EXERCISE 2:
SENSITIVITY OF d_N/d_S RATIO TO ASSUMPTIONS

Objectives: You will investigate the sensitivity of your estimate of ω to the transition/transversion ratio (κ), and to the assumed model for codon frequencies (π_i's). After you collect the required data you will determine which assumptions yield the largest and smallest values of S, and what is the effect on ω.

Step-by-step guide:

1. Find the files for Exercise 2 on the workshop web-site (codeml.ctl.ex2, seqfile.txt) and familiarize yourself with them. It would be best to create a new directory for exercise 2.

2. Run CODEML using the settings in the control file for exercise 2. Familiarize yourself with the results. In addition to the likelihood score you must be able to identify the part of the result file that provides estimates of the following:

    a. Number of synonymous or nonsynonymous sites (S and N)

    b. Synonymous and nonsynonymous rates (d_S and d_N)

3. As in exercise 1, you will need to change the control files and re-run CODEML. The objective is to compute the likelihood of the example dataset under different model assumptions. To do this you must:

    a. 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 codon frequencies (via "`CodonFreq = `") and kappa (via "`kappa =`" and "`fix_kappa =`").

4. Repeat step 3 for each set of assumptions about codon frequencies and kappa given as comments at the bottom of the example control file.

5. In your favorite spreadsheet program create a table like "Table E2" in the slides and fill it in with your results.

6. Use your table to determine which assumptions yield the largest and smallest values of S. What is the effect on omega?