

Assignment IV – 28 Oct 2004

You'll need to use PAUP\* for the following problems. PAUP is installed on iNquiry (<http://inquiry.egg.isu.edu>). Data required for the following problem may be obtained from the course website but you may also use your own data. The problem requires the generation of some data – you will also provide a brief interpretation (1-2 paragraphs) of the final results. **The assignment is due on Thursday, November 4.**



### ***The Strepsiptera Problem, pt 2.***

In Assignment III, you found that the 18S gene in Strepsiptera has experienced many more substitutions than the same molecule in most other insects. It seems that the flies have also experienced an exceptional number of substitutions in the 18S genes. Because of these long branches, the Strepsiptera and the flies tend to form a clade when a phylogeny is inferred using NJ with a biased distanced estimate or when using MP. When an unbiased estimate is used with NJ, the Strepsiptera tend to cluster with the beetles. Traditionally, we consider the Strepsiptera more related to the beetles than to other insects, although these relationships are still unresolved. Can using ML help clear this up?

1. Given the 18S genes of insects, consisting of 13 taxa and 770bp, use ModelTest to identify the AIC unbiased estimate of substitution rates for the given data set. Set likelihood parameters (Lset) to the AIC recommended settings.
2. Construct a NJ tree with JC. Calculate the  $-\ln L$  score for this tree using the AIC likelihood settings.
3. Construct a NJ tree using your AIC unbiased estimator. Calculate the  $-\ln L$  score for this tree using the AIC likelihood settings.
4. Find the ML tree by a heuristic search using a model approximating JC (equal base frequencies, one substitution rate, equal rates among sites). Calculate the  $-\ln L$  score for this tree using the AIC likelihood settings.
5. Find the ML tree by a heuristic search using the model suggested by AIC in ModelTest. Calculate the  $-\ln L$  score for this tree using the AIC likelihood settings.
6. Include in your report:
  - a. NJ (JC) tree,  $-\ln L$ :
  - b. NJ (AIC model) tree,  $-\ln L$ :
  - c. ML (JC) tree,  $-\ln L$ :
  - d. ML (AIC model) tree,  $-\ln L$ :
  - e. Print the ML (JC) tree; include branch lengths.
  - f. Print the ML (AIC model) tree; include branch lengths.
  - g. Discussion (1-2 paragraphs); these questions might help guide your comments: Is ML immune to the “long-branch attraction” problem that causes difficulties with MP? What placement of the Strepsiptera is supported by the highest likelihood value? How is the topology of the ML tree different than the topology of the NJ tree constructed using the same model? Given the NJ, MP, and ML trees you've constructed for the insects, if you were to put a single tree in a publication, which would you choose and how would you defend your decision?