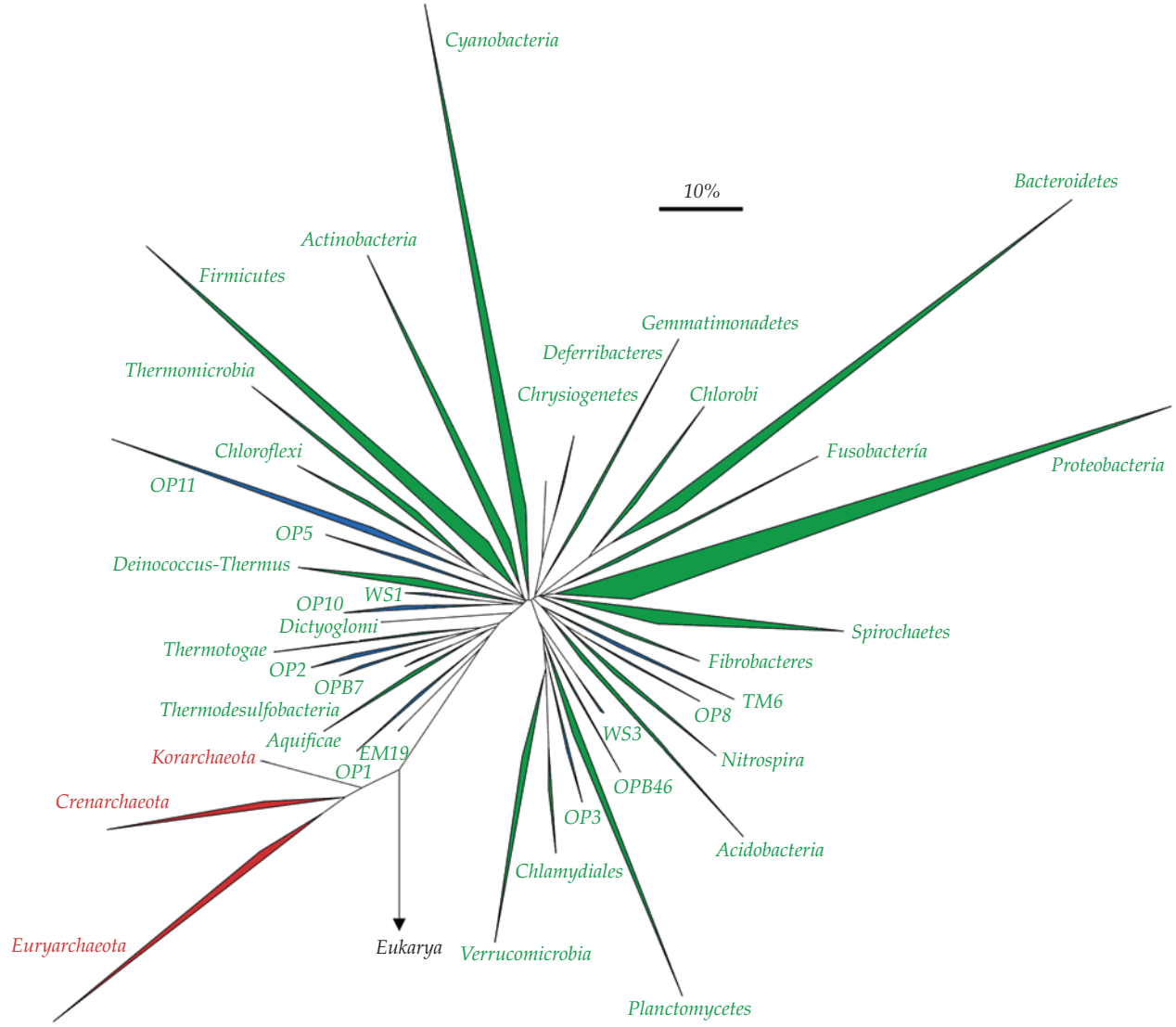


Phylogenetic reconstruction based on the analysis of 16S ribosomal RNA of the domains with a prokaryotic cellular organization: Archaea (red) and Bacteria (green). All phyla represented only by sequences obtained from environmental gene libraries are indicated in blue. They are new candidate phyla, which are still not represented in a single pure culture in a laboratory. If we exclude the most deeply branching orders of phyla (for example, Aquificae, Thermodesulfobacteria, Thermotogae, and Dictyoglomi), it has not been possible to resolve the order of branches in the rest of the phyla with any of the sets of phylogenetic markers in use (16S rRNA, 23S rRNA, gyrB, ATPase, elongation factor Tu, etc.). The bar indicates a 10% divergence in the sequence.

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

Archaea and Bacteria

The Prokaryotic Cell Organization

Ramon Rosselló-Móra and Josefa Antón

3

SUMMARY Archaea and Bacteria are the two domains that include all unicellular organisms without a nucleus, which are also known as "prokaryotes." These microorganisms represent the most abundant group of living beings in the biosphere, and whose biomass may be equivalent to that of all plants. However, the number of described archaeal and bacterial species is fewer than 12,000 (as of 2014), less than 0.5% of all species described to date in the global calculation of living organisms. Because of their ubiquity and their phylogenetic and metabolic diversity, it is speculated that there are actually over 10 million species. Archaea and Bacteria colonize every niche that supports life; they have been developing and modifying the planet for approximately 3.5 billion years; and they are probably the most important organisms in maintaining the biosphere's homeostasis. The greatest problem in describing new species is the need to have pure laboratory cultures available for research. Nevertheless, the new molecular methodologies are helping discern species and permit acquisition of a great deal of information on the genetic potential and the true diversity hidden in natural environments.

The term prokaryote refers to a type of cellular organization without a nucleus. The "prokaryotes" are a paraphyletic group; that is to say, they include organisms with very different evolutionary histories and without a single common ancestor. As shown in the phylogenetic reconstruction that begins this chapter, of the three domains into which organisms can be classified (Eucarya, Archaea, and Bacteria), two (Archaea and Bacteria) are formed exclusively by microorganisms with a prokaryotic cellular organization. These two phylogenetic groups are also the two taxonomic categories of greatest rank that include all known "prokaryotes." Sometimes, the terms "prokaryote" and bacteria are used interchangeably, although this is not correct from a phylogenetic point of view.

Although there are important exceptions, Archaea and Bacteria are normally unicellular organisms, small in size (between 2 and 10 μm), whose nucleoplasm (or genophore) is not physically separated from the cytoplasm by an endomembrane system. The plasma membrane often has a complex topology, with vesicular, lamellar, or tubular intrusions into the cytoplasm. Self-replicating vacuoles and organelles independent of the plasma membrane are relatively rare. Respiratory and photosynthetic functions are associated with the plasma membrane (with the exception of the cyanobacteria, in which they are associated with thylakoids). The ribosomes of Archaea and Bacteria are generally of the 70S type and are dispersed in the cytoplasm (there is no endoplasmic reticulum where the

What is a "prokaryote"?

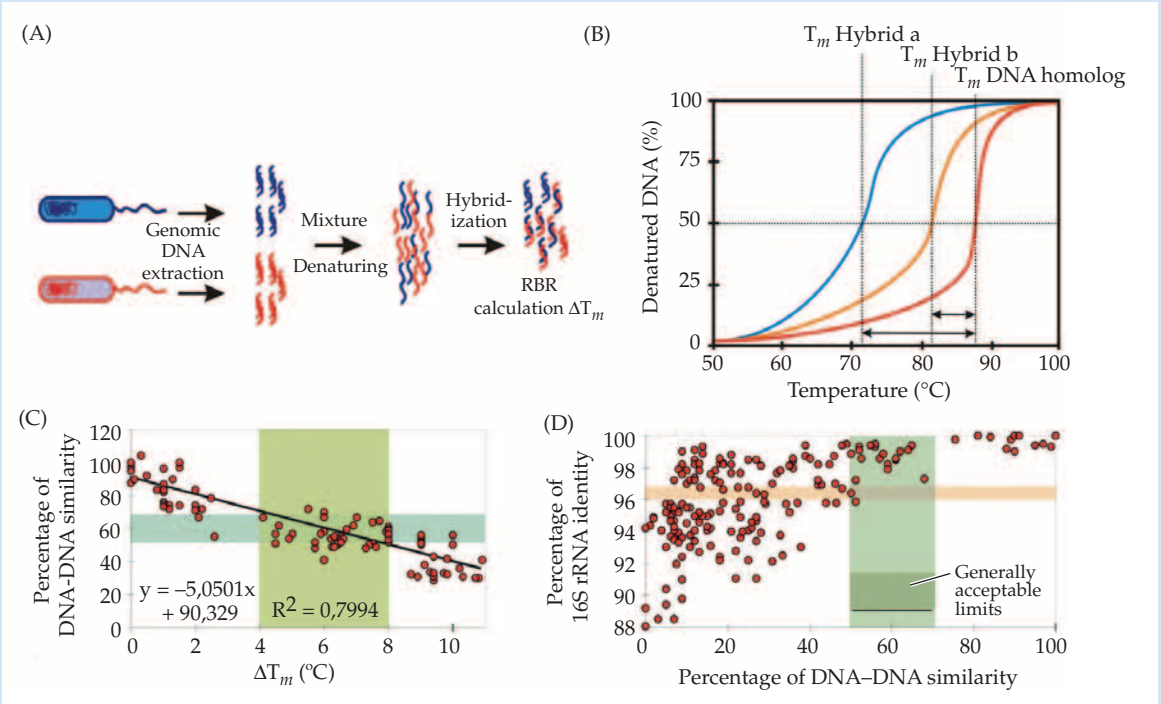
The term *prokaryote* refers to a type of cellular organization. All cells characterized by absence of a differentiated membrane-bound nucleus separating the genome from the cytoplasm, and which do not form differentiated tissues, are called prokaryotes, from the Greek *pro* (before) and *karyon* (nut or nucleus). However, in some cases they do have structures resembling the eukaryotic nucleus (for example, *Planctomycetes*). The "Prokaryote" group is defined, therefore, by the absence (not the presence) of a character, and covers two independent, paraphyletic lineages. The term "prokaryote" does not name a taxonomically valid category, but rather groups two domains (Bacteria and Archaea), which, together with the Eucarya, represent the entirety of cellular diversity in the biosphere.

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

BOX 3.1 Species of Archaea and Bacteria

As with eukaryotes, the species is the basic unit from which the entire taxonomic scheme of Archaea and Bacteria has been built. Classification of these unicellular organisms was established taking the hierarchical system conceived by Linnaeus for plants and animals as a model. However, due to their intrinsic characteristics, genetic as well as physiological and ecological, the parameters useful for delineating species in plants and animals are not useful in Archaea and Bacteria. In general, Archaea and Bacteria reproduce asexually and clonally, as well as apparently having worldwide distribution. Therefore, the concept of sexual isolation, useful in plants and animals, is not applicable to prokaryotes. A species is considered to be a group of organisms of monophyletic origin, displaying genomic and phenotypic coherence that makes them distinguishable from other similar groups. The need to start from pure cultures poses the greatest limitation for classifying species, because at present it is almost impossible to obtain enough information directly from an environmental sample to classify Archaea and Bacteria.

Phylogenetic coherence of members of a species is demonstrated by analysis and comparison of housekeeping genes, universal among prokaryotes, and especially by means of the gene that codes for 16S ribosomal DNA. Genomic coherence is verified by analyzing parameters such as the G+C content and the hybridization among genomes, known as DNA–DNA hybridization (DDH). It consists of quantifying, after hybridization of two DNAs, either (a) the percentage of hybrid chains formed, known as relative binding ratio (RBR), or (b) the thermal stability of the hybrid in comparison with the homolog DNA (ΔT_m). Both parameters are correlated (c), and in general, an elevated RBR (> 75%) correlates to small differences in the annealing temperature of hybrids and homologs (ΔT_m < 4ºC). That means that the more two genomes resemble one another, the more genes they have in common and the more similar are their sequences, thus more hybrids are formed than would be otherwise. Similarly, the more the thermally stable genes resemble one another, the more stable is the hybrid formed; thus, two similar genomes will give rise to hybrids with very similar annealing temperatures. There is also a correlation between DDH and the identity of the rRNA sequence (D). E. Stackebrandt and B. M. Goebel proposed establishing a value of 97%, later narrowed to 98.7%, in the identity of this sequence as a boundary for delineating species. However, due to the enormous diversity in phylogenetic groups and ecological niches, this gene's rate of mutation is not synchronous, and therefore, this rule is not always followed (d). At present, there is an attempt to replace DDH by multilocus sequence analysis (MLSA), which concatenates sequences of several essential genes with unique copies in the genome. The resulting alignments contain many more homologous positions than those based on a single gene. The intent is to thus minimize biases in choice of a single gene for verifying phylogenetic homogeneity. However, this technique is still in a development stage and is not yet used very frequently for description of new species. The parameter of choice that would replace DDH is the average nucleotide identity (ANI). This parameter measures the percentage of identity between homologous stretches of DNA in a pairwise genome comparison. The ANI value between 94–96% would mirror the DDH value of 70% (see Figure 3.1). Phenotypic coherence is observed by applying metabolic and enzymatic tests, and by measuring chemotaxonomic parameters such as profiles of fatty acids, polyamines, respiratory quinone compositions, polar lipids, and so forth.



UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

Basic terms

DDH (DNA–DNA hybridization): Whole genome hybridization technique used in delineating species of prokaryotes.

FISH (Fluorescence in situ hybridization): Optical microscopy technique combining the use of fluorochromes that bind DNA and phylogenetic probes targeting ribosomal RNA. This technique is used to reveal the structure of natural microbial communities.

ICSP (International Committee on Systematics of Prokaryotes): Scientific community dealing with the discussion and resolution of conflicts arising in the classification of "prokaryotes" and watching for the correct application of the Bacteriological Code.

ITS (Internal transcribed spacer): DNA fragment contained between the 16S and 23S rRNA genes in operons of Archaea and Bacteria. They generally contain noncoding DNA, as well as some transfer RNA genes.

MLSA (Multilocus sequence analysis): Also known as MLST (multilocus sequencing typing). System for comparing among different strains or species based on sequence analysis of several essential genes; normally a minimum of seven.

Northern blot: Technique combining the use of RNA extracts bound to a physical support by hybridization with specific marker probes.

PFGE (Pulsed Field Gel Electrophoresis): Electrophoresis technique allowing the separation of high-molecular weight DNA fragments that could not be separated by means of conventional electrophoresis.

RAPD (Randomly Amplified Polymorphic DNA): PCR amplification technique using randomly generated sieves. In general, the patterns of bands obtained are characteristic of a single genome and permit identification of clonal variants.

ribosomes are anchored). Archaea and Bacteria capture nutrients dissolved in the medium through their cellular envelopes, a process called osmotrophy. There are no known cases of either phagotrophy or the capture of nutrient particles by formation of pseudopods or digestive vacuoles. The cell is normally enclosed in a rigid cellular wall, although this is not a universal trait, because examples are known of both Archaea and Bacteria without cell walls. The cells can be immobile or have a swimming movement, through flagella or by sliding over surfaces. Most are unicellular organisms, but on occasions they may form filamentous, mycelial, or colonial structures. Although there are some exceptions, the possibilities of morphological differentiation are few and limited to anchorage structures, resting states, and modifications in cellular morphological characteristics. Bacteria and archaea experience gene transfer and recombination mechanisms, but their reproduction never involves gametogenesis or the formation of zygotes. Morphological differentiation between eukaryote and prokaryote cellular organization is thus simple, although in a few cases it can be complicated by the presence of attributes such as hyphae in *Actinomycetes* that make them resemble filamentous fungi, or the flagellar tufts of some bacteria that might be confused with eukaryote flagella.

From the beginning of microbial taxonomy at the end of the nineteenth century to the first molecular phylogenies at the end of the 1970s in the twentieth century, classification of Archaea and Bacteria was based essentially on the observation of phenotypic similarities by comparison of metabolic characters, and on genomic similarities by means of **DNA–DNA hybridization** (or **DDH**, now to be replaced by the average nucleotide identity, which reproduces in silico pairwise genome comparisons; **Figure 3.1**) by determining the GC content (**Box 3.1**). In the case of Archaea and Bacteria, both parameters

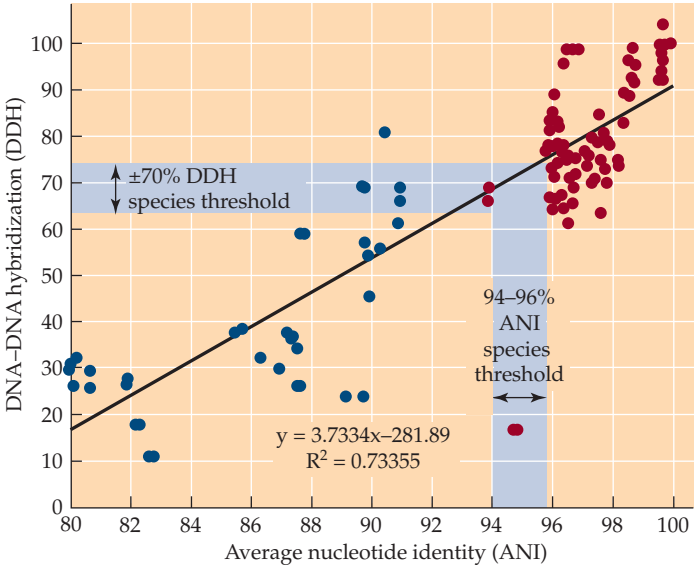


Figure 3.1 Reciprocity between DNA–DNA hybridization (DDH) and the in silico average nucleotide identity (ANI) results. The species thresholds for whole-genome comparisons were observed to occur at about 70% DDH. These results find their reciprocity at 94–96% ANI values when comparing orthologous DNA stretches of two genomes. DDH is experimentally error prone, which may explain the outliers found.

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

compensate for the scarce information provided by the morphological and ontogenetic aspects (for example, development stages or cellular cycles) that have been useful in configuring the classification scheme of eukaryotes. In particular, delineation of categories higher than species and genus has been complicated by the lack of universally applicable objective criteria for all Archaea and Bacteria. The development of methodology for manipulating, sequencing and comparing the ribosomal RNA sequence (originally 5S and later 16S, or its coding gene) permitted establishment of a generally applicable criterion, and with it, greater objectivity in delineating hierarchical categories above species. In fact, the decade of the 1990s was characterized by significant activity in reclassifying taxa. Genera such as *Pseudomonas* were reclassified into several orders, with their respective families and genera. At present, consolidation of the taxonomy of Archaea and Bacteria has been achieved by the use of these and other phylogenetic markers. The new classifications are essentially based on phylogenetic reconstructions that, in general, seem to correlate with comparative studies among genomes. With the exception of a few scattered cases of genera and species, 16S RNA molecular phylogeny has enabled classification of Archaea and Bacteria into taxa formed by organisms of monophyletic origin. However, some genera and species constituted by paraphyletic or polyphyletic aggregations of organisms do appear.

The first phylogenetic reconstruction based on comparison of genes coding for the small subunit of ribosomal RNA, carried out in 1977, separated the cellular organisms into three distinct lineages named *domains*. According to this reconstruction, the set of "prokaryotes" was divided into two phylogenetic lines (see opening tree): the one encompassing the true bacteria, or the domain of Bacteria (initially named Eubacteria); and the domain of Archaea (initially named Archaeobacteria). The most surprising conclusion from this research was that Archaea, at least in regard to replication mechanisms, might have a monophyletic origin with the eukaryotes (the domain Eucarya). This is why the category of "Prokaryote" is considered a paraphyletic group.

Members of Archaea and Bacteria share a significant number of genetic characters, such as covalently closed circular DNA chromosomes; the presence of plasmids, ribosomes of similar size, the organization of genes in operons, among others; and physiological characters, such as the capacity for anaerobic respiration, nitrogen fixing, chemolithoautotrophy, rhodopsin-based photosynthesis, and thermophilia, among others. However, Archaea and Eucarya also share relevant genetic traits such as the presence of histones, the initiator methionine tRNA, the similarity in the RNA polymerases, and a promoter structure in the form of a TATA box. These last facts reinforce the hypothesis that, at least in regard to genetic mechanisms, Eucarya and Archaea share a common ancestor.

Abundance and Diversity of Archaea and Bacteria

Archaea and Bacteria are ubiquitous. Their presence has been detected in a multitude of natural and artificial environments, from several thousand meters high in the atmosphere to the deepest layers of the earth's crust in which life can exist. It is estimated that they first appeared on Earth some 3500 million years ago, more than 1500 million years before the Eucarya. Throughout this period, Archaea and Bacteria dominated the biosphere and transformed it, and they are in great part responsible for its current characteristics. As summarized in **Table 3.1**, it is estimated that the number of Archaea and Bacteria present in the biosphere is on the order of 10^{30} individuals, a value that surpasses the abundance of eukaryotes by orders of magnitude. Furthermore, the biomass they account for is almost equivalent to the total plant biomass. However, only about 10% of these individuals colonize environments to which humans have relatively easy access. The remaining 90% would be found in the subsurface, mostly in anaerobic conditions and with low metabolic activity. For this reason, these organisms' diversity, both metabolic and genetic, is expected to be very significant. In the gut of a human being, for example, there are more Archaea and Bacteria (between 10^{13} and 10^{14}) than there are eukaryotic cells in the body. Moreover, the combined genome (i.e.,

TABLE 3.1 Estimated values, in number of individuals and biomass, of Archaea and Bacteria in different environments of the biosphere[1,2]

Habitat	No. of cells ($\times 10^{28}$)	Grams of carbon ($\times 10^{15}$)
Aquatic (oceanic and terrestrial)	12	2.2
Oceanic subsurface	355	303
Soil	26	26
Terrestrial subsurface	25–250	22–215
Total	415–640	353–546
Plant	–	652

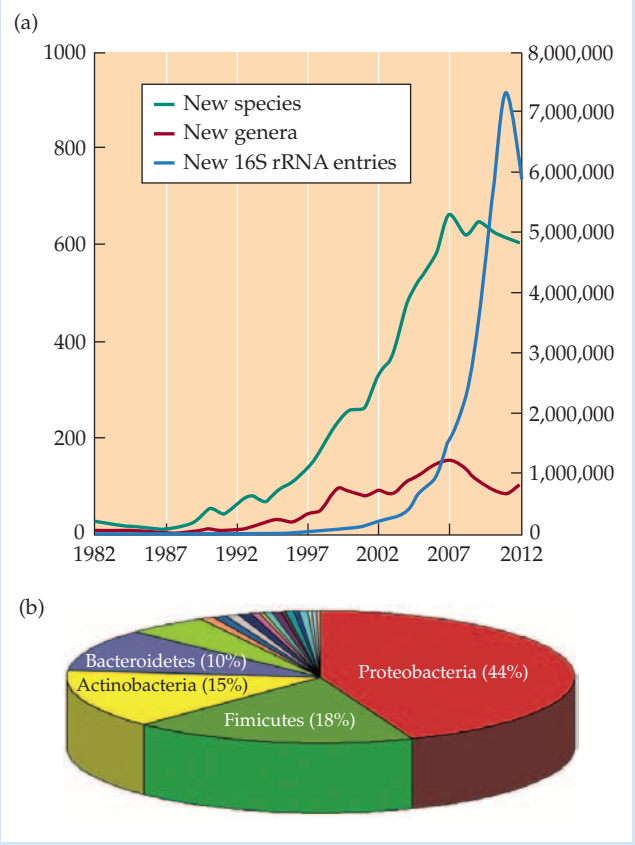
[1] Data from Whitman et al., 1998. *Proc. Natl. Acad. Sci. USA* 95: 6578–6583.

[2] Recently, such numbers have been presumed to be overestimated; the current absolute numbers are calculated in a range between 9.2–31.7 × 10^{29} cells. Kallmeyer et al., 2012. *Proc. Natl. Acad. Sci. USA* 109: 16213–16216.

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

BOX 3.2 Increase in new taxa and new 16S rRNA gene sequences in public repositories.

The speed in new taxa descriptions for Archaea and Bacteria had been enhanced by the use of 16S rRNA gene data to reconstruct phylogenies. Since the early 1990s there has been a parallel increase in new taxa descriptions and the deposit of new 16S rRNA gene sequences in public repositories (**Figure a**). The current (2014) rate of new descriptions is about 600–700 species per year. However, the deposit of new sequences in public repositories has increased exponentially to the point where currently there are about 700,000 entries per year. Most of the entries are environmental sequences unrelated to pure laboratory cultures. In the 16S rRNA gene databases, less than 1% are of cultured organisms, and less than 0.1% of type strains of described species. In 2010, the taxonomic classification of Archaea and Bacteria comprised 8602 species classified into 1779 genera, 285 families, 115 orders, 52 classes, 29 phyla, and 2 domains. The basic conditions for classifying new taxa are: having representatives available as pure laboratory cultures, and one selected strain being deposited in two international culture collections. Such premises, together with the need to study genealogy, genetics, and phenotype, are determinant for the speed of new classifications. Additionally, four phyla encompass nearly 90% of all the genera (and species) described: Proteobacteria (classically Gram-negative bacteria); Firmicutes and Actinobacteria (classicallly Gram-positive bacteria); and Bacteroidetes (**Figure b**).



metagenome) of these gut microorganisms contains on the order of 100 times more genes than those present in the human genome.

As reported in **Box 3.2**, the number of species of Archaea and Bacteria described to date (as of 2014) is fewer than 12,000, in contrast to the number of eukaryotic species described (**Table 3.2**) in which the arthropods alone number more than one million. Archaea and Bacteria amount to less than 1% of the total of all species described, in spite of being the most diverse group (as can be deduced from their abundance; their ecological, metabolic, and phylogenetic diversity; and their age in the biosphere). In this past decade, the increase in the descriptions of Archaea and Bacteria is notable. In fact, between 500 and 700 new species are described every year, but even following this dynamic, it would take 2000 years to arrive at numbers similar to those for arthropods. The main limiting factor in the description of species of such microorganisms is the need to obtain pure cultures in the laboratory. As has already been stated, the morphological characteristics of Archaea and Bacteria are not very

TABLE 3.2 Total number of species described at the end of the 20th century. Values shown in thousands of units[1]

Group	Species described (× 1000)	Percentage of total	Estimated species (× 1000)
Plants	270	15.4	300–500
Chordates	45	2.6	50–55
Arthropods	1065	60.8	2375–101,200
Molluscs	70	4	100–200
Nematodes	25	1.5	100–1000
"Protists"	40	2.3	60–200
"Algae"	40	2.3	150–1000
Fungi	75	4.3	200–9900
Archaea and Bacteria	4.9	0.3	50–3000
Viruses	4	0.2	50–1000
Other	115	6.6	200–800
Total	1753	100	3635–111,655

[1] Data from Bull and Stach. 2004. *Microbial Diversity and Bioprospecting*, pp. 15–28.

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

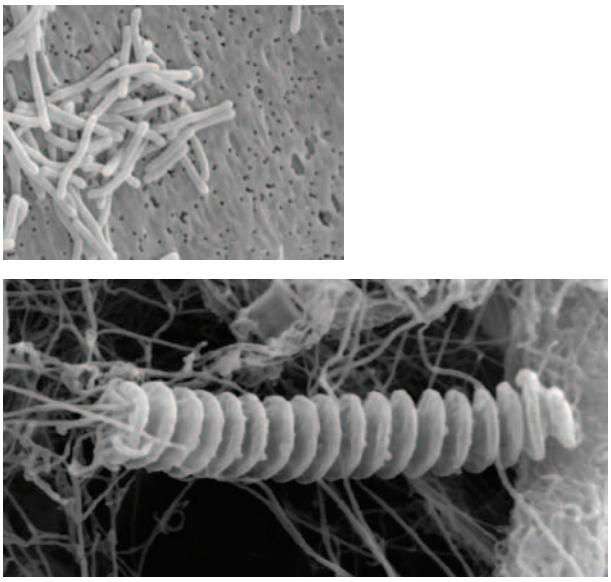


Figure 3.2 The morphology of Archaea and Bacteria offers little in the description of diversity. In general, most of the microorganisms are in the shape of rods (top; elongated structures of different lengths and diameters) or cocci (bottom; spheres of different sizes). There are a few cases of more complex shapes, such as spirals, hyphae, or peduncles. In any case, simply observing an organism does not guarantee identifying it.

informative (**Figure 3.2**); therefore, useful information for classifying them cannot be obtained through simple microscopic observation of natural samples. Due to their reduced size, and the complexity of species present in a single ecosystem (it is speculated that a cubic centimeter of water can host a genomic diversity equivalent to some 200 different genomes; and the same volume of soil, up to 4000), information about the metabolism or genetic complexity of a single microorganism cannot be obtained directly from its own environment. It is for that reason that it is necessary to isolate individual cells in the laboratory, and get these to multiply, to generate sufficient biomass to carry out experiments. However, this is not a trivial task, since most microorganisms are reluctant to be cultured by conventional techniques (**Table 3.3**). This is due to the fact that, in general, the culture media available have been designed for culturing pathogens or microorganisms related to human health. Although these culture media and the strategies followed in microbiology have been of considerable value in the study of infectious diseases and human health- and food-related issues, they have not been very successful for recovering most microorganisms present in nature, because in most cases less than 1% of the total of organisms present in a sample are successfully grown in a laboratory. Therefore, one of the most daunting challenges in microbiology is culturing new microorganisms (**Figure 3.3**). Application of molecular biology techniques to the study of microbial communities is making it possible not only to understand the diversity hidden in the ecosystems, but also to bring pure cultures of microorganisms into the laboratory and thus exploit their biotechnological possibilities.

TABLE 3.3 Possibility of culturing Archaea and Bacteria determined as a percentage of the culturable organisms compared to the total number of cells in a sample[1]

Habitat	Cultivability (%)
Sea water	0.001–0.1
Fresh water	0.25
Mesotrophic lake	0.1–1
Uncontaminated estuarine waters	0.1–3
Activated sludge from sewage treatment systems	1–15
Sediments	0.25
Soils	0.3

[1] Data from Amann et al., 1995. *Microbiol. Rev.* 59: 143–169.

Once a microorganism has been successfully isolated in the laboratory so that it is possible to culture and maintain it, identification and characterization must be done by observing its metabolism and analyzing its genetic characters. The task of obtaining information from Archaea and Bacteria requires applying molecular and biochemical techniques, which, in some cases, demand relatively complex equipment. In contrast to many types of eukaryotic organisms, the requirements of time, equipment, and infrastructure for an exhaustive description of a taxon are costly. Obviously, isolating organisms as pure cultures in the laboratory is an essential requirement for describing a new species of Archaea or Bacteria, and this makes description more difficult than it is for many eukaryotic organisms, which can be described from material preserved in formaldehyde or dried.

Classification and Collections of Cultures

Although there is no microbiological society dedicated specifically to the taxonomy of Archaea and Bacteria, nor is there an official classification, there is an official nomenclature with its own rules and standards. Classification of microbes has been by means of articles in various scientific journals, whose criteria for classification were established based on experience and the classifications originated by pioneers in the research. Nevertheless, there are two entities that, by consensus among microbiologists, have taken responsibility for

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

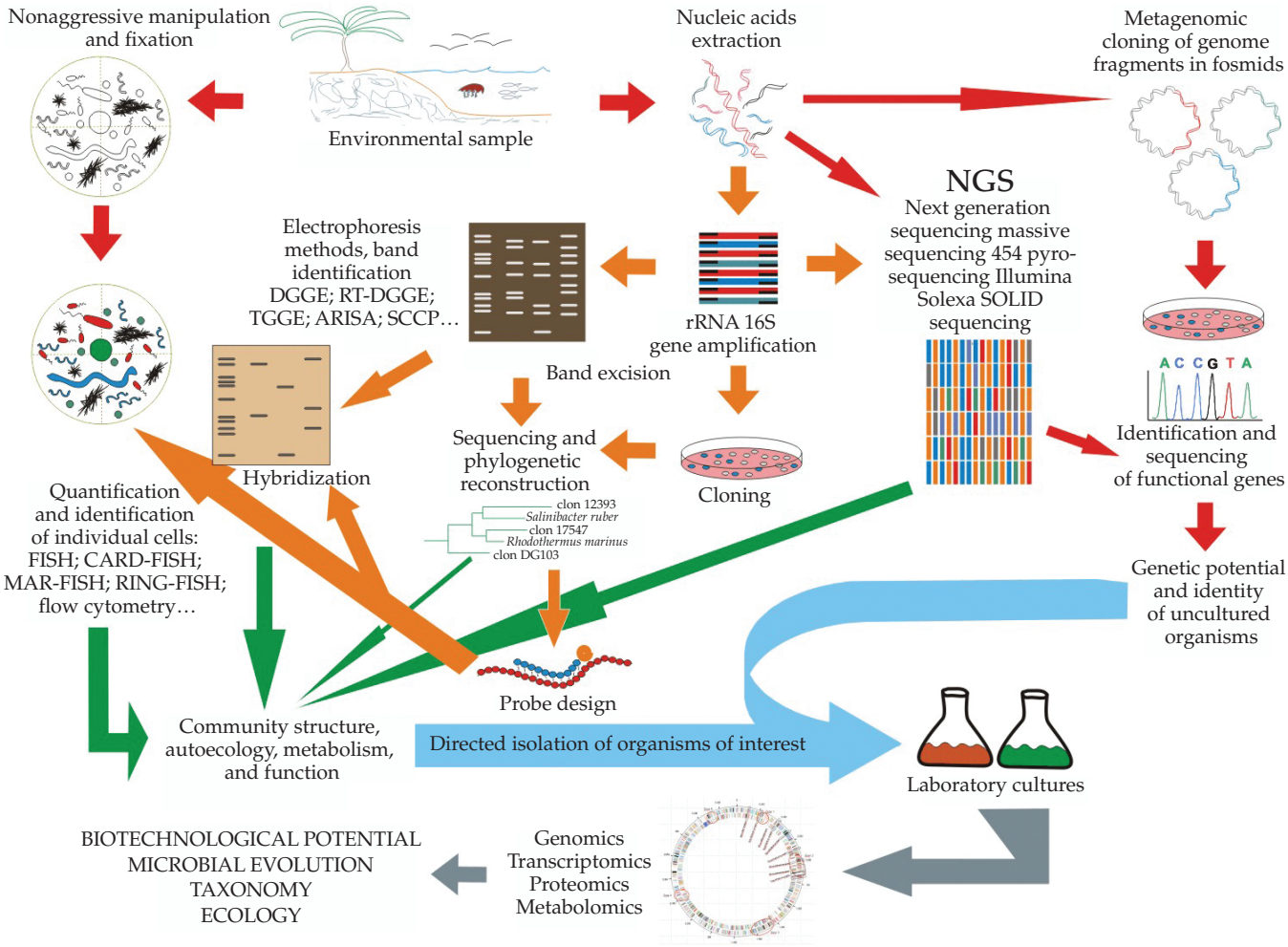


Figure 3.3 Strategy for studying microbial communities. To quantify and identify populations in their natural environment, it has been necessary to develop molecular techniques that combine gene sequencing with probe design. One of the final goals of these studies is to grow organisms in the laboratory with significant metabolic, genetic, or ecological capabilities.

organizing and objectively analyzing the classification of Archaea and Bacteria. One is the International Journal of Systematic and Evolutionary Microbiology, in which most new classifications are published and the names validated in the form of a list of classifications made in other specialized journals, such as *Systematic and Applied Microbiology* and *Archives of Microbiology*. The other compilation of Archaea and Bacteria is the *Bergey's Manual of Systematic Bacteriology*, now in its second edition.

As mentioned in the introduction, the classification criteria have been developing empirically, in parallel to the developments of the available identification methods. Nevertheless, experience has been making the criteria objective and also consolidating classification techniques and parameters, such as DDH and rRNA sequencing, among others. Within the framework of the International Union of Microbiological Societies, an **International Committee of Prokaryote Systematics** (**ICSP**) has been formed, which has taken responsibility for evaluating the validity of classifications and managing the development of classification guidelines. Ultimately, despite the lack of an official body with authority over the taxonomy of Archaea and Bacteria, a consensus has indeed been reached among microbiologists on how this is to be governed. At this time, the most up-to-date and reliable information on classification is found on www.bacterio.net.

Among the general requirements for any new classification is the obligation to deposit a live specimen for each species in two independent culture collections in two different nations. The most representative strain of each species, selected by the investigator as type material, must be deposited with organizations responsible for storing, perpetuating, and providing other researchers with the living material in their custody. Culture collections play a fundamental role both in guaranteeing the availability of type cultures to other researchers and in preserving the biological material in optimal conditions. The obligation to make depos-

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

its of these strains, which remain publicly available, guarantees that the parameters established in the classification can be confirmed by other scientists. These microbe banks are essential, because they help to guarantee the stability of the classification schema and permit reevaluation of already established classifications.

The Domain Archaea

Of the two domains of microorganisms with prokaryotic cellular organization, Archaea is the one with the least taxa to date, encompassing 4% of all described species of "prokaryotes." Its name comes from the fact that the majority of cultured members in this category are organisms that live in extreme (from an anthropocentric point of view) environments, and many of them carry out a chemolithoautotrophic metabolism characterized by oxidation–reduction reactions of sulfur compounds, a metabolism that is thought to be related to the origin of life. In fact, the Archaea include the most extreme thermophilic organisms known. For example, *Pyrolobus fumarii*, with an optimum growth between 90 and 113°C, is one of the most themophilic organisms known. Moreover, this domain also includes an entire class of halophilic organisms (Halobacteria) whose members grow only in hypersaline conditions—many of them even in saturated salt concentrations. The metabolism of the members of Archaea is diverse, ranging from chemolithoautotrophy, in which compounds such as iron and sulfur can be used either as donors or receptors of electrons (anaerobic respiration), to the strict heterotrophy of the Halobacteria, which are strict aerobes. Furthermore, a nonchlorophyllic photosynthesis based on rhodopsins was described for the first time in these last organisms. For a long time, this was considered a characteristic exclusive to this domain, but in the last decade, members of Bacteria that also have this capability have been described, such as several marine bacteria and the extreme halophile *Salinibacter ruber*. A unique metabolism in Archaea is that of methanogenesis, an energy-obtaining process resulting in the formation of methane as a by-product. Methanogenic organisms, called the methanogens, are strict anaerobes and very frequently autotrophs.

Archaea is formed by three phyla, of which two, Crenarchaeota and Euryarchaeota (three, if we consider the Thaumarchaea; see next paragraph), include both cultured and to-date noncultured microorganisms that have been detected in natural samples using molecular techniques (**Figure 3.4**). The third phylum, Korarchaeota, is formed only by sequences that have been obtained by means of molecular ecology techniques.

The Crenarchaeota includes most of the hyperthermophilic organisms, among them the most extreme examples. Many consider them organisms with metabolisms very similar to those the first microorganisms to populate the Earth must have had. With a few exceptions, the cultured members of this phylum are obligate anaerobes, having either chemolithoautotrophic or heterotrophic metabolisms (or both, as in the case of *Sulfolobus*, which can switch between them). However, when the techniques of molecular microbial ecology began to be used in the early 1990s, the abundant presence of Crenarchaeota in marine waters and soils became evident, casting doubt on this domain's extremophilic character. In 2005, isolation and pure culture of *Nitrosopumilus maritimus*, the first non-extremophile Crenarchaeota, was achieved. Isolated from the Seattle Aquarium, it has both an autotrophic and a chemolithotrophic metabolism, using ammonium as a source of energy. This discovery, together with the estimated contribution of Archaea to the fixation of CO_2 in oceans, has demonstrated the importance of this group of microorganisms in an environment that is not considered extreme. In 2008, a new phylum of deep-branching

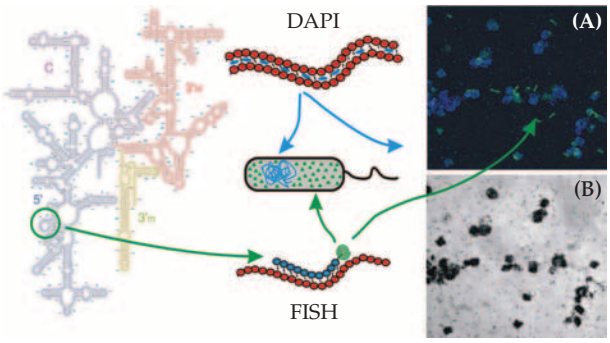


Figure 3.4 Identification of cells by fluorescence in situ hybridization (FISH) and observation of metabolism by means of radioactive isotopes. To determine the identity of the microorganisms present in a sample, we turn to the use of fluorochrome-labeled phylogenetic probes. (A) All biological structures that stain blue with DAPI, a fluorescent compound that binds with double-stranded DNA, are identified as microorganisms. Cell identity is determined with probes specifically designed and complementary to their ribosomal RNA. Each ribosome binds with a fluorochrome-labeled probe. The intensity of the light emitted in each cell in which a positive hybridization is produced is proportional to the number of ribosomes present. Due to the interspecies variability of these sequences, probes can be designed targeting taxonomic groups at different hierarchical levels: domain, family, genus, species, etc. The secondary structure of the sequence of 16S rRNA can be found at rna.ucsc.edu/rnacenter/ribosome_images.html. (B) Additionally, FISH can be combined with microautoradiography; the community is supplied with a radioactive-marked substrate (generally ^{14}C), which is incorporated into the biomass of the microorganisms consuming it. In this technique, the preparation of the sample for FISH incorporates a photographic film that allows radioactive impregnation. The cells that have incorporated the ^{14}C are shown with a silver precipitate. In this manner, identity and metabolism are learned simultaneously.

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

archaeans was proposed, the Thaumarchaea, which included those lineages of nitrifying archaeans previous affiliated with the Crenarchaeota. Members of this group have been detected in a wide variety of mesophilic and extreme ecosystems, including marine and fresh waters, soils, and hot environments.

The Euryarchaeota includes microorganisms of very diverse metabolisms and ecological niches. All of the strict halophilic organisms forming the class of Halobacteria are in this phylum. These are considered to be the most evolved of the Euryarchaeota. They have an aerobic heterotrophic metabolism, although some may obtain energy through a proton pump activated by light: bacteriorhodopsin. This photosynthetic system without chlorophyll is the simplest known mechanism for using solar energy. The most abundant and most significant group of Euryarchaeota, both in numbers of species and in environmental relevance, is the group of methanogens, characterized by generating methane (CH_4) as a byproduct of their metabolism. It is estimated that 500 million tons of this gas, with its enormously powerful greenhouse effect, are released into the atmosphere per year. Most of the methane is produced in the subsoil, in anaerobic conditions. However, a frequent habitat of methanogens, in addition to subsoil, is the gut of mammals, especially ruminants. There are also methanogens among the inhabitants of the gut in some humans.

One of the most intriguing groups within the Euryarchaeota is the group of Thermoplasmatales. This order contains thermophilic organisms without cell walls, such as *Ferroplasma* and *Thermoplasma*, or acidophiles such as *Picrophilus*, which can grow in extremely acid environments with a pH as low as 0.06. Similarly, within the same phylum there are significant orders of hyperthermophilic microorganisms such as Thermococcales and Archaeoglobales, with anaerobic chemolithoautotrophic metabolism as a whole. Finally, as occurred with the Crenarchaeota, organisms in the Euryarchaeota phylum have been found, through independent culturing techniques, to be widely distributed in marine environments.

The phylum of Korarchaeota is formed only by environmental sequences from genomes of the gene for 16S rRNA, obtained from DNA extracted from the Obsidian Pool hot spring in Yellowstone Park. In spite of consisting of only environmental sequences, mixed cultures of Korarchaeota have been achieved in the laboratory, where it coexists with other anaerobic thermophiles.

The Domain Bacteria

Bacteria is the broadest domain and includes most species and genera with a prokaryotic cellular organization. At this time, there are 26 taxonomically recognized phyla in Bacteria. However, at least 14 additional phyla have been observed that are represented only by environmental sequences that do not correspond to those of any laboratory-cultured microorganism. Currently, in the database of sequences of the 16S rRNA gene, there are over 3,500,000 different entries for members of this domain. These represent 95% of the "prokaryote" sequences and 80% of the total entries in the small ribosome subunit present in public databases (including those that correspond to the eukaryotes). In any case, the majority of these sequences are environmental clones for which there is no cultured representative. From observation of this significant diversity of sequences, it is inferred that the potential number of microorganisms still to be described is enormous.

As Box 3.2 shows, almost 90% of classified genera in the domain are concentrated within four phyla: Proteobacteria (44%), Firmicutes (18%) and Actinobacteria (15%)—these two are both Gram-positive bacteria—and Bacteroidetes (10%). One of the reasons for such an imbalance in the relationship among taxa is that most of the culture media traditionally used for studying microorganisms were designed in their day for culturing organisms pathogenic or saprophytic to humans and animals, which are much more highly represented in these four phylogenetic groups.

The domain Bacteria is very diverse in reference to its metabolism, genetics, and ecology. Members of this domain are found colonizing practically all environments where life is possible. This is due in part to this group's great metabolic diversity, in which organisms exist with extremely different mechanisms for obtaining nutrients and energy. In fact, unlike Archaea, the only metabolism not present in this domain is methanogenesis. However, only in Bacteria are there microorganisms capable of performing chlorophyll-based photosynthesis, whether oxygenic or anoxygenic. In fact, the photosynthetic capability of some eukaryotes is due to the presence of chloroplasts in their cells, which are definitely endosymbiotic Cyanobacteria.

The most significant bacterial taxa, whether for the number of taxa they include, or for their physiological or ecological characteristics, are discussed below.

Proteobacteria

This is perhaps the most important phylum from a clinical viewpoint, and the one that covers most of the bacterial species described. It was originally called the "phylum of purple bacteria" because it includes anoxygenic photosynthetic organisms (with a single photosystem in which water is not hydrolyzed) such as *Rhodospirillum*, whose pigments give them a purple or reddish color. This phylum of Gram-negative bacteria is subdivided into five subclasses (alpha, beta, gamma,

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

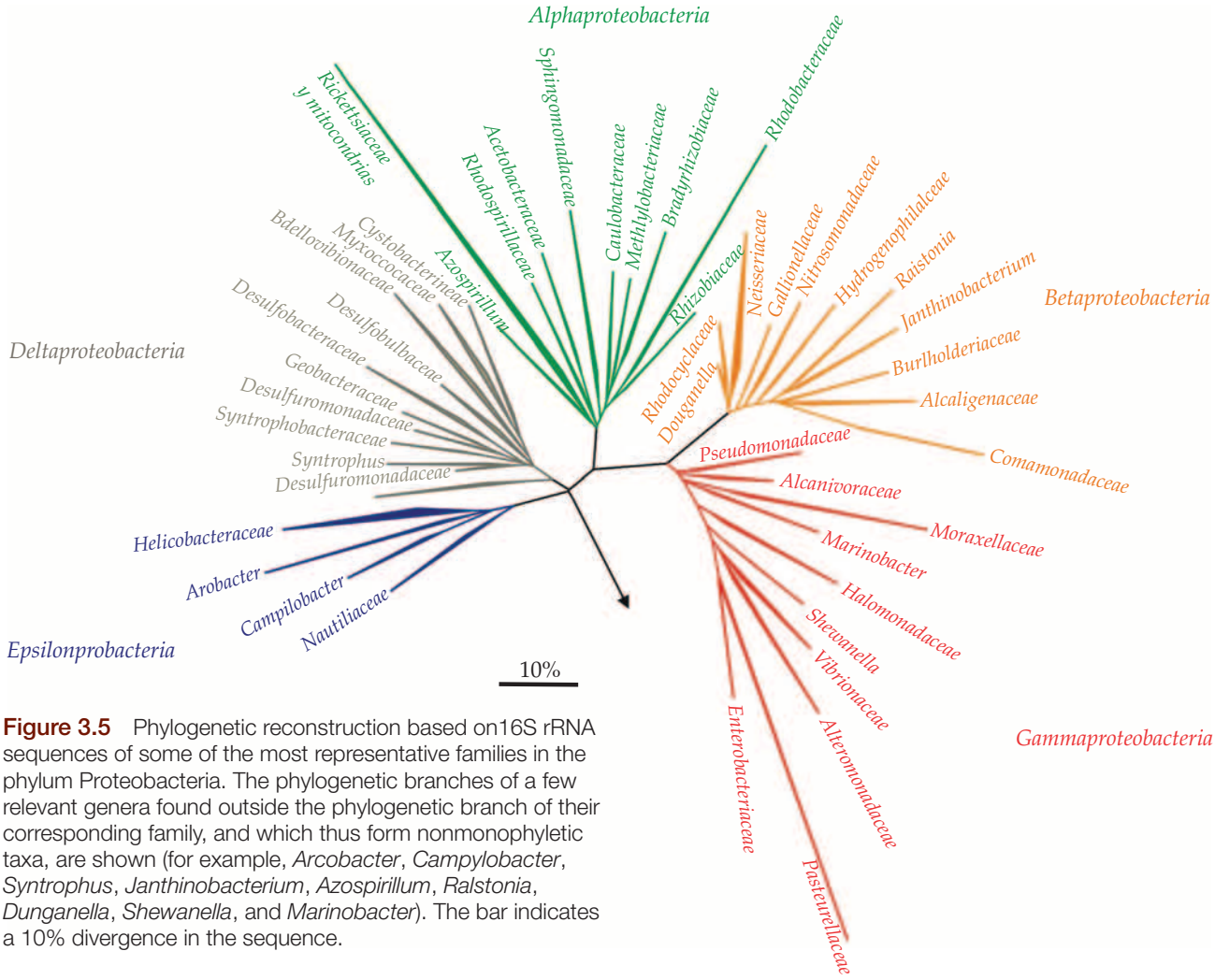


Figure 3.5 Phylogenetic reconstruction based on16S rRNA sequences of some of the most representative families in the phylum Proteobacteria. The phylogenetic branches of a few relevant genera found outside the phylogenetic branch of their corresponding family, and which thus form nonmonophyletic taxa, are shown (for example, *Arcobacter*, *Campylobacter*, *Syntrophus*, *Janthinobacterium*, *Azospirillum*, *Ralstonia*, *Dunganella*, *Shewanella*, and *Marinobacter*). The bar indicates a 10% divergence in the sequence.

delta, and epsilon; **Figure 3.5**). Its members are found colonizing all types of environments, from sea water to the guts of animals; and considered as a group, their metabolisms vary greatly, from photoautotrophy (for example, *Chromatium*) and chemolithoautotrophy (e.g., *Ectothiorhodospira*) to photoheterotrophy (e.g., *Rhodospirillum*), and strict heterotrophy (e.g., *Escherichia*). In this group are organisms especially relevant to biogeochemical cycles, such as nitrogen-fixing microbes (e.g., *Rhizobium*), ammonia-oxidizing bacteria (e.g., *Nitrosomonas*), and bacteria that oxidize sulfur (e.g., *Acidithiobacillus*, which is also capable of growing at pH 2). There are proteobacteria that can respire different inorganic compounds, such as sulfates (*Desulfovibrio*), iron (*Shewanella*), nitrates (*Pseudomonas*), and manganese (*Arcobacter*), and there are strict fermenters (*Enterobacter*). Major human pathogens are also found in this group (*Salmonella*, *Neisseria*, etc.), animal pathogens (*Campylobacter*), and plant pathogens (*Xhanthomonas*). Organisms in this phylum can also be found associated as endosymbionts in animals (as is the case of the oligochaete *Olavius algarvensis*), or in protists (as is the case of *Paramecium caudatum*). In a large part of these symbioses, the symbionts' contribution to the host metabolism is essential. The classic pattern of endosymbiosis of a proteobacterium with a eukaryote would be that of mitochondria, whose origin is clearly localized within the group of Alphaproteobacteria.

Cyanobacteria

This is a very homogeneous phylum from a metabolic standpoint. It consists of obligate and autotrophic oxygenic photosynthetic organisms, with two photosystems and membrane systems in the form of thylakoids, where the systems of capturing and channeling light energy are contained. This phylum is, however, very heterogeneous in regard to its members' morphological characteristics. In fact, the classification of Cyanobacteria was under the authority of botanists until the late twentieth century and, in contrast to the rest of Archaea and Bacteria, it is fundamentally based on observable morphological traits. Cyanobacteria display a huge diversity in shapes, from unicellu-

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

lar spheres (*Gloeothece*), to spiral-shaped multicellular chains (*Arthrospira*) and branched filaments (*Fischerella*). This morphological abundance has, not without a great deal of controversy, permitted identification of the oldest microbial fossils, dating from 3500 million years ago, as members of this group. Furthermore, and because they are the only group of Archaea and Bacteria capable of hydrolyzing water to free molecular oxygen, they are thought to be the cause of the increase in oxygen concentration in the atmosphere occurring during the Archeozoic and Proterozoic periods of the Precambrian. It is proposed that this led to formation of the ozone layer, which in turn led to the colonization of the terrestrial surface of the earth, and the massive development of aerobic life in the biosphere. Additionally, many Cyanobacteria can also fix nitrogen; thus some of them, such as the filamentous *Trichodesmium*, play an important role in the availability of nitrogen in the ocean. A member of this phylum, *Prochlorococcus*, is the most abundant photosynthetic organism on the planet and one of the principal fixers of CO_2 in oligotrophic marine environments. Finally, this phylum includes the microorganism considered to have given rise to chloroplasts in plants through endosymbiosis. If the chloroplast is considered a cyanobacterium, the capacity to photosynthesize is exclusive to the Bacteria.

Chlorobi

Among the 16S rRNA sequences assigned to this phylum, the percentage corresponding to cultured microorganisms is relatively small (six genera). These obligate and anoxygenic photosynthetic organisms are in general autotrophs, although some are capable of photoheterotrophic growth. Like the photosynthetic Proteobacteria, they use H_2S as a source of electrons. They are ubiquitous organisms, frequently associated with oxic–anoxic transition zones in lakes and seas. Normally these bacteria require light of very low intensity; therefore they occupy zones in which, although solar radiation still penetrates, other phototrophic organisms cannot grow. In some cases it has been observed that they can form consortia with sulfate-reducing Proteobacteria (for example, *Chlorochromatium aggregatum*), acting as a biological unit.

Bacteroidetes

This phylum is formed by a ubiquitous, heterogeneous, and very diverse group of organisms that colonize natural environments from water and marine sediments to the intestinal tracts of animals. The genus *Bacteroides*, which gives the phylum its name, is a group of obligate anaerobes, fermenters frequently found in the gut of humans and other animals. In fact, Bacteroidetes, together with Firmicutes, are the most abundant groups of organisms in the human gut. However, other groups, such as the Cytophaga or Flavobacteria, are mostly aerobic organisms and identified in general as polymer-degrading microbes. They are producers of exoenzymes that take part in the primary degradation of macromolecules in aerobic aquatic environments; for that reason they are important components of plankton. Moreover, using molecular techniques, it has been possible to see members of these groups appearing in marine sediments, where they can be carrying out fermentation processes. Finally, by means of molecular techniques and culturing, Bacteroidetes (e.g., *Salinibacter*) have been detected in hypersaline environments around the world, where the presence of members of the bacterial domain had been considered irrelevant in comparison to the Archaea. In general, data collected through the study of different ecosystems indicate that Bacteroidetes could be more environmentally relevant than had first been expected.

Firmicutes

This phylum, together with Actinobacteria, forms the group of Gram-positive organisms. For almost a century, Bacteria were classified into two large groups, the Gram-positive and the Gram-negative, according to their different responses to the Gram stain. This stain was developed by Hans Christian Gram in 1884 to show differential characteristics in the cell walls of bacteria. Every organism whose peptidoglycan layer can be stained with crystal violet is considered Gram-positive, while organisms that cannot retain this colorant in their walls after a decolorizing wash with alcohol are Gram-negative. Until the development of molecular phylogeny, Gram-positive and Gram-negative bacteria were considered two differentiated taxa in the Bacteria. Now, two monophyletic phyla are recognized (Firmicutes and Actinobacteria) that include all Gram-positive bacteria, while the Gram-negatives are distributed throughout the rest of the phyla. Firmicutes is also a very diverse group of organisms, and also ubiquitous and especially relevant in soils and anaerobic environments. Many of them are fermenters and some are obligate anaerobes (such as *Clostridium* and *Heliobacterium*). Included in this group are the majority of the endospore-forming organisms. Endospores are differentiated cells capable of withstanding extremely adverse environmental conditions, such as desiccation and high temperatures. The Firmicutes include organisms important to the food industry, such as the lactic bacteria (*Lactobacillus* and *Leuconostoc*). They are very important in transforming and conserving foods, such as dairy products and cured meats, among others. Nevertheless, the same phylum also contains very significant pathogenic organisms, such as *Streptococcus* (the cause of scarlet fever and sore throats), *Clostrid-*

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

ium (the cause of tetanus and botulism), and *Bacillus anthracis* (the cause of anthrax and known as a potential bacteriological weapon). However, there are also organisms of great biotechnological interest among them, such as *Bacillus thuringiensis*, which synthesizes insecticidal compounds that are used extensively as biological insecticides. Within this phylum is the order of Mollicutes, characterized by the lack of a cell wall. These are considered the organisms with the smallest genome within the domain that have the capacity for autonomous growth. Their genome is smaller than that of most Archaea and Bacteria (0.58 Mb in the case of *Mycoplasma genitalium*), even smaller than that of parasitic or obligate symbiont organisms such as *Chlamydia* and *Rickettsia*. Mollicutes are especially found distributed as pathogens, commensals, and even symbionts of plants and animals, and some of them, such as *Spiroplasma*, appear to cospeciate with their insect hosts.

Actinobacteria

Actinobacteria are also known as high G+C Gram-positives, although very recently low G+C Actinobacteria, yet uncultured, have been discovered in environmental surveys. They have a thick and rigid cell wall, and some produce special waxes and lipids (for example, mycolipids in *Mycobacterium*), which confer a high degree of hydrophobicity to the cell. They are generally aerobic, ubiquitous organisms; many of them are important in soils (e.g., *Arthrobacter* and *Streptomyces*). The name of the phylum comes from the genera *Actinomyces* and *Streptomyces*, which have a filamentous shape, are very similar to fungal mycelia, live primarily in soils, and in general, produce antibiotics (*Streptomyces*, with more than 500 species described, is an important producer). Although the majority of members of this phylum are inhabitants of soil, important human pathogens are also included in the group, such as the causal agent of diphtheria (*Corynebacterium*) and those of tuberculosis and leprosy (*Mycobacterium*). Likewise, the phylum also holds innocuous and ubiquitous organisms frequently suspended in air, such as *Microbacterium*.

Spirochaetes

This is a very homogeneous phylum with regard to morphology and cellular structure. As its name indicates, spirochetes have a spiral shape, occasionally with narrowly coiled spirals. They are very long organisms (some 0.75 μm in diameter and 250 μm long), and their motility presents some peculiar characteristics, due to the presence of endoflagella in the periplasm. The phylum includes both free-living microorganisms in sediments and aquatic environments, animal parasites, and commensals. Among them are important pathogens such as *Treponema pallidum* (responsible for syphilis), *Borrelia* (responsible for Lyme disease or recurrent fevers transmitted through ticks), and *Leptospira* (responsible for leptospirosis). Members of this phylum had been found as symbionts of molluscs and protozoa, and saprophytes in the oral human microbiota.

Planctomycetes

This phylum owes its name to being originally created in order to embrace certain marine planktonic bacteria. However, Planctomycetes have subsequently been found in great numbers in soil. The members of this phylum lack peptidoglycan in their cell walls, and they frequently have morphologically diverse cell structures, with spicules that permit them to be easily identified. In some cases, they have complex life cycles. They also generally exhibit intracellular compartmentation. It is the only group of Archaea and Bacteria (i.e., *Gemmata*) that has the genome enclosed in a compartmented structure that is reminiscent of the nucleus of eukaryotes. This group includes a microorganism capable of oxidizing ammonia in the absence of oxygen, a process of great ecological relevance.

Identification and Affiliation of Uncultured Organisms

As has been stated, most of the Archaea and Bacteria in the biosphere have yet to be adequately described, due mainly to the need to obtain pure cultures in the laboratory. These microorganisms are found in relatively dense communities, in which the number of distinct genomes per cubic centimeter can range from 150 to over 4000. This great diversity, together with the reduced size of the microorganisms and the scant information their morphological characters provide, makes direct microscopic observation inadequate for ecological studies. The development of molecular biology, and especially of techniques utilizing 16S rRNA, has permitted understanding of the structure and dynamics of microbial communities in natural environments (see Figure 3.3). The design of phylogenetic probes targeting the ribosome directly has been used to quantify taxa in natural samples through techniques such as **northern blot** and fluorescence microscopy (see Figure 3.4). The fact that it is possible to find regions in the 16S rRNA sequence exclusive to the group of organisms to be identified has been put to advantage in the design of specific probes. In this way, it has been possible to design probes targeting members of each one of the domains (*Bacteria*, *Archaea*, and *Eucarya*), and probes restricted to taxa in distinct hierarchical ranks. Probes have been designed in this way, for example, for different subclasses of Proteobacteria, for different genera of sulfate-reducers,

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

and for phylogenetic groups that were never cultured and were represented only by sequences in the form of clones. In general, and especially through the use of fluorescent probes in situ (**fluorescence in situ hybridization**, or **FISH**), natural microbial communities of very diverse origin have been quantified. But there are basically two problems that hamper the development of phylogenetic probe technology: the first lies in their specificity, and the second in the quality of the sequences. The specificity of probes designed is analyzed against the database available at that time. In the last 20 years, the database of 16S rRNA sequences has grown from 473 sequences to more than over 3,500,000 (according to www.arb-silva.de) in a process that has paralleled the development and interest in the specific molecular microbial ecology. This yearly increase of thousands of deposited sequences (currently about 700,000 per year; see Box 3.2) means that many probes that had once been considered exclusive for a phylum might be unspecific, and therefore, they need to be designed once again. However, one of the greatest problems lies in the fact that many of the sequences deposited are of low quality and generally short, with a high number of indeterminations. In fact, 50% of all sequences can be considered of insufficient quality for any phylogenetic or ecological study. Because of this, the useful database in reality has approximately 1,500,000 sequences as of 2014.

In spite of the difficulties implicit in the growth of the database, the use of more novel molecular techniques has revealed the morphological, ecological, and in some cases, metabolic characteristics of organisms that have not yet been cultured. For these organisms, from which a minimum of characters can be detected that guarantee their identification, a provisional status has been established within the taxonomic scheme of the Archaea and Bacteria. They are assigned a status of *candidatus*, with a provisional name that would be validated in the case of a pure culture being obtained. This status has been used to provisionally classify the largest unicellular microorganisms known to date (e.g., *Thiomargarita namibiensis*), microorganisms with anaerobic chemolithoautotrophic metabolisms, or pathogenic intracellular parasites (e.g., *Rhabdochlamydia porcellionis*). Pure laboratory culture of the extreme halophile *Salinibacter* was also achieved because of the understanding of some of its metabolic parameters, revealed by studies of molecular ecology.

Recently, mainly due to the decrease of sequencing costs and the increase of available bioinformatics tools and platforms for analysis, metagenomic approaches began substituting the rRNA approach to study microbial communities and characterize uncultured microorganisms. Basically, a metagenomic analysis consists of the sequencing of nucleic acids directly extracted from a natural sample. These nucleic acids correspond to the sum of all the genomes (the metagenome) of the organisms contained in the sample. One of the most outstanding achievements of metagenomics was the discovery of the light-harvesting proton pump proteorhodopsin in marine environments, which has turned out to be one of the most ubiquitous and abundant proteins in the ocean. Currently there are several large metagenomic projects underway, including The Human Microbiome Project (commonfund.nih.gov/hmp) and TARA Oceans (oceans.taraexpeditions.org), among others. Finally, recently developed techniques, such as single-cell genomics, are allowing the analysis of uncultured microbes at the level of individual cells.

Biogeography and Speciation

Archaea and Bacteria are ubiquitous and are found colonizing any environment that supports life. One of the most intriguing questions in microbial ecology lies in clarifying whether mechanisms of geographic isolation occur that allow for allopatric speciation of these microorganisms. In 1936, Baas Becking formulated his famous statement, "Everything is everywhere, but the environment selects," an affirmation that has determined the view of microbial diversity around the world; in fact, members of the same species are not infrequently isolated in samples from far distant environments. Genetic tools have been put to use for some years now to learn whether the organisms found in geographically distant locations owe their differences to mechanisms of speciation linked to distance and geographic isolation. These techniques compare the diversity among genomes indirectly, by methods such as **pulsed field gel electrophoresis** (**PFGE**) or **random amplification of polymorphic DNA** (or **RAPD**). Studies have also been conducted on diversity in the sequence of the spacer (**internal transcribed spacer**, or **ITS**) among the genes that code for 16S and 23S rRNA. At present, sequencing and concatenation of the alignment of essential genes is being combined (**multilocus sequence analysis**, or **MLSA**). However, a clear genetic segregation linked to geographic isolation has only been demonstrated for extremophilic organisms, such as *Sulfolobus* (Archaea), and for some thermophilic Cyanobacteria (Bacteria). These microorganisms inhabit environments very restricted in environmental characteristics and geographically widely separated. It remains to be understood whether the dispersion mechanisms of microorganisms are so swift that they do not permit a genetic divergence linked to geographic isolation to be established, or whether the genetic parameters being measured are, themselves, not sufficiently resolute for detecting an allopatric speciation.

One of the most important factors that might be masking the processes of speciation due to geographic

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

isolation is the capacity of Archaea and Bacteria for genetic exchange. The mechanisms of horizontal gene transfer have been known for decades. It has been shown that acquisition of exogenous genes can be performed by means of an exchange of plasmids (conjugation), by active incorporation of free DNA (natural transformation), or by means of infection by phages (transduction). These processes can take place among organisms in very different categories, and in contrast to what occurs in eukaryotes, they are not limited to members of the same species. The turn of the century brought with it a significant effort to sequence complete Archaea and Bacteria genomes. This year, more than 20,000 "prokaryote" genomes have been sequenced and annotated, most of which are of Bacteria (www.ncbi.nlm.nih.gov/genomes/lproks.cgi). In spite of the bias toward the sequencing of pathogenic organisms, with a relatively low proportion of organisms of environmental origin, the first observations have provoked important questions about the processes of speciation. Comparing the composition and the similarities of orthologous genes among genomes, some researchers speculate that the process of genetic exchange could even have breached the borders between domains, and that the genealogy of Archaea and Bacteria could be blurred by considerable gene transfer. Nevertheless, other researchers consider that its importance is being overrated due to a base of data that is still scanty, and that very probably the classical Darwinian model of lineages based on genetic drift, duplications, and deletions is indeed the dominant model of evolution in Archaea and Bacteria. For the purpose of accommodating intraspecies genomic incongruences with the identity of a species, it has been proposed that this identity be conferred through a group of invariable (core) genes in the taxon, which would be accompanied by a group of auxiliary genes whose presence would not be necessary for the organism's identity. A true understanding of the importance of horizontal gene transfer on microbial evolution will have to wait for the sequencing of many more genomes.

Basic Bibliography

Amann, R., Ludwig, W., and Schleifer, K. H. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews* 59: 143–169.

Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H., and Stackebrandt, E. 2007. *The Prokaryotes*, 3rd edition. New York, NY: Springer-Verlag.

Garrity, G. M., Boone, D. R., and Castenholz, R. W. 2001. *Bergey's Manual of Systematic Bacteriology*, 2nd edition. New York, NY: Springer-Verlag.

Lapage, S. P., Sneath, P. H. A., Lessel, E. F., Skerman, V. B. D., Seeliger, H. P. R., and Clark, W. A. (eds.). 1992. *International Code of Nomenclature of Bacteria*, 1990 revision. Washington, DC: ASM Press.

Whitman, W. B., Coleman, D. C., and Wiebe, W. J. 1998. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America* 9: 6578–6583.

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.

UNCORRECTED PAGE PROOFS © 2014 Sinauer Associates, Inc. This material cannot be copied, reproduced, manufactured or disseminated in any form without express written permission from the publisher.