

EXERCISE 3:

TEST HYPOTHESES ABOUT MOLECULAR EVOLUTION OF *LDH*

Objectives: Use three LRTs to evaluate the following hypotheses: (1) the mutation rate of *Ldh-C* has increased relative to *Ldh-A*, (2) a burst of positive selection for functional divergence occurred following the duplication event that gave rise to *Ldh-C*, and (3) there was a long term shift in selective constraints following the duplication event that gave rise to *Ldh-C*.

Step-by-step guide:

1. Obtain the files for Exercise 3 from the course web-site. The tree files represent different hypotheses denoted H0, H1, H2 & H3. These hypotheses represent the following concepts:
 - a. H0: homogeneous selection pressure over the tree
 - b. H1: episodic change in selection pressure in *Ldh-C* (only in the branch that immediately follows the gene duplication event).
 - c. H2: Long term shift in selection pressure in *Ldh-C* only; *Ldh-C* has a permanent change in selection pressure (as compared to its ancestors) whereas *Ldh-A* remains subject to the ancestral level of selection pressure.
 - d. H3: Long term shift in selection on both *Ldh-C* and *Ldh-A*; those lineages are subject to selection pressures different from each other and from the ancestor.
2. Run CODEML using the settings in the control file for Exercise 3. Familiarize yourself with the results. In addition to the likelihood score you must be able to identify the branch-specific estimates of the omega parameter. (In the first run, the branch specific values for omega will all be the same. In later runs there will be differences among some branches).
3. As in the previous exercises, you will need to change the control files and re-run CODEML. The objective is to compute the likelihood, and estimate omega parameters, under different models of how selection pressure changes in different parts of the tree. Because the relevant model information is contained in the tree file, you will need several tree files (obtained from the course web site) and change the control file so that it reads the different tree files.
 - a. As always, you should change the name of the main result file (via `"outfile ="` in the control file) or you will overwrite your previous results.
 - b. Change the model assumptions about branch specific omega values by changing the tree files (via `"treefile ="` and `"model ="`) set within the control file.
4. Repeat step 3 for each of the four tree files that have been provided to you. Again, keep track of your results by using a table like "Table E3" shown in the slides. In addition, carry out likelihood ratio tests (LRT) of the hypotheses below. See the lecture notes for additional details about the LRT. Use 1 degree of freedom to obtain the *P*-value for each LRT.
 - a. H0 vs. H1
 - b. H0 vs. H2
 - c. H2 vs. H3