

Genetics and Evolution IB 201 06- Lecture 2

Reconstructing Phylogenetic Trees

The next two lectures consider phylogenetics and the methods of reconstructing the history of relationships among organisms. This knowledge sets the groundwork for all the other subdisciplines in evolutionary biology, including evolutionary genetics. Phylogenetic trees are constructed from currently living (extant) organisms, and we use species as the basic units in this reconstruction. We will discuss later exactly what a species is, most of you already have an idea.

Species are not independent entities—they are connected by **descent from a common ancestor**. They are connected by a lineage, a genealogy.

A genealogy or lineage of species is known as a **phylogeny**, and the scientific discipline concerned with reconstructing the phylogenies of organisms is called **phylogenetics** (a part of the overall discipline called Systematics).

I. Terms

- *Phylogenetic inference*: seeks to recover the historical (genealogical) relationships among organisms
- *Phylogenetics*: study of historical/genealogical relationships among organisms (taxa)
- *Phylogeny*: graphical representation of historical/genealogical relationships; a branching diagram
- *Tree topology*: (phylogenetic tree, cladogram): a branching diagram that represents the genealogical relationships among taxa
- *Taxon* (pl. taxa): a group of organisms with an assigned name; can be a species name (*Bombus impatiens*), family name (e.g., Apidae), order name (e.g., Hymenoptera).

A tree is represented by

1. *Tips*: the species used in constructing the tree (currently extant, or extinct)
2. *Branches*: ancestral populations of a species over time
3. *Nodes*: the point where a species splits into two or more descendant populations

II. Phylogenetic Principles

1.

2.

3.

4.

III. More Terms

- *Homology*: correspondence (morphological, molecular, etc.) inherited through common ancestry
- *Apomorphy*: derived character
- *Synapomorphy*: shared, derived character—informative for phylogenetic estimation
- *Autapomorphy*: uniquely derived character—non-informative for phylogeny
- *Symplesiomorphy*: shared ancestral character—non-informative for phylogeny
- Western World view before 1859

IV. How to Build a Phylogenetic Tree: Simple Example (5 characters, 4 taxa)

A. Consider *homologous* similarity of structures vs. *convergent* similarity

B. How to combine information from different characters to infer an overall phylogeny?

- Construction of data matrix

- Establish *ancestral* versus *derived* characters
 - **Outgroup comparison**

VI. What to do if characters conflict?

A. Optimality criterion—Maximum Parsimony

VII. Additional Aspects of Phylogeny Reconstruction

Parsimony (developed in part by J. S. Farris in the 1970s—developed the computer algorithms)

- seeks to minimize the total amount of evolutionary change that has occurred
- assumes that homoplasy is rare
- seeks to find the tree that requires the fewest number of character state changes.
- the tree of choice is the shortest tree, often called the *maximum parsimony* tree or the *shortest* tree.

The most parsimonious tree will minimize the amount of homoplasy inferred for the data, which may be the best estimate of relationships among the taxa of interest

However, parsimony may not always result in the best estimate of relationships, particularly if there have been many convergences and reversals in character states for many of the characters.

Other tree reconstruction methods have been developed, including:

- *Distance Methods* (often called phenetic methods)
- **Maximum Likelihood Methods**

VIII. Additional Terms

- *Character*: any phenotypic trait of an organism (tongue length, head shape, nucleotide position #50, degree of social behavior, and so on).
- *Character state*: variant (presumed to be homologous) forms of a character; coded as 0s, 1s, 2s, etc., if morphological trait, coded as G, A, T, or C, if characters are DNA sequences (also can code as different amino acids), and so on.
- *Homoplasy*:
 - *Convergence*: correspondence or similarity between characters which is not due to common ancestry; could be due to sharing similar environment. Common in nature and often a source of error in phylogenetic estimation
 - *Parallelism*: development of the same derived state in separate lineages from the same or a similar ancestral state—similar concept as convergence.
 -
- *Character*: any phenotypic trait of an organism (tongue length, head shape, nucleotide position #50, degree of social behavior, and so on).

Ref: Read your text to learn more about these methods (F&H pp 549-558; 567- top 571; 437-450).