

PHYLOGENETIC CLASSIFICATION

A classification is an orderly arrangement of organisms (as taxa) or objects in a hierarchical series. The kind of relationships revealed in a classification depends on the criteria or characters used in its construction. If characters are quantified and statistically assessed on the basis of overall similarity, the relationships are called phenetic. If the characters are analyzed in a way that detects their evolutionary or genealogical development, the relationships are called phylogenetic. Although most classifications are, to a greater or lesser extent, artificial, systematists try to produce classifications that are natural. Artificial classifications are structured for convenience, using readily identifiable characters that are not necessarily indicative of phylogenetic relationships. Natural classifications are based on hypothetical phylogenetic relationships that attempt to document that all members of a particular group share a common ancestor.

In modern systematic study, researchers aim to construct classifications that are phylogenetic. To do this, many characters need to be observed, recorded, and analyzed for all the organisms or taxa included in an investigation. Factors to consider in the phylogeny of organisms include common ancestry, evolutionary divergence, and inequality of evolutionary rates. The difficulties in assessing phylogenies usually result because of parallel, convergent, or reticulate evolution. **Parallel** evolution refers to the independent acquisition of similar character states in distinct organisms or taxa that are derived from a common ancestor. For example, in the genus *Sedum*, white flowers appear in several different species that do not share the most immediate recent ancestor; in this case, white flowers have developed in parallel among several species groups. **Convergent** evolution refers to the independent acquisition of similar character states in organisms or taxa that are derived from different ancestors. For example, stem succulence has arisen in several unrelated plant families, such as Cactaceae and Euphorbiaceae. Parallelisms and convergences make taxa appear to be more closely allied than they really are. **Reticulate** evolution involves hybridization followed by divergence of the hybrid forms from the parents; anastomosing networks result that complicate the determination of phylogenies.

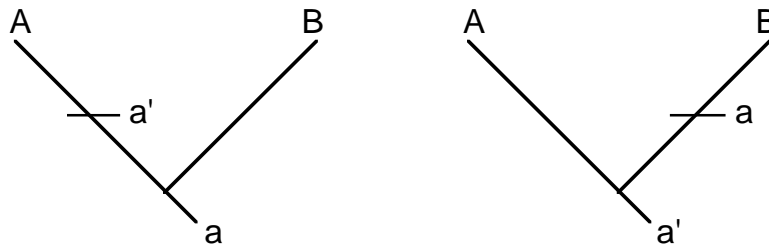
Phylogenetic classifications may be monophyletic, polyphyletic, or paraphyletic. **Monophyletic** refers to one ancestry, and exists when all the included organisms in a group share a common ancestor and none of the descendants of that ancestor are excluded. When the included organisms are derived from two or more distinct ancestors, the classification is **polyphyletic**. When a classification does not include all the descendants of the most recent common ancestor, it is termed **paraphyletic**. Phylogenetic classifications should be based on monophyletic groups, whenever possible.

Before we can construct a phylogeny of organisms or taxa, we first must assess the extent of variation of single characters. It also is important to determine if a particular character is phenotypically plastic (variable with respect to different environmental conditions). To be useful in systematics, characters need to show some variation, such

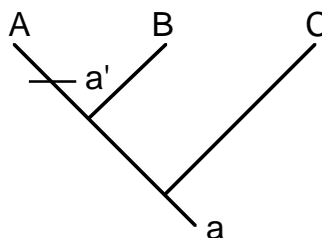
as red, pink, or white petals. In this case, petal color is the character and red, pink, and white represent states of the character. To use petal color in a phylogenetic analysis, we need to determine character polarity, which identifies the ancestral or plesiomorphic state from the derived or apomorphic state.

The determination of **polarity** for a character is not easy, and often requires the use of information from several sources. Developmental or ontogenetic studies and the fossil record may provide clues, if not substantial data, on the evolutionary history of organs or tissues, but for the majority of characters, such information will not be available and another means of analysis must be used. **Ingroup** analysis uses only the characters of the immediate group under study and assumes that the frequency of a given character state is correlated with its polarity (ancestral vs. derived state). The objection to the ingroup method is that commonness of a character state does not necessarily reflect its polarity; a common character may be recently derived or ancestral. Ingroup analysis may be useful in determining similarities within a group, but is generally unsatisfactory for phylogenetic classifications because of its failure to account for character polarity. The principal method used for assessing polarity is **outgroup** analysis.

Outgroup analysis typically involves the group under study (ingroup) and one or more groups that are more distantly related. The character states for each taxon needs to be documented before the analysis can be made. For a given character, the states for each of the taxa are determined. If the same character state is found in the outgroup and some members of the ingroup, then that character state is ancestral (plesiomorphic) and the other state is derived (apomorphic). In outgroup analysis, phylogenetic relationships are determined by the shared derived characters (synapomorphies). In the following example, there are two taxa (A and B) and two character states (a and a') of a single character. Without any further information it is impossible to know which character state is derived and which is ancestral.



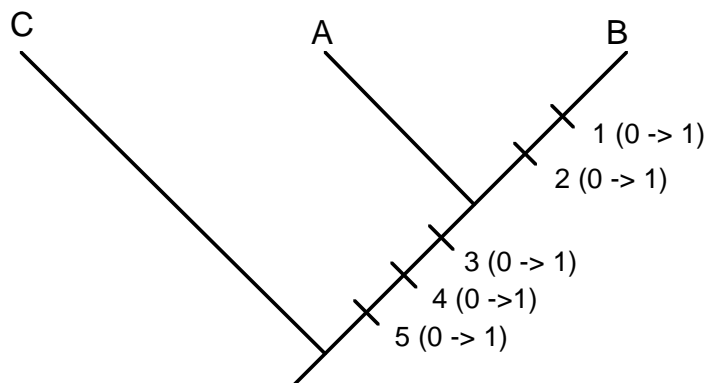
A third taxon (C), sister to A and B, is needed to determine character polarity. If a is the character state in C, then a' is derived.



To demonstrate the phylogenetic relationship of A to B, an outgroup needs to be selected that shares a more ancient common ancestor with these two groups. In this example, C will serve as the outgroup. The distribution of character states in the ingroup. First, a list of characters that are used to distinguish the families is generated, and a matrix of the character state distribution in all three families is prepared. "0" is used to indicate the plesiomorphic or ancestral condition and "1" is used for the apomorphic or derived character state.

	C	A	B
1. leaf arrangement 0:3-ranked 1:2-ranked	0	0	1
2. leaf sheath 0:closed 1:open	0	0	1
3. inflorescence 0:panicle, umbel 1:spikelet	0	1	1
4. perianth 0:tepals 1:absent or modified	0	1	1
5. fruit 0:several seeds/fr 1:1 seed/fr	0	1	1

Then, a phylogenetic tree can be constructed, based on the occurrence of apomorphic or derived character states. The simplest tree, which is the one with the fewest steps in construction, is considered to be the most parsimonious or the one most likely to reflect the evolution of the group.



In this tree, characters 1 and 2 may help differentiate species 'B' from species 'A', but do not provide evidence of relationships among any species. However, characters 3, 4, and 5 all indicate a common ancestry for 'A' and 'B' to the exclusion of 'C.'

Cladistics is a method commonly used to determine phylogenies or ancestor-descendant relationships in monophyletic groupings of taxa using outgroup comparisons. The term is derived from clade, which refers to a monophyletic group. A branching pattern (cladogram), which depicts the divergence of the taxa based on the distribution of shared derived character states, may be constructed.

To carry out the cladistic method, the following need to be accomplished: establish monophyly for the study group, document the character variation for the study group, recognize the most closely related groups (sister groups or outgroups) and document the characters in those groups, and determine character polarity by comparison of outgroups with the study group. After the character polarities can be determined, branching patterns of a cladogram can be constructed. If several outgroups are used, and several cladograms are constructed, the parts of the cladograms that are similar or congruent are considered to be the most reliable. In modern computer programs designed to facilitate analysis of complex data sets, a branching network, or unrooted tree, is first constructed based on the data from the ingroup and outgroup simultaneously. Then the tree is 'rooted' along the branch connecting the outgroup(s) to the rest of the tree. Because our knowledge of characters is rarely, if ever, complete, we need to remember that all classifications remain hypotheses of relationships.