# Molecular Phylogenetic Questions

Please note that most of the following questions, but not all, have a clear-cut answer. Others are intended just to stimulate thinking and discussion.

**1)** Which of the following methods

Maximum Parsimony (MP) • Distance Methods • Maximum Likelihood (ML) • Bayesian Analysis

could be used to infer trees from

       a) morphological data;

       b) molecular data;

       c) molecular data under a molecular clock?

**2)** Suppose you have a fully resolved (binary) tree with N terminal taxa (leaves). How many

       a) branches;

       b) internal branches;

       c) internal nodes

does the tree contain? (Suggestion: Draw a simple tree, count, and look what happens if you successively add taxa.)

**3)** Draw the fundamental tree of all organisms (three OTUs). How must this tree be rooted? Why? Who is LUCA?

**4)** Suppose you computed a pair-wise distance matrix for N taxa. How many entries (do not count double entries and diagonal zeros) does the matrix have?

**5)** Order the following methods according to increasing computation time:

MP, heuristic search • Neighbor Joining • ML, heuristic search • UPGMA • Minimum Evolution, heuristic search • Least Squares, heuristic search

assuming identical heuristic search options, respectively. Try it with PAUP* if you are not sure!

**6)** In which cases does UPGMA infer a rooted tree? Why?

**7)** Suppose you used MP as well as ML (after finding an appropriate substitution model with Modeltest) for analysing your data by heuristic search and bootstrapping. What is your interpretation if a certain Clade A

       a) occurs in both MP and MP trees, but receives no bootstrap support, respectively;

       b) is well supported in both MP and MP bootstrap analyses;

       c) is well supported in MP(ML) bootstrap analysis, but unsupported in the ML(MP) bootstrap analysis, although it occurs in the other original ML(MP) tree;

d) is well supported in MP(ML) bootstrap analysis, but contradicting bootstrap support is revealed with ML(MP) bootstrap analysis;

e) does never occur and receives conflicting support in both MP and ML analyses?

**8)** In which cases should you add

(a) more taxa;

(b) more characters

to your dataset?

**9)** Are uncorrected ("p") distances usually

(a) higher;

(b) as high as;

(c) smaller

than distances corrected under a model of site substitution?

**10)** Which problems can occur if a substitution model has

(a) not enough;

(b) too many

parameters?

**11)** Suppose you are analyzing the following five taxa as well as the three states A, B, and C of a single morphological character which are distributed as follows:

| taxon1 | A |
|--------|---|
| taxon2 | A |
| taxon3 | B |
| taxon4 | B |
| taxon5 | C |

Current taxonomy arranges the taxa like this:

12idae: taxon1, taxon2

34idae: taxon3, taxon4

5idae: taxon5

Assume also that the molecular tree your infer is ((((taxon1,taxon2),taxon3),taxon4),taxon5); and has to be rooted with 5idae. Are your molecular results

(a) in accordance with taxonomy;

(b) in accordance with morphology?