3
Microbial Systematics

Aharon Oren

CONTENTS

Abstract This chapter aims at presenting an overview of different aspects of the classification and nomenclature of the prokaryotes, i.e., the domains *Bacteria* and *Archaea*. Concepts of systematics, taxonomy, classification, nomenclature, and identification are discussed. The number of species of prokaryotes – *Bacteria* and *Archaea* combined – is surprisingly small, 8,226 as of November 2009. It is obvious that the true number of different species much larger: we probably know less than 1% of the types of prokaryotes in Nature. Classification of prokaryotes and description of species based on a polyphasic approach that includes phenotypic as well as genotypic properties. Nomenclature is governed by the rules of the Bacteriological Code (International Code of Nomenclature of Prokaryotes), as determined by the International Committee on Systematics of Prokaryotes.

Key Words Prokaryotes • *Bacteria* • *Archaea* • systematics • taxonomy • nomenclature • bacteriological code • culture collections • small-subunit rRNA.

From: *Handbook of Environmental Engineering, Volume 10: Environmental Biotechnology*
Edited by: L. K. Wang et al., DOI: 10.1007/978-1-60327-140-0_3

1. INTRODUCTION

Environmental biotechnologists and environmental engineers often deal with microorganisms and especially with prokaryotes. These are the organisms responsible for majority of the biodegradative processes of organic carbon in nature, as well as for important reactions in the cycles of the elements including nitrification, denitrification, dissimilatory sulfate reduction, sulfide oxidation, and many others.

It is therefore inevitable that the environmental engineer will encounter many names of microorganisms – prokaryotic as well as eukaryotic – that are involved in the processes studied. The engineer will also be challenged from time to time to identify certain microorganisms when describing the phenomena occurring in the ecosystem under study, as well as to provide information on the organism(s) that causes environmental problems and/or those that provide possible solutions to those problems.

This chapter aims at presenting an overview of different aspects of the classification and nomenclature of the prokaryotes, i.e., the domains *Bacteria* and *Archaea*. These groups are by far the most important as far as their metabolic diversity is concerned. Their classification and nomenclature are governed by the same rules. The classification of the eukaryotic microorganisms in part follows the rules of botanical taxonomy (fungi, yeasts, algae) and in part the zoological taxonomy (protozoa and larger animals). The formal rules used in the nomenclature of these groups are greatly different from those used for the prokaryotes.

The highly diverse prokaryotic world has presented serious challenges to the many microbiologists who have attempted to achieve a satisfactory classification in the past. Many of the problems are yet to be solved, and as a result, the classification schemes have been subject to frequent changes with the increase of our understanding of the physiology and the molecular properties of the different groups of microorganisms. Prokaryote systematics is a highly dynamic science in which the concepts rapidly change in accordance with the development of new techniques and approaches. There is no official classification of prokaryotes, as will be shown in the sections below (this in contrast with prokaryote nomenclature, which is governed by a series of internationally approved rules and regulations). The basic unit of the classification of all living organisms is the species, and it may be surprising to many to read that in the case of the prokaryotes there is no universally recognized definition of the species. Prokaryotic systematics thus lacks a firm theoretical basis. Microbiologists work with a species concept that is much broader than that used in the disciplines of botany and zoology. The number of species of prokaryotes (*Bacteria* and *Archaea* combined, i.e., two out of the three domains of life – *Bacteria*, *Archaea*, and *Eucarya*) that have been described and named may even seem surprisingly small, 8,226 as of November 3, 2009 (1). This number is extremely small compared to the more than a million of insects described, and about 22,000 members of the *Orchidaceae*, a single family of vascular plants. In the following we will explore several aspects of classification, nomenclature, and identification of prokaryotes. We will present evidence for the existence of many more species of bacteria on Earth than the about 8,000 species documented thus far; experimental data suggest the existence of at least two orders of magnitude more species, and possibly even many more. Today, it is generally accepted that less than 1% of the prokaryote species that inhabit our planet have been named.

It is not feasible in a single chapter to provide in-depth information and complete practical guidelines on how to identify, name, and classify prokaryotes. However, environmental engineers and other professionals working in related fields are sometimes faced with the need to identify strains of microorganisms. This chapter intends to summarize some of the current concepts in the field of prokaryote systematics and will refer the reader to more detailed sources of information on the subject.

2. SYSTEMATICS, TAXONOMY, AND NOMENCLATURE OF PROKARYOTES

2.1. General Definitions

Systematics is the scientific study of organisms with the ultimate object of characterizing and arranging them in an orderly manner. The terms "systematics" and "taxonomy" are sometimes used as synonyms. The term "taxonomy" is often defined as the theory and practice of classifying organisms into groups (taxa) on the basis of similarities and relationships. Systematics generally signifies a broader concept that includes the evolutionary and phylogenetic relationships of the organisms studied. Taxonomy can be subdivided into three disciplines: classification, nomenclature, and identification (2). Classification is the orderly arrangement of units into groups. There are many ways in which living organisms can be arranged in groups, and taxonomy is therefore a most subjective branch of science. Classification is purpose-oriented, and there may be different ways that lead toward successful classification (3).

Nomenclature is the assignment of names to the taxonomic groups defined during classification. The rules of nomenclature of living organisms are agreed upon internationally, and they are laid down in three documents: the Botanical Code, the Zoological Code, and the Bacteriological Code, each dealing with a specific group. Here, we will discuss only the International Code of Nomenclature of Bacteria (the Bacteriological Code) (4), renamed as International Code of Nomenclature of Prokaryotes in 2000, and its implications for the nomenclature of prokaryotes. Similar to plants and animals, prokaryotes are given a genus and a species name according to the binomial system introduced by Linnaeus in the middle of the eighteenth century. For example, the genus *Bacillus*, a genus of Gram-positive aerobic endospore-forming bacteria, presently (November 2009) contains about 160 species; examples include *Bacillus subtilis* (a species from which valuable proteolytic enzymes are produced), *Bacillus anthracis* (the anthrax bacillus), and *Bacillus thuringiensis* (a species of which certain strains produce potent substances that kill mosquito larvae) (1).

The nomenclature of prokaryotes is subject to changes, and there are many examples of species that have been renamed, moved to other existing genera, or reclassified in newly established genera in accordance to new insights. For example, the species formerly known as *Bacillus polymyxa* and *Bacillus stearothermophilus* have been renamed *Paenibacillus polymyxa* and *Geobacillus stearothermophilus*, respectively (1), when it became desirable to split up the genus *Bacillus* based on an increased understanding of the phylogenetic relationships within the aerobic Gram-positive endospore-forming bacteria, mainly on the basis of sequence analysis of small-subunit ribosomal RNA (see also Sect. 6.1).

Identification is the practical use of classification schemes and the labeling schemes provided by nomenclature to establish the identity of isolated microorganisms as members of

previously identified species. Identification can be defined as the practical use of classification and nomenclature to determine the identity of an isolate as a member of an established taxon or as a member of a previously unidentified species.

It should be stressed here once more that there is no "official" classification of prokaryotes, but there is an official nomenclature, regulated by internationally agreed-upon rules. The classification provided in *Bergey's Manual of Systematic Bacteriology* (5) (see also Sect. 7.1) is widely used among microbiologists, but was never intended to obtain official status.

The classification used in Bergey's Manual divides the prokaryotes into two domains: the *Bacteria* and the *Archaea*. The domain is therewith the highest taxonomic rank. According to the same classification scheme, all eukaryotic organisms – microorganisms and macroorganisms, plants, and animals – belong to the domain *Eucarya*. The domains, *Bacteria* and *Archaea,* are divided into phyla, which each encompass one or more classes (6, 7). The classes are divided into orders, which are subdivided into families, genera, and species. As stated previously, the species is the basic unit of taxonomy. Sometimes, subdivision of a species into subspecies and infra-subspecific units is necessary.

An important term in bacterial taxonomy and nomenclature is that of *type strain*. For each species, a type strain has been designated, which is the name-bearer of that species and is the reference specimen for the name. Whenever a new prokaryote species is described, the authors are obliged to deposit the type strain of the species in at least two publicly accessible culture collections located in different countries for safekeeping (see Sect. 5), and make subcultures available to any interested scientist for further study. Identification of unknown isolates should use such type strains of recognized species for comparison. The terms "strain" and "isolate" refer to the descendants of a single isolation in pure culture. They are usually made up of a succession of cultures ultimately derived from an initial single colony. For each genus, a type species is defined, and for each family and each order, a type genus is designated.

2.2. *The Definition of the Prokaryote Species*

To the botanist and the zoologist, the definition of the concept "species" presents relatively few problems. In the plant and the animal world, a species is generally defined as a population of individuals that can interbreed under natural conditions, produce fertile offspring, and that is reproductively isolated from other populations. Such a definition is useless in the case of the prokaryotes, as these show no sexual reproduction.

There is no general consensus about the definition of the concept of the species, i.e., the basic taxonomic unit in the prokaryote world (8–11). Definitions found in the literature may for example circumscribe the species as "a distinct group of strains that have certain distinguishing features and that generally bear a close resemblance to one another in the more essential features of organization," or "an assemblage of clonal populations that share a high degree of phenotypic similarity, coupled with an appreciable dissimilarity from other assemblages of the same general kind." Such definitions provide little practical information on how close that resemblance and similarity should be for two strains to be classified in the same species, what features of organization should be considered essential, and what degree of dissimilarity is required for two strains to warrant classification in different species.

The delineation of species according to such definitions is therefore highly subjective. Some definitions found in the literature stress this subjectivity to an even larger extent, such as that of a species as "a group of organisms defined more or less subjectively by the criteria chosen by the taxonomist to show to best advantage as far as possible and putting into practice his individual concept of what a species is" (12).

It is obvious that a simpler, more pragmatic species definition is required to enable the design of classification schemes and the establishment of a nomenclature that can be widely used. Based on the experience of the last 20 years, such a pragmatic definition of the prokaryote species has indeed emerged. This species concept is based on the recommendations published in 1987 by a committee of experts (9). These recommendations were recently confirmed and extended by a new ad-hoc committee (10). The species concept is based on a polyphasic approach (see also Sect. 3.3), which includes description of diagnostic phenotypic features combined with genomic properties. Consistency of phenotypic and genomic characters is required to generate a useful classification system for the prokaryotes (13). It is recommended that a distinct genospecies (i.e., a species discernible only by nucleic acid comparisons) that cannot be differentiated from another genospecies on the basis of any known phenotypic property not be named until some phenotypic differentiating property is found (9). Individually, many of the phenotypic and chemotaxonomic characters used as diagnostic properties are insufficient to delineate species, but together they provide sufficient descriptive information to allow the definition of a species. Description of a species should ideally be based on a comparative study of a large number of isolates to define also the degree of variation of certain properties within the defined boundaries of the species. In practice, however, more than 80% of the new species descriptions that have been published in recent years were based on the study of single isolates. Unfortunately, such species descriptions that are based on one or a few strains only can lead to improper phenotypic circumscription of taxa, making the identification of new isolates as members of a taxon problematic. From time to time, formal proposals have been made to the International Committee for Systematics of Prokaryotes/International Committee for Systematic Bacteriology (see Sect. 4.3) to define a minimum number of isolates necessary as the basis of a description of a new species. However, such proposals have never been formally approved. If they would have, many extremely interesting *Bacteria* and *Archaea* that have been proposed and approved as new species in recent years could not have been described simply because it had not been feasible to obtain more isolates. Our understanding of the prokaryote world would have been much poorer as a result.

The genomic properties to be determined for the delineation of species are based on comparisons of the complete genomes using techniques of DNA similarity determination by DNA–DNA hybridization and/or assessment of the difference in the melting temperature between the homologous and the heterologous DNA hybrids (7, 11). These tests have to be performed under carefully standardized conditions to give reproducible results. The widely accepted criterion defines a prokaryotic species as a group of strains, including the type strain, that share at least 70% total genome DNA–DNA hybridization and have less than $5^{\circ}\mathrm{C}\ \Delta T_{\mathrm{m}}$ (= the difference in the melting temperature between the homologous and the heterologous hybrids formed under standard conditions). The delineation value of 70%, as introduced around 1987 (9), is artificial, but has proven satisfactory in most cases. DNA relatedness

values between 30 and 70% point to a moderate degree of relationship – often parallel to the extent of the genus. There are cases in the literature of species and even genera that share more than 70% DNA–DNA similarity. A well-known case is that of the genera *Escherichia* and *Shigella*, enteric bacteria that share more than 85% similarity. However, for pragmatic reasons, the separation into two species is maintained. At the level of genera or higher taxa, the resolving power of DNA–DNA similarity assays is limited.

It should be stated that DNA–DNA hybridization assays are seldom routinely performed in most laboratories. The protocol includes labeling of the DNA (generally with radioisotopes), shearing, and denaturing the sheared DNA, whereafter the labeled denatured DNA is mixed with excess of unlabeled DNA from the second organism. The mixture is then cooled and allowed to reanneal under carefully controlled conditions. Duplex DNA is then separated from any unhybridized DNA remaining and the amount of bound DNA quantified. A control experiment with homologous DNA is included and its results of the heterologous DNA binding are normalized with respect to the homologous control. Different protocols exist, and nonradioactive methods have also been introduced. The procedure is time-consuming, and allows for pair-wise comparisons only, making comparisons of large numbers of strains cumbersome. A recent reappraisal of the currently used methods in bacterial taxonomy to delineate species concluded that, despite certain drawbacks with respect to reproducibility and workability, DNA–DNA hybridization is still the best criterion for species delineation. Its great advantage is its universal applicability. A disadvantage is that the method gives no indication of which genes contribute to or detract from the similarity. In spite of the usefulness of DNA–DNA similarity determinations, phenotypic properties including chemotaxonomic markers will remain essential to describe new species (10), and such phenotypic properties will always be essential as diagnostic markers to be used when new isolates are to be identified.

Given that there is no clear definition of the prokaryotic species, the guidelines for the delineation of genera or higher taxonomic levels within the prokaryotes are even less clear. The genus may be defined as "a collection of species with many characters in common," but the extent of the shared characters that should exist for species to be classified in a single genus is largely a matter of personal judgment. There is, however, a general consensus that the division into higher taxonomic levels should reflect phylogenetic relationships. As explained in Sect. 6.1, sequence analysis of small-subunit ribosomal RNA (16S rRNA in prokaryotes, 18S rRNA for eukaryotes) has provided a large extent of insight in the phylogenetic relationships among microorganisms. The construction of small subunit rRNA-based phylogenetic trees provides the taxonomist with a powerful technique to determine the phylogenetic position of an isolate to the level of family and genus (14–16). Generally, there is a good correlation between the DNA–DNA similarity and the similarity of the 16S rRNA gene sequence (17). A DNA–DNA similarity of less than 70% generally corresponds with less than 97% 16S rRNA sequence identity. Species classified in a single genus generally share at least 93–95% identity in their 16S rRNA gene sequence. However, at the species level, the 16S rRNA-based methods lack the necessary resolving power, and then DNA–DNA reassociation experiments are still required (18) (see also Sect. 3.3). There are known cases in which two distinctly different species with DNA–DNA similarity of less than 50% have identical 16S rRNA gene sequences.

Compared to the species concept in the plant and the animal world, the bacterial species concept is exceptionally broad. When we would apply the "70% DNA–DNA similarity" criterion to delineate species of higher eukaryotes, the number of recognized species would decrease dramatically. For example, humans would not only belong to the same species as chimpanzees (98.4% DNA–DNA relatedness), gorillas (97.7%), and orangutans (96.5%), but even lemurs (78%) would be classified in the same species! The prokaryotic species, thus, encompasses species that may be highly different on the genetic level (8).

The availability of the methodology to determine complete genome sequences of prokaryotes (generally ranging in length between 2×10^6 and 6×10^6 base pairs) will probably change the rules and concepts used in delineating species in the future. Until now, the sequencing and annotation of bacterial genomes is a time-consuming and costly process, however, the number of complete bacterial and archaeal genome sequences that are available in public data bases is rapidly increasing (952 and 71, respectively, as of November 2009). With the rapid developments in automated genome sequencing technology, the day may come when the publication of the complete genome sequence of the type strain may become obligatory for the description of a new species. In the meantime, the comparison of the existing genome sequences has taught us much important information not only about the structure of the prokaryote genome, but also about the extent of possible lateral gene transfer between prokaryotes, not only at the level of species and genera, but even at the level of phyla and domains (*Archaea–Bacteria*, *Bacteria–Eucarya*, *Archaea–Eucarya*). The question has rightfully been asked whether is still would it be possible to delineate species at all if indeed genes move freely from species to species, even between completely unrelated ones. The conclusion of almost a decade of studies since the first complete bacterial genome sequence was published is, that indeed prokaryotes can "capture" new genes from other organisms, sometimes extremely distantly related. However, each species still appears to have its genetic individuality, and it is surely not so that life is a common gene pool, shared more or less randomly between all organisms that inhabit our planet. There are barriers that prevent a too extensive exchange of genes, and as a result it is still possible to formulate a species concept in the prokaryote world. As stressed above, the DNA–DNA hybridization method is based on the similarities of complete genomes, not on the presence or absence of single genes, so that the result of the test will not be greatly affected by the lateral transfer of a few genes obtained from other organisms.

2.3. The Number of Prokaryotes that Have Been Described

Bacterial nomenclature saw a new beginning in 1980. In view of the confusion that had arisen by that time, with many species of prokaryotes being known under different synonyms, it was decided that the thousands of bacterial names that had been published in the past in the greatly dispersed scientific literature would lose their validity, with the exception of approved lists of about 2,500 species names that were published in that year (19). Since that date, only those names of new prokaryote taxa published in the lists in the *International Journal of Systematic Bacteriology* (from 2000 onwards renamed as the *International Journal of Systematic and Evolutionary Microbiology*, see also Sect. 4.4) obtain standing in the nomenclature of prokaryotes. Such new names can either be published in the form of original articles published

Table 3.1
The Number of Prokaryotes described (*Bacteria* and *Archaea* Combined) with names with standing in Prokaryote nomenclature, as of November 3, 2009. Derived from www.bacterio.cict.fr

Number of phyla[a]	26
Number of classes[b]	70
Number of orders	116
Number of families	253
Number of genera	1,732
Number of species	8,226

[a] The term "Phylum" is not covered by the Bacteriological Code (4).
[b] The term "Class" is not covered by the Bacteriological Code (4).

in that journal, or by including the proposed new names in the "Validation Lists" of species that had before been described ("effectively published") in other scientific journals. This rule has greatly simplified bacterial taxonomy and nomenclature, and it is thus easy to keep track of the number of published species and their names.

Table 3.1 gives an overview of the number of names of taxa (species, genera, families, orders, classes, phyla, and domains) in the prokaryote world, *Bacteria* and *Archaea* combined, that have been validly described by November 3, 2009. These numbers are updated bimonthly with the publication of each new issue of the *International Journal of Systematic and Evolutionary Microbiology*. The "List of Bacterial Names with Standing in Nomenclature" website http://www.bacterio.cict.fr, maintained by Dr. Jean Euzéby (1) provides updated information on the number of name of species (and subspecies) and on the higher taxa of prokaryotes with standing in nomenclature (see also Sect. 4.5).

3. CLASSIFICATION OF PROKARYOTES

Historically, phenotypic properties have dominated the classification schemes of bacteria. In the past, identification was primarily based on properties such as cell morphology, staining properties (the Gram stain; acid-fastness stain), the ability to grow on certain carbon sources, excretion of certain end products, presence or absence of certain enzymatic activities, temperature and pH range of growth, etc. The first genomic property to be included in species descriptions of prokaryotes was the determination of the guanine + cytosine (G + C) percentage in the DNA. This property was introduced in the early 1960s, and has retained its value in bacterial taxonomy ever since, to the extent that no description of a new species is complete without it. Even if the amount of information obtained by the determination of the G + C percentage is limited (it does not provide any information on where in the genome these guanine and cytosine bases are found), it has a distinct advantage that it characterizes the complete genome, not a small part of it that may have a special function. With the technological advancement in DNA sequencing, more genotypic properties were found to be of value in bacterial characterization, classification, and identification. The best known example is of course the sequence determination of the small subunit ribosomal RNA

Table 3.2
Some properties used in identification and polyphasic taxonomy of prokaryotes (20)

Genotypic information
Properties based on the total DNA
Determination of the mol% Guanine + Cytosine
Restriction patterns: Restriction fragment length polymorphism (RFLP), Low frequency restriction fragment analysis (PFGE)
DNA–DNA hybridization
Determination of the genome size
Properties based on DNA segments
Polymerase chain reaction (PCR) based DNA fingerprinting: ribotyping, Amplified rDNA restriction analysis (ARDRA), Randomly amplified polymorphic DNA (RAPD), Amplified fragment length polymorphism (AFLP), Arbitrarily primed PCR, repetitive element sequence-based PCR
DNA sequencing of selected genes
Use of specific probes for detection of selected genes
RNA-based properties
RNA sequencing
Determination of low molecular weight RNA profiles
Phenotypic information
Protein-based properties
Electrophoretic patterns (one- or two-dimensional) of total cellular or cell envelope proteins
Enzyme patterns (multilocus enzyme electrophoresis)
Chemotaxonomic markers
Cellular fatty acids (detected as fatty acid methyl esters)
Detection and characterization of mycolic acids
Polar lipid characterization
Identification of respiratory quinones
Identification of cellular polyamines
Characterization of the cell wall and of extracellular polysaccharides
Phenotypically expressed properties
Cell morphology
Physiological properties such as the range of substrates used
Enzymological tests
Serological characterization using monoclonal or polyclonal antibodies

(for prokaryotes: 16S rRNA). As discussed above (see Sect. 2.2), DNA–DNA hybridization similarity, another genomic property, has obtained special status when it has to be decided whether two strains do or do not belong to the same species. Bacterial taxonomy nowadays can be described as "polyphasic," i.e., involving both phenotypic and genomic traits (13, 20). The sections below will discuss the different methods that are currently employed in this polyphasic approach (summarized in Table 3.2), the kind of information that is obtained using each of the common tests, and the relative merits of the different approaches.

3.1. Genotypic Properties Used in Prokaryote Classification

A look at the upper part of Table 3.2 shows that many genotypic and genomic properties have found their way into modern prokaryote taxonomy. Some of these are properties determined by the complete genome – notably the G + C base ratios and the DNA–DNA hybridization methods. Other methods specifically target a special molecule, which often is 16S rRNA (universally applicable; 18S rRNA for eukaryotes) or specific genes that are present in a certain group of *Bacteria* or *Archaea* only and that provide taxonomically valuable information. Examples of the latter are *amoA* (coding for a subunit of the ammonium monooxygenase that catalyzes the first step in ammonium oxidation during autotrophic nitrification), *nifH* (encoding nitrogenase reductase, one of the enzymes essential for nitrogen fixation), and *dsr*, the gene encoding the dissimilatory bisulfite reductase involved in the dissimilatory reduction of sulfate to sulfide of sulfate reducers.

The G + C base ratio varies over a wide range in the prokaryote world, from approximately 20–80%. While an identical G + C content of two species' DNA does not prove the existence of any relationship (after all, the same overall G + C content can be obtained with many different sequences), a large difference (e.g., of 5 mol% or more) in G + C content unequivocally shows that the isolates cannot be closely related. A number of different techniques are currently in use for the experimental determination of the G + C percentage – in addition of course to complete genome sequencing, which will automatically yield the desired information; there are methods based on thermal denaturation profiles, centrifugation methods that assess the buoyant density of the DNA, and HPLC methods that determine the amount of each nucleotide after hydrolysis of the DNA. The determination of the buoyant density is relatively seldom performed nowadays, as analytical ultracentrifuges are operated in only a few laboratories. The method is based on the principle that the higher the G + C content of the DNA, the higher the buoyant density of the DNA is in a CsCl gradient obtained by means of high-velocity centrifugation. More common is the determination of the thermal denaturation profile of the DNA. As the triple hydrogen bond between G–C pairs is stronger than the double hydrogen bond between A–T, the higher the G + C content of the DNA double helix is, the higher the temperature at which the two stands of the DNA will separate, a phenomenon that can be monitored by the increase in the absorbance at 254 nm that accompanies thermal denaturation. The HPLC method, based on quantification of the fragments obtained after enzymatic hydrolysis of the DNA, has gained much popularity since it was first introduced in 1989. Whatever method is used, different types of reference DNA of known G + C content should be included in the tests for calibration, and information about the method used should be provided (see the example in Table 3.3, showing a recently published description of a new species within the genus *Halorubrum* (Domain *Archaea*, Phylum *Euryarchaeota*, Class *Halobacteria*, Order *Halobacterales* (*Halomebiales*), Family *Halobacteriaceae*) (21).

Determination of 16S rRNA gene sequences has become an essential part of any species description as well. When large numbers of strains should be compared, individual sequencing of the 16S rRNA genes of all isolates is often not feasible. In such cases, shortcuts can be introduced such as ribotyping (see Table 3.2), which is an identification method based on the fragmentation pattern when the genomic DNA is cut by specific restriction enzymes, the

Table 3.3
Example of a description of a new Prokaryote species based on polyphasic taxonomy

	Comments
Description of *Halorubrum terrestre* sp. nov.	sp. nov. = species nova = new species
Halorubrum terrestre (ter.res'tre. L. neut. adj. *terrestre* of the soil, from which the strains were isolated)	The etymology of the specific epithet proposed, in this case a Latin adjective
Cells are pleomorphic, flat and disc-shaped, 1.0–1.5 × 1.5–2.5 μm in size. Motile. Gas vacuoles not produced	The morphological properties of the species, as observed microscopically
Colonies are orange-red	Description of special properties such as pigmentation, as observed in colonies on agar plates or in liquid culture
Growth occurs in media that contain 15–30% NaCl, with optimum growth at 25% NaCl. Growth occurs between 28 and 50°C (optimum at 37–45°C) and pH 5–9 (optimum, 7.5)	The physical and chemical conditions required for growth: salt concentration, temperature, pH
Chemo-organotrophic; aerobic; oxidase- and catalase-positive	The mode of metabolism and the relation of the species to molecular oxygen
Acid is produced from glycerol, but not from arabinose, fructose, galactose, glucose, lactose, maltose, sucrose, or trehalose. Nitrate is not reduced to nitrite. Indole is not produced from tryptophan. Voges–Proskauer test is negative. Starch, gelatin, and casein are not hydrolyzed. H_2S is not produced. Arginine dihydrolase, lysine decarboxylase, and ornithine decarboxylase are not produced. The following compounds are not used as sole carbon and energy sources: arabinose, cellobiose, aesculin, fructose, fucose, gluconolactone, glucose, glucosamine, inulin, mannose, melibiose, raffinose, rhamnose, ribose, sucrose, trehalose, xylose, adonitol, dulcitol, erythritol, ethanol, glycerol, mannitol, *meso*-inositol, propanol, sorbitol, α-aminovalerate, butyrate, caprylate, citrate, fumarate, glutamate, glycerate, 2-oxoglutarate, malate, malonate, oxalate, propionate, saccharate, and tartrate. The following compounds are not used as sole carbon, nitrogen or energy sources: L-alanine, L-arginine, L-asparagine, betaine, creatine, L-glutamine, glycine, L-histidine, L-lysine, L-methionine, L-ornithine, L-proline, putrescine, sarcosine, L-serine, L-threonine, and L-valine	Description of the substrates on which the species can grow and those that do not support growth, as well as the result of selected enzymatic tests that provide information on the metabolic abilities of the species
Susceptible to anisomycin, bacitracin, and novobiocin; resistant to ampicillin, chloramphenicol, kanamycin, nalidixic acid, penicillin G, polymyxin, streptomycin, and tetracycline	Information on the sensitivity of the species to different antibiotics and other antibacterial compounds

(*Continued*)

Table 3.3
(Continued)

	Comments
Polar lipids are $C_{20}C_{20}$ derivatives of phosphatidylglycerol, phosphatidylglycerol phosphate methyl ester, phosphatidylglycerol sulfate and a sulfated diglycosyl diether	Chemotaxonomic information on the types of lipids present, with emphasis on those lipids that are diagnostic for different genera within the *Halobacteriaceae*
DNA G + C content is 64.2–64.9 mol% (T_m method)	The range of Guanine + Cytosine content of the genomic DNA of the isolates of the new species, including information on the methodology used for its determination (thermal denaturation)
The type strain is $4p^T$ (=VKM B-1739^T = JCM 10247^T)	The designated type strain (indicated with a superscript capital T) and its accession number in two public culture collections (VKM = All-Russian Collection of Microorganisms, Russian Academy of Sciences, Pushchino, Russia; JCM = Japanese Collection of Microorganisms, Saitama, Japan)
DNA G + C content of this strain is 64.4 mol% (T_m method)	The Guanine + Cytosine content of the genomic DNA of the type strain of the new species, including information on the methodology used for its determination (thermal denaturation)
Isolated from saline soils	The habitat from which the strains investigated were isolated
The GenBank/EMBL/DDBJ accession number for 16S rRNA gene sequence of *Halorubrum terrestre* strain VKM B-1739^T is AB09016	The accession number of the 16S rRNA gene nucleotide sequence of the type strain, as deposited in the public gene sequences databases

fragments separated by electrophoresis, and then hybridized with a probe specific for 16S rRNA genes. Each species has its unique characteristic restriction pattern, and computerized databases exist in which the restriction patterns of many species are stored. The method is much more rapid than full sequencing of the 16S rRNA gene, and is very specific. There are other current DNA fingerprinting methods, listed in Table 3.2, such as randomly amplified polymorphic DNA (RAPD) strain typing (22).

3.2. Phenotypic Properties Used in Prokaryote Classification

A large number of phenotypic properties have been found useful in the description, classification, and identification of bacteria. Table 3.2 shows some of the most important ones that

are generally applicable. There are many other properties that are relevant for specific groups of microorganisms.

Among the phenotypic properties valuable for the description of bacterial isolates and of value in bacterial systematics and taxonomy, one can name:

- Cell shape, cell size, motility, mode of flagellation
- Colonial morphology
- Gram-staining behavior, the ultrastructure, and the chemical structure of the cell wall (presence of peptidoglycan, type of peptidoglycan, especially the nature of the peptides linking the polysaccharide chains), presence of an outer membrane, presence of hopanoids and teichoic acids in the cell wall
- The presence of exopolysaccharides and their structure
- Formation of endospores
- Presence of cellular inclusions such as gas vesicles, storage materials, etc., and other ultrastructural characteristics
- Pigmentation and characterization of the chemical nature of the pigments present (bacteriochlorophylls, chlorophylls, phycobiliproteins, etc. in phototrophic *Bacteria*, carotenoids both in phototrophic and in nonphototrophic prokaryotes, other classes of pigments)
- The nature of the membrane lipids (ether-linked in *Archaea*, ester-linked in *Bacteria*), and the types of fatty acids present (generally determined as fatty acid methyl esters by gas chromatography). Some fatty acids are very valuable as diagnostic markers. Rigorous standardization of the methodology used is essential, and cultures to be compared should be grown under identical conditions, as growth conditions may have a strong influence on the fatty acids pattern
- For respiratory organisms – the chemical nature of the respiratory quinones
- The types of polyamines present in the cell
- Temperature, pH, and salinity range and optimum for growth
- The types of metabolism performed by the cells: aerobic, anaerobic, chemoorganotrophic, phototrophic, chemolithotrophic, etc., with information on the range of energy sources, electron donors, electron acceptors, and carbon sources used. Miniaturized standardized tests, such as those provided by the BIOLOG^{R} system, are often helpful. The latter consists of microtiter plates with wells that contain potential growth substrates and a redox indicator. Utilization of the carbon source will cause a color change of the indicator, enabling the simultaneous determination of the utilization of 95 different substrates within a short time. The procedure can also easily be automatized. Other commercially available test kits (such as, e.g., the API system for identification of enteric bacteria, which consists of miniaturized test tubes with different media and reagents) enable the rapid determination of other physiological properties such as the production of certain enzymes, growth and acid formation on specific carbon sources, and others.
- Special nutritional requirements
- Excretion of exoenzymes (amylases, proteases, lipases, etc.)
- Presence of special diagnostic enzymes
- Susceptibility to a range of antibiotics and other antibacterial substances; susceptibility to attack by specific bacteriophages
- Immunological properties (reactions with specific antibodies, etc.)

It must be stressed that the results of individual phenotypic tests are insufficient to provide information on the identity of bacterial isolates and on the genetic relatedness of strains. However, integration of the results of a large number of phenotypic tests provides reliable descriptive information enabling to recognize prokaryote taxa. This is illustrated in Table 3.3,

which provides an example of a recently published species description that includes a large number of phenotypic tests of diagnostic value. It is clear that description of an isolate as a member of a new species requires extensive documentation of its properties, morphological as well as physiological, 16S rRNA sequence determination, assessment of the base composition of the DNA, determination of DNA–DNA similarity with its closest relatives, and documentation of as many other characteristics as possible.

A special type of taxonomy that is based on the comparison of many bacterial strains for a large number of mainly phenotypic properties is the so-called numerical taxonomy (23), also sometimes referred to as phenetic or Adansonian taxonomy. In this type of taxonomic characterization, developed in the 1950s based on multivariate analysis, the strains are divided into clusters on the basis of large numbers of tests followed by computer-assisted evaluation of the results. These tests should belong to different categories (morphology, physiology, biochemistry, serology, etc.) and should be independent of each other. An important principle of this type of classification is that it is based on equal weighting of each of the characters determined, so that each of the many tests (100–200 tests are ideally included in such studies, and at least 50–60 are required to yield relevant results) has an equal importance in the final outcome, rather than stressing the importance of certain traits. Following data collection of the different tests, the results are coded as positive or negative, and the resemblance between the strains is expressed in a similarity matrix. This matrix is then analyzed for taxonomic structure, and the strains are arranged in groups (so-called phenons, equivalent to species in many cases). On the basis of the delineation of these phenons, the diagnostic characters of each group can be identified, and these can then later be used for the preparation of identification schemes that will enable to place additional strains within the established framework.

3.3. The Polyphasic Approach Toward Prokaryote Classification

As stated in Sect. 2.2, there are no generally agreed-upon rules to delineate genera, except the notion that genera should reflect phylogenetic relationships. Also, the delineation of species is still problematic in the prokaryote world. Nowadays, there is a broad consensus among microbial taxonomists that phylogenetic data are of superior value for the delineation of genera and species, but that "polyphasic" definition of the taxa is required to describe and define taxa at the genus and the species level and to differentiate them from their neighbors (13, 20).

The polyphasic approach to taxonomy uses a combination of a variety of different phenotypic and genetic properties to establish a classification of microorganisms. It is the most obvious strategy to collect a maximum amount of direct and indirect information about the total genome. The term "polyphasic" was introduced in 1970 to describe taxonomy that assembles and accumulates multiple sources of information, based on genetic–phylogenetic as well as on phenotypic data and ecological properties. Nowadays, polyphasic taxonomy refers to a consensus type of taxonomy and aims to utilize all the available data in delineating consensus groups. The more properties are included in the descriptions (see Table 3.2), the more robust and stable the resulting classification schemes will be. Different properties have different resolving power; some are species-specific, while others are valuable to discriminate genera, families, and orders. Descriptions of species using the polyphasic approach should

reflect the phylogenetic relationships of the species with other related taxa, include total genome DNA–DNA hybridization to determine the genomic relationships with related taxa, and also provide further descriptive genomic, phenotypic, and chemotaxonomic information.

4. NAMING OF PROKARYOTES

4.1. The Binomial System of Naming Prokaryotes

Similar to the eukaryotic organisms, the species of prokaryotes have generic names and specific epithets derived from Latin, latinized Greek or latinized modern words or names (24, 25). The binomial system that was introduced by Linnaeus for the plant world in the 1750s is used for the domains *Bacteria* and *Archaea* as well. In many cases, do these names give some information about the properties of the organism such as shape, color, habitat, or physiology. Others have been named in honor of famous microbiologists in the past and present. Three examples of bacteria and their etymology: the name *Streptococcus pneumoniae* is derived from the Greek adjective streptos – στρεπτος, twisted, flexible, the Greek noun kokkos – κοκκος designating a seed, a berry, and the Greek noun pneuma – πνευμα – breath, from where the neo-Latin pneumonia. The nitrifying, nitrite-oxidizing bacterium *Nitrobacter winogradskyi* derives its name from the Latin noun nitrum, nitrate, and bacter, being an equivalent for the Greek noun bakterion – βακτεριον, rod, staff; the species was named in honor of Sergei Winogradsky, who in the 1880s discovered the nitrifying bacteria and formulated the concept of chemolithotrophy. Finally, the name *Acidithiobacillus ferrooxidans* provides considerable information on the physiology of the organism: the Latin adjective acidus means acidic; thion – θιον is Greek for sulfur, bacillus is derived from the Latin diminutive noun bacillum, a small rod, ferrum is the Latin form of iron, and the neo-Latin verb oxidare means to oxidize. The name, thus, signifies that the organism is a rod-shaped bacterium that lives in acidic environments, is involved in sulfur transformation, and oxidizes iron compounds.

Each description of a new species (see also Table 3.3) should include a proposal for its name and explain the etymology of that name. The naming of bacteria is subject to many rules and recommendations, both derived from linguistic constraints and from scientific considerations. Practical recommendations on how to name a new prokaryote can be found in a number of treatises on the subject (24, 25).

The phylum *Cyanobacteria* presents special nomenclature problems, as the group is also included by the botanists under the rules of the Botanical Code as *Cyanophyta* or blue-green algae. This group consists of organisms with a prokaryotic cell structure that display oxygenic photosynthesis, i.e., physiologically they resemble the eukaryotic algae and the higher plants. The rules of botanical nomenclature are very different from those of the Bacteriological Code, as the botanical types are not axenic live cultures such as are required under the Bacteriological Code, but descriptions and material preserved in herbaria. The result is a highly confusing and unsatisfactory situation in which many "species" appear under different names in the literature. The 1980 approved lists of bacterial names (19) did not contain any names of *Cyanobacteria*, and only very few names of species have since been validly published in the *International Journal of Systematic Bacteriology/International Journal of Systematic and*

Evolutionary Microbiology. The latest edition of Bergey's Manual (5) (see also Sect. 7.1) does not divide the phylum into classes, orders, and families, but instead provides a provisional division of the group into subsections, each subsection consisting of several "form-genera." It is to be expected that the nomenclature problems of this group will not very soon find a solution that will satisfy bacteriologists and botanists alike.

4.2. The Bacteriological Code

The nomenclature of the prokaryotes, *Bacteria* as well as *Archaea*, is regulated by the rules of the Bacteriological Code – The International Code of Nomenclature of Bacteria, as approved at the Ninth International Congress of Microbiology, Moscow, 1966, revised in 1990 (4), and amended at the meetings of the International Committee for Systematics of Prokaryotes (before 2000: International Committee of Systematic Bacteriology (see Sect. 4.3). A new revised version of the Code is presently in preparation – to be renamed International Code of Nomenclature of Prokaryotes since 2000). The Bacteriological Code presents the formal framework according to which prokaryotes are named and according to which existing names can be changed or rejected. It covers the rules for the naming of species (and subspecies), genera, families, and orders of prokaryotes. No provisions are made by the Code for the naming of the higher taxa: class, phylum, and domain.

One aspect in which the Bacteriological Code differs from the Botanical Code and the Zoological Code, the latter two being similar documents that regulate nomenclature in the plant and the animal world, is the rule (Rule no. 24a) according to which a new start was made in prokaryote nomenclature with the publication of the 1980 approved lists of names (19), therewith abolishing the need to search the older, and often very confusing literature (see also Sect. 2.3). In contrast, botanists who want to establish whether a certain plant has been recorded before in the literature often have to search the scientific journals and books as far back as 1753, the year in which Linnaeus published his "Species Plantarum." Zoologists face a similar problem when they must decide whether an animal they discovered may be described as a new species.

4.3. The International Committee on Systematics of Prokaryotes

The rules that regulate the nomenclature of prokaryotes, as published in the Bacteriological Code, are set by the International Committee on Systematics of Prokaryotes (ICSP) (before 2000: the International Committee of Systematic Bacteriology). This committee is a constituent part of the International Union of Microbiological Societies (IUMS). The committee meets at least once every 3 years – at the time of the IUMS congresses. The committee discusses nomenclatural problems that have arisen in different groups of prokaryotes and proposes changes and amendments to the rules of the Bacteriological Code. The minutes of the committee's meetings are published in the *International Journal of Systematic and Evolutionary Microbiology* (see Sect. 4.4). Information on the committee, its members, and its different subcommittees can be found in its Web site: http://www.the-icsp.org.

The ICSP has established several subcommittees. One important subcommittee is the Judicial Commission, a committee that deals with problematic cases in bacterial nomenclature and renders judicial decisions in instances of controversy about the validity of a name, identity

of type strains, and cases of emerging problems with the interpretation of the rules of the Bacteriological Code. It may propose amendments to the Code and consider exceptions that may be needed to certain rules. The decisions of the Judicial Commission need to obtain approval of the ICSP to obtain standing. Problems can be brought to the attention of the Judicial Commission by submission of a "Request for an Opinion" to be published in the *International Journal of Systematic and Evolutionary Microbiology*.

The ICSP has established a number of taxonomic subcommittees that discuss nomenclatural problems of specific groups of prokaryotes. Currently, there are 28 such subcommittees, dealing with groups such as *Bacillus* and related organisms, phototrophic bacteria, methanogenic *Archaea*, Gram-negative anaerobic rods, *Bifidobacterium*, *Lactobacillus* and related organisms, the genus *Mycobacterium*, staphylococci and streptococci, and others. Details about the membership and about the taxa covered by these subcommittees can be found in the web site of the ICSP. The minutes of the meetings of these subcommittees are published as well in the *International Journal of Systematic and Evolutionary Microbiology*.

4.4. The International Journal of Systematic and Evolutionary Microbiology

The *International Journal of Systematic and Evolutionary Microbiology* (prior to 2000: *International Journal of Systematic Bacteriology*) is the journal in which new names of taxa of prokaryotes must be published in order to obtain standing in the nomenclature; since 1976 validation is only possible by publication in this journal. It is also possible to describe the new species or other taxa in another scientific journal ("effective publication"), but the new name will not obtain standing in the nomenclature until it has been included in the "Validation Lists" of names first published elsewhere. Such validation lists are included regularly in the issues of the journal, which presently appears monthly.

In addition to being the platform for describing new prokaryote taxa, the *International Journal of Systematic and Evolutionary Microbiology* publishes articles that address taxonomy, phylogeny and evolution of prokaryotes (as well as of fungi and some other groups of eukaryotic protists). In addition, it contains the minutes of the meetings of the International Committee of Systematics of Prokaryotes, its Judicial Commission and its taxonomic subcommittees, as well as correspondence relating to bacterial nomenclature such as "Request for an Opinion" documents with queries to the Judicial Commission (see Sect. 4.3).

4.5. Information on Nomenclature of Prokaryotes on the Internet

The web site www.bacterio.cict.fr, maintained by Dr. Jean Euzéby of the University of Toulouse, France (1), contains a wealth of information on all names of prokaryotes that have standing in the nomenclature. The site is updated monthly with the publication of the latest issue of the *International Journal of Systematic and Evolutionary Microbiology*. It provides important information on the type strains of each species and on any name changes and current or past nomenclature problems. In addition, the site contains a great deal of additional information relevant to prokaryote nomenclature. For example, the information on the total number of species with standing in the nomenclature as given in Table 3.1 was derived from this web site.

In addition, much information on specific groups of microorganisms can be found in the Web pages of several of the International Committee on Systematics of Prokaryotes taxonomic subcommittees (www.the-icsp.org; see also Sect. 4.3).

5. CULTURE COLLECTIONS OF PROKARYOTES AND THEIR IMPORTANCE IN TAXONOMY AND IDENTIFICATION

As explained in Sect. 2.1, the type strain of each newly described species must be deposited in at least two publicly accessible culture collections located in different countries, so that the strain will be preserved and made available for further study by any interested scientist. Such culture collections are extremely important for the preservation of microbial biodiversity, and their importance for microbial taxonomy cannot be overestimated when it comes, for example, to referencing strains that should be used when comparing new isolates (26). Culture collections generally preserve bacterial strains either in dry, lyophilized form, to be revived by wetting and suspension in suitable growth medium, or frozen in liquid nitrogen. In addition to storing and distributing publicly available strains, culture collections may provide safekeeping facilities for patented strains of microorganisms. Many also provide characterization and identification facilities.

A list of culture collections that maintain cultures of prokaryotes can be found in the www.bacterio.cict.fr Web site (see Sect. 4.5), and a representative list is found in Vol. 1 of the second edition of *Bergey's Manual of Systematic Bacteriology* (26) (see also Sect. 7.1). Some of these culture collections maintain Web sites that, in addition to the strain catalogs and technical details about depositing and ordering strains, provide extensive information about the history and the nomenclature of the strains, recipes for media in which the isolates can be grown, and much more. The Web sites of the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Braunschweig, Germany – www.dsmz.de) and the American Type Culture Collection (ATCC, Manassas, VA, USA – www.atcc.org) are especially useful in this respect.

6. SMALL-SUBUNIT rRNA-BASED CLASSIFICATION OF PROKARYOTES

Before the late 1970s, no methods were available that would enable linking bacterial systematics and taxonomy with bacterial phylogeny. Thereafter, molecular methods have been introduced, based on sequencing of genes that can be used as "evolutionary clocks," i.e., that provide information on evolution in the prokaryote world. The principle of molecular taxonomy and of phylogenetic tree reconstruction is based on the concept that biological macromolecules can be used as evolutionary chronometers that measure evolutionary change. Mutations that have occurred during the course of time have become fixed in the populations, resulting in diversity in sequences of nucleotides in genes and of amino acids in proteins. Evolutionary distances can thus be measured by differences in nucleotide or amino acid sequence of monomers in homologous macromolecules. As explained in Sect. 3.3, one of the goals of polyphasic taxonomy is that the classification of the genera should reflect phylogenetic relationships. Thanks to the availability of the appropriate techniques, notably the sequencing

of the genes encoding the small-subunit (16S) ribosomal RNA, the fulfillment of this goal is now within reach.

6.1. 16S rRNA as a Phylogenetic Marker

When Carl Woese started to compare nucleotide sequences of small-subunit ribosomal RNA (16S rRNA in prokaryotes, 18S rRNA in eukaryotes) at the University of Illinois in the late 1970s, few people realized that the approach used and the results obtained would within one to two decades revolutionize our views not only about prokaryote evolution, but also about classification of prokaryotes (27). It was to Woese's credit that he realized early on that the ribosomal RNA molecules are the best suited to serve as molecular chronometers to track down the course of prokaryote phylogeny (14–16). As each cell has ribosomes, ribosomal RNA is universally distributed and functionally homologous, and the rate of change in these molecules has proven sufficiently slow to be useful to reconstruct the phylogeny of organisms that exist on Earth for three and a half billion years at least. Common ancestry, genetic stability, appropriate size, and the presence of independently evolving domains within the molecules are other properties that make the 16S and 18S rRNAs ideally suited to serve as phylogenetic markers.

Prokaryote ribosomes contain three molecules of RNA, 5S rRNA, 16S rRNA, and 23S rRNA, with about 120, 1,540, and 2,900 nucleotides, respectively. The latter two are large enough to contain sufficient information for the purpose of phylogenetic tree reconstruction. Technically, the sequencing of the 16S rRNA (or the gene that codes for it) is easier than sequencing of the much longer 23S rRNA. The results obtained from analysis of the 23S rRNA gene sequencing generally confirm those obtained with the 16S rRNA gene. Accordingly, 16S rRNA has become the molecule of choice for general use in phylogenetic studies of prokaryotes.

Determination of phylogenetic relationships, based on 16S and to a lesser extent on 23S rRNA sequence similarities, has become routine procedure in bacterial taxonomy. Sequencing of the genes generally follows their prior amplification by PCR. No characterization of new species of prokaryotes is nowadays complete without presentation of the 16S rRNA gene sequence and deposition of this sequence in a public database, such as the GenBank, to make it available to the scientific world (see also Table 3.2). There is also a specialized database, the Ribosomal Data Base Project (http://rdp.msu.edu), which presently (November 10, 2009) contains 1,235,044 16S rRNA gene sequences, both of cultured organisms and of sequences recovered from DNA isolated from the mixed community present in the environment (see Sect. 9). New sequences can be compared by aligning them with those present in the database, and phylogenetic trees can then be computed using statistical methods. Figure 3.1 shows an example of the phylogenetic tree of the prokaryotes obtained on the basis of 16S rRNA gene sequence comparisons (based on sequences derived from cultured species only; a few phyla of *Bacteria* that are currently represented by only very few cultured species are not shown), and Table 3.4 summarizes the properties of the most important phylogenetic groups that have emerged from these comparisons. It should be noted that the exact topology of such trees may to some extent depend on the computational procedure followed. A discussion of the algorithms used for the calculation of the tree topologies is outside the scope of this chapter.

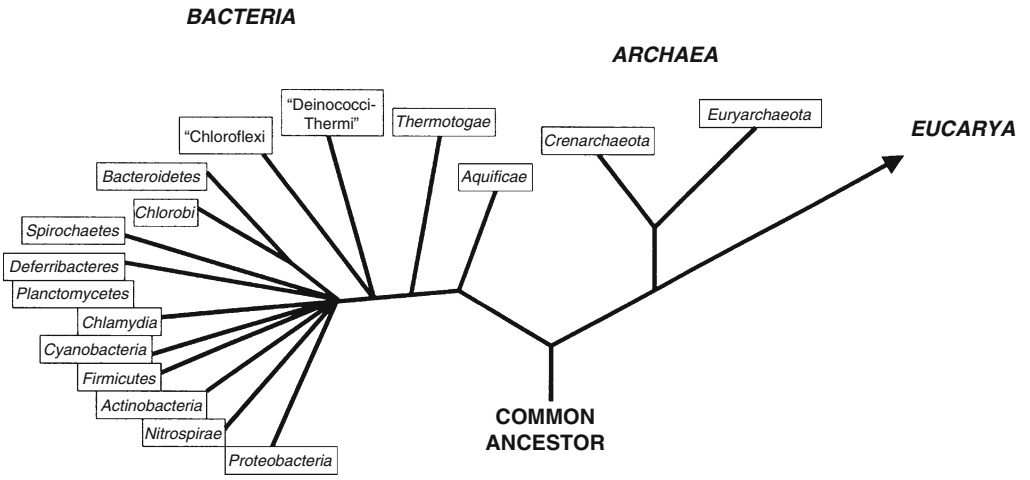

Fig. 3.1. Schematic rooted phylogenetic tree, based on small subunit ribosomal RNA comparisons, showing the most important phyla of *Bacteria* and *Archaea*.

Overall, it can be stated that the resulting phylogenetic trees are rather robust constructs. Their topology is largely confirmed by comparative analyses of other conserved molecules (elongation factors, the β-subunit of ATP synthase, and other proteins).

The existence of this extensive and ever-growing database enables us to rapidly place any newly isolated prokaryote within the phylogenetic classification scheme at least to the level of the genus. As stressed earlier (see Sect. 2.2), the resolving power of 16S rRNA gene as a taxonomic marker is insufficient to allow identification to the species level.

One of the most important concepts that emerged from 16S rRNA and 18S rRNA sequence comparisons, already from the very beginning of these studies in Carl Woese's laboratory in the 1970s, is that the prokaryote world is not phylogenetically homogeneous. On the basis of the small subunit ribosomal RNA sequences, the fundamental division of the forms of life that inhabit planet Earth should be not into two groups (prokaryotes and eukaryotes), as was customary at the time, but into three groups, the eukaryotes and two groups ("primary kingdoms," now called domains) of prokaryotes, which among each other show as little phylogenetic relationship as each of these domains with the eukaryotes. Woese originally named these prokaryotes, *Eubacteria* and *Archaebacteria*. The names currently used are *Bacteria* and *Archaea* for the prokaryotic domains, and *Eucarya* for the eukaryotic domain (see also Fig. 3.1) (14–16). The following sections explore the differences between the *Bacteria* and the *Archaea*, and provide an overview of the properties of the different groups (at least at the level of phyla) within each of these two basic domains of prokaryote life. Phylogenetic analyses based on conserved protein sequences generally support a closer relationship between the *Archaea* and the *Eucarya* than between the *Archaea* and the *Bacteria*, such as is also suggested by the tree topology shown in Fig. 3.1. The concept that the prokaryotes are phylogenetically heterogeneous was only slowly accepted by the scientific world. However, nowadays, the basic division of life into three domains, separating the *Archaea* from the *Bacteria* is now accepted

Table 3.4
The major divisions of the Prokaryote world – the domains, phyla, and classes as given in the classification proposed in the second edition of Bergey's manual of systematic bacteriology (5)

Domain	Phylum	Class	Main representatives and further information
Archaea			
	Crenarchaeota		
		Thermoprotei	All cultured representatives are extremely thermophilic, mostly anaerobic isolates from hydrothermal vents, hot springs, and other thermal environments. Some well-known species: *Pyrodictium*, *Sulfolobus*
	Euryarchaeota		
		Methanobacteria	Strictly anaerobic methanogenic prokaryotes (e.g., *Methanobacterium*, *Methanobrevibacter*), including a few thermophiles (*Methanothermus*)
		Methanococci	Strictly anaerobic methanogenic prokaryotes (e.g., *Methanococcus, Methanospirillum*, *Methanosarcina*), including a few thermophiles (*Methanocaldococcus*)
		Halobacteria	Extremely halophilic aerobes or facultative anaerobes such as *Halobacterium* and *Haloarcula*; some are alkaliphilic (e.g., *Natronobacterium*)
		Thermoplasmata	Thermophilic organisms such as *Thermoplasma*, and thermoacidophiles (*Picrophilus*)
		Archaeoglobi	Anaerobic thermophiles with chemoautotrophic or heterotrophic metabolism, using sulfate (*Archaeoglobus*) or nitrate (*Ferroglobus*) as electron acceptor
		Methanopyri	A separate lineage of thermophilic methanogens (*Methanopyrus*)
Bacteria			
	Aquificae		
		Aquificae	Moderately thermophilic to hyperthermophilic chemolithotrophs and chemoorganotrophs (*Aquifex*, *Hydrogenobacter*)
	Thermotogae		
		Thermotogae	Extreme thermophilic anaerobes with an outer sheath-like envelope (*Thermotoga*, *Geotoga*)

(Continued)

Table 3.4
(Continued)

Domain	Phylum	Class	Main representatives and further information
	Thermodesulfo-bacteria		
		Thermodesulfo-bacteria	Currently represented by a single, deep-branching genus *Thermodesulfobacterium*, of sulfate-reducing anaerobes
	"Deinococcus-Thermus"		
		Deinococci	A group of chemoorganotrophic aerobes that includes *Deinococcus*, a genus of radiation-resistant cocci, and *Thermus*, a genus of thermophilic bacteria from hot springs
	Chrysiogenetes		
		Chrysiogenetes	A deep lineage within the *Bacteria*, currently represented by a single genus *Chrysiogenes* with a single species that grows anaerobically by oxidizing acetate with arsenate as electron acceptor
	Chloroflexi		
		"Chloroflexi"	Gram-negative bacteria showing gliding motility. Some (*Chloroflexus*) are anoxygenic phototrophs, others (*Herpetosiphon*) are aerobic heterotrophs
	Thermomicrobia		
		Thermomicrobia	A phylum currently represented by a single genus *Thermomicrobium* with a single species of hydrocarbon-utilizing thermophiles
	Nitrospirae		
		"Nitrospira"	A heterogeneous group that includes nitrite-oxidizing autotrophs (*Nitrospira*), iron-oxidizers (*Leptospirillum*), and thermophilic sulfate reducers (*Thermodesulfovibrio*)
	Deferribacteres		
		Deferribacteres	Heterotrophic anaerobes (*Deferribacter*, *Geovibrio*), which use oxidized iron, manganese, or nitrate as electron acceptor

Cyanobacteria		
	Cyanobacteria	Aerobic oxygenic phototrophs, unicellular as well as filamentous, including many types that fix molecular nitrogen. Different classification schemes exist, mostly based on the Botanical Code. Well-known species include the unicellular *Synechococcus* and *Microcystis* and filamentous types such as *Oscillatoria, Phormidium, Anabaena*, and *Nostoc*. Few species have yet been named under the Bacteriological Code
Chlorobi		
	"Chlorobia"	Anaerobic oxygenic phototrophs that oxidize reduced sulfur compounds (*Chlorobium*, *Prosthecochloris*, and others)
Proteobacteria		
	"Alphaproteo-bacteria"	A large and very heterogeneous group of Gram-negative bacteria, including anoxygenic photoheterotrophs (*Rhodospirillum*, *Rhodobacter*), aerobic heterotrophs (*Paracoccus*, *Caulobacter*, *Acetobacter*), parasites/pathogens (*Rickettsia*), symbionts (*Rhizobium*), chemoautotrophs (*Nitrobacter*), and many others
	"Betaproteo-bacteria"	Another large and very heterogeneous group of Gram-negative bacteria, which includes heterotrophs (*Spirillum*, *Zoogloea*) chemoautotrophs (*Thiobacillus*, *Nitrosomonas*), and anoxygenic photoheterotrophs (*Rhodoferax*)
	"Gammaproteo-bacteria"	The third subgroup of the Proteobacteria, no less heterogeneous than the first two. It has anoxygenic photoautotrophic representatives (e.g., *Chromatium*), chemolithotrophic ammonium oxidizers (e.g., *Nitrosococcus*), chemolithotrophic sulfur oxidizers (*Beggiatoa*), heterotrophic aerobes (*Pseudomonas*, *Azotobacter*, *Vibrio*), fermentative anaerobes (*Ruminobacter*), the enteric bacteria (*Escherichia*, *Salmonella*, *Proteus*), and many others
	"Deltaproteo-bacteria"	Another physiologically heterogeneous group that includes most dissimilatory sulfate-reducing bacteria (*Desulfovibrio*, *Desulfobacter*), certain nitrifying chemoautotrophs (*Nitrospina*), iron-reducing bacteria (*Geobacter*), anaerobic obligatory syntrophic bacteria (*Syntrophobacter*), predatory bacteria (*Bdellovibrio*), and others

(Continued)

Table 3.4
(Continued)

Domain	Phylum	Class	Main representatives and further information
		"Epsilon-proteobacteria"	A small subgroup of the *Proteobacteria* that contains a number of interesting organisms such as *Helicobacter*, the causative agent of gastric ulcers, *Thiovulum*, a sulfur-oxidizing bacterium, and *Campylobacter*, a genus of intestinal bacteria
	Firmicutes		
		"Clostridia"	Anaerobic fermentative Gram-positive bacteria, some forming endospores, including genera such as *Clostridium*, *Ruminococcus*, *Peptostreptococcus*, endospore-forming dissimilatory sulfate reducers of the genus *Desulfotomaculum*, and anoxygenic phototrophs (*Heliobacterium*)
		Mollicutes	A group of mostly wall-less pathogens such as *Mycoplasma, Acholeplasma*, and *Erysipelothrix*
		"Bacilli"	This group of Gram-positive bacteria includes the aerobic endospore-forming organisms (*Bacillus* and related species), and the lactic acid bacteria (*Lactobacillus*, *Leuconostoc*, *Streptococcus*)
	Actinobacteria		
		Actinobacteria	A large group of mostly aerobic heterotrophic microorganisms, many of them growing in the form of mycelia (*Steptomyces*, *Actinomyces*), pathogens such as *Corynebacterium* and *Mycobacterium*, fermentative anaerobes (*Propionibacterium*), symbiotic nitrogen fixers (*Frankia*), and others
	Planctomycetes		
		"Planctomycetacia"	A small group of mostly aerobic heterotrophic bacteria of unusual cellular structure (*Plancomyces*, *Pirellula*). It includes the anaerobic chemoautotrophic organisms that oxidize ammonium ions with nitrite as electron acceptor
	Chlamydiae		
		"Chlamydiae"	A group of intracellular parasites and pathogens (*Chlamydia*)
	Spirochaetes		
		"Spirochaetes"	Aerobic and anaerobic organisms with an unusual mode of motility. The group includes free-living species (*Spirochaeta*), symbionts (*Cristispira*), as well as pathogens (*Treponema*, *Borrelia*)

Fibrobacteres		
	"Fibrobacteres"	A deep lineage within the *Bacteria*, presently consisting of a single genus, *Fibrobacter* – anaerobic fermentative organisms found in the rumen of ruminant animals
Acidobacteria		
	"Acidobacteria"	A small group of acidophilic organotrophs (*Acidobacterium*) and iron reducers (*Geothrix*)
Bacteroidetes		
	"Bacteroides"	Anaerobic fermentative bacteria (*Bacteroides* and others)
	"Flavobacteria"	Generally aerobic, often pigmented heterotrophs such as *Flavobacterium*
	"Sphingobacteria"	Heterotrophic aerobic bacteria involved in aerobic degradation, often of polymeric substrates (*Cytophaga*, *Flexibacter*, and the thermophilic *Rhodothermus*)
Fusobacteria		
	"Fusobacteria"	A small group of fermentative anaerobes of different metabolic types (*Fusobacterium*, *Propionigenium*)
Verrucomicrobia		
	Verrucomicrobiae	A small group of prosthecate bacteria, including *Verrucomicrobium* and *Prosthecobacter*, forming a deep linage within the phylogenetic tree of the *Bacteria*
Dictyoglomi		
	"Dictyoglomi"	A linage of *Bacteria* thus far represented only by *Dictyoglomus*, an anaerobic thermophilic organotroph

Names of taxa in italic type have been formally described in the literature and have standing in the bacteriological nomenclature. It should be noted that the terms "Domain", "Phylum," and "Class" are not covered by the Bacteriological Code (4). Names printed in roman type in quotation marks are newly proposed names yet without standing in the nomenclature.

by nearly all bacteriologists. Alternative views are, however, proposed from time to time (28). A discussion of these alternative models of prokaryote evolution is outside the scope of this chapter.

Phylogenetic trees, such as the one shown in Fig. 3.1, do not provide an answer to the question of the nature of the universal ancestor of all life forms on Earth. One can also not infer from these trees whether the *Archaea* are a more ancient group than the *Bacteria*, or whether the prokaryotes formed an ancestral stage that led to the development of the primitive eukaryotic cell. It is, however, well established that the mitochondria found in most eukaryotic cells have an ancestry that can be traced back to the *Proteobacteria*, one of the phyla within the domain *Bacteria*. Similarly, the origin of the chloroplasts, the organelles responsible for light harvesting, and autotrophic fixation of carbon dioxide in algae and higher plants, is within the *Cyanobacteria* phylum of the *Bacteria*. These organelles are the size of prokaryotes, contain their own DNA, and have ribosomes that resemble in size and properties those of the *Bacteria* rather than those of the eukaryotic cell. Genome sequence analyses of representatives of the three domains have made it clear that phylogenetically speaking, the eukaryotic cell is thus a chimera, which includes components derived from the *Eucarya*, the *Bacteria*, and also the *Archaea*.

6.2. The Differences Between Bacteria and Archaea

If indeed the *Archaea* are phylogenetically so distant from the *Bacteria* as it appears from the 16S rRNA gene sequence-based trees (Fig. 3.1), it should be expected that this great evolutionary distance would also find its expression in a large number of other properties, including phenotypic ones. Morphologically, the *Archaea* and the *Bacteria* are very similar, and also at the level of the cell ultrastructure, there are no obvious differences between typical representatives of both groups. Because of this apparent morphological and structural similarity of the members of the two domains, the three-domain model of Woese met with much skepticism in the first years. However, when more in-depth comparative studies were made of the representatives of the two domains, it became clear that indeed there are far-reaching differences between *Bacteria* and *Archaea*, not only in the nucleotide sequences of ribosomal RNAs, but also in many fundamental properties of the cell, including the structure of the cell wall, the type of lipids in the membrane, the properties of the transcription mechanism of DNA to form RNA, the details of operation of the protein synthesis machinery of the ribosome, sensitivity to different antibiotics, and others. Table 3.5 summarizes the most important of those differences that define the two prokaryotic domains.

One of the most prominent differences between the members of the archaeal and the bacterial domains is the structure of the cell wall. With a few exceptions, the cell wall of the *Bacteria* contains peptidoglycan, either in a thin layer and accompanied by other cell wall and outer membrane layers in the Gram-negative members, or a thick layer in the Gram-positives. Peptidoglycan is altogether absent in the archaeal domain. Accordingly, the archaeal cell wall is not lysed by the enzyme lysozyme, and none of the *Archaea* is susceptible to penicillin and other β-lactam antibiotics that inhibit the cross-linking of the polysaccharide chains by peptide chains in peptidoglycan. Another striking difference between the domains is the structure of the membrane lipids. The lipids of the *Bacteria* closely resemble the lipids

Table 3.5
The major differences between the *Archaea* and the *Bacteria* (29)

Property	*Archaea*	*Bacteria*	Comments
Presence of peptidoglycan in the cell wall	Absent	Present in most	Peptidoglycan is absent in the eukaryote world as well
Sensitivity to penicillin and other [β]-lactam antibiotics	Resistant	Mostly sensitive	β-Lactam antibiotics inhibit the cross-linking of the polysaccharide chains during the formation of peptidoglycan, a component absent in the *Archaea*
Nature of the membrane lipids	Diphytanyl (C_{20}) glycerol diethers or dibiphytanyl (C_{40}) diglycerol tetraethers (branched long-chain alcohols – phytanols – ether linked to glycerol)	Glycerol esters of aliphatic fatty acids (generally straight-chain, C_{16} and C_{18}-dominated)	The *Bacteria* have the same types of lipids as the eukaryotic micro- and macroorganisms. This type of lipids is missing altogether in the *Archaea*. Ether-linked fatty acids are only rarely found in the domain *Bacteria*
Antibiotic sensitivity of the protein-synthesizing machinery of the ribosome	Chloramphenicol- and kanamycin-resistant, often anisomycin-sensitive	Chloramphenicol- and kanamycin-sensitive, anisomycin resistant	The eucaryal protein synthesis resembles that of the *Archaea* in the sensitivity to the named antibiotics. However, *Archaea* are not inhibited by cycloheximide, a potent inhibitor of protein synthesis by the eukaryote ribosome

(Continued)

Table 3.5
(Continued)

Property	*Archaea*	*Bacteria*	Comments
First amino acid to initiate a polypeptide chain during protein synthesis	Methionine	*N*-formylmethionine	Methionine
Sensitivity of the ADP-ribosylation of the peptide elongation factor EF-2 to diphtheria toxin	Sensitive	Insensitive	The eukaryotic ribosome is susceptible to inhibition by diphtheria toxin as well
Structure of the promoter for initiation of RNA synthesis	TATA box	−10 and −35 sequences (Pribnow box)	The eukaryotic promoters contain the TATA box, similar to the *Archaea*
Properties of the DNA-dependent RNA polymerases	Multicomponent enzymes, containing 8–12 polypeptides; not inhibited by rifampicin and streptolydigin	$a_2\beta\beta'\sigma$ type; inhibited by rifampicin and streptolydigin	In the eukaryotes, multicomponent enzymes (12–14 polypeptides) are found
Special properties of the tRNAs	The "common arm" of the tRNA usually contains ribothymidine	The "common arm" of the tRNA usually contains pseudouridine or 1-methylpseudouridine	

of the *Eucarya*, with generally straight-chain aliphatic fatty acids bound to glycerol by ester bonds. The *Archaea*, on the other hand, contain lipids in which the chemical bond between the glycerol moieties and the hydrophobic chains is an ether instead of an ester bond. Moreover, the hydrophobic chains are not straight-chain 16- or 18-carbon fatty acids, but isoprenoid branched chains with generally 20 carbons. Many *Archaea*, notably the hyperthermophilic species, have instead of a lipid bilayer a lipid monolayer membrane in which two glycerol moieties are linked by 40-carbon isoprenoid (biphytanyl) chains, providing a highly stable membrane with covalent bonds spanning over its whole width.

There are additional differences between *Archaea* and *Bacteria*, as listed in Table 3.5. Some of these are connected with the protein synthesis machinery, and they result in different sensitivities to antibiotics. Others are located in the mechanism of transcription.

6.3. An Overview of the Bacteria

The classification scheme of the prokaryotes into two domains, *Bacteria* and *Archaea*, as given in the latest edition of Bergey's Manual (5) (see Sect. 7.1) can be conveniently used as a framework to provide an overview of the different groups of prokaryotic organisms (Table 3.5). It must be again stressed more that this classification should not be considered as an "official" classification of the prokaryotes, as such an official classification does not exist (see Sect. 2.1).

Volume 1 of the latest edition of Bergey's Manual (2001) divides the domain *Bacteria* into 23 phyla and 31 classes. Some of these phyla contain as yet a few species only, which have obtained their special status on the basis of their highly divergent 16S rRNA gene sequences. Other phyla contain many hundreds of species – examples are the *Proteobacteria* and the *Actinobacteria*. Some phyla and classes consist of physiologically and/or morphologically similar groups of microorganisms. Well-known examples are the *Thermotogae* and the *Aquificae*, which consist entirely of thermophiles, the *Chlorobi*, which are all anoxygenic phototrophic prokaryotes, the *Spirochaetes*, spiral-shaped cells with a characteristic mode of motility due to the unusual way the flagella are inserted, and the *Cyanobacteria*, which are all oxygenic phototrophs. It cannot be excluded that these phyla may prove more diverse when more representatives of the groups will be isolated in the future. Other phyla and classes are very heterogeneous from the aspect of the physiology of their members. Most of the five classes of the phylum *Proteobacteria* contain obligatory and facultative aerobic, obligatory anaerobic, photoautotrophic, photoheterotrophic, and chemolithotrophic organisms of highly diverse morphology, range of substrates uses as carbon and energy sources, etc. The only reason why these had been brought together in one phylum or class is the similarity in 16S rRNA gene sequence, on which this particular classification scheme is heavily based. Also, the phyla, *Firmicutes* and *Bacteroidetes,* contain species that are differ greatly in their physiological and other phenotypic properties.

Most of the Gram-positive bacteria are clustered in the phylum *Firmicutes*, but this phylum also contains species that show a negative Gram stain reaction, such as the wall-less *Mollicutes* (*Mycoplasma* and relatives). Gram-negative organisms are found all over the phylogenetic tree of the *Bacteria*. The Gram stain is thus of little value to assess the position of a bacterial isolate in the phylogenetically based classification scheme of Bergey's Manual. Possession of

chlorophyll or bacteriochlorophyll derivatives and a phototrophic life style alone is also insufficient to place any isolate in its proper place in the system. Chlorophyll-based photosynthesis has not yet been encountered in the domain *Archaea*; however, the property is widespread in the bacterial domain, no less than five of the phyla contain phototrophs. Chlorophyll *a*-based oxygenic photosynthesis is found in the *Cyanobacteria*, and bacteriochlorophyll-based anoxygenic photosynthesis is found in all known representatives of the *Chlorobi*, in most species of the *Chloroflexi*, in many representatives of the *Proteobacteria* (within the classes "Alphaproteobacteria," the "Betaproteobacteria," and the "Gammaproteobacteria" – classes that all contain very heterogeneous assemblages of species as far as their physiology is concerned), and in a few genera of the family "Heliobacteriaceae" within the phylum *Firmicutes*.

6.4. An Overview of the Archaea

The Bergey's Manual classification scheme (5), the domain *Archaea* is subdivided into two phyla – the *Crenarchaeota* with one class, and the *Euryarchaeota* with seven classes. With the exception of most methanogens (*Euryarchaeota*), all cultured representatives of the domain *Archaea* are extremophiles that inhabit environments with extremely high temperatures (these include many methanogens as well), often combined with growth at low pH, or environments characterized by a high salt concentration, in many cases combined with high pH values. Physiologically, the group is very heterogeneous. One property that is notably absent from the archaeal domain is that of chlorophyll-based photosynthesis, a feature so widespread in the bacterial domain. Use of light energy is, however, possible in some extremely halophilic representatives of the family *Halobacteriaceae* (*Euryarchaeota*), based on light absorption by bacteriorhodopsin with the formation of a transmembrane proton gradient. Until recently, this type of light utilization was even considered unique to the archaeal domain. However, with the discovery of proteorhodopsin – a similar pigment that is found in the membrane of certain (yet to be cultured) representatives of the *Proteobacteria* – this claim can no longer be maintained.

All cultured members of the *Crenarchaeota* (class *Thermoprotei*, orders *Thermoproteales*, *Sulfolobales*, and *Desulfurococcales*) are thermophilic. The group contains aerobes as well as anaerobes. Many representatives obtain their energy by oxidizing hydrogen or organic compounds while using elemental sulfur as terminal electron acceptor in respiration; others reduce sulfur compounds. Some are chemoautotrophs; others require organic carbon sources.

In recent years, it has become clear that the *Crenarchaeota* phylum contains nonextremophilic representatives as well, and these appear to be widespread. Sequencing of 16S rRNA genes isolated from marine bacterioplankton showed that a substantial fraction of the prokaryotic community in the open sea consists of *Archaea*, *Crenarchaeota* as well as *Euryarchaeota* (30). A marine crenarchaeote named ("*Candidatus* Nitrosopumilus maritimus") has recently been isolated; it is an autotrophic ammonia-oxidizing organism. Little can yet be said about the physiology and the function of other groups of Archaea in the marine environment.

In the phylum *Euryarchaeota,* a greater phenotypical diversity is encountered than in the *Crenarchaeota*. Three out of the seven classes presently recognized (the *Methanobacteria*, the *Methanococci*, and the *Methanopyri*) consist of methanogenic anaerobes. The *Halobacteria*

generally have an aerobic life style, and their habitat is restricted to hypersaline environments, typically from 150 g/L salt up to saturation. The three remaining classes (*Thermoplasmata*, *Thermococci*, and *Archaeoglobi*) are all thermophilic, and most live as anaerobic or facultative aerobic heterotrophs. The *Archaeoglobi* are heterotrophs or chemoautotrophs, and perform anaerobic respiration with sulfate or nitrate as terminal electron acceptors. As stated earlier, the open sea contains large communities of *Euryarchaeota* as well. These have not yet been brought into culture, and their mode of metabolism is as yet unknown.

7. SOURCES OF INFORMATION ON PROKARYOTE SYSTEMATICS

Except for the general textbooks of microbiology (29), there are two major sources of valuable information on classification of prokaryotes and on the properties of each group: *Bergey's Manual of Systematic Bacteriology* and *The Prokaryotes*.

7.1. Bergey's Manual of Systematic Bacteriology

Bergey's Manual of Systematic Bacteriology (5) and its predecessor, *Bergey's Manual of Determinative Bacteriology,* have served microbiologists since 1923. Eight editions of *Bergey's Manual of Determinative Bacteriology* have been published between 1923 and 1974. The first edition of *Bergey's Manual of Systematic Bacteriology* was published in four volumes between 1984 and 1990. The first volume of the second edition (see also http://www.bergeys.com), which covers the *Archaea* and the deep branching and phototrophic *Bacteria*, was released in 2001 (5). The second volume, covering the *Proteobacteria* was released in 2005 and the third volume on the *Firmicutes* in 2009; two more volumes are scheduled to follow in the coming years.

Bergey's Manual provides formal descriptions of all prokaryote taxa described to date, including both phenotypic and genetic information – all in accordance with the polyphasic approach outlined in Sect. 3.3. It also supplies much useful information that enables the microbiologist to identify his isolates. The older editions (*Bergey's Manual of Determinative Bacteriology*) provided extensive keys for the identification of bacterial isolates, resembling the dichotomous identification keys found in plant identification manuals. This approach has largely been abandoned in the two editions of *Bergey's Manual of Systematic Bacteriology*. Here, bacterial and archaeal classification is primarily based on 16S rRNA gene sequence comparisons (5, 31). The classification scheme shown in Table 3.5 is based on the latest edition of the manual. Although this classification scheme does not represent an "official" classification of prokaryotes, it has been adopted by the scientific community more or less as a "consensus" framework to classify prokaryotes. Tables for differentiation of the various taxa are included in each chapter to make the information accessible for identification purposes too.

7.2. The Prokaryotes

The Prokaryotes (32) is another extremely valuable resource of information on bacterial diversity and systematics. This handbook is now in its third edition. The first edition (1981) was published in two volumes, the second (1992) in four. The third edition was first published online only, but a printed edition in 7 volumes, encompassing 259 chapters on more than

7,500 pages, was published in 2006. These chapters cover all aspects of the biology of the prokaryotes, provide useful practical suggestions for growing and handling them, and include recipes of growth media for the cultivation of each group. The wealth of information supplied is also very useful as an aid toward identification of new isolates.

8. IDENTIFICATION OF PROKARYOTE ISOLATES

Environmental microbiologists are often faced with the need to identify bacteria isolated from the environment. From the earlier sections in this chapter, it should be apparent that identification of prokaryotes is not a simple procedure. There are no straightforward identification tables such as those existing for the higher plants. The older editions of *Bergey's Manual* indeed contained such dichotomous identification keys, but nowadays such keys are no longer satisfactory. Characterization of species, and therefore identification as well, is now based on the polyphasic approach to taxonomy (13, 20, 33) (see Sect. 3.3). The old identification tables are, however, still of considerable use for specialized groups of bacteria such as pathogens in clinical microbiology and potential pathogens in public health microbiology. Such identification schemes should not be considered as classification schemes, and they serve for practical purposes only (22). In such identification schemes, key phenotypic characteristics should be chosen for testing so that they can be easily determined by most microbiology laboratories. It is important that the identification depend on a pattern of several properties, not merely one or a few characteristics. It is also desirable that the determination of those characteristics chosen for an identification scheme be relatively inexpensive, and that the tests to be performed will give results in a short time.

In most cases, identification of a prokaryote begins at the level of domain, to descend to the level of phyla, classes, orders, and families, to finally narrow down to the level of the genus and the species. Different kinds of information are necessary in each step, as exemplified in Table 3.6. This table shows the place of a single bacterium, *Streptococcus pneumoniae*, in the taxonomic hierarchy, and includes information on the criteria on the basis of which the species can be classified in each of the taxonomic ranks. It may be noted that 16S rRNA gene sequence information is of great value to place an organism within the higher ranks, but less so at the species level.

In accordance with the polyphasic approach (13, 20, 33), there are many properties that should be investigated to obtain a proper identification. These include morphological characters (cell shape and size, the Gram-reaction, cell inclusions, presence and nature of the surface layers, including extracellular capsules), information on motility (presence of flagella, their number and the way they are inserted into the cell, gliding movement), the mode of nutrition (assimilatory metabolism) and energy generation (dissimilatory metabolism), the cells' relationship to molecular oxygen, temperature, pH, tolerance toward and requirement for salt, and many others. Miniaturized tests, such as the BIOLOG® and the API system, are often very useful. Additional tests of value toward the identification of the isolate may be its sensitivity toward different antibiotics, as well as immunological properties. Genotypic information is often essential. Notably, the 16S rRNA gene sequence is a very valuable tool for placing any isolate in the proper place in the classification scheme, at least down to the

Table 3.6
The species *Streptococcus pneumoniae* in the taxonomic hierarchy

Taxonomic division	Name	Properties	Confirmed by
Domain	*Bacteria*	Prokaryotic cell structure; Ribosomal RNA sequences typical of the *Bacteria*	Microscopy; 16S rRNA gene sequence; Presence of peptidoglycan in the cell wall; Lipids containing straight-chain fatty acids bound by ester bonds to glycerol
Phylum	*Firmicutes*	Gram-positive cell wall structure; Ribosomal RNA sequences typical of the *Firmicutes*	Gram stain and analysis of the cell wall structure; 16S rRNA gene sequence
Class	"Bacilli"	Aerobic or anaerobic – aerotolerant life style; Ribosomal RNA sequences typical of the "Bacilli"	16S rRNA gene sequence
Order	"Lactobacillales"	Fermentative metabolism with aerotolerant growth; Ribosomal RNA sequences typical of the "Lactobacillales"	Analysis of products formed during growth under aerobic and anaerobic conditions; 16S rRNA gene sequence
Family	*Streptococcaceae*	Spherical cells in pairs or chains or tetrads; Fermentative metabolism with fastidious nutritional demands; Ribosomal RNA sequences typical of the family *Streptococcaceae*	Microscopy; Analysis of fermentation products; Tests for growth on simple and complex media; 16S rRNA gene sequence
Genus	*Streptococcus*	Cell division in one plane resulting in pairs and chains; Homolactic type of fermentation; Morphological properties; Ribosomal RNA sequences typical of the genus *Streptococcus*	Microscopy; Analysis of fermentation products; Microscopy; 16S rRNA gene sequence; G + C content of DNA
Species	*Streptococcus pneumoniae*	Cells typically in pairs; Distal ends of each pair of cells tend to be pointed or lance-shaped; Cells often surrounded by a polysaccharide capsule; Sodium hippurate not hydrolyzed; Inulin not fermented; Characteristic reaction on blood agar; Inhabits the respiratory tract of man and animals	Microscopy; Fermentation tests on different substrates; DNA–DNA hybridization; Serological properties

family and genus level. For final identification of the species, DNA–DNA hybridization tests are the ultimate tool to decide whether two isolates should be classified in the same species. More specific tests, such as serotyping of phage typing, may be necessary for certain groups of microorganisms to obtain a reliable identification.

It is essential to ensure that pure cultures be used when the above tests are performed, otherwise contamination by other microorganisms makes the results of any of these tests meaningless. Another important rule is that the test methods should be carefully standardized, and known "positive" and "negative" controls should be included, so that the result obtained with the unknown isolate can be compared with the behavior of known organisms. For identification to the species level, it is essential that the unknown be compared with the type strains of the related species (see Sect. 2.1).

9. THE NUMBER OF DIFFERENT SPECIES OF PROKARYOTES IN NATURE

All the information provided in the preceding sections, and all classification and identification schemes for the prokaryotes, including those given in *Bergey's Manual*, are based on those 8,226 species of prokaryotes that have been isolated, characterized, and whose names have been validly published. New species are constantly being discovered. Thus, 593, 631, and 598 new species names have been validated in 2006, 2007, and 2008, respectively.

Nowadays, it is well established that the over 8,000 prokaryote species cultured and described form only a small part of the true number of species extant in nature. It is known for a long time that our cultivation techniques enable growth of only a small fraction of the microorganisms present in any natural sample. Comparison of the number of bacterial colonies that appear on agar plates or the number of bacteria determined with other growth-dependent methods with the number of bacteria that can be observed by direct examination of the sample by microscopic techniques invariably shows that only a small fraction of those organisms present can be cultured. The discrepancy between the viable counts and the total microscopic counts is in many cases a difference of several orders of magnitude. This observation, known as "the great plate count anomaly," shows that our cultivation methods are inappropriate for growing all prokaryotes (8). It also suggests that among those many organisms that do not form colonies on agar plates or cannot be cultured with any other available technique may be many novel species with unknown characteristics. Nature can cultivate all microorganisms, but the microbiologists still have much to learn about the proper methods to bring even the numerically dominant *Bacteria* and *Archaea* into culture. Our isolation and cultivation methods, which to a large extent are based on the procedures introduced by Robert Koch and his coworkers in the 1880s, are obviously not suitable for many prokaryotes.

Introduction of 16S rRNA gene sequencing methods into environmental microbiological studies has confirmed that we know only a small fraction of the number of species of prokaryotes. Moreover, it can now be ascertained that, in most cases, we even do not know the identity of those microorganisms that are numerically dominant in common environments such as soil, seawater, rivers, lakes, etc. Characterization of the microbial community in complex ecosystems nowadays often includes the sequencing of 16S rRNA genes present in DNA extracted from the natural community. In a typical experiment, DNA is isolated from the

biomass present in the sample. The genes encoding 16S rRNA are then amplified by use of the polymerase chain reaction (PCR), either using universal primers (i.e., primers that will enable amplification of all 16S rRNA genes, from all prokaryotes and even eukaryotes, or primers that target specific groups such as the domain *Bacteria*, the class "Betaproteobacteria,") or a certain family or genus within that class. The products obtained are then separated using electrophoretic techniques, purified, and sequenced, often following an additional cloning step to further increase the amount of material and to improve its purity (34). It must be stressed here that there are many potential problems with the methodology that can distort the result of the analysis, and the technology still has many limitations that should be recognized (35).

Comparison of the sequences obtained with the known 16S rRNA gene sequences of the bacterial and archaeal species, as present in the GenBank and in the Ribosomal Data Base (http://rdp.cme.msu.edu, see Sect. 6.1), almost invariably shows that the sequences obtained differ from those of the type strains of the established species. In most cases, the differences are so substantial that the organism that harbors the 16S rRNA gene sequence characterized would deserve classification in a new genus and often in a new order, a new class or sometimes even a new phylum, if only the organism that harbors this 16S rRNA gene could be isolated in culture, characterized, deposited in culture collections, and named in accordance with the rules of the Bacteriological Code. The Ribosomal Database is full of such "environmental 16S rRNA gene sequences" recovered from natural samples that represent yet unknown species. It only seldom occurs that a complete match is found between a small subunit rRNA gene sequence isolated from nature and any of the type strains of the over 8,000 described and named species of prokaryotes. This shows that indeed the numerically dominant organisms in the environment studied are not those species that have been named and are available from culture collections. Based on such studies, it can be estimated that the number of prokaryote species described and named thus far is at most 1–2% of the true number of bacterial species extant, as based on the current species concept. The true fraction may even be much lower than that (34).

A completely different line of evidence that enables an estimate of the true diversity of the microbial communities in water, soil, and other ecosystems comes from measurements of the renaturation kinetics of DNA extracted from the community following thermal denaturation (36). When thermally denatured DNA derived from different organisms is mixed and cooled down, the average time needed for a single DNA strand to find its homolog depends on the frequency in which that homolog occurs within the mixture, and thus on the number of other genomes present. Based on the results of such studies, Dykhuizen (37) calculated that there may be at least 10^9 different bacterial species on Earth that differ from each other sufficiently to meet the species delineation criteria explained in Sect. 2.2. According to his estimations, there may be more than half a million different bacterial species in a single 30 g sample of forest soil.

The extensive database of 16S rRNA gene sequences allows for the design of specific probes that enable the detection of specific groups of prokaryotes in the natural environment. The popular technique of fluorescence in situ hybridization ("FISH") is based on the design of 16S rRNA-targeted probes that are labeled with a fluorescent marker. The cells are fixed so that they become permeable to the probes. After reaction with the probes, the excess nonfixed

probe is washed off, and the samples are then examined in the fluorescence microscope. Probes of different specificity can be designed, from general, domain- or class-specific probes to probes that allow discriminating between members of the same genus or species, provided that the stringency of the washing procedure is sufficient to differentiate between very similar sequences. The technique allows obtaining information on the spatial distribution of different types of microorganisms within complex ecosystems (38, 39).

These culture-independent small-subunit rRNA-based techniques have shown that there are major groups of microorganisms in nature that are present in very large numbers in the most common ecosystems, but of which we do not have a single representative in culture (40, 41). As outlined in Sect. 6.4, the *Archaea* are presently classified in two phyla, the *Euryarchaeota*, which are either obligatory anaerobic methanogens, extreme halophiles, or thermophiles/thermoacidophiles, and the *Crenarchaeota*, all known representatives of which are extreme thermophiles. It is now well documented that the world's oceans contain large amounts of *Archaea* of both phyla. *Archaea*-related 16S rRNA gene sequences belonging to different groups, and different from the sequences of the cultured *Archaea*, are consistently being amplified from DNA extracted from marine picoplankton. Moreover, fluorescent in situ hybridization using *Archaea*- and *Bacteria*-specific probes has shown that about 30% of all prokaryotes in the oceans belong to the archaeal domain (30). The domain *Archaea* thus consists not only of extremophiles and methanogens. However, we have very little information on the physiology of these extremely abundant marine *Archaea*, and only recently has the first representative of marine *Crenarchaeota* been isolated: the ammonia-oxidizing autotroph "*Candidatus* Nitrosopumilus maritimus. More archaeal rRNA gene sequences unrelated to any of the cultured groups have been recovered from other environments. A lineage designated "Korarchaeota" is present in certain hot springs. Based on the 16S rRNA gene sequences, the group is sufficiently different from the *Crenarchaeota* and the *Euryarchaeota* to obtain the status of a new phylum. No members of this group have yet been cultured. Similarly, many groups of *Bacteria*, differing from the cultured ones at the phylum level, have been recognized on the basis of environmental 16S rRNA gene sequences, and are awaiting to be isolated (40).

Unfortunately, the 16S rRNA gene sequence alone does not provide any information on the physiological properties of all these yet uncultured prokaryotes. Isolation and characterization of these abundant microorganisms, which thus far have eluded all attempts toward their cultivation, continues to be a major challenge to the microbiologist who wants to know those prokaryotes that to a large extent determine the properties of the ecosystem studied.

10. CONCLUSIONS

The description of the present status of prokaryote taxonomy and classification, as given in the preceding sections, shows that prokaryote classification is not at all straightforward. First of all, there is no clear species concept for the prokaryotes; all classification systems have to be based on some kind of a consensus of what a prokaryote species is and how to discriminate it from all other species. We have further seen that the number of species of *Archaea* and *Bacteria* together that have been described and whose names have obtained standing in the nomenclature under the Bacteriological Code is relatively small – little over 8,200.

Comparison of the properties of new isolates with these species described in the literature enables their identification as members of a recognized species or, if different from all other species, as members of a new, yet to be described species. A polyphasic approach, including determination of both genetic and phenotypic properties, to compare strains is essential to properly assign them to genera and species within the existing classification schemes.

Even the identification of all species of microorganisms that can be isolated from any ecosystem using the currently available techniques will not provide a reliable picture of the microbial diversity in that ecosystem. The applications of molecular biology techniques, especially those that target 16S rRNAs or the genes encoding them, have unequivocally shown that those microorganisms that have been cultured thus far form only a small fraction of the true prokaryotic diversity in any ecosystem. Generally, the numerically dominant types belong to species that are still waiting to be isolated. Those 8,226 described and named species probably represent no more than 1–2% of the true number of bacterial species, possibly even much less.

Finally, it must be stressed that classification of microorganisms is a dynamic process, and that our views of how the prokaryotes can best be classified are constantly changing with the advancement of our knowledge. As stated before, there is no "official" classification of prokaryotes, and "consensus" classifications such as those given in Table 3.4 are constantly subject to change.

NOMENCLATURE

DNA = Deoxyribonucleic acid
RNA = Ribonucleic acid
rDNA = DNA coding for ribosomal RNA
rRNA = Ribosomal RNA
S = Svedberg unit of sedimentation, equal to 10^{-13} s

REFERENCES

1. 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
2. Oren A (2008) Systematics of Archaea and Bacteria, In: Biological Science. Fundamentals and Systematics, Vol. 2. Encyclopedia of Life Support Systems (Minelli A, Contrafatto G, eds). United Nations Educational, Scientific and Cultural Organization, pp. 164–188
3. Brenner DJ, Staley JT, Krieg NR (2001) Classification of procaryotic organisms and the concept of bacterial speciation. In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 27–31
4. Lapage SP, Sneath PHA, Lessel EF Jr, Skerman VBD, Seelinger HPR, ClarkWA (eds) (1992) International Code of Nomenclature of Bacteria (1990 revision). American Society of Microbiology, Washington, DC

5. Garrity GM (editor-in-chief), Boone DR, Castenholz RW (eds) (2001) Bergey's Manual of Systematic Bacteriology, 2nd ed., Vol. 1. The Archaea and the Deeply Branching and Phototrophic Bacteria. Springer-Verlag, New York
6. Sneath PHA (2001) Bacterial nomenclature. In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 83–88
7. Stackebrandt E (2000) Defining taxonomic ranks. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (eds) The Prokaryotes: an evolving electronic resource for the microbiological community, 3rd edn, release 3.1, 20 January 2000. Springer, New York, http://link.springer-ny.com/link/service/books/10125/
8. Oren A (2004) Prokaryote diversity and taxonomy: present status and future challenges. Philos Trans R Soc Lond B Biol Sci 559:623–638
9. Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP, Trüper HG (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int J Syst Bacteriol 37:463–464
10. Stackebrandt E, Frederiksen W, Garrity GM, Grimont PAD, Kämpfer P, Maiden MCJ, Nesme X, Rosseló-Mora R, Swings J, Trüper HG, Vauterin L, Ward AC, Whitman WB (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
11. Rosselló-Mora R, Amann R (2001) The species concept for prokaryotes. FEMS Microbiol Rev 25:39–67
12. Cowan ST (1968) A dictionary of microbial taxonomic usage. Oliver & Boyd, Edinburgh
13. 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
14. Woese CR (1987) Bacterial evolution. Microbiol Rev 51:221–271
15. Woese CR (1992) Prokaryote systematics: the evolution of a science. In: Balows A, Trüper HG, Dworkin M, Harder W, Schleifer K-H (eds) The Prokaryotes. A handbook on the biology of bacteria: ecophysiology, isolation, identification, applications, 2nd edn, vol 1. Springer, Berlin, pp 3–18
16. Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal of the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87:4576–4579
17. Stackebrandt E, Goebel BM (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
18. Fox GE, Wisotzkey JD, Jurtshuk P Jr (1992) How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int J Syst Bacteriol 42:166–170
19. Skerman VBD, McGowan V, Sneath PHA (1980) Approved lists of bacterial names. Int J Syst Bacteriol 30:225–420
20. Gillis M, Vandamme P, De Vos P, Swings J, Kersters K (2001) Polyphasic taxonomy, In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 43–48
21. Ventosa A, Gutiérrez MC, Kamekura M, Zvyagintseva IS, Oren A (2004) Taxonomic study of *Halorubrum distributum* and proposal of *Halorubrum terrestre* sp. nov. Int J Syst Evol Microbiol 54:389–392

22. Krieg NR (2001) Identification of Prokaryotes. In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 33–38
23. Sneath PHA (2001) Numerical taxonomy. In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 39–42
24. Trüper HG (1999) How to name a prokaryote? Etymological considerations, proposals and practical advice in prokaryotic nomenclature. FEMS Microbiol Rev 23:231–249
25. Trüper HG (2001) Etymology in nomenclature of procaryotes. In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 89–99
26. Labeda DP (2001) Culture collections: an essential resource for microbiology. In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 111–113
27. Ludwig W, Schleifer K-H (1994) Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. FEMS Microbiol Rev 15:155–173
28. Gupta RS (1998) What are archaebacteria: 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
29. Madigan MT, Martinko JM, Dunlap PV, Clark DP (2009) Brock Biology of Microorganisms. 12th ed. Pearson-Benjamin Cummings, San Francisco
30. Karner MB, DeLong EF, Karl DM (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. Nature 409:507–510
31. Ludwig W, Klenk H-P (2001) Overview: a phylogenetic backbone and taxonomic framework for procaryotic systematics. In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 49–65
32. Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (eds) (2006) The Prokaryotes. 3rd ed. Springer, New York
33. Trüper HG, Schleifer K-H (2006) Prokaryote characterization and identification, In: The Prokaryotes. (Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E eds), Springer, New York, Vol. 1, pp. 58–79
34. Amann RI, Ludwig W, Schleifer K-H (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59:143–169
35. 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
36. Torsvik V, Sørheim R, Goksøyr J (1996) Total bacterial diversity in soil and sediment communities – a review. J Ind Microbiol 17:170–178
37. Dykhuizen DE (1998) Santa Rosalia revisited: why are there so many species of bacteria? Antonie van Leeuwenhoek 73:25–33
38. DeLong EF, Wickham GS, Pace NR (1989) Phylogenetic stains: ribosomal RNA-based probes for the identification of single microbial cells. Science 243:1360–1363
39. Amann R, Schleifer K-H (2001) Nucleic acid probes and their application in environmental microbiology. In: Boone DR, Castenholz RW (eds), Garrity GM (editor-in-chief) Bergey's manual

of systematic bacteriology, 2nd edn, vol 1. The Archaea and the deeply branching and phototrophic Bacteria. Springer, New York, pp 67–82
40. Hugenholtz P, Goebel BM, Pace NR (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J Bacteriol 180:4765–4774
41. Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, Giovannoni SJ (2002) SAR11 clade dominates ocean surface bacterioplankton communities. Nature 420:806–810