

PRACTICE PROBLEMS 2 ANSWERS

Problem 5. The tree shows two cases of horizontal transfer of the gene. One was from a bacterium related to the ancestor of *Yersinia pestis* and *Vibrio cholerae* to a common ancestor of the three mammals, or vice versa. The other was from a relative or ancestor of *Pasteurella multocida* or *Haemophilus influenzae* to *Trichomonas* or an ancestor thereof. The mechanism of transfer in these cases is unknown.

Problem 6.

$$f(\text{AdhF}) = 0.40 + 0.32/2 = 0.56 \quad f(\text{AdhS}) = 0.28 + 0.32/2 = 0.44$$

Problem 7.

$$\text{AcpH-1H} = 2 \times 0.995 \times 0.005 = 0.00995$$

$$\begin{aligned} \text{AcpH-2H} &= 1 - (0.009^2 + 0.066^2 + 0.882^2 + 0.014^2 + 0.005^2 + 0.024^2) \\ &= 0.217 \end{aligned}$$

$$\text{Est-5 H} = 1 - (0.483^2 + 0.396^2 + 0.122^2) = 0.595$$

Problem 8. To answer this problem, I went through the sequences looking for the restriction site and found two, underlined below. I also found and underlines sites where there was diversity for a base pair and underlined those sites on the top sequence. Then I made a summary of the differences, using dashes in the sequences where they were identical. The + indicates the first restriction site was present (I didn't have to worry about the second site because all sequences had it, so by restriction alone no one would know it was there at all). I also translated the sequences using a program called DNA Strider, but it would be easy to do by hand because you only have to translate the codons in which there is sequence diversity.

```

1  AAAACGACATTAACAGCAGCAATAACTATGACTTTAGCAGCACGTGGAAACTCCGTAGGTAAAAAATATGAAGAA
2  AAAACGACATTAACAACAGCAATAACTATGACTTTAGCAGCACGTGGAAACTCCGTAGGTAAAAAATATGAAGAA
3  AAAACGACATTAACATCAGCAATAACTATGACTTTAGCAGCACGTGGAAACTCCGTAGGTAAAAAATATGAAGAA
4  AAAACGACATTAACAGCAGCAATAACTATGACTTTAGCAGCACGTGGAAACTCCGTAGGTAAAAAATATGAAGAC
5  AAAACGACATTAACAGCAGCAATAACTATGACTTTAGCAGCACGTGGAAACTCCGTAGGTAAAAAATATGAAGAC

```

```

1  -----G+-----A-----A
2  -----A-----A-----A
3  -----T-----A-----A
4  -----T+-----G-----C
5  -----G+-----A-----C

```

```

1  ala  glu
2  thr  glu
3  ser  glu
4  ala  asp
5  ala  asp

```

ala, thr, ser have neutral R
glu, asp have negative R

Answers:

Restriction with *BbvI* (cuts GCAGC): two alleles, 1+4+5, 2+3

DNA sequence: five alleles, 1, 2, 3, 4, 5

Enzyme electrophoresis: only one allele if separate on charge;
four alleles: 1, 2, 3, 4+5 if separate on any amino acid difference

Problem 9. In this problem, remember that the fixation probability of an allele is x at the time when its frequency is x ; as the frequency changes, so does the fixation probability.

- (a) $1/(2 \times 5 \times 10^4) = 10^{-5}$
- (b) 0.1
- (c) 0.8
- (d) 0.2×10^{-5} , 0.1, and 0.8

Problem 10.

- (a) 4×10^3 , 4×10^5 , 4×10^6 generations
- (b) $2N = 2 \times 10^8$ generations
- (c) The coalescent time is the same as the fixation time, or in this case 2×10^8 generations

Problem 11. If you tried to calculate the answers for (a), (b), and (c) using $H = 4N_e u$, you quickly ran into problems because the answers were 0.2, 20, and 2000 respectively. This is absurd because H is the probability that two randomly chosen genes are different alleles, and a probability can't be greater than 1. Instead you have to use $H/(1 +)$.

- (a) $0.2/1.2 = 0.167$
- (b) $20/21 = 0.952$
- (c) 10^3 , because only the H for this effective population size is reasonable; the other value is much higher than is typical for mammals, indeed higher than has ever been observed for any organism, so far as I know.
- (d) $= 4 \times 10^5 \times 5 \times 10^{-9} = 0.002$
- (e) 0 (The first codon is always ATG so the first base is always A.)

Problem 12. The numbers were chosen so that $N_e s = 1$ and $N_e/N = 1$, for easy calculation. Note that $2N_e s = 2$, so that selection is weak. The calculations can be done by using the formula $E = Nu[F_a(\text{freq. advantageous mut's}) + F_n(\text{freq. neutral mut's}) + F_d(\text{freq. detrimental mut's})]$. I got: $F_a = 0.002310724$, $F_n = 0.001$, $F_d = 0.000313349$, while the frequencies of advantageous and detrimental mutations were 0.1×0.8 and 0.9×0.8 in (a), and 0.1×0.2 and 0.9×0.2 in (b). Then the final answers were:

- (a) 3.1×10^{-9}
- (b) 4.5×10^{-9}

Both results are below the strictly neutral rate which would be 5×10^{-9} ; the second case is very close to neutral. Think about what the answers would be if only 1 out of 100 selected mutations are detrimental; this is a distinct possibility.

Problem 13. $E = K/2T = 0.744/160 \times 10^6 = 4.65 \times 10^{-9}$ bp substitutions per site per year. This is also a good estimate of the mutation rate, given that synonymous mutations are neutral or nearly so.

Problem 14. The synonymous mutation rate is essentially the rate of neutral substitution in mammals. Consequently these numbers are also equal to the mutation rate, which is much higher in the mitochondrion. There is evidence that the mitochondrial DNA polymerase is error-prone; that mitochondria lack some of the usual repair systems; and that the mitochondrion is rich in mutagenic compounds produced as a byproduct of metabolism.

Problem 15. (a) 5×10^{-3} (b) $T = d/2E = 5/10^3 \times 2 \times 5 \times 10^{-9} = 5 \times 10^5$ years (c) 5×10^5 years