

"PRINCIPLES OF PHYLOGENETICS: ECOLOGY AND EVOLUTION"

Integrative Biology 200B
University of California, Berkeley

Spring 2007
B.D. Mishler

April 3, 2007. **Comparing Cladograms**

-- There are many reasons why one would want to compare cladograms, falling into two basic categories:

-- *Within an analysis of one clade*; e.g., equally or nearly equally parsimonious (or likely) trees, trees resulting from different character partitions, OTU selections, models of evolution, or methods of analysis, and comparisons with trees from the literature.

-- *Comparing analyses of different clades*, e.g., gene family evolution, migration between populations, vicariance biogeography, host/ parasite relationships, symbiosis, community evolution, or any long-term ecological association

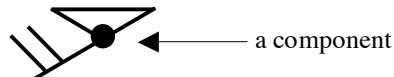
-- See following page for a simple intuitive example to motivate this discussion.

-- Methodology for comparing cladograms:

(1) consensus techniques (strict, semi-strict, majority rule, Adams) -- finding shared signal among trees (examples to be gone over in lab today)

(2) tree-to-tree distance metrics (e.g., NNI interchange metric, partition metrics, agreement subtrees; examples to be gone over in lab today)

(3) component analysis (more next week in biogeography) -- finding individual statements of relationship that are shared among trees, basically a node relating some taxa to the exclusion of others.



(4) parsimony approaches -- representing the grouping information in separate trees as characters in a matrix (Brooks' approach -- to be gone over below)

(5) maximum likelihood approaches (parametric bootstrapping) -- comparing alternative trees or alternative models of evolution for your data.

(6) supertrees -- combining different trees, e.g., using Brooks parsimony (called "matrix representation" in the literature).

Brooks Parsimony

(see Brooks & McLennan, 1991; Brooks 1981, Syst. Zool. 30:229; Wiley 1988, Syst. Zool. 37:271; and see Kluge 1988, Syst. Zool. 37:315 for some suggested modifications)

Steps:

1. Cladogram of parasite group
 2. Cladogram of host group
 3. Cladogram of parasite group is taken as a completely polarized, multistate transformation series -- recoded by additive binary coding
 4. make new data matrix with hosts as OTU's and parasite clades as characters
 5. construct new host cladogram from this matrix
 6. compare this new host cladogram (derived from cladogram of the parasite group) with the original host cladogram (derived from host characters) and with host cladograms based on other parasite groups, if possible. Congruence is taken as evidence of common cause (shared history); incongruence (homoplasy) is taken as due to separate causes (e.g., host-switching or extinction)
- Problem resolution:
1. when more than one parasite occurs in given host, codes are combined (e.g., host E in example combines codes of 10 & 5)
 2. when all members of parasite clade are missing from a host taxon, host coded with "?"
 3. when one parasite occurs in more than one host, codes are combined (e.g., if parasite species 10 occurred in hosts A,B,& E) -- note that this is controversial -- some have suggested downweighting such species, even eliminating them from the analysis entirely.