

# Bi/GE105: Evolution

## Homework 4

### Due Date: Friday, February 21, 2014

“Problems worthy of attack prove their worth by hitting back.” - Piet Hein

#### 1. Population Genetics.

In the same style in which we have done the written parts of assignments in previous homeworks, provide a one-paragraph summary of the goals and the insights that emerge from population genetics.

#### 2. Genetic Drift and Population Size.

In class, we gave a cursory examination to the question of how many generations it takes to fix an allele from genetic drift as a function of the population size. In this problem, we return to our simulations of genetic drift to try and provide a more accurate answer to that question using our simple “toy model” of two alleles at a single locus.

Using the function “*num\_gens\_to\_fixation.m*”, compute the number of generations until fixation for the following population sizes: [5, 10, 50, 100, 500, 1000]. Please explain what the simulation is doing and why genetic drift occurs at all. Make sure to run the function at least 10 times for each population size (using a “for loop” is recommended). Make a plot of the average number of generations to fixation vs population size. Describe the implications of your results for small, isolated populations of organisms (like those living in the Galapagos).

#### 3. Population Genetics and Selection.

In class we talked about how allele frequencies change over time in the presence of selection. In class, we derived the change in allele frequencies in terms of the average excess of fitness.

(a) In the first part of the problem fill in all of the details for the arguments made in class and derive the change in the frequencies  $\Delta p$  and  $\Delta q$ . Just as

a reminder, we showed that

$$\Delta p = \frac{p}{\bar{w}} a_1 \quad (1)$$

where

$$a_1 = p(w_{11} - \bar{w}) + q(w_{12} - \bar{w}) \quad (2)$$

is the average excess of fitness and we defined the mean fitness as

$$\bar{w} = p^2 w_{11} + 2pqw_{12} + q^2 w_{22}. \quad (3)$$

Essentially, provide a clear derivation of the way in which allele frequencies change from one generation to the next in the case where selection is in effect. Find both  $\Delta p$  and  $\Delta q$ . Given those results for  $\Delta p$  and  $\Delta q$ , should we expect that  $\Delta p + \Delta q = 0$  and if so, are your results for the changes in allele frequencies consistent with that constraint? To demonstrate this, you need to actually calculate  $\Delta p + \Delta q$ .

(b) As a toy model for how allele frequencies associated with the sickle cell trait change assume that the fitnesses are  $w_{AA} = 0.9$ ,  $w_{AS} = 1.0$  and  $w_{SS} = 0.2$ , where we use  $A$  for the wild-type allele and  $S$  for the allele associated with sickle-cell. Imagine that the  $S$  allele begins with a very low frequency,  $q$ . First, give an intuitive argument for what you expect the *sign* of  $\Delta p$  and  $\Delta q$  to be in the next generation of reproduction. Then, by assuming that  $q = 0.001$ , find  $\Delta p$  and  $\Delta q$  by using the formulae you derived in the first part of the problem. Make sure you identify the mean fitness and the average excess of fitness. Do  $\Delta p$  and  $\Delta q$  have the signs you expect intuitively? Given this result, explain what will happen over time to the allele frequencies.

#### 4. Mutation-Selection Balance.

Another topic discussed in a cursory fashion in class was the way in which mutation and selection compete to determine allele frequencies. In this problem we are going to work out the mathematics of this effect explicitly. Our strategy is similar in spirit to what we did in class where we introduced the idea of an urn from which we draw our alleles  $A_1$  and  $A_2$ . We imagine that the allele  $A_1$  is mutated into  $A_2$  with probability  $\mu$  each generation and that the rate of mutation of  $A_2$  back to  $A_1$  is negligibly small. This scenario could be realized if we think of  $A_2$  really as the set of all the ways that the gene allele  $A_1$  can be broken. There are thus many ways to break our allele but

only one way that mutation can restore it and hence this rate is assumed small.

For the simple case in which there are only two possible alleles, we consider the process in two steps. To be concrete, consider the case in which the fitnesses are  $w_{11} = 1$ ,  $w_{12} = 1$  and  $w_{22} = 1 - s$ . First, show that we have

$$p_{\text{after selection}} = \frac{p}{1 - q^2 s}. \quad (4)$$

We already worked this out in class, so your job is simply to recapitulate that discussion and to explain all of the steps leading up to that equation, both mathematically and conceptually. Now we need to impose the results of mutation. Impose the mutation by computing

$$p' = p_{\text{after mutation/selection}} = (1 - \mu) \frac{p}{1 - q^2 s}. \quad (5)$$

Explain what this equation means and how it captures the amendment of the allele frequency  $p$  due to mutation. For the case when mutation and selection exactly balance, we have  $p' = p$ . This implies that we have

$$p = (1 - \mu) \frac{p}{1 - q^2 s}. \quad (6)$$

Show that this steady state condition implies that

$$q = \sqrt{\frac{\mu}{s}}. \quad (7)$$

Examine your result qualitatively and provide an argument as to how increasing either mutation rate or selection coefficient will alter allele frequencies. Does this equation make sense intuitively? Why is this result referred to as “mutation-selection balance”.