular Identification, Systematics and Population Structure of Prokaryotes*, pp. 23–50. Edited by E. Stackebrandt. Heidelberg, Germany: Springer Verlag.

**Rütters, H., Sass, H., Cypionka, H. & Rullkötter, J. (2001).** Monalkylether phospholipids in the sulfate-reducing bacteria *Desulfosarcina vaiabilis* and *Desulforhabdus aminigenus*. *Arch Microbiol* **176**, 435–442.

**Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988).** Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491.

**Sanger, F., Nicklen, S. & Coulson, A. R. (1977).** DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463–5467.

**Scherer, P. & Kneifel, H. (1983).** Distribution of polyamines in methanogenic bacteria. *J Bacteriol* **154**, 1315–1322.

**Shirling, E. B. & Gottlieb, D. (1966).** Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol* **16**, 313–340.

**Schleifer, K.-H. & Kandler, O. (1972).** Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* **36**, 407–477.

**Schleifer, K.-H, Steber, J. & Mayer, H. (1982).** Chemical composition and structure of the cell wall of *Halococcus morrhuae*. *Zbl Bakt Hyg I Abt Orig C* **3**, 171–178.

**Schumann, P., Kämpfer, P., Busse, H.-J. & Evtushenko, L. I. for the Subcommittee on the Taxonomy of the Suborder Micrococcineae of the International Committee on Systematics of Prokaryotes (2009).** Proposed minimal standards for describing new genera and species of the suborder *Micrococcineae*. *Int J Syst Evol Microbiol* **59**, 1823–1849.

**Smibert, R. M. & Krieg, N. R. (1994).** Phenotypic characterization. In *Methods for General and Molecular Bacteriology*, pp. 607–654. Edited by P. Gerhardt, R. G. E. Murray, W. A. Wood & N. R. Krieg. Washington, DC: American Society for Microbiology.

**Sneath, P. H. A. & Sokal, R. R. (1973).** *Numerical Taxonomy*. W. H. San Francisco: Freeman and Company.

**Snel, B., Bork, P. & Huynen, M. A. (1999).** Genome phylogeny based on gene content. *Nat Genet* **21**, 108–110.

**Sprott, G. D., Ekiel, I. & Dicaire, C. (1990).** Novel, acid-labile, hydroxydiether lipid cores in methanogenic bacteria. *J Biol Chem* **265**, 13735–13740.

**Tamaoka, J. & Komagata, K. (1984).** Determination of DNA base composition by reversed-phase high-performance liquid chromatography. *FEMS Microbiol Lett* **25**, 125–128.

**Tindall, B. J. (1999).** Misunderstanding the Bacteriological Code. *Int J Syst Bacteriol* **49**, 1313–1316.

**Tindall, B. J. (2005).** Respiratory lipoquinones as biomarkers. In *Molecular Microbial Ecology Manual*, Section 4.1.5, Supplement 1, 2nd edn. Edited by A. Akkermans, F. de Bruijn & D. van Elsas. Dordrecht, Netherlands: Kluwer Publishers.

**Tindall, B. J. (2008).** Confirmation of deposit, but confirmation of what? *Int J Syst Evol Microbiol* **58**, 1785–1787.

**Tindall, B. J. & Garrity, G. M. (2008).** Proposals to clarify how type strains are deposited and made available to the scientific community

for the purpose of systematic research. *Int J Syst Evol Microbiol* **58**, 1987–1990.

**Tindall, B. J., Kämpfer, P., Euzéby, J. P. & Oren, A. (2006).** Valid publication of names of prokaryotes according to the rules of nomenclature: past history and current practice. *Int J Syst Evol Microbiol* **56**, 2715–2720.

**Tindall, B. J., Sikorski, J., Smibert, R. M. & Krieg, N. R. (2007).** Phenotypic characterization and the principles of comparative systematics. In *Methods for General and Molecular Microbiology,* pp. 330–393. Edited by C. A. Reddy, T. J. Beveridge, J. A. Breznak, G. Marzluf, T. M. Schmidt & L. R. Snyder. Washington DC: American Society for Microbiology.

**Tindall, B. J., De Vos, P. & Trüper, H. G. (2008).** Judicial Commission of the International Committee on Systematics of Prokaryotes; XIth International (IUMS) Congress of Bacteriology and Applied Microbiology: Minutes of the meetings, 23, 24 and 27 July 2005, San Francisco, CA, USA. *Int J Syst Evol Microbiol* **58**, 1737–1745.

**Torkko, P., Katila, M.-L. & Kontro, M. (2003).** Gas-chromatographic lipid profiles in identification of currently known slowly growing environmental mycobacteria. *J Med Microbiol* **52**, 315–323.

**Tornabene, T. G. & Langworthy, T. A. (1979).** Diphytanyl and dibiphytanyl glycerol ether lipids of methanogenic archaebacteria. *Science* **203**, 51–53.

**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-Móra, 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.

**Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E. & 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.

**Weckesser, J. & Mayer, H. (1988).** Different lipid A types in lipopolysaccharides of phototrophic and related non-phototrophic bacteria. *FEMS Microbiol Rev* **4**, 143–153.

**Wiegel, J. (1981).** Distinction between the Gram reaction and the Gram type of bacteria. *Int J Syst Bacteriol* **31**, 88.

**Willumsen, P., Karlson, U., Stackebrandt, E. & Kroppenstedt, R. M. (2001).** *Mycobacterium frederiksbergense* sp. nov., a novel polycyclic aromatic hydrocarbon-degrading *Mycobacterium* species. *Int J Syst Evol Microbiol* **51**, 1715–1722.

**Wolf, Y. I., Rogozin, I. B., Grishin, N. V., Tatusov, R. L. & Koonin, E. V. (2001).** Genome trees constructed using five different approaches suggest new major bacterial clades. *BMC Evol Biol* **1**, 8.

**Yarza, P., Richter, M., Peplies, J., Euzéby, J., Amann, R., Schleifer, K. H., Ludwig, W., Glöckner, F. O. & Rosselló-Móra, R. (2008).** The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst Appl Microbiol* **31**, 241–250.

**Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G. & Mamadou Daffé, M. (2008).** Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J Bacteriol* **190**, 5672–5680.