Laboratory 1 Part 2: Forces of evolution

The purpose of this part is to explore the effects of various combinations of natural selection, genetic drift, migration, and mutation. For partial guidance, we present several questions below that you should specifically answer. While we have not covered genetic drift in any detail, for the purposes of this part, you can envision drift to be the effect of sampling variation when the population size is less than 'arbitrarily large'. The object of this exercise is really to give you some feel for how gene frequencies are expected to change in natural populations. In Part 2a, we use *Populus*; in Part 2b, we use instead the *PopG* model, because it allows one to more easily study the effect of migration along with all the other forces.

Part 2a. Brief Synopsis of Assumptions and Questions

For all simulations and problems in this part make the following assumptions. Assume that coat color in a certain strain of mice is controlled by one gene with 2 alleles. One allele codes for black coats (*A* allele), and the other codes for white coats (*a* allele). In the population you find 3 coat phenotypes: black (*AA*), gray (heterozygotes – *Aa*), and white (*aa*). Now, assume we have a stable population of mice living on an island with no owls. For convenience, let's assume that there are just as many "A" alleles in the population as "a" alleles (unless otherwise noted), and the population starts out in Hardy-Weinberg equilibrium.

2a.1 Simulation 1. Here we will assume we have a very large, isolated mouse population with no appreciable mutations in coat color alleles, and random mating. When owls find their way to the island, it suddenly becomes somewhat more dangerous to be a white mouse. We want to know how the mouse population evolves in response to this selection pressure. How strong does selection have to be in order for there to be a response to it?

1. Open up Populus and go to the Selection Models. Choose Selection on a Diallelic Autosomal Locus (by the way, what is a diallelic autosomal locus?).
2. Set plot options to "genotypic frequencies vs. *t*."
3. Choose "Fitness" (rather than "Selection"). Fitness is expressed relative to other genotypes.
    a. For the fitness of *AA*, enter 1.0.
    b. For the fitness of *Aa*, enter 1.0.
    c. For the fitness of *aa*, enter 0.7
4. For initial conditions, choose one initial frequency and enter 0.5. Set number of generations at 130.
5. Hit "view."
6. If you select "6 Initial Frequencies" the plot shows *p* vs. *t* for 6 computer-generated initial frequencies of the *A* allele. However, you can't plot genotypic frequencies vs. time for this choice. If you want to examine genotype frequencies for different initial conditions, you must enter them one at a time (see question e below).
7. Save copies of the most relevant graphs. (For your purposes – you won't have to send them to me!)
8. Answer the following questions.
    a. Identify the lines representing the 3 genotypes. What happens to each one?
    b. If *AA* and *Aa* have equal fitness, why does the frequency of *AA* go up and the frequency of Aa go down?

c. If *aa* is bad, why doesn't that genotype disappear entirely? Why doesn't the *a* allele disappear? In fact, go back to the Plot Options box and check "*p* vs. *t*". This shows how the allele frequency (*p* = frequency of allele A) changes over time. What do you see?

d. What does this simulation tell us about the relationship between fitness and genotypic frequency?

e. Change the initial frequency of the *A* allele to 0.1 (leave everything else the same). In other words, we're assuming that for whatever reason, white mice outnumber dark mice on the island prior to the arrival of owls. So, why does the *aa* line start so high and drop so fast? Why does *Aa* increase, then decrease?

g. Plot "*p* vs. *t*" for this scenario. What does this tell you about how selection can work?

2a.2 Simulation 2. Here we will simulate the same large, isolated population of mice with no appreciable mutation in coat color alleles, random mating, and where individuals with white coats are spotted most frequently by predators, individuals with black coats are the next most frequently spotted, and gray individuals are rarely spotted by predators. In the same *Populus* model, Selection on a Diallelic Autosomal Locus, set everything up as before, except that this time, set the fitness of the *AA* allele at 0.9 (with *Aa* at 1.0 and aa at 0.7).

Questions:

a. What is the equilibrium condition?

b. What are the major differences between this simulation and the previous one?

2a.3 Simulation 3. Now, we'll simulate the same isolated population of mice, but with a small population size. We will assume random mating, no appreciable mutation in coat color alleles, and no differential survival among coat color phenotypes.

1. Open up *Populus*, then go to the Mendelian-Genetics-Genetic Drift models. Choose the Monte Carlo tab. Make sure the default settings read:
    a. Runtime = 3N generations
    b. Loci = 6
    c. Initial frequency = 0.5
    d. Population size = 500
2. Hit View. Each color follows the trajectory through time of the frequency of a particular allele (think of them as six randomly chosen independent loci within the genome). Note that these are neutral alleles (i.e., there is no selection acting on them, and they confer no survival or reproductive advantage relative to other alleles at that locus).
3. Run 18 trials, six with population size of 500, six at N = 50, and six at N = 5.
4. For each trial, record the following information
    a. Trial #
    b. Population size (N)
    c. Generations to first fixation (of any allele)
    d. Color of first to fixation
    e. Fixation ratio (up/down – i.e., was it fixed or lost from the population?)
5. Answer the following questions:
    a. Within each trial, did each of the 6 loci behave similarly? Why or why not?
    b. Did each color loci behave similarly across the iterations? Why or why not?

c. Were particular colors most likely to be the first to go to fixation? Why or why not?

d. Within a given population size, how much did time to first fixation vary?

e. How does population size affect time to first fixation?

f. If these loci are *neutral* with respect to selection, why are they changing in frequency over time? Why are some alleles winners and others losers?

g. Many people confuse small population size effects with drift. Genetic drift is one effect of small population size. One easy way to remember drift is that the colored lines were drifting randomly around on the plot. That random drifting *is* genetic drift. Note that drift occurs even in large populations, but is more dramatic and consequential in small populations.

h. These simulations show changes in gene frequency over time. Isn't that the definition of evolution? Were we watching evolution? Explain your answer.

i. For a given population, can you precisely predict when loss of genetic variation (fixation) will occur?

2a.4. Simulation 4. Drift vs.Selection. In the real world, drift and selection often operate simultaneously. In fact, drift and selection are probably the two most important agents of evolutionary change. But do they necessarily work hand in hand? Consider again our island mice. With the arrival of owls, the selective regime is against those very common white mice (but note that it's not lethal to be a white mouse – they do 90% as well as darker mice since they hide well in dense island vegetation).

1. Go to Mendelian Genetics-Drift and Selection. Alter the default settings to read:
   a. N = 500, $p = 0.1$, Generations = 500
   b. $AA = 1$, $Aa = 1$, $aa = 0.9$
2. Before you run the simulation, consider: in the absence of drift, do you expect the *a* allele to go extinct in such a population? Explain your answer. (Be fair! Don't peek!)
3. Now hit View and see what happens over 500 generations. Hit View 5 more times, each time seeing what happens. What did happen? Was it the same every time? Why did the *A* allele go to fixation in this exercise?
4. So … selection is pushing the frequency of the *A* allele upwards. But unlike what we saw in Simulation 1, it is not a smooth monotonic increase. The increase is jerky. That drifting line *is* genetic drift in action.
5. Now change the population size to 50 and hit View. What happens in this smaller population?
6. Hit View 5 more times. Are your results similar to what you saw in the population of 500?
7. When the population size was 500, you undoubtedly saw the frequency of *A* make its way upward (jerkily) until it eventually (at least often) became fixed in the population. And it probably did this when the population was at 50, too. But when the population was 50, you probably occasionally saw the frequency of A actually decline and hit zero (that is, a became fixed in the population), even though *a* **was being selected against**.
8. Were the mice on the island evolving? If so, what mechanism was responsible?
9. What do these results say about the power of selection and drift in small and large populations?

Part 2b. Systematic exploration of the forces – a summary study

OK, now that you've had some practice investigating the basics we would like to give you some feel for how gene frequencies are expected to change in natural populations. To that end, we'd like you to *explore* the effects of various combinations of natural selection, genetic drift, migration, and mutation. For this purpose, you will use the *PopG* program – it is not as flexible as *Populus*, but it does one thing well: one-locus, two allele simulations, with selection, mutation, migration, and drift. Drift is built-in because you specify the population size. (Please refer to the web page cited for instructions on how to use *PopG*.)

You might want to start with no mutation and no migration and look at the effects of natural selection including overdominance, as suggested below. Then perhaps look at the effect of mutation versus selection. You may also want to examine the effect of genetic drift alone by making all fitnesses equal, and the sizes of the other effects necessary to have a noticeable effect on the outcome in the face of drift. Be creative, but with some ability to compare results with other cases that differ in useful ways. After you're done, write up what you've found out in 2-5 pages. To provide a bit of framework, we pose the following specific questions, which you should address in your report. You should make sure your report does answer each of them – but in addition, goes beyond just a point-by-point answer to attempt some generalizations. (In particular, I don't provide any focused questions on drift – but you should still investigate that and include a discussion in your report.)

Selection

Consider a number of populations (8-10 is fine) that have their population size (N) being 1000 individuals, and no mutation.

1. Try cases with no mutation and no migration and all fitness values set to 1.0 (i.e. there is no selection). Does genetic drift in a population of size 1000 accomplish roughly the same changes in 1000 generations as genetic drift in a population of size 100 does in 100 generations? By running a number of simulations, check whether the probability that an allele is fixed by genetic drift is equal to its initial frequency in the populations – is this true or not?
2. Try a case with no mutation and migration, with the *A* allele favored by natural selection (with fitness of the *AA* genotype set highest and fitness of the *aa* genotype set lowest). Start with a small frequency of *A*. Is it always fixed? If one starts with a single copy of the allele, how does the probability that *A* is fixed compare with the selection coefficient favoring it in the heterozygote (compared to the fitness of the *aa* genotype? Is this fixation probability larger than the one you would get with the same initial frequency with no selection?
3. Try overdominance (*Aa* having the highest fitness). Does the gene frequency converge towards an equilibrium? Why does it vary from this equilibrium frequency? How large do the selection coefficients have to be to cause the gene frequency to stay away from fixation or loss for large amounts of time? In particular, what happens when the fitness values of *AA*, *Aa*, and *aa* are 0.99, 1.0, and 0.99, respectively?
4. Try underdominance (*Aa* having the lowest fitness). Is there a starting gene frequency that will result in some populations heading for fixation, and others heading for loss? If you add a small amount of migration, what will happen if you add a small amount of mutation in both directions?
5. Comparison: With migration but no selection or mutation, how much migration is needed to make the gene frequency curves from questions 1-4 above all wind up quite similar to each other? How much is needed to make them all end up at the same gene frequency?

Mutation alone & mutation vs. selection

6. With mutation but no selection or migration, how much mutation is needed to cause the gene frequencies to converge on a mutational equilibrium gene frequency? How does this value relate to the population size?

7. If an allele is selected against, can you set up mutation rates that will maintain it at low frequency in the population?

8. Also try different migration rates from 0 up to 0.1. For each, start the population with gene frequency 0.5. Describe the outcomes. I am not looking for detailed graphs of all runs but a general verbal description. However, distinguish between what happens with the different parameter values.

Migration alone & migration vs. selection

9. Consider the number of individuals migrating into a typical population in one generation. What is it in terms of the population size N and the migration rate *m*?

10. Can you find a value of this number of new migrants, above which the outcome is different than below it (i.e. some value that is roughly the border between two kinds of behavior of the simulations)? What is the value?

11. With nonzero migration rates, do you ever find that the populations permanently fix for different alleles?

Drift alone & drift vs. the other forces.

Please don't forget to explore this! PopG makes it easy to investigate this interaction…when would you judge that a population becomes 'arbitrarily large' in this two allele case?

(This is the end of Part 2 of the lab; please go back to the web page and continue with Part 3.)