**Problem Set # 6 – Due Fri., Nov. 29**

This is the sixth and final problem sets that will count towards your final grade. It is due **at the start of lecture** on Friday, Nov. 29. You may work in groups of up to three people. Please hand in one assignment per group with up to three names listed. Late assignments will be penalized 20% per day (or part thereof). **Staple** multiple pages together –no paper clips or folded corners as sheets inevitably get lost. **Show your work.** Incorrect answers with some correct work will receive part marks; correct answers with no work shown may not receive full marks.

1. Simulate the action of natural selection at a single locus with two alleles here: [http://evolutiongenetics.georgetown.edu/simulations/naturalselection](http://evolutiongenetics.georgetown.edu/simulations/naturalselection/). Set the genotype frequencies and relative fitnesses for the different cases below using the sliders (click on the slider ‘arrow’ and then use the arrow keys to move it back and forth). You may have to increase or decrease the number of generations to see what is going on.

**a) Genotype frequencies: *P*AA = 0.25, *P*Aa = 0.5, *P*aa = 0.25; Relative fitnesses: *w*AA = 1; *w*Aa = 0.9; *w*aa = 0.9**

i) What form of selection is this (e.g. directional selection for/against *A*/*a*, heterozygote advantage, heterozygote disadvantage)?

ii) Is *a* fully dominant, partially dominant/recessive, or fully recessive?

iii) What are the selection and dominance coefficients (i.e. *s* and *h*) associated with the *a* allele in this fitness set?

iv) Examine the plots. What happens to population mean fitness as allele frequency changes under selection? How does the maximum mean fitness relate to the equilibrium allele frequency?

v) Does one of the alleles eventually fix and if so, which one? Explain why fixation does or does not occur.

vi) Change the relative fitness of the heterozygote to 1. How does the rate at which *a* changes in frequency compare with that above where the heterozygote had a fitness of 0.9. Explain (1 sentence).

**b) Genotype frequencies: *P*AA = 0, *P*Aa = 0.5, *P*aa = 0.5; Relative fitnesses: *w*AA = 1; *w*Aa = 0.9; *w*aa = 0.8**

i) What form of selection is this (e.g. directional selection for/against *A*/*a*, heterozygote advantage, heterozygote disadvantage)?

ii) Is *a* fully dominant, partially dominant/recessive, or fully recessive?

iii) What are the selection and dominance coefficients (i.e. *s* and *h*) associated with allele *a* in this fitness set?

iv) Does one of the alleles eventually fix and if so, which one? Explain why fixation does or does not occur.

v) In the plot of genotypes frequencies, why do heterozygotes increase in frequency for the first 10-15 generations before decreasing?

**c)** **Genotype frequencies: *P*AA = 0.25, *P*Aa = 0.5, *P*aa = 0.25; Relative fitnesses: *w*AA = 1; *w*Aa = 0.8; *w*aa = 0.9**

i) What form of selection is this (e.g. directional selection for/against *A*/*a*, heterozygote advantage, heterozygote disadvantage)?

ii) Does one of the alleles eventually fix and if so, which one?

iii) Change the starting genotype frequencies to: *P*AA = 0, *P*Aa = 0.5, *P*aa = 0.5 Does one of the alleles eventually fix and if so, which one? What determines which outcome occurs?

iv) Identify all equilibria and their stabilities.

v) Examine the plot of $\overbar{w}$ (average fitness of the population) vs. *p*, from iii above. This is sometimes called a “fitness surface” or “fitness landscape”. Natural selection changes genotype frequencies in a way that increases the population mean fitness (i.e. causes the population to ‘climb uphill’ from its current position). What does this plot say about natural selection’s ability to maximize population mean fitness?

2. Given the phylogeny indicated below:

1. Explain, in three sentences or less, the basic concept behind a relative rate test for consistency in nucleotide substitution rates in the lineages leading from X to B and from X to A.
2. Perform this test using the sequence data below. State the null hypothesis, provide the final chi-square value, the degrees of freedom, and the associated P-value, and state your conclusion as to whether you reject or accept the null hypothesis. You can get a p-value here: <https://www.socscistatistics.com/pvalues/chidistribution.aspx>



|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A | ATG | CTA | GCG | TGC | ATG | CTA | GCA | ATA | ATA | GCC | CTA | ATC |
| B | ATG | CTG | GCA | TGT | ATG | CCA | GCA | ATA | ATG | GCT | CCG | ATT |
| C | ATA | CCA | GCA | TGC | ATG | CTA | GCG | ATA | ATA | GCA | CTA | ATC |

3. The human immunodeficiency virus HIV causes AIDS. In this virus, the rate of nucleotide evolution has been estimated at approximately 0.01 substitutions per synonymous site per year. Two viruses isolated in Zaire and San Francisco in 1983 differ in 1/3 of their synonymous sites. Estimate the **YEAR** in which these viruses last shared a common ancestor.

4. In *Drosophila melanogaster*, the phenotype *curly wings* is due to a mutant allele *Cy* that is lethal when homozygous. A population is established with an initial frequency of *Cy* equal to 0.168. Denoting + as the wild-type (i.e. non-mutant) allele at this locus, calculate the expected frequency of *Cy* in the next generation if the relative fitness of the ++ homozygote to the *Cy*/+ heterozygote is:

a) 1 : 1

b) 1 : 0.5

c) Briefly explain why the change in frequency of the *Cy* allele is greater in one case than the other.

5. In molecular genetic terms, what is a ‘multiple hit’ and why do we need to account for them when estimating substitution rates (or time since the most recent common ancestor).