20 | PHYLOGENIES AND THE HISTORY OF LIFE

Figure 20.1 A bee's life is very different from a flower's, but the two organisms are related. Both are members of the domain Eukarya and have cells containing many similar organelles, genes, and proteins. (credit: modification of work by John Beetham)

Chapter Outline

Introduction

This bee and *Echinacea* flower (**Figure 20.1**) could not look more different, yet they are related, as are all living organisms on Earth. By following pathways of similarities and changes—both visible and genetic—scientists seek to map the evolutionary past of how life developed from single-celled organisms to the tremendous collection of creatures that have germinated, crawled, floated, swum, flown, and walked on this planet.

20.1 | Organizing Life on Earth

By the end of this section, you will be able to do the following:

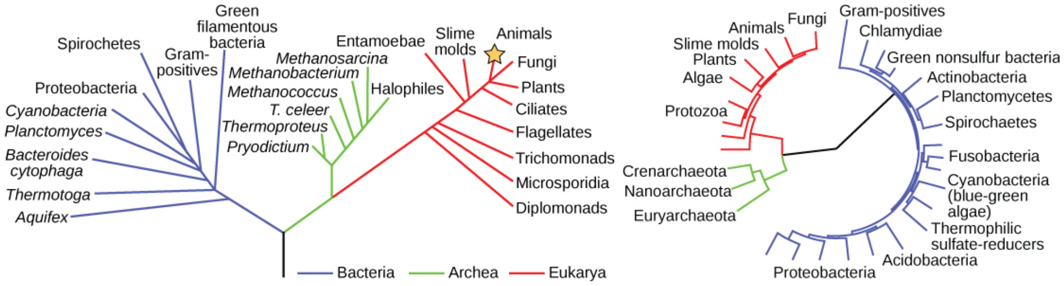
- Discuss the need for a comprehensive classification system
- List the different levels of the taxonomic classification system
- Describe how systematics and taxonomy relate to phylogeny
- Discuss a phylogenetic tree's components and purpose

In scientific terms, **phylogeny** is the evolutionary history and relationship of an organism or group of organisms. A phylogeny describes the organisim's relationships, such as from which organisms it may have evolved, or to which species it is most closely related. Phylogenetic relationships provide information on shared ancestry but not necessarily on how organisms are similar or different.

Phylogenetic Trees

Scientists use a tool called a phylogenetic tree to show the evolutionary pathways and connections among organisms. A **phylogenetic tree** is a diagram used to reflect evolutionary relationships among organisms or groups of organisms. Scientists consider phylogenetic trees to be a hypothesis of the evolutionary past since one cannot go back to confirm the proposed relationships. In other words, we can construct a “tree of life” to illustrate when different organisms evolved and to show the relationships among different organisms (**Figure 20.2**).

Unlike a taxonomic classification diagram, we can read a phylogenetic tree like a map of evolutionary history. Many phylogenetic trees have a single lineage at the base representing a common ancestor. Scientists call such trees **rooted**, which means there is a single ancestral lineage (typically drawn from the bottom or left) to which all organisms represented in the diagram relate. Notice in the rooted phylogenetic tree that the three domains—Bacteria, Archaea, and Eukarya—diverge from a single point and branch off. The small branch that plants and animals (including humans) occupy in this diagram shows how recent and miniscule these groups are compared with other organisms. Unrooted trees do not show a common ancestor but do show relationships among species.



(a) Rooted phylogenetic tree (b) Unrooted phylogenetic tree

Figure 20.2 Both of these phylogenetic trees show the relationship of the three domains of life—Bacteria, Archaea, and Eukarya—but the (a) rooted tree attempts to identify when various species diverged from a common ancestor while the (b) unrooted tree does not. (credit a: modification of work by Eric Gaba)

In a rooted tree, the branching indicates evolutionary relationships (**Figure 20.3**). The point where a split occurs, a **branch point**, represents where a single lineage evolved into a distinct new one. We call a lineage that evolved early from the root that remains unbranched a **basal taxon**. We call two lineages stemming from the same branch point **sister taxa**. A branch with more than two lineages is a **polytomy** and serves to illustrate where scientists have not definitively determined all of the relationships. Note that although sister taxa and polytomy do share an ancestor, it does not mean that the groups of organisms split or evolved from each other. Organisms in two taxa may have split at a specific branch point, but neither taxon gave rise to the other.

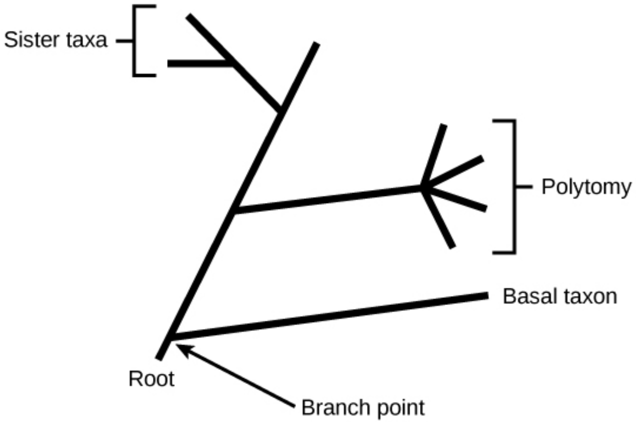


Figure 20.3 A phylogenetic tree's root indicates that an ancestral lineage gave rise to all organisms on the tree. A branch point indicates where two lineages diverged. A lineage that evolved early and remains unbranched is a basal taxon. When two lineages stem from the same branch point, they are sister taxa. A branch with more than two lineages is a polytomy.

The diagrams above can serve as a pathway to understanding evolutionary history. We can trace the pathway from the origin of life to any individual species by navigating through the evolutionary branches between the two points. Also, by starting with a single species and tracing back towards the "trunk" of the tree, one can discover species' ancestors, as well as where lineages share a common ancestry. In addition, we can use the tree to study entire groups of organisms.

Another point to mention on phylogenetic tree structure is that rotation at branch points does not change the information. For example, if a branch point rotated and the taxon order changed, this would not alter the information because each taxon's evolution from the branch point was independent of the other.

Many disciplines within the study of biology contribute to understanding how past and present life evolved over time; these disciplines together contribute to building, updating, and maintaining the “tree of life.” **Systematics** is the field that scientists use to organize and classify organisms based on evolutionary relationships. Researchers may use data from fossils, from studying the body part structures, or molecules that an organism uses, and DNA analysis. By combining data from many sources, scientists can construct an organism's phylogeny Since phylogenetic trees are hypotheses, they will continue to change as researchers discover new types of life and learn new information.

Limitations of Phylogenetic Trees

It may be easy to assume that more closely related organisms look more alike, and while this is often the case, it is not always true. If two closely related lineages evolved under significantly varied surroundings, it is possible for the two groups to appear more different than other groups that are not as closely related. For example, the phylogenetic tree in Figure 20.4 shows that lizards and rabbits both have amniotic eggs; whereas, frogs do not. Yet lizards and frogs appear more similar than lizards and rabbits.

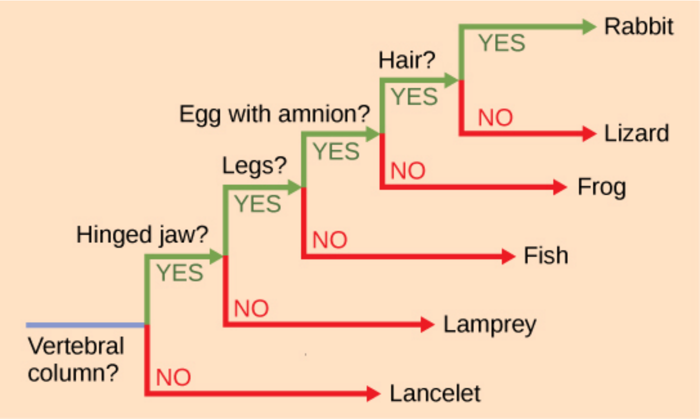


Figure 20.4 An organism that lacked a vertebral column roots this ladder-like phylogenetic tree of vertebrates. At each branch point, scientists place organisms with different characters in different groups based on shared characteristics.

Another aspect of phylogenetic trees is that, unless otherwise indicated, the branches do not account for length of time, only the evolutionary order. In other words, a branch's length does not typically mean more time passed, nor does a short branch mean less time passed— unless specified on the diagram. For example, in Figure 20.4, the tree does not indicate how much time passed between the evolution of amniotic eggs and hair. What the tree does show is the order in which things took place. Again using Figure 20.4, the tree shows that the oldest trait is the vertebral column, followed by hinged jaws, and so forth. Remember that any phylogenetic tree is a part of the greater whole, and like a real tree, it does not grow in only one direction after a new branch develops. Thus, for the organisms in Figure 20.4, just because a vertebral column evolved does not mean that invertebrate evolution ceased. It only means that a new branch formed. Also, groups that are not closely related, but evolve under similar conditions, may appear more phenotypically similar to each other than to a close relative.

LINK TO LEARNING

Head to this **website (http://openstaxcollege.org/l/tree_of_life)** to see interactive exercises that allow you to explore the evolutionary relationships among species.

Classification Levels

Taxonomy (which literally means “arrangement law”) is the science of classifying organisms to construct internationally shared classification systems with each organism placed into increasingly more inclusive groupings. Think about a grocery store's organization. One large space is divided into departments, such as produce, dairy, and meats. Then each department further divides into aisles, then each aisle into categories and brands, and then finally a single product. We call this organization from larger to smaller, more specific categories a hierarchical system.

The taxonomic classification system (also called the Linnaean system after its inventor, Carl Linnaeus, a Swedish botanist, zoologist, and physician) uses a hierarchical model. Moving from the point of origin, the groups become more specific, until one branch ends as a single species. For example, after the common beginning of all life, scientists divide organisms into three large categories called domains: Bacteria, Archaea, and Eukarya. Within each domain is a second category called a **kingdom**. After kingdoms, the subsequent categories of increasing specificity are: **phylum**, **class**, **order**, **family**, **genus**, and **species** (**Figure 20.5**).

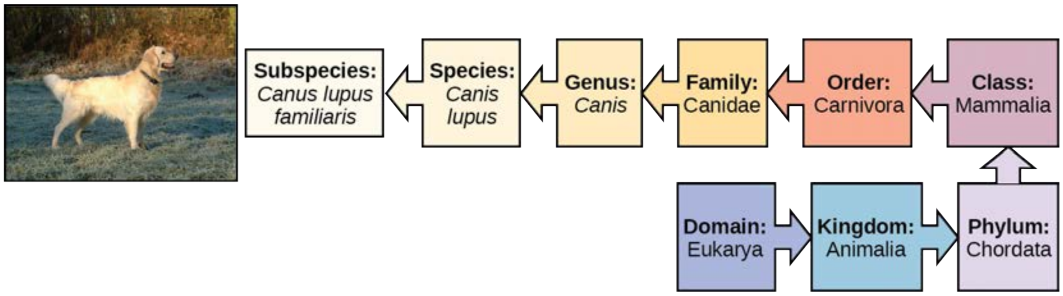


Figure 20.5 The taxonomic classification system uses a hierarchical model to organize living organisms into increasingly specific categories. The common dog, *Canis lupus familiaris*, is a subspecies of *Canis lupus*, which also includes the wolf and dingo. (credit “dog”: modification of work by Janneke Vreugdenhil)

The kingdom Animalia stems from the Eukarya domain. **Figure 20.5** above shows the classification for the common dog. Therefore, the full name of an organism technically has eight terms. For the dog it is: Eukarya, Animalia, Chordata, Mammalia, Carnivora, Canidae, *Canis,* and *lupus*. Notice that each name is capitalized except for species, and the genus and species names are italicized. Scientists generally refer to an organism only by its genus and species, which is its two-word scientific name, or **binomial nomenclature**. Therefore, the scientific name of the dog is *Canis lupus*. The name at each level is also a **taxon**. In other words, dogs are in order Carnivora. Carnivora is the name of the taxon at the order level; Canidae is the taxon at the family level, and so forth. Organisms also have a common name that people typically use, in this case, dog. Note that the dog is additionally a subspecies: the “*familiaris*” in *Canis lupus familiaris.* Subspecies are members of the same species that are capable of mating and reproducing viable offspring, but they are separate subspecies due to geographic or behavioral isolation or other factors.

Figure 20.6 shows how the levels move toward specificity with other organisms. Notice how the dog shares a domain with the widest diversity of organisms, including plants and butterflies. At each sublevel, the organisms become more similar because they are more closely related. Historically, scientists classified organisms using characteristics, but as DNA technology developed, they have determined more precise phylogenies.

visual CONNECTION

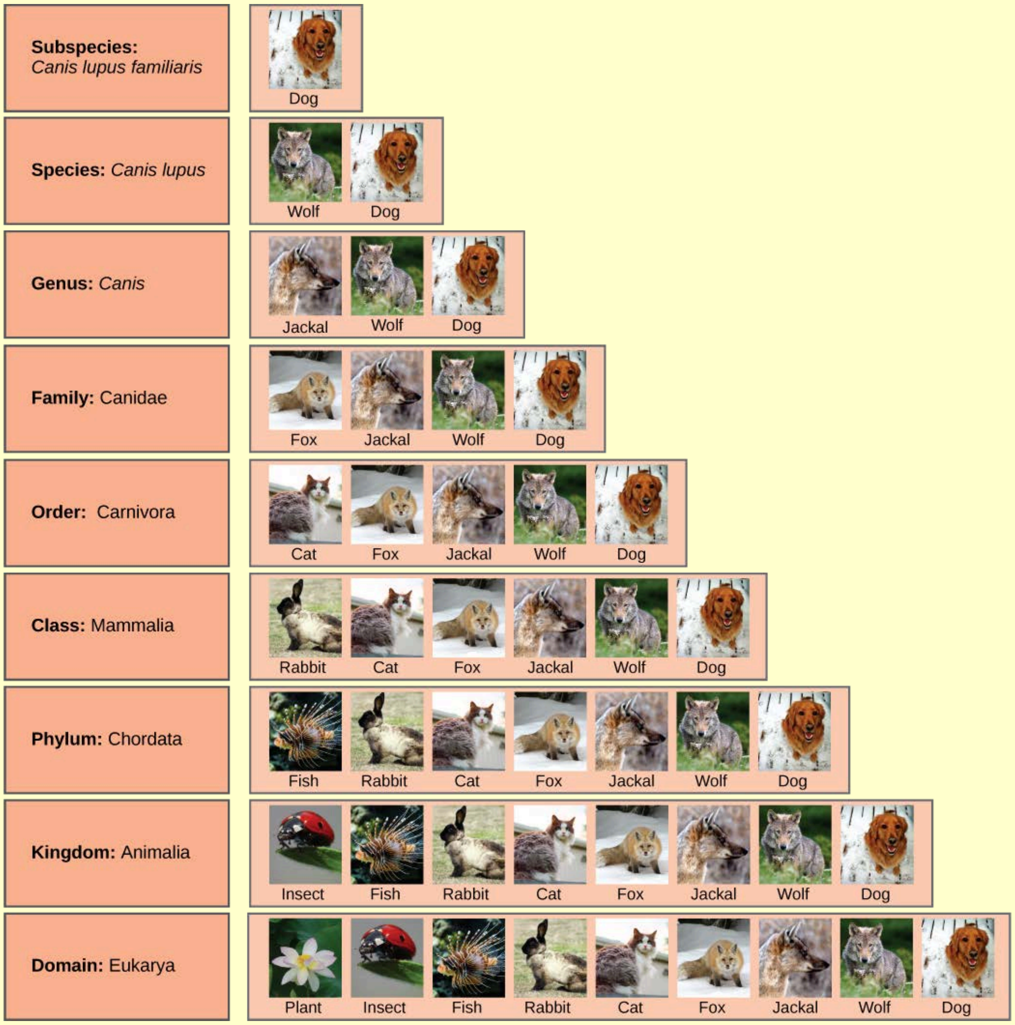


Figure 20.6 At each sublevel in the taxonomic classification system, organisms become more similar. Dogs and wolves are the same species because they can breed and produce viable offspring, but they are different enough to be classified as different subspecies. (credit “plant”: modification of work by "berduchwal"/Flickr; credit “insect”: modification of work by Jon Sullivan; credit “fish”: modification of work by Christian Mehlführer; credit “rabbit”: modification of work by Aidan Wojtas; credit “cat”: modification of work by Jonathan Lidbeck; credit “fox”: modification of work by Kevin Bacher, NPS; credit “jackal”: modification of work by Thomas A. Hermann, NBII, USGS; credit “wolf”: modification of work by Robert Dewar; credit “dog”: modification of work by "digital_image_fan"/Flickr)

At what levels are cats and dogs part of the same group?

LINK TO LEARNING

Visit this **website (http://openstaxcollege.org/l/classify_life)** to classify three organisms—bear, orchid, and sea cucumber—from kingdom to species. To launch the game, under Classifying Life, click the picture of the bear or the Launch Interactive button.

Recent genetic analysis and other advancements have found that some earlier phylogenetic classifications do not align with the evolutionary past; therefore, researchers must make changes and updates as new discoveries occur. Recall that phylogenetic trees are hypotheses and are modified as data becomes available. In addition, classification historically has focused on grouping organisms mainly by shared characteristics and does not necessarily illustrate how the various groups relate to each other from an evolutionary perspective. For example, despite the fact that a hippopotamus resembles a pig more than a whale, the hippopotamus may be the whale's closest living relative.

20.2 | Determining Evolutionary Relationships

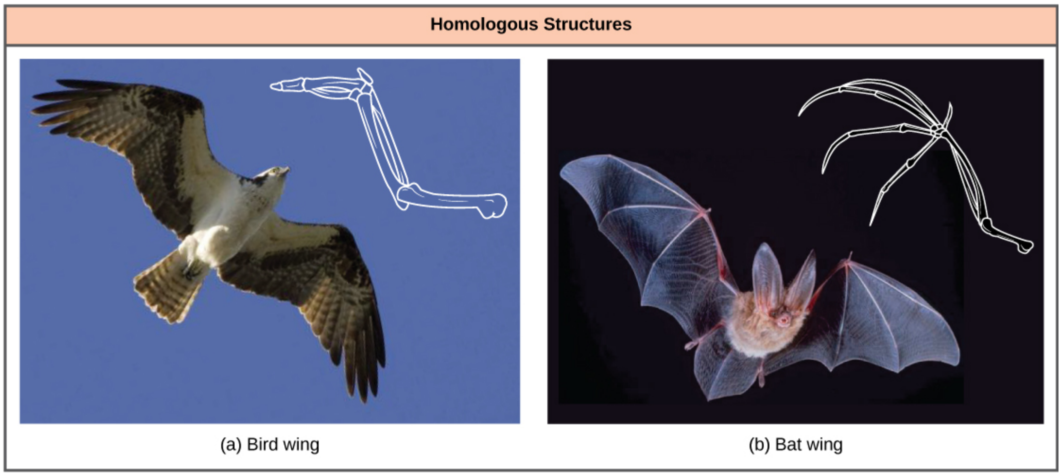
By the end of this section, you will be able to do the following:

- Compare homologous and analogous traits
- Discuss the purpose of cladistics
- Describe maximum parsimony

Scientists must collect accurate information that allows them to make evolutionary connections among organisms. Similar to detective work, scientists must use evidence to uncover the facts. In the case of phylogeny, evolutionary investigations focus on two types of evidence: morphologic (form and function) and genetic.

Two Options for Similarities

In general, organisms that share similar physical features and genomes are more closely related than those that do not. We refer to such features that overlap both morphologically (in form) and genetically as homologous structures. They stem from developmental similarities that are based on evolution. For example, the bones in bat and bird wings have homologous structures (**Figure 20.7**).



(a) Bird wing
(b) Bat wing

Figure 20.7 Bat and bird wings are homologous structures, indicating that bats and birds share a common evolutionary past. (credit a: modification of work by Steve Hillebrand, USFWS; credit b: modification of work by U.S. DOI BLM)

Notice it is not simply a single bone, but rather a grouping of several bones arranged in a similar way. The more complex the feature, the more likely any kind of overlap is due to a common evolutionary past. Imagine two people from different countries both inventing a car with all the same parts and in exactly the same arrangement without any previous or shared knowledge. That outcome would be highly improbable. However, if two people both invented a hammer, we can reasonably conclude that both could have the original idea without the help of the other. The same relationship between complexity and shared evolutionary history is true for homologous structures in organisms.

Misleading Appearances

Some organisms may be very closely related, even though a minor genetic change caused a major morphological difference to make them look quite different. Similarly, unrelated organisms may be distantly related, but appear very much alike. This usually happens because both organisms were in common adaptations that evolved within similar environmental conditions. When similar characteristics occur because of environmental constraints and not due to a close evolutionary relationship, it is an **analogy** or homoplasy. For example, insects use wings to fly like bats and birds, but the wing structure and embryonic origin is completely different. These are analogous structures (Figure 20.8).

Similar traits can be either homologous or analogous. Homologous structures share a similar embryonic origin. Analogous organs have a similar function. For example, the bones in a whale's front flipper are homologous to the bones in the human arm. These structures are not analogous. A butterfly or bird's wings are analogous but not homologous. Some structures are both analogous and homologous: bird and bat wings are both homologous and analogous. Scientists must determine which type of similarity a feature exhibits to decipher the organisms' phylogeny.

(a) Bat wing (b) Bird wing

(c) Insect wing

Figure 20.8 The (c) wing of a honeybee is similar in shape to a (b) bird wing and (a) bat wing, and it serves the same function. However, the honeybee wing is not composed of bones and has a distinctly different structure and embryonic origin. These wing types (insect versus bat and bird) illustrate an analogy—similar structures that do not share an evolutionary history. (credit a: modification of work by U.S. DOI BLM; credit b: modification of work by Steve Hillebrand, USFWS; credit c: modification of work by Jon Sullivan)

LINK TO LEARNING

This website (http://openstaxcollege.org/l/relationships) has several examples to show how appearances can be misleading in understanding organisms' phylogenetic relationships.

Molecular Comparisons

The advancement of DNA technology has given rise to **molecular systematics**, which is use of molecular data in taxonomy and biological geography (biogeography). New computer programs not only confirm many earlier classified organisms, but also uncover previously made errors. As with physical characteristics, even the DNA sequence can be tricky to read in some cases. For some situations, two very closely related organisms can appear unrelated if a mutation occurred that caused a shift in the genetic code. Inserting or deleting a mutation would move each nucleotide base over one place, causing two similar codes to appear unrelated.

Sometimes two segments of DNA code in distantly related organisms randomly share a high percentage of bases in the same locations, causing these organisms to appear closely related when they are not. For both of these situations, computer technologies help identify the actual relationships, and, ultimately, the coupled use of both morphologic and molecular information is more effective in determining phylogeny.

evolution CONNECTION

Why Does Phylogeny Matter?

Evolutionary biologists could list many reasons why understanding phylogeny is important to everyday life in human society. For botanists, phylogeny acts as a guide to discovering new plants that can be used to benefit people. Think of all the ways humans use plants—food, medicine, and clothing are a few examples. If a plant contains a compound that is effective in treating cancer, scientists might want to examine all of the compounds for other useful drugs.

A research team in China identified a DNA segment that they thought to be common to some medicinal plants in the family Fabaceae (the legume family). They worked to identify which species had this segment (Figure 20.9). After testing plant species in this family, the team found a DNA marker (a known location on a chromosome that enabled them to identify the species) present. Then, using the DNA to uncover phylogenetic relationships, the team could identify whether a newly discovered plant was in this family and assess its potential medicinal properties.

Figure 20.9 *Dalbergia sissoo (D. sissoo)* is in the Fabaceae, or legume family. Scientists found that *D. sissoo* shares a DNA marker with species within the Fabaceae family that have antifungal properties. Subsequently, researchers found that *D. sissoo* had fungicidal activity, supporting the idea that DNA markers are useful to screen plants with potential medicinal properties.

Building Phylogenetic Trees

How do scientists construct phylogenetic trees? After they sort the homologous and analogous traits, scientists often organize the homologous traits using **cladistics**. This system sorts organisms into clades: groups of organisms that descended from a single ancestor. For example, in Figure 20.10, all the organisms in the orange region evolved from a single ancestor that had amniotic eggs. Consequently, these organisms also have amniotic eggs and make a single clade, or a **monophyletic group**. Clades must include all descendants from a branch point.

visual CONNECTION

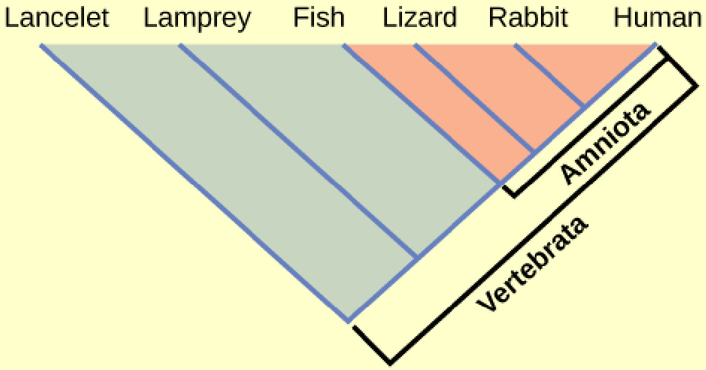


Figure 20.10 Lizards, rabbits, and humans all descend from a common ancestor that had an amniotic egg. Thus, lizards, rabbits, and humans all belong to the clade Amniota. Vertebrata is a larger clade that also includes fish and lamprey.

Which animals in this figure belong to a clade that includes animals with hair? Which evolved first, hair or the amniotic egg?

Clades can vary in size depending on which branch point one references. The important factor is that all organisms in the clade or monophyletic group stem from a single point on the tree. You can remember this because monophyletic breaks down into “mono,” meaning one, and “phyletic,” meaning evolutionary relationship. Figure 20.11 shows various clade examples. Notice how each clade comes from a single point; whereas, the non-clade groups show branches that do not share a single point.

visual CONNECTION

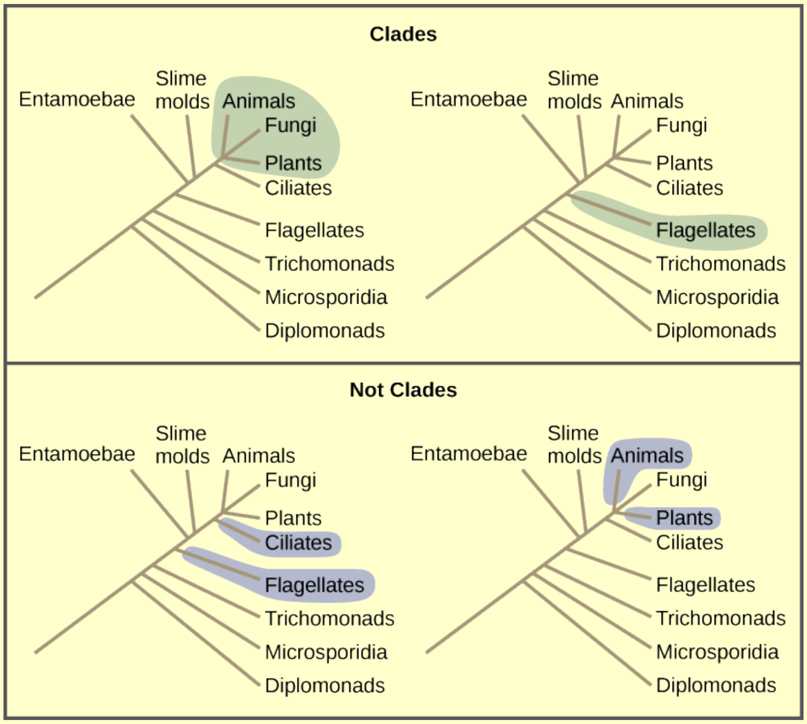


Figure 20.11 All the organisms within a clade stem from a single point on the tree. A clade may contain multiple groups, as in the case of animals, fungi and plants, or a single group, as in the case of flagellates. Groups that diverge at a different branch point, or that do not include all groups in a single branch point, are not clades.

What is the largest clade in this diagram?

Shared Characteristics

Organisms evolve from common ancestors and then diversify. Scientists use the phrase “descent with modification” because even though related organisms have many of the same characteristics and genetic codes, changes occur. This pattern repeats as one goes through the phylogenetic tree of life:

1. A change in an organism's genetic makeup leads to a new trait which becomes prevalent in the group.
2. Many organisms descend from this point and have this trait.
3. New variations continue to arise: some are adaptive and persist, leading to new traits.
4. With new traits, a new branch point is determined (go back to step 1 and repeat).

If a characteristic is found in the ancestor of a group, it is considered a **shared ancestral character** because all of the organisms in the taxon or clade have that trait. The vertebrate in Figure 20.10 is a shared ancestral character. Now consider the amniotic egg characteristic in the same figure. Only some of the organisms in Figure 20.10 have this trait, and to those that do, it is called a **shared derived character** because this trait derived at some point but does not include all of the ancestors in the tree.

The tricky aspect to shared ancestral and shared derived characters is that these terms are relative. We can consider the same trait one or the other depending on the particular diagram that we use. Returning to Figure 20.10, note that the amniotic egg is a shared ancestral character for the Amniota clade, while having hair is a shared derived character for some organisms in this group. These terms help scientists distinguish between clades in building phylogenetic trees.

Choosing the Right Relationships

Imagine being the person responsible for organizing all department store items properly—an overwhelming task. Organizing the evolutionary relationships of all life on Earth proves much more difficult: scientists must span enormous blocks of time and work with information from long-extinct organisms. Trying to decipher the proper connections, especially given the presence of homologies and analogies, makes the task of building an accurate tree of life extraordinarily difficult. Add to that advancing DNA technology, which now provides large quantities of genetic sequences for researchers to use and analzye. Taxonomy is a subjective discipline: many organisms have more than one connection to each other, so each taxonomist will decide the order of connections.

To aid in the tremendous task of describing phylogenies accurately, scientists often use the concept of **maximum parsimony**, which means that events occurred in the simplest, most obvious way. For example, if a group of people entered a forest preserve to hike, based on the principle of maximum parsimony, one could predict that most would hike on established trails rather than forge new ones.

For scientists deciphering evolutionary pathways, the same idea is used: the pathway of evolution probably includes the fewest major events that coincide with the evidence at hand. Starting with all of the homologous traits in a group of organisms, scientists look for the most obvious and simple order of evolutionary events that led to the occurrence of those traits.

LINK TO LEARNING

Head to this **website (http://openstaxcollege.org/l/using_parsimony)** to learn how researchers use maximum parsimony to create phylogenetic trees.

These tools and concepts are only a few strategies scientists use to tackle the task of revealing the evolutionary history of life on Earth. Recently, newer technologies have uncovered surprising discoveries with unexpected relationships, such as the fact that people seem to be more closely related to fungi than fungi are to plants. Sound unbelievable? As the information about DNA sequences grows, scientists will become closer to mapping the evolutionary history of all life on Earth.

20.3 | Perspectives on the Phylogenetic Tree

By the end of this section, you will be able to do the following:

- Describe horizontal gene transfer
- Illustrate how prokaryotes and eukaryotes transfer genes horizontally
- Identify the web and ring models of phylogenetic relationships and describe how they differ from the original phylogenetic tree concept

Phylogenetic modeling concepts are constantly changing. It is one of the most dynamic fields of study in all biology. Over the last several decades, new research has challenged scientists' ideas about how organisms are related. The scientific community has proposed new models of these relationships.

Many phylogenetic trees are models of the evolutionary relationship among species. Phylogenetic trees originated with Charles Darwin, who sketched the first phylogenetic tree in 1837 (**Figure 20.12a**). This served as a prototype for subsequent studies for more than a century. The phylogenetic tree concept with a single trunk representing a common ancestor, with the branches representing the divergence of species from this ancestor, fits well with the structure of many common trees, such as the oak (**Figure 20.12b**). However, evidence from modern DNA sequence analysis and newly developed computer algorithms has caused skepticism about the standard tree model's validity in the scientific community.

(a) (b)

Figure 20.12 The (a) concept of the “tree of life” dates to an 1837 Charles Darwin sketch. Like an (b) oak tree, the “tree of life” has a single trunk and many branches. (credit b: modification of work by "Amada44"/Wikimedia Commons)

Limitations to the Classic Model

Classical thinking about prokaryotic evolution, included in the classic tree model, is that species evolve clonally. That is, they produce offspring themselves with only random mutations causing the descent into the variety of modern-day and extinct species known to science. This view is somewhat complicated in eukaryotes that reproduce sexually, but the laws of Mendelian genetics explain the variation in offspring, again, to be a result of a mutation within the species. Scientists did not consider the concept of genes transferring between unrelated species as a possibility until relatively recently. Horizontal gene transfer (HGT), or lateral gene transfer, is the transfer of genes between unrelated species. HGT is an ever-present phenomenon, with many evolutionists postulating a major role for this process in evolution, thus complicating the simple tree model. Genes pass between species which are only distantly related using standard phylogeny, thus adding a layer of complexity to understanding phylogenetic relationships.

The various ways that HGT occurs in prokaryotes is important to understanding phylogenies. Although at present some do not view HGT as important to eukaryotic evolution, HGT does occur in this domain as well. Finally, as an example of the ultimate gene transfer, some scientists have proposed genome fusion theories between symbiotic or endosymbiotic organisms to explain an event of great importance—the evolution of the first eukaryotic cell, without which humans could not have come into existence.

Horizontal Gene Transfer

Horizontal gene transfer (HGT) is the introduction of genetic material from one species to another species by mechanisms other than the vertical transmission from parent(s) to offspring. These transfers allow even distantly related species to share genes, influencing their phenotypes. Scientists believe that HGT is more prevalent in prokaryotes, but that this process transfers only about 2% of the prokaryotic genome. Some researchers believe such estimates are premature: we must view the actual importance of HGT to evolutionary processes as a work in progress. As scientists investigate this phenomenon more thoroughly, they may reveal more HGT transfer. Many scientists believe that HGT and mutation are (especially in prokaryotes) a significant source of genetic variation, which is the raw material in the natural selection process. These transfers may occur between any two species that share an intimate relationship (Table 20.1).

Prokaryotic and Eukaryotic HGT Mechanisms Summary

	Mechanism	Mode of Transmission	Example
Prokaryotes	transformation	DNA uptake	many prokaryotes
	transduction	bacteriophage (virus)	bacteria

Table 20.1

Prokaryotic and Eukaryotic HGT Mechanisms Summary

	Mechanism	Mode of Transmission	Example
	conjugation	pilus	many prokaryotes
	gene transfer agents	phage-like particles	purple non-sulfur bacteria
Eukaryotes	from food organisms	unknown	aphid
	jumping genes	transposons	rice and millet plants
	epiphytes/parasites	unknown	yew tree fungi
	from viral infections		

Table 20.1

HGT in Prokaryotes

HGT mechanisms are quite common in the Bacteria and Archaea domains, thus significantly changing the way scientists view their evolution. The majority of evolutionary models, such as in the Endosymbiont Theory, propose that eukaryotes descended from multiple prokaryotes, which makes HGT all the more important to understanding the phylogenetic relationships of all extant and extinct species. The Endosymbiont Theory purports that the eukaryotes' mitochondria and the green plants' chloroplasts and flagellates originated as free-living prokaryotes that invaded primitive eukaryotic cells and become established as permanent symbionts in the cytoplasm.

Microbiology students are well aware that genes transfer among common bacteria. These gene transfers between species are the major mechanism whereby bacteria acquire resistance to antibiotics. Classically, scientists believe that three different mechanisms drive such transfers.

1. Transformation: bacteria takes up naked DNA
2. Transduction: a virus transfers the genes
3. Conjugation: a hollow tube, or pilus transfers genes between organisms

More recently, scientists have discovered a fourth gene transfer mechanism between prokaryotes. Small, virus-like particles, or **gene transfer agents (GTAs)** transfer random genomic segments from one prokaryote species to another. GTAs are responsible for genetic changes, sometimes at a very high frequency compared to other evolutionary processes. Scientists characterized the first GTA in 1974 using purple, non-sulfur bacteria. These GTAs, which are most likely bacteriophages that lost the ability to reproduce on their own, carry random DNA pieces from one organism to another. Controlled studies using marine bacteria have demonstrated GTAs' ability to act with high frequency. Scientists have estimated gene transfer events in marine prokaryotes, either by GTAs or by viruses, to be as high as 10^{13} per year in the Mediterranean Sea alone. GTAs and viruses are efficient HGT vehicles with a major impact on prokaryotic evolution.

As a consequence of this modern DNA analysis, the idea that eukaryotes evolved directly from Archaea has fallen out of favor. While eukaryotes share many features that are absent in bacteria, such as the TATA box (located in many genes' promoter region), the discovery that some eukaryotic genes were more homologous with bacterial DNA than Archaea DNA made this idea less tenable. Furthermore, scientists have proposed genome fusion from Archaea and Bacteria by endosymbiosis as the ultimate event in eukaryotic evolution.

HGT in Eukaryotes

Although it is easy to see how prokaryotes exchange genetic material by HGT, scientists initially thought that this process was absent in eukaryotes. After all, prokaryotes are but single cells exposed directly to their environment; whereas, the multicellular organisms' sex cells are usually sequestered in protected parts of the body. It follows from this idea that the gene transfers between multicellular eukaryotes should be more difficult. Scientists believe this process is rarer in eukaryotes and has a much smaller evolutionary impact than in prokaryotes. In spite of this, HGT between distantly related organisms is evident in several eukaryotic species, and it is possible that scientists will discover more examples in the future.

In plants, researchers have observed gene transfer in species that cannot cross-pollinate by normal means. Transposons or “jumping genes” have shown a transfer between rice and millet plant species. Furthermore, fungal species feeding on yew trees, from which the anti-cancer drug TAXOL® is derived from the bark, have acquired the ability to make taxol themselves, a clear example of gene transfer.

In animals, a particularly interesting example of HGT occurs within the aphid species (**Figure 20.13**). Aphids are insects that vary in color based on carotenoid content. Carotenoids are pigments that a variety of plants, fungi, and microbes produce, and they serve a variety of functions in animals, who obtain these chemicals from their food. Humans require carotenoids to synthesize vitamin A, and we obtain them by eating orange fruits and vegetables: carrots, apricots, mangoes, and sweet potatoes. Alternatively, aphids have acquired the ability to make the carotenoids on their own. According to DNA analysis, this ability is due to fungal genes transferring into the insect by HGT, presumably as the insect consumed fungi for food. A carotenoid enzyme, or desaturase, is responsible for the red coloration in certain aphids, and when mutation activates this gene, the aphids revert to their more common green color (**Figure 20.13**).

Figure 20.13 (a) Red aphids get their color from red carotenoid pigment. Genes necessary to make this pigment are present in certain fungi, and scientists speculate that aphids acquired these genes through HGT after consuming fungi for food. If mutation inactivates the genes for making carotenoids, the aphids revert back to (b) their green color. Red coloration makes the aphids considerably more conspicuous to predators, but evidence suggests that red aphids are more resistant to insecticides than green ones. Thus, red aphids may be more fit to survive in some environments than green ones. (credit a: modification of work by Benny Mazur; credit b: modification of work by Mick Talbot)

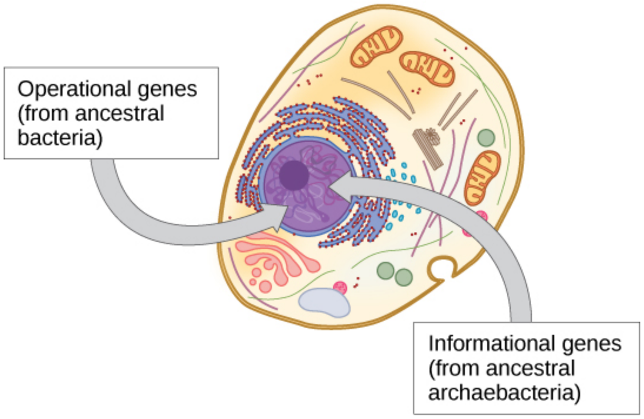
Genome Fusion and Eukaryote Evolution

Scientists believe the ultimate in HGT occurs through **genome fusion** between different prokaryote species when two symbiotic organisms become endosymbiotic. This occurs when one species is taken inside another species' cytoplasm, which ultimately results in a genome consisting of genes from both the endosymbiont and the host. This mechanism is an aspect of the Endosymbiont Theory, which most biologists accept as the mechanism whereby eukaryotic cells obtained their mitochondria and chloroplasts. However, the role of endosymbiosis in developing the nucleus is more controversial. Scientists believe that nuclear and mitochondrial DNA have different (separate) evolutionary origins, with the mitochondrial DNA derived from the bacteria's circular genomes ancient prokaryotic cells engulfed. We can regard mitochondrial DNA as the smallest chromosome. Interestingly enough, mitochondrial DNA is inherited only from the mother. The mitochondrial DNA degrades in sperm when the sperm degrades in the fertilized egg or in other instances when the mitochondria located in the sperm's flagellum fails to enter the egg.

Within the past decade, James Lake of the UCLA/NASA Astrobiology Institute proposed that the genome fusion process is responsible for the evolution of the first eukaryotic cells (**Figure 20.14a**). Using DNA analysis and a new mathematical algorithm, conditioned reconstruction (CR), his laboratory proposed that eukaryotic cells developed from an endosymbiotic gene fusion between two species, one an Archaea and the other a Bacteria. As mentioned, some eukaryotic genes resemble those of Archaea; whereas, others resemble those from Bacteria. An endosymbiotic fusion event, such as Lake has proposed, would clearly explain this observation. Alternatively, this work is new and the CR algorithm is relatively unsubstantiated, which causes many scientists to resist this hypothesis.

Lake's more recent work (**Figure 20.14b**) proposes that gram-negative bacteria, which are unique within their domain in that they contain two lipid bilayer membranes, resulted from an endosymbiotic fusion of archaeal and bacterial species. The double membrane would be a direct result of the endosymbiosis, with the endosymbiont picking up the second membrane from the host as it was internalized. Scientists have also used this mechanism to explain the double membranes in mitochondria and chloroplasts. Lake’s work is not without skepticism, and the biological science community still debates his ideas. In addition to Lake’s hypothesis, there are several other competing theories as to the origin of eukaryotes. How did the eukaryotic nucleus evolve? One theory is that the prokaryotic cells produced an additional membrane that surrounded the bacterial chromosome. Some bacteria have the DNA enclosed by two membranes; however, there is no evidence of a nucleolus or nuclear pores. Other proteobacteria also have membrane-bound chromosomes. If the eukaryotic nucleus evolved this way, we would expect one of the two types of prokaryotes to be more closely related to eukaryotes.

(a) Genome fusion by endosymbiosis



(b) Endosymbiotic formation of Gram-negative bacteria

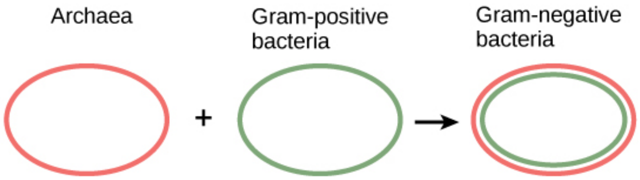


Figure 20.14 The scientific community now widely accepts the theory that mitochondria and chloroplasts are endosymbiotic in origin. More controversial is the proposal that (a) the eukaryotic nucleus resulted from fusing archaeal and bacterial genomes, and that (b) Gram-negative bacteria, which have two membranes, resulted from fusing Archaea and Gram-positive bacteria, each of which has a single membrane.

The **nucleus-first** hypothesis proposes that the nucleus evolved in prokaryotes first (**Figure 20.15a**), followed by a later fusion of the new eukaryote with bacteria that became mitochondria. The **mitochondria-first** hypothesis proposes that mitochondria were first established in a prokaryotic host (**Figure 20.15b**), which subsequently acquired a nucleus, by fusion or other mechanisms, to become the first eukaryotic cell. Most interestingly, the **eukaryote-first** hypothesis proposes that prokaryotes actually evolved from eukaryotes by losing genes and complexity (**Figure 20.15c**). All of these hypotheses are testable. Only time and more experimentation will determine which hypothesis data best supports.

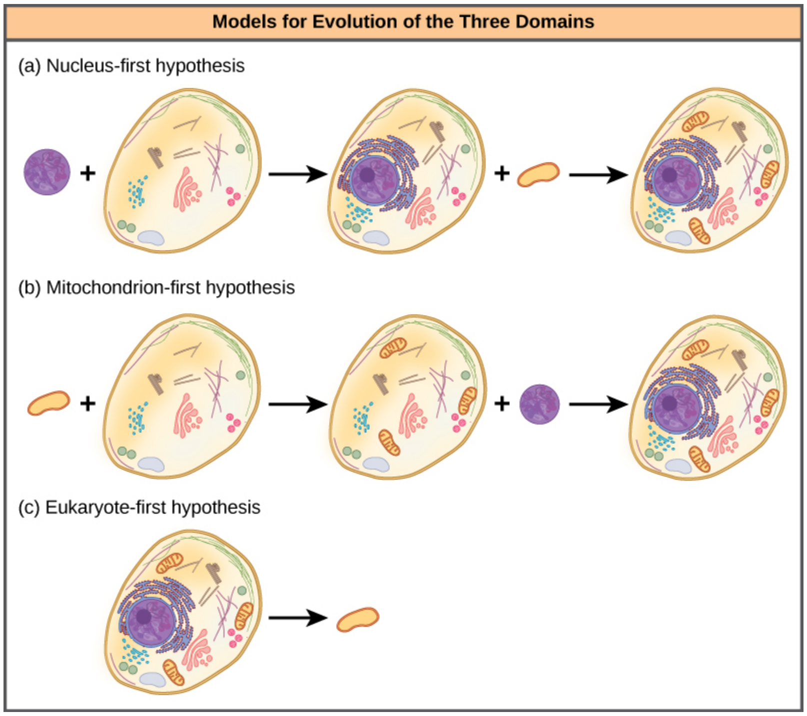


Figure 20.15 Three alternate hypotheses of eukaryotic and prokaryotic evolution are (a) the nucleus-first hypothesis, (b) the mitochondrion-first hypothesis, and (c) the eukaryote-first hypothesis.

Web and Network Models

Recognizing the importance of HGT, especially in prokaryote evolution, has caused some to propose abandoning the classic “tree of life” model. In 1999, W. Ford Doolittle proposed a phylogenetic model that resembles a web or a network more than a tree. The hypothesis is that eukaryotes evolved not from a single prokaryotic ancestor, but from a pool of many species that were sharing genes by HGT mechanisms. As **Figure 20.16a** shows, some individual prokaryotes were responsible for transferring the bacteria that caused mitochondrial development to the new eukaryotes; whereas, other species transferred the bacteria that gave rise to chloroplasts. Scientists often call this model the “ **web of life**.” In an effort to save the tree analogy, some have proposed using the *Ficus* tree (**Figure 20.16b**) with its multiple trunks as a phylogenetic way to represent a diminished evolutionary role for HGT.

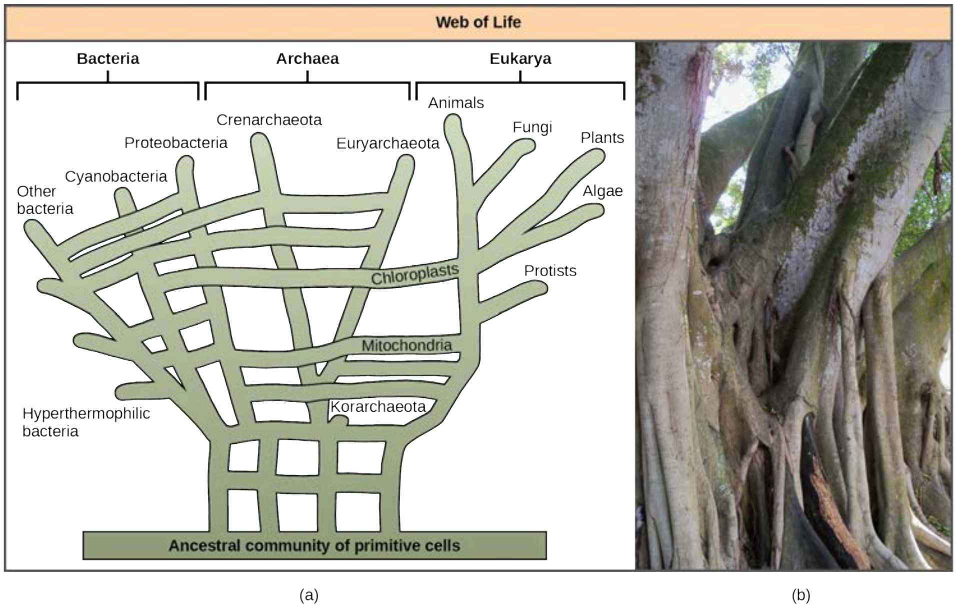


Figure 20.16 In W. Ford Doolittle's (a) phylogenetic model, the “tree of life” arose from a community of ancestral cells, has multiple trunks, and has connections between branches where horizontal gene transfer has occurred. Visually, this concept is better represented by (b) the multi-trunked **Ficus** than by an oak's single trunk similar to Darwin's tree in **Figure 20.12**. (credit b: modification of work by "psyberartist"/Flickr)

Ring of Life Models

Others have proposed abandoning any tree-like model of phylogeny in favor of a ring structure, the so-called “ **ring of life**” (**Figure 20.17**). This is a phylogenetic model where all three domains of life evolved from a pool of primitive prokaryotes. Lake, again using the conditioned reconstruction algorithm, proposes a ring-like model in which species of all three domains—Archaea, Bacteria, and Eukarya—evolved from a single pool of gene-swapping prokaryotes. His laboratory proposes that this structure is the best fit for data from extensive DNA analyses performed in his laboratory, and that the ring model is the only one that adequately takes HGT and genomic fusion into account. However, other phylogeneticists remain highly skeptical of this model.

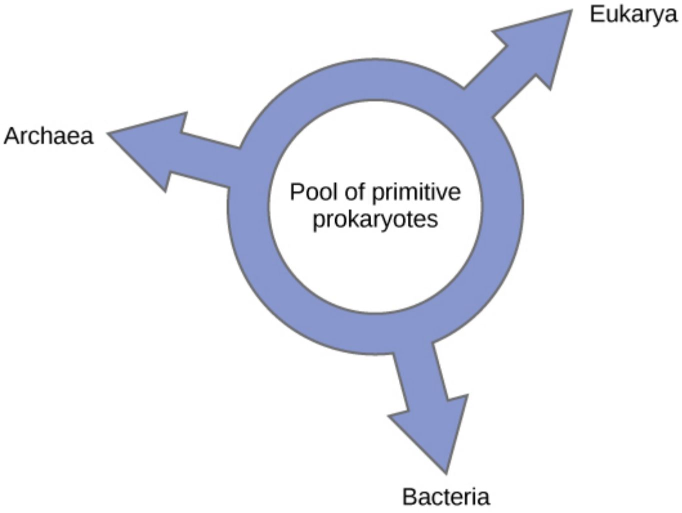


Figure 20.17 According to the “ring of life” phylogenetic model, the three domains of life evolved from a pool of primitive prokaryotes.

In summary, we must modify Darwin's “tree of life” model to include HGT. Does this mean abandoning the tree model completely? Even Lake argues that scientists should attempt to modify the tree model to allow it to accurately fit his data, and only the inability to do so will sway people toward his ring proposal.

This doesn’t mean a tree, web, or a ring will correlate completely to an accurate description of phylogenetic relationships of life. A consequence of the new thinking about phylogenetic models is the idea that Darwin's original phylogenetic tree concept is too simple, but made sense based on what scientists knew at the time. However, the search for a more useful model moves on: each model serves as hypotheses to test with the possibility of developing new models. This is how science advances. Researchers use these models as visualizations to help construct hypothetical evolutionary relationships and understand the massive amount of data that requires analysis.