PROBLEM SET 2 KEY EVOLUTIONARY BIOLOGY FALL 2017

Mutation, Selection, Migration, Drift (20 pts total)

1) In class we discussed some prion diseases including the infamous Kuru. There are many other prion diseases found in wildlife and livestock that occasionally are transferred to humans. One alarming prion disease known as Chronic Wasting Disease (CWD) has been found in a number of North American deer populations and can be transfer to humans via the consumption of infected animals. In some deer populations resistance seems to be due to the presence of an allele, **R** at a single locus in which the normal nonresistant allele may be denoted **N**. In the **absence** of CWD, the scaled fitnesses of the **NN**, **NR**, and **RR** genotypes have been estimated as 1.00, 0.61, and 0.28 respectively. In the **presence** of CWD, the scaled fitnesses have been estimated as 0.23, 1.00, and 0.78 respectively. (5 pts total)

a) Calculate the selection coefficient (s) against the **RR** homozygous genotype when CWD is not present. Calculate the dominance coefficient (h). (2 pts)

W_{NN}	W_{NR}	W_{RR}
1.0	0.61	0.28

$W_{NN} = 1$; $W_{NR} = 1 - hs$; $W_{RR} = 1 - s$

$s = 1 - 0.28 = 0.72$
$1 - hs = 0.61$, $h = (1-0.61)/0.72 = 0.54$

b) Calculate the equilibrium frequency of **R** in the presence of CWD. (1 pt)

In the presence of CWD the form of selection operating is Heterozygote Advantage (overdominance).

Solve for the selection coefficients *s* and *t*

W_{NN}	W_{NR}	W_{RR}
0.23	1.0	0.78

$W_{NN} = 1 - s$; $W_{NR} = 1$; $W_{RR} = 1 - t$

$s = 1 - 0.23 = 0.77$
$t = 1 - 0.77 = 0.22$

(note, be careful of the sign of s & t. With overdominant selection (heterozygote advantage) s & t cause a decrease in fitness and are negative. However, as long as they both have the same sign the estimate of equilibrium frequency will not be affected. When estimating the equilibrium frequency of an allele the selection coefficient against the homozygote for the opposite allele is in the numerator.

To estimate the equilibrium frequencies:

$$\hat{p}_N = \frac{t}{(s+t)} = \frac{-0.22}{(-0.77-0.22)} = 0.22$$

$$\hat{q}_R = \frac{s}{(s+t)} = \frac{-0.77}{(-0.77-0.22)} = 0.78$$

c) In the absence of selection from CWD, the effects of the **R** allele on fitness is negative and its frequency in a population would be determined by the balance between mutation and selection. Given the probability that **N** mutates to **R** at a rate of 4×10^{-5}, what are the frequencies of **N** and **R** at mutation-selection equilibrium? How does this equilibrium value compare to the equilibrium

values under CWD selection? (1 pt)

For a partially recessive deleterious allele the equilibrium frequency at mutation–selection balance is:

$$\hat{q}_R = \frac{u}{hs} = \frac{4x10^{-5}}{(0.54)(0.72)} = 1.03x10^{-4}$$

$$\hat{p}_N = 1 - (1.03x10^{-4}) = 0.9998$$

In the absence of CWD, and under mutation-selection balance, the R allele will be maintained at very low frequencies. In contrast, when CWD is present there in a much higher equilibrium frequency of the R allele.

d) A small amount of dominance can have a major effect in reducing the equilibrium frequency of a harmful allele when there is a mutation-selection balance. To confirm this for yourself, imagine the ***R*** allele is completely recessive to the **N** allele when no CWD is present. In this case, the scaled fitness values of the **NN**, **NR**, and **RR** genotypes are 1.00, 1.00, and 0.28 respectively. How does the equilibrium value of the two alleles change when R is completely recessive? (1 pt)

For a deleterious completely recessive allele the equilibrium frequency in mutation-selection balance is:

$$\hat{q}_R = \sqrt{\frac{u}{s}} = \sqrt{\frac{4x10^{-5}}{0.72}} = 0.0075$$

When the R allele is partially dominant it has a higher equilibrium frequency than when it is completely recessive. When the R allele is completely recessive it is 72 times more common than when it is partially dominant. The "load" of completely recessive deleterious alleles in a population is typically higher than the "load " of partially recessive or dominant deleterious alleles.

2) Conservation managers are concerned about the CWD susceptibility in a deer population that has a very low frequency of the R allele so they are considering introducing individuals from another population with a higher frequency of this allele. Before they start mixing these populations they want to estimate the level of gene flow between them. They gather data from a single allozyme locus with 2 alleles and determine the frequency of these 2 alleles in each population. The data are shown below. (5 pts total)

a) Using these data, and assuming that no selection operates on this allozyme locus, and the populations are in Hardy-Weinberg equilibrium, calculate F_{ST}. (2 pts)

	p	***q***	***2pq***
Population 1	0.9	0.1	0.18
Population 2	0.3	0.7	0.42
Average frequency =	0.6	0.4	

Average expected heterozgosity within populations H_s= (0.18+0.42)/2 = 0.30
Total expected heterozygosity H_t= 2(0.6*0.4) = 0.48

$F_{st} = (H_t - H_s)/H_t$= (0.48-0.30)/0.48 = 0.375

b) Assuming the necessary assumptions are met, what is the effective number of migrants (N_m) among these populations each generation? (2 pts)

$F_{st} = 1/(1+4N_m)$ =>N_m = 0.42 or ~1 migrant every 2 generations

c) Assuming other genes in these populations related to local adaptation have a similar F_{ST} as this allozyme locus, do you think it is a good idea to mix these populations? Why or why not? (1 pt)

Depends. Mixing these populations will increase the heterozygosity of each (2pq will go up). Higher levels of genetic variation allow populations to evolve more rapidly under directional selection. So, from this perspective mixing is a good strategy to accommodate future selective challenges. However, if the differences in allele frequencies reflect different selective pressures (i.e., are locally adapted) in the two populations then mixing will homogenize the populations and make them less fit under their current selective pressures.

3) Declines in available habitat due to deforestation have led to dramatic population declines in many Malaysian primate species including Orangutans. To provide a demographic buffer, a conservation organization in Sumatra maintains a captive colony of 80 Orangutans. (4 pts total)

a) If there are 60 female and 20 male Orangutans in this colony, what is the effective population size (N_e)? (2 pts)

$$N_e = \frac{4N_mN_f}{N_m + N_f} = \frac{4(20)(60)}{20 + 60} = 60$$

b) One big concern in any captive breeding program is that genetic variation will be lost over time via random genetic drift and that this loss will cause a decrease in fitness. If the original heterozygosity in the colony was 0.60, what proportion of the original heterozygosity would be lost after 15 generations? Assume the colony is kept at its current size and sex-ratio. (2 pts)

The heterozygosity remaining after 15 generations is given by the following equation:

$$H_t = \left[1 - \frac{1}{2N_e}\right]^t * H_0 = \left[1 - \frac{1}{2(60)}\right]^{15} * 0.60 = 0.53$$

The proportion of heterozygosity lost is 1- (0.53/0.60) = 0.12 ~12%

You could also use the exponential form of the equation.

4) In class we discussed some nice examples of coloration being under selection including peppered moths in Kettlewell's experiments on industrial melanism and King & Lawson's water snakes. Often these color patterns or color intensities are under selection from predation. Another nice example of this phenomenon is found in lizards that live in the White Sands National Monument in New Mexico. Lighter individuals are more cryptic in this environment. To investigate how selection is operating on coloration in these lizards, graduate students from Notre Dame conducted a mark-recapture experiment similar to Kettlewell's (1973) classic study. Use the data from this mark-recapture experiment (given below) to answer the following questions. Assume that lizard skin coloration is controlled by a single locus with two alleles (L & D) and that you can directly infer the underlying genotype of an individual by its skin color. In this wild population, there is incomplete dominance so that all three genotypes can be distinguished based on skin color. (6 points total)

Dark skin color= DD homozygote
Medium skin color= LD heterozytgote
Light skin color= LL homozygote

Genotype	Marked Sample	Recaptured sample
DD	200	50
DL	900	250

LL	400	200
Total (N)	1500	500

a) What is the survival rate for each genotype? (2 pts)

To estimate the survival rate, the expected number of individuals for each genotype must be calculated from the frequency in the original sample.

Genotype	Frequency in the Original Sample
DD	200/1500 = 0.13
DL	900/1500 = 0.6
LL	400/1500 = 0.27

Genotype	Expected in the Recaptured Sample (1pt)
DD	500(0.13) = 67
DL	500(0.6) = 300
LL	500(0.27) = 133

Survival Rate = Ratio of Observed to Expected (O/E) (1pt)

DD	50/66 = 0.75
DL	250/300 = 0.83
LL	200/133 = 1.5

b) What form of selection is operating on this locus? (1 pt)

Directional

c) What are the selection (**s**) and the dominance (**h**) coefficients? (2 pts)

To estimate the selection and dominance coefficients the survival rates need to be scaled to one and then using the general framework for directional selection the coefficients can be estimated.

Fitness = Survival Rate scaled to the best genotype

DD	0.75/1.5 = 0.5
DL	0.83/1.5 = 0.56
LL	1.5/1.5 = 1.0

LL	DL	DD
1	1-hs	1-s

s = 1 - 0.5 = 0.5 (1 pt)
hs = 1 – 0.56 = 0.44
h = .0.44/0.5 = 0.88 (1 pt)

d) Assuming the frequency of the light-skinned allele (L) in this population is Freq(L) = p = 0.70 what is the population mean fitness ($\overline{W}$)? (1 pt)

HWE Frequencies

LL $(0.7)^2 = 0.49$
DL $2(0.7)(0.3) = 0.42$
DD $(0.3)^2 = 0.09$

$\bar{W} = p^2(W_{LL}) + 2pq(W_{DL}) + q^2(W_{DD})$= 0.49(1.0) + 0.42(0.56) + 0.09(0.5) = 0.77