

SEGMENT THREE LECTURE SIX: INFERRING EVOLUTIONARY HISTORIES

Phylogenetic methods are used to determine evolutionary trees from almost any kind of biological information. Two broad types of biological information:

- 1) **Structural/Morphological** characters can be used to infer a tree. These characters are often coded as binary data (presence-absence; Fig. 26.11a/b).
- 2) **DNA** characters are generally more abundant (including entire genomes in some studies) than morphological characters and are often used in contemporary phylogenetics. Each homologous nucleotide site is a character, for which there are four possible **character states** (A, C, G, or T).

Phylogenetic methods

- 1) **Distance methods** convert the character state data into distances between all pairs of species. Species with the least distance separating them are clustered together. Distance methods are readily programmed and can be executed quickly. A distance tree can be a useful first approximation to a phylogeny.
- 2) **Maximum Likelihood** models the evolutionary process. Incorporated into the model are parameters such as mutation rates. The method maximizes the likelihood statistic calculated under a particular evolutionary model so that the optimal tree is the tree with the highest likelihood. These methods are useful for phylogenetic analysis of DNA sequences.
- 3) **Maximum parsimony** minimizes the number of changes (mutations) on a tree, so that the optimal tree is the shortest tree (Fig. 26.15). Ancestral states are inferred from **outgroup taxa**, which are organisms known to have diverged before the species being studied or the “ingroup” (e.g., the outgroup of vertebrates is the lancelet in Fig 26.11b).

Shared Ancestral Characters are those features found in a group and its (even distant) ancestors (e.g. backbones in mammals).

Shared Derived Characters are unique features found in a group, but not in its ancestors (e.g. placentas in some mammals). These characters define branch points in parsimony trees.

Phylogenetic Trees are hypotheses of relationships and can be further tested. For example, for a given tree, a biologist can predict that features shared by closely related organisms were also found in their common ancestor, an arrangement called **phylogenetic bracketing**. The best hypothesis fits all of the available data.