Prokaryote diversity and taxonomy: current status and future challenges

Aharon Oren

The Institute of Life Sciences, and the Moshe Shilo Minerva Center for Marine Biogeochemistry, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel (orena@shum.cc.huji.ac.il)

The prokaryotes are by far the most abundant organisms inhabiting planet Earth. They are also by far the most diverse, both metabolically and phylogenetically; they encompass the Bacteria and the Archaea, two out of the three major divisions of living organisms. The current prokaryote species classification is based on a combination of genomic and phenotypic properties. The recommended cut-off value of 70% DNA–DNA similarity to delineate species signifies an extremely broad species definition for the prokaryotes compared with the higher eukaryotes. The number of validly named species of prokaryotes is currently slightly more than 6200. However, on the basis of small-subunit rDNA characterization of whole communities and other approaches, the more exact number of species present can be inferred to be at least two orders of magnitude larger. Classic culturing methods based on colony formation on agar are generally unsatisfactory for the recovery of bacteria from the environment. Many of the most abundant prokaryotes in nature have not yet been brought into culture. Some of these may thrive by means of as yet unknown modes of energy generation. Several novel methods have recently enabled the isolation of some interesting organisms of environmental significance. A better coverage of the prokaryote diversity on Earth depends on such innovative approaches, combined with appropriate funding.

Keywords: prokaryotes; Archaea; Bacteria; biodiversity; taxonomy

1. INTRODUCTION: THE ABUNDANCE OF BACTERIA AND ARCHAEA ON EARTH

The prokaryotes are by far the most abundant organisms inhabiting planet Earth. They are also phylogenetically the most diverse, as two out of the currently recognized three major divisions (domains) of living organisms (Bacteria, Archaea and Eucarya) consist of prokaryotes (figure 1). They thus represent a large proportion of life's genetic diversity. Moreover, the prokaryotes are metabolically far more diverse than the eukaryotic organisms, and they are responsible for many of the key processes in the biogeochemical cycling on Earth (Lengeler *et al.* 1999; Madigan *et al.* 2003).

This article attempts to provide an overview of our current understanding of prokaryote diversity and of the present status of the taxonomy of the Bacteria and the Archaea, with special emphasis on recent developments in these fields, while attempting to identify future challenges and directions. Only in recent years have we started to obtain some insight into the true diversity of the prokaryotes. It has become clear that only a small fraction of all prokaryote species have been brought into culture and characterized. Current estimates range from less than one, or one-tenth of a per cent, to even much less according to some opinions (Torsvik *et al.* 1996, 1998; Dykhuizen 1998; DeLong & Pace 2001). The exploration and description of the prokaryotic diversity requires a

taxonomic framework defining species, genera and higher-order taxa of Bacteria and Archaea. It is difficult to discuss prokaryote taxonomy, including our current views as to how a prokaryote species should be defined, without a proper understanding of the diversity within the prokaryote world. Prokaryote diversity is therefore closely linked to prokaryote taxonomy.

It is fitting to start this overview by emphasizing the quantitative importance of the prokaryotes among living organisms. Almost 50 years ago, Kluyver and van Niel had already recognized that approximately one-half of the 'living protoplasm' on Earth is microbial: '... the total weight of microbial protoplasm on earth exceeds that of animal protoplasm by many times. Ignoring the microbe would obviously mean that a very considerable part—perhaps almost one-half—of the living protoplasm on earth is left out of consideration' (Kluyver & van Niel 1956, pp. 3–4). Whitman *et al.* (1998) attempted to make an inventory of prokaryotes on Earth, based on the latest estimates of their numbers in different ecosystems (table 1). Of $4\text{--}6 \times 10^{30}$ prokaryotic cells, most are found in the open ocean, in soil and in oceanic and terrestrial subsurfaces (1.2×10^{29} , 2.6×10^{29} , 3.5×10^{30} and $0.25\text{--}2.5 \times 10^{30}$, respectively). Typical densities in continental shelf waters and in the upper 200 m of the ocean are *ca.* 5×10^5 cells ml^{-1} , decreasing to 5×10^4 cells ml^{-1} in the deep ocean. Freshwater and saline lakes generally contain *ca.* 10^6 cells ml^{-1} . The Bacteria and Archaea present in the gastrointestinal tracts of animals contribute relatively little: the number of prokaryotes found in the bovine rumen (2.9×10^{24} in 1.3×10^9 animals) is 4–6 orders of magnitude less than the numbers found in soil, the subsurface and seawater.

One contribution of 19 to a Theme Issue 'Taxonomy for the twenty-first century'.

Figure 1. Universal, small-subunit rRNA-based phylogenetic tree, showing the three domains: Bacteria, Archaea and Eucarya. The exact branching order of the 'Korarchaeota' and the 'Nanoarchaeota' is uncertain.

Table 1. Number and biomass of prokaryotes in the world.
(Modified from Whitman *et al.* (1998) and reproduced with permission.)

environment	number of prokaryotic cells $\times 10^{28}$	carbon in prokaryote biomass ($\times 10^{15}$ g)
aquatic habitats	12	2.2
oceanic subsurface	355	303
soil	26	26
terrestrial subsurface	25–250	22–215
total	415–640	353–546

The prokaryotes present in the colon of 5.6×10^9 humans contribute 3.9×10^{23} cells.

With an estimated total biomass of $350\text{--}550 \times 10^{15}$ g of carbon, prokaryotic cells contain as much as 60–100% of the estimated total carbon in plants (including extracellular material such as cell walls and structural polymers; Whitman *et al.* 1998). Inclusion of the prokaryotic carbon in global models of the carbon cycle will therefore almost double the estimates of the amount of carbon stored in living organisms. In addition, the prokaryotes contain $85\text{--}130 \times 10^{15}$ g of nitrogen and $9\text{--}14 \times 10^{15}$ g of phosphorus, i.e. approximately 10-fold more than found in plants. Turnover of prokaryotic biomass is rapid; the production rate for all prokaryotes on Earth is estimated at 1.7×10^{30} cells yr^{-1} . Rates are especially high in the open ocean: 8.2×10^{29} cells of marine heterotrophs per year in marine systems in the upper 200 m, corresponding to a turnover time of 16 days. For comparison, prokaryote production in soil was estimated to be 1×10^{29} cells yr^{-1} , with a turnover time of 900 days. The conclusion by Whitman *et al.* (1998, p. 6582) that 'Given prokaryotes' numerical abundance and importance in biogeochemical transformations, the absence of detailed knowledge of prokaryotic

diversity is a major omission in our knowledge of life on Earth' thus appears fully justified.

When molecular approaches based on sequencing of small-subunit rRNA (16S in prokaryotes, 18S in eukaryotes) were introduced by Carl Woese in the late 1970s, it became clear that the prokaryotes do not form a single, phylogenetically coherent group, but consist of two fundamentally different groups, the divisions (domains) Archaea (formerly called Archaeobacteria) and Bacteria (formerly named Eubacteria) (Woese 1987, 1992; Woese *et al.* 1990). This splitting-up of the prokaryotes into Archaea and Bacteria is now generally accepted. An in-depth discussion of alternative models describing prokaryote phylogeny (e.g. Gupta 1998) is outside the scope of this review. The Archaea and the Bacteria differ in many properties. These include the structure of the cell wall, the nature of the membrane lipids, the properties of the protein synthesizing system, sensitivity to different antibiotics, and many others. Initially, the Archaea were considered as a somewhat exotic group of microorganisms, mostly restricted to extreme environments. However, based on molecular ecological approaches, it is now clear that Archaea are widespread (Olsen 1994), and

are also abundant in less extreme ecosystems such as seawater. Karner *et al.* (2001) estimated the global oceans to harbour *ca.* 1.3×10^{28} archaeal and 3.1×10^{28} bacterial cells, numbers that total less than the total oceanic prokaryote number of 1.2×10^{29} given by Whitman *et al.* (1998) (see also § 4).

2. THE DEFINITION OF A BACTERIAL SPECIES

From the early days of microbiology as a science, microbiologists realized the difficulties in establishing a satisfactory classification system for prokaryotes. Bacteria and Archaea are morphologically simple and undistinguished in most cases, so morphological characters are of little use. Moreover, a useful fossil record is altogether lacking and the existing fossils are phylogenetically uninformative.

The description of species of prokaryotes is based on living cultures, and one isolate is designated as the nomenclatural type. The basis of the taxonomic hierarchy is the species. However, the concept of a prokaryote species still lacks a theoretical basis, and all existing definitions are pragmatic ones, such as, for example: 'A species consists of an assemblage of individuals (or, in micro-organisms, of clonal populations) that share a high degree of phenotypic similarity, coupled with an appreciable dissimilarity from other assemblages of the same general kind', or: 'a collection of strains showing a high degree of overall similarity, compared to other, related groups of strains' (Colwell *et al.* 1995). Ward (1998) called for the establishment of a 'natural' species concept for the prokaryotes, based on 'evolutionary species', being defined as lineages evolving separately from others and with their own unitary evolutionary roles and tendencies. The species definition is so extremely subjective because one cannot accurately determine and define such concepts as 'a close resemblance', 'essential features', or how many 'distinguishing features' are sufficient to create a species.

In-depth essays on the current thoughts on the species concept for prokaryotes were recently published by Dykhuizen & Green (1991), Istock *et al.* (1996), Palys *et al.* (1997), Stackebrandt (2000) and Rosselló-Mora & Amann (2001). The species should ideally be described as 'a category that circumscribes a (preferable) genomically coherent group of individual isolates/strains sharing a high degree of similarity in (many) independent features, comparatively tested under highly standardized conditions' (the phylo-phenetic species concept). Our present species concept is, to a large extent, based on the recommendations published in 1987 by a committee of experts (Wayne *et al.* 1987), recently confirmed and extended by a new *ad hoc* committee (Stackebrandt *et al.* 2002). Species classification is based on a combination of diagnostic phenotypic features and genomic properties. It is widely accepted that an adequate classification of prokaryotes, in particular of the lower taxonomic ranks such as species, can only be achieved when a 'polyphasic approach' is undertaken in which both genomic and phenotypic characteristics are taken into account. Relevant genomic properties are primarily the base ratio in the DNA (expressed as the molar percentage of G+C) (not for species of an established genus) and the DNA–DNA hybridization similarity as measured by reassociation. Consistency

of phenotypic and genomic characters is required to generate a useful classification system for the prokaryotes (Vandamme *et al.* 1996). Individually, many of the phenotypic and chemotaxonomic characters used are insufficient as parameters to delineate species, yet as a whole they provide descriptive information enabling us to recognize taxa.

Genotypic clustering has largely replaced phenotypic clustering for demarcating bacterial species (Wayne *et al.* 1987). Probably the most widely accepted pragmatic definition of a prokaryotic species defines a species as a group of strains, including the type strain, that share *ca.* 70% or higher total genome DNA–DNA hybridization and have less than 5 °C ΔT_m (the difference in the melting temperature between the homologous and the heterologous hybrids formed under standard conditions, being a reflection of the thermal stability of DNA duplexes). Thus, total genome DNA–DNA hybridization values are the key parameter in this species delineation, as DNA reassociation values are an indirect expression of the genome sequence identity. The species concept is thus based on whole genome similarities. Unfortunately, DNA–DNA similarity studies are time-consuming and they rely on pairwise comparisons, thus experiments are normally performed with a relatively small set of organisms. The method gives no indication of which genes contribute to or detract from the similarity. At the level of genera or higher taxa, DNA–DNA has limited resolving power: a 50% reassociation level may mean that the strains evolved isochronically from each other or that one of the two lacks a significant portion of the genome.

The arbitrarily determined value of *ca.* 70% DNA–DNA similarity to delineate species signifies an extremely broad species definition for the prokaryotes compared with, for example, the higher eukaryotes. For comparison, humans and chimpanzees share 98.4% relatedness on the basis of DNA–DNA hybridization; the similarity of human and gorilla or orang-utan DNA is 97.7% and 96.5%, respectively, and based on the 70% DNA–DNA similarity criterion even humans and lemurs should be classified within the same species (78% similarity). Many of the extremely important properties of prokaryotes vary at the subspecies level. Features such as antibiotic production and virulence of pathogens are strain dependent and vary greatly between members of the same species. Thus, the prokaryotic species includes organisms that are functionally as well as genetically different, and should therefore be compared with some deeper eukaryotic classification level such as the tribe.

The 70%–5 °C ΔT_m boundary has proven to be satisfactory in most cases, but there are some cases in which the 70% criterion is too conservative for prokaryotes. Thus, *Neisseria gonorrhoeae* and *N. meningitidis* are considered separate species even though the percentage of DNA that reannealed was 93%. Based on DNA hybridization alone, *N. lactamica* and *N. polysaccharaea* would also have to be included in this species (Hoke & Vedros 1982).

Another genotypic characteristic that is extensively used in prokaryote taxonomy is the nucleotide sequence of the small-subunit (16S) rRNA (see also § 4). Generally there is a good correlation between the DNA–DNA similarity and the similarity of the 16S rDNA sequence, 50% DNA–DNA pairing corresponding to *ca.* 99% 16S rDNA similarity. However, comparative studies clearly revealed the

limitations of the 16S rDNA sequence analysis in the determination of relationships at the strain level, for which DNA–DNA reassociation experiments still constitute the superior method (Stackebrandt & Goebel 1994). It must be stressed, however, that the 16S rDNA sequence alone should not be used to delineate species, and there are even cases known in which two distinctly different species share an identical 16S rDNA sequence. Such is, for example, the case for *Sporosarcina globispora* and *S. psychrophila* (DNA–DNA similarity less than 50%; Fox *et al.* 1992). In an analysis of the genera *Aeromonas* and *Plesiomonas* (γ -Proteobacteria) the levels of the relationships derived from analysis of rDNA sequence data were in marked disagreement with the results of chromosomal DNA–DNA pairing experiments (Martínez-Murcia *et al.* 1992).

The development of new methodological approaches and the increased insight into microbial population structure (see also § 4), was the incentive to form an *ad hoc* committee which convened in Gent, Belgium, in 2002 to reassess the bacterial species concept (Stackebrandt *et al.* 2002). It was concluded that despite certain drawbacks with respect to reproducibility, workability, and rigid application of DNA–DNA hybridization values for species delineation, the currently used system is sound. The current species definition is pragmatic, operational and universally applicable. The committee considered several techniques of great promise: for example, sequencing of several housekeeping genes encoding universal metabolic functions (see Palys *et al.* 1997; Lan & Reeves 2001), DNA profiling, use of DNA arrays, etc. However, it was also stressed that phenotypic properties, including chemotaxonomic markers, will remain important diagnostic properties in any species description.

The occurrence of horizontal gene transfer between bacteria potentially has an impact on the bacterial species concept; if genes indeed move freely from cell to cell, would it still be possible to delineate species at all? (For an in-depth discussion see Doolittle (1999).) Bacteria can ‘capture’ new gene loci from other organisms, sometimes extremely distantly related. This may occur as a side effect of homologous recombination, whereby a heterologous gene from a donor is integrated together with closely flanking homologous DNA. Alternatively, heterologous genes may be integrated together with a transposable element introduced in the recipient cell on a plasmid or a phage (Cohan 2001). Analysis of the increasing numbers of complete bacterial genomes that have been sequenced has shown many presumptive cases of horizontal gene transfer, and some of these lateral gene transfer events may have occurred 1–2 billion years ago. However, the overall picture that emerges from the genomic approach is that the prokaryotic species concept need not be troubled by the problem of horizontal gene exchange as the species concept is based on whole genome similarities, and not on the properties of single genes.

According to the rules of the International Code of Nomenclature of Bacteria (Bacteriological Code) (see § 3), one representative strain should be designated as the type strain when describing a (new) bacterial or archaeal species. It is always desirable to study many strains on which the species description is to be based. Poor descriptions based on a small set of strains can lead to improper phenotypic circumscription of taxa, thus hindering the

identification of new isolates. Sneath (1977) suggested that a minimum of 10, or preferably 25, isolates should be studied. However, a survey of the new species descriptions that appeared in the *International Journal of Systematic and Evolutionary Microbiology* in 2002 shows that more than half of the new species described were based on a single isolate. A proposal by Christensen *et al.* (2001) to make the study of more than one strain obligatory for the description of new species was not implemented by the International Committee for Systematics of Prokaryotes on the grounds that it is not always feasible to obtain more isolates, more strains can always be added later in an amended description, and no description at all would be a loss to science.

The questions of whether clusters of strains observed in the bacterial world share the dynamic properties attributed to the eukaryote species, and whether the mechanisms driving the origin of new species in bacteria might be shared with the eukaryotes, were addressed by Cohan (2001). He concluded that, despite basic differences in the nature of the genetic exchange between bacteria and eukaryotes, bacterial species share many of the fundamental properties of eukaryotic species. In the case of the prokaryotes, recombination depends on vectors such as bacteriophage-mediated transduction or plasmid-mediated conjugation, and is therefore limited by the host ranges of the respective vectors. Moreover, restriction endonuclease activity can greatly reduce the rate of recombination. Cohan further assessed the potential of the formation of new species within the prokaryote world as a result of the formation of new ‘ecotypes’. On the basis of multilocus sequence typing he defined an ecotype as a set of strains using the same or very similar ecological niches, such that an adaptive mutant from within the ecotype outcompetes all other strains of the same ecotype. The potential for speciation is very large indeed for the following reasons.

- (i) Speciation in prokaryotes requires only ecological divergence (versus both reproductive and ecological divergence in sexual eukaryotes).
- (ii) The extremely large population size of prokaryotes makes rare mutation and recombination events much more accessible to a population than is the case for macro-organisms.
- (iii) A prokaryote species is open to gene transfers from many taxa, even those distantly related, so it can take up existing adaptations from anywhere in the prokaryote world.
- (iv) Recombination events have a localized nature, as only a small fraction of the donor’s genome is integrated.

This allows for transfer of a useful adaptation without co-transfer of other donor segments that would be deleterious for the recipient. This is in contrast with the processes of meiosis and fertilization in eukaryotes, which yield hybrids that are a 1 : 1 mix of both genomes. The physiology of bacteria may thus be more modular than is the case for macro-organisms in that a new adaptation might be accommodated with very little fitness cost (Cohan 2001).

There are no clear-cut recommendations for the delineation of prokaryote genera or higher taxonomic levels.

The Bacteriological Code (Lapage *et al.* 1992) recognizes genera, families and orders, but does not concern itself with higher levels (classes, kingdoms, etc.). The ranks of genera, families and orders are determined by subjective decision-making by bacteriologists. The genus may be operationally defined as 'a collection of species with many characters in common', but what constitutes a genus is purely a matter of personal judgement, and clear guidelines are lacking. The availability of molecular data has not changed the definition of a genus.

3. HOW MANY PROKARYOTE SPECIES HAVE BEEN DESCRIBED?

As is apparent from the above discussion on the species concept for the prokaryotes, the description of one or more isolates as members of a new species requires extensive documentation of their morphological and physiological properties, as well as determination of 16S rDNA sequences, DNA base composition, DNA–DNA similarity with the species' closest relatives and information about many other characteristics. A living culture, established as the type strain, has to be deposited in at least two public culture collections located in different countries.

Prokaryote nomenclature is regulated by the International Committee on Systematics of Prokaryotes (ICSP; <http://www.the-icsp.org>) (formerly: the International Committee on Systematics of Bacteria). The rules of nomenclature appear in the International Code of Nomenclature of Bacteria (Bacteriological Code) (soon to be renamed the International Code of Nomenclature of Prokaryotes), which covers both the Archaea and the Bacteria (Lapage *et al.* (1992); a new edition is currently in preparation). The ICSP has established taxonomic subcommittees on different groups of bacteria. Its Judicial Commission considers amendments to the Bacteriological Code and any exceptions that may be needed to specific rules.

A list of approved names of prokaryotes, encompassing ca. 2500 species, was published in 1980 (Skerman *et al.* 1980). With the publication of that list of names approved for retention in the new bacteriological nomenclature, all other names that had appeared in the literature before that date had lost their validity. Thus, a new start was made in bacteriological nomenclature by discarding the tens of thousands of names that had appeared in the literature of the past. New names with validity in prokaryote nomenclature can only be added by publication in the *International Journal of Systematic and Evolutionary Microbiology* (before 2000: the *International Journal of Systematic Bacteriology*), either in the form of an original article or in the 'Validation Lists' of species effectively published in other journals.

As of 9 July 2003, the number of validly published names of prokaryotes, Archaea and Bacteria combined, was 6205, belonging to 1174 genera and 157 families (for updates see <http://www.bacterio.cict.fr>) (Euzéby 1997). This number is extremely small when compared with the numbers of plants, animals, fungi, etc. described in the literature. For comparison, almost a million insect species have been recognized.

Extensive information on the prokaryote species described can be found in two handbooks: *Bergey's manual*

and *The prokaryotes*. *The prokaryotes*, first published in 1981, is now in its third edition (Dworkin *et al.* 1999), which is exclusively published online (<http://www.prokaryotes.com>). This manual provides in-depth information on all aspects of the life of the Bacteria and the Archaea, including their taxonomy. Eight editions of *Bergey's manual of determinative bacteriology* have been published between 1923 and 1974, each edition being updated with the newly discovered bacterial species and revised according to the changing concepts on bacterial taxonomy. Under the new name *Bergey's manual of systematic bacteriology*, it was published in four volumes between 1984 and 1990. The second edition of this manual (<http://www.cme.msu.edu/bergeys>) will appear in five volumes, the first of which was released in 2001 (Garrity 2001). *Bergey's manual* provides an operative division of the prokaryote world into two domains (Archaea and Bacteria), phyla (two for the Archaea, 23 for the Bacteria), classes, orders, families, genera, species and subspecies (Brenner *et al.* 2001; Garrity 2001). This list does not represent an official classification of the prokaryotes into higher taxa; such an official classification does not exist, as the Bacteriological Code does not deal with taxonomic ranks above the class. However, the *Bergey* system is widely adopted by bacteriologists. The classification proposed in *Bergey's manual* is largely based on small-subunit rRNA sequences (Ludwig & Schleifer 1994; Ludwig & Klenk 2001). Some of the phyla in the *Bergey* classification are phenotypically homogeneous (e.g. the Aquificae, the Thermotogae, the Chloroflexi, the Chlorobi, the Cyanobacteria, and the Spirochaetes), whereas other phyla are very heterogeneous. It may be noted that those phyla that are phenotypically homogeneous contain relatively few species. It is well possible that further sampling in the future may prove these phyla to be more diverse than currently recognized. The greatest diversity is found within the phylum Proteobacteria, which contains aerobic, anaerobic, photoautotrophic, photoheterotrophic and chemolithotrophic representatives. Considerable phenotypic diversity is also found in the phyla Firmicutes and Bacteroidetes.

The cyanobacteria form a special case in taxonomy and nomenclature, as they are also being claimed by botanists as being part of the plant world, and as such fall under the rules of the International Code of Botanical Nomenclature (Botanical Code). In botanical classification systems they are named Cyanophyta (blue-green algae). Whereas the rules of the Bacteriological Code require living type cultures, species are described under the Botanical Code according to preserved herbarium specimens, photographs, etc. The nomenclature of the cyanobacteria is extremely confusing, and many species appear under two or more different names in the literature. The 2001 edition of *Bergey's manual* refrains from dividing this phylum into classes, orders and families, but divides the group into five subsections, each consisting of several 'form-genera', awaiting a definitive solution for the nomenclature of the cyanobacteria which has to be acceptable to both bacterial and botanical taxonomists (see also Knapp *et al.* 2004).

In those cases in which a prokaryote can be described in sufficient detail to warrant establishment of a novel taxon, including information about its morphology, 16S rDNA sequence and ecological niche, but cannot (yet) be

cultured in pure culture, the rank of 'Candidatus' has been established, a waiting position for putative taxa (Murray & Stackebrandt 1995). As of 9 July 2003, the above-mentioned bacterial nomenclature Web site <http://www.bacteria.cict.fr> lists 46 such 'Candidati'.

Recently, there have been extensive discussions in the 'Correspondence' pages of *Nature* about the desirability of online taxonomic databases towards the description of all living organisms on Earth (Bisby *et al.* 2002; Godfray 2002; Lee 2002). For the prokaryotes (at least for those members of the prokaryotes that have already been isolated and named), this goal has essentially been achieved (Oren & Stackebrandt 2002). The <http://www.bacterio.cict.fr> Web site contains the approved lists of bacterial names and provides information on all bacterial names with validity in the nomenclature, including their nomenclatural history, basonyms, etc. The problem of synonyms hardly exists any longer for the prokaryotes; the ICSP with its Judicial Commission and the different taxonomic subcommittees have virtually solved all these problems. There are, however, a few cases in which a valid prokaryote name is also included in the lists of names of plants or animals. Thus, the name *Bacillus* both describes an aerobic endospore-forming bacterium and the stick insect (order Phasmida); *Proteus* is both a genus of Proteobacteria and a cave-dwelling amphibian. In such cases coordination between prokaryote and eukaryote taxonomists will be required when a single universal taxonomic database is to be prepared.

4. HOW MANY PROKARYOTE SPECIES ARE THERE ON EARTH?

The small number of bacterial and archaeal species already described greatly underestimates the true number of species in nature. Our isolation and cultivation methods are still, to a large extent, based on the procedures introduced by Robert Koch and his co-workers in the 1880s. These methods have enabled the isolation and characterization of *Bacillus anthracis* or *Mycobacterium tuberculosis* and many other species, but may not be suitable for many other prokaryotes. Furthermore, many prokaryotes live in close symbiotic associations, forming consortia whose members depend on each other for existence, and can therefore not be isolated in pure culture. Still others are obligate symbionts or parasites of animals and cannot (yet) be cultured in the absence of their host.

How little we know of the true diversity of prokaryotes in nature became clear when molecular techniques based on small-subunit rRNA sequencing were first applied to whole ecosystems. In this approach, 16S rRNA genes are amplified from DNA isolated from biomass collected from the environment, or alternatively 16S rDNA is prepared from 16S rRNA purified from the environmental sample using reverse transcriptase (Amann *et al.* 1995; Embley & Stackebrandt 1997). These approaches to obtain information about prokaryote biodiversity have become extremely popular. A survey of the 2002 issues of the journals *Applied and Environmental Microbiology* (section Microbial Ecology), *FEMS Microbiology Ecology* and *Microbial Ecology* shows that, respectively, 37%, 24% and 22% of the articles published used some variant of this methodology. The first papers using this approach were

published in 1990, describing studies performed on the bacterioplankton of the Sargasso Sea (Giovannoni *et al.* 1990a,b; Britschgi & Giovannoni 1991) and a cyanobacterial mat in a thermal spring (Ward *et al.* 1990). It became immediately clear that the sequences obtained directly from the environment nearly always differ from those present in the database of 16S rDNA sequences of named prokaryote species; only seldom does it occur that a sequence recovered from the environment fully matches that of a known species. The environmental sequences differ from the known sequences to the extent that they may warrant establishment of new species, genera, families, orders, and even phyla, if only we were able to obtain the organisms harbouring these sequences in pure culture so that they can be studied and described. The large number of new and diverse sequences thus obtained led to the conclusion that the number of prokaryote species already described and named is, at most, 1–2% of the true number of bacterial species extant, as based on the current species concept. Some estimates are even lower; extensive surveys of 16S rDNA variation from environmental samples suggest that culturable bacteria amount to less than 0.1% of the species in a given environment (Hugenholtz *et al.* 1998; Ward *et al.* 1998; Dojka *et al.* 2000; DeLong & Pace 2001). There are even many higher taxa that can be detected by using the 16S rDNA approach that have no cultured representatives: the paper by Liesack & Stackebrandt (1992) on bacterial diversity in an Australian soil is probably the first study documenting this. A phylogenetic tree of the domain Bacteria in which the environmental sequences were included showed more than 40 deep-branching lineages that may correspond to phyla in phylogenetic depth; 13 of these lineages are not yet represented by cultured members (Pace 1997; Hugenholtz *et al.* 1998). The isolation and characterization of such novel micro-organisms, whose existence could not even be predicted a few years ago, is now a major challenge of microbiological science. The 16S rDNA sequence itself provides little or no information on the possible function of the organism in the ecosystem, even when organisms with very similar sequences are extant in culture.

The information gained by sequencing 16S rRNA or 16S rRNA genes recovered from the environment can be further used to obtain information about the distribution of the organisms harbouring these sequences. One popular approach is FISH, in which 16S rRNA-targeted fluorescently labelled probes are used to specifically detect cells with the corresponding rRNA structure. This approach, pioneered by DeLong *et al.* (1989) and often used in combination with other staining methods (Hicks *et al.* 1992), has proved extremely valuable to obtain information about the spatial distribution of specific types of uncultured micro-organisms (Antón *et al.* 2000; Dojka *et al.* 2000; Ramsing *et al.* 2000), and can also be used in combination with flow cytometry to enumerate cells that carry a specific 16S rRNA phylotype.

One of the most important results obtained using the 16S rDNA sequencing approach is the recognition of some as yet uncultured bacteria as being numerically dominant in major ecosystems such as the world ocean. A sequence of an α -Proteobacteria phylotype designated SAR11, first identified in the oligotrophic Sargasso Sea (Giovannoni *et al.* 1990a) now appears to belong to the

most abundant micro-organism on Earth, an organism that accounts for a third or more of the cells present in sea surface waters and nearly a fifth of the cells present in the mesopelagic zone of the ocean. Members of the SAR11 clade may represent as much as 50% of the total surface microbial community in some regions (Morris *et al.* 2002). The number of SAR11 cells in the world ocean can be estimated at 2.4×10^{28} . SAR11 cells account for 18% of the total bacterial biomass in ocean surface waters (upper 200 m), and overall contribute *ca.* 12% of the total marine prokaryote biomass. FISH studies showed that the organism that harbours the SAR11 phylotype is a small ($0.4 \mu\text{m} \times 0.2 \mu\text{m}$) curved rod. The biogeochemical role of the SAR11 clade remains uncertain, but this microbial group is obviously among the most successful organisms on Earth. SAR11 has recently been brought into culture, using a novel approach (Rappé *et al.* 2002; see also § 9). As yet, cell densities obtained in culture are extremely low, but it may be expected that more will soon be learned about the properties of this numerically important but elusive organism.

Analysis of small-subunit rRNA genes has also greatly increased our insight into the diversity within the domain Archaea. Thus, many new archaeal 16S rDNA sequences were detected in 'Jim's Black Pool', a hot spring in Yellowstone National Park (Barns *et al.* 1994). Sequencing of 16S rRNA genes cloned from DNA isolated from biomass collected from 'Obsidian Pool', another Yellowstone hot spring, yielded many novel lineages of Crenarchaeota. Of special interest are those sequences representing a major kingdom(phylum)-level branch not yet represented by any cultured representative, provisionally named the 'Korarchaeota'. This allegedly new phylum of Archaea is located closer to the presumed root of the phylogenetic tree than any known organism (Barns *et al.* 1996). Yet another new kingdom(phylum)-level branch within the Archaea (the 'Nanoarchaeota') was recently discovered when the 16S rDNA of a $0.4 \mu\text{m}$ small spherical archaeon that grows attached to the surface of *Ignicoccus*, a thermophilic member of the Crenarchaeota from a submarine hot vent, was sequenced (Huber *et al.* 2002). The organism was provisionally named 'Nanoarchaeum equitans'. Small-subunit rRNA sequences very similar to those of 'Nanoarchaeum' have been found in a variety of high temperature biotopes, suggesting that members of this new phylum may be widely distributed (Hohn *et al.* 2002). These findings suggest that the microbial diversity is much greater than previously estimated, even under the most extreme conditions.

Even more exciting is the finding that archaeal 16S rDNA sequences abound in DNA isolated from 'non-extreme' environments, including ocean water, lake water and forest soil. The representatives of the Archaea known in the 1980s were a fairly bizarre collection of microbes, most of which inhabit extreme environments such as hypersaline lakes and boiling neutral and acid springs. The finding that Archaea abound in the aerobic cold marine environment in the early 1990s (DeLong 1992; Fuhrman *et al.* 1992) came, therefore, as a tremendous surprise. Sequences of Euryarchaeota and of Crenarchaeota (a group represented thus far by cultured extreme thermophiles only) are consistently being recovered. Similar sequences, not closely related to any cultured species of

Archaea, have been obtained from marine environments worldwide. The presence of Archaea in the marine system has since been confirmed on the basis of specific lipid biomarkers (e.g. Wuchter *et al.* 2003). Recently, some information has been obtained about the physiology of the marine Crenarchaeota and Euryarchaeota (see § 7), but we remain far removed from an understanding of their role in the marine ecosystem. One particular type of marine crenarchaeote (designated 'Cenarchaeum symbiosum') is a symbiont associated with a marine sponge (Preston *et al.* 1996; Stein & Simon 1996). Quantitative estimates based on FISH and related techniques indicate that Archaea represent *ca.* 30% of all prokaryote cells in the oceans, and in some areas they comprise almost half of the prokaryotes detected. A study performed at the Hawaii Ocean Times-series station showed the crenarchaeotal community to increase with depth, with an especially sharp increase between 100 and 150 m, below which they represented up to 39% of the total picoplankton numbers (Karner *et al.* 2001). Despite their numerical abundance, none of the marine planktonic Archaea has yet been cultured. A wealth of novel sequences of Crenarchaeota and Euryarchaeota have also been recovered from other ecosystems, including groundwater from an oxic basalt aquifer in Idaho, USA (O'Connell *et al.* 2003) and from deep gold mines in South Africa (Takai *et al.* 2001), to give just two examples.

The small-subunit rRNA sequencing approach has also shown us that the marine environment contains many 'picoeukaryotes' of *ca.* 1–2 μm size, of types that were not previously recognized. Their initial discovery was a by-product of a study of bacterial 16S rRNAs, which yielded many unusual plastid-derived 16S sequences (Rappé *et al.* 1998). These new types of organism, belonging to novel lineages, presumably include heterotrophic, mixotrophic, as well as phototrophic types, and they are probably widespread (Moreira & López-García 2002).

It should be taken into account that the approach has many pitfalls and problems, and its limitations should be clearly recognized. Each step in the procedure is a source of bias, which will lead to a distorted view of the 'real world'. The bias starts with the method used to collect the sample. The protocol used to lyse the cells and extract their DNA may introduce many artefacts; too drastic procedures will fracture the DNA and disrupt relevant genes, whereas more gentle methods may be insufficient to break the more resistant cells within the community, and the DNA used for amplification or cloning will thus represent part of the community only. PCR-based protocols further introduce problems, not only because of the chance of formation of chimaeric molecules, but also because of selective priming leading to differential amplification of different sequences, and the possible presence of inhibitory substances that inhibit the PCR reaction. An excellent overview of the problems that can be encountered when using the technology was given by von Wintzingerode *et al.* (1997). When the rRNA from the community is taken as the starting material it is important to realize that the number of ribosomes per cell is highly variable (commonly between 10^3 and 10^5 ribosomes per cell), both among different species and within cells of any species, dependent of the growth rate. When the community DNA is used as a starting point, it should be taken into account

that many bacteria have multiple copies of the rRNA operon—in some cases up to 10 or more, and such organisms may therefore be detected at an increased frequency. A drawback of all 16S rRNA-based approaches is that, even when several hundred clones are examined, populations of organisms that are present in low numbers may escape detection.

In many micro-organisms that have multiple copies of the rRNA operon there may be considerable sequence heterogeneity. Thus, *Paenibacillus polymyxa* has at least 12 copies of the 16S rRNA gene with extensive sequence heterogeneity of 10 variant nucleotide positions within a 347 base-pair fragment investigated (Nübel *et al.* 1996). Fifteen different sequences were found in 16S rRNA genes of *Clostridium paradoxum*. Most of the cloned genes contained intervening sequences, located in the variable region I of the 16S rDNA and varying in length from 129 to 131 nucleotides. These intervening sequences were absent from the mature rRNA (Rainey *et al.* 1996). Significant heterogeneity was also found in the two genes encoding 16S rRNA in the archaeon *Haloarcula marismortui*, which differ by 74 nucleotide substitutions, thus exhibiting 5% overall sequence divergence (Mylvaganam & Dennis 1992). FISH staining showed that each cell contains both types of ribosome (Amann *et al.* 2000). The two 16S rRNA genes of *Thermobispora bispora* differ in 98 nucleotide positions (6.4% sequence divergence), in addition to the presence of six regions of deletion–insertion (Wang *et al.* 1997). Because of these observations, the finding of different 16S rDNA sequences does not always imply the presence of different bacterial species.

An entirely different method used to assess the biodiversity in microbial ecosystems is based on the measurement of the renaturation rate of thermally denatured DNA extracted from the community (Torsvik *et al.* 1990a,b, 1996, 1998). The rationale of the approach is that the average time it takes for a single DNA strand to find its homologue depends on its frequency within the background of all other genomes present. By measuring the rate of annealing of heterogeneous DNA extracted from forest soil, and based on the generally accepted species definition (DNA–DNA similarity of more than 70%), the reassociation studies with DNA extracted from forest soil indicated a complexity comparable to at least 4000 genomes (Torsvik *et al.* 1990a). Moreover, it was calculated that none of the species present dominated the community, and that the most abundant species present represented less than 1% of the individuals in this soil (Torsvik *et al.* 1998). Extrapolating from such results, Dykhuizen (1998) argued that worldwide there must be at least 10^9 bacterial species. He estimated the number of bacterial species in 30 g of forest soil at over half a million. Assuming that there are at least 2000 different types of community that differ from each other, and are as complex as Norwegian forest soil, the number of one billion is easily exceeded. With the high numbers of prokaryotes present, their high growth rates, and their excellent survival under adverse conditions, speciation in bacteria is simple and extinction is difficult, leading to an ever-increasing number of species over time (Dykhuizen 1998).

If, in addition, most eukaryotes bear their own species-specific bacterial pathogens, the number of bacterial

species would be very large indeed. For example, *Streptococcus pyogenes* and some related streptococci are known to infect humans only. Almost a million insect species have been recognized, and each individual insect harbours millions to billions of bacteria. At least 10% of insect, tick and mite species harbour obligate, and probably species-specific symbionts that have coevolved with their hosts (Dasch *et al.* 1984).

Based on the different experimental approaches and theoretical considerations presented above, it is now evident that the 6205 prokaryote species described thus far form only a very small fraction of the total number of prokaryote species in nature.

5. ARE BACTERIA DISTRIBUTED UBIQUITOUSLY, AND DO ENDEMIC BACTERIAL SPECIES EXIST?

Because of the small size of prokaryotes and the ease with which they are distributed from place to place, the famous statement by Baas-Becking (1934) that ‘Everything is everywhere, the environment selects’ would imply that Bacteria and Archaea may be cosmopolitan in their distribution, and that endemism can hardly be expected to occur in the prokaryote world—with the obvious exception of species-specific symbionts and pathogens of animals or plants endemic to certain geographical regions (see also Finlay (2004) for a discussion of protists).

Little is known about the biogeography of bacteria, and the question of whether there are endemic bacteria has seldom been asked. Endemism would imply that a given organism has resided in an area long enough to have formed a cluster of phylogenetically related groups or clades. Use of molecular techniques now enables the examination of the distribution of bacterial species. However, it may not be possible to unequivocally prove endemism in bacteria: not finding an organism does not necessarily mean that it is not there!

In some well-documented cases it indeed appears that prokaryote species may have a worldwide distribution. As mentioned above, SAR11, the small marine bacterium now known to be the most abundant micro-organism on Earth, is one example. Identical 16S rDNA sequences representing this type of organism can be recovered from marine environments all over the globe. The same is true for a few other abundant phylotypes. Another well-researched case is the mat-forming cyanobacterium *Microcoleus chthonoplastes*. Isolates from the coastal areas of Germany, Spain, Egypt, the USA and Mexico are virtually identical in 16S rDNA sequence as well in phenotypic properties (Garcia-Pichel *et al.* 1996). A strain of the thermophilic archaeon *Archaeoglobus fulgidus* isolated from a North Atlantic oil field showed 100% DNA–DNA reassociation with the type strain of the species that was isolated from Italy (Beeder *et al.* 1994).

In other cases, however, indications for endemism in the prokaryote world have been obtained. A comparison of the Proteobacteria isolated from sea ice in the Arctic and the Antarctic regions showed consistent distinct differences. Geographically, these communities are isolated, communication is not feasible as the sea surface waters of the oceans are too warm for psychrophiles to survive, and transit with the cold deep current is far too slow. Comparison of two *Octadecabacter* isolates (α -Proteobacteria), one

Table 2. Culturability determined as a percentage of culturable bacteria (measured as colony-forming units) in comparison with total cell counts.
(From Amann *et al.* (1995) and reproduced with permission.)

habitat	culturability (%)
seawater	0.001–0.1
freshwater	0.25
mesotrophic lake	0.1–1
unpolluted estuarine waters	0.1–3
activated sludge	1–15
sediments	0.25
soil	0.3

obtained from the Arctic and one from the Antarctic, showed the strains to be closely related, but sufficiently different on the level of DNA–DNA reassociation to warrant classification as separate species. A similar phenomenon was found for *Polaribacter* isolates from both polar areas. At present, no species have been identified that occur both in the Arctic and the Antarctic, indicating the possibility of endemism (Staley 1999).

Another case that may rule against cosmopolitan distribution of bacteria was documented by Fulthorpe *et al.* (1998). Twenty-four soil samples from six different regions on five continents were enriched with 3-chlorobenzoate. Genotypes of the 3-chlorobenzoate mineralizers were determined by repetitive extragenic palindromic PCR genomic fingerprinting and by restriction digests of the 16S rRNA genes (ARDRA). The collection of 150 isolates contained 48 genotypes and 44 ARDRA types, which formed seven distinct clusters. Most (91%) of the genotypes were unique to the sites from which they were isolated, and each genotype was found only in one region. All but one of the ARDRA types were found in one region only. These results suggest that the ability to mineralize 3-chlorobenzoate is distributed among very different genotypes and that these genotypes are not globally dispersed. Soil bacteria may thus have a limited distribution. If this is a general phenomenon, the numbers of prokaryote species may indeed be as large as those calculated by Dykhuizen (1998).

6. WHY ARE WE NOT ABLE TO GROW ALL LIVING PROKARYOTES?

Comparison of the number of bacteria detected in different ecosystems and the number of colony-forming bacteria recovered from such ecosystems consistently shows that we are able to grow only a few per cent at best of the prokaryotes present, and generally even much less (table 2). The figures quoted in table 2 may be somewhat low, and higher colony yields can be achieved under optimal conditions. However, even then colony yields above 5–10% are seldom obtained. This phenomenon, referred to as ‘the great plate count anomaly’ (Staley & Konopka 1985) may be due to a combination of causes. The cultivation methods used may be unsuitable for the many new species present whose properties and nutritional demands are unknown. Indeed, as documented above, most bacteria in nature belong to species that have not yet been cultivated. Nature can cultivate all extant micro-organisms,

but microbiologists still have much to learn about the proper methods to bring even the numerically dominant Bacteria and Archaea into culture. It is becoming more and more clear that the types of rich medium developed in the nineteenth century by medical microbiologists for the isolation of human pathogens are highly unsuitable for the recovery of chemo-organotrophic micro-organisms from the oligotrophic environment. Low nutrient media such as, for example, the R2A agar medium marketed by Difco, generally give higher counting efficiencies. For some micro-organisms even such poor media may have far too high organic carbon concentrations (see § 9). Another possible cause for the ‘great plate count anomaly’ may be that many micro-organisms live in consortia. Being inter-dependent on each other, they cannot be cultured alone.

An entirely different reason for low colony recovery may be that bacteria that otherwise easily form colonies on nutrient media are present in a dormant state (VBNC). It is now well documented that many bacteria may enter a non-culturable state upon exposure to unfavourable conditions (too high or too low salinity, oxidative stress, low temperature, etc.). The phenomenon has been described for organisms such as *Vibrio cholerae*, *V. vulnificus*, *Salmonella enteritidis*, and others (Roszak *et al.* 1984; Colwell *et al.* 1985; Roszak & Colwell 1987; Oliver *et al.* 1991). McDougald & Kjelleberg (1999) listed 35 species belonging to 17 genera of Proteobacteria in which the phenomenon has been documented. Little is known about the properties of these VBNC cells, except for the fact that they still exhibit signs of metabolic activity and thus viability, and even infectivity in the case of pathogenic species, but will not form colonies when plated on conventional agar media. There are few convincing reports of their resuscitation; although it has been documented that heat shock treatment may induce resuscitation of dormant *V. cholerae* cells. Nevertheless, it is apparent that the formation of VBNC cells may be an important adaptive survival strategy response to environmental stress. Entry into the VBNC state may be a programmed response. In the first state, in which culturability is lost, cellular integrity and intact nucleic acids are maintained. Later, cellular integrity is gradually lost and nucleic acids are degraded, eventually leading to loss of viability. Others have described the VBNC state as a moribund condition in which cells become progressively debilitated until cell death finally occurs. McDougald & Kjelleberg (1999) and McDougald *et al.* (1998) have reviewed the current views on the phenomenon, which may be one of the causes of the ‘great plate count anomaly’.

7. NOVEL TYPES OF METABOLISM IN PROKARYOTES

It is often assumed that since the days of Sergei Winogradsky and Martinus Beijerinck we understand the roles that the prokaryotes play in the biogeochemical cycles in nature. Even when it is now recognized that most species that perform these transformations are still waiting to be isolated, our insight into the nature of the transformations themselves is generally believed to be quite complete.

We are also deceived here, as in recent years several novel modes of metabolism have been discovered, some of them of major biogeochemical importance, which had

thus far remained undetected. In some cases the discoveries were made by chance (e.g. the finding of light-harvesting retinal proteins in marine Proteobacteria; Béjà *et al.* 2000); in other cases they were the result of an intensive search (anaerobic methane oxidation; Boetius *et al.* 2000); in yet other cases the process was discovered in nature long after its possible occurrence was predicted on thermodynamical grounds (anaerobic ammonium oxidation).

The discovery of anaerobic autotrophic ammonium oxidation in an anaerobic bioreactor around 1995 (for a review see Jetten *et al.* (1998)) was preceded by the theoretical prediction of the possible occurrence of the process by Engelbert Broda in 1977. He calculated that the following two processes are thermodynamically favourable, and theoretically should be able to support autotrophic growth:

No organisms performing either process were known at the time, and the known ammonium oxidizing bacteria were all aerobes, requiring molecular oxygen, not only as the terminal electron acceptor in respiration, but also as substrate in the first step of ammonium oxidation, catalysed by ammonia monooxygenase. The newly discovered 'anammox' bacteria use nitrite as an electron acceptor according to the first of the two above reactions. The bacteria performing the process are phylogenetic neighbours of the *Planctomyces* group. The biochemistry of anaerobic ammonium oxidation is completely different from aerobic nitrification or anaerobic denitrification, and involves hydrazine and hydroxylamine as intermediates. The 'anammox' process is not just a laboratory curiosity; in certain marine systems the pathway may be responsible for a substantial loss of bound nitrogen as N_2 (Dalsgaard *et al.* 2003; Kuypers *et al.* 2003). The discovery of the 'anammox' process shows that (nearly) all reactions that are thermodynamically feasible and can yield sufficient energy for growth are exploited by nature (the second process predicted by Broda on theoretical grounds—anoxigenic photosynthesis with ammonium as electron donor—is still awaiting discovery).

Another process of global biogeochemical importance that was only recently characterized at the level of the responsible organisms is anaerobic methane oxidation. Indications for the occurrence of the process have been accumulating for several decades, and it has been predicted that sulphate is the electron acceptor:

The process is now known to be performed by a consortium of Archaea that oxidize methane by some kind of 'reversed methanogenesis' in a reaction that is thermodynamically unfavourable under standard conditions, and sulphate-reducing Bacteria related to *Desulfosarcina* or *Desulfococcus* (Boetius *et al.* 2000). At least two distinct groups of Archaea may participate in such methane-oxidizing consortia (Orphan *et al.* 2002). The nature of the

intermediate transferred between the partners of the consortium is still unknown, but on thermodynamical and kinetic grounds formate was speculated to be the best candidate (Sørensen *et al.* 2001).

Until recently, the presence of a light-driven retinal-based proton pump (bacteriorhodopsin) was believed to be a unique feature of certain extremely halophilic Archaea (*Halobacterium* and relatives). However, analysis of DNA extracted from marine bacterioplankton cloned in bacterial artificial chromosomes has shown genes for bacteriorhodopsin-like proteins (now termed proteorhodopsins) in abundant but as yet uncultured marine members of the γ -Proteobacteria (the SAR86 cluster) (Béjà *et al.* 2000). At the time this finding was completely unexpected. However, it was subsequently found that proteorhodopsins are indeed found in the bacterioplankton in the sea and that these pigments function to make light energy available to the marine bacterial community (Béjà *et al.* 2001).

Using the same approach it was discovered that the biota of the oxic ocean harbour a great diversity of anoxygenic phototrophs whose light-harvesting system is based on bacteriochlorophyll *a* (Béjà *et al.* 2002b). This finding is in agreement with spectroscopic measurements in which variable fluorescence transients monitored with an infrared fast repetition rate fluorometer provided evidence for bacterial photosynthetic electron transport. On the basis of the vertical distribution of bacteriochlorophyll *a*, and the number of infrared fluorescent cells, it was concluded that photosynthetically competent anoxygenic phototrophic bacteria may comprise at least 11% of the total microbial community in the upper ocean and account for up to 5% of the photosynthetic electron transport in the surface layers of the ocean (Kolber *et al.* 2000, 2001). Although the cultured marine aerobic anoxygenic phototrophs (*Erythrobacter*, *Roseobacter*) belong to the α -Proteobacteria, screening of bacterial artificial chromosomes containing environmental DNA led to the identification of multiple, distantly related photosynthetically active bacterial groups. Surprisingly, no sequences of *Erythrobacter*, one of the more commonly cultured marine bacteriochlorophyll-*a*-containing bacteria, were found. A few sequences recovered were related to *Roseobacter*, but others belonged to as yet unknown photosynthetic organisms possibly affiliated with the β - and the γ -Proteobacteria (Béjà *et al.* 2002b). It may be mentioned here that in the marine environment small prokaryotic oxygenic phototrophs of the cyanobacterial genera *Prochlorococcus* and *Synechococcus*, organisms that have only recently been recognized, are extremely abundant and are responsible for a significant part of the primary production in the sea (Waterbury *et al.* 1979; Partensky *et al.* 1999; Scanlan & West 2002).

Such genomic analyses of large DNA fragments isolated from the environment may show us the physiological potential of uncultured organisms. Thus, analysis of a 38.5 kbp recombinant fosmid clone derived from an uncultured marine crenarchaeote (identified as such on the basis of the 16S rDNA gene present in the clone) yielded the first information on additional genes present in the organism (Stein *et al.* 1996). Béjà *et al.* (2002a) extended these studies by sequencing and annotating five fosmids containing genome fragments derived from group

I Crenarchaeota from Antarctic surface waters and deeper waters of the temperate North Pacific. Some information was recently obtained on the possible mode of metabolism of the ubiquitous marine Crenarchaeota: they most probably lead a chemoautotrophic life. When seawater was incubated in the dark in the presence of ^{13}C -labelled bicarbonate, specific Crenarchaeota-associated biphytanyl membrane lipids became selectively labelled. Although the nature of the electron donor used by these presumptive chemoautotrophs is not yet known, it is already clear that these as yet uncultured but abundant organisms represent a significant global sink for inorganic carbon and may form an important component of the global carbon cycle (Wuchter *et al.* 2003).

Microscopic techniques combining FISH and microautoradiography have recently been developed to obtain information on the specific role that different members of microbial communities may play (Lee *et al.* 1999). Using this approach it was suggested that the as yet uncultured marine Euryarchaeota might take up amino acids (Ouverney & Fuhrman 2000).

These and additional studies, disclosing some of the physiological attributes of organisms that resist cultivation, may open the way towards the rational design of enrichment and isolation methods that may eventually enable us to study such organisms in culture and characterize their properties.

8. ARE THERE ENDANGERED PROKARYOTES?

Prokaryotes have never yet been the subject of concern of environmental activists who lobby for the protection of endangered animal and plant species. In fact, little is known as to what extent there are 'endangered' prokaryotes on Earth, and no specific bacteria are considered to be naturally endangered (Staley 1997). Perhaps there are a few such organisms that live in restricted geographical areas whose habitats may be endangered. Possible examples are, for example, hot springs and other thermal areas. However, there may be many more endangered prokaryote species than is generally realized. If indeed every animal carries specific symbionts or pathogens, then with the extinction of the animal these symbionts are also doomed to extinction. This is especially true in the case of insects, which generally carry symbiotic bacteria. A few prokaryotic micro-organisms are intentionally endangered because of public health considerations. Thus, *Mycobacterium leprae*, the causative agent of leprosy, has been placed on the World Health Organization's list for species to be eradicated.

Species extinction should be considered to be a natural process, whereby one species is replaced by another through natural selection pressures. Staley (1997) documented the case of *Simonsiella muelleri*, a harmless inhabitant of the oral cavities of humans, cats, dogs, sheep and other mammals. *Simonsiella muelleri* is nowadays rarely encountered in the mouth of people from western countries. Apparently this organism has been largely eliminated from humans by modern dental hygiene practices and/or diet, and only humans from remote areas, such as native Americans from Alaska, appear to carry it. *Simonsiella muelleri* may thus be threatened with extinction. The

question of whether its disappearance should be considered a loss remains open.

9. DEVELOPMENT OF NOVEL ISOLATION TECHNIQUES FOR PROKARYOTES

It is now obvious that the well-established isolation and cultivation methods are inadequate to grow many or even most of the prokaryotes found in nature. New approaches are therefore needed towards their isolation. Several novel methods have emerged in recent years and these have already enabled the isolation of some interesting organisms.

One exciting development is the use of 'laser optical tweezers', a micromanipulation method in which a single bacterial cell can be manipulated in a capillary tube. This optical trapping with infrared lasers was developed in the 1980s (Ashkin *et al.* 1987), and is now being applied to bacterial isolations. Thus, an aggregate-forming coccoid thermophilic archaeon from 'Obsidian Pool', Yellowstone National Park, which was earlier known only from its morphology and its 16S rDNA sequence, was isolated from an enrichment culture dominated by rod-shaped thermophiles (Huber *et al.* 1995). The organism was later described as *Thermosphaera aggregans* (Huber *et al.* 1998).

Dilution cultures in sterile filtered unamended seawater, or seawater amended with very low concentrations of nutrients, have been used in attempts to isolate abundant types of oligotrophic marine prokaryote. After many weeks or months of incubation, several small oligotrophic bacteria were isolated (Button *et al.* 1993), and one of these, tentatively identified as a *Sphingomonas* sp., has been studied in further depth (Schut *et al.* 1993). This 'dilution to extinction' approach using unamended seawater was further developed by Connon & Giovannoni (2002) as a high-throughput method using microtitre plates and cell arrays on microscope slides for FISH staining to lower detection sensitivity. Thus, 0.1–0.2 ml aliquots of culture are sufficient for enumeration, and densities as low as 10^3 cells ml^{-1} can be detected. Using unamended seawater or seawater supplemented with very low concentrations of ammonium, phosphate and/or organic nutrients, up to 14% of the bacteria from coastal seawater could be cultured—a great increase over the values of 0.1–0.001% obtained using conventional methods (table 2). This approach recently led to the isolation of the elusive SAR11, the most abundant of all marine bacteria (Rappé *et al.* 2002; see also § 4). Cell densities achieved never exceeded $6\text{--}7 \times 10^5$ ml^{-1} . The organism, designated with the (malformed) name '*Candidatus* Pelagibacter ubique', is now awaiting a full taxonomic description.

Slowly but surely some of those micro-organisms that are numerically dominant in their environment, but that were unknown except for their 16S rDNA sequence, are now being obtained in culture. The case of SAR11 is one example. Another example is *Salinibacter*, an extremely halophilic member of the Bacteria which was originally recognized as an important component of the microbial community in saltern crystallizer ponds on the basis of FISH technology (Antón *et al.* 2000) and was later isolated by hybridizing colonies with specific 16S rRNA-based probes or by polar lipid analyses of selected colonies (Antón *et al.* 2002). Other organisms of special interest,

such as the proteorhodopsin-containing marine SAR86 or the thermophilic Korarchaeota inhabiting hot springs, still defy the microbiologists' attempts at their isolation.

10. FUNDING FOR MICROBIAL DIVERSITY STUDIES

Appropriate funding is needed for the development of the kind of innovative techniques required to obtain a reliable picture of the true microbial biodiversity in nature. Thus far, the level of funding for studies of prokaryote diversity has greatly lagged behind the funds available for plant and animal biodiversity studies. This may be due, to some extent, to the old notion that the number of prokaryote species is so small. As late as 1988 it was stated that 'Two of these kingdoms, the prokaryotic monerans and the eukaryotic protists, comprise microscopic unicellular organisms, and together they account for something like 5% of recorded living species' (May 1988, p. 1445). Now it is clear that the prokaryotes comprise two of life's three primary domains, that they are mainly responsible for the functioning of the biogeochemical cycles on Earth, and that they constitute an enormous reservoir of untapped biotechnological potential. However, the still common underestimation of the true diversity of prokaryotes has obvious negative effects on the distribution of research funding (Ward 1998; Garrity *et al.* 1999; Triplett 1999).

However, there is hope for the future. The US National Science Foundation has recognized the importance of microbial diversity, and in 1999 launched a programme establishing 'microbial observatories' (see <http://www.nsf.gov/pubs/awards/mo1999.htm>; accessed 17 April 2003). The goal of the 'microbial observatories' activity is to discover previously unknown microbes and to describe and characterize microbial diversity, phylogenetic relationships, interactions and other novel properties. In the framework of this exciting programme, funded with almost US\$25 million, 31 grants have been awarded for funding for up to 5 years, and a wide range of environments are being studied. It can only be hoped that this kind of activity will continue in the future and that other science funding agencies will take up the challenge of establishing similar funds to advance our understanding of microbial diversity on Earth.

11. CONCLUSIONS

Prokaryotic taxonomy is a dynamic science. The recognition of the Archaea as a separate third domain of life has revolutionized our views on prokaryote diversity and classification in the late 1970s and the 1980s. The development of techniques to characterize natural bacterial communities on the basis of 16S rDNA sequences has shown us how limited our understanding of the true diversity of prokaryotes in nature is. Only recently have we started to realize that we do not even know those Bacteria and Archaea that are most abundant in common ecosystems, aquatic as well as terrestrial. Novel approaches are slowly emerging towards the cultivation of some of the most intriguing yet uncultivated micro-organisms, but it will probably require much additional time and effort (depending on appropriate funding) until all the prokaryotes on Earth have been properly described and named

and arranged in a satisfactory classification scheme, and until the function of each species in its ecological niche is fully understood. New species are currently being described at a relatively slow rate: the numbers of new species added to the Notification Lists and the Validation Lists in the *International Journal of Systematic and Evolutionary Microbiology* in 2000, 2001 and 2002 were 236, 274 and 304, respectively. These are small numbers compared with the tens or probably even hundreds of thousands of species that are waiting to be described.

Now, 150 years after Ferdinand Cohn first started to name bacterial species and 80 years after the first edition of *Bergey's manual of determinative bacteriology* was published, taxonomy of prokaryotes still lacks a solid theoretical basis, and there is still no firm definition of the basic taxonomic unit, the species, used in their classification. Despite these obvious problems, a classification scheme has emerged that is now generally accepted.

According to the current species concept, bacterial taxonomy is, to a large extent, based on total genome comparison, as DNA-DNA similarity is one of the main criteria to delineate species. It may be expected that methods based on DNA chip technology and proteomics will change the way of identification of prokaryotes. Although sequencing of complete bacterial genomes is still not a routine procedure, more and more complete genome sequences of Bacteria and Archaea will become available in the near future. At the time of writing (April 2003) complete genomes of 16 species of Archaea and 81 species of Bacteria were available (see <http://mbgd.denome.ad.jp> and <http://www.tigr.org/tigr-scripts/CMR2/CMRGenomes.spl>; both accessed 1 May 2003). It is difficult to predict whether in the not too distant future whole genome sequences may become an integral part of the characterization of new species. However, comparative genomics will probably enable us to obtain a better species concept and to refine our approaches to prokaryote taxonomy.

The author thanks Erko Stackebrandt (DSMZ, Braunschweig) and two anonymous reviewers for critical reading of the manuscript and valuable comments.

REFERENCES

- Amann, G., Stetter, K. O., Llobet-Brossa, E., Amann, R. & Antón, J. 2000 Direct proof for the presence and expression of two 5% different 16S rRNA genes in individual cells of *Haloarcula marismortui*. *Extremophiles* **4**, 373–376.
- Amann, R. L., Ludwig, W. & Schleifer, K.-H. 1995 Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**, 143–169.
- Antón, J., Rosselló-Mora, R., Rodríguez-Valera, F. & Amann, R. 2000 Extremely halophilic bacteria in crystallizer ponds from solar salterns. *Appl. Environ. Microbiol.* **66**, 3052–3057.
- Antón, J., Oren, A., Benlloch, S., Rodríguez-Valera, F., Amann, R. & Rosselló-Mora, R. 2002 *Salinibacter ruber* gen. nov., sp. nov., a novel extremely halophilic member of the Bacteria from saltern crystallizer ponds. *Int. J. Syst. Evol. Microbiol.* **52**, 485–491.
- Ashkin, A., Dziedzic, J. M. & Yamane, T. 1987 Optical trapping and manipulation of single cells using infrared laser beams. *Nature* **330**, 769–771.
- Baas-Becking, L. G. M. 1934 *Geobiologie of inleiding tot de milieukunde*. Den Haag, The Netherlands: W. P. van Stockum en Zoon.

- Barns, S. M., Fundyga, R. E., Jeffries, M. W. & Pace, N. R. 1994 Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc. Natl Acad. Sci. USA* **91**, 1609–1613.
- Barns, S. M., Delwiche, C. F., Palmer, J. D. & Pace, N. R. 1996 Perspectives on archaeal diversity, thermophily and monophyly from environmental rDNA sequences. *Proc. Natl Acad. Sci. USA* **93**, 9188–9193.
- Beeder, J. R., Nielsen, K., Rosnes, J. T., Torsvik, T. & Lien, T. 1994 *Archaeoglobus fulgidus* isolated from hot North Sea oil field waters. *Appl. Environ. Microbiol.* **60**, 1227–1234.
- Béjà, O., Spudich, E. N., Spudich, J. L., Leclerc, M. & DeLong, E. F. 2000 Proteorhodopsin in the ocean. *Nature* **411**, 786–789.
- Béjà, O. (and 11 others) 2001 Bacterial rhodopsin: evidence for a new type of photosynthesis in the ocean. *Science* **289**, 1902–1906.
- Béjà, O., Koonin, E. V., Arvind, L., Taylor, L. T., Seitz, H., Stein, J. L., Bensen, D. C., Feldman, R. A., Swanson, R. V. & DeLong, E. F. 2002a Comparative genomic analysis of archaeal genotypic variants in a single population and in two different oceanic provinces. *Appl. Environ. Microbiol.* **68**, 335–345.
- Béjà, O., Suzuki, M. T., Heidelberg, J. F., Neson, W. C., Preston, C. M., Hamada, T., Eisen, J. A., Fraser, C. M. & DeLong, E. F. 2002b Unsuspected diversity among marine aerobic anoxygenic phototrophs. *Nature* **415**, 630–633.
- Bisby, F. A., Shimura, J., Ruggiero, M., Edwards, J. & Haeuser, C. 2002 Taxonomy, at the click of a mouse. *Nature* **418**, 367.
- Boetius, A., Ravensschlag, K., Schubert, C. J., Rickert, D., Widdel, F., Gieseke, A., Amann, R., Jørgensen, B. B., Witte, U. & Pfannkuche, O. 2000 A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**, 623–626.
- Brenner, D. J., Staley, J. T. & Krieg, N. R. 2001 Classification of procaryotic organisms and the concept of bacterial speciation. In *Bergey's manual of systematic bacteriology*, 2nd edn, vol. 1 (ed. D. R. Boone & R. W. Castenholz; editor-in-chief G. M. Garrity), pp. 27–31. New York: Springer.
- Britschgi, T. B. & Giovannoni, S. J. 1991 Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. *Appl. Environ. Microbiol.* **57**, 1707–1713.
- Broda, E. 1977 Two kinds of lithotrophs missing in nature. *Zeitschr. Allg. Mikrobiol.* **17**, 491–493.
- Button, D. K., Schut, P., Quang, R., Martin, R. & Robertson, B. R. 1993 Viability and isolation of marine bacteria by dilution culture: theory, procedures, and initial results. *Appl. Environ. Microbiol.* **59**, 881–891.
- Christensen, H., Bisgaard, M., Fredricksen, W., Møtters, 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.
- Cohan, F. M. 2001 Bacterial species and speciation. *Syst. Biol.* **50**, 513–524.
- Colwell, R. R., Brayton, P. R., Grimes, D. J., Roszak, D. R., Huq, S. A. & Palmer, L. M. 1985 Viable but nonculturable *Vibrio cholerae* and related pathogens in the environment: implications for the release of genetically engineered microorganisms. *Biotechnology* **3**, 817–820.
- Colwell, R. R., Clayton, R. A., Ortiz-Conde, B. A., Jacobs, D. & Russek-Cohen, E. 1995 The microbial species concept and biodiversity. In *Microbial diversity and ecosystem function* (ed. D. Allsopp, R. R. Colwell & D. L. Hawksworth), pp. 3–15. Wallingford: CAB International.
- Connon, S. A. & Giovannoni, S. J. 2002 High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl. Environ. Microbiol.* **68**, 3878–3885.
- Dalsgaard, T., Canfield, D. E., Petersen, J., Thamdrup, B. & Acuña-González, J. 2003 N₂ production by the anammox reaction in the anoxic water column of Golfo Dulce, Costa Rica. *Nature* **422**, 606–608.
- Dasch, G. A., Weiss, E. & Chang, K.-P. 1984 Endosymbionts of insects. In *Bergey's manual of systematic bacteriology*, vol. 1 (ed. N. R. Krieg & J. G. Holt), pp. 811–833. Baltimore, MD: The Williams & Wilkins Co.
- DeLong, E. F. 1992 Archaea in coastal marine environments. *Proc. Natl Acad. Sci. USA* **89**, 5685–5689.
- DeLong, E. F. & Pace, N. R. 2001 Environmental diversity of bacteria and archaea. *Syst. Biol.* **50**, 470–478.
- DeLong, E. F., Wickham, G. S. & Pace, N. R. 1989 Phylogenetic stains: ribosomal RNA-based probes for the identification of single microbial cells. *Science* **243**, 1360–1363.
- Dojka, M. K., Harris, J. K. & Pace, N. R. 2000 Expanding the known diversity distribution of an uncultured phylogenetic division of Bacteria. *Appl. Environ. Microbiol.* **66**, 1617–1621.
- Doolittle, W. F. 1999 Phylogenetic classification and the universal tree. *Science* **284**, 2124–2128.
- Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H. & Stackebrandt, E. (eds.) 1999 *The Prokaryotes: an evolving electronic resource for the microbiological community*, 3rd edn. New York: Springer. See <http://link.springer-ny.com/link/service/books/10125/>.
- Dykhuizen, D. E. 1998 Santa Rosalia revisited: why are there so many species of bacteria? *Antonie van Leeuwenhoek* **73**, 25–33.
- Dykhuizen, D. E. & Green, L. 1991 Recombination in *Escherichia coli* and the definition of biological species. *J. Bacteriol.* **173**, 7257–7268.
- Embley, T. M. & Stackebrandt, E. 1997 Species in practice: exploring uncultured prokaryote diversity in natural samples. In *Species: the units of biodiversity* (ed. M. F. Claridge, H. A. Dawah & M. R. Wilson), pp. 61–81. London: Chapman & Hall.
- Euzéby, J. 1997 List of bacterial names with standing in nomenclature: a folder available on the internet (URL: <http://www.bacterio.cict.fr/>). *Int. J. Syst. Bacteriol.* **47**, 590–592.
- Finlay, B. J. 2004 Protist taxonomy: an ecological perspective. *Phil. Trans. R. Soc. Lond. B* **359**, 599–610. (DOI 10.1098/rstb.2003.1450.)
- Fox, G. E., Wisotzkey, J. D. & Jurtshuk Jr, P. 1992 How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int. J. Syst. Bacteriol.* **42**, 166–170.
- Fuhrman, J. A., McCallum, K. & Davis, A. A. 1992 Novel major archaeobacterial group from marine plankton. *Nature* **356**, 148–149.
- Fulthorpe, R. R., Rhodes, A. N. & Tiedje, J. M. 1998 High level of endemicity of 3-chlorobenzoate-degrading soil bacteria. *Appl. Environ. Microbiol.* **64**, 1620–1627.
- García-Pichel, F., Prufert-Bebout, L. & Muyzer, G. 1996 Phenotypic and phylogenetic analyses show *Microcoleus chthonoplastes* to be a cosmopolitan cyanobacterium. *Appl. Environ. Microbiol.* **62**, 3284–3291.
- Garrity, G. M. (editor-in-chief) 2001 *Bergey's manual of systematic bacteriology*, 2nd edn, vol. 1 (ed. D. R. Boone & R. W. Castenholz). New York: Springer.
- Garrity, G. M., Boone, D. R., Brenner, D. J., Krieg, N. R. & Staley, J. T. 1999 The need for a new dimension of environmental prokaryotic biology in the NSF. *ASM News* **65**, 459.

- Giovannoni, S. J., Britschgi, T. B., Moyer, C. L. & Field, K. G. 1990a Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**, 60–63.
- Giovannoni, S. J., DeLong, E. F., Schmidt, T. M. & Pace, N. R. 1990b Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. *Appl. Environ. Microbiol.* **56**, 2572–2575.
- Godfray, H. C. L. 2002 Challenges for taxonomy. The discipline will have to reinvent itself if it is to survive and flourish. *Nature* **417**, 17–19.
- Gupta, R. S. 1998 What are archaeobacteria: life's third domain or monoderm prokaryotes related to Gram-positive bacteria? A new proposal for the classification of prokaryotic organisms. *Mol. Microbiol.* **29**, 695–707.
- Hicks, R. E., Amann, R. I. & Stahl, D. A. 1992 Dual staining of natural bacterioplankton with 4',6-diamidino-2-phenylindole and fluorescent oligonucleotide probes targeting kingdom-level 16S rRNA sequences. *Appl. Environ. Microbiol.* **58**, 2158–2163.
- Hohn, M. J., Hedlund, B. P. & Huber, H. 2002 Detection of 16S rDNA sequences representing the novel phylum 'Nanoarchaeota': indication for a wide distribution in high temperature biotopes. *Syst. Appl. Microbiol.* **25**, 551–554.
- Hoke, C. & Vedros, N. A. 1982 Taxonomy of the Neisseriaceae: deoxyribonucleic acid base composition, interspecies transformation, and deoxyribonucleic acid hybridization. *Int. J. Syst. Bacteriol.* **32**, 57–66.
- Huber, R., Burggraff, S., Mayer, T., Barns, S. M., Rossmagel, P. & Stetter, K. O. 1995 Isolation of a hyperthermophilic archaeum predicted by *in situ* RNA analysis. *Nature* **376**, 57–58.
- Huber, R., Dyba, D., Huber, H., Burggraf, S. & Rachel, R. 1998 Sulfur-inhibited *Thermosphaera aggregans* sp. nov., a new genus of hyperthermophilic archaea isolated after its prediction from environmentally derived 16S rRNA sequences. *Int. J. Syst. Bacteriol.* **48**, 31–38.
- Huber, H., Hohn, M. J., Rachel, R., Fuchs, T., Wimmer, V. C. & Stetter, K. O. 2002 A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**, 63–67.
- Hugenholtz, P., Goebel, B. M. & Pace, N. R. 1998 Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol.* **180**, 4765–4774.
- Istock, C. A., Bell, J. A., Ferguson, N. & Istock, N. L. 1996 Bacterial species and evolution: theoretical and practical perspectives. *J. Ind. Microbiol.* **17**, 137–150.
- Jetten, M. S. M., Strous, M., van de Pas-Schoonen, K. T., Schalk, J., van Dongen, U. G. J. M., van de Graaf, A. A., Logemann, S., Muyzer, G., van Loosdrecht, M. C. M. & Kuenen, J. G. 1998 The anaerobic oxidation of ammonium. *FEMS Microbiol. Rev.* **22**, 421–437.
- Karner, M. B., DeLong, E. F. & Karl, D. M. 2001 Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**, 507–510.
- Kluyver, A. J. & van Niel, C. B. 1956 *The microbe's contribution to biology*. Cambridge, MA: Harvard University Press.
- Knapp, S., Lamas, G., Nic Lughadha, E. & Novarino, G. 2004 Stability or stasis in the names of organisms: the evolving codes of nomenclature. *Phil. Trans. R. Soc. Lond. B* **359**, 611–622. (DOI 10.1098/rstb.2003.1445.)
- Kolber, Z. S., Van Dover, C. L., Niederman, R. A. & Falkowski, P. G. 2000 Bacterial photosynthesis in surface water of the open ocean. *Nature* **407**, 177–179.
- Kolber, Z. S., Plumley, F. G., Lang, A. S., Beatty, J. T., Blankenship, R. E., VanDover, C. L., Vetriani, C., Kobelzik, M., Rathgeber, C. & Falkowski, P. G. 2001 Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean. *Science* **292**, 2492–2495.
- Kuypers, M. M. M., Slikkers, A. O., Lavik, G., Schmid, M., Jørgensen, B. B., Kuenen, J. G., Sinninghe Damsté, J. S., Strous, M. & Jetten, M. S. M. 2003 Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. *Nature* **422**, 608–611.
- Lan, R. & Reeves, P. R. 2001 When does a clone deserve a name? A perspective on bacterial species based on population genetics. *Trends Microbiol.* **9**, 419–424.
- Lapage, S. P., Sneath, P. H. A., Jr Lessel, E. F., Skerman, V. B. D., Seelinger, H. P. R. & Clark, W. A. (eds) 1992 *International Code of Nomenclature of Bacteria (1990 revision)*. Washington, DC: American Society of Microbiology.
- Lee, M. S. Y. 2002 Online database could end taxonomic anarchy. *Nature* **417**, 787–788.
- Lee, N., Nielsen, P. H., Andreasen, K.-H., Juretschko, S., Nielsen, J. L., Schleifer, K.-H. & Wagner, M. 1999 Combination of fluorescent *in situ* hybridization and microautoradiography—a new tool for structure–function analyses in microbial ecology. *Appl. Environ. Microbiol.* **65**, 1289–1297.
- Lengeler, J. W., Drews, G. & Schlegel, H. G. (eds) 1999 *Biology of the prokaryotes*. Stuttgart: Thieme.
- Liesack, W. & Stackebrandt, E. 1992 Occurrence of novel groups of the domain Bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. *J. Bacteriol.* **174**, 5072–5078.
- Ludwig W. & Klenk H.-P. 2001 Overview: a phylogenetic backbone and taxonomic framework for procaryotic systematics. In *Bergey's manual of systematic bacteriology*, 2nd edn, vol. 1 (ed. D. R. Boone & R. W. Castenholz; editor-in-chief G. M. Garrity), pp. 49–65. New York: Springer.
- Ludwig, W. & Schleifer, K.-H. 1994 Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiol. Rev.* **15**, 155–173.
- McDougald, D. & Kjelleberg, S. 1999 New perspectives on the viable but nonculturable response. *Biologia* **54**, 617–623.
- McDougald, D., Rice, S. A., Weichert, D. & Kjelleberg, S. 1998 Nonculturability: adaptation or debilitation? *FEMS Microbiol. Ecol.* **25**, 1–9.
- Madigan, M. T., Martinko, J. M. & Parker, J. 2003 *Brock biology of microorganisms*, 10th edn. Upper Saddle River, NJ: Prentice-Hall.
- Martinez-Murcia, A. J., Benlloch, S. & Collins, M. D. 1992 Phylogenetic interrelationships of members of the genera *Aeromonas* and *Plesiomonas* as determined by 16S ribosomal DNA sequencing—lack of congruence with results of DNA–DNA hybridizations. *Int. J. Syst. Bacteriol.* **42**, 412–421.
- May, R. M. 1988 How many species are there on Earth? *Science* **241**, 1441–1449.
- Moreira, D. & López-García, P. 2002 The molecular ecology of microbial eukaryotes unveils a hidden world. *Trends Microbiol.* **10**, 31–38.
- Morris, R. M., Rappé, M. S., Connon, S. A., Vergin, K. L., Siebold, W. A., Carlson, C. A. & Giovannoni, S. J. 2002 SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**, 806–810.
- Murray, R. G. E. & Stackebrandt, E. 1995 Taxonomic note: implementation of the provisional status *Candidatus* for incompletely described prokaryotes. *Int. J. Syst. Bacteriol.* **45**, 186–187.
- Mylvaganam, S. & Dennis, P. P. 1992 Sequence heterogeneity between the two genes encoding 16S rRNA from the halophilic archaeobacterium *Haloarcula marismortui*. *Genetics* **130**, 399–410.
- Nübel, U., Engelen, B., Felske, A., Schnaidr, J., Wieshuber, A., Amann, R. I., Ludwig, W. & Backhaus, H. 1996 Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. *J. Bacteriol.* **178**, 5636–5643.

- O'Connell, S. P., Lehman, R. M., Snoeyenbos-West, O., Winston, V. D., Cummings, D. E., Watwood, M. E. & Colwell, F. S. 2003 Detection of Euryarchaeota and Crenarchaeota in an oxic basalt aquifer. *FEMS Microbiol. Ecol.* **44**, 165–173.
- Oliver, J. D., Nilsson, L. & Kjelleberg, S. 1991 Formation of nonculturable *Vibrio vulnificus* cells and its relationship to the starvation state. *Appl. Environ. Microbiol.* **57**, 2640–2644.
- Olsen, G. J. 1994 Archaea, Archaea, everywhere. *Nature* **371**, 657–658.
- Oren, A. & Stackebrandt, E. 2002 Prokaryote taxonomy online: challenges ahead. *Nature* **419**, 15.
- Orphan, V. J., House, C. H., Hinrichs, K.-U., McKeegan, K. D. & DeLong, E. F. 2002 Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. *Proc. Natl Acad. Sci. USA* **99**, 7663–7668.
- Ouverney, C. C. & Fuhrman, J. A. 2000 Marine planktonic archaea take up amino acids. *Appl. Environ. Microbiol.* **66**, 4829–4833.
- Pace, N. R. 1997 A molecular view of microbial diversity and the biosphere. *Science* **276**, 734–740.
- Palys, T. F., Cohan, F. M. & Nakamura, L. K. 1997 Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int. J. Syst. Bacteriol.* **47**, 1145–1156.
- Partensky, F., Hess, W. R. & Vaulot, D. 1999 *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* **63**, 106–127.
- Preston, C. M., Wu, K. Y., Molinski, T. F. & DeLong, E. F. 1996 A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc. Natl Acad. Sci. USA* **93**, 6241–6246.
- Rainey, F. A., Ward-Rainey, N. L., Janssen, P. H., Hippe, H. & Stackebrandt, E. 1996 *Clostridium paradoxum* DSM 7308^T contains multiple 16S rRNA genes with heterogeneous intervening sequences. *Microbiology* **142**, 2087–2095.
- Ramsing, N. B., Ferris, M. J. & Ward, D. M. 2000 Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobacterial mat. *Appl. Environ. Microbiol.* **66**, 1038–1049.
- Rappé, M. S., Suzuki, M. T., Vergin, K. L. & Giovannoni, S. J. 1998 Phylogenetic diversity of ultraplankton plastid small-subunit rRNA genes recovered in environmental nucleic acid samples from the Pacific and Atlantic coasts of the United States. *Appl. Environ. Microbiol.* **64**, 294–303.
- Rappé, M. S., Connon, S. A., Vergin, K. L. & Giovannoni, S. J. 2002 Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* **418**, 630–633.
- Rosselló-Mora, R. & Amann, R. 2001 The species concept for prokaryotes. *FEMS Microbiol. Rev.* **25**, 39–67.
- Roszak, D. B. & Colwell, R. R. 1987 Survival strategies of bacteria in the natural environment. *Microbiol. Rev.* **51**, 365–379.
- Roszak, D. B., Grimes, D. J. & Colwell, R. R. 1984 Viable but non-recoverable stage of *Salmonella enteritidis* in aquatic systems. *Can. J. Microbiol.* **30**, 334–338.
- Scanlan, D. J. & West, N. J. 2002 Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. *FEMS Microbiol. Ecol.* **40**, 1–12.
- Schut, F., deVries, E. J., Gottschal, J. C., Robertson, B. R., Harder, W., Prins, R. A. & Button, D. K. 1993 Isolation of typical marine bacteria by dilution culture: growth, maintenance, and characteristics of isolates under laboratory conditions. *Appl. Environ. Microbiol.* **59**, 2150–2160.
- Skerman, V. B. D., McGowan, V. & Sneath, P. H. A. 1980 Approved lists of bacterial names. *Int. J. Syst. Bacteriol.* **30**, 225–420.
- Sneath, P. H. A. 1977 The maintenance of large numbers of strains of microorganisms and the implications for culture collections. *FEMS Microbiol. Lett.* **1**, 333–334.
- Sørensen, K. B., Finster, K. & Ramsing, N. B. 2001 Thermodynamic and kinetic requirements in anaerobic methane oxidizing consortia exclude hydrogen, acetate, and methanol as possible electron shuttles. *Microbiol. Ecol.* **42**, 1–10.
- Stackebrandt, E. 2000 Defining taxonomic ranks. In *The Prokaryotes: an evolving electronic resource for the microbiological community*, 3rd edn, release 3.1, 20 January 2000. New York: Springer. See <http://link.springer-ny.com/link/service/books/10125/>.
- Stackebrandt, E. & Goebel, B. M. 1994 Taxonomic note: a place for DNA:DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.* **44**, 846–849.
- Stackebrandt, E. (and 12 others) 2002 Report of the *ad hoc* committee for the re-evaluation of the species definition in bacteriology. *Int. J. Syst. Evol. Microbiol.* **52**, 1045–1047.
- Staley, J. T. 1997 Biodiversity: are microbial species threatened? *Curr. Opin. Biotechnol.* **8**, 340–345.
- Staley, J. T. 1999 Bacterial biodiversity: a time for place. *ASM News* **65**, 681–687.
- Staley, J. T. & Konopka, A. 1985 Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *A. Rev. Microbiol.* **39**, 1379–1384.
- Stein, J. L. & Simon, M. I. 1996 Archaeal ubiquity. *Proc. Natl Acad. Sci. USA* **93**, 6228–6230.
- Stein, J. L., Marsh, T. L., Wu, K. Y., Shizuya, H. & DeLong, E. F. 1996 Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *J. Bacteriol.* **178**, 591–599.
- Takai, K., Moser, D. P., DeFlaun, M., Onstott, T. C. & Fredrickson, J. K. 2001 Archaeal diversity in waters from deep South African gold mines. *Appl. Environ. Microbiol.* **67**, 5750–5760.
- Torsvik, V., Goksøyr, J. & Daane, F. L. 1990a High diversity of DNA in soil bacteria. *Appl. Environ. Microbiol.* **56**, 782–787.
- Torsvik, V., Salte, K., Sørheim, R. & Goksøyr, J. 1990b Comparison of phenotypic diversity and DNA heterogeneity in a population of soil bacteria. *Appl. Environ. Microbiol.* **56**, 776–781.
- Torsvik, V., Sørheim, R. & Goksøyr, J. 1996 Total bacterial diversity in soil and sediment communities—a review. *J. Ind. Microbiol.* **17**, 170–178.
- Torsvik, V., Daane, F. L., Sandaa, R. A. & Øverås, L. 1998 Novel techniques for analysing microbial diversity in natural and perturbed environments. *J. Biotechnol.* **64**, 53–62.
- Tripplett, E. W. 1999 Proposal and justification for a new division of environmental prokaryotic biology in the NSF. *ASM News* **65**, 63–64.
- Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K. & Swings, J. 1996 Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol. Rev.* **60**, 407–438.
- von Wintzingerode, F., Göbel, U. & Stackebrandt, E. 1997 Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* **21**, 213–229.
- Wang, Y., Zhang, Z. & Narandakumar, R. 1997 The actinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. *J. Bacteriol.* **179**, 3270–3276.
- Ward, D. M. 1998 A natural species concept for prokaryotes. *Curr. Opin. Microbiol.* **1**, 271–277.
- Ward, D. M., Weller, R. & Bateson, M. M. 1990 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**, 63–65.

- Ward, D. M., Ferris, M. J., Nold, S. C. & Bateson, M. M. 1998 A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol. Mol. Biol. Rev.* **62**, 1353–1370.
- Waterbury, J. B., Watson, S. W., Guillard, R. R. & Brand, L. E. 1979 Widespread occurrence of a unicellular, marine, planktonic, cyanobacterium. *Nature* **277**, 293–294.
- Wayne, L. G. (and 11 others) 1987 Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol.* **37**, 463–464.
- Whitman, W. B., Coleman, D. C. & Wiebe, W. J. 1998 Prokaryotes: the unseen majority. *Proc. Natl Acad. Sci. USA* **95**, 6578–6583.
- Woese, C. R. 1987 Bacterial evolution. *Microbiol. Rev.* **51**, 221–271.
- Woese, C. R. 1992 Prokaryote systematics: the evolution of a science. In *The prokaryotes. A handbook on the biology of bacteria: ecophysiology, isolation, identification, applications*, 2nd edn, vol. 1 (ed. A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K.-H. Schleifer), pp. 3–18. Berlin: Springer.
- Woese, C. R., Kandler, O. & Wheelis, M. L. 1990 Towards a natural system of organisms: proposal of the domains Archaea, Bacteria and Eucarya. *Proc. Natl Acad. Sci. USA* **87**, 4576–4579.
- Wuchter, C., Schouten, S., Boschker, H. T. S. & Sinninghe Damsté, J. S. 2003 Bicarbonate uptake by marine Crenarchaeota. *FEMS Microbiol. Lett.* **219**, 203–207.

GLOSSARY

- ARDRA: amplified ribosomal DNA restriction analysis
FISH: fluorescence *in situ* hybridization
VBNC: viable but non-culturable