Phylogeny and the Tree of Life
Chapter 26

• Objectives
  • Explain the following characteristics of the Linnaean system of classification:
    a. binomial nomenclature
    b. hierarchical classification
  • List the major taxonomic categories from most to least inclusive.
  • Explain the justification for taxonomy based on a PhyloCode
  • Explain the importance of distinguishing between homology and analogy

• Distinguish between the following terms: monophyletic, paraphyletic, and polyphyletic groups; shared ancestral and shared derived characters; orthologous and paralogous genes
• Define horizontal gene transfer and explain how it complicates phylogenetic trees
• Explain molecular clocks and discuss their limitations
Investigating the Tree of Life

• This chapter describes how biologists trace phylogeny - the evolutionary history of a species or group of related species
• Biologists draw on the fossil record which provides information about ancient organisms
• Biologists also use systematics as an analytical approach to understanding the diversity and relationships of organisms, both present-day and extinct

Phylogenies Show Evolutionary Relationships

• Phylogenetic systematics connects classification with evolutionary history
• Taxonomy is the ordered division of organisms into categories based on a set of characteristics used to assess similarities and differences

Binomial Nomenclature

• Binomial nomenclature is the two-part format of the scientific name of an organism and was developed by Carolus Linnaeus
  – the binomial name of an organism or scientific epithet is latinized and includes the genus and species
Hierarchical Classification

- Linnaeus also 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

Linking Classification and Phylogeny

- Systematists depict evolutionary relationships in branching phylogenetic trees
  - each branch point represents the divergence of two species
  - “deeper” branch points represent progressively greater amounts of divergence
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 polytomy is a branch from which more than two groups emerge.
What We Can and Cannot Learn from Phylogenetic Trees

- Phylogenetic trees do show patterns of descent
- Phylogenetic trees do not indicate when species evolved or how much genetic change occurred in a lineage
- It shouldn’t 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
Tracing Phylogenies

- Phylogenies are based on common ancestries inferred from fossil, morphological, and molecular evidence.

Morphological and Molecular Homologies

- In addition to fossil organisms, phylogenetic history can be inferred from certain morphological and molecular similarities among living organisms.
- In general, organisms that share very similar morphologies or similar DNA sequences are likely to be more closely related than organisms with vastly different structures or sequences.
Sorting Homology from Analogy

- A potential misconception in constructing a phylogeny is similarity due to convergent evolution, called analogy, rather than shared ancestry
  - 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.
Cladistics

• Cladistics groups organisms by common descent
  – Clades can be nested within larger clades, but not all groupings or organisms qualify as clades
    • A valid clade is monophyletic signifying that it consists of the ancestor species and all its descendants
    • A paraphyletic clade is a grouping that consists of an ancestral species and some, but not all, of the descendants
    • A polyphyletic grouping includes numerous types of organisms that lack a common ancestor

• Polyphyletic groups are distinguished from paraphyletic groups by the fact that they do not include the most recent common ancestor
  – Biologists avoid defining polyphyletic groups and instead reclassify organisms if evidence suggests they are polyphyletic
Shared Primitive and Shared Derived Characteristics

- In cladistic analysis clades are defined by their evolutionary novelties
  - A shared primitive character is a homologous structure that predates the branching of a particular clade from other members of that clade
    - It is shared beyond the taxon we are trying to define
  - A shared derived character is an evolutionary novelty unique to a particular clade

Outgroups

- Systematists use a method called outgroup comparison to differentiate between shared derived and shared primitive characteristics
  - As a basis of comparison we need to designate an outgroup which is a species or group of species that is closely related to the ingroup, the various species we are studying
• Outgroup comparison is based on the assumption that homologies present in both the outgroup and ingroup must be primitive characters that predate the divergence of both groups from a common ancestor
  – The outgroup comparison enables us to focus on just those characters that were derived at the various branch points in the evolution of a clade

Phylogenetic Trees and Timing

• Any chronology represented by the branching pattern of a phylogenetic tree is relative rather than absolute in terms of representing the timing of divergences
  – Phylograms provide information about sequences of events
  – Ultrametric trees present information about the timing of events
Phylograms

- In a phylogram the length of a branch in a cladogram reflects the number of genetic changes that have taken place in a particular DNA or RNA sequence in that lineage.

Ultrametric Trees

- In an ultrametric tree the branching pattern is the same as in a phylogram, but all the branches that can be traced from the common ancestor to the present are of equal length.
Systematists can never be sure of finding the single best tree in a large data set. The possibilities can be narrowed by applying the principles of maximum parsimony and maximum likelihood.

Among phylogenetic hypotheses, the most parsimionous tree is the one that requires the fewest evolutionary events to have occurred in the form of shared derived characters.
• 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.
Phylogenetic Trees as Hypotheses

- The best hypotheses for phylogenetic trees are those that fit the most data: morphological, molecular, and fossil
  - Sometimes there is compelling evidence that the best hypothesis is not the most parsimonious

Phylogenetic bracketing allows us to predict features of an ancestor from features of its descendents
  - This has been applied to infer features of dinosaurs from their descendents: birds and crocodiles
Phylogeny and the Genome

- Much of an organism’s evolutionary history is documented in its genome
  - Comparing nucleic acids or other molecules to infer relatedness is a valuable tool 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 is one of the most important types of mutation in evolution because it increases the number of genes in the genome, providing further opportunities for evolutionary changes
  - Orthologous genes are genes found in a single copy in the genome and can diverge only once speciation has taken place
  - Paralogous genes result from gene duplication, so they are found in more than one copy in the genome and can diverge within the clade that carries them, often adding new functions
Genome Evolution

• Orthologous genes are widespread and extend across many widely varied species
  – The widespread consistency in total gene number in organisms of varying complexity indicates that genes in complex organisms are extremely versatile and that each gene can perform many functions

Molecular Clocks

• Molecular clocks help track evolutionary time
  – The molecular clock is a yardstick for measuring the absolute time of evolutionary change based on the observation that some genes and other regions of genomes appear to evolve at constant rates
    • 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
Neutral Theory

- Neutral theory states that much evolutionary change in genes and proteins has no effect on fitness and therefore is not influenced by Darwinian selection
  - the rate of molecular change in these genes and proteins should be regular like a clock

Difficulties 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
The Origin of HIV: Applying a Molecular Clock

- Phylogenetic analysis shows that HIV is descended from viruses that infect chimpanzees and other primates
  - A comparison of HIV samples from throughout the epidemic has shown that the virus has evolved in a remarkably clocklike fashion

The Universal 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
Horizontal Gene Transfer

- 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 complicates efforts to build a tree of life