1 Bridging 200 Years of Bacterial Classification

Ramon Rosselló-Móra[1] and Erko Stackebrandt[2,*]
[1]*Grup de Microbiologia Marina, Department Animal and Microbial Diversity, IMEDEA (CSIC-UIB), Esporles, Spain;* [2]*Kneitlingen, OT Ampleben, Germany*

Introduction

The history of bacterial systematics reflects scientific progress in which approaches (often developed in non-biological disciplines) were adopted to improve the accuracy of bacterial taxonomy and to develop comprehensible relationships. What started two centuries ago with a blurred vision of the simplest properties of the bacterial cell became more focused once pure cultures were achievable. At an amazing speed, their existence as both harmful pathogens for humans, plants and animals and as beneficial partners in agriculture, food and industrial applications was disclosed. Microbiology as a scientific discipline in its own right was established, although the historical connection to botany was still recognizable until recently. In parallel with the enormous avalanche of strains, names and accompanying data already generated at the dawn of microbiology, attempts were made to understand the origin of bacteria, and to order them into a system that depicted the relationships among themselves and to other living organisms. A large number of systems were outlined with hardly any of them being identical to another. It was not until far into the 20th century that two major events occurred that would change our perception of how to reject the plethora of bacterial names and classification systems. The first was the establishment of the Approved List of Bacterial Names (Skerman *et al.*, 1980) that brought order into the huge number of synonyms based on obscure species descriptions; the second was the perception that proper classification needed to be based upon phylogenetic relationships. Within a short period of 20 years the tree of life began to unfold, and what originally was founded on a single evolutionary conservative gene has now been extended to genome sequences. Moreover, the era in which pure cultures were needed to assess phylogenetic novelty has been extended to the direct application of metagenome and microbiome studies to environmental samples. As a result, the 'traditional' bacteriologists are confronted with the interest of 'molecular systematists' in giving names to as yet uncultured organisms and even to genomes or fragments thereof. No doubt, without a mutual understanding of the rationale and purpose of naming the individual entities (from genes to cultured cells) confusion is bound to occur, especially if different entities of a given organism are named differently. This chapter will briefly take a historical view of the major steps in bacterial systematics leading to the first reconciliation workshop in 1987 and a re-evaluation of the species concept in 2002.

*erko@dsmz.de

 Trends in the Systematics of Bacteria and Fungi (eds. P. Bridge, D. Smith and E. Stackebrandt)

New challenges and concepts developed since then will be outlined.

The Historical Perspective

The changing consideration of bacterial taxonomic assessment

The early era

Microorganisms became visible by means of the introduction of Antony van Leeuwenhoek's microscope in the mid-17th century. He and Lazzaro Spallanzani objected to the idea of spontaneous generation of microorganisms, and paved the way for the future development of the science of microbiology. It took about another 100 years before the first report on classification and characterization of microorganisms was attempted (Müller, 1786). Another century of optical and instrumental improvement passed before the name '*Bacterium*' as a genus was introduced as a scientific word by Christian Gottfried Ehrenberg (1838). As Ehrenberg stated: 'eine Milchstrasse der kleinsten Organisation geht duch die Gattungen Monas, Vibrio, Bacterium…' ('a Milky Way of the smallest organization runs through the genera Monas, Vibrio, Bacterium…'). However, while at that time their occurrence and relationships in different niches became apparent to the early microbiologists, neither their ecological function nor their mutual relationships revealed themselves to the observer.

It is a part of human nature to arrange subjects into categories, no matter how small the number and how difficult the selection criteria. Being mostly botanist by education is not surprising that, in the absence of facts about their phylogeny, the nature of bacteria (based on nothing more than basic morphological descriptions) could not be properly evaluated by these early systematists. The apparent similarity of mainly unicellular forms and separation by fission allowed the early microbial systematists to deduce that bacteria derived from animals (van Leeuwenhoek, Müller, Ehrenberg), and later from fungi, classifying them as Schizophytae (Cohn, 1872), a system in which bacteria were placed together with 'Spaltalgen' (today named cyanobacteria).

This early period which Paul de Kruif (1926) characterizes so well in his famous novel *Microbe Hunters* is dominated by discoveries in the medical field. But this era also saw the exploration into other scientific fields, such as immunology, chemotherapy, physiology, industrial microbiology, biochemistry and the study of metabolisms, leading to a rapid growth in knowledge about the properties of microorganisms and their interactions with biotic and abiotic matter.

In the second half of the 19th century the deposition and exchange of microbial cultures; the organization of international conferences; and the accessibility of scientific literature was limited, and it is not surprising that individual systematists generated a plethora of systems. Migula (1900) compared about 30 such systems, published between 1836 and 1894. At the dawn of microbiology, and even later, the same microorganism was given different names as individual researchers based their taxonomy on different properties. It was not until 1980 that the Approved List of Bacterial Names (Skerman *et al.*, 1980) brought nomenclatural order into bacteriology by starting a new date for bacterial nomenclature. The tens of thousands of names for bacterial species were reduced to about 2500 names that could be linked unambiguously to a previously defined name of a bacterial species. Nevertheless, despite the methodological shortcomings of taxon descriptions, the early 19th century must be considered a period of great scientific achievement and accurate observations, as some of the names given to bacteria with particular morphologies in the first half of the 19th century (e.g. *Spirillum* [Ehrenberg, 1835], *Spirochaeta* [Ehrenberg, 1835] and *Sarcina* [Goodsir, 1842]) were recognized as valid and included in the Approved List.

To recognize the historic development of species characterization and the changes in affiliating genera to higher taxa we use the fate of *Vibrio cholerae* as an example, starting from the first description by Filipo Pacini (1854) through a series of progress reports as laid down in Bergey's Manual of Determinative Bacteriology and later of Bergey's Manual of Systematic Bacteriology. Almost any other taxon described before the late 20th century was prone to the same fate as the genus *Vibrio* and its type species *V. cholerae*.

A witness of scientific progress: Bergey's Manual of Determinative Bacteriology

About 20 years after the publication of Ehrenberg (1832) a *Vibrio*-shaped bacterium was deduced

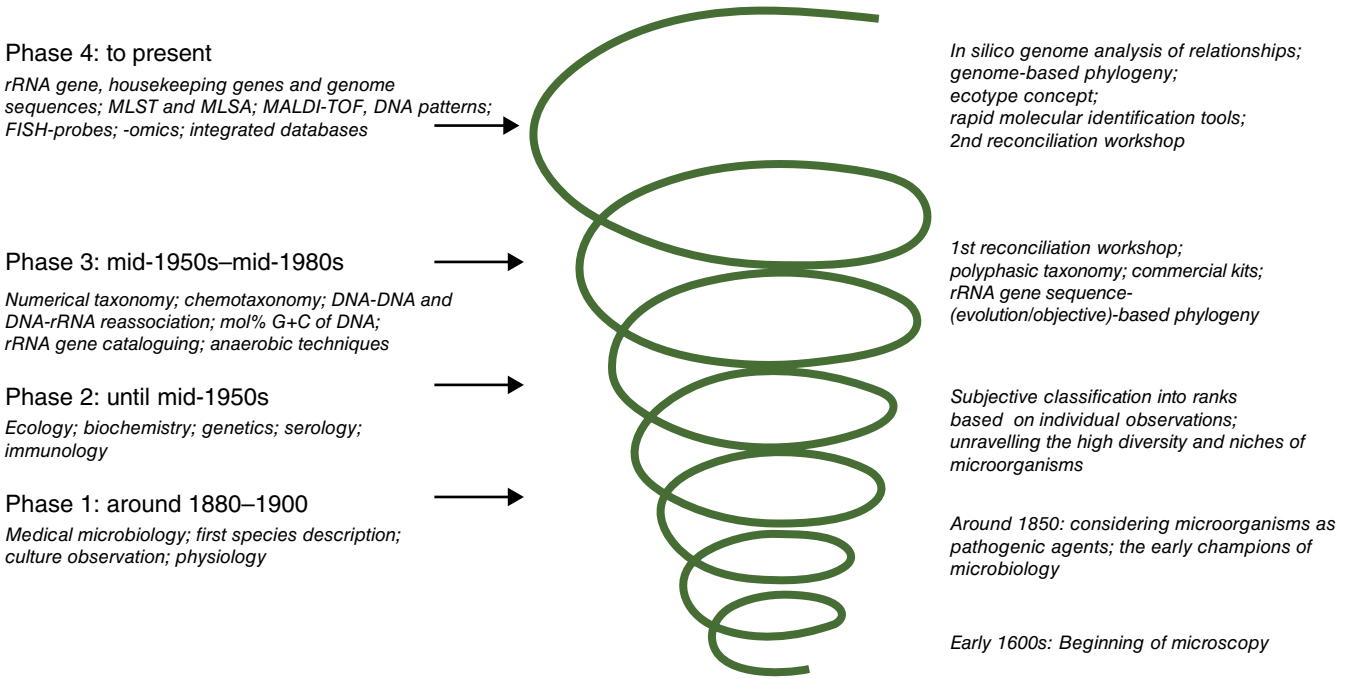

Fig. 1.1. Development of taxonomic approaches through different historical phases. MLST: Multilocus Sequence Typing; FISH: Fluorescent in situ Hybridization.

as the most likely causative agent of cholera disease and was named *V. cholerae* by Pacini (1854), although for several decades Robert Koch was listed as the discoverer of this bacterium. For some of the reasons indicated above, strains of this species or similar organisms were named (among others) *Bacillus cholerae-asiaticae* (Pacini, 1854) Trevisan, 1884; *Spirillum cholerae-asiaticae* (Trevisan, 1884) Zopf, 1885; *Microspira comma* Schröter, 1886; *Spirillum cholerae* (Pacini, 1854) Macé, 1889; *Vibrio comma* (Schröter, 1886) Blanchard, 1906; or *Liquidivibrio cholerae* (Pacini, 1854) Orla-Jensen, 1909 (see www.bacterio.net/vibrio.html, accessed 4 July 2020).

After World War I the dominance of microbial systematics and taxonomy shifted from Europe to the USA. Initiated by the Society of American Bacteriologists, David Hendricks Bergey was appointed chairman of an editorial board in charge of publication of a Manual; under the name *Bergey's Manual of Determinative Bacteriology*, this book and its several editions remained the international reference work and benchmark for bacterial taxonomy. Since 1923 this Manual has served as the main data depository for identification and systematics. No other textbook on bacterial properties is a more accurate witness of taxonomic innovation and progress in identification. Its editors chose a useful, reasonably stable, artificial classification rather than trying to place systematics within a phylogenetic framework (cited in Sapp, 2005). This was not, however, consequently put into practice, as seen in the affiliation of the genus *Vibrio* with the family *Spirillaceae* in the 1st edition (Bergey *et al.*, 1923). The number of properties included for *V. comma* (first Koch, then Schröter were given as references) is small and restricted to culture observations such as shape, flagellation, reaction of growth and pigmentation in different media, a few physiological properties and habitat. The name and taxonomic affiliation, as well as the descriptive criteria, remained basically unchanged in the 2nd to the 5th edition (Bergey, 1925, 1930, 1935; Bergey *et al.*, 1939), except that the genus *Vibrio* was placed in the family *Pseudomonadaceae* in the 5th edition and in the listing of a few more physiological reactions (acid production from carbohydrates). The stagnation seen in microbial taxonomy in the first 30 years of the 20th century contrasts with the progess witnessed in the 'Golden Age of Microbiology' before and around the turn of the 19th century. Although DNA was discovered around 1860, bacterial transformation experiments were published in the 1920s, and chromosomes were identified as the cellular structures responsible for heredity, the structure of DNA had still not been deciphered and genetics was in its infancy. The absence of a clearly defined nucleus let Copeland (1938) propose a separate kingdom – Monera – for the bacteria (and cyanobacteria), bringing them to the same level as those of animals, plants and protoctista. A historic view on the development of the highest systematic ranks (kingdoms, domains) and a summary of the thoughts of leading scientists on the likeliness (and purpose) of constructing a phylogenetic framework for bacteria has been given by Sapp (2005). The 1940s and 1950s saw the development of electron microscopy; the evidence that DNA, not protein, is the genetic material; and the double helix structure of DNA was proposed. As progress in these scientific eras was not evaluated for application in bacterial systematics, the range of taxonomic properties remained basically the same until the early 1970s. van Niel's statement (1955) about the inappropriateness of phenotypic data to order taxa into a hierarchic system and to replace the binominal system of species by common names is a clear testimony to the frustration with bacterial classification prevailing in these decades.

Still concentrating on the cultural properties of *V. comma*, the description of this species in the 6th and 7th editions (Breed *et al.*, 1948 and 1957, respectively) were almost identical but included physiological properties additional to those in the previous editions. More important, information on the antigen structure (O, somatic and H, flagellar) led to the clustering of *V. cholerae* strains into groups and subgroups. While in the 6th edition the genus *Vibrio* was affiliated to the family *Spirilleae*, it was a member of the family *Spirillaceae* in the 7th edition.

The 1960s and 1970s saw a boom in the introduction of taxonomic methods, ranging from the molecular (DNA-DNA and DNA-rRNA hybridization, mol% G+C of DNA) over chemotaxonomic (peptidoglycan, fatty acid, isoprenoid quinone, lipid A, polar lipid) to the phenotypic level (numerical taxonomy) as well as the design of novel cultivation strategies (anaerobic techniques). Naming the key authors of these many approaches would be beyond the framework of

this chapter (see Chapters 9, 10, 13, 15 and 16). However, Sokal and Sneath (1963) should be mentioned here, as computer-assisted numerical taxonomy was the favoured classification of microbes of the 1960s and 1970s. The need to optimally characterize bacterial isolates by the many taxonomic techniques available was coined 'polyphasic taxonomy' by Colwell (1970), a term still very much in use today.

The 8th edition (Buchanan and Gibbons, 1974) included a few changes from the first seven editions, in that (i) the name *V. comma* was replaced by *V. cholerae* Pacini 1854; (ii) the genus *Vibrio* was placed into the family *Vibrionaceae*, listed under Part III 'Gram-negative facultative anaerobic rods' and the rank *Schizomycetes* was abolished; and (iii) the first genomic marker, here the DNA base composition, was included. The range of phenotypic properties was significantly expanded and used in the differentiation from other *Vibrio* species. Molecular approaches not to be directly applied to identification (e.g. DNA-DNA hybridization data) were mentioned under 'Further Comments'.

In the mid-1970s the taxonomic toolbox was well filled with a wide range of methods suited to characterize strains and to delineate closely related species from each other. Chemotaxonomy provided a significant number of genomically stable properties which especially facilitated the characterization of Gram-positive bacteria. The era of numerical taxonomy ended when the inability of phenotypic data to unravel natural relationships became obvious by nucleic acid hybridization. Nevertheless, phenotypic data still played a significant role in species identification and characterization, as later seen by the introduction of commercial systems such as BIOLOG, VITEK or API. Although DNA-rRNA hybridization allowed the detection of broader genetic relationships, it was mainly used for some Gram-negative taxa and the resolution power ended at the suprageneric level. But the fate of bacterial systematics changed rapidly as, almost unnoticed by the majority of taxonomists, a few scientists began to develop methods that would fundamentally change our idea of the natural relationships of bacteria, tracing them back to the origin of life.

A 9th edition was published in 1994, between the release of the two editions of *Bergey's Manual of Systematic Bacteriology* (see below), as a true identification reference book in which the taxonomic data over the past 20 years were added to the 8th edition. Some changes in family composition, biovar characteristics of *V. cholerae* and an extensive list of phenotypic traits of old and newly added genus members were included.

The dawn of unravelling the evolution of Prokaryotes

The majority of the younger readers of this chapter may not be aware of the significance of the term 'oligonucleotide sequencing' as they have been educated in times of rapid genome sequencing. What is done today by sophisticated machines and a range of treeing algorithms at an incredible speed not foreseen 50 years ago, has been a labour- and intellectually intensive exercise. Devised in the 1960s, the oligonucleotide sequence approach was the only method able to assemble entire RNA species, such as small microbial ribosomal or transfer RNA, or the genomes of single-stranded RNA bacteriophages. These molecules had a number of intrinsic properties, such as small size, slow evolution of rrn genes, the availability of site-specific RNases motives, the lack of complementary strands (Heather and Chain, 2016) and the ease at which bulk RNA could be labelled and isolated, and so at that time had a head start over the use of sequencing DNA, which began at the beginning of the 1970s. Originally developed by analytical chemists, the use of RNA oligomers was the basis on which to test the hypothesis of Zuckerkandl and Pauling (1965) that 'semantides' – the sequences of information carrying macromolecules – serve as the basis for deciphering molecular phylogenies (Sogin *et al.*, 1971). Once the method of two-dimensional separation of P^{32}, labelled short 16S rRNA fragments (6 to ~24 nt long), and the 'mastermind' approach of reassembling these T1 RNase fragments from subsets of even smaller ones generated by RNases of different specificity (U2, A) had been established and applied to Bacteria and Archaea (Sogin *et al.*, 1971; Uchida *et al.*, 1974; Woese *et al.*, 1975), the door opened to what culminated in 1977 in the definition of three primary kingdoms (Woese and Fox, 1977). First sceptically considered by the majority of taxonomists, the convincing phylogenetic

relationships led to a rethink in terms of the use of molecular approaches and in the genomic relationships among living matter. The paradigm shift in the understanding of natural relationships embraced and clarified diverse aspects of bacterial taxonomy, systematics, phylogeny and evolution. It later extended to bacterial identification and, crossing kingdom/domain boundaries, to yeast and fungi, algae and protozoa up to the ranks of Plantae and Animalia. Within this period of rapid change in concepts of affiliating lower to higher ranks, the contribution of a few scientists laying the breathtaking foundation of revolutionary changes has been well covered by Sapp (2009). No stone was left unturned (taxonomically speaking), and no historic grail was considered to be untouchable. A change in the taxonomist's consideration of systematics resulted in a dramatic rearrangement of higher taxa in the domain Bacteria and laid the basis for a hierarchic structure of the domain Archaea (Woese *et al.*, 1990). Speculations on the ancestral, primitive stage of the living cell tried to explain the differences in members of the three 'primary kingdoms'. In addition to the numerous reclassification of species taxa like *Firmicutes* (Gibbons and Murray, 1978) and *Proteobacteria* (Stackebrandt *et al.*, 1988) were introduced the origin of mycoplasmas (Woese *et al.*, 1984) and *Cyanobacteria* (Woese *et al.*, 1985) linked to the root of *Firmicutes* (Gram-positive bacteria); the significance of phenotypic and ecological forces to shape bacterial taxa (Zavarzin *et al.*, 1991) were explained and chemotaxonomic traits (Goodfellow *et al.*, 1988) were evaluated in an evolutionary context. The period of 'oligonucleotide sequencing' was replaced, first by reverse sequencing of rRNA genes, then shortly afterwards by cycle sequencing of rDNA for taxonomic purposes (Sanger *et al.*, 1975; Mullis *et al.*, 1986). Since then, sequencing of rRNA genes has become a universal approach to placing a new isolate into a phylogenetic framework, making taxonomy an objective pursuit and less subject to personal judgements about relationships.

Reconciliation of bacterial taxonomy

The influence of 16S rRNA and 16S rDNA gene sequencing in rearranging the place of taxa as outlined in *Bergey's Manual of Determinative Bacteriology* raised the concern of some taxonomists about the power of a single methodological approach that may result in nomenclatural disruption and the abolishment of phenotypic and chemotaxonomic traits. As a consequence, an ad hoc committee, including experts from a wide range of taxonomic fields, met to make recommendations about the future of bacterial systematics (Wayne *et al.*, 1987). In summary, the committee acknowledged, among other things: (i) that a single formal classification system appears to be adequate; (ii) that the genome sequence would be the reference standard for phylogeny and that phylogeny should determine taxonomy; (iii) the need to search for molecules other than rrn genes to verify the findings based on a ribosomal RNA species; (iv) that there should be no designation of higher ranks without chemotaxonomic and sequence data support; and (v) that DNA-DNA hybridization is the superior approach for species delineation. The latter notion was included because the sequences of rrn genes is too conserved to discriminate between closely related species. These recommendations were expanded by subsequent working groups, concentrating on specific taxa (Murray *et al.*, 1990) or on the species level (Stackebrandt *et al.*, 2002), expressing the opinion that the 'polyphasic approach' is superior over emphasis on a few methods and data, and that recent advances in molecular techniques (e.g. MLSA, riboprinting), should be regularly evaluated for their application in shaping taxonomic ranks (see Chapters 10, 11, 13).

Changing gear: *Bergey's Manual of Systematic Bacteriology*

In light of the ongoing creation of names for phylogenetically defined higher taxa, in 1984 Bergey's Trust changed the name of the Manual from 'Determinative Bacteriology' to 'Systematic Bacteriology' to reflect the advancement in phylogenetic considerations. The 1st edition (Holt, 1984) divided the kingdom Prokaryotae into four phylogenetically defined divisions, but the lower ranks were still based on phenotypic descriptions. In contrast, the 2nd edition (Brenner *et al.*, 2005, for Volume 2, *Proteobacteria*) made a

courageous step forward and included a complete hierarchic system of taxa described until then, and consequently continued the naming of higher taxa up to the Phylum level. For species including *V. cholerae* (Phylum *Proteobacteria*, class *Gammaproteobacteria*, order *Vibrionales*, family *Vibrionaceae*), the ecological, phylogenetic and molecular characterization data (e.g. plasmids, bacteriophages, bacteriocins, antigenic structure, gene-specific probes, Restriction Fragment Length Polymorphism) were referenced, and new identification data were included to specifically differentiate the ranks of species and genus.

Some Considerations on Taxonomy and the Misunderstandings

As Cowan (1965) stated, taxonomy is a discipline that is dedicated to three major objectives: (i) classification; that is, the establishment of a structured system of categories of living entities by means of their natural relationships; (ii) nomenclature; that is, the procedure of naming taxa in accordance with scientifically established rules; and (iii) identification, the major goal of taxonomy that focuses on the recognition of members of classified taxa by means of a series of diagnostic features. The system created needs to be *universal* (applying to all members that the system should embrace), *operative* (must work and not be so complexly generated that is unworkable) and *predictive* (ideally, one should be able to predict genetic, phenotypic and even ecologic traits when hearing or reading a name) (Rosselló-Móra, 2012). Thus, taxonomy is about being accurate but also pragmatic, as all disciplines (in our case of bacteriology) must be able to make use of it with no major complications. And this is one of the major problems that taxonomists are disputing among themselves, as well with the scientific community: the disjunction between being precise and operational, between constructing a system for everyone and showing fine-tuned natural relations. However, the major problem of prokaryotic taxonomy is that only nomenclature is official, while there is no official classification, nor classification requirements. The Bacteriological Code, also known as the *International Code of Nomenclature of Prokaryotes* (ICNP; Parker *et al.*, 2019), was first created in 1958 and reviewed in 1990, 1992 and 2006) (Lapage *et al.*, 1992; Parker *et al.*, 2019), and is the official document showing how names must be constructed and prioritized once a species is published. The new taxonomic descriptions must be first published in an international, peer-reviewed journal. Then, upon publication, the protologue (or formal description in which the etymology, diagnostic features and type material are indicated) is evaluated by the list editors of the *International Journal of Systematic and Evolutionary Microbiology* (*IJSEM*) who validate the name if formulated in accordance with the Bacteriological Code. Only the manuscripts published in *IJSEM* are automatically validated and listed in its notification lists. All other classifications in journals other than *IJSEM* must be evaluated by the responsible *IJSEM* list editors who generate the validation lists. Unfortunately, and despite the editors' reminder to submit a request for validation to *IJSEM*, many names remain invalid as a request has never been submitted (Oren *et al.*, 2018). Both notification and validation lists become the 'official' lists of accepted names, and ultimately are what most scientists believe is 'the official taxonomy' (see Chapter 3).

Theoretically, the classification of a taxon is just a matter of the scientific opinion of a taxonomist, without restricting the criteria to circumscribe taxa and giving all freedom. Ultimately, this is a judgement of an expert, and as such is referred to in the Bacteriological Code (Parker *et al.*, 2019). However, this freedom is only partially true, because to catalogue a name in the validation or notification lists, the taxonomic description must be first published in an international, peer-reviewed journal. Thus, the first restrictions to the scientific freedom originate from the consensus (or opinion) of a specific subcommittee of scientists (if any) who outline the minimal standards for the classification of the microorganisms they cover as experts. They also come from the referees and editors who request specific parameters and standards (beyond the naming rules) for taxa descriptions (Konstantinidis *et al.*, 2018). The name will ultimately be accepted if the description meets the requirements established in the Bacteriological Code, in which the deposition of a pure culture in two different strain collections is an indispensable requisite. No other form of type or reference material is

allowed. Altogether, the highly restrictive way bacteriologists proceed with the classification of bacterial taxa and valid publication of the new names is excellent in guaranteeing taxonomic accuracy, but ultimately one must admit that there is no freedom to decide what a taxon is (see Chapter 3).

The Paramount Relevance of the Genomic Data

As mentioned in the historical perspective, the early age of microbial taxonomy was based on phenotypic distinction of taxa, and the hierarchy was mostly reflecting physiological and morphological similarities and divergences. However, major breakthroughs occurred with the discovery that DNA was the genetic information store of living cellular organisms, and of the intrinsic chemical properties of this molecule. Already in the 1960s the DNA-DNA hybridization experiments were designed to evaluate the genomic similarities between classified taxa (Rosselló-Móra *et al.*, 2011). Soon, it was observed that often phenotypically coherent groups of strains also formed genomically coherent groups based on either reassociation kinetics or percentage of complementarity. The many laboratory experiments done in the second half of the 20th century showed that a difference smaller than 5°C in the melting temperature between hybrids and homologous DNA (ΔTm) or higher than 70% in DNA-DNA similarity could be the inclusive border of members of the same species. Actually, these two parameters have been the gold standard for circumscribing species (despite the fact that some taxonomists do not accept them), as for several decades they provided the only approach that would allow numerically measured boundaries. Almost all new classifications using two or more strains were evaluated with one of the different DNA-DNA hybridization techniques (Rosselló-Móra *et al.*, 2011). Only in cases where a single strain was serving as type material, and its 16S rRNA gene identity with the closest relative type strain was below 98.7% (Stackebrandt and Ebers, 2006), was the new classification without genomic comparisons tolerated.

The influence of the species circumscription by means of genomic data has been of paramount relevance since the first measurements, and has gained significance as the feasibility of sequencing genomes at relative low costs allowed in silico comparisons. As indicated below, the wet-lab studies using error-prone DNA-DNA hybridization techniques have been substituted by a series of more precise in silico measured overall genome-related indices (OGRI; Chun *et al.*, 2014). The provision of draft or complete genome sequences of the type strain has become compulsory for any new classification, demanded by journals publishing taxonomic papers, such as *Systematic and Applied Microbiology* (compulsory since 2014), *Archives of Microbiology* and *Current Microbiology* (since 2017) or *IJSEM* (since 2019). Although not well accepted, modern bacterial classification has a strong genetic and genomic basis. Phenotyping has been losing relevance, and most of the tests used are either unnecessary and/or uninformative in many taxonomic papers (Sutcliffe, 2015). Like it or not, the current classification of higher taxa is only driven by molecular phylogenies based on 16S rRNA genes or, more recently, by analyses of essential genes (Parks *et al.*, 2018), with the fine-tuned species and genus circumscriptions using OGRI. Recent developments in high-throughput sequencing platforms, and bioinformatics allowing the determination of almost complete genomes retrieved from culture-independent methods, challenge the future of microbial taxonomy, as discussed below.

Recent Innovations

The evolution of the classification of prokaryotes has evolved in parallel with technological developments (see above and Rosselló-Móra, 2012); and prokaryotic taxonomy has always pioneered the use and development of molecular approaches to discern among distinct, but closely related organisms. GC mol% content, followed by DNA-DNA hybridization techniques, and molecular phylogenies using 16S rRNA gene analysis, have been the major advances in genotyping that allowed the current classification with high stability. Early in this century, multilocus sequence analysis (MLSA) was foreseen as an alternative to DNA-DNA hybridization (Stackebrandt *et al.*, 2002), but mostly has been used as an alternative to complement molecular phylogenies.

However, the major breakthrough in the taxonomy for Bacteria and Archaea occurred with the development of high-throughput sequencing technologies. The capability to fully sequence genomes at a relative low cost, together with the development of a myriad of bioinformatic tools to analyse and compare them, has revealed a completely new dimension in classification. Again, genotyping has been the major source of innovative and precise measurements that can numerically circumscribe taxa. As indicated above, high-throughput sequencing technologies have allowed in silico genome to genome comparisons, and several OGRI (Chun *et al.*, 2014) have been developed. The simplest and therefore most straightforward parameters are both the average nucleotide identity (ANI; Konstantinidis and Tiedje, 2005a) and average amino acid identity (AAI; Konstantinidis and Tiedje, 2005b), as the former calculate the mean in identity between orthologous pairs of DNA fragments and the latter the mean of similarity among orthologous DNA-translated proteins. All other OGRI are slightly more complex, as they use additional genome parameters to generate a value, but ultimately all of them render similarly valuable information. Since its development (Konstantinidis and Tiedje, 2005a) ANI has been used extensively to circumscribe species. In the dawn of the in silico analyses using the extant genomic information, the commonly used range of similarity around the 70% DNA-DNA hybridization boundary could be equivalent to an ANI range of 94%–96% (Richter and Rosselló-Móra, 2009). However, the existence of discrete biological units (which we call 'species') has been often questioned, as the boundaries were considered pragmatic, artificial and of no biological significance (Rosselló-Móra, 2012). With the increasing number of genomes in the databases, it was possible to observe a bimodal ANI curve among the species of a genus reinforcing the existence of a fuzzy zone between 92% and 96% ANI (Rosselló-Móra and Amann, 2015). Stronger proof that 'species' exist, however, came from the thousands (close to 100,000 genomes) of pairwise comparisons reflecting that, between a species and its closest relative species, there was a gap in the ANI values (ANI-gap) showing a clear jump (Jain *et al.*, 2018). The demonstration with a rational biological foundation of the existence of what bacteriologists call 'species' is perhaps the most important breakthrough in the taxonomy of prokaryotes since the beginning of the 19th century, and will most probably have a paramount influence in the very near future.

AAI is also a parameter that has been shown to be very informative in the classification of novel taxa. The much more conservative nature of the translated nucleic acid information to proteins has revealed resolution at the level of higher taxa, and especially at the genus level (Konstantinidis and Tiedje, 2005b). The boundary of 70% AAI can serve as a very plausible threshold to discern two genera, similar to the previously used 94% threshold of 16S rRNA gene identity (Yarza *et al.*, 2014). The artificial nature of the present hierarchical classification system (Rosselló-Móra, 2012), results in a much fuzzier circumscription of higher taxa than that of species and genus, but there is still a correlation between how taxonomists have classified families, orders and classes, and the AAI and 16S rRNA gene identity thresholds (Konstantinidis and Tiedje, 2005b; Yarza *et al.*, 2014). It is remarkable that the taxonomic rank Phylum has been thoroughly used in molecular phylogenies and molecular ecology, but this category has no standing in nomenclature (Whitman *et al.*, 2018). Yet there is a definitive need to implement this category in the code of nomenclature for prokaryotes as it is possibly, after the rank of species, the most popular category used among molecular microbial ecologists.

In addition to OGRI, the high-throughput sequencing platforms have brought the possibility to expand phylogenetic analyses beyond the single gene, as 16S rRNA or a small set of essential single copy genes (MLSA). Access to the almost complete gene content of genomes has allowed genealogical reconstruction using the shared core of genes among a group of organisms (phylogenomics), or even reconstruction of a prokaryotic-wide phylogeny using the universally conserved housekeeping genes (Parks *et al.*, 2018). The impact of the genomic taxonomy approach is enormous, and has influenced even the structure of *Bergey's Manual of Systematics of Archaea and Bacteria* that will implement the taxonomic structure proposed in the Genome Taxonomy Database (GTDB) (www.bergeys.org/). The broad range of effects on genomic taxonomy has resulted in the proposal of some relevant changes in hitherto accepted classification.

The influence of the 16S rRNA gene sequence is predominant in systematics, however; its highly conserved nature and the large database guarantees its dominant role in prokaryotic classification, and it will not be abandoned in the short term. As the sequencing approaches and bioinformatic tools improve, and the genome databases such as GTDP (Parks *et al.*, 2018) or Microbial Genomes Atlas (Rodriguez-R *et al.*, 2018) expand, most probably the relevance of the small subunit ribosomal RNA gene will slowly decrease.

Altogether, the influence of genotyping – and especially that of the basic genomic comparisons – has guided the structure of the current classification system (see Chapters 10 and 15) and, somehow, phenotyping has been left aside and relegated to just the minimal standards needed for classification. Often most of the phenotypic data provided in the current taxonomic descriptions are of very low relevance and do not really include truly valuable diagnostic traits (Sutcliffe *et al.*, 2012; Sutcliffe, 2015). To overcome this lack of solid phenotypic data, an effort is needed to implement the high-throughput technologies originally developed in chemistry, also known as metabolomics (Rosselló-Móra *et al.*, 2008). There is a series of different mass spectrometry and fine-tuned technologies that could be implemented. For example, large molecules (between 1000 and 10000D) could be clearly detected using matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS), and small molecules (between 50 and 2000D) using high-resolution spectrometry such as Fourier-transform ion cyclotron resonance (FT-ICR MS). The latter has been used already to determine metabolite profiles to distinguish between different strains of the same species (e.g. Antón *et al.*, 2013), along the growth curve of the same strain (Brito-Echeverria *et al.*, 2011), or even to show the biogeographic patterns of members of the same species (Rosselló-Móra *et al.*, 2008). Such metabolomics have never, however, been used in taxonomy. On the other hand, MALDI-TOF has already gained relevance in systematics (Welker and Moore, 2011), and also in medical diagnostics. The relatively specific whole-cell profiles, mostly originating from ribosomal proteins, show patterns that could be diagnostic for members of the same species or of very close species (see Chapters 7 and 8). This approach has been already well used in classifying and/or identifying thousands of strains in a single study (Viver *et al.*, 2015; Alejandre-Colomo *et al.*, 2020). The cost efficiency of this technology allows the large-scale screening of strains and their classification in accordance with their profiles. The cumulative database of profiles will, in the future, allow rapid classification of new isolates, as well as precise identification of those that may be new and unclassified, facilitating the discovery of taxonomic novelty. Moreover, it helps in generating collections of organisms in which single species can be represented by several strains, avoiding the problems of species described only using a single strain (Christensen *et al.*, 2001). A strategy to assess the large-scale cultivation of microorganisms from the human gut resulted in the isolation of thousands of strains on about 200 variations of culture media (later reduced to about 20) and led to the development of the so-called Culturomics approach (Lagier *et al.*, 2016). In accelerating the process of identification, and to provide a shortcut through the lengthy traditional species description, MALDI-TOF MS represents the beginning of phenotypic characterization of the microbiome, followed by rRNA gene sequence analysis of putative novel species. In accord with the idea of phylotaxonomy, the genome sequence of such species is generated and compared to those of the nearest related type strains, based upon 16S rRNA gene sequence similarity. A few basic phenotypic traits, and whole-cell fatty acid composition, complemented the spectrum of data which led to a significant increase in the recording of novel species from the human gut (Lagier *et al.*, 2018).

The identification of a huge number of strains from hosts or the environment is not only a source of novelty, but also a good way to expand the collection of isolated organisms in pure culture that can serve as a source for biotechnological applications. The fact that not all members of the same species isolated from the same source are genomically identical (e.g. Gevers *et al.*, 2005; Antón *et al.*, 2013), enables large collections of strains of the same taxon to be very useful material in understanding genomic diversity. Another benefit of high-throughput sequencing in systematics is in promoting our understanding of how genes show the blueprint of classified taxa (Munoz *et al.*, 2020).

Taxonomy Needs to Change Its Path

The relevance of high-throughput approaches in the study of the diversity of prokaryotes contrasts with the relatively low pace in the description of new species. At the time of writing, about 14,000 species had been classified and close to 19,000 names had been proposed (in accordance with the List of Prokaryotic Names with Standing in Nomenclature [LPSN] figures; www.bacterio.net/; Parte, 2018). It is commonly acknowledged that the number of species in the environment is orders of magnitude higher, and to some extent we hypothesized a range of between 1 to 10 million (Yarza *et al.*, 2014). The relatively low pace of bacterial taxonomy is due to the entire, complex procedure of classifying new taxa, which has several bottlenecks that need to be resolved. In the first instance, only names given to pure cultures can be validated. The Bacteriological Code requires that the reference material (in the form of a type strain) has to be deposited in two international culture collections, to preserve the living genetic material, and to make it available to the scientific community (Parker *et al.*, 2019).The deposition of strains in collections alone takes some time, as these have to certificate the identity of the deposit; other collections implement legal constraints in sharing material that may make the valid publication of a name problematic (Oren *et al.*, 2018).

The pace of species descriptions (Fig. 1.2) has witnessed a steady growth since the 16S rRNA gene sequence allowed the recognition of novelty in the genealogical tree. The arithmetic growth reached a plateau between 2007 and 2014 (with about 600 species descriptions/year), and a significant jump in new descriptions occurred in 2015 (with about 900 species descriptions/year) and in the following years. Basically, the reasons were the increase in numbers of publications (due to the appearance of novel journals publishing taxonomic papers), and the increased number of papers, especially from countries that recognized the value of indigenous resources. But the yearly number of papers dealing with species is still limited, and also the policies of some journals may lead them to drastically reduce the number of taxonomic papers (Sutcliffe, 2019). Altogether, this may have a negative impact on the pace of new taxonomic proposals. One of the main reasons for the current fluctuations is the rise of single-strain species descriptions (SSSD; Tamames and Rosselló-Móra,

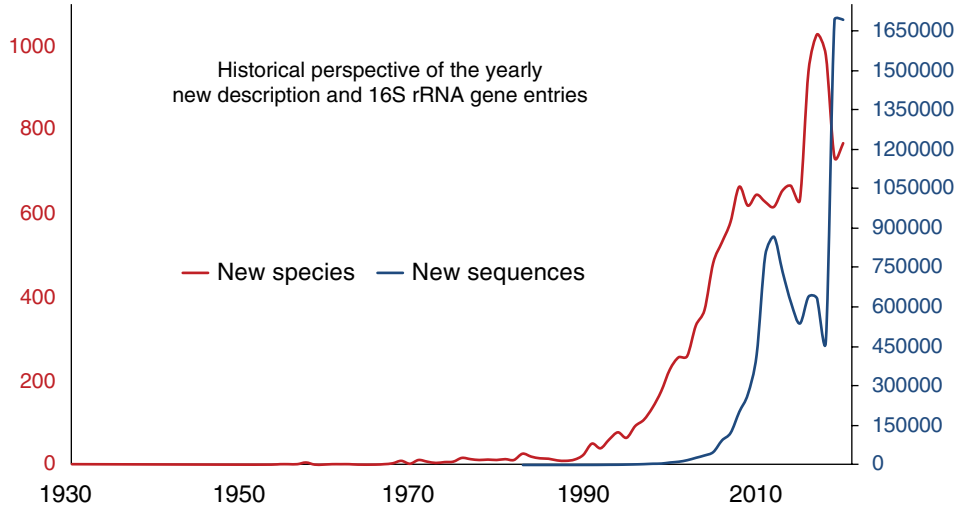

Fig. 1.2. Yearly number of new species descriptions and new 16S rRNA gene entries in public repositories. The primary Y-axis (left) in red relates to the number of described taxa (red line), and the secondary Y-axis (right) relates to the number of gene entries in public repositories (blue line) in accordance with the SILVA database (www.arb-silva.de). Owing to the lack of releases in 2018, the numbers of gene entries for 2018 and 2019 are given as the mean in yearly increase between SILVA SSU 132 and 138 (Quast *et al.*, 2013).

2012), often with reduced relevance owing to the very sparse information content of the descriptions (Christensen *et al.*, 2001). This, together with 'one-species one-paper' practices (Sutcliffe, 2019), has strongly impacted the citation indices of the journals publishing these papers. The impact has had such an influence that many journals adhered to the Declaration on Research Assessment (DORA; https://sfdora.org/) and avoided the use of altmetrics in their journal advertisements (Parish *et al.*, 2019).

It seems that there is a conflict between fostering an increase in new classifications, and the journal policies, owing to the scientific relevance of their publications. Actually, the cataloguing of new taxa may well be disassociated from the standard publication procedures. Journals could centre their publication basis on highly relevant classifications. As Sutcliffe (2019) stated, '…it seems unfeasible, given the sheer scale of microbial diversity, that a "one species, one paper" publishing model can be sustained indefinitely, particularly if there is to be a shift towards "high throughput" approaches to the circumscription and valid naming of prokaryotic taxa'. Thus, it is clear that a new path for taxonomic cataloguing is needed. To this purpose, the Digital Protologue Database (DPD; Rosselló-Móra *et al.*, 2017a,b) was created. In the first instance this was to generate a repository of taxonomic descriptions in an orderly manner in a database-based format that could be cumulative and searchable following the current bioinformatics requirements. The protologues were laid out in a way that was reminiscent of the gene and genome entries in public repositories, formatted in fields of searchable information and universal to all new classifications. Each entry was given a unique TaxoNumber. But, in the second instance (and as future perspective), it was suggested that the DPD was to be a repository of descriptions, each having been given a digital object identifier (DOI) entry. It was to be citable independently from publishing journals, and ideally run by one of the managing entities of the International Nucleotide Sequence Database Collaboration (INSDC) that would guarantee stability. The DPD was warmly welcomed by other journals publishing taxonomic papers, such as *Archives of Microbiology* (Stackebrandt and Smith, 2017a), *Current Microbiology* (Stackebrandt and Smith, 2017b) and *New Microbes New Infections* (Drancourt and Fournier, 2018). In its 2-year life over 1000 entries and 750 registered users were accomplished. However, the inability to secure the support of the editorial board of the *IJSEM* (because of its dominant role in publishing > 80% of all new names) forced closure of the DPD (Rosselló-Móra and Sutcliffe, 2019a,b). The authors believe that it is only a question of time before the classification of microorganisms will be automatized (Rosselló-Móra and Whitman, 2019), especially if we want to speed up the process of cataloguing the expected vast amount of as yet unclassified taxa. The DPD, in one form or another, will then prevail. We foresee that the global centralization of data, with taxonomic descriptions and their interlinking genomic, genetic and phenotypic data, will be the choice of the future; the one-species one-paper practices will then go extinct.

However, the major bottleneck to cataloguing prokaryotic species is cultivation. In general, discussion about the vast majority of prokaryotes being recalcitrant to cultivation (Konstantinidis *et al.*, 2017) is commonplace, but is one of the major problems if we want to classify the entire diversity at a relatively fast pace. Traditionally, since the introduction of molecular microbial ecology tools, uncultured microorganisms were detected solely by means of 16S rRNA amplified and cloned gene sequences, and by application of phylogenetic probes directed to the ribosomal RNA sequences either using northern blots or fluorescence in situ hybridization (Amann *et al.*, 1995). The failure to retrieve accurate genetic and phenotypic data from uncultivated microorganisms made their stable classification and nomenclature impossible. Only some microorganisms with conspicuous characters such as size, inclusions or lifestyle could be identified, and for these the provisional category of *Candidatus* was created (Murray and Stackebrandt, 1995). The disadvantage of this provisional category is that it has no standing in the Code; and, derived from this, the names given have no priority (Whitman *et al.*, 2019). The lack of priority means that if someone isolates a representative of a *Candidatus* as a pure culture, any new name could be published, while the earlier given name would be denied. The lack of a stable framework seems to have discouraged molecular ecologists to 'formally' name the uncultivated prokaryotes they detected and, in about 20 years of existence, only about < 500 *Candidatus* taxa were proposed (Oren, 2017).

But things have changed, and the high-throughput sequencing strategies have brought a completely new dimension to microbial systematics (see Chapters 10, 11, 13, 15). The metagenomic approaches, together with the new bioinformatic tools, have allowed the segregation of single-species genomes from complex metagenome pools of sequences. The metagenome assembled genomes (MAGs) represent the pooled mosaic of the genomes of the coexisting populations or strains of a single species thriving in the same sample (Konstantinidis *et al.*, 2017). On the other hand, single cell amplified genomes (SAGs) can render genomic information on single strains without a cultivation step (Hedlund *et al.*, 2015). This giant step in microbial molecular ecology has allowed the retrieval of genetic information from uncultured organisms that is at least of equivalent quality as that derived from cultured organisms (Konstantinidis and Rosselló-Móra, 2015). The most important parameters for a taxonomic classification that infers phylogeny and calculates OGRI can be readily be determined, and with high accuracy, using these culture-free methods. Essential gene analyses to infer global phylogenies are now feasible, and therefore the new taxonomic framework already includes uncultured MAGs and SAGs (Parks *et al.*, 2018). Similarly, global genome analyses to precisely circumscribe species and genera using ANI and AAI can now also include MAGs and SAGs (Rodriguez-R *et al.*, 2018) with which the ANI-gap has been demonstrated (see above; Jain *et al.*, 2018). Metabolism can also be inferred from the genome, and although it is clear that the presence of genes may not always correlate with gene expression (Bisgaard *et al.*, 2019), the phenotypic inference can always be proven by means of extant sophisticated techniques using radiolabelled (Rosselló-Móra *et al.*, 2003) or stable isotope-labelled compounds (Musat *et al.*, 2016), as well as metatranscriptomics (Zuñiga *et al.*, 2017) and metaproteomics (Armengaud, 2016).

The revolution of high-throughput techniques and their application in systematics has led to the radical proposal for the Bacteriological Code to allow DNA sequences to become type material as an alternative to the deposit of living cultures in international collections (Whitman, 2015, 2016). If DNA could become type material, the classification of all (not only fastidious) microorganisms would be facilitated and the door would be opened to classifying uncultured taxa for which high quality MAGs or SAGs had been retrieved (Konstantinidis *et al.*, 2017). The most important benefit of cataloguing the as yet uncultured taxa using the same classification standards as used for those that are cultured is the provision of a stable nomenclature that would avoid anarchic designations and the synonymy that often arises owing to the lack of rules. However, this proposal has not always been welcome by taxonomists (Oren and Garrity, 2018; Bisgaard *et al.*, 2019; Overmann *et al.*, 2019). There is an uneasiness that nomenclatural chaos will be created arising from the avalanche of names based on the classification of MAGs and SAGs; there is also great concern that these sequences do not represent extant species genomes, but non-existent chimeras resulting from insufficiently powerful bioinformatic tools. We are at the dawn of the high-throughput bioinformatics era, and in the future the accuracy of the new approaches will improve enormously. But, even today, there is evidence indicating that MAGs share ANI values > 97% with the genomes of isolates: for example, *Haloquadratum walsbyi* (Viver *et al.*, 2019), *Salinibacter ruber* (Ramos-Barbero *et al.*, 2019), *Escherichia coli* (Almeida *et al.*, 2019; Peña-Gonzalez *et al.*, 2019) and *Candidatus* 'Macondimonas diazotropica' (Karthikeyan *et al.*, 2019). Moreover, the fact that, in a given sample, different populations of the same species coexist with different genomic content (Antón *et al.*, 2013), the retrieved MAGs would represent the core genome of the species that thrived when the sample was taken. Actually, there is an advantage with this approach, as it is most likely that the genomic blueprint of what a species is relies on the shared set of genes, and a MAG will better represent the species than the single isolate of the common SSSDs.

Conclusions: Reconciliation or Divorce

Whether DNA is accepted as alternative type material remains independent of whether the code of nomenclature will consider MAGs and SAGs as nomenclatural types. Thus, both cases can be treated independently by the International Committee on Systematics for Prokaryotes (ICSP;

www.the-icsp.org). Several proposals have been forwarded to the committee for its evaluation and vote on the use of DNA as type material (Whitman, 2016) and the acceptance of Phylum as the highest category (Whitman *et al.*, 2018). At present, both requests can only apply to pure cultures but, to allow the names of MAGs and SAGs to gain priority, a third proposal was raised for them (Whitman *et al.*, 2019). At the time of writing, none of the proposals had been voted for by the ICSP, and so none of the relevant topics are either accepted for or denied implementation in the Bacteriological Code. Microbial systematics is now at the tipping point between reconciliation or divorce among classical taxonomists and molecular microbial ecologists.

The result of the opinion of the ICSP will definitively clarify the path of prokaryotic taxonomy, and the decision is especially important for the cataloguing of the uncultured organisms. As already proposed (Konstantinidis *et al.*, 2017), either the Bacteriological Code is adapted to the new winds of molecular systematics, thus allowing MAGs and SAGs to be designated with stable names with priority, or an alternative nomenclatural code should be created by microbial molecular ecologists. A parallel code for the uncultured organisms would result in an alternative taxonomy exclusively for DNA-based classifications that would run independently of the decisions of the ICSP. But this divorce would surely have many other negative effects. To this purpose, taxonomists and molecular ecologists have again reinforced the need for an urgent solution (Murray *et al.*, 2020). The preferred Plan A would include a common nomenclature for cultured and uncultured taxa, fostering a harmonious classification. The less desirable Plan B would include a new code allowing uncultured taxa to be stably named (the UnCode) that would have at least the effect of halting chaos.

Never before has the future of the prokaryotic taxonomy been at such a critical point. The uncertain future, and whether Plan A or Plan B will prevail, depends totally on the wisdom of the ICSP. Perhaps by the time this book is published the situation will have been clarified, but the current situation is as full of uncertainty as of excitement, and none of the scenarios can be predicted.

In April 2020, a majority of the ICSP members decided to reject the proposals to use DNA as type material. Therefore, this rejection leaves the only path for microbial ecologists to go through Plan B. Only time will reveal whether this was the best decision for the future of the taxonomy for prokaryotes.

Acknowledgements

RRM acknowledges the financial support from the Spanish Ministry of Science and Innovation with the projects CLG2015_66686-C3-1-P, PGC2018-096956-B-C41, PRX18/00048 and RTC-2017-6405-1, also supported with European Regional Development Fund (FEDER).

References

Alejandre-Colomo, C., Harder, J., Fuchs, B.M., Rosselló-Móra, R. and Amann, R. (2020) High-throughput cultivation of heterotrophic bacteria along a spring phytoplankton bloom in the North Sea. *Systematic and Applied Microbiology* 43: 126066. https://doi.org/10.1016/j.syapm.2020.126066

Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A., Lawley, T.D. and Finn, R.D. (2019) A new genomic blueprint of the human gut microbiota. *Nature* 568, 499–504. https://doi.org/10.1038/s41586-019-0965-1

Amann, R.I., Ludwig, W. and Schleifer, K.-H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiology Reviews* 59, 143–169. https://doi.org/10.1128/MMBR.59.1.143-169.1995

Antón, J., Lucio, M., Peña, A., Cifuentes, A., Brito-Echeverría, J., Moritz, F., Tziotis, D., López, C., Urdiain, M., Schmitt-Kopplin, P. and Rosselló-móra, R. (2013) High metabolomic microdiversity within co-occurring isolates of the extremely halophilic bacterium *salinibacter ruber*. *PLoS ONE* 8(5): e64701. https://doi.org/10.1371/journal.pone.0064701

Armengaud, J. (2016) Next-generation proteomics faces new challenges in environmental biotechnology. *Current Opinion in Biotechnology* 38, 174–182. https://doi.org/10.1016/j.copbio.2016.02.025

Bergey, D.H. (1925) *Bergey's Manual of Determinative Bacteriology* 2nd edn. Williams & Wilkins Co., Baltimore, Maryland.
Bergey, D.H. (1930) *Bergey's Manual of Determinative Bacteriology* 3rd edn. Williams & Wilkins Co., Baltimore, Maryland.
Bergey, D.H. (1934) *Bergey's Manual of Determinative Bacteriology* 4th edn. Williams & Wilkins Co., Baltimore, Maryland. https://doi.org/10.1097/00000441-193408000-00030
Bergey, D.H., Harrison, F.C., Breed, R.S., Hammer, B.W. and Huntoon, F.M. (1923) *Bergey's Manual of Determinative Bacteriology*. Williams & Wilkins Co., Baltimore, Maryland.
Bergey, D.H., Breed, R.S., Murray, E.G.D. and Hitchens, A.P. (1939) *Bergey's Manual of Determinative Bacteriology* 5th edn. Williams & Wilkins Co., Baltimore, Maryland.
Bisgaard, M., Christensen, H., Clermont, D., Dijkshoorn, L., Janda, J.M., Moore, E.R.B., Nemec, A., Norskov-Lauritsen, N., Overmann, J. and Reubsaet, F.A.G. (2019) The use of genomic DNA sequences as type material for valid publication of bacterial species names will have severe implications for clinical microbiology and related disciplines. *Diagnostic Microbiolology and Infectious Diseases* 95, 102–103. https://doi.org/10.1016/j.diagmicrobio.2019.03.007
Breed, R.S., Murray, E.G.D. and Hitchens, A.P. (1948) *Bergey's Manual of Determinative Bacteriology* 6th edn. Williams & Wilkins Co., Baltimore, Maryland.
Breed, R.S., Murray, E.G.D. and Smith N.R. (1957) *Bergey's Manual of Determinative Bacteriology* 7th edn. Williams & Wilkins Co., Baltimore, Maryland.
Brenner, D.J., Krieg, N.R., Staley, J.T. and Garrity, G.M. (eds) (2005) Proteobacteria. *Bergey's Manual of Systematic Bacteriology* 2nd edn, vol. 2, parts A, B and C. Springer-Verlag, New York, NY. https://doi.org/10.1007/0-387-28021-9
Brito-Echeverría, J., Lucio, M., López-López, A., Antón, J., Schmitt-Kopplin, P. and Rosselló-Móra, R. (2011) Response to adverse conditions in two strains of the extremely halophilic species *Salinibacter ruber*. *Extremophiles* 15, 379–389. https://doi.org/10.1007/s00792-011-0366-3
Buchanan, R.E. and Gibbons, N.E. (1974) *Bergey's Manual of Determinative Bacteriology* 8th edn. Williams & Wilkins Co., Baltimore, Maryland.
Christensen, H., Bisgaard, M., Frederiksen, W., Mutters, R., Kuhnert, P. and Olsen, J.E. (2001) Taxonomic note: is characterization of a single isolate sufficient for valid publication of a new genus or species? Proposal on reformulation of recommendation 30b of the Bacteriological Code. *International Journal of Systematic and Evolutionary Microbiology* 51, 2221–2225. https://doi.org/10.1099/00207713-51-6-2221
Chun, J. and Rainey, F.A. (2014) Integrating genomics into the taxonomy and systematics of Bacteria and Archaea. *International Journal of Systematic and Evolutionary Microbiology* 64, 316–324. https://doi.org/10.1099/ijs.0.054171-0
Cohn, F. (1872) Untersuchungen über Bacterien. *Beiträge zur Biologie der Pflanzen* 1(2), 127–224.
Colwell, R.R. (1970) Polyphasic taxonomy of the genus *Vibrio*: Numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species. *Journal of Bacteriology* 104, 410–433. https://doi.org/10.1128/JB.104.1.410-433.1970
Copeland, H. (1938) The kingdoms of organisms. *Quarterly Review of Biology* 13, 383–420. https://doi.org/10.1086/394568
Cowan, S.T. (1965) Principles and practice of bacterial taxonomy - a forward look. *Journal of General Microbiology* 39, 148–159. https://doi.org/10.1099/00221287-39-1-143
De Kruif, P. (1926) *Microbe Hunters*. 1996 Reprint. Houghton Mifflin Harcourt, USA. New https://doi.org/10.2307/3221690
Drancourt, M., Fournier, P.E. (2018) New species announcement 2.1. *New Microbes New Infections* 25, 48. https://doi.org/10.1016/j.nmni.2018.06.009
Ehrenberg, C.G. (1832) Beiträge zur Kenntnis der Organization der Infusorien und ihrer geographischen Verbreitung besonders in Sibirien. *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin* 1830, 1–88. https://doi.org/10.5962/bhl.title.143632
Ehrenberg, C.G. (1835) Dritter Beitrag zur Erkenntniss grosser Organisation in der Richtung des kleinsten Raumes. *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin aus den Jahren* 1833-1835 143–336.
Ehrenberg, C.G. (1838) Die Infusionsthierchen als vollkommene Organismen. Ein Blick in das tiefere organische Leben der Natur. Voss, Leipzig 1838. https://doi.org/10.5962/bhl.title.97605
Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F.L., Swings, J. (2005) Opinion: Re-evaluating prokaryotic species. *Nature Reviews Microbiology* 3, 733–739. https://doi:10.1038/nrmicro1236

Gibbons, N.E. and Murray, R.G.E. (1978) Proposals concerning the higher taxa of bacteria. *International Journal of Systematic Bacteriology* 28, 1–6. https://doi.org/10.1099/00207713-28-1-1

Goodfellow, M., Stackebrandt, E. and Kroppenstedt, R.M. (1988) Chemotaxonomy and actinomycete systematics In: Okami. Y., Beppu, T. and Ogawara, H. (Eds) *Biology of Actinomycetes*. Japan Scientific Societies Press, Tokyo, Japan, pp. 233–238.

Goodsir, J. (1842) History of a case in which a fluid periodically ejected from the stomach contained vegetable organisms of an undescribed form. *Edinburgh Medical Surgery Journal* 57, 430–443.

Heather, J.M. and Chain, B. (2016) The sequence of sequencers: The history of sequencing DNA. *Genomics* 1-8. https://doi.org/10.1016/j.ygeno.2015.11.003

Hedlund, B.P., Dodsworth, J.A., Staley, J.T. (2015) The changing landscape of microbial biodiversity exploration and its implications for systematics. *Systematic and Applied Microbiology* 38, 231–236. https://doi.org/10.1016/j.syapm.2015.03.003

Holt, J.G.B (1994) *Bergey's Manual of Systematic Bacteriology* 1st edn. Williams & Wilkins Co., Baltimore, Maryland.

Jain, C., Rodriguez-R, L.M., Phyllippy, A.M., Konstantinidis, K.T. and Aluru, S. (2018) High-throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications* 9, 5114. https://doi.org/10.1038/s41467-018-07641-9

Karthikeyan, S., Rodriguez-R, J.M., Heritier-Robbins, P., Kim, M., Overholt, W.A., Gaby, J.C., Hatt, J.K., Spain, J., Rosselló-Móra, R., Huettel, M., Kostka, J. and Konstantinidis, K. (2019) "Candidatus Macondimonas diazotrophica" a novel gammaproteobacterial genus dominating crude-oil-contaminated coastal sediments. *The ISME Journal* 13, 2129–2134. https://doi.org/10.1038/s41396-019-0400-5

Konstantinidis, K. and Rosselló-Móra, R. (2015) Classifying the uncultivated microbial majority: a place for metagenomic data in the Candidatus proposal. *Systematic and Applied Microbiology* 38, 223–230. https://doi.org/10.1016/j.syapm.2015.01.001

Konstantinidis, K. and Tiedje, J.M. (2005a) Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences* USA 102, 2567–2592. https://doi.org/10.1073/pnas.0409727102

Konstantinidis, K.T. and Tiedje, J.M. (2005b) Towards a genome-based taxonomy for prokaryotes. *Journal of Bacteriology* 187, 6258–6264. https://doi.org/10.1128/JB.187.18.6258-6264.2005

Konstantinidis, K., Rosselló-Móra, R. and Amann, R. (2017) Uncultivated microbes in need of their own taxonomy. *The ISME Journal* 11, 2399–2406. https://doi.org/10.1038/ismej.2017.113

Konstantinidis, K., Rosselló-Móra, R. and Amann, R. (2018) Reply to the commentary "Uncultivated microbes-in need of their own nomenclature?". *The ISME Journal* 12, 653–654. https://doi.org/10.1038/s41396-017-0011-y

Lagier, J.C., Khelaifia, S. and Tidjani-Alou, M. (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. *Nature Microbiology* 1, 16203. https://doi.org/10.1038/nmicrobiol.2016.203

Lagier, J.C., Dubourg, G., Million, M., Cadoret, F., Bilen, M., Fenollar, F., Levasseur, A., Rolain, J.M., Fournier, P.E. and Raoult, D. (2018) Culturing the human microbiota and culturomics. *Nature Reviews Microbiology* 16, 540–550. https://doi.org/10.1038/s41579-018-0041-0

Lapage, S.P., Sneath, P.H.A., Lessel, E.F., Skerman, V.B.D., Seeliger, H.P.R. and Clark, W.A. (1992) *International Code of Nomenclature of Bacteria*. American Society for Microbiology, Washington, D.C.

Migula, W. (1900) *System der Bakterien*. Jena Gustav Fischer Verlag.

Müller, O.F. (1786) Animalcula infusoria; fluvia tilia et marina. Hauniae: Typis Nicolai Mölleri (in Latin).

Mullis, K.F., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harbor Symposium in Quantitative Biology* 51, 263–273. https://doi.org/10.1101/SQB.1986.051.01.032

Munoz, R., Teeling, H., Amann, R. and Rosselló-Móra, R. (2020) Ancestry and adaptive radiation of Bacteroidetes as assessed by comparative genomics. *Systematic and Applied Microbiology* 43, 126065. https://doi.org/10.1016/j.syapm.2020.126065

Murray, A.E., Freudenstein, J., Gribaldo, S., Hatzenpichler, R., Hedlund, B.P., Hugenholtz, P., Kämpfer, P., Konstantinidis, K.T., Lane, C., Papke, R.T., Parks, D.H., Reysenbach, A-L, Rosselló-Móra, R., Stott, M., Sutcliffe, I.C., Thrash, J.C., Venter, S.N., Whitman, W.B. *et al*. (2020) Roadmap for naming uncultivated Archaea and Bacteria. *Nature Microbiology* 5, 987–994. https://doi.org/10.1038/s41564-020-0733-x.

Murray, R.G. and Stackebrandt, E. (1995) Taxonomic note: implementation of the provisional status Candidatus for incompletely described procaryotes. *International Journal of Systematic Bacteriology* 45, 186–187. https://doi.org/10.1099/00207713-45-1-186

Murray, R.G.E., Brenner, D.J., Colwell, R.R., DeVos, P., Goodfellow, M., Grimont, P.A.D., Pfennig, N., Stackebrandt, E. and Zavarzin, G.A. (1990). Report of the ad hoc Committee on Approaches to Taxonomy within the Proteobacteria. *International Journal of Systematic Bacteriology* 40, 213–215. https://doi.org/10.1099/00207713-40-2-213

Musat, N., Musat, F., Weber, P.K. and Pett-Ridge, J. (2016) Tracking microbial interactions with NanoSIMS. *Current Opinion in Biotechnology* 41, 114–121. https://doi.org/10.1016/j.copbio.2016.06.007

Oren, A. (2017) A plea for linguistic accuracy - also for Candidatus taxa. *International Journal of Systematic and Evolutionary Microbiology* 67, 1085–1094. https://doi.org/10.1099/ijsem.0.001715

Oren, A. and Garrity, G. (2018) Commentary: Uncultivated microbes - in the need of their own nomenclature? *The ISME Journal* 12, 309–311. https://doi.org/10.1038/ismej.2017.188

Oren, A., Garrity, G.M. and Parte, A.C. (2018) Why are so many effectively published names of prokaryotic taxa never validated? *International Journal of Systematic and Evolutionary Microbiology* 68, 2125–2129. https://doi.org/10.1099/ijsem.0.002851

Overmann, J., Huang, S., Nübel, U., Hahnke, R.L. and Tindall, B.J. (2019) Relevance of phenotypic information for the taxonomy of not-yet-cultured microorganisms. *Systematic and Applied Microbiology* 42, 22–29. https://doi.org/10.1016/j.syapm.2018.08.009

Pacini, F. (1854) Osservazione microscopiche e deduzioni patologiche sul cholera asiatico. *Gazette Medicale de Italiana Toscano Firenze* 6, 405–412.

Parish, T., Harris, M., Fry, N., Mathee, K., Trujillo, M.E., Bentley, S. and Thomson, N. (2019) DORA Editorial. *International Journal of Systematic and Evolutionary Microbiology*. 69, 1–2. https://doi.org/10.1099/ijsem.0.003172

Parker, C.T., Tindall, B.J. and Garrity, G.M. (2019) International Code of Nomenclature of Prokaryotes. *International Journal of Systematic and Evolutionary Microbiology* 69, S1–S111. https://doi.org/10.1099/ijsem.0.000778

Parks, D., Chuvochina, M., Waite, D.V., Rinke, C., Skarshewsky, A., Chaumeil, P.A. and Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology* 36, 996–1004. https://doi.org/10.1038/nbt.4229

Parte, A.C. (2018) LPSN - list of prokaryotic names with standing in nomenclature (bacterio.net), 20 years on. *International Journal of Systematic and Evolutionary Microbiology*. 68, 1825–1829. https://doi.org/10.1099/ijsem.0.002786

Peña-Gonzalez, A., Soto-giron, M.J., Smith, S., Sistrunk, J., Montero, L., Paez, M., Ortega, E., Hatt, J.K., Cevallos, W., Trueba, G., Levy, K. and Konstantinidis, K. (2019) Metagenomic signatures of gut infections caused by different *Escherichia coli* pathotypes. *Applied and Environmental Microbiology* 85: e01820-19. https://doi.org/10.1128/AEM.01820-19

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41, D590–D596. https://doi.org/10.1093/nar/gks1219

Ramos-Barbero, M.D., Martin-Cuadrado, A.-B., Viver,T., Santos, F., Martinez-Garcia, M. and Anton, J. (2019) Recovering microbial genomes from metagenomes in hypersaline environments. The good, the bad and the ugly. *Systematic and Applied Microbiology* 42, 30–40. https://doi.org/10.1016/j.syapm.2018.11.001

Richter, M. and Rosselló-Móra, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences USA* 106, 19126–19131. https://doi.org/10.1073/pnas.0906412106

Rodriguez-R, L.M., Gunturu, S., Harvey, W.T., Rosselló-Móra, R., Tiedje, J.M., Cole, J.R. and Konstantinidis, K.T. (2018) The microbial genomes atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic Acids Research* 46(W1), W282–W288. https://doi.org/10.1093/nar/gky467

Rosselló-Móra, R. (2012) Towards a taxonomy of Bacteria and Archaea based on interactive and cumulative data repositories. *Environmental Microbiology* 14, 318–334. https://doi.org/10.1111/j.1462-2920.2011.02599.x

Rosselló-Móra, R. and Amann, R. (2015) Past and future species definitions for Bacteria and Archaea. *Systematic and Applied Microbiology* 38, 209–216. https://doi.org/10.1016/j.syapm.2015.02.001

Rosselló-Móra, R. and Sutcliffe, I. (2019a) Reflections on the introduction of the Digital Protologue Database - a partial success? *Systematic and Applied Microbiology* 42, 1–2 https://doi.org/10.1016/j.syapm.2018.12.002

Rosselló-Móra, R. and Sutcliffe, I. (2019b) Reflections on the introduction of the Digital Protologue Database - a partial success? *Antonie van Leeuwenhoek* 112, 141–143. https://doi.org/10.1007/s10482-018-01221-z

Rosselló-Móra, R. and Whitman, W.B. (2019) Dialogue on the nomenclature and classification of prokaryotes. *Systematic and Applied Microbiology* 42, 5–14. https://doi.org/10.1016/j.syapm.2018.07.002

Rosselló-Mora, R., Lee, N., Antón, J. and Wagner, M. (2003) Substrate uptake in extremely halophilic microbial communities revealed by microautoradiography and fluorescence in situ hybridisation. *Extremophiles* 5, 409–413. https://doi.org/10.1007/s00792-003-0336-5

Rosselló-Móra, R., Lucio, M., Peña, A., Brito-Echeverría, J., López-López, A., Valens-Vadell, M., Frommberger, M., Antón, J. and Schmitt-Kopplin, P. (2008) Metabolic evidence for biogeographic isolation of the extremophilic bacterium *Salinibacter ruber*. *The ISME Journal* 2, 242–253. https://doi.org/10.1038/ismej.2007.93

Rosselló-Móra, R., Urdiain, M. and López-López, A. (2011) DNA-DNA hybridization. In: Rainey, F. and Oren, A. (eds) *Methods in Microbiology* Volume 38. Elsevier Ltd. pp. 325–347. https://doi.org/10.1016/B978-0-12-387730-7.00015-2

Rosselló-Móra, R., Trujillo, M. and Sutcliffe, I.C. (2017a) Introducing a digital protologue: a timely move towards a database-driven systematics of archaea and bacteria. *Systematic and Applied Microbiology* 40, 121–122. https://doi.org/10.1016/j.syapm.2017.02.001

Rosselló-Móra, R., Trujillo, M. and Sutcliffe, I.C. (2017b) Introducing a digital protologue: a timely move towards a database-driven systematics of archaea and bacteria. *Antonie van Leewenhoek* 110, 455–456. https://doi.org/10.1007/s10482-017-0841-7

Sanger, F. and Coulson, A.R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* 93, 441–448. https://doi.org/10.1016/0022-2836(75)90213-2

Sapp, J. (2005) The Prokaryote-Eukaryote Dichotomy: Meanings and Mythology. *Microbiology and Molecular Biology Reviews* 69, 292–305. https://doi.org/10.1128/MMBR.69.2.292-305.2005

Sapp, J. (2009) *The New Foundation of Evolution: On the Tree of Life*. Oxford University Press, New York.

Sogin, S.J., Sogin, M.L. and Woese, C.R. (1971) Phylogenetic measurement in procaryotes by primary structural characterization. *Journal of Molecular Evolution* 1, 173–184. https://doi.org/10.1007/BF01659163

Sogin, S.J., Sogin, M.L. and Woese, C.R. (1972) Phylogenetic measurement in procaryotes by primary structural characterization. *Journal of Molecular Evolution* 1, 173–184.

Sokal, R.R. and Sneath, P.H.A. (1963) *Principles of Numerical Taxonomy*. W.H. Freeman, San Francisco, CA.

Stackebrandt, E. and Ebers, J. (2006) Taxonomic parameter revisited: tarnished gold standards. *Microbiology Today* 33, 152–155.

Stackebrandt, E. and Smith D. (2017a) Expanding the 'Digital Protologue' database (DPD) to 'Archives of Microbiology': an offer to scientists and science. *Archives of Microbiology* 199, 519–520. https://doi.org/10.1007/s00203-017-1369-y

Stackebrandt, E. and Smith, D. (2017b) Expanding the 'Digital Protologue' database (DPD) to 'Current Microbiology': an offer to scientists and science. *Current Microbiolology* 74, 1003–1004. https://doi.org/10.1007/s00284-017-1290-2

Stackebrandt, E, Murray, R.G.E. and Trüper, H.G. (1988) Proteobacteria classis nov., a name for the phylogenetic taxon that includes the purple bacteria and their relatives. *International Journal of Systematic Bacteriology* 38, 321–325. https://doi.org/10.1099/00207713-38-3-321

Stackebrandt, E., Frederiksen, W., Garrity, G.M., Grimont, P.A.D., Kämpfer, P., Maiden, M.C.J., Nesme, X., Rosselló-Móra, R., Swings, J., Trüper, H.G., Vauterin, L., Ward, A.C. and Whitman, W.B. (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology* 52, 1043–1052. https://doi.org/10.1099/00207713-52-3-1043

Sutcliffe, I.C. (2015) Challenging the anthropocentric emphasis on phenotypic testing in prokaryotic species descriptions: rip it up and start again. *Frontiers in Genetics* 6, 218. https://doi.org/10.3389/fgene.2015.00218

Sutcliffe, I.C. (2019) Valediction: descriptions of novel prokaryotic taxa published in Antonie van Leeuwenhoek-change in editorial policy and a signpost to the future? *Antonie Van Leeuwenhoek* 112, 1281–1282. https://doi.org/10.1007/s10482-019-01311-6

Sutcliffe, I.C., Trujillo, M.E. and Goodfellow, M. (2012) A call to arms for systematists: revitalising the purpose and practices underpinning the description of novel microbial taxa. *Antonie Van Leeuwenhoek* 101, 13–20. https://doi.org/10.1007/s10482-011-9664-0

Tamames, J. and Rosselló-Móra, R. (2012) On the fitness of microbial taxonomy. *Trends in Microbiology* 20, 514–516. https://doi.org/10.1016/j.tim.2012.08.012

Uchida, T., Bonen, L., Schaup, H.W., Lewis, B.J., Zablen, L. and Woese, C. (1974) The use of ribonuclease U2 in RNA sequence determination. Some corrections in the catalog of oligomers produced by ribonuclease T1 digestion of *Escherichia coli* 16S ribosomal RNA. *Journal of Molecular Evolution* 3, 63–77. https://doi.org/10.1007/BF01795977

van Niel, C.B. (1955) Classification and taxonomy of the bacteria and blue green algae. In: Kessel, E.L. (ed.), *A Century of Progress in the Natural Sciences, 1853–1953*. California Academy of Sciences, San Francisco, CA, pp. 89–114.

Viver, T., Cifuentes, A., Díaz, S., Rodríguez-Valdecantos, G., González, B., Antón, J. and Rosselló-Móra, R. (2015) Diversity of extremely halophilic cultivable prokaryotes in Mediterranean, Atlantic and Pacific solar salterns: evidence that unexplored sites constitute sources of cultivable novelty. *Systematic and Applied Microbiology* 38, 266–275. https://doi.org/10.1016/j.syapm.2015.02.002

Viver, T., Orellana, L.H., Díaz, S., Urdiain, M., Ramos-Barbero, M.D., Gonzalez-Pastor, J.E., Oren, A., Hatt, J.K., Amann, R., Anton, J., Konstantinidis, K.T. and Rosselló-Móra, R. (2019) Predominance of deterministic microbial community dynamics in salterns exposed to different light intensities. *Environmental Microbiology* 21, 4300–4315. https://doi.org/10.1111/1462-2920.14790

Wayne, L., 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., Stackebrandt, E., Starr, M.P. and Trüper, H.G. (1987) International Committee on Systematic Bacteriology: Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic Bacteriology* 37, 463–464. https://doi.org/10.1099/00207713-37-4-463

Welker, M. and Moore, E.R.B. (2011) Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. *Systematic and Applied Microbiology* 34, 2–11. https://doi.org/10.1016/j.syapm.2010.11.013

Whitman, W.B. (2015) Genome sequences as the type material for taxonomic descriptions of prokaryotes. *Systematic and Applied Microbiology* 38, 217–222. https://doi.org/10.1016/j.syapm.2015.02.003

Whitman, W.B. (2016) Modest proposals to expand the type material for naming of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology* 66, 2108–2112. https://doi.org/10.1099/ijsem.0.000980

Whitman, W.B., Oren, A., Chuvochina, M., da Costa, M.S., Garrity, G.M., Rainey, F.A., Rosselló-Móra, R., Schink, B., Sutcliffe, I., Trujillo, M.E. and Ventura, S. (2018) Proposal of the suffix -ota to denote phyla. Addendum to 'proposal to include the Rank of phylum in the International Code of Nomenclature of Prokaryotes'. *IJSEM* 68, 967–969. https://doi.org/10.1099/ijsem.0.002593

Whitman, W.B., Sutcliffe, I.C. and Rosselló-Móra, R. (2019) Proposal for changes in the International Code of Nomenclature of Prokaryotes: granting priority to Candidatus names. *International Journal of Systematic and Evolutionary Microbiology* 69, 2174–2175. https://doi.org/10.1099/ijsem.0.003419

Woese, C.R., Sogin, M.L., Bonen, L. and Stahl, D. (1975) Sequence studies on 16S ribosomal RNA from a blue-green alga. *Journal of Molecular Evolution* 4, 307–315. https://doi.org/10.1007/BF01732533

Woese, C.R., and Fox, G.E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences USA* 74, 5088–5090. https://doi.org/10.1073/pnas.74.11.5088

Woese, C.R., Stackebrandt, E. and Ludwig, W. (1984) What are mycoplasmas: the relationship of tempo and mode in bacterial evolution. *Journal of Molecular Evolution* 21, 305–316. https://doi.org/10.1007/BF02115648

Woese, C.R., Debrunner-Vossbrinck, B.A., Oyaizu, H., Stackebrandt, E. and Ludwig, W. (1985) Gram-positive bacteria: possible photosynthetic ancestry. *Science* 229, 762–765. https://doi.org/10.1126/science.11539659

Woese, C.R., Kandler, O. and Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences USA* 87, 4576–4579. https://doi.org/10.1073/pnas.87.12.4576

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., Whitman, W., Euzéby, J., Amann, R. and Rosselló-Móra, R. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews in Microbiology* 12, 635–645. https://doi.org/10.1038/nrmicro3330

Zavarzin, G.A., Stackebrandt, E. and Murray, R.G. (1991) A correlation of phylogenetic diversity in the Proteobacteria with the influences of ecological forces. *Canadian Journal of Microbiology* 37, 1–6. https://doi.org/10.1139/m91-001

Zuckerkandl, E. and Pauling, L. (1965) Evolutionary divergence and convergence in proteins. In: Bryson, V. and Vogel, H.J., (eds) *Evolving Genes and Proteins*. Academic Press, New York, pp. 97–166. https://doi.org/10.1016/B978-1-4832-2734-4.50017-6

Zuñiga, C., Zaramela, L. and Zengler, K. (2017) Elucidation of complexity and prediction of interactions in microbial communities. *Microbial Biotechnology* 6, 1500–1522. https://doi.org/10.1111/1751-7915.12855