Chapter 26
Phylogeny and the Tree of Life
Overview: Investigating the Tree of Life

• Legless lizards have evolved independently in several different groups
• **Phylogeny** is the evolutionary history of a species or group of related species
• The discipline of **systematics** classifies organisms and determines their evolutionary relationships
• Systematists use fossil, molecular, and genetic data to infer evolutionary relationships
Concept 26.1: Phylogenies show evolutionary relationships

- **Taxonomy** is the ordered division and naming of organisms
Binomial Nomenclature

• In the 18th century, Carolus Linnaeus published a system of taxonomy based on resemblances
• Two key features of his system remain useful today: two-part names for species and hierarchical classification
• The two-part scientific name of a species is called a **binomial**

• The first part of the name is the **genus**

• The second part, called the specific epithet, is unique for each species within the genus

• The first letter of the genus is capitalized, and the entire species name is italicized

• Both parts together name the species (not the specific epithet alone)
Hierarchical Classification

- Linnaeus introduced a system for grouping species in increasingly broad categories.
- The taxonomic groups from broad to narrow are domain, kingdom, phylum, class, order, family, genus, and species.
- A taxonomic unit at any level of hierarchy is called a taxon.
- The broader taxa are not comparable between lineages.
  - For example, an order of snails has less genetic diversity than an order of mammals.
Figure 26.3

Species:
*Panthera pardus*

Genus:
*Panthera*

Family:
Felidae

Order:
Carnivora

Class:
Mammalia

Phylum:
Chordata

Domain:
Animalia

Domain:
Archaea

Domain:
Bacteria

Domain:
Eukarya
Linking Classification and Phylogeny

• Systematists depict evolutionary relationships in branching **phylogenetic trees**
Figure 26.4

Order | Family | Genus | Species
--- | --- | --- | ---
Canivora | Felidae | Panthera | *Panthera pardus* (leopard)
 | Mustelidae | Taxidea | *Taxidea taxus* (American badger)
 | | Lutra | *Lutra lutra* (European otter)
 | Canidae | Canis | *Canis latrans* (coyote)
 | | | *Canis lupus* (gray wolf)
• Linnaean classification and phylogeny can differ from each other

• Systematists have proposed the **PhyloCode**, which recognizes only groups that include a common ancestor and all its descendants
• A phylogenetic tree represents a hypothesis about evolutionary relationships
• Each **branch point** represents the divergence of two species
• **Sister taxa** are groups that share an immediate common ancestor
• A **rooted** tree includes a branch to represent the last common ancestor of all taxa in the tree
• A **basal taxon** diverges early in the history of a group and originates near the common ancestor of the group
• A **polytomy** is a branch from which more than two groups emerge
Figure 26.5

Branch point:
where lineages diverge

This branch point represents the common ancestor of taxa A–G.

This branch point forms a polytomy: an unresolved pattern of divergence.

ANCESTRAL LINEAGE

Sister taxa

Basal taxon

Taxon A

Taxon B

Taxon C

Taxon D

Taxon E

Taxon F

Taxon G
What We Can and Cannot Learn from Phylogenetic Trees

• Phylogenetic trees show patterns of descent, not phenotypic similarity
• Phylogenetic trees do not indicate when species evolved or how much change occurred in a lineage
• It should not be assumed that a taxon evolved from the taxon next to it
Applying Phylogenies

- Phylogeny provides important information about similar characteristics in closely related species
- A phylogeny was used to identify the species of whale from which “whale meat” originated
RESULTS

Minke (Southern Hemisphere)
- Unknowns #1a, 2, 3, 4, 5, 6, 7, 8

Minke (North Atlantic)
- Unknown #9

Humpback (North Atlantic)
- Humpback (North Pacific)
- Unknown #1b

Gray

Blue

- Unknowns #10, 11, 12
- Unknown #13

Fin (Mediterranean)
- Fin (Iceland)
Concept 26.2: Phylogenies are inferred from morphological and molecular data

- To infer phylogenies, systematists gather information about morphologies, genes, and biochemistry of living organisms
Morphological and Molecular Homologies

• Phenotypic and genetic similarities due to shared ancestry are called homologies

• Organisms with similar morphologies or DNA sequences are likely to be more closely related than organisms with different structures or sequences
Sorting Homology from Analogy

• When constructing a phylogeny, systematists need to distinguish whether a similarity is the result of homology or **analogy**
• Homology is similarity due to shared ancestry
• Analogy is similarity due to convergent evolution
• Convergent evolution occurs when similar environmental pressures and natural selection produce similar (analogous) adaptations in organisms from different evolutionary lineages.
• Bat and bird wings are homologous as forelimbs, but analogous as functional wings.
• Analogous structures or molecular sequences that evolved independently are also called **homoplasies**.
• Homology can be distinguished from analogy by comparing fossil evidence and the degree of complexity.
• The more complex two similar structures are, the more likely it is that they are homologous.
Evaluating Molecular Homologies

• Systematists use computer programs and mathematical tools when analyzing comparable DNA segments from different organisms
Figure 26.8-1

1. C C A T C A G A G T C C
2. C C A T C A G A G T C C
Figure 26.8-2

1. Deletion

2. Insertion

GTA
Figure 26.8-3

1

1 C C A T C A G A G T C C
2 C C A T C A G A G T C C

2

1 C C A T C A G A G T C C
2 C C A T C A G A G T C C

2

1 C C A T C A G A G T C C
2 C C A T C A G A G T C C

3

1 C C A T C A A G T C C
2 C C A T G T A C A G A G T C C
Figure 26.8-4

Deletion

Insertion

G T A

© 2011 Pearson Education, Inc.
• It is also important to distinguish homology from analogy in molecular similarities
• Mathematical tools help to identify molecular homoplasies, or coincidences
• **Molecular systematics** uses DNA and other molecular data to determine evolutionary relationships
Concept 26.3: Shared characters are used to construct phylogenetic trees

• Once homologous characters have been identified, they can be used to infer a phylogeny
Cladistics

- **Cladistics** groups organisms by common descent
- A **clade** is a group of species that includes an ancestral species and all its descendants
- Clades can be nested in larger clades, but not all groupings of organisms qualify as clades
A valid clade is **monophyletic**, signifying that it consists of the ancestor species and all its descendants.
(a) Monophyletic group (clade)

(b) Paraphyletic group

(c) Polyphyletic group
Figure 26.10a

(a) Monophyletic group (clade)

Group I
• A **paraphyletic** grouping consists of an ancestral species and some, but not all, of the descendants
Figure 26.10b

(b) Paraphyletic group

Group II

A
B
C
D
E
F
G

© 2011 Pearson Education, Inc.
A polyphyletic grouping consists of various species with different ancestors.
(c) Polyphyletic group
Shared Ancestral and Shared Derived Characters

- In comparison with its ancestor, an organism has both shared and different characteristics
• A **shared ancestral character** is a character that originated in an ancestor of the taxon
• A **shared derived character** is an evolutionary novelty unique to a particular clade
• A character can be both ancestral and derived, depending on the context
Inferring Phylogenies Using Derived Characters

- When inferring evolutionary relationships, it is useful to know in which clade a shared derived character first appeared.
Figure 26.11

(a) Character table

<table>
<thead>
<tr>
<th>CHARACTERS</th>
<th>Lancelet (outgroup)</th>
<th>Lamprey</th>
<th>Bass</th>
<th>Frog</th>
<th>Turtle</th>
<th>Leopard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertebral column (backbone)</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hinged jaws</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Four walking legs</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Amnion</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hair</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

© 2011 Pearson Education, Inc.

(b) Phylogenetic tree

- Lancelet (outgroup)
- Lamprey
- Bass
- Frog
- Turtle
- Leopard
• An **outgroup** is a species or group of species that is closely related to the **ingroup**, the various species being studied.

• The outgroup is a group that has diverged before the ingroup.

• Systematists compare each ingroup species with the outgroup to differentiate between shared derived and shared ancestral characteristics.
• Characters shared by the outgroup and ingroup are ancestral characters that predate the divergence of both groups from a common ancestor
Phylogenetic Trees with Proportional Branch Lengths

• In some trees, the length of a branch can reflect the number of genetic changes that have taken place in a particular DNA sequence in that lineage.
Figure 26.12

- Drosophila
- Lancelet
- Zebrafish
- Frog
- Chicken
- Human
- Mouse
In other trees, branch length can represent chronological time, and branching points can be determined from the fossil record.
Millions of years ago:

- 542: PALEOZOIC
- 251: MESOZOIC
- 65.5: CENozoic
- Present

Taxa:
- Drosophila
- Lancelet
- Zebrafish
- Frog
- Chicken
- Human
- Mouse
Maximum Parsimony and Maximum Likelihood

- Systematists can never be sure of finding the best tree in a large data set
- They narrow possibilities by applying the principles of maximum parsimony and maximum likelihood
• Maximum parsimony assumes that the tree that requires the fewest evolutionary events (appearances of shared derived characters) is the most likely

• The principle of maximum likelihood states that, given certain rules about how DNA changes over time, a tree can be found that reflects the most likely sequence of evolutionary events
Figure 26.14

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
<th>Mushroom</th>
<th>Tulip</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>0</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td>Mushroom</td>
<td>0</td>
<td>40%</td>
<td></td>
</tr>
<tr>
<td>Tulip</td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

(a) Percentage differences between sequences

(b) Comparison of possible trees

Tree 1: More likely

Tree 2: Less likely

© 2011 Pearson Education, Inc.
• Computer programs are used to search for trees that are parsimonious and likely
Three phylogenetic hypotheses:

1  

2  

3  

4  

RESULTS

6 events 7 events 7 events
Phylogenetic Trees as Hypotheses

• The best hypotheses for phylogenetic trees fit the most data: morphological, molecular, and fossil

• Phylogenetic bracketing allows us to predict features of an ancestor from features of its descendants
  – For example, phylogenetic bracketing allows us to infer characteristics of dinosaurs
Figure 26.16

Common ancestor of crocodilians, dinosaurs, and birds

- Lizards and snakes
- Crocodilians
- Ornithischian dinosaurs
- Saurischian dinosaurs
- Birds

© 2011 Pearson Education, Inc.
• Birds and crocodiles share several features: four-chambered hearts, song, nest building, and brooding
• These characteristics likely evolved in a common ancestor and were shared by all of its descendants, including dinosaurs
• The fossil record supports nest building and brooding in dinosaurs
Figure 26.17

(a) Fossil remains of *Oviraptor* and eggs

(b) Artist’s reconstruction of the dinosaur’s posture based on the fossil findings
Concept 26.4: An organism’s evolutionary history is documented in its genome

- Comparing nucleic acids or other molecules to infer relatedness is a valuable approach for tracing organisms’ evolutionary history.
- DNA that codes for rRNA changes relatively slowly and is useful for investigating branching points hundreds of millions of years ago.
- mtDNA evolves rapidly and can be used to explore recent evolutionary events.
Gene Duplications and Gene Families

• Gene duplication increases the number of genes in the genome, providing more opportunities for evolutionary changes
• Repeated gene duplications result in gene families
• Like homologous genes, duplicated genes can be traced to a common ancestor
• **Orthologous genes** are found in a single copy in the genome and are homologous between species.
• They can diverge only after speciation occurs.
Formation of orthologous genes: a product of speciation

Ancestral gene

Ancestral species

Speciation with divergence of gene

Orthologous genes

Species A  Species B

Formation of paralogous genes: within a species

Ancestral gene

Species C

Gene duplication and divergence

Paralogous genes

Species C after many generations
Formation of orthologous genes: a product of speciation

Figure 26.18a

- Ancestral gene
- Ancestral species
- Speciation with divergence of gene
- Orthologous genes

Species A
Species B
• **Paralogous genes** result from gene duplication, so are found in more than one copy in the genome.
• They can diverge within the clade that carries them and often evolve new functions.
Formation of paralogous genes: within a species

Species C

Gene duplication and divergence

Paralogous genes

Species C after many generations
Genome Evolution

- Orthologous genes are widespread and extend across many widely varied species
  - For example, humans and mice diverged about 65 million years ago, and 99% of our genes are orthologous
Gene number and the complexity of an organism are not strongly linked.

For example, humans have only four times as many genes as yeast, a single-celled eukaryote.

Genes in complex organisms appear to be very versatile, and each gene can perform many functions.
Concept 26.5: Molecular clocks help track evolutionary time

- To extend molecular phylogenies beyond the fossil record, we must make an assumption about how change occurs over time
Molecular Clocks

• A molecular clock uses constant rates of evolution in some genes to estimate the absolute time of evolutionary change
• In orthologous genes, nucleotide substitutions are proportional to the time since they last shared a common ancestor
• In paralogous genes, nucleotide substitutions are proportional to the time since the genes became duplicated
• Molecular clocks are calibrated against branches whose dates are known from the fossil record
• Individual genes vary in how clocklike they are
Neutral Theory

• **Neutral theory** states that much evolutionary change in genes and proteins has no effect on fitness and is not influenced by natural selection
• It states that the rate of molecular change in these genes and proteins should be regular like a clock
Problems with Molecular Clocks

• The molecular clock does not run as smoothly as neutral theory predicts
• Irregularities result from natural selection in which some DNA changes are favored over others
• Estimates of evolutionary divergences older than the fossil record have a high degree of uncertainty
• The use of multiple genes may improve estimates
Applying a Molecular Clock: The Origin of HIV

- Phylogenetic analysis shows that HIV is descended from viruses that infect chimpanzees and other primates
- HIV spread to humans more than once
- Comparison of HIV samples shows that the virus evolved in a very clocklike way
- Application of a molecular clock to one strain of HIV suggests that that strain spread to humans during the 1930s
Figure 26.20

Index of base changes between HIV gene sequences

Year

HIV

Adjusted best-fit line (accounts for uncertain dates of HIV sequences)

Range
Concept 26.6: New information continues to revise our understanding of the tree of life

- Recently, we have gained insight into the very deepest branches of the tree of life through molecular systematics
From Two Kingdoms to Three Domains

- Early taxonomists classified all species as either plants or animals.
- Later, five kingdoms were recognized: Monera (prokaryotes), Protista, Plantae, Fungi, and Animalia.
- More recently, the three-domain system has been adopted: Bacteria, Archaea, and Eukarya.
- The three-domain system is supported by data from many sequenced genomes.
A Simple Tree of All Life

- The tree of life suggests that eukaryotes and archaea are more closely related to each other than to bacteria
- The tree of life is based largely on rRNA genes, as these have evolved slowly
• There have been substantial interchanges of genes between organisms in different domains
• **Horizontal gene transfer** is the movement of genes from one genome to another
• Horizontal gene transfer occurs by exchange of transposable elements and plasmids, viral infection, and fusion of organisms
• Horizontal gene transfer complicates efforts to build a tree of life
Billions of years ago

4 3 2 1 0

Bacteria
Eukarya
Archaea

© 2011 Pearson Education, Inc.
Is the Tree of Life Really a Ring?

- Some researchers suggest that eukaryotes arose as a fusion between a bacterium and archaean.
- If so, early evolutionary relationships might be better depicted by a ring of life instead of a tree of life.
Figure 26.23

Archaea

Eukarya

Bacteria