**BIOL 502 Population Genetics**

**Spring 2017 Midterm Exam 2 (Take-home)**

**Instructor: Dr. Arun Sethuraman**

**Due on Turnitin, Tuesday, April 18th, 2017 at 4 PM.**

**No late submissions/email submissions will be accepted.**

1. Consider a stepping-stone model of population subdivision, and migration, such that only subpopulations that are directly adjacent to each other are able to exchange genes, in a single direction, i.e. Subpopulation 1 -> Subpopulation 2 -> Subpopulation 3. This can often be generalized as the “isolation-by-distance” model, or the “serial-founder” model, wherein one subpopulation of individuals “migrate” and establish a new subpopulation, and so on. Consider a single, biallelic genetic locus, with two alleles A and a. Derive the allele frequencies after one generation of migration in all three subpopulations, and show mathematically that subpopulation 1 and subpopulation 2 are more genetically similar to each other than subpopulation 1 and subpopulation 3. (20 points).

2. Consider an inbreeding population in which a deleterious mutation, *a*, at a locus *A* arises at the rate of μ per generation. This mutation *a* can also back-mutate to the wildtype allele *A* at the same rate. (a) What would be the expected genotype frequencies of the *AA*, *Aa*, *aa* genotypes after one generation of inbreeding and mutation? (10 points)

(b) Homozygotes with this mutation have much lower fitness than the homozygote wildtype, and the heterozygote has intermediate fitness. Consider a scenario where the mean fitness of AA, Aa and aa is 1, 1-hs, 1-s respectively, where s is the selection coefficient, h is the degree of dominance. Mathematically, when does inbreeding depression occur in this population? (10 points)

3. Show graphically (using R, write the simulation and please include your code and plot(s) with the exam) how an advantageous (s > 0) and deleterious allele (s < 0) go to fixation in populations of varying sizes. Use your own combinations of N. (10 x 2 = 20 points)

4. Pick a Mendelian genetic disorder locus below according to the first letter of your last name. Do some research on the associated gene’s allele frequency in human populations, and its dominant/recessive nature, as well as the associated disease. (a) Given the information you have collected (please include references – there will be zero partial credit if you don’t add a reference/citation), what would be the expected frequency of observing the disorder in a randomly mating human population? (10 points) (b) What would be the expected frequency of observing this same disorder in a population that is comprised only of offspring of first cousin mating? (10 points) You’re welcome to get your allele frequency data for any SNP at these loci from any database online (e.x. ALFRED - <https://alfred.med.yale.edu/>, or OMIM – [www.omim.org](http://www.omim.org), or dbSNP, ClinVar, Genbank, etc, as long as you add references, and compute your numbers accordingly).

A-E: BRCA1

F-J: P53

K-O: IDDM

P-T: PSORS1

U-Z: PDCD1

5. Read the Holsinger and Weir (2009) paper posted on Cougar Courses (under March 28-Apri 3, Holsinger and Weir 2009 – Interpreting Fst), and Mobegi et al. 2014 (under March 7-13, Mobegi et al. 2014). Based on these papers, answer/interpret the following:

a) Figure 6 from Mobegi et al. 2014, showing genome-wide Fst between the Guinean population and the Gambian population of mosquitoes. (10 points)

b) Describe two definitions and interpretations of Fst in your own words. (2 x 5 = 10 points)