

Phylogeny & Tree Terminology Part I

Nodes – Can be external or internal. These are the branching points on the tree.

Branches – aka edges, pathway or connection between nodes.

External Nodes – aka leaves or OTUs, represent sequences or data that exists for extant organisms.

Internal Nodes – these represent hypothetical or inferred ancestors.

Taxon – type of organism (ranked in many ways, species, class, subdivision).
“Taxa” is plural for taxon.

OTUs – Operational Taxonomic Units. Recycled ecology term based on groups defined in a reproducible and taxonomic sense. For example, used once restriction data has been obtained and similarities have been compared.

Phylotypes – Grouping or classification designation that also contains evolutionary history or phylogenetic information. For example, used once sequence data has been obtained and correctly compared.

Phylogeny – The true evolutionary tree we are trying to estimate.

Ingroup – The group under study (should be monophyletic).

Outgroup – a group used in analysis to establish polarity of characters. Usually a sister group, but may be further out, more than one may be used.

Tree – graphical representation/mathematical structure used to model or map actual evolutionary history. Comprised of nodes and branches. Must be rooted to show phylogeny.

Root – Ancestor for the entire tree, when known. Shows direction of change and therefore a phylogeny is established.

Phylogeny & Tree Terminology Part II

Homology – Resemblance due to inheritance from common ancestry, i.e., truly linked traits or characteristics passed down from ancestor to offspring.

Orthologous – Homology that arises by divergence and/or speciation.

Paralogous – Homology that arises from gene duplication.

Homoplasy – Resemblance due to reversals, parallelism and/or convergence and NOT due to common ancestry, i.e., nature trying to trick you yet again.

Cladogram – Dendrogram of branching diagram constructed by cladistic principles and methods (need not be rooted). Neither axis means anything.

Additive tree – Depicts the amount of evolutionary change to have occurred along the different branches. One axis is used to measure evolutionary change.

Ultrametric tree – Depicts the actual times of divergence under a constant clock assumption. One axis is used to measure time.

Consensus tree – trees that represent the commonality (if any) among a set of trees generated by a phylogenetic analysis, e.g., Strict, Adam's, Majority rule, etc. Often used in conjunction with bootstrapping.

Bootstrapping – Method of determining the robustness of a dataset by resampling with replacement. Evaluates whether the data supports the tree NOT if the tree supports the data.

Polytomy – represents either simultaneous divergence of multiple sequences and thereby taxa (i.e., hard polytomy) or more likely the inability to resolve certain branches (bifurcations) due to insufficient or conflicting data (i.e., soft polytomy). aka “star phylogeny”

Evolutionary Distance – the number of steps or evolutionary changes required to explain the data given a particular tree topology. Obtained by summing the number of character state changes along each branch and internode of the tree.

Phylogeny & Tree Terminology Part III

Clade, Lineage, or Cluster – A monophyletic group.

Cladistics Terms:

Apomorphy – Derived character state.

Plesiomorphy – Ancestral character state.

Autapomorphy – Derived character state found only in one clade.

Synapomorphy – Shared-derived character state, the heart of cladistic theory. Make sure you know this one and why it is the ultimate in cladistic character types!

Monophyletic – A group of organisms that have a common evolutionary ancestry (history). Members share a set of relationships through common ancestry that are not shared with any other taxa outside the group.

Paraphyletic – Groups based on plesiomorphies or shared ancestral characteristics. Exclude taxa with autapomorphies, e.g., reptiles.

Polyphyletic – Groups based erroneously on convergent characters, e.g., vultures or eyes.

Systematic error – Making the wrong choices with respect to analyzing data using improper algorithms and/or using incorrect assumptions about the data.

Stochastic or Random Error – Errors associated with the fact that the sequences in question no longer contain the full level of phylogenetic information that might have once been available. This can only be avoided with an infinite amount of data.