EEB 135/235: POPULATION GENETICS - HOMEWORK 7

Note: For full credit you must show your work.

(1) (10 points) Describe the differences between the neutral theory of evolution and the nearly neutral theory of evolution.

(2) (10 points) (a) Explain why under the neutral model, the rate of substitution is independent of the population size. (b) If the rate of substitutions is 10^{-5} per site per generation, what is the expected time between substitutions?

(3) (20 points, From Nielsen & Slatkin) Suppose that a mutation rate is 2×10^{-9} per site per year.

　(a) What is the rate of substitution of deleterious mutations per million years if the selection coefficient against them is 0.001 in a population with 10,000, 1,000, or 100 individuals?

　(b) What fraction of the neutral rate (2×10^{-9} per site per year) are the substitution rates that you computed from (a)?

(4) (15 points) Because of redundancy in the genetic code, some changes to sequence do not lead to changes in amino acids. Such changes are called "synonymous" changes, and others that do change the amino acid are called "non-synonymous." If one aligns two genes between species, one can look to see whether differences are synonymous or nonsynonymous. Let's assume that synonymous changes are completely neutral, while nonsynonymous changes are not.

　(a) Suppose that at all synonymous sites, the observed pattern of differences leads to an estimated rate of substitution of 2×10^{-8} per site per generation. What would you estimate the total mutation rate per site to be?

　(b) Suppose that at non-synonymous sites the estimate of the rate of substitution is 5×10^{-9} per site per generation. Using your answer from (a), what is your estimate of the fraction of neutral mutations at non-synonymous sites?

　(c) Suppose instead that at non-synonymous sites the estimate of the rate of substitution is 5×10^{-8} per site per generation. What might explain the non-synonymous to synonymous rate ratio observed?

(5) (15 points) When comparing the amino acid sequence of the α-hemoglobin molecule between dogs and kangaroos, you find that 13.5% of the sites have different amino acids. Assume that the α-hemoglobin molecule has been undergoing amino acid replacement at a rate of 10^{-9} amino acid replacements per amino acid site per year. Estimate the divergence time between kangaroo and dog. Is this estimate the split time between the populations or the coalescent time of the lineages? Please explain.

(6) (25 pts) Fay et al. (2002) surveyed polymorphism at 45 genes in *D. melanogaster* and in *D. simulans*. The segregating sites are classified as either synonymous or nonsynonymous. The sequences of these genes in the two Drosophila species were compared to assess the number of substitutions.

The results are summarized in the following McDonald-Kreitman table. Assume that the number of synonymous sites of these 45 genes is Ls = 8,000 and the number of nonsynonymous sites is Lns = 20,000.

	Polymorphisms	Substitutions
Synonymous	224	950
Nonsynonymous	65	598

(a) Based on the ratio of the observed number of non-synonymous and synonymous substitutions alone, and assuming synonymous substitutions are neutral, compute K_A/K_S.

(b) Using the estimate of K_A/K_S from (a), is there evidence this gene may have been under positive selection (i.e. is $K_A/K_S > 1$)?

(c) Using all of the data (both substitutions and polymorphism), does the gene appear to be evolving neutrally? To assess this, compute the neutrality index (NI).

(d) Is the NI index consistent with positive selection? If so, estimate α, the fraction of nonsynonymous substitutions that are due to adaptive substitutions.

(e) Is it possible for K_A/K_S as estimated in parts (a and b), to be less than 1 (arguing for selective constraint on the gene), and have evidence for positive selection based on NI (parts c and d)? Why or why not?

(7) (10 pts, graduate/Extra credit): Based on the paper by Justin Fay (Weighing the evidence for adaptation at the molecular level; *Trends in Genetics*; http://www.genetics.wustl.edu/jflab/fay11_inpress.pdf), list and briefly explain two possible reasons why some species show evidence for abundant positive selection using the McDonald-Kreitman test and why other species show little such evidence.