

## 147

## 20.2 Phylogenies are inferred from morphological and molecular data

**Morphological and Molecular Homologies** Homologies are similarities due to shared ancestry. The more similar organisms' morphologies or DNA sequences are, the more likely those organisms are closely related.

**Sorting Homology from Analogy** **Analogy** is similarity due to convergent evolution, in which unrelated species develop similar features because natural selection has led to similar adaptations. Analogous structures are also called **homoplasies**. In general, when two complex structures share many similar features, it is more likely that those structures were inherited from a common ancestor. And when genes from different organisms share many nucleotide sequences, it is more likely that the genes are homologous.

### FOCUS QUESTION 20.2

What two complications may make it difficult to determine phylogenetic relationships based on morphological similarities between species? Give examples.

**Evaluating Molecular Homologies** How are molecular homologies identified? Molecular comparisons are often complicated by insertion or deletion mutations that change the lengths of homologous regions of DNA. Computer programs can identify and align homologous DNA segments properly for nucleotide comparisons. Statistical tools help distinguish "distant" homologies in divergent sequences from coincidental molecular homoplasies.

## 20.3 Shared characters are used to construct phylogenetic trees

**Cladistics** What is the most common methodology used to reconstruct phylogeny? **Cladistics** uses common ancestry to classify organisms. A **clade** consists of an ancestral species and all of its descendant species. Such a clade is **monophyletic**. A **paraphyletic** group excludes some species that share a common ancestor with other species in the group, and a **polyphyletic** group includes several groups with different ancestors.

How are clades identified? **Shared ancestral characters** are found in a particular clade but originated

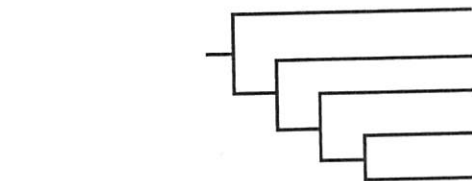
in an ancestor that is not a member of that clade. **Shared derived characters** are unique to a particular clade.

To determine the branching sequence of a group of related species, the group is compared to an **outgroup**, a species or group of species that diverged before the group being studied (the **ingroup**). A comparison of the characters that are present in each taxon of the ingroup indicates the sequence in which shared derived characters evolved and determines the branch points used to produce a phylogenetic tree.

### FOCUS QUESTION 20.3

Place the taxa (outgroup, A, B, C, and D) on the following phylogenetic tree based on the presence or absence of the characters 1–4 as shown in the table. Indicate before each branch point the number for the shared derived character that evolved in the ancestor of the clade.

	Taxa				
	Outgroup O	A	B	C	D
1	0	1	1	1	1
2	0	0	1	0	1
3	0	1	1	0	1
4	0	0	1	0	0



**Phylogenetic Trees with Proportional Branch Lengths** The branching pattern of most phylogenetic trees is relative rather than absolute, indicating only the order in which members of each clade last shared a common ancestor. The branch length of some trees can be scaled to reflect rates of evolutionary change (for instance, number of changes in DNA sequences) or time (using the dates of branch points as indicated in the fossil record).

**Maximum Parsimony** Systematists use morphological characters or molecular comparisons to choose among many possible phylogenetic trees using the principle of **maximum parsimony**—that the smallest number of evolutionary changes is the simplest explanation and thus the best hypothesis to consider first. Computer programs search for the most parsimonious trees.

**Phylogenetic Trees as Hypotheses** A phylogenetic tree represents the best hypothesis of the relationships among a set of species; the more data that can be compared, the more reliable the tree becomes.

Phylogenetic hypotheses can be used to make and test predictions. *Phylogenetic bracketing* predicts that features shared by two closely related organisms will be present in their common ancestor and all its descendants. Fossil discoveries, for example, support the prediction that dinosaurs, as descendants of the common ancestor of birds and crocodiles, built nests and brooded their eggs.

### FOCUS QUESTION 20.4

According to the principle of parsimony, the evolution of the four-chambered heart should place birds and mammals in the same clade. Why does the most accepted evolutionary tree show them as separate branches from the reptilian line?

### 20.4 Molecular clocks help track evolutionary time

**Molecular Clocks** Some regions of DNA appear to evolve at constant rates, and comparisons of the number of nucleotide substitutions in related genes can serve as **molecular clocks** to estimate the time since two species branched from their common ancestor.

Some genes appear to have a reliable average rate of evolution. Graphs that plot nucleotide or amino acid differences against the times for known evolutionary branch points can be used to estimate phylogenetic branchings that are not evident from the fossil record.

### FOCUS QUESTION 20.5

Assuming that harmful mutations are removed quickly from the gene pool, but neutral mutational changes that have little effect on fitness should occur at a constant rate, explain why different genes might have a different molecular clock rate.

Natural selection, which favors some DNA changes over others, may disrupt the smooth running of the molecular clock. Molecular clocks may be less reliable when used to date evolutionary divergences that occurred billions of years ago. When molecular clocks are calibrated using many genes, fluctuations due to natural selection may average out. For example, molecular and fossil-based estimates of divergence times in vertebrate evolution agree closely.

**Applying a Molecular Clock: The Origin of HIV** By comparing nucleotide sequences of HIV from samples taken at various times during the epidemic, researchers have observed a remarkably consistent rate of evolution. They estimate that the HIV-1 M strain first infected humans in the 1930s.

### 20.5 New information continues to revise our understanding of evolutionary history

**From Two Kingdoms to Three Domains** Historically, taxonomists divided the diversity of life into two kingdoms—plants and animals. In the five-kingdom system, the prokaryotes were set apart from the eukaryotes and placed in kingdom Monera. The Protista contained mostly unicellular eukaryotes while kingdoms Plantae, Fungi, and Animalia consisted of multicellular eukaryotes.

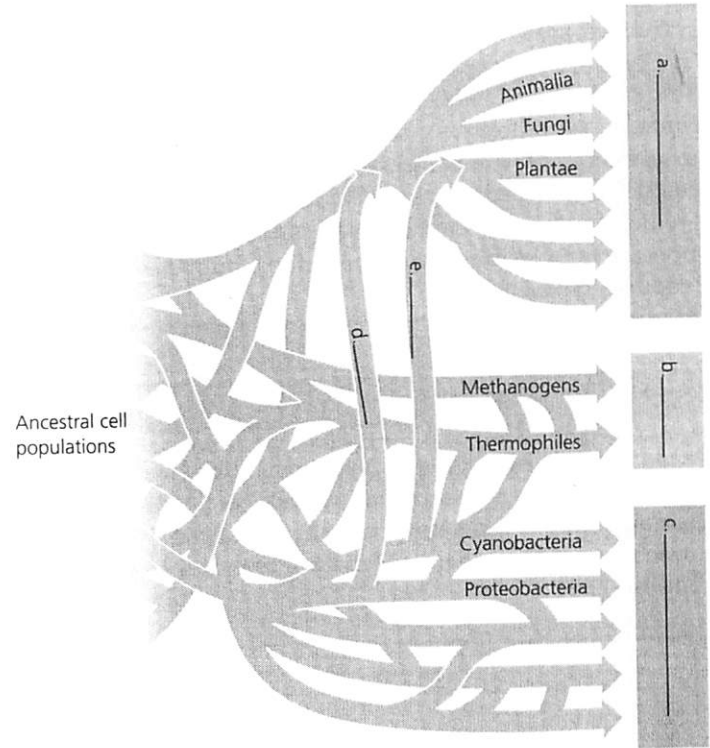
The current three-domain system creates a taxon above the kingdom level. The domains Bacteria and Archaea, although both consisting of single-celled prokaryotes, differ in many characteristics. The domain Eukarya includes all the eukaryotes: plants, fungi, animals, and many groups of mostly single-celled organisms.

**The Important Role of Horizontal Gene Transfer** The rRNA genes evolve so slowly that they have been used as the basis for constructing a tree of life. The tree's first major split represents the divergence of the bacteria from the other two domains. Analysis of other genes, however, suggests that eukaryotes share a more recent common ancestor with bacteria than with archaea.

Genome comparisons from the three domains indicate that **horizontal gene transfer**, perhaps through transposable elements, plasmid exchange, and viral infection, or even through fusions of different organisms, occurred during the early history of life. As a result, some scientists suggest that the early history of life is best represented as a tangled network of branches.

## FOCUS QUESTION 20.6

This tangled web of life represents a current hypothesis on the origin of the three domains. Identify the domains **a**, **b**, and **c**, and the two major episodes of horizontal gene transfer labeled **d** and **e**.



## Word Roots

- analog-** = proportion (*analogy*: similarity between two species due to convergent evolution)
- bi-** = two; **nom-** = name (*binomial*: two-part latinized format for naming of a species, consisting of the genus and specific epithet)
- clad-** = branch (*clade*: a group of species that includes an ancestral species and all its descendants)
- homo-** = like, resembling (*homoplasy*: a similar structure or molecular sequence that has evolved independently in two species)
- mono-** = one (*monophyletic*: pertaining to a group of taxa that consists of a common ancestor and all its descendants)
- parsi-** = few (*principle of parsimony*: the premise that a theory about nature should be the simplest explanation that is consistent with the facts)
- phylo-** = tribe; **-geny** = origin (*phylogeny*: the evolutionary history of a species or group of related species)

## Structure Your Knowledge

1. Draw a phylogenetic tree that best represents the relationships among taxa A–E described as follows: Taxon A is the basal taxon. B and C are sister taxa. They share a more recent common ancestor with D than with E.
2. Describe some of the tools that systematists use for constructing phylogenetic trees.
3. What is a molecular clock and how can it be used?

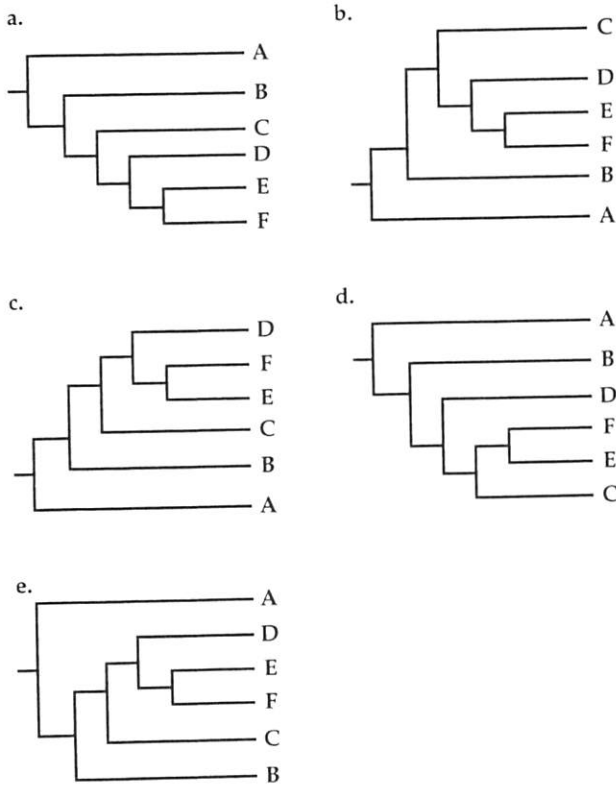
## Test Your Knowledge

**MULTIPLE CHOICE:** Choose the one best answer.

1. Related families are grouped into the next-highest taxon, which is called a(n)
  - a. class.
  - b. phylum.
  - c. order.
  - d. genus.
  - e. kingdom.



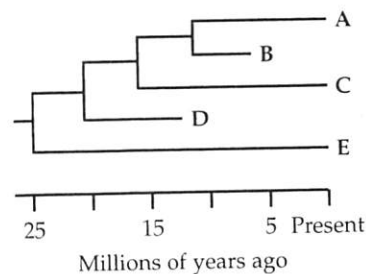
2. Four of the following trees describe the same phylogenetic relationships among taxa A, B, C, D, E, and F. Which tree shows a different phylogeny?



3. Convergent evolution may occur
- when ancestral structures are co-opted for new functions.
  - when homologous structures are adapted for different functions.
  - from adaptive radiation.
  - when species are widely separated geographically.
  - when species have similar ecological roles.
4. Which of the following provides the best example of analogous structures?
- forelimbs of bat and mole
  - treelike and shrublike silversword plants of Hawaii
  - wings of bee and hummingbird
  - skulls of apes and humans
  - hindlegs of Australian and North American moles
5. Which of the following is a shared derived character for monotreme, marsupial, and eutherian (placental) mammals?
- parental care
  - internal fertilization
  - amnion
  - production of milk for young
  - complete embryonic development inside a uterus

6. Analysis of which of the following data sets would produce the most reliable phylogenetic tree?
- DNA sequences of many homologous genes
  - the fossil record
  - morphology
  - homoplasies
  - Using all of these would produce the best supported phylogeny.
7. A taxon such as the class Reptilia—which does not include its relatives, the birds—is
- really an order.
  - a clade.
  - monophyletic.
  - paraphyletic.
  - polyphyletic.

The following phylogenetic tree includes the dates of divergence for taxa A through E as determined from the fossil record. Use this tree to answer questions 8 through 12.



8. Which letter represents the basal taxon?
9. Which letter(s) refer to extinct group(s)?
- A
  - B
  - C
  - E and D
  - B and D
10. Which two extant taxa are most closely related?
- A and B
  - A and C
  - A and E
  - C and D
  - C and E
11. How many million years ago did taxa A and D last share a common ancestor?
- 25
  - 20
  - 15
  - 10
  - cannot tell
12. The technique that enables biologists to make predictions about certain characteristics of taxon B by comparing shared derived characters of taxa A and C is called
- cladistics.
  - systematics.
  - phylogenetic bracketing.
  - maximum parsimony.
  - taxonomy.

13. The greatest number of shared derived characters should be found in two organisms that were traditionally placed in the same
- order.
  - domain.
  - family.
  - class.
  - phylum.
14. Which of the following approaches would allow a biologist studying the evolution of four similar species of birds to choose the best phylogenetic tree from all possible phylogenies?
- Draw the simplest tree and choose that one.
  - From a comparison of DNA sequences, determine the number of evolutionary events required for each tree and then choose the most parsimonious tree.
  - Compare the entire genomes of all species; the two most similar genomes are the two species that are most closely related.
  - Determine which species can interbreed; those that can interbreed evolved from a common ancestor most recently.
  - Choose the tree that has the most evolutionary changes, as this would be the most likely explanation for how these very similar birds evolved into four distinct species.
15. Which of the following segments of DNA would likely have the fastest molecular clock rate?
- noncoding DNA that has a regulatory function
  - a pseudogene (gene that has lost sequences needed for expression)
  - a gene for an essential enzyme
  - a gene for rRNA
  - a gene for a cytochrome involved in cellular respiration
16. Which of the following is the *best* description of our current hypothesis of the tree of life?
- The tree of life consists of three domains: Bacteria, Archaea, and Eukarya.
  - The base of the tree of life is still uncertain because the molecular clock is not accurate for evolutionary events that occurred that long ago.
  - The domain Archaea is known to be the first branch; domains Bacteria and Eukarya are more closely related to each other.
  - There was substantial horizontal gene transfer between different organisms during the early history of life.
  - Both a and d represent our current hypothesis.