

Write your answers and show your work only in the space provided for each question. I will not grade anything written anywhere else, and I will not grade anything that is written illegibly. When calculations are necessary to get an answer, you must show the calculations to get credit. All quantitative answers must have units (e.g. 27 substitutions/site X year). Partial credit will be given when appropriate.

1. (20) Below are electrophoretic genotypes of the *Adh* gene and numbers of individuals observed in a sample of 35 phlox plants.

(a) Calculate the allele frequencies and the expected and observed heterozygosity.

| <u>genotype</u> | <u>number</u> |
|------------------------------------|---------------|
| Adh ^a /Adh ^a | 2 |
| Adh ^a /Adh ^b | 5 |
| Adh ^b /Adh ^b | 12 |
| Adh ^b /Adh ^c | 10 |
| Adh ^c /Adh ^c | 5 |
| Adh ^a /Adh ^c | <u>1</u> |
| | 35 |

f(Adh^a) _____ f(Adh^b) _____ f(Adh^c) _____
observed heterozygosity _____ expected heterozygosity _____

(b) Suppose both alleles of the gene had been sequenced from each of these 35 plants. Would you expect the expected heterozygosity calculated from sequence data to be greater, the same as, or less than what you calculated above? Briefly explain your answer.

2. (21) In the distant (?) future, a medical genetic engineer treats a woman who wishes to have a normal child in spite of the fact that she has sickle-cell anemia. He takes an egg from the woman and replaces the Hb^S -globin gene with a slightly modified Hb^A -globin gene. The egg is fertilized *in vitro* with a sperm from a normal donor and a child is reared to reproductive maturity. That individual, hereafter called F_1 , is phenotypically normal but heterozygous for the modified -globin gene.

(a) Assume that for human beings at that time, $N = N_e = 5 \times 10^9$. What is the probability that the modified -globin gene will be fixed in the human population?

(b) What is the probability that it will be lost?

(c) What is the probability that the modified -globin gene will be lost in the first generation, i.e. will never appear in the progeny of the F_1 , if the F_1 has 2 children?

3. (21) A segment of the *yfg* gene was amplified and sequenced from two different species and aligned; the fossil record showed that these two species diverged 220 million years ago. Both sequences are from an exon and begin at the first position in a codon.

```
ACTTTAACA---GCAGCAATCACAATGCAGACTTTT  
ACACTAACACCTGCAGCTATTACAATG---ACATTA
```

(a) Calculate the sequence divergence, corrected for multiple hits, between these sequences. Show all stages in your work.

(b) Calculate the rate of sequence divergence.

4. (8) In a population of 100,000 diploid organisms in which $N_e = N$. If the mutation rate is $u = 10^{-9}$ mutations per base pair per generation, what proportion of bases is expected to be different in the two alleles of the *mfg* gene in one individual organism? (Assume the population is random mating and all mutations are effectively neutral.)

5. (21) Suppose that you compared the rates and patterns of sequence evolution in a number of protein-coding genes in two species of bacteria, A and B. You found that the synonymous substitution rates are similar in the two species, but the nonsynonymous rates are higher for each gene in species A. For each of the following, indicate whether it could explain this difference, and why.

(a) effective population size

(b) mutation rate

(c) selection

6. (9) Briefly explain how transposable elements contribute to the C-value paradox.
