This homework is primarily to see how much population genetics you remember from previous courses. If you find it difficult, please don't panic; we will cover all of this material in more detail during the course. Please pay particular attention to **boldface** terms as they are meant in a technical sense. Also, be sure not to round so much that the effect you are looking for disappears! (I recommend at least 4 digits after the decimal.)

1. Cystic fibrosis (CF) is a disease caused by a mutation at the CFTR gene. Mutant homozygotes are gravely ill and prior to modern medical treatment would normally die before reproducing. Heterozygotes ("carriers") appear normal.

   A village was founded by 98 normal homozygotes and two heterozygotes.

   (a) (1 pt) What are the **allele frequencies** of the normal and disease alleles at the time of founding? *pA=0.99, pa=0.01*

   (b) (1 pt) In the next generation, assuming random mating, what are the expected frequencies of the three **genotypes** (homozygous normal, heterozygous, homozygous diseased) at birth? *pAA=0.9801, pAa=0.0198, paa=0.0001. Be sure not to round too much here.*

   (c) (1 pt) What are the frequencies in survivors of that generation at age 50 (assuming that all CF homozygotes die, and no one else dies)? *pAA=0.9802, pAa=0.0198, paa=0.0. Be especially careful not to round too much here, or you will lose the effect of the selection. Also, be sure to divide through by the new total. Note: a few students interpreted this question as "What proportion of the population survives?" and I did give credit for correct answers to that question.*

   (d) (2 pts) If no further mutations to create CF alleles occur, what would be the expected long-term fate of the CF allele? You may want to draw a rough picture of the change in allele frequencies over time. *The CF allele will decrease due to lethality in the homozygote. However, it will decrease more and more slowly as it becomes rarer. The common wrong answer was that it would stay the same, perhaps because the numbers got rounded too severely and looked stable. But if CF alleles are being removed and not replaced, it must decrease.*

   (e) (2 pts) If the village had a tiny population such as 20, would your predictions about changes in CF allele frequency be any different? *With a population this tiny, genetic drift will be very strong. The CF allele could disappear instantly or could even increase in frequency. The average outcome (decrease) is the same but surprising outcomes will be much more frequent.*

   (f) (2 pts) A particular village is found to have a very high frequency of CF. What are two possible explanations? *Genetic drift (random chance) in a small population. Founder effect: the village was founded by a small number of people, one or more of whom were carriers, so the initial frequency was high. Heterozygote advantage: the heterozygote is good in some way and this compensates for the bad homozygote. Nonrandom mating: if the villagers have a marriage pattern which avoids creating homozygotes, they can avoid the selection. (Or if they have a mating pattern which creates homozygotes, they will briefly have a lot of CF cases before the allele is removed by selection.) Population bottleneck (much like founder effect).*

   Two less likely explanations are mutation–a mutation rate high enough to do this against the force of selection would probably require the village to be in Chernobyl–and some environmental difference which makes CF homozygotes viable in this particular location.

2. A salmon researcher found the following genotype frequencies of a protein-coding gene (it has two alleles, A and B):

Genotype	Frequency
AA	0.45
AB	0.05
BB	0.50

   (a) (1 pt) What are the **allele frequencies**? *pA=0.475, pB=0.525*

   (b) (1 pt) What **genotype frequencies** would be expected in a randomly mating neutral population (i.e. what are the Hardy-Weinberg proportions)? *pAA=0.2256, pAB=0.4988, pBB=0.2756*

   (c) (3 pts) Suggest at least three reasons why this gene might not exhibit Hardy-Weinberg proportions. *Small population size. Small sample size (perhaps this is not even a significant deviation–we would need to know the sample size to tell). Admixture between a mostly AA population and a mostly BB population; these could even be separate*

species. Strong selection against the heterozygote. Mating preference for same-allele matings. Inbreeding (preferentially mating with relatives). Presence of an unobserved allele. A less good explanation is mutation: mutation would not tend to eliminate heterozygotes, nor to make such dramatically skewed frequencies.

(d) (1 pt) Can you do a statistical test, based on the given data, of whether these deviations from H-W are significant? If so, how? If not, why not? *No, because we don't know the sample size. These frequencies could reflect 20 sampled individuals or 20,000, and the significance of the results would be vastly greater in the second case. If we did know the sample size, a χ^2 test would be appropriate.*

3. A mutant allele of the Duffy blood-type gene, DuffyO, grants complete immunity to one form of malaria, *Plasmodium vivax.* DuffyO has a frequency close to 100% in sub-Saharan Africa and is nearly unknown elsewhere.

(a) (1 pt) Give a plausible explanation for the high frequency of DuffyO in Africa? *Selection driven by malaria.*

(b) (2 pts) The area of the chromosome close to Duffy has almost no variation in sub-Saharan Africans; in non-Africans it is about as variable as the rest of the genome. As far as we can tell there are no functional genes or control elements in this area. Why is it so lacking in diversity in Africans? *The DuffyO allele arose relatively recently and swept through Africa. Thus, in this part of the genome all Africans have a recent common ancestor, so they show little variation. Non-Africans have a much more distant common ancestor for this region, and have had time to accumulate variation.*

This is considered strong evidence that DuffyO is a new allele and Duffy+ is the ancestral form. The alternative theory that DuffyO is the ancestral form and non-Africans evolved Duffy+ would predict more variability in Africa, not less.

It is unlikely that this effect is caused by selection for keeping the region around Duffy unchanged, given that there are no genes there. Selection usually acts on genes.

(c) (2 pts) *Plasmodium vivax* is currently found in Asia, Europe and the Americas, but NOT in most of Africa. Speculate on the past history of *P. vivax*. It may be helpful to know that human hosts are essential to its life cycle. *If P. vivax drove the spread of DuffyO, it must have been in Africa in the past. It is not there now, presumably because DuffyO has destroyed its ability to infect African humans, but some P. vivax populations escaped into the rest of the world where DuffyO is not yet prevalent. We might expect DuffyO to be increasing in other parts of the world now, but this will take time.*

An alternative explanation is that DuffyO became common in Africa for some other reason, and the protection against P. vivax is a coincidence. Under this hypothesis, perhaps P. vivax was never in Africa. Malaria does not fossilize well so these two theories are hard to test. Malaria organisms found in Egyptian mummies are P. falciparum, not P. vivax–however, this does not prove absence of P. vivax.