16

Microbial Evolution and Systematics

Fluorescent dyes bound to specific nucleic acid probes can differentiate cells in natural samples that are morphologically similar but phylogenetically distinct.

A unifying theme in all of biology is **evolution**. By deploying its major tools of descent through modification and selection of the fittest, evolution has affected all life on Earth, from the first self-replicating entities, be they cells or otherwise, to the modern cells we see today. Since its origin, Earth has undergone a continuous process of physical and geological change, eventually establishing conditions conducive to the origin of life. After microbial life appeared, Earth continued to present it with new opportunities and challenges. As microbial metabolisms and physiologies evolved in response, microbial activities changed planet Earth in significant ways to yield the biosphere we see today.

This chapter focuses on the evolution of microbial life, from the origins of the earliest cells and metabolisms to the microbial diversity we see today. Methods for discerning evolutionary relationships among modern-day descendants of early microbial lineages are a major theme. Overall, the goal of this chapter is to provide an evolutionary and systematic framework for the diversity of contemporary microbial life that we will explore in the next four chapters.

I Early Earth and the Origin and Diversification of Life

In these first few sections, we consider the possible conditions under which life arose, the processes that might have given rise to the first cellular life, its divergence into two evolutionary lineages, ***Bacteria*** and ***Archaea***, and the later formation, through endosymbiosis, of a third lineage, the ***Eukarya***. Although much about these events and processes remains speculative, geological and molecular evidence has combined to build a plausible scenario for how life might have arisen and diversified.

16.1 Formation and Early History of Earth

Before considering how life arose, we need to go back even farther, and ask how Earth itself formed.

Origin of Earth

Earth is thought to have formed about 4.5 billion years ago, based on analyses of slowly decaying radioactive isotopes. Our planet and the other planets of our solar system arose from materials making up a disc-shaped nebular cloud of dust and gases released by the supernova of a massive old star. As a new star—our sun—formed within this cloud, it began to compact, undergo nuclear fusion, and release large amounts of energy in the form of heat and light. Materials left in the nebular cloud began to clump and fuse due to collisions and gravitational attractions, forming tiny accretions that gradually grew larger to form clumps that eventually coalesced into planets. Energy released in this process heated the emerging Earth as it formed, as did energy released by radioactive decay within the condensing materials, forming a planet Earth of fiery hot magma. As Earth cooled over time, a metallic core, rocky mantle, and a thin lower-density surface crust formed.

The inhospitable conditions of early Earth, characterized by a molten surface under intense bombardment from space by asteroids and other objects, are thought to have persisted for over 500 million years. Water on Earth originated from innumerable collisions with icy comets and asteroids and from volcanic outgassing of the planet's interior. At this time, due to the heat, water would have been present only as water vapor. No rocks dating to the origin of planet Earth have yet been discovered, presumably because they have undergone geological metamorphosis. However, ancient sedimentary rocks, which formed under liquid water, have been found in several locations on Earth. Some of the oldest sedimentary rocks discovered thus far are in southwestern Greenland; these rocks date to about 3.86 billion years ago. The sedimentary composition of these rocks indicates by that time Earth had at least cooled sufficiently ($<100°C$) for the water vapor to have condensed and formed the early oceans.

Even more ancient materials, crystals of the mineral zircon ($ZrSiO_4$), however, have been discovered, and these materials give us a glimpse of even earlier conditions on Earth. Impurities trapped in the crystals and the mineral's isotopic ratios of oxygen (Section 22.8) indicate that Earth cooled much earlier than previously believed, with solid crust forming and water condensing into oceans perhaps as early as 4.3 billion years ago. The presence of liquid water implies that conditions could have been compatible with life within a couple of hundred million years after Earth was formed.

Evidence for Microbial Life on Early Earth

The fossilized remains of cells and the isotopically "light" carbon abundant in these rocks provide evidence for early microbial life (we discuss the use of isotopic analyses of carbon and sulfur as indications of living processes in Section 22.8). Some ancient rocks contain what appear to be bacteria-like microfossils, typically simple rods or cocci (**Figure 16.1**).

In rocks 3.5 billion years old or younger, microbial formations called **stromatolites** are common. Stromatolites are microbial mats consisting of layers of filamentous prokaryotes and trapped

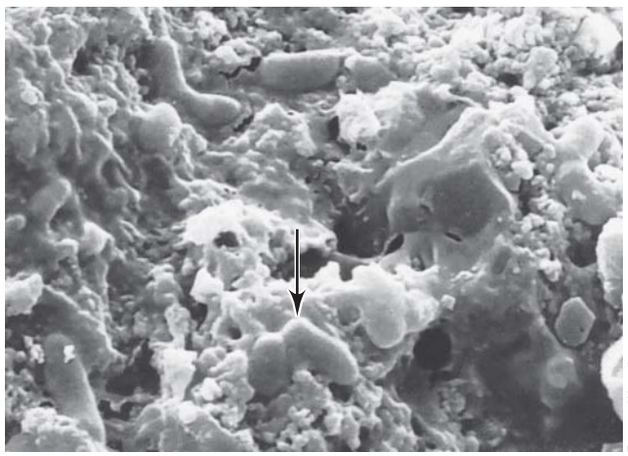


Figure 16.1 **Ancient microbial life.** Scanning electron micrograph of microfossil bacteria from 3.45 billion-year-old rocks of the Barberton Greenstone Belt, South Africa. Note the rod-shaped bacteria (arrow) attached to particles of mineral matter. The cells are about 0.7 μm in diameter.

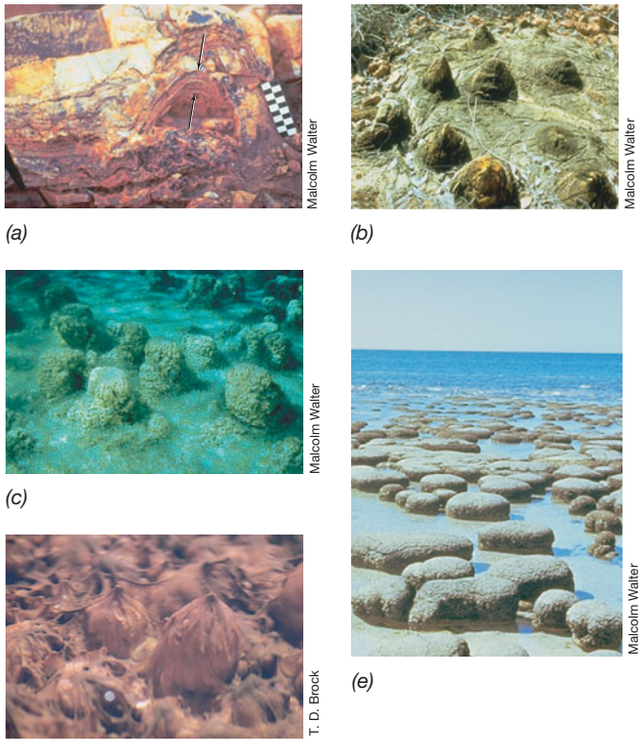


Figure 16.2 Ancient and modern stromatolites. *(a)* The oldest known stromatolite, found in a rock about 3.5 billion years old, from the Warrawoona Group in Western Australia. Shown is a vertical section through the laminated structure preserved in the rock. Arrows point to the laminated layers. *(b)* Stromatolites of conical shape from 1.6 billion-year-old dolomite rock from northern Australia. *(c)* Modern stromatolites in Shark Bay, Western Australia. *(d)* Modern stromatolites composed of thermophilic cyanobacteria growing in a thermal pool in Yellowstone National Park. Each structure is about 2 cm high. *(e)* Another view of modern and very large stromatolites from Shark Bay. Individual structures are 0.5–1 m in diameter.

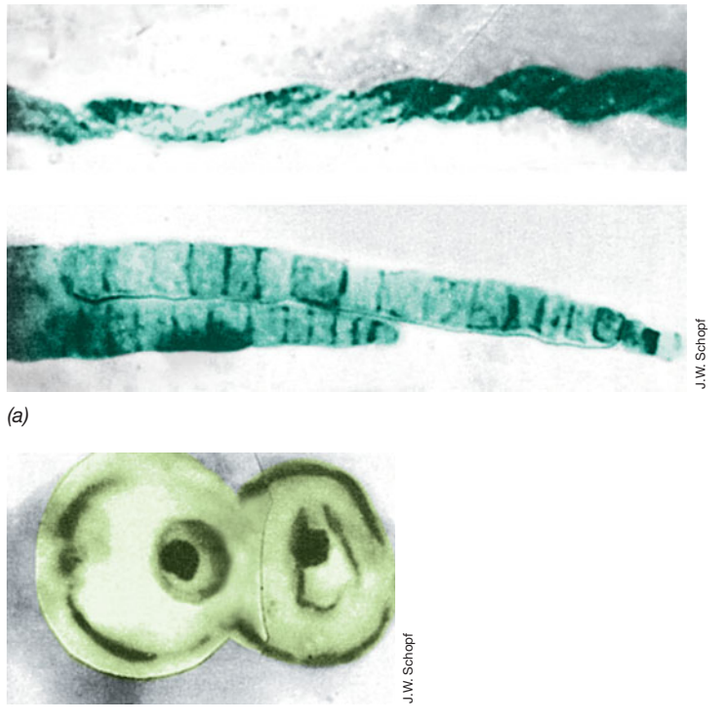


Figure 16.3 More recent fossil bacteria and eukaryotes. *(a)* One billion-year-old microfossils from central Australia that resemble modern filamentous cyanobacteria. Cell diameters, 5–7 μm. *(b)* Microfossils of eukaryotic cells from the same rock formation. The cellular structure is similar to that of certain modern green algae, such as *Chlorella* species. Cell diameter, about 15 μm. Color was added to make cell form more apparent.

mineral materials; they may become fossilized (**Figure 16.2*a*, *b***) (we discuss microbial mats in Section 23.5). What kind of organisms were these ancient stromatolitic bacteria? By comparing ancient stromatolites with modern stromatolites growing in shallow marine basins (Figure 16.2*c* and *e*) or in hot springs (Figure 16.2*d*; ↻ Figure 23.9*b*), we can see it is likely that ancient stromatolites formed from filamentous phototrophic bacteria, such as ancestors of the green nonsulfur bacterium *Chloroflexus* (↻ Section 18.18). **Figure 16.3** shows photomicrographs of thin sections of more recent rocks containing microfossils that appear remarkably similar to modern species of cyanobacteria and green algae, both of which are oxygenic phototrophs (↻ Sections 18.7 and 20.20). The age of these microfossils, about 1 billion years, is well within the time frame that such organisms were thought to be present on Earth (↻ Figure 1.6).

In summary, microfossil evidence strongly suggests that microbial life was present within at least 1 billion years of the formation of Earth and probably somewhat earlier, and that by that time, microorganisms had already attained an impressive diversity of morphological forms. We tackle the issue of how life first evolved from nonliving materials in the next section. But regardless of when self-replicating life forms first appeared, the process of evolution began at the same time, selecting for improvements that would eventually lead to microbial cells' inhabiting every ecosystem on Earth that was chemically and physically compatible with life.

MiniQuiz

- How did planet Earth form?
- What evidence is there that microbial life was present on Earth 3 billion years ago?
- What do crystals of the mineral zircon tell us about conditions for early life?

16.2 Origin of Cellular Life

Here we consider the issue of how living organisms might have originated, focusing on two questions: (1) How might the first cells have arisen? (2) What might those early cells have been like? But along the way, we will consider the likely possibility that self-replicating RNAs preceded cellular life and how these molecules may have paved the way for cellular life.

Surface Origin Hypothesis

One hypothesis for the origin of life holds that the first membrane-enclosed, self-replicating cells arose out of a primordial soup rich in organic and inorganic compounds in a "warm little pond," as Charles Darwin suggested in *On the Origin of Species*—in other words, life arose on Earth's surface. Although there is experimental evidence that organic precursors to living cells can form spontaneously under certain conditions, surface conditions on early Earth are thought to have been hostile to both life and its inorganic and organic precursors. The dramatic temperature fluctuations and mixing resulting from meteor impacts, dust clouds, and storms, along with intense ultraviolet radiation, make a surface origin for life unlikely.

Subsurface Origin Hypothesis

A more likely hypothesis is that life originated at hydrothermal springs on the ocean floor, well below Earth's surface, where conditions would have been much less hostile and much more stable. A steady and abundant supply of energy in the form of reduced inorganic compounds, for example, hydrogen (H_2) and hydrogen sulfide (H_2S), would have been available at these spring sites. When the very warm (90–100°C) hydrothermal water flowed up through the crust and mixed with cooler, iron-containing and more oxidized oceanic waters, precipitates of colloidal pyrite (FeS), silicates, carbonates, and magnesium-containing montmorillonite clays formed. These precipitates built up into structured mounds with gel-like adsorptive surfaces, semipermeable enclosures, and pores. *Serpentinization*, the abiotic process by which Fe/Mg silicates (serpentines) react with other minerals and H_2, was a likely source of the first organic compounds, such as hydrocarbons and fatty acids. These could then have reacted with iron and nickel sulfide minerals to eventually form amino acids, simple peptides, sugars, and nitrogenous bases (**Figure 16.4**).

With phosphate from seawater, nucleotides such as AMP and ATP could have been formed and polymerized into RNA by montmorillonite clay, a material known to catalyze such reactions. The flow of H_2 and H_2S from the crust provided steady sources of electrons for this prebiotic chemistry, and the process was perhaps powered by redox and pH gradients developed across semipermeable FeS membrane-like surfaces, providing a prebiotic proton motive force (↻ Section 4.10).

An important point to keep in mind here is that before life appeared on Earth, organic precursors of life would not have been consumed by organisms, as they would be today. So the possibility that millions of years ago organic matter accumulated to levels where self-replicating entities emerged, is not an unreasonable hypothesis.

An RNA World and Protein Synthesis

The synthesis and concentration of organic compounds by prebiotic chemistry set the stage for self-replicating systems, the precursors to cellular life. How might self-replicating systems have arisen? One possibility is that there was an early *RNA world*, in which the first self-replicating systems were molecules of RNA (Figure 16.4). Although fragile, RNA could have survived in the cooler temperatures where the gel-like precipitates formed at ocean floor warm springs. Because RNA can bind small molecules, such as ATP and other nucleotides, and has catalytic activity (ribozymes, ↻ Section 7.8), RNA might have catalyzed its own synthesis from the available sugars, bases, and phosphate.

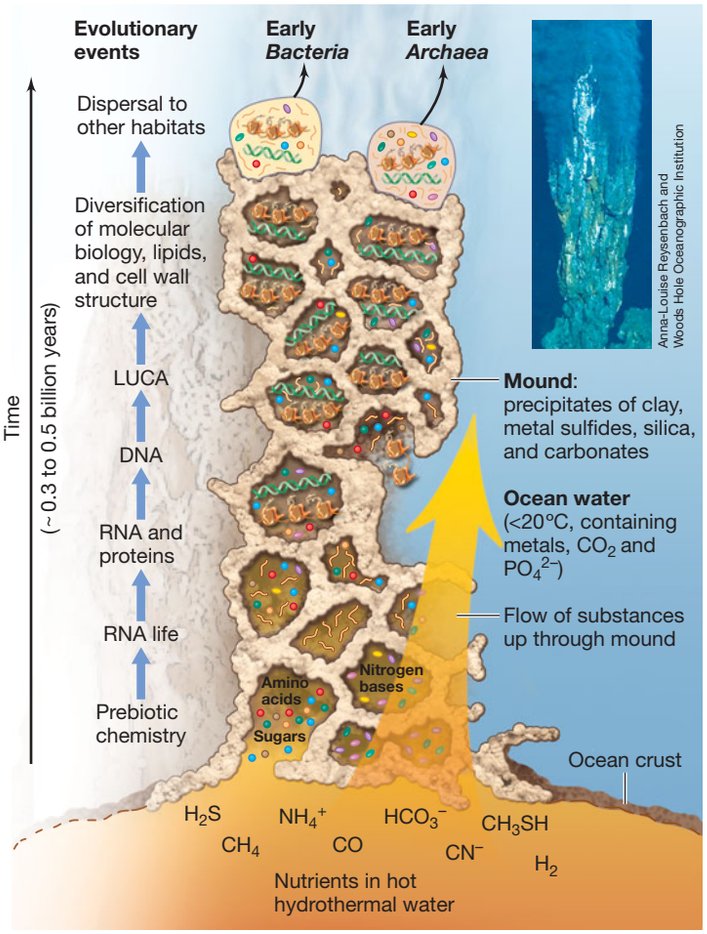


Figure 16.4 Submarine mounds and their possible link to the origin of life. Model of the interior of a hydrothermal mound with hypothesized transitions from prebiotic chemistry to cellular life depicted. Key milestones are self-replicating RNA, enzymatic activity of proteins, and DNA taking on a genetic coding function, leading to early cellular life. This was followed by diversification of molecular biology and biochemistry, eventually giving rise to early *Bacteria* and *Archaea*. LUCA, last universal common ancestor. Inset: photo of an actual hydrothermal mound. Hot mineral-rich hydrothermal fluid mixes with cooler, more oxidized, ocean water, forming precipitates. The mound is composed of precipitates of Fe and S compounds, clays, silicates, and carbonates.

RNA also can bind other molecules, such as amino acids, catalyzing the synthesis of primitive proteins. As different proteins were made and then accumulated in the RNA world, they coated the inner surfaces of the hydrothermal mounds. Later, as different types of proteins emerged, some with catalytic abilities, proteins began to take over the catalytic role of RNAs (Figure 16.4). Eventually, DNA, a molecule more stable than RNA and therefore a better repository of genetic (coding) information, arose and assumed the template role for RNA synthesis. This three-part system—DNA, RNA, and protein—became fixed early on as the fittest solution to biological information processing. Following these steps, one can envision a time of intensive biochemical

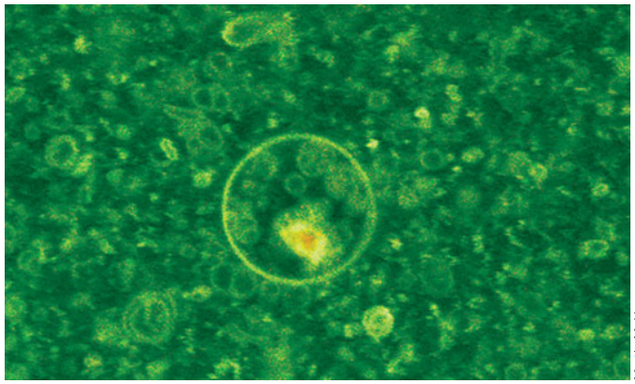


Figure 16.5 **Lipid vesicles made in the laboratory from the fatty acid myristic acid and RNA.** The vesicle itself stains green, and the RNA complexed inside the vesicle stains red. Vesicle synthesis is catalyzed by the surfaces of montmorillonite clay particles.

innovation and experimentation in which much of the structural and functional machinery of these earliest self-replicating systems was invented and refined by natural selection.

Lipid Membranes and Cellular Life

Anther important step in the emergence of cellular life was the synthesis of phospholipid membrane vesicles that could enclose the evolving biochemical and replication machinery. Proteins embedded in the lipids would have made the vesicles semipermeable and thus able to shuttle nutrients and wastes across the membrane, setting the stage for the evolution of energy-conserving processes and ATP synthesis. By entrapping RNA and DNA, these lipoprotein vesicles, which may have been similar to montmorillonite clay vesicles that can be synthesized in the laboratory (**Figure 16.5**), may have enclosed the first self-replicating entities, partitioning the biochemical machinery in a unit not unlike the cells we know today.

From this population of structurally very simple early cells, referred to as the *last universal common ancestor* (*LUCA*), cellular life began to evolve in two distinct directions, possibly in response to physiochemical differences in their most successful niches (Figure 16.4). These two populations of cells would have then undergone strong selection for improvements in transport, metabolism, motility, energy conservation, and the many other structural and functional aspects we associate with cells today. In the two lineages similar overarching processes evolved, but many of the underlying details differed. For example, the two populations evolved different lipids, cell walls, specialized metabolisms, and enzymatic machinery for nucleic acid replication and protein synthesis. As natural selection continued, these two prokaryotic lineages, the *Bacteria* and the *Archaea*, became ever more distinct, displaying the characteristic properties we associate with each lineage today (see Table 16.1).

Early Metabolism

From the time of formation, the early ocean and all of Earth was anoxic. Molecular oxygen (O_2) did not appear in any significant quantities until oxygenic photosynthesis by cyanobacteria

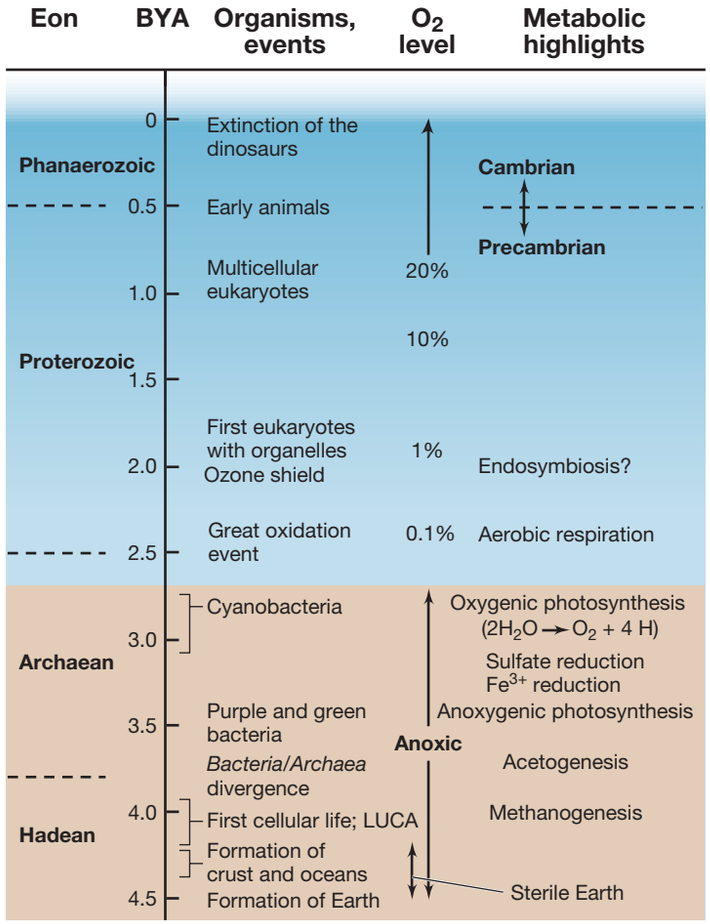


Figure 16.6 **Major landmarks in biological evolution, Earth's changing geochemistry, and microbial metabolic diversification.** The maximum time for the origin of life is fixed by the time of the origin of Earth, and the minimum time for the origin of oxygenic photosynthesis is fixed by the Great Oxidation Event, about 2.4 billion years ago (BYA). Note how the oxygenation of the atmosphere from cyanobacterial metabolism was a gradual process, occurring over a period of about 2 billion years. *Bacteria* respiring at low O_2 levels likely dominated Earth for a billion years or so before Earth's atmosphere reached current levels of oxygen. Compare this figure with the introduction to the antiquity of life on Earth shown in Figure 1.6.

evolved (**Figure 16.6**). Thus, the energy-generating metabolism of primitive cells would have been exclusively anaerobic and would likely have had to be heat-stable because of the temperature of early Earth. Carbon metabolism may well have been autotrophic because consumption of abiotically formed organic compounds for cellular material probably would have exhausted these compounds relatively quickly. The possibility that the use of CO_2 as a carbon source (autotrophy) was an early physiological lifestyle is also supported by the metabolism of many of the earliest lineages on the phylogenetic tree of life (see Figure 16.16); for example, the genera *Aquifex* (*Bacteria*) and *Pyrolobus* (*Archaea*) are autotrophs and, not surprisingly, are also hyperthermophiles.

It is widely thought that H_2 was a major fuel for energy metabolism of early cells. This hypothesis is also supported by the tree

of life, in that virtually all of the earliest branching organisms in the *Bacteria* and *Archaea* lineages use H_2 as an electron donor in energy metabolism. Abiotic reactions between iron sulfide minerals and hydrogen sulfide have been proposed as a source of the needed H_2:

$$FeS + H_2S \rightarrow FeS_2 + H_2 \qquad \Delta G^{0\prime} = -42 kJ$$

Also, ferrous iron (Fe^{2+}) can reduce protons to H_2 in the presence of ultraviolet radiation as an energy source. Regardless of the source, H_2 could have fueled a primitive ATPase in the cytoplasmic membrane of early cells to yield ATP (**Figure 16.7**). However, with H_2 as an electron donor, an electron acceptor would also have been required to form a redox pair (↻ Section 4.6); this could have been elemental sulfur (S^0). As shown in Figure 16.7, the oxidation of H_2 with the reduction of S^0 to yield H_2S is exergonic and would likely have required few enzymes. Moreover, because of the abundance of H_2 and sulfur compounds on early Earth, this scheme would have provided cells with a nearly limitless supply of energy.

These early forms of chemolithotrophic metabolism driven by H_2 would likely have supported the production of large amounts of organic compounds from autotrophic CO_2 fixation. Over time, these organic materials would have accumulated and could have provided the environment needed for the appearance of new chemoorganotrophic bacteria with diverse metabolic strategies to conserve energy from organic compounds; metabolic diversity (Chapter 14) would have been off and running.

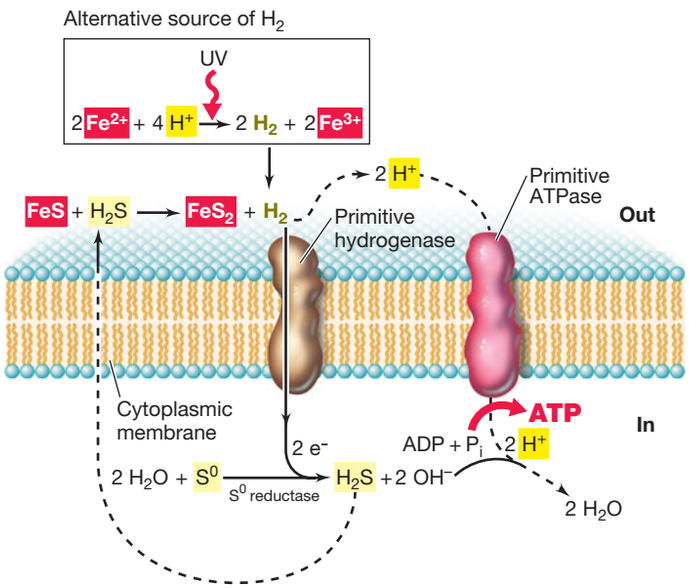


Figure 16.7 A possible energy-generating scheme for primitive cells. Formation of pyrite leads to H_2 production and S^0 reduction, which fuels a primitive ATPase. Note how H_2S plays only a catalytic role; the net substrates would be FeS and S^0. Also note how few different proteins would be required. $\Delta G^{0\prime} = -42$ kJ for the reaction $FeS + H_2S \rightarrow FeS_2 + H_2$. An alternative source of H_2 could have been the UV-catalyzed reduction of H^+ by Fe^{2+} as shown.

MiniQuiz

- What roles did the mounds of mineral-rich materials at warm hydrothermal springs play in the origin of life?
- What important cell structure was necessary for life to proceed from an RNA world to cellular life?
- How could cells have obtained energy from $FeS + H_2S$?

16.3 Microbial Diversification: Consequences for Earth's Biosphere

Following the origin of cells and the development of early forms of energy and carbon metabolism, microbial life underwent a long process of metabolic diversification, taking advantage of the various and abundant resources available on Earth. As particular resources were consumed and became limiting, evolution selected for more efficient and novel metabolisms. Also, microbial life, through its metabolic activity, altered the biosphere, depleting some resources and creating others through the production of waste products and cellular material. Here we examine the scope of metabolic diversification and focus on one key metabolic waste product in particular, molecular oxygen (O_2), a molecule that had a profound impact on the further evolution of life on Earth.

Metabolic Diversification

Geological and molecular data allow us to look back in time to gain insight into microbial diversification. Molecular evidence, in contrast to geological materials, which can be examined directly, is indirect; phylogenies are based on comparisons of DNA sequences that only estimate when the ancestors of modern bacteria first appeared. In Section 16.6 we describe molecular clocks and DNA sequence-based analysis, but here we use molecular information to estimate a timescale for the appearance of the major metabolic groups of bacteria.

LUCA, the last universal common ancestor, may have existed as early as 4.3 billion years ago (Figure 16.6). Molecular evidence suggests that ancestors of modern-day *Bacteria* and *Archaea* had already diverged by 3.8–3.7 billion years ago (↻ Figure 1.6*b*). As these lineages diverged, they developed distinct metabolisms. Early *Bacteria* may have used H_2 and CO_2 to produce acetate, or ferrous iron (Fe^{2+}) compounds, for energy generation, as noted above. At the same time, early *Archaea* developed the ability to use H_2 and CO_2, or possibly acetate as it accumulated, as substrates for methanogenesis (the production of methane, CH_4), according to the following formulas:

$$4\,H_2 + CO_2 \rightarrow CH_4 + 2\,H_2O$$
$$H_3CCOO^- + H_2O \rightarrow CH_4 + HCO_3^-$$

Phototrophy (↻ Sections 13.1–13.5) arose somewhat later, about 3.3 billion years ago, and apparently only in *Bacteria*. The ability to use solar radiation as an energy source allowed phototrophs to diversify extensively. With the exception of the early-branching hyperthermophilic *Bacteria* (genera such as *Aquifex* and *Thermotoga*), the common ancestor of all other *Bacteria* appears to be an anaerobic phototroph, possibly similar to *Chloroflexus*. About

UNIT 6

2.7–3 billion years ago, the cyanobacteria lineage developed a photosystem that could use H_2O in place of H_2S for photosynthetic reduction of CO_2, releasing O_2 instead of elemental sulfur (S^0) as a waste product; the evolution of this process opened up many new metabolic possibilities, in particular aerobic respirations (Figure 16.6).

The Rise of Oxygen: Banded Iron Formations

Molecular and chemical evidence indicates that oxygenic photosynthesis first appeared on Earth about 300 million years before significant levels of O_2 appeared in the atmosphere. By 2.5 billion years ago, O_2 levels had risen to one part per million, a tiny amount by present-day standards, but enough to initiate what has come to be called the *Great Oxidation Event* (Figure 16.6). What delayed the buildup of O_2 for so long?

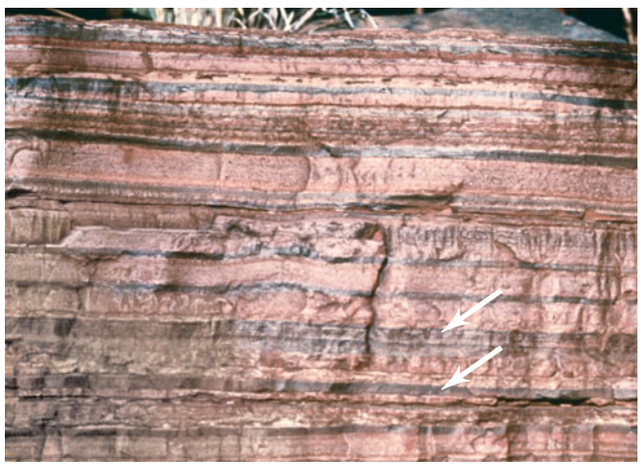
The O_2 that cyanobacteria produced did not begin to accumulate in the atmosphere until it first reacted and consumed the bulk of reduced materials, especially reduced iron minerals such as FeS and FeS_2, in the oceans; these materials oxidize slowly but spontaneously with O_2. The Fe^{3+} produced from the oxidation of these minerals became a prominent marker in the geological record. Much of the iron in rocks of Precambrian origin (>0.5 billion years ago, see Figure 16.6) exists in **banded iron formations** (**Figure 16.8**), laminated sedimentary rocks formed in deposits of iron- and silica-rich materials. The metabolism of cyanobacteria yielded O_2 that oxidized Fe^{2+} to Fe^{3+}. The Fe^{3+} formed various iron oxides that accumulated in layers as banded iron formations (Figure 16.8). Once the abundant Fe^{2+} on Earth was consumed, the stage was set for O_2 to accumulate in the atmosphere, but not until 800–900 million years ago did atmospheric O_2 accumulate to present-day levels (~21%, Figure 16.6).

New Metabolisms and the Ozone Shield

As O_2 accumulated on Earth, the atmosphere gradually changed from anoxic to oxic (Figure 16.6). Species of *Bacteria* and *Archaea* unable to adapt to this change were increasingly restricted to anoxic habitats because of the toxicity of O_2 and because it oxidized the reduced substances upon which their metabolisms were dependent. However, the oxic atmosphere also created conditions for the evolution of various new metabolic pathways, such as sulfate reduction, nitrification, and the various other chemolithotrophic processes (Chapters 13 and 14). Prokaryotes that evolved the ability to respire O_2 gained a tremendous energetic advantage because of the high reduction potential of the O_2/H_2O couple (↻ Section 4.6) and so were capable of producing larger cell populations from a given amount of resources than were anaerobic organisms. Larger and rapidly growing cell populations increased the chances for natural selection of new types of metabolic schemes.

As Earth became more oxic, organelle-containing eukaryotic microorganisms arose (Section 16.4), and the rise in O_2 spurred their rapid evolution. The oldest microfossils known to be eukaryotic because they have recognizable nuclei are about 2 billion years old. Multicellular and increasingly complex microfossils of algae are evident from 1.9 to 1.4 billion years ago. By 0.6 billion years ago, with O_2 near present-day levels, large multicellular organisms, the Ediacaran fauna, were present in the sea (Figure 16.6). In a relatively short time, multicellular eukaryotes diversified into the ancestors of modern-day algae, plants, fungi, and animals (↻ Section 16.8).

An important consequence of O_2 for the evolution of life was the formation of *ozone* (O_3), a gas that provides a barrier preventing much of the intense ultraviolet (UV) radiation of the sun from reaching the Earth. When O_2 is subject to UV radiation, it is converted to O_3, which strongly absorbs wavelengths up to 300 nm. Until an ozone shield developed in Earth's upper atmosphere, evolution could have continued only beneath the ocean surface and in protected terrestrial environments where organisms were not exposed to the lethal DNA damage from the sun's intense UV radiation. However, as Earth developed an ozone shield, organisms could range over the surface of Earth, exploiting new habitats and evolving ever-greater diversity. Figure 16.6 summarizes some landmarks in biological evolution and Earth's geochemistry as Earth transitioned from an anoxic to a highly oxic planet.



John M. Hayes

Figure 16.8 **Banded iron formations.** An exposed cliff made of sedimentary rock about 10 m in height in Western Australia contains layers of iron oxides (arrows) interspersed with layers containing iron silicates and other silica materials. The iron oxides contain iron in the ferric (Fe^{3+}) form produced from ferrous iron (Fe^{2+}) primarily by the oxygen released by cyanobacterial photosynthesis.

MiniQuiz

- Why is the advent of cyanobacteria considered a critical step in evolution?
- In what oxidation state is iron present in banded iron formations?
- What role did ozone play in biological evolution, and how did cyanobacteria make the production of ozone possible?

16.4 Endosymbiotic Origin of Eukaryotes

Up to about 2 billion years ago, all cells apparently lacked a membrane-enclosed nucleus and organelles, the key characteristics of eukaryotic cells (domain *Eukarya*). Here we consider the origin of the *Eukarya* and show how eukaryotes are genetic chimeras containing genes from at least two different phylogenetic domains.

Endosymbiosis

The lineages that gave rise to modern-day *Bacteria* and *Archaea* had existed as the only life forms on our planet for about 2 billion years before eukaryotes appeared (Figure 16.6). This timing tells us that the origin of eukaryotes came after the rise in atmospheric O_2, the development of respiratory metabolism and photosynthesis in *Bacteria,* and the evolution of enzymes such as superoxide dismutase (↻ Section 5.18) that could detoxify the oxygen radicals generated as a by-product of aerobic respiration. How might *Eukarya* have arisen and in what ways did the availability of oxygen influence evolution?

A well-supported explanation for the origin of the eukaryotic cell is the **endosymbiotic hypothesis**. The hypothesis posits that the mitochondria of modern-day eukaryotes arose from the stable incorporation of a respiring bacterium into other cells, and that chloroplasts similarly arose from the incorporation of a cyanobacterium-like organism that carried out oxygenic photosynthesis. Oxygen was almost certainly a driving force in endosymbiosis through its consumption in energy metabolism by the ancestor of the mitochondrion and its production in photosynthesis by the ancestor of the chloroplast. The greater amounts of energy released by aerobic respiration undoubtedly contributed to rapid evolution of eukaryotes, as did the ability to exploit sunlight for energy.

The overall physiology and metabolism of mitochondria and chloroplasts and the sequence and structures of their genomes support the endosymbiosis hypothesis. For example, both mitochondria and chloroplasts contain ribosomes of prokaryotic size (70S) and have 16S ribosomal RNA gene sequences (Section 16.6) characteristic of certain *Bacteria*. Moreover, the same antibiotics that inhibit ribosome function in free-living *Bacteria* inhibit ribosome function in these organelles. Mitochondria and chloroplasts also contain small amounts of DNA arranged in a covalently closed, circular form, typical of *Bacteria* (↻ Section 2.6). Indeed, these and many other telltale signs of *Bacteria* are present in organelles from modern eukaryotic cells (↻ Section 20.4).

There are, however, two other questions germane to how the eukaryotic cell arose: (1) What kind of cell was it that acquired endosymbionts? (2) How did the nuclear membrane form?

Formation of the Eukaryotic Cell

Two hypotheses have been put forward to explain the formation of the eukaryotic cell (**Figure 16.9**). In one, eukaryotes began as a nucleus-bearing cell that later acquired mitochondria and chloroplasts by endosymbiosis (Figure 16.9*a*). In this hypothesis, the nucleus-bearing cell line arose in a lineage of cells that split from the *Archaea*; the nucleus is thought to have arisen in this cell line during evolutionary experimentation with increasing cell and genome size, probably in response to oxic events that were transforming the geochemistry of Earth (Section 16.3). However, a major problem with this hypothesis is that it does not easily account for the fact that *Bacteria* and *Eukarya* have similar membrane lipids, in contrast to those of *Archaea* (↻ Section 3.3).

The second hypothesis, called the *hydrogen hypothesis,* proposes that the eukaryotic cell arose from an association between a H_2-producing species of *Bacteria,* the symbiont, which eventually gave rise to the mitochondrion, and a species of H_2-consuming *Archaea,* the host (Figure 16.9*b*). In this

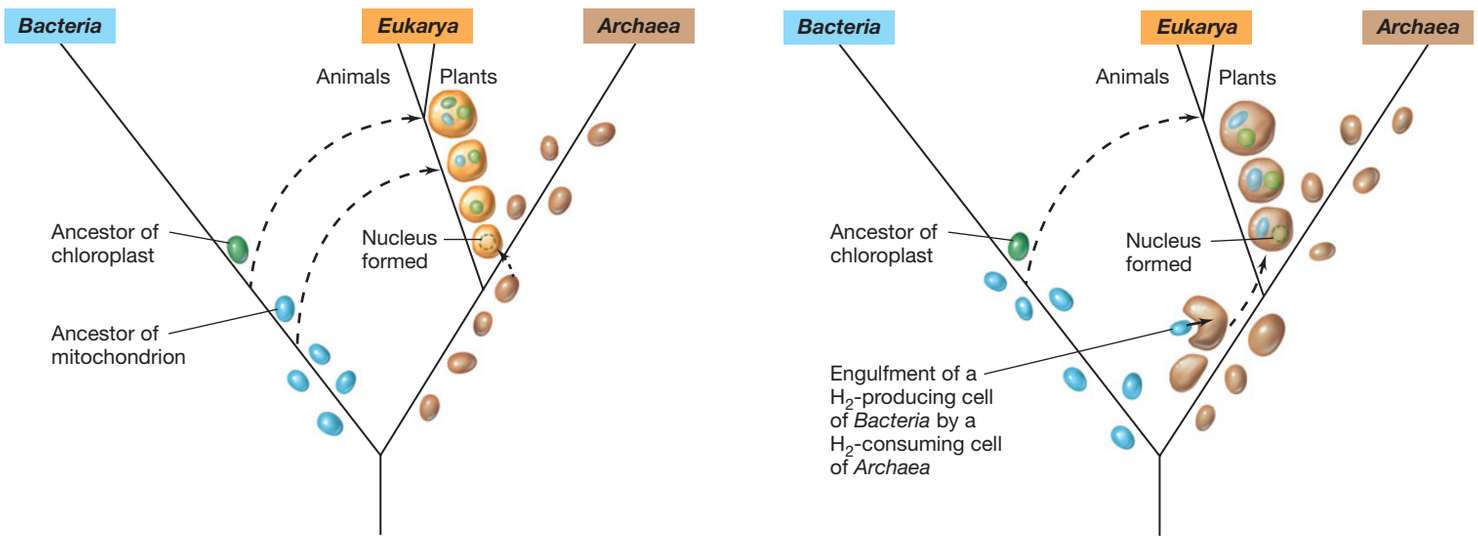


Figure 16.9 **Endosymbiotic models for the origin of the eukaryotic cell.** *(a)* The nucleated line diverged from the archaeal line and later acquired by endosymbiosis the bacterial ancestor of the mitochondrion and then the cyanobacterial ancestor of the chloroplast, at which point the nucleated line diverged into the lineages giving rise to plants and animals. *(b)* The hydrogen hypothesis. The bacterial ancestor of the mitochondrion was taken up endosymbiotically by a species of *Archaea* and the nucleus developed later followed by the endosymbiotic acquisition of the cyanobacterial ancestor of the chloroplast. Note the position of the mitochondrion and chloroplast on the universal phylogenetic tree shown in Figure 16.16.

hypothesis, the nucleus arose after genes for lipid synthesis were transferred from the symbiont to the host. This led to the synthesis of lipids containing fatty acids by the host, lipids that may have been more conducive to the formation of internal membranes, such as the nuclear membrane system (↻ Section 20.1). The simultaneous increase in size of the host genome led to sequestering DNA within a membrane, which organized it and made replication and gene expression more efficient. Later, this mitochondrion-containing, nucleated cell line acquired chloroplasts by endosymbiosis, leading to the first phototrophic eukaryotes (Figure 16.9*b*).

Both hypotheses to explain the origin of eukaryotes point to the eukaryotic cell as a genetic chimera, a cell made up of genes from both *Bacteria* and *Archaea*. However, the hydrogen hypothesis nicely accounts for the observation that eukaryotic cells contain bacterial (rather than archaeal) lipids, yet share with *Archaea* many molecular features of transcription and translation (see Table 16.1 and Chapter 7). The hydrogen hypothesis predicts that aspects of energy metabolism—that is, ATP-producing pathways in mitochondria, hydrogenosomes (degenerate mitochondria, ↻ Section 20.2), and the cytoplasm, as well as glycolytic enzymes in the cytoplasm—should be more similar in eukaryotes and *Bacteria* than in eukaryotes and *Archaea*, and research has shown this to be true.

Consequences of the Evolution of the Modern Eukaryote

No matter how the eukaryotic cell arose, the appearance of eukaryotes was a major step in evolution, creating complex cells with new capabilities powered by a respiratory organelle and, in phototrophic cells, a photosynthetic organelle as well. Like the origin of the first cells from abiotic materials and the diversification of *Bacteria* and *Archaea*, evolution of the eukaryotic cell with its many individual components probably took long periods of time and had many dead ends. Like the explosion in diversity of *Bacteria* and *Archaea*, each step along the way to the modern eukaryotic cell created new opportunities for variation to arise and natural selection to work, with some functions being discarded while others were being refined, eventually producing a totally new model for cellular life, a model upon which complex multicellular organisms could be built. Indeed, the period from about 2 billion years ago to the present saw the rise and diversification of unicellular eukaryotic microorganisms, the origin of multicellularity, and the appearance of structurally complex plant, animal, and fungal life (Figures 16.6, 16.9, and see Figure 16.16).

MiniQuiz

- What evidence supports the idea that the eukaryotic mitochondrion and chloroplast were once free-living members of the domain *Bacteria*?
- Why does the hydrogen hypothesis for endosymbiosis best account for the properties of modern eukaryotes?
- In what ways are modern eukaryotes a combination of attributes of *Bacteria* and *Archaea*?

II Microbial Evolution

We begin here by reviewing the evolutionary process. We consider how scientists reconstruct the evolutionary history of life using methods of molecular genetics, and will see, as summarized in the universal tree of life, how microorganisms are related to each other and to other living things.

16.5 The Evolutionary Process

Evolution, the process by which organisms undergo descent with modification, is driven by mutation and selection. In this Darwinian view of life, all organisms are related through descent from an ancestor that lived in the past. We have outlined a hypothesis for the origin of the most distant of those ancestors, the last universal common ancestor (LUCA, Section 16.2). Since the time of LUCA, life has undergone an extensive process of change as new kinds of organisms arose from other kinds existing in the past. Evolution has also led to the loss of life forms, with organisms less able to compete becoming extinct over time. Evolution accounts not only for the tremendous diversity we see today, but also for the high level of complexity in modern organisms. Indeed, no organism living today is primitive. All extant life forms are modern organisms, well adapted to and successful in their ecological niches, having arisen by evolution under the pressure of natural selection.

Genomic Changes

DNA sequence variation can arise in the genome of an organism from mutations including the loss or gain of whole genes. Mutations, which arise from errors in replication and from certain external factors such as ultraviolet radiation, are essential for life to evolve through natural selection. Adaptive mutations are those that improve the **fitness** of an organism, increasing its survival capacity or reproductive success compared with that of competing organisms. By contrast, harmful mutations lower an organism's fitness. Most mutations, however, are neutral, neither benefiting nor causing harm to the organism, and over time these mutations can accumulate in an organism's genome.

Recall that prokaryotes are genetically haploid; this affects their evolution because mutations in prokaryotic cells are not "covered" by a second copy of the gene, as they are in diploid organisms, but are instead immediately expressed. However, the process of *gene duplication* (↻ Section 12.10) can set the stage for the origin of new functions as mutations in the duplicated sequence encode proteins that differ in greater and greater ways from the original protein. Mutations can also lead to gene loss, which eliminates from the cell the gene product and any competitive benefit accruing from it. Extreme cases of gene loss are often part of the evolutionary history of obligate symbionts and parasites, organisms that receive their essential nutrients from their hosts (↻ Section 25.9).

Another process can also bring about heritable changes in the sequence of an organism's genome: **Horizontal gene transfer.** This process can bring in genes from near or distantly related lineages as cells exchange genes by any of several mechanisms (Chapters 10 and 12).

Selection and the Rapidity of Evolution in Prokaryotes

Regardless of whether a mutation or other change in a genome is neutral, beneficial, or harmful, these changes provide the opportunity for selection of genetically new organisms whose genomes have greater or fewer capacities. As environmental changes create new habitats, cells are presented with new conditions under which they may either survive and successfully compete for nutrients or become extinct. The heritable variation present in a population of cells provides the raw material for natural selection (see Figure 16.25). That is, reproduction of those individuals bearing mutations beneficial under the new circumstances is favored. Moreover, because bacteria typically form large populations that can increase in number quite rapidly, evolutionary events in bacterial populations can also occur quite rapidly. A classic example of this can be seen in laboratory experiments with purple phototrophic bacteria such as *Rhodobacter*, organisms that can grow both chemotrophically and phototrophically. When cultured under anoxic conditions, the cells produce bacteriochlorophyll and carotenoids. In the light, these pigments allow for photosynthetic reactions that lead to ATP synthesis (↻ Section 13.5). However, when cultured under anoxic conditions in darkness, the cells still make pigments because anoxia is the signal that triggers their synthesis. Is there a selective advantage (or disadvantage) to this metabolic strategy?

In nature, if dark-growing cells of phototrophic purple bacteria do not see light right away, they may see light a bit later, and by synthesizing pigments in the dark they are prepared to begin photosynthesis immediately when light returns; thus there is a selective advantage to this strategy. But when serially subcultured in darkness in the laboratory, making pigments that cannot be used is a metabolic disadvantage, and mutants incapable of photosynthesis quickly take over the population (**Figure 16.10**). These mutants no longer carry the burden of making all (or in some mutants any) of the photosynthetic pigments of the wild-type organisms. Since in darkness such pigments would be useless to these cells anyway, the mutants grow faster than any remaining wild-type cells. Although these mutants have reduced phototrophic capacities or in some cases have completely lost the ability to grow phototrophically (see photo inset in Figure 16.10), in permanent darkness they quickly become the fittest organisms in the population and therefore enjoy the greatest reproductive success. Such mutations affecting photosynthesis occur at the same rate in the light as in the dark, but in the light the selection for phototrophy is so strong that such mutants are quickly lost from the population.

The transitions shown in the experiment of Figure 16.10, which occurred in a matter of a few days, remind us of how fast evolutionary pressures can shift even major properties (such as metabolic strategies) of a microbial cell population. In accordance with evolutionary theory, the environment (in this case darkness) selected the fittest organisms for further propagation; that is, those cells whose dark growth rate was maximal. Cells unable to maintain such rapid growth rates are replaced, and eventually, a homogeneous population exists of cells that grow best under the given set of conditions.

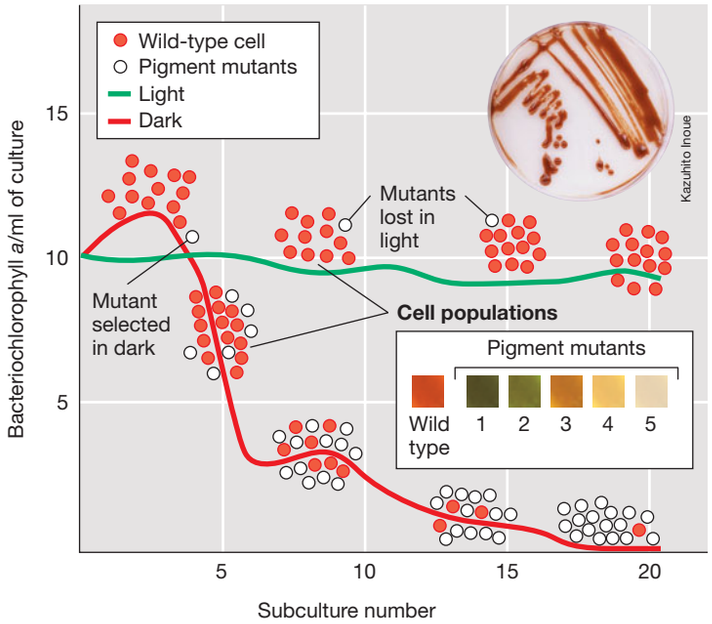


Figure 16.10 Survival of the fittest and natural selection in a population of phototrophic purple bacteria. Serial subculture of *Rhodobacter capsulatus* under anoxic dark conditions quickly selects for nonphototrophic mutants that outcompete and grow faster than cells still making bacteriochlorophyll and carotenoids. Photos: top, plate culture showing colonies of phototrophic cells of *R. capsulatus*; bottom, close-up photos of colonies of wild type and five pigment mutants (1–5) obtained during serial dark subculture. Wild-type cells are reddish-brown from their assortment of carotenoid pigments. The color of mutant colonies reflects the absence (or reduced synthesis) of one or more carotenoids. Mutant strain 5 lacked bacteriochlorophyll and was no longer able to grow phototrophically. Mutant strains 1–4 could grow phototrophically but at reduced growth rates from the wild type. Data adapted from Madigan, M.T., et. al. 1982. *J. Bacteriol. 150*: 1422–1429.

MiniQuiz

- How can gene duplication assist the evolutionary process?
- How does the accumulation of mutations set the stage for selection?
- In the experiment of Figure 16.10, why did the dark cell population lose its pigments?

16.6 Evolutionary Analysis: Theoretical Aspects

The evolutionary history of a group of organisms is called its **phylogeny**, and a major goal of evolutionary analysis is to understand phylogenetic relationships. Because we do not have direct knowledge of the path of microbial evolution, phylogeny is inferred indirectly from nucleotide sequence data. Our premises are that (1) all organisms are related by descent, and (2) that the sequence of DNA in a cell's genome is a record of the organism's ancestry. Because evolution is a process of inherited nucleotide sequence change, comparative analyses of DNA sequences allow

us to reconstruct phylogenetic histories. Here, we examine some of the ways in which this is carried out.

Genes Employed in Phylogenetic Analysis

Various genes are used in molecular phylogenetic studies of microorganisms. Most widely used and useful for defining relationships in prokaryotes is the gene encoding **16S ribosomal RNA (rRNA)** (**Figure 16.11**) and its counterpart in eukaryotes, 18S rRNA, parts of the small subunit of the ribosome (↻ Section 6.19). These **small subunit rRNA (SSU rRNA)** genes have been used extensively for sequence-based evolutionary analysis as pioneered by Carl Woese, an American scientist, in the 1970s. SSU rRNA genes are excellent candidates for phylogenetic analysis because they are (1) universally distributed, (2) functionally constant, (3) sufficiently conserved (that is, slowly changing), and (4) of adequate length to provide a deep view of evolutionary relationships. A large and constantly growing database of SSU rRNA gene sequences exists. For example, the **Ribosomal Database Project** (RDP; http://rdp.cme.msu.edu) contains a collection of such sequences, now numbering over 1.3 million, and provides computational programs for analytical purposes.

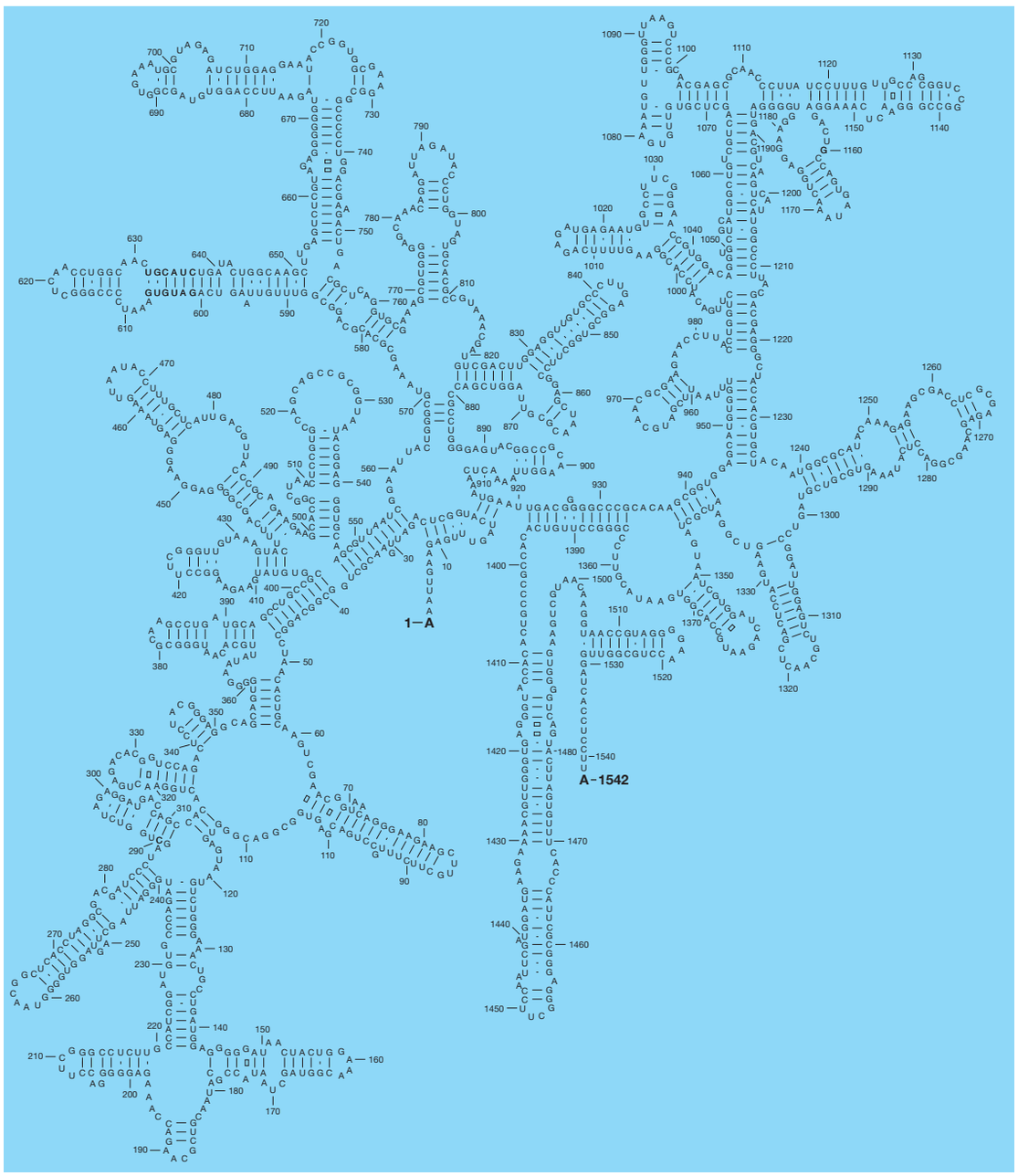


Figure 16.11 **Ribosomal RNA (rRNA).** Primary and secondary structure of 16S rRNA from *Escherichia coli* (*Bacteria*). The 16S rRNA from *Archaea* is similar in secondary structure (folding) but has numerous differences in primary structure (sequence).

Along with SSU genes, those for several highly conserved proteins have been used effectively in phylogenetic analysis, including genes encoding protein synthesis elongation factor EF-Tu (↻ Section 6.19), heat shock protein Hsp60 (↻ Section 8.11), and several transfer RNA (tRNA) synthetases (↻ Section 6.18). Although the highly conserved SSU genes are particularly useful for deep evolutionary analysis, the amount of variation present in SSU rRNA gene sequences is often insufficient to discriminate among closely related species. In Section 16.11 we discuss ways of bypassing this problem by using genes whose sequences have diverged more than the 16S rRNA gene, consequently revealing distinctions between closely related bacteria, and by using multiple genes simultaneously for evolutionary analyses.

Molecular Clocks

An unresolved question in phylogenetics is whether DNA (and protein) sequences change at a constant rate. The approach to answering the question focuses on pairs of homologous sequences—that is, sequences of shared evolutionary ancestry that encode functionally equivalent molecules. If sequences do change at a constant rate, such pairs would serve as an approximate **molecular clock**, allowing the time in the past when the two sequences diverged from a common ancestral sequence to be estimated. Major assumptions of the molecular clock approach are that nucleotide changes accumulate in a sequence in proportion to time, that such changes generally are neutral and do not interfere with gene function, and that they are random.

The molecular clock approach has been used to estimate the time of divergence of distantly related organisms, such as the domains *Archaea* and *Eukarya* (about 2.8 billion years ago, Figure 1.6*b*), as well as closely related organisms, such as the enteric bacteria *Escherichia coli* and *Salmonella typhimurium* (about 120–140 million years ago). These data have also been combined with evidence from the geological record on isotopes and specific biological markers to approximate when different metabolic patterns emerged in bacteria (Section 16.3 and Figure 16.6).

The main problem with the molecular clock approach, however, is that DNA sequences do change at different rates, which means that direct and reliable correlations to a timescale will be difficult to make. However, much of phylogenetic analysis is concerned with *relative* relationships among organisms, shown by their branching order on phylogenetic trees. These relationships are generally discernible from molecular sequence analyses regardless of whether different sequences change at similar rates, so the accuracy of the molecular clock approach is not a major concern.

MiniQuiz

- List three reasons that SSU rRNA genes are suitable for phylogenetic analyses.
- What information does the Ribosomal Database Project provide?
- What value do molecular clocks have in phylogenetic analysis?

16.7 Evolutionary Analysis: Analytical Methods

As we have seen, modern phylogenetics is based on nucleotide sequence comparisons, for which specific methods have been developed. We consider these methods here.

Obtaining DNA Sequences

Phylogenetic analysis using DNA sequences relies heavily on the polymerase chain reaction (PCR) to obtain sufficient copies of a gene for reliable sequencing (↻ Section 6.11). Specific oligonucleotide primers have been designed that bind to the ends of the gene of interest, or to DNA flanking the gene, allowing DNA polymerase to bind to and copy the gene. The source of DNA bearing a gene of interest typically is genomic DNA purified from particular bacterial strains, but could be DNA extracted from an environmental sample (Section 16.9). The PCR product is visualized by agarose gel electrophoresis, excised from the gel, extracted and purified from the agarose, and then sequenced, often using the same oligonucleotides as primers for the sequencing reactions. These steps are summarized in **Figure 16.12**.

An important aspect of PCR amplification is *primer design*, which is a matter of deciding which sequence to use to amplify a specific gene and then actually constructing the sequence. Standard primers exist for many highly conserved genes, such as the SSU rRNA genes (Figure 16.12). Primers are available that are domain-specific and can be used to amplify an SSU gene from any organism in a given domain. Other primers can be designed that are lineage-specific, or even more restrictive. At the other extreme, "universal" primers are available that will amplify SSU genes from *any* organism, prokaryote or eukaryote. Primer design is both an art and a science and often requires computational analyses, along with some trial and error, to construct primers that will effectively amplify the gene of interest.

Sequence Alignment

Phylogenetic analysis is based on *homology*, that is, analysis of DNA sequences that are related by common ancestry. Once the DNA sequence of a gene is obtained, the next step in phylogenetic analysis is to align that sequence with homologous sequences from other organisms. By doing this, nucleotide mismatches and insertions and deletions, some of which may be phylogenetically informative, can be pinpointed.

Figure 16.13 shows an example of sequence alignment. The web-based algorithm BLAST (*B*asic *L*ocal *A*lignment *S*earch *T*ool) of the National Institutes of Health (http://www.ncbi.nlm.nih.gov/BLAST) aligns sequences automatically and can identify genes homologous to a specific sequence from among the many thousands already sequenced. Related sequences are then downloaded from GenBank (http://www.ncbi.nih.gov/Genbank), which is an annotated collection of all publicly available DNA sequences, and aligned. Proper sequence alignment is critical to phylogenetic analysis because the assignment of mismatches and gaps caused by deletions is in effect an explicit hypothesis of how the sequences have diverged from a common ancestral sequence. Genes that encode proteins usually are aligned with the aid of their inferred

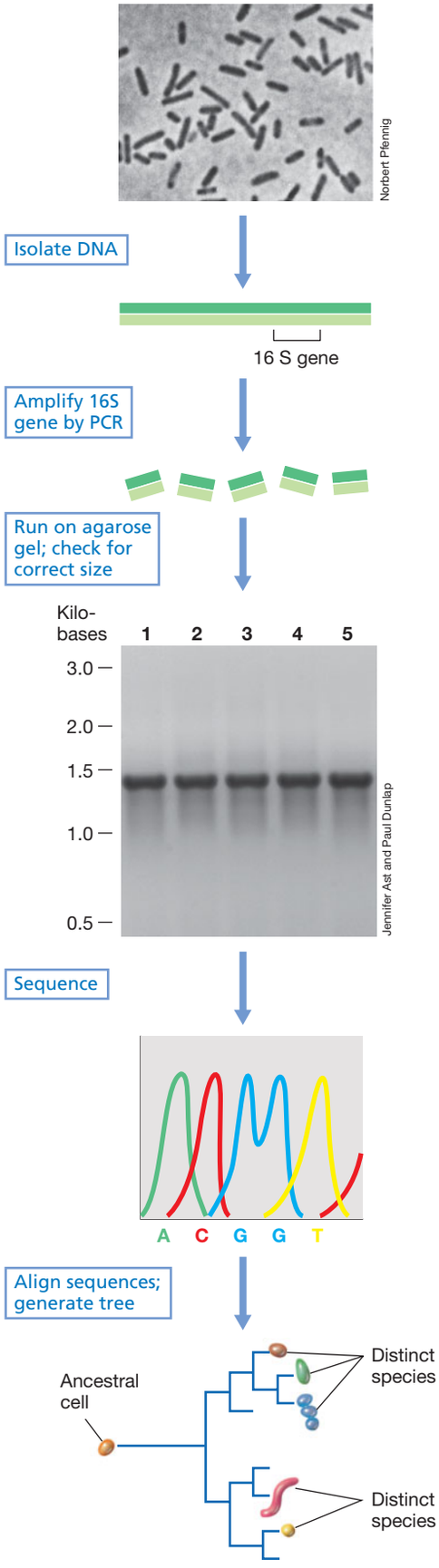


Figure 16.12 PCR amplification of the 16S rRNA gene. Following DNA isolation, primers complementary to the ends of the 16S rRNA (see Figure 16.11) are used to PCR-amplify the 16S rRNA gene from genomic DNA of five different unknown bacterial strains and the products are run on an agarose gel (photo). The bands of amplified DNA are approximately 1465 nucleotides in length. Positions of DNA kilobase size markers are indicated at the left. Excision from the gel and purification of these PCR products is followed by sequencing and analysis to identify the bacteria.

amino acid sequences. Other genes, such as those encoding 16S rRNA, can often be aligned by inspection or through the use of computer programs designed to minimize the number of mismatches and gaps. Secondary structure, the folding of the 16S rRNA (Figure 16.11), is also helpful in making accurate gene alignments because base mismatches that show up in the secondary structure of highly conserved regions of the molecule readily signal alignment errors.

Phylogenetic Trees

Reconstructing evolutionary history from observed nucleotide sequence differences includes construction of a phylogenetic tree, which is a graphic depiction of the relationships among sequences of the organisms under study, much like a family tree. A phylogenetic tree is composed of *nodes* and *branches* (**Figure 16.14**). The tips of the branches represent species that exist now and from which the sequence data were obtained. The nodes are points in evolution where an ancestor diverged into two new organisms, each of which then began to evolve along its separate pathway. The branches define both the order of descent and the ancestry of the nodes, whereas the branch length represents the number of changes that have occurred along that branch.

Phylogenetic trees can be constructed that are either *unrooted,* showing the relative relationships among the organisms under study but not the evolutionary path leading from an ancestor to a strain (Figure 16.14*a*), or *rooted,* in which case the unique path from an ancestor to each strain is defined (Figure 16.14*b*, *c*). Trees are rooted by the inclusion in the analysis of an *outgroup,* an organism that is less closely related to the organisms under study than the organisms are to each other, but that shares with them homologs of the gene under study.

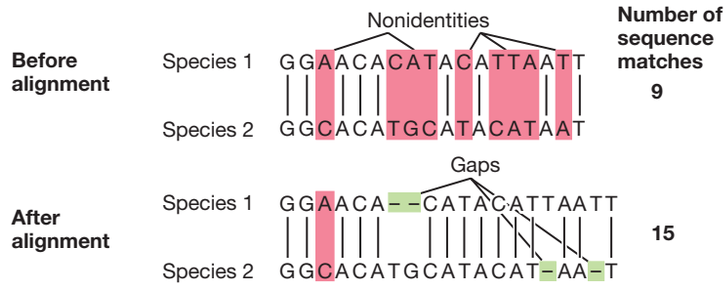


Figure 16.13 Alignment of DNA sequences. Sequences for a hypothetical region of a gene are shown for two organisms, before alignment and after the insertion of gaps to improve the matchup of nucleotides, indicated by the vertical lines showing identical nucleotides in the two sequences. The insertion of gaps in the sequences substantially improves the alignment.

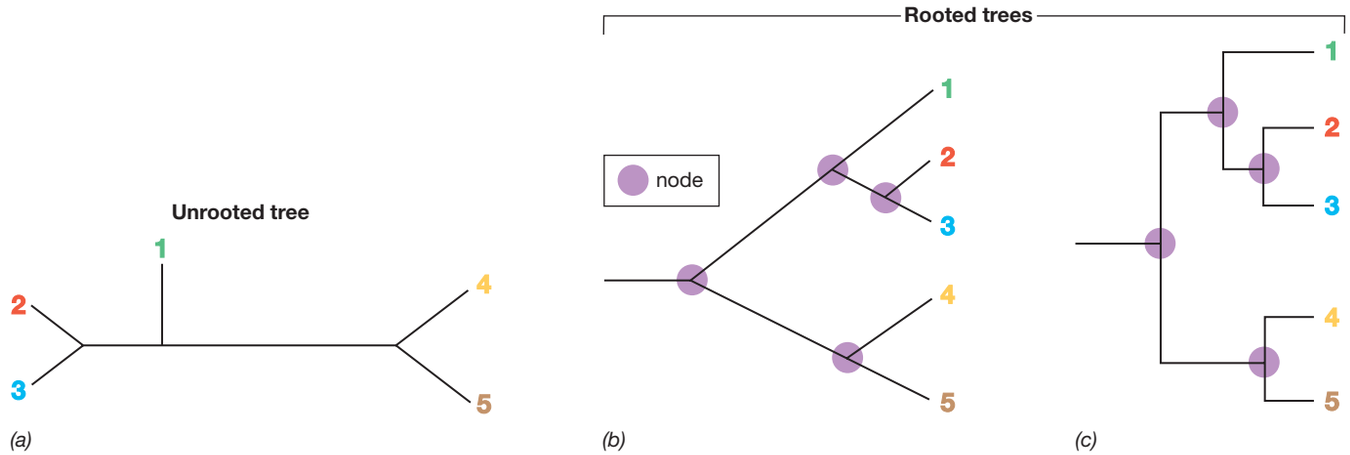


Figure 16.14 **Phylogenetic trees.** Unrooted *(a)* and rooted *(b, c)* forms of a phylogenetic tree are shown. The tips of the branches are species (or strains) and the nodes are ancestors. Ancestral relationships are revealed by the branching order in rooted trees.

In its most basic form, a phylogenetic tree is a depiction of lines of descent, and the relationship between two organisms therefore should be read in terms of common ancestry. That is, the more recently two species shared a common ancestor, the more closely related they are. The rooted trees in Figure 16.14*b* and *c* illustrate this point. Species 2 is more closely related to species 3 than it is to species 1 because 2 and 3 share a more recent common ancestor than do 2 and 1.

Tree Construction

Modern evolutionary analysis uses character-state methods, also called *cladistics*, for tree construction. Character-state methods define phylogenetic relationships by examining changes in nucleotides at particular positions in the sequence, using those characters that are *phylogenetically informative*. These are characters that define a **monophyletic** group; that is, a group that has descended from one ancestor. **Figure 16.15** describes how phylogenetically informative characters are recognized in aligned sequences. Computer-based analysis of these changes generates a phylogenetic tree, or *cladogram*.

A widely used cladistic method is *parsimony,* which is based on the assumption that evolution is most likely to have proceeded by the path requiring fewest changes. Computer algorithms based on parsimony provide a way of identifying the tree with the smallest number of character changes. Other cladistic methods, *maximum likelihood* and *Bayesian analysis,* proceed like parsimony, but they differ by assuming a model of evolution, for example, that certain kinds of nucleotide changes occur more often than others. Inexpensive computer applications, such as PAUP (*P*hylogenetic *A*nalysis *U*nder *P*arsimony, and Other Methods), guidebooks, and web-accessible tutorials are available for learning the basic procedures of cladistic analysis and tree construction.

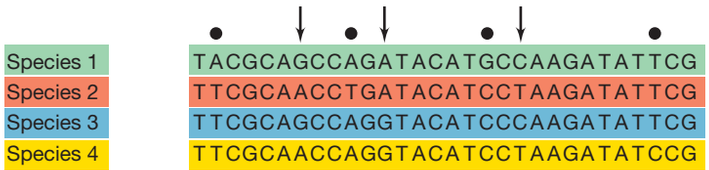


Figure 16.15 **Identification of phylogenetically informative sites.** Aligned sequences for four species are shown. Invariant sites are unmarked, and phylogenetically neutral sites are indicated by dots. Phylogenetically informative sites, varying in at least two of the sequences, are marked with arrows.

MiniQuiz

- How are DNA sequences obtained for phylogenetic analysis?
- What does a phylogenetic tree depict?
- Why is sequence alignment critical to phylogenetic analysis?

16.8 Microbial Phylogeny

Biologists previously grouped living organisms into five *kingdoms*: plants, animals, fungi, protists, and bacteria. DNA sequence-based phylogenetic analysis, on the other hand, has revealed that the five kingdoms do not represent five primary evolutionary lines. Instead, as previously outlined in Chapter 2, cellular life on Earth has evolved along *three* primary lineages, called **domains**. Two of these domains, the *Bacteria* and the *Archaea,* are exclusively composed of prokaryotic cells. The *Eukarya* contains the eukaryotes (**Figure 16.16**), including the plants, animals, fungi, and protists.

An SSU rRNA Gene–Based Phylogeny of Life

The **universal phylogenetic tree** based on small subunit rRNA genes (Figure 16.16) is a genealogy of all life on Earth. It depicts the evolutionary history of all cells and clearly reveals the three domains. The root of the universal tree represents a point in time when all extant life on Earth shared a common ancestor, the last universal common ancestor, LUCA (Figure 16.16).

The three-domain concept is also supported by sequence analysis of several other genes shared among all organisms. Analysis of

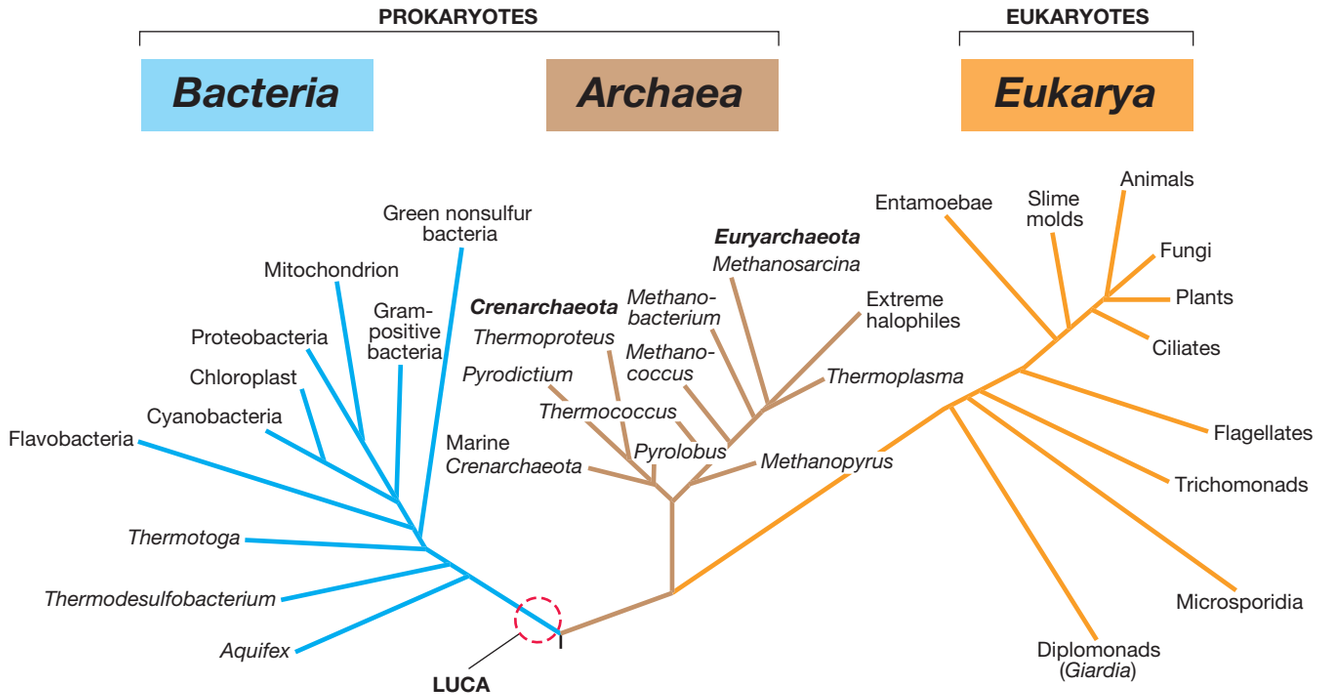


Figure 16.16 Universal phylogenetic tree as determined from comparative SSU rRNA gene sequence analysis. Only a few key organisms or lineages are shown in each domain. At least 80 lineages of *Bacteria* have now been identified although many of these have not yet been cultured. LUCA, last universal common ancestor.

over 30 genes present in nearly 200 species of *Bacteria, Archaea,* and *Eukarya* whose genomes have been completely sequenced confirms the distinct separation between these three lines of descent. Although branching orders and relationships among some lineages within the domains will likely be revised as more genetic data are obtained, analysis of multiple genes from genomic studies (Section 16.11) supports the basic structure of life proposed by Woese based on sequence analysis of SSU rRNA genes.

The presence of genes in common in *Bacteria, Archaea,* and *Eukarya*, of which there are many examples (↻ Section 7.4 and Figure 7.5), raises an interesting question. If these lineages diverged from each other so long ago from a common ancestor, how is it they share so many genes? One hypothesis is that early in the history of life, before the primary domains had diverged, horizontal gene transfer (Chapters 10 and 12) was extensive. During this time, genes encoding proteins that conferred exceptional fitness, for example, genes for the core cellular functions of transcription and translation, were promiscuously transferred among a population of primitive organisms derived from a common ancestral cell. If true, this would explain why, as genome analyses have shown, all cells regardless of domain have many core functional genes in common, more than would be expected if all cells shared a primitive common ancestor (Figure 16.16).

But what about the unique genes present in each domain, of which there are several examples as well (↻ Section 7.4)? It is hypothesized that over time barriers to unrestricted horizontal gene transfer evolved, perhaps from the selective colonization of habitats (thereby generating reproductive isolation) or as the result of structural barriers that in some way prevented free genetic exchange. As a result, the previously genetically promiscuous population slowly began to sort out into the primary lines of evolutionary descent, the *Bacteria* and *Archaea* (Figure 16.16). As each lineage continued to evolve, certain unique biological traits became fixed within each group. Then, about 2.8 billion years ago, the *Archaea* and *Eukarya* diverged as distinct domains (↻ Figure 1.6). Today, after a total of nearly 4 billion years of microbial evolution, we see the grand result: three domains of cellular life that on the one hand share many common features, but on the other hand, display distinctive evolutionary histories of their own. **Table 16.1** summarizes some major characteristics of the three domains.

Bacteria

Among *Bacteria,* at least 80 lineages (called phyla, singular **phylum**, or divisions) have been discovered thus far; only some key ones are shown in the universal tree in Figure 16.16. The *Bacteria* are discussed in detail in Chapters 17 and 18. Many lineages of *Bacteria* are known only from environmental sequences (phylotypes, Sections 16.9 and 22.5). Although some lineages are characterized by unique phenotypic traits, such as the morphology of the spirochetes or the physiology of the cyanobacteria, most major groups of *Bacteria* consist of species that, although specifically related from a phylogenetic standpoint, lack strong phenotypic cohesiveness. The largest group, the ***Proteobacteria***, is a good example of this, as collectively this group shows all known forms of microbial physiology.

Table 16.1 *Major characteristics of Bacteria, Archaea,* and *Eukarya*[a]

Characteristic	*Bacteria*	*Archaea*	*Eukarya*
Morphological and genetic			
Prokaryotic cell structure	Yes	Yes	No
Cell wall	Peptidoglycan	No peptidoglycan	No peptidoglycan
Membrane lipids	Ester-linked	Ether-linked	Ester-linked
Membrane-enclosed nucleus	Absent	Absent	Present
DNA present in covalently closed and circular form	Yes	Yes	No
Histone proteins present	No	Yes	Yes
RNA polymerases (↻ Figure 7.2)	One (4 subunits)	One (8–12 subunits)	Three (12–14 subunits each)
Ribosomes (mass)	70S	70S	80S
Initiator tRNA	Formylmethionine	Methionine	Methionine
Introns in most genes	No	No	Yes
Operons	Yes	Yes	No
Capping and poly(A) tailing of mRNA	No	No	Yes
Plasmids	Yes	Yes	Rare
Sensitivity to chloramphenicol, streptomycin, kanamycin, and penicillin	Yes	No	No
Physiological/special structures			
Dissimilative reduction of S^0 or SO_4^{2-} to H_2S, or Fe^{3+} to Fe^{2+}	Yes	Yes	No
Nitrification (ammonia oxidation)	Yes	Yes	No
Chlorophyll-based photosynthesis	Yes	No	Yes (in chloroplasts)
Denitrification	Yes	Yes	No
Nitrogen fixation	Yes	Yes	No
Rhodopsin-based energy metabolism	Yes	Yes	No
Chemolithotrophy (Fe, NH_3, S, H_2)	Yes	Yes	No
Endospores	Yes	No	No
Gas vesicles	Yes	Yes	No
Synthesis of carbon storage granules composed of poly-β-hydroxyalkanoates	Yes	Yes	No
Growth above 70°C	Yes	Yes	No
Growth above 100°C	No	Yes	No

[a]Note that for many features only particular representatives within a domain show the property.

The major eukaryotic organelles clearly originated from within the domain *Bacteria*, the mitochondrion from within the *Proteobacteria* and the chloroplast from within the cyanobacteria (Figure 16.16). As we discussed earlier, eukaryotic organelles originated from endosymbiotic events (Figure 16.9) that shaped the modern eukaryotic cell as a genetic chimera containing genes from two or more phylogenetic lineages.

Archaea

The domain *Archaea* consists of two major phyla, the *Crenarchaeota* and *Euryarchaeota* (Figure 16.16). We discuss *Archaea* in detail in Chapter 19. Branching close to the root of the universal tree are hyperthermophilic species of *Crenarchaeota,* such as *Pyrolobus* (Figure 16.16). These are followed by the phylum *Euryarchaeota*, which includes the methane-producing (methanogenic) *Archaea* and the extreme halophiles and extreme acidophiles, such as *Thermoplasma* (Figure 16.16). As in the tree of *Bacteria*, there are some lineages of *Archaea* known only from the sampling of rRNA genes from the environment (Section 16.9). This list keeps expanding as more habitats are specifically sampled for archaeal diversity. It has become clear that cultured species of *Archaea*, primarily obtained from extreme environments such as hot springs, saline lakes, acidic soils, and the like, have many relatives in more moderate habitats such as freshwater lakes, streams, agricultural soils, and the oceans; at this point we have only limited knowledge of the activities and metabolic strategies of the *Archaea* that inhabit nonextreme environments.

Eukarya

Phylogenetic trees of species in the domain *Eukarya* have been constructed from comparative sequence analysis of the 18S rRNA gene, the functional equivalent of the 16S rRNA gene. In Chapter 20 where we consider microbial eukaryotes in detail, we will see that the SSU phylogenetic picture of eukaryotes is probably inaccurate. The 18S tree shows some "early-branching" microbial eukaryotes, such as the microsporidia and the diplomonads (Figure 16.16). By contrast, the position of these organisms on multigene phylogenetic trees (Section 16.11) is quite different, and shows them to have arisen during a burst of evolutionary radiation that led to most lineages of microbial eukaryotes (↻ Figure 20.12). It is likely that this burst in eukaryotic evolution was triggered by the onset of oxic conditions on Earth and subsequent development of the ozone shield (Section 16.3). The latter would have greatly expanded the number of surface habitats available for colonization. Nevertheless, although 18S rRNA sequencing appears to give a skewed view of eukaryotic microbial evolution, it still clearly sorts the eukaryotes out as a distinct domain of life with evolutionary roots more closely tied to the *Archaea* than to the *Bacteria* (Figure 16.16).

MiniQuiz

- How does the SSU rRNA tree of life differ from the grouping of life based on five kingdoms?
- What kinds of evidence support the three-domain concept of life?
- How does the universal tree in Figure 16.16 support the hypothesis of endosymbiosis (Figure 16.9)?

16.9 Applications of SSU rRNA Phylogenetic Methods

Many research tools make use of small subunit (SSU) rRNA gene sequencing. These tools include rRNA probes, used in both microbial ecology and diagnostic medicine, and DNA fingerprinting.

Phylogenetic Probes and FISH

Recall that a probe is a strand of nucleic acid that can be labeled and used to hybridize to a complementary nucleic acid (↻ Section 11.2). Probes can be general or specific. For example, universal SSU rRNA probes are available that bind by complementary base pairing to conserved sequences in the rRNA of all organisms, regardless of domain. By contrast, specific probes can be designed that react only with the ribosomes of species in a single domain. Such **phylogenetic probes** can also be designed to target lineages within a domain, such as members of particular families, genera, or even species.

The binding of probes to cellular ribosomes can be seen microscopically if a fluorescent dye is attached to the probes. When cells are treated with the appropriate reagents, their membranes become permeable and allow penetration of the probe–dye mixture. After hybridization of the probe directly to rRNA in ribosomes, the cells become uniformly fluorescent and

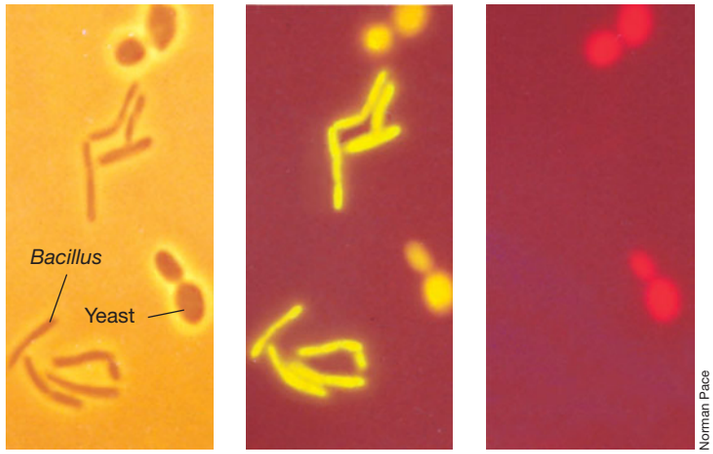


Figure 16.17 Fluorescently labeled rRNA probes: Phylogenetic stains. *(a)* Phase-contrast photomicrograph of cells of *Bacillus megaterium* (rod, *Bacteria*) and the yeast *Saccharomyces cerevisiae* (oval cells, *Eukarya*). *(b)* Same field; cells stained with a yellow-green universal rRNA probe (this probe reacts with species from any domain). *(c)* Same field; cells stained with a eukaryal probe (only cells of *S. cerevisiae* react). Cells of *B. megaterium* are about 1.5 μm in diameter and cells of *S. cerevisiae* are about 6 μm in diameter.

can be observed under a fluorescent microscope (**Figure 16.17**). This technique is called **FISH** (*f*luorescent *i*n *s*itu *h*ybridization) and can be applied to cells in culture or in a natural environment (the term *in situ* means "in the environment"). In essence, FISH is a *phylogenetic stain*.

FISH technology is widely used in microbial ecology and clinical diagnostics. In ecology, FISH can be used for the microscopic identification and tracking of organisms directly in the environment. FISH also offers a method for assessing the composition of microbial communities directly by microscopy (↻ Section 22.4). In clinical diagnostics, FISH has been used for the rapid identification of specific pathogens from patient specimens. The technique circumvents the need to grow an organism in culture. Instead, microscopic examination of a specimen can confirm the presence of a specific pathogen, thus facilitating a rapid diagnosis and treatment. By contrast, isolation and identification of pathogens by classical means typically takes 24–48 hours and can take much longer.

Microbial Community Analysis

Polymerase chain reaction (PCR)-amplified rRNA genes (Figure 16.12) need not originate from a pure culture grown in the laboratory. Using methods described in detail in Chapter 22, a phylogenetic "snapshot" of a natural microbial community can be taken using PCR to amplify the genes encoding SSU rRNA from organisms in that community. Such genes can easily be sorted out, sequenced, and aligned. From these data, a phylogenetic tree can be constructed of sequences that depict the different rRNA genes present in the natural community. From this tree, the presence of specific organisms can be inferred even though none of them were actually cultivated or otherwise identified. Such

microbial community analyses, a major tool of microbial ecology research today, have revealed many key features of microbial community structure and microbial interactions.

Ribotyping

Information from rRNA-based phylogenetic analyses also finds application in a technique for bacterial identification called **ribotyping**. Unlike comparative sequencing methods, however, ribotyping does not require sequencing. Instead, it generates a specific pattern of bands, a kind of *DNA fingerprint* called a *ribotype*, when DNA from an organism is digested by a restriction enzyme and the fragments are separated by gel electrophoresis and probed with an rRNA gene probe (**Figure 16.18**). Differences between organisms in the sequence of their 16S rRNA genes translate into the presence or absence of sites cut by different restriction endonucleases (Section 11.1). The ribotype of a particular organism may therefore be unique and diagnostic, allowing identification of different species and even different strains of a species if there are differences in their SSU rRNA gene sequences.

In ribotyping, following digestion and separation DNA fragments are transferred from the gel onto nylon membranes and hybridized with a labeled rRNA gene probe. The pattern of the fragments on the gel is then digitized, and compared with patterns of reference organisms in a computer database (Figure 16.18). Ribotyping is highly specific and rapid because it bypasses the PCR, sequencing, sequence alignment, and sequence analysis steps of SSU rRNA phylogenetic analysis (Figure 16.12). For these reasons, ribotyping has found many applications in clinical diagnostics and the microbial analyses of food, water, and beverages.

MiniQuiz

- How can oligonucleotide probes be made visible under the microscope? What is this technology called?
- What kinds of questions can be addressed using microbial community analysis?
- How is ribotyping able to distinguish between different bacteria?

Figure 16.18 **Ribotyping.** Ribotype results for four different lactic acid bacteria. DNA was taken from a colony of each bacterium, digested into fragments by restriction enzymes, separated by gel electrophoresis, and then probed with a 16S rRNA gene probe. For each species the electrophoresis produced a unique pattern of bands. Variations in position and intensity of the bands are important in identification.

III Microbial Systematics

Systematics is the study of the diversity of organisms and their relationships. It links together phylogeny, just discussed, with **taxonomy**, in which organisms are characterized, named, and placed into groups according to several defined criteria. Bacterial taxonomy traditionally has focused on practical aspects of identification and description, activities that have relied heavily on phenotypic comparisons. At present, the growing use of genetic information, especially DNA sequence data, is increasingly allowing taxonomy to reflect phylogenetic relationships as well.

Bacterial taxonomy has changed substantially in the past few decades, embracing a combination of methods for the identification of bacteria and description of new species. This *polyphasic approach* to taxonomy uses three kinds of methods—phenotypic, genotypic, and phylogenetic—for the identification and description of bacteria. Phenotypic analysis examines the morphological, metabolic, physiological, and chemical characteristics of the cell. Genotypic analysis considers characteristics of the genome. These two kinds of analysis group organisms based on similarities. They are complemented by phylogenetic analysis, which seeks to place organisms within an evolutionary framework.

16.10 Phenotypic Analysis: Fatty Acid Methyl Esters (FAME)

The observable characteristics—the phenotype—of a bacterium provide many traits that can be used to differentiate species. Typically, for either describing a new species or identifying a bacterium, several of these traits are determined for the organism of interest. The results are then compared with phenotypes of known organisms, either examined in parallel with the unknowns or from published information. The specific traits used depend on the kind of organism, and which traits are chosen for testing may arise from the investigator's purpose and from substantial prior knowledge of the bacterial group to which the new organism likely belongs. For example, in applied situations, such as in clinical diagnostic microbiology, where identification may be an end in itself and time is of the essence, a well-defined subset of traits is typically used that quickly discriminates between likely possibilities. **Table 16.2** lists general categories and examples of some phenotypic traits used in identifications and species descriptions, and we examine one of these traits here.

The types and proportions of fatty acids present in cytoplasmic membrane lipids and the outer membrane lipids of gram-negative bacteria are major phenotypic traits of interest. The technique for identifying these fatty acids has been nicknamed **FAME**, for *f*atty *a*cid *m*ethyl *e*ster, and is in widespread use in clinical, public health, and food and water-inspection laboratories where pathogens routinely must be identified. FAME analyses are also widely used in the characterization of new species of bacteria.

The fatty acid composition of *Bacteria* varies from species to species in chain length and in the presence or absence of double

Table 16.2 *Some phenotypic characteristics of taxonomic value*

Category	*Characteristics*
Morphology	Colony morphology; Gram reaction; cell size and shape; pattern of flagellation; presence of spores, inclusion bodies (e.g., PHB,[a] glycogen, or polyphosphate granules, gas vesicles, magnetosomes); capsules, S-layers or slime layers; stalks or appendages; fruiting-body formation
Motility	Nonmotile; gliding motility; swimming (flagellar) motility; swarming; motile by gas vesicles
Metabolism	Mechanism of energy conservation (phototroph, chemoorganotroph, chemolithotroph); utilization of individual carbon, nitrogen, or sulfur compounds; fermentation of sugars; nitrogen fixation; growth factor requirements
Physiology	Temperature, pH, and salt ranges for growth; response to oxygen (aerobic, facultative, anaerobic); presence of catalase or oxidase; production of extracellular enzymes
Cell lipid chemistry	Fatty acids[b]; polar lipids; respiratory quinones
Cell wall chemistry	Presence or absence of peptidoglycan; amino acid composition of cross-links; presence or absence of cross-link interbridge
Other traits	Pigments; luminescence; antibiotic sensitivity; serotype; production of unique compounds, for example, antibiotics

[a]PHB, poly-β-hydroxybutyric acid (Section 3.10).
[b]Figure 16.19

bonds, rings, branched chains, or hydroxy groups (**Figure 16.19*a***). Hence, a fatty acid profile can often identify a particular bacterial species. For the analyses, fatty acids extracted from cell hydrolysates of a culture grown under standardized conditions are chemically derivatized to form their corresponding methyl esters. These now volatile derivatives are then identified by gas chromatography. A chromatogram showing the types and amounts of fatty acids from the unknown bacterium is then compared with a database containing the fatty acid profiles of thousands of reference bacteria grown under the same conditions. The best matches to that of the unknown are then selected (Figure 16.19*b*).

As a phenotypic trait for species identification and description, FAME does have some drawbacks. In particular, FAME analyses require rigid standardization because fatty acid profiles of an organism, like many other phenotypic traits, can vary as a function of temperature, growth phase (exponential versus stationary), and to a lesser extent, growth medium. Thus, for consistent results, it is necessary to grow the unknown organism on a specific medium and at a specific temperature for comparison of its fatty acid profile with those of organisms from the database that have been grown in the same way. For many organisms this is impossible, of course, and thus FAME analyses are limited to those organisms that can be grown under the specified conditions. In addition, the extent of variation in FAME profiles among strains of a species, a necessary consideration in studies to discriminate between species, is not yet well documented.

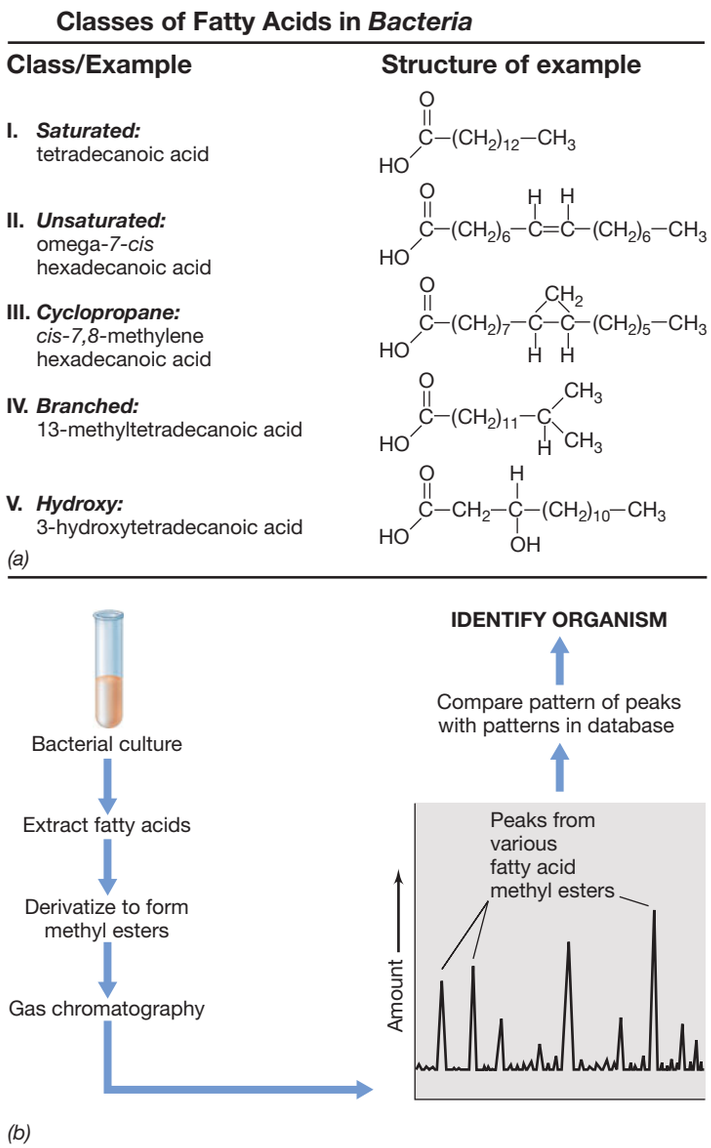


Figure 16.19 **Fatty acid methyl ester (FAME) analysis in bacterial identification.** *(a)* Classes of fatty acids in *Bacteria*. Only a single example is given of each class, but in fact, more than 200 different fatty acids are known from bacterial sources. A methyl ester contains a methyl group (CH_3) in place of the proton on the carboxylic acid group (COOH) of the fatty acid. *(b)* Procedure. Each peak from the gas chromatograph is due to one particular fatty acid methyl ester, and the peak height is proportional to the amount.

MiniQuiz

- What is FAME analysis?
- What are some of the drawbacks of FAME analysis?

16.11 Genotypic Analyses

Comparative analysis of the genome provides many traits for discriminating between species of bacteria. Genotypic analysis has particular appeal in microbial taxonomy because of the insights it provides at the DNA level. The method of genotypic analysis used depends on the question(s) posed, with DNA–DNA hybridization and DNA profiling among the more commonly used in microbial taxonomy.

DNA–DNA Hybridization

When two organisms share many identical or highly similar genes, their DNAs are expected to hybridize in approximate proportion to the similarities in their DNA sequences. For this reason, measurement of **DNA–DNA hybridization** between the genomes of two organisms provides a rough index of their similarity to each other. DNA–DNA hybridization therefore is useful for differentiating between organisms as a complement to small subunit rRNA gene sequencing.

We discussed the theory and methodology of nucleic acid hybridization in Section 11.2. In a hybridization experiment, genomic DNA isolated from one organism is made radioactive with radioactive phosphorus (^{32}P) or tritium (^{3}H), sheared to a relatively small size, heated to separate the two strands, and mixed with an excess of unlabeled DNA prepared in the same way from a second organism (**Figure 16.20**). The DNA mixture is then cooled to allow the single strands to reanneal. The double-stranded DNA is separated from any remaining unhybridized DNA. Following this, the amount of radioactivity in the hybridized DNA is determined and compared with the control, which is taken as 100% (Figure 16.20). Several nonradioactive DNA labeling systems are also available (↻ Section 11.2).

DNA–DNA hybridization is a sensitive method for revealing subtle differences in the genomes of two organisms and is therefore often useful for differentiating very similar organisms. Although there is no fixed convention as to how much hybridization between two DNAs is necessary to assign two organisms to the same taxonomic rank, hybridization values of 70% or greater are recommended as evidence that two isolates are the same species. Values of at least 25% are required to argue that two organisms are in the same genus (Figure 16.20*c*). DNAs from more distantly related organisms, for example, *Clostridium* (gram-positive) and *Salmonella* (gram-negative), would hybridize at only background levels, 10% or less.

GC Ratios

Another method that has been used to compare and describe bacteria is the **GC ratio** of their DNA. The GC ratio is the percentage of guanine (G) plus cytosine (C) in an organism's genomic DNA. GC ratios vary over a wide range, with values as low as 17% and as high as nearly 80% among species of *Bacteria* and *Archaea,* a range that is somewhat broader than for eukaryotes. It is typically the case that if two organisms' GC ratios differ by more than about 5%, they have few DNA sequences in common and are therefore unlikely to be closely related. However, two organisms can have identical GC ratios and yet be unrelated because very different nucleotide sequences are possible from DNA of the same overall base composition. In this case, the identical GC ratios are taxonomically misleading. Because gene sequence data are increasingly easy to obtain, GC ratios are applied less commonly in bacterial taxonomy than in the past.

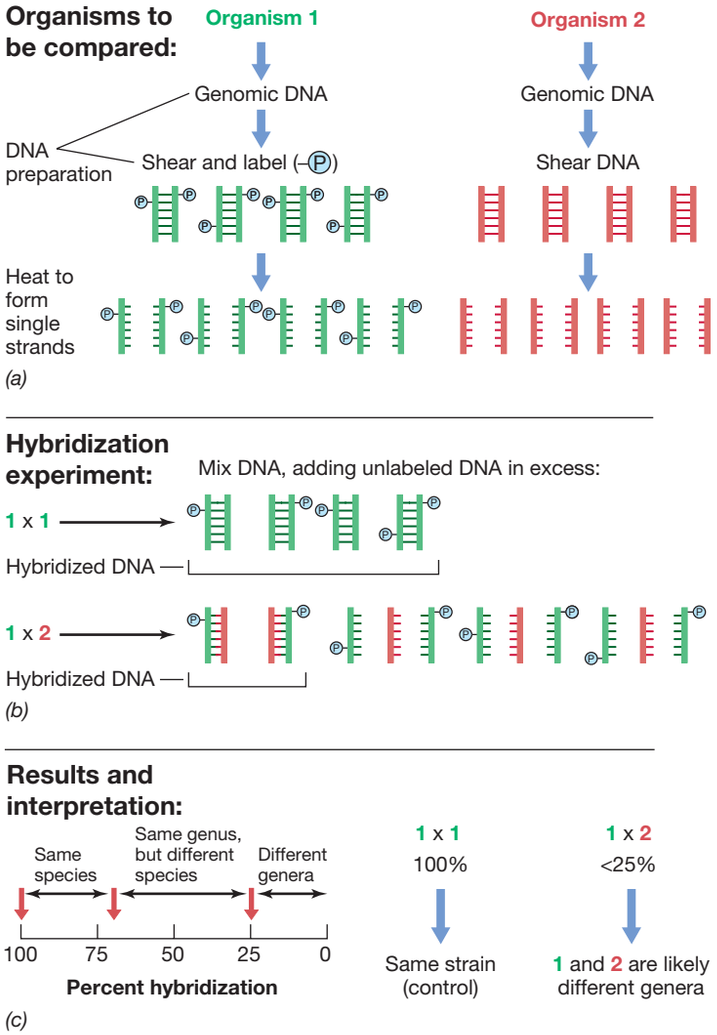


Figure 16.20 Genomic hybridization as a taxonomic tool.
(a) Genomic DNA is isolated from test organisms. One of the DNAs is labeled (shown here as radioactive phosphate in the DNA of Organism 1). *(b)* Excess unlabeled DNA is added to prevent labeled DNA from reannealing with itself. Following hybridization, hybridized DNA is separated from unhybridized DNA. Radioactivity in the hybridized DNA is measured. *(c)* Radioactivity in the control (Organism 1 DNA hybridizing to itself) is taken as the 100% hybridization value.

DNA Profiling Methods

There are several methods that generate DNA fragment patterns for analysis of genotypic similarity among bacterial strains. One of these DNA profiling methods, ribotyping, was described earlier. Other commonly used methods for rapid genotyping of bacteria

Table 16.3 *Some genotypic methods used in bacterial taxonomy*

Method	*Description/application*
DNA–DNA hybridization	Genome-wide comparison of sequence similarity. Useful for distinguishing species within a genus
DNA profiling	Ribotyping (Section 16.9), AFLP, rep-PCR (Figure 16.21). Rapid method to distinguish between species and strains within a species
Multilocus sequence typing	Strain typing using DNA sequences of multiple genes (Figure 16.22). High resolution, useful for distinguishing even very closely related strains within a species
GC ratio	Percentage of guanine–cytosine base pairs in the genome. If the GC ratio of two organisms differs by more than about 5%, they cannot be closely related, but organisms with similar or even identical GC ratios may be unrelated. Not much used now in taxonomy because of poor resolution
Multiple-gene or whole genome phylogenetic analyses	Application of cladistic methods to subsets of genes or to whole genomes from the organisms to be compared. Yields better phylogenetic picture than single-gene analyses

include *repetitive extragenic palindromic PCR* (*rep-PCR*) and *amplified fragment length polymorphism* (*AFLP*) (**Table 16.3**). In contrast to ribotyping, which focuses on a single gene, rep-PCR and AFLP assay for variations in DNA sequence throughout the genome.

The rep-PCR method is based on the presence of highly conserved repetitive DNA elements interspersed randomly around the bacterial chromosome. The number and positions of these elements differ between strains of a species that have diverged in genome sequence. Oligonucleotide primers designed to be complementary to these elements enable PCR amplification of elements from different genomic fragments that can be visualized by gel electrophoresis as patterns of bands. The patterns differ among different strains, giving what amounts to strain-specific DNA "fingerprints" (**Figure 16.21**).

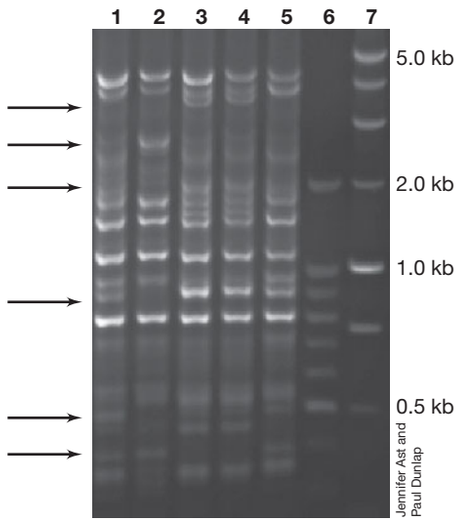


Figure 16.21 DNA fingerprinting with rep-PCR. Genomic DNAs from five strains (1–5) of a single species of bacteria were PCR-amplified using specific primers called *rep* (repetitive extragenic palindromic); the PCR products were separated in an agarose gel on the basis of size to generate DNA fingerprints. Arrows indicate some of the differing bands. Strains 3 and 4 have very similar DNA profiles. Lanes 6 and 7 are 100-bp and 1-kbp DNA size markers, respectively, used for estimating sizes of the DNA fragments.

AFLP is based on the digestion of genomic DNA with one or two restriction enzymes and selective PCR amplification of the resulting fragments, which are then separated by agarose gel electrophoresis. Strain-specific banding patterns similar to those of rep-PCR or other DNA fingerprinting methods are generated, with the large number of bands giving a high degree of discrimination between strains within a species. A technique similar to AFLP called T-RFLP (*terminal restriction fragment length polymorphism*) is widely used in phylogenetic analyses of natural microbial communities (Section 22.5).

Multilocus Sequence Typing

One of the limitations of both rRNA gene sequence analysis and ribotyping (but not of strain typing with rep-PCR or AFLP) is that these analyses focus on only a single gene, which may not provide sufficient information for unequivocal discrimination of bacterial strains. **Multilocus sequence typing (MLST)** circumvents this problem and is a powerful technique for characterizing strains within a species.

MLST consists of sequencing several different "housekeeping" genes from an organism and comparing their sequences with sequences of the same genes from different strains of the same organism. Housekeeping genes encode essential functions in cells and are located on the chromosome rather than on a plasmid. For each gene, an approximately 450-base-pair sequence is amplified using PCR and is then sequenced. Each nucleotide along the sequence is compared and differences are noted. Each difference, or sequence variant, is called an **allele** and is assigned a number. The strain being studied is then assigned a series of numbers as its allelic profile, or multilocus sequence type. In MLST, strains with identical sequences for a given gene have the same allele number for that gene, and two strains with identical sequences for all the genes have the same allelic profile (and would be considered identical by this method). The relatedness between each allelic profile is expressed in a dendrogram of linkage distances that vary from 0 (strains are identical) to 1 (strains are only distantly related, if at all) (**Figure 16.22**).

MLST has sufficient resolving power to distinguish among even very closely related strains. In practice, strains can be discriminated on the basis of a single nucleotide change in just one

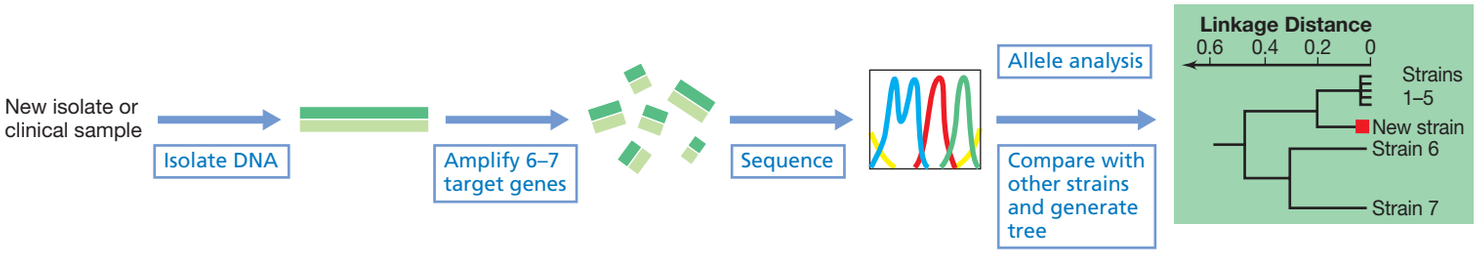


Figure 16.22 **Multilocus sequence typing.** Steps in MLST leading to a similarity phenogram are shown. Strains 1–5 are virtually identical, whereas strains 6 and 7 are distinct from one another and from strains 1–5.

of the analyzed genes. MLST is not useful, however, for comparing organisms above the species level; its resolution is too sensitive to yield meaningful information for grouping higher-order taxa such as genera and families.

MLST has found its greatest use in clinical microbiology, where it has been used to differentiate strains of a particular pathogen. This is important because some strains within a species—*Escherichia coli* K-12, for example—may be harmless, whereas others, such as strain O157:H7, can cause serious and even fatal infections (↻ Section 36.9). MLST is also widely used in epidemiological studies to track a virulent strain of a bacterial pathogen as it moves through a population, and in environmental studies to define the geographic distributions of strains.

Multigene and Whole Genome Analyses

Sequence comparisons of particular genes can provide valuable insight for taxonomy as well as phylogeny. The 16S rRNA gene, for example, the importance of which in microbial phylogeny was described in Section 16.6, has proven exceptionally useful in taxonomy as well, serving as a "gold standard" for the identification and description of new species. Other highly conserved genes, such as *recA,* which encodes a recombinase protein, and *gyrB,* which encodes a DNA gyrase protein, also can be useful for distinguishing bacteria at the species level. But for many reasons, including the facts that single-gene analyses give only a very limited genomic snapshot and that some genes may have been subject to horizontal gene flow that could lead to incorrect taxonomic conclusions, multigene and whole genomic analyses are becoming popular in microbial systematics.

The use of multiple genes for the identification and description of bacteria can avoid problems associated with reliance on individual genes. Multigene sequence analysis is similar to MLST (Figure 16.22), except that complete or nearly complete gene sequences are obtained and comparisons are made using cladistic methods (Section 16.7 and see Figure 16.24). By sequencing several functionally unrelated genes, one can obtain a more representative sampling of the genome than is possible with a single gene, and instances of horizontal gene transfer can be detected and those genes excluded from further consideration. Analyses of whole genome sequences provide an even greater depth of genotypic analysis. For example, differences between species in genome structure, including size and number of chromosomes, their GC content, and whether the chromosomes are linear or circular may have taxonomic significance. Comparative analysis of gene content (presence or absence of genes) and the order of genes in the genome can also provide insights.

MiniQuiz

- What is DNA fingerprinting and how is it useful for distinguishing bacteria?
- Hybridization of 90% of two organisms' DNA indicates that they are _____?
- How do AFLP and MLST differ from ribotyping?
- What advantages do multigene and whole genome analyses have over single-gene analyses?

16.12 The Species Concept in Microbiology

At present, there is no universally accepted concept of **species** for prokaryotes. Microbial systematics combines phenotypic, genotypic, and sequence-based phylogenetic data within a framework of standards and guidelines for describing and identifying prokaryotes, but the issue of what actually constitutes a prokaryotic species remains controversial. Because species are the fundamental units of biological diversity, how the concept of species is defined in microbiology determines how we distinguish and classify the units of diversity that make up the microbial world.

Current Definition of Prokaryotic Species

A prokaryotic species is defined operationally as a group of strains sharing a high degree of similarity in several independent traits. Traits currently considered most important for grouping strains together as a species include 70% or greater genomic DNA–DNA hybridization and 97% or greater identity (<3% difference) in 16S rRNA gene sequence (Sections 16.6 and 16.11). Experimental data suggest that these two criteria are valid, reliable, and consistent in identifying new species of prokaryotes (**Figure 16.23**). Based on genotypic criteria such as these, over 7000 species of *Bacteria* and *Archaea* have been formally recognized. What criteria should be used to define a genus, the next highest taxon, is more a matter of judgment, but 16S rRNA gene sequence differences of more than 5% from all other organisms is considered good evidence that an organism constitutes its own genus. Above the level of genus to the family, order, and other ranks of higher taxa, no consensus ribosomal RNA sequence-based criteria exist for delineating

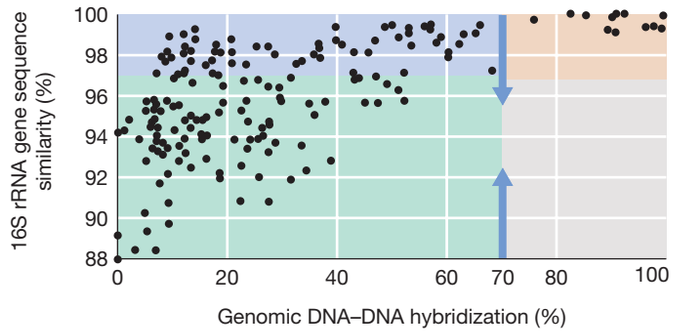


Figure 16.23 Relationship between 16S rRNA gene sequence similarity and genomic DNA–DNA hybridization for pairs of organisms. These data are the results from several independent experiments with various species of the domain *Bacteria*. Points in the darker tan region at the upper right represent pairs for which 16S rRNA gene sequence similarity and genomic hybridization were both very high; thus, in each case, the two organisms tested were clearly the same species. Points in the green box represent pairs that appear to be different species, and both methods show this. The blue box shows examples of pairs that seem to be different species as measured by genomic DNA–DNA hybridization, but not by 16S rRNA gene sequence. Note that above 70% DNA–DNA hybridization, no 16S rRNA gene similarities were found that were less than 97%. Data from Rosselló-Mora, R., and R. Amann. 2001. *FEMS Microbiol. Revs. 25*:39–67.

these ranks. **Table 16.4** gives an example of species definition in practice, listing relevant traits for the classification of the phototrophic purple bacterium *Allochromatium warmingii* from the domain down to the species level.

The Biological and Phylogenetic Species Concepts

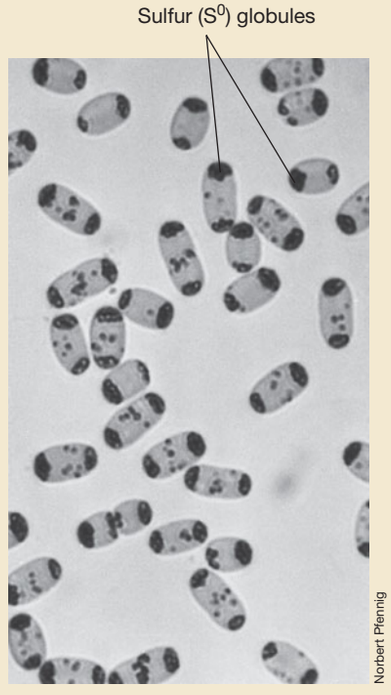
The *biological species concept* posits that a species is an interbreeding population of organisms that is reproductively isolated from other interbreeding populations; it is widely accepted as effective for defining species of eukaryotic organisms. However, the biological species concept is not meaningful for *Bacteria* and *Archaea* because they are haploid organisms that do not reproduce sexually.

An alternative to the biological species concept suitable for haploid organisms is the *phylogenetic species concept*. This concept defines a prokaryotic species as a group of strains that cluster closely with each other and are distinct from other groups of strains based on multiple-gene cladistic analyses (Sections 16.7 and 16.11). An example of such analysis using six genes from three species of the bacterium *Photobacterium* is shown in **Figure 16.24**.

The DNA sequence of the 16S rRNA gene has diverged relatively little throughout evolutionary history and therefore provides good family- and genus-level resolution. But 16S rRNA

Table 16.4 ***Taxonomic hierarchy for the purple sulfur bacterium*** **Allochromatium warmingii**

Taxon	*Name*	*Properties*	*Confirmed by*
Domain	*Bacteria*	Bacterial cells; rRNA gene sequences typical of *Bacteria*	Microscopy; 16S rRNA gene sequence analysis; presence of unique biomarkers, for example, peptidoglycan
Phylum	*Proteobacteria*	rRNA gene sequence typical of *Proteobacteria*	16S rRNA gene sequence analysis
Class	*Gammaproteobacteria*	Gram-negative bacteria; rRNA sequence typical of *Gammaproteobacteria*	Gram-staining, microscopy
Order	*Chromatiales*	Phototrophic purple bacteria	Characteristic pigments (↻ Figure 13.3)
Family	*Chromatiaceae*	Purple sulfur bacteria	Ability to oxidize H_2S and store S^0 within cells; microscopic observation of S^0 (see photo); 16S rRNA gene sequence
Genus	*Allochromatium*	Rod-shaped purple sulfur bacteria; <95% 16S gene sequence identity with all other genera	Microscopy (see photo)
Species	*warmingii*	Cells 3.5–4.0 μm × 5–11 μm; storage of sulfur mainly in poles of cell (see photo); <97% 16S gene sequence identity with all other species	Cell size measured microscopically with a micrometer; observation of polar position of S^0 globules in cells (see photo); 16S rRNA gene sequence



Photomicrograph of cells of the purple sulfur bacterium *Allochromatium warmingii*.

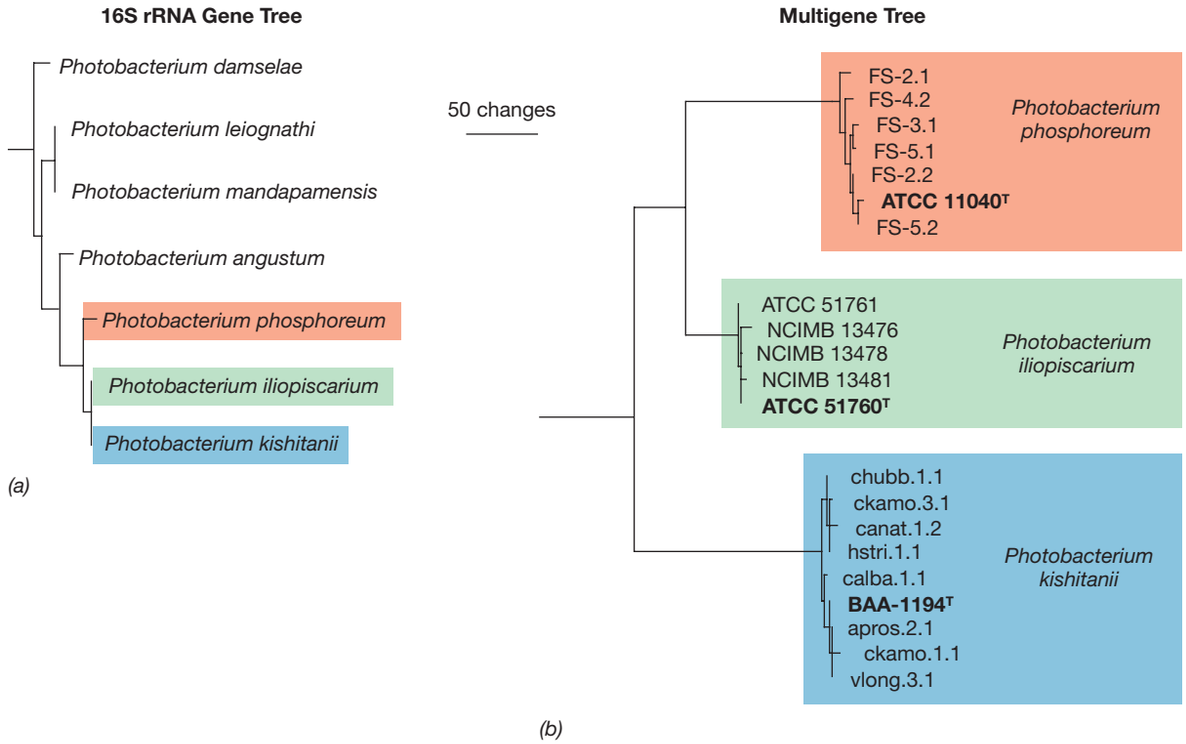


Figure 16.24 Multigene phylogenetic analysis. A phylogeny is shown for species in the genus *Photobacterium* (*Gammaproteobacteria*). *(a)* 16S rRNA gene tree, showing the species to be poorly resolved. *(b)* Multigene analysis based on combined parsimony analysis of the 16S rRNA gene, *gyrB*, and *luxABFE* genes in three *Photobacterium* species. Because the *gyrB* and *luxABFE* sequences diverge more than the 16S rRNA sequence, multigene analysis clearly resolves the 21 different strains analyzed into three distinct clades (phylogenetic species), *P. phosphoreum* (7 strains), *P. iliopiscarium* (5 strains), and *P. kishitanii* (9 strains). The scale bar indicates the branch length equal to a total of 50 nucleotide changes. The type strain of each species is designated with a superscript T appended to the strain designation and shown in bold. Phylogenetic analyses courtesy of Tory Hendy and Paul V. Dunlap.

UNIT 6

gene sequence analyses do not necessarily provide good species-level resolution when sequences differ very little, as is the case here (Figure 16.24*a*). For better species-level resolution *gyrB*, the gene encoding DNA gyrase subunit B, and *luxABFE*, a series of genes encoding luminescence enzymes (↻ Figure 1.1), were used, in addition to the 16S rRNA gene. The *gyrB* and *luxABFE* genes are less functionally constrained than the 16S rRNA gene, meaning that their sequences can vary more without a loss of function of the proteins they encode. The multigene analysis clearly resolves the strains of *Photobacterium* into three distinct evolutionary clades, and each clade can be considered a phylogenetic species (Figure 16.24*b*). In this way, multigene phylogenetic analyses and the phylogenetic species concept can be used to distinguish bacterial species that cannot be resolved by rRNA gene sequence analyses alone.

Speciation in Prokaryotes

How do new prokaryotic species arise? A likely possibility is by the process of periodic purges and selection within cell populations. Imagine a population of bacteria that originated from a single cell and that occupies a particular niche in a habitat. In theory, these cells are genetically identical. If cells in this population share a particular resource (for example, a key nutrient), the population is considered an **ecotype**.

Different ecotypes can coexist in a habitat, but each is only most successful within its prime niche in the habitat. However, within each ecotype, genes mutate at random over time as the cells grow. Most of these mutations are neutral and have no effect. However, if there is a beneficial mutation (one that increases fitness) in a cell in one of the ecotypes, that cell will produce more progeny over time, and this will purge the population of the original, less well-adapted cells (see Figure 16.10). Repeated rounds of mutation and selection in this ecotype lead it to become more and more distinct genetically from the other ecotypes. Then, given enough time, cells in this lineage will carry a sufficiently large set of unique traits that they emerge as their own species (**Figure 16.25**). Selection of strains bearing beneficial mutations can proceed gradually, or it can occur quite suddenly due to rapid environmental change. Note that this series of events *within* an ecotype has no effect on *other* ecotypes, because different ecotypes do not compete for the same resources (Figure 16.25).

It is also possible that a new genetic capacity in an ecotype may arise from genes obtained from cells of another ecotype by horizontal gene transfer, rather than from mutation and selection. The extent of horizontal gene transfer among bacteria is variable. Genome sequence analyses have revealed examples in which horizontal transfer of genes has apparently been frequent and others in which it has been rare (↻ Section 12.11). However,

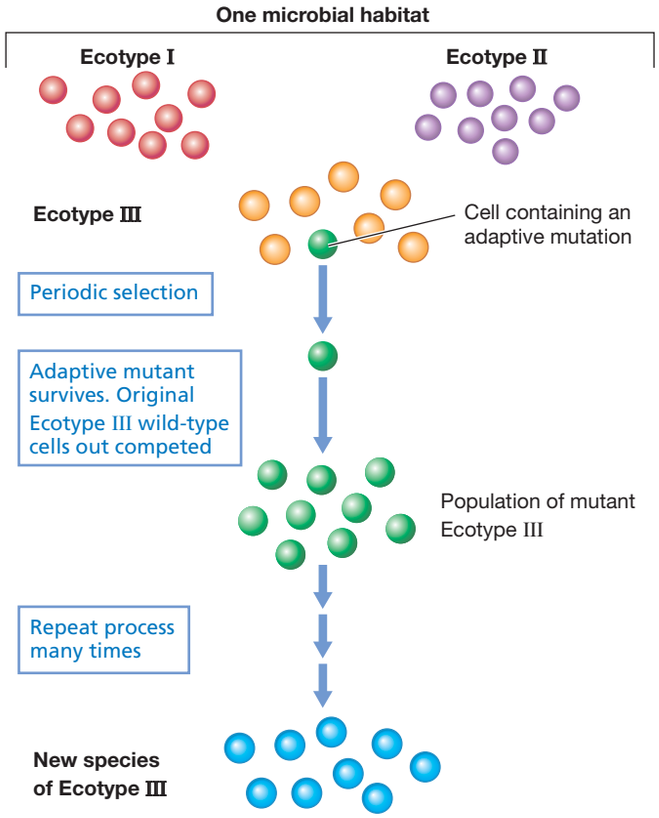


Figure 16.25 **A model for bacterial speciation.** Several ecotypes can coexist in a single microbial habitat, each occupying its own prime ecological niche. A cell within an ecotype that has a beneficial mutation may grow to become a population that eventually replaces the original ecotype. As this is repeated within a given ecotype, a genetically distinct population of cells arises that is a new species. Because other ecotypes do not compete for the same resources, they are unaffected by genetic and selection events outside their prime niche.

despite the potential impact of horizontal gene transfer for the instant acquisition of new metabolic capacities, speciation in *Bacteria* and *Archaea* is likely driven primarily by mutation and periodic selection (Figure 16.25). This is because horizontal gene transfer often confers upon the recipient cell only temporary benefits because the transferred genes will likely be lost if there is insufficient selective pressure to retain them.

How Many Prokaryotic Species Are There?

The result of nearly 4 billion years of evolution is the prokaryotic world we see today (Figure 16.16). Microbial taxonomists agree that no firm estimate of the number of prokaryotic species can be given at present, in part because of uncertainty about what defines a species. However, they also agree that in the final analysis, this number will be very large. Over 7000 species of *Bacteria* and *Archaea* are already known, based primarily on 16S rRNA gene sequencing, and thousands more, perhaps as many as 100,000–1,000,000, are thought likely to exist. If we were to attack this problem using multigene analyses, which provide better species-level resolution than 16S rRNA analyses alone (Figure 16.24), the species estimate would increase by one to two orders of magnitude.

Microbial community analyses (Sections 16.9, 22.5, and 22.6) indicate that we have only scratched the surface in our ability to culture the diversity of *Bacteria* and *Archaea* in nature. With the future application of more powerful tools—both molecular and cultural—for revealing diversity, it is likely that the already impressive list of species known will grow even larger. But the reality today is that an accurate estimate of prokaryotic species is simply out of reach of our current understanding and technology.

MiniQuiz

- How do the biological and phylogenetic species concepts differ? Which is suitable for prokaryotes and why?
- What is an ecotype?
- How many species of *Bacteria* and *Archaea* are already known? How many likely exist?

16.13 Classification and Nomenclature

We conclude this chapter with a brief description of how *Bacteria* and *Archaea* are classified and named; the science of *taxonomy*. Information is also presented on culture collections, which serve as repositories for scientific deposition of microbial cultures, on some key taxonomic resources available for microbiology, and the procedures for naming new species. The formal description of a new prokaryotic species and deposition of living cultures into a culture collection form an important foundation for prokaryotic systematics.

Taxa and Naming of Prokaryotes

Classification is the organization of organisms into progressively more inclusive groups on the basis of either phenotypic similarity or evolutionary relationship. The hierarchical nature of classification was shown in Table 16.4. A species is made up of one to several strains, and similar species are grouped into genera (singular, genus). Similar genera are grouped into families, families into orders, orders into classes, up to the domain, the highest-level taxon.

Nomenclature is the actual naming of organisms and follows the **binomial system** of nomenclature devised by the Swedish medical doctor and botanist, Carl Linnaeus, and used throughout biology; organisms are given genus names and species epithets. The names are Latin or Latinized Greek derivations, often descriptive of some key property of the organism, and are printed in *italics*. By classifying organisms into groups and naming them, we order the natural microbial world and make it possible to communicate effectively about all aspects of particular organisms, including their behavior, ecology, physiology, pathogenesis, and evolutionary relationships. The creation of new names must follow the rules described in *The International Code of Nomenclature of Bacteria*. This source presents the formal framework by which *Bacteria* and *Archaea* are to be officially named and the procedures by which existing names can be changed, for example, when new data warrants taxonomic rearrangements.

Table 16.5 *Some national microbial culture collections*

Collection	*Name*	*Location*	*Web address*
ATCC	American Type Culture Collection	Manassas, Virginia	**http://www.atcc.org**
BCCM/LMG	Belgium Coordinated Collection of Microorganisms	Ghent, Belgium	**http://bccm.belspo.be**
CIP	Collection de l'Institut Pasteur	Paris, France	**http://www.pasteur.fr**
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen	Braunschweig, Germany	**http://www.dsmz.de**
JCM	Japan Collection of Microorganisms	Saitama, Japan	**http://www.jcm.riken.go.jp**
NCCB	Netherlands Culture Collection of Bacteria	Utrecht, The Netherlands	**http://www.cbs.knaw.nl/nccb**
NCIMB	National Collection of Industrial, Marine and Food Bacteria	Aberdeen, Scotland	**http://www.ncimb.com**

Bergey's Manual and *The Prokaryotes*

Because taxonomy is largely a matter of scientific judgment, there is no "official" classification of *Bacteria* and *Archaea*. Presently, the classification system most widely accepted by microbiologists is that of *Bergey's Manual of Systematic Bacteriology,* a major taxonomic treatment of *Bacteria* and *Archaea* (see Appendix 2 for a list of genera and higher-order taxa from *Bergey's Manual*). Widely used, *Bergey's Manual* has served the community of microbiologists since 1923 and is a compendium of information on all recognized prokaryotes. Each chapter, written by experts, contains tables, figures, and other systematic information useful for identification purposes.

A second major source in bacterial diversity is *The Prokaryotes,* a reference that provides detailed information on the enrichment, isolation, and culture of *Bacteria* and *Archaea*. This work is available online by subscription through university libraries. Collectively, *Bergey's Manual* and *The Prokaryotes* offer microbiologists both the concepts as well as the details of the biology of *Bacteria* and *Archaea* as we know it today; they are the primary resources for microbiologists characterizing newly isolated organisms.

Culture Collections

National microbial culture collections (**Table 16.5**) are an important foundation of microbial systematics. These permanent collections catalog and store microorganisms and provide them upon request, usually for a fee, to researchers in academia, medicine, and industry. The collections play an important role in protecting microbial biodiversity, just as museums do in preserving plant and animal specimens for future study. However, unlike museums, which maintain collections of chemically preserved or dried, dead specimens, microbial culture collections store microorganisms as *viable cultures*, typically frozen or in a freeze-dried state. These storage methods maintain the cells indefinitely in a living state.

A related and key role of culture collections is as repositories for *type strains*. When a new species of bacteria is described in a scientific journal, a strain is designated as the nomenclatural type of the taxon for future taxonomic comparison with other strains of that species. Deposition of this type strain in the national culture collections of at least two countries, thereby making the strain publicly available, is a prerequisite for validation of the new species name. Some of the large national culture collections are listed in Table 16.5. Their websites contain searchable databases of strain holdings, together with information on the environmental sources of strains and publications on them.

Describing New Species

When a new prokaryote is isolated from nature and thought to be unique, a decision must be made as to whether it is sufficiently different from other prokaryotes to be described as a new taxon. To achieve formal validation of taxonomic standing as a new genus or species, a detailed description of the organism's characteristics and distinguishing traits, along with its proposed name, must be published, and, as just mentioned, viable cultures of the organism must be deposited in at least two international culture collections (Table 16.5). The manuscript describing and naming a new taxon undergoes peer review before publication. A major vehicle for the description of new taxa is the *International Journal of Systematic and Evolutionary Microbiology* (*IJSEM*), the official publication of record for the taxonomy and classification of *Bacteria* and *Archaea*. In each issue, the *IJSEM* publishes an approved list of newly validated names. By providing validation of newly proposed names, publication in *IJSEM* paves the way for their inclusion in *Bergey's Manual of Systematic Bacteriology*. Two websites provide listings of valid, approved bacterial names: List of Prokaryotic Names with Standing in Nomenclature (http://www.bacterio.cict.fr), and Bacterial Nomenclature Up-to-Date (http://www.dsmz.de/bactnom/bactname.htm).

The International Committee on Systematics of Prokaryotes (ICSP) is responsible for overseeing nomenclature and taxonomy of *Bacteria* and *Archaea*. The ICSP oversees the publication of *IJSEM* and the *International Code of Nomenclature of Bacteria*, and it gives guidance to several subcommittees that establish and revise standards for the description of new species in the different groups of prokaryotes.

MiniQuiz

- What roles do culture collections play in microbial systematics?
- What is the *IJSEM* and what taxonomic function does it fulfill?
- Why might viable cell cultures be of more use in microbial taxonomy than preserved specimens?

Big Ideas

16.1
Planet Earth is about 4.5 billion years old. The first evidence for microbial life can be found in rocks 3.86 billion years old. In rocks 3.5 billion years old or younger, microbial formations called stromatolites are abundant and show extensive microbial diversification.

16.2
Life may have first arisen at submerged hydrothermal springs, and the first self-replicating life forms may have been RNAs. Eventually, DNA evolved and the DNA plus RNA plus protein model for cellular life was fixed. Early microbial metabolism was anaerobic and likely chemolithotrophic, exploiting abundant abiotic sources of H_2, FeS, and H_2S. The earliest carbon metabolism may have been autotrophic.

16.3
Early *Bacteria* and *Archaea* diverged from a common ancestor as long as 4 billion years ago. Microbial metabolism diversified on early Earth with the evolution of methanogenesis and anoxygenic photosynthesis. Oxygenic photosynthesis eventually led to an oxic Earth, banded iron formations, and great bursts in metabolic and cellular evolution.

16.4
The eukaryotic cell developed from endosymbiotic events. In the most likely scenario, a H_2-producing species of *Bacteria* was incorporated as an endosymbiont into a H_2-consuming species of *Archaea* (the host). The modern eukaryotic cell is a chimera with genes and characteristics from both *Bacteria* and *Archaea*.

16.5
Evolution is descent with modification. Natural selection works by favoring the survival and reproductive success of organisms that by chance have mutations that confer high fitness under the existing environmental conditions.

16.6
Phylogeny, the evolutionary history of life, can be reconstructed through analysis of homologous DNA sequences. Genes encoding SSU rRNAs have been used as molecular clocks to construct a phylogeny of all organisms, prokaryotes as well as eukaryotes.

16.7
Analytical methods for evolutionary analysis include sequence alignment and construction of phylogenetic trees that, if rooted, indicate a path of evolution based on common ancestry. Character-state methods such as parsimony are commonly used for tree construction.

16.8
Life on Earth evolved in three major directions, forming the domains *Bacteria*, *Archaea*, and *Eukarya*. Each domain contains several major lineages. The universal tree of life shows that the two prokaryotic domains, *Bacteria* and *Archaea*, split from each other eons ago, and that *Eukarya* split from *Archaea* later in the history of life.

16.9
Phylogenetic analyses of SSU rRNA genes have led to the development of research tools useful in ecology and medicine. Key among these is FISH, which uses fluorescently labeled phylogenetic probes to identify organisms in a natural sample. Other key methods include microbial community analysis and ribotyping.

16.10
Systematics is the study of the diversity and relationships of living organisms. Polyphasic taxonomy is based on phenotypic, genotypic, and phylogenetic information. Phenotypic traits useful in taxonomy include morphology, motility, metabolism, and cell chemistry, especially lipid analyses.

16.11
Genotypic analysis examines traits of the genome. Bacterial species can be distinguished genotypically on the basis of DNA–DNA hybridization, DNA profiling, MLST, multigene or whole genome analyses, and by the GC content of their DNA.

16.12
At present a prokaryotic species is defined operationally based on shared genetic and phenotypic traits. The biological species concept is unsuitable for prokaryotes because their mode of reproduction is not sexual. New species of prokaryotes arise from periodic purging and selection within an ecotype, and the number of distinct species of prokaryotes in nature is surely enormous.

16.13
Formal recognition of a new prokaryotic species requires depositing a sample of the organism in culture collections and publishing the new species name and description. *Bergey's Manual of Systematic Bacteriology* and *The Prokaryotes* are major taxonomic compilations of *Bacteria* and *Archaea*.

Review of Key Terms

Allele a sequence variant of a given gene
Archaea phylogenetically related prokaryotes distinct from *Bacteria*
Bacteria phylogenetically related prokaryotes distinct from *Archaea*
Banded iron formation iron oxide–rich ancient sedimentary rocks containing zones of oxidized iron (Fe^{3+}) formed by oxidation of Fe^{2+} by O_2 produced by cyanobacteria
Binomial system the system devised by the Swedish scientist Carl Linnaeus for naming living organisms in which an organism is given a genus name and a species epithet
Cladistics phylogenetic methods that group organisms by their evolutionary relationships, not by their phenotypic similarities
Domain in a taxonomic sense, the highest level of biological classification
DNA–DNA hybridization the experimental determination of genomic similarity by measuring the extent of hybridization of DNA from one organism with that of another
Ecotype a population of genetically identical cells sharing a particular resource within an ecological niche
Endosymbiotic hypothesis the idea that a chemoorganotrophic bacterium and a cyanobacterium were stably incorporated into another cell type to give rise, respectively, to the mitochondria and chloroplasts of modern-day eukaryotes
Eukarya all eukaryotes: algae, protists, fungi, slime molds, plants, and animals
Evolution descent with modification; DNA sequence variation and the inheritance of that variation
FAME fatty acid methyl ester; a technique for identifying microorganisms from their fatty acids
FISH fluorescent *in situ* hybridization; a staining technique for phylogenetic studies
Fitness the capacity of an organism to survive and reproduce as compared to that of competing organisms
GC ratio in DNA from an organism, the percentage of the total nucleic acid that consists of guanine and cytosine bases
Horizontal gene transfer the transfer of DNA from one cell to another, possibly distantly related, cell
Molecular clock a DNA sequence, such as the gene for rRNA, that can be used as a comparative temporal measure of evolutionary divergence
Monophyletic in phylogeny, a group descended from one ancestor
Multilocus sequence typing (MLST) a taxonomic tool for classifying organisms from gene sequence variations in several housekeeping genes
Phylogenetic probe an oligonucleotide, sometimes made fluorescent by attachment of a dye, complementary in sequence to some sequence in rRNA
Phylogeny the evolutionary history of an organism
Phylum a major lineage of cells in one of the three domains of life
Proteobacteria a large group of phylogenetically related, gram-negative *Bacteria*
Ribosomal Database Project (RDP) a large database of small subunit (SSU) rRNA sequences that can be retrieved electronically and used in comparative rRNA sequence studies
Ribotyping a means of identifying microorganisms from analysis of DNA fragments generated from restriction enzyme digestion of the genes encoding their 16S rRNA
16S rRNA a large polynucleotide (~1500 bases) that functions as part of the small subunit of the ribosome of *Bacteria* and *Archaea* and from whose gene sequence evolutionary information can be obtained; its eukaryotic counterpart is 18S rRNA
Small subunit (SSU) rRNA RNA from the 30S ribosomal subunit of *Bacteria* and *Archaea* or the 40S ribosomal subunit of eukaryotes; that is, 16S or 18S rRNA, respectively
Species defined in microbiology as a group of strains that all share the same major properties and differ in one or more significant properties from other groups of strains; defined phylogenetically as a monophyletic, exclusive group based on DNA sequence analyses
Stromatolite a laminated microbial mat, typically built from layers of filamentous *Bacteria* and other microorganisms, which can become fossilized
Systematics the study of the diversity of organisms and their relationships; includes taxonomy and phylogeny
Taxonomy the science of identification, classification, and nomenclature
Universal phylogenetic tree a tree that shows the positions of representatives of all domains of cells

Review Questions

1. What is the age of planet Earth? When did the oceans form, and what is the age of the earliest known microfossils (Section 16.1)?
2. Under what conditions did life likely originate? What were the steps leading from prebiotic chemistry to living cells (Section 16.2)?
3. What kind of energy and carbon metabolisms likely characterized early cellular life (Section 16.2)?
4. Why was the evolution of cyanobacteria of such importance to the further evolution of life on Earth? What component of the geological record is used to date the evolution of cyanobacteria (Section 16.3)?
5. What is the hydrogen hypothesis and how does it relate to the endosymbiotic origin of the eukaryotic cell (Section 16.4)?
6. What does the phrase "descent with modification" imply about natural relationships among living organisms (Section 16.5)?
7. Why are SSU rRNA genes good choices for phylogenetic studies, and what are their limitations (Section 16.6)?
8. Describe the steps for determining an SSU phylogeny of three bacteria you have isolated from nature (Section 16.7).
9. What major evolutionary finding emerged from the study of rRNA sequences? How has this discovery supported the endosymbiotic theory of eukaryotic origins (Section 16.8)?
10. What major physiological and biochemical properties do *Archaea* share with *Eukarya* or with *Bacteria* (Section 16.8)?
11. What is FISH technology? Give an example of how it could be used (Section 16.9).
12. What major phenotypic and genotypic properties are used to classify organisms in bacterial taxonomy (Sections 16.10 and 16.11)?
13. What is measured in FAME analyses (Section 16.10)?

14. How does 16S rRNA gene sequence analysis differ from multilocus sequence typing as an identification tool (Section 16.11)?
15. How is multigene phylogenetic analysis an improvement over analyses based on individual genes (Sections 16.11 and 16.12)?
16. How is it thought that new bacterial species arise? How many bacterial species are there? Why don't we know this number more precisely (Section 16.12)?
17. What roles do microbial culture collections play in microbial systematics (Section 16.13)?

Application Questions

1. Compare and contrast the physical and chemical conditions on Earth at the time life first arose with conditions today. From a physiological standpoint, discuss at least two reasons why *animals* could not have existed on early Earth.
2. In what ways has microbial metabolism altered Earth's biosphere? How might life on Earth be different if oxygenic photosynthesis had not evolved?
3. For the following sequences, identify the phylogenetically informative sites (assume that the sequences are properly aligned). Identify also the phylogenetically neutral sites and those that are invariant.

   Taxon 1: TCCGTACGTTA

   Taxon 2: TCCCCACGGTT

   Taxon 3: TCGGTACCGTA

   Taxon 4: TCGGTACCGTA
4. Imagine that you have been given several bacterial strains from various countries around the world and that all the strains are thought to cause the same gastrointestinal disease and to be genetically identical. Upon carrying out a DNA fingerprint analysis of the strains, you find that four different strain types are present. What methods could you use to test whether the different strains are actually members of the same species?
5. Imagine that you have discovered a new form of microbial life, one that appears to represent a fourth domain. How would you go about characterizing the new organism and determining if it actually is evolutionarily distinct from *Bacteria, Archaea,* and *Eukarya*?

Need more practice? Test your understanding with Quantitative Questions; access additional study tools including tutorials, animations, and videos; and then test your knowledge with chapter quizzes and practice tests at **www.microbiologyplace.com**.