Name:_____

Simulation of Population Genetic Processes

9/11/23

Quantification of population genetic structure has importance within many biological disciplines including human genetics, medicine, evolution, ecology, and conservation. In general, the field of population genetics is focused on determining the processes responsible for shaping spatial patterns of allelic variation. Four primary mechanisms are responsible for controlling allele frequencies within populations: **mutation, gene flow, natural selection, and genetic drift**. Non-random mating (including sexual selection) does not directly affect allele frequencies, but can influence genotype frequencies within populations.

Key Definitions:

Microevolution - changes in population allele frequencies over multiple generations

Mutation – the spontaneous creation of a new allele in a population

Somatic mutations – mutations generally restricted to an individual's fitness

Germline mutations – mutations that contribute to genetic variation within a population

Gene flow – the transfer of alleles from one population to another

Natural selection – the influence of the environment on relative fitness of individuals in a population. Some alleles in a population might be selected for, whereas others will be selected against.

Genetic drift – changes in allele frequencies due to mating efficiency and randomness associated with allelic segregation during gametogenesis and meiosis.

Non-random mating – sometimes considered a microevolutionary force due to changes in genotype frequencies

Phenotype- expression of the genotype in a given environment

Artificial selection – people select which organisms get to reproduce based on desirable traits

Codominant – type of inheritance in which two alleles of the same gene are expressed separately to yield different traits in an individual

---

It is important to make the distinction between what are called **neutral microevolutionary processes** versus **adaptive microevolutionary processes**. Natural selection is considered an adaptive process because it alters allele frequencies by 'selecting' for certain alleles that confer a fitness advantage to individuals possessing them. For example, beak morphology in Darwin's finches differs to best utilize different food sources that naturally

occur on different islands of the Galapagos. Conversely, **genetic drift, gene flow, and mutation are often considered neutral processes because they change allele frequencies through non-adaptive mechanisms.**

In real populations it is important to note that many times multiple microevolutionary processes act in synergy to control allele frequencies. For example, a new random mutation may originate in a population of fish leading to a new allele. This new mutation may confer some selective advantage over other alleles already present in the population. Because individuals possessing this allele are more fit, natural selection may increase the frequency of the new allele (Fig 1). Highly advantageous alleles are expected to reach **fixation** (all individuals are homozygous for a given locus) while disadvantageous alleles are **loss**. However, sometimes even advantageous alleles are loss to stochastic process of drift. Remember, **genetic drift changes allele frequencies by random processes**. For example, who mates with whom and which alleles are passed to which gametes during meiosis. In general, genetic drift has a more profound influence in smaller populations.

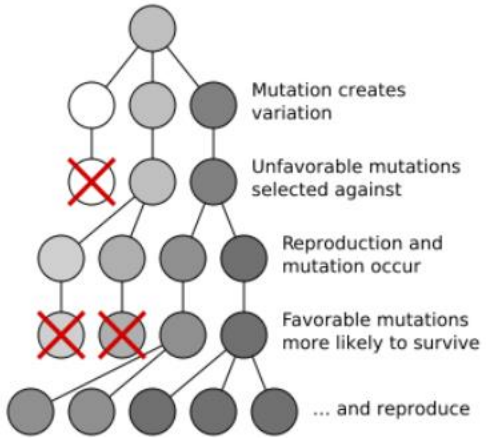

Fig 1. Example of how mutation and natural selection may interact to shape allele frequencies in populations.

Researchers have argued for decades on the relative importance of neutral versus adaptive processes in evolution. The Darwinian view posits that natural selection (i.e. adaptive processes) is the primary mechanism to explain allele frequencies and the evolution of biodiversity. This was a widely accepted hypothesis since Charles Darwin came up with his theory of evolution by natural selection. Decades later, a biologist and researcher named Motoo Kimura suggested that most evolutionary changes are neutral and do not confer a fitness advantage or disadvantage to organisms. Kimura's **neutral theory of molecular evolution** placed a much stronger emphasis on genetic drift to explain the evolution of life on Earth. Because few germline *de novo* mutations are found in functional regions (i.e. protein-coding regions) of the genome, the majority of mutations are considered to be under neutral or "effectively neutral" selection while the minority are under strong natural selection. For the majority of organisms, the most likely answer is that **both drift and selection are important in shaping allele frequencies in populations**, and both processes are vital in shaping the evolution of biodiversity.

Population genetic analyses have the expectation that genotype frequencies can be predicted from the allele frequencies in the absence of evolution. This mathematical equation, known as the **Hardy-Weinberg (HW) equilibrium** ($p^2 + 2pq + q^2 = 1$; p/q=proportion of p/q homozygotes, pq = proportion of heterzygotes), expects that the population has 1) random mating, 2) no immigration or emigration (closed system), 3)

large population size, 4) Mendelian inheritance, 5) no mutations, and 6) all genotypes are equally fit. If all of the Hardy-Weinberg assumptions are met, we can conclude that the population is not evolving. In reality, populations are always evolving due to one or multiple microevolutionary forces that change allele frequencies from generation to generation. Significant deviations of HW equilibrium can reveal aspects of the population like underlying substructure.

Activity 1: Simulating the influence of genetic drift

Please navigate to (link is posted on Canvas):

http://virtualbiologylab.org/ModelsHTML5/PopGenFishbowl/PopGenFishbowl.html

We are going to use Koi, domesticated carp from Japan, to explore patterns of drift, selection and mutation on allele frequencies in a population. Wild carp are dark brown while the domesticated breeds are brightly colored and patterned. These color patterns, or **phenotypes**, have emerged as a byproduct of **artificial selection** on the different color traits. Many genes are involved in the color and scale patterns with various inheritance patterns (not all Mendelian). The Kohaku variety has **codominant** color trait whereby the heterozygotes has a speckled pattern (allele Rr) compared to the solid homozygote variants (RR = orange, rr= white).

We will simulate **genetic drift**, or random change in allele frequencies, by creating a population of organisms and following it for several generations. To begin, leave all parameters at their default values. Note: Migration rate is 0, so we are dealing with a closed population, the mutation rate is 0 (alleles won't change by chance), and the phenotypes have the same fitness values (random mating).

To begin a simulation, click the blue icon with the curved arrow. The koi population you designed will begin moving around in their pond. Speed up time to the maximum (16x),

then click on the "To Data" Arrow to watch what happens. Let it run for about 100 generations-- this will only take a few minutes. Then press pause and assess the composition of the pond.

What are the frequencies of the genotypes rr, Rr, and RR in your pond?

rr=

Rr=

RR=

Repeat (click the blue arrow icon to restart the simulation). What are the frequencies of the genotypes rr, Rr, and RR in your pond?

rr=

Rr=

RR=

Are they the same or different from the previous simulation? Why?

Repeat for an additional 3 trials (100 gens each). Record the results below.

	1	2	3	4	5
RR					
Rr					
rr					

Are there any trends among the different simulations trials?

Go back to the settings window and change the initial size and carrying capacity from 200 to 50. What happened to the R allele after 100 generations? Run it once more, does it give the same result?

Next, change the initial population size from 50 to 500 and run the simulation. How does the initial size parameter influence the frequency of the R allele in each population? Does the R allele become fixed or lost? How does this compare to the simulation using a population size and carrying capacity of 50?

Why does drift have a larger impact on small populations?

Activity 2: Simulating the influence of natural selection

Return to the window and restore the default settings (200 initial and carrying capacity size). We will now test how natural selection might influence the frequency of the R allele. To minimize the influence of drift in these simulations specify a large population size of 500. When examining natural selection, we usually specify the relative fitness (w) of the different genotypes. By default, a value of 1.0 is assigned to the genotype with the highest fitness, and the remaining genotypes are assigned values relative to the value of 1.0. Let's start by assigning the following relative fitness values: wRR = 1.0 wRr = 0.7 wrr = 0.3

Run the simulation for 25 generations using the relative fitness values above. What happens to the frequency of the R allele?

Next, run the simulation for 250 generations. You should notice that the R allele eventually becomes fixed in each population. At approximately which generation does the R allele become fixed?

In the previous example the homozygous dominant genotype exhibited the highest fitness. What would the simulation look like in a case of **overdominance** (i.e. heterozygote advantage)? Let's change the relative fitness values to the following: wRR = 0.6 wRr = 1.0 wrr = 0.4

Create a new simulation by first resetting the default parameters. Enter the relative fitness values above to model our example of overdominance. Again, play with the population size to determine the influence of drift. Leave all other settings as the defaults. What do

you notice about the frequency of the R allele over the course of 25 generations? What about for 100 generations?

How do your simulations of overdominance compare to your simulations assuming that the homozygous dominant genotype had the highest fitness?

Activity 3: Simulating the influence of mutation

Mutation is the ultimate source of genetic variation in populations. Mutation can create new alleles in populations, which can then be acted upon by other microevolutionary processes such as drift and selection. It's important to note that although mutation can create new alleles, it does not substantially alter allele frequencies because mutation is a relatively infrequent event. Mutation rate is commonly represented by the Greek symbol μ and can be expressed in units such as generations or time (e.g. per million years). For example, the typical mutation rate of vertebrate mitochondrial DNA (mtDNA) ranges between about 0.8 – 2.0 substitutions per site per million years. Site refers to a specific base position in an alignment of homologous DNA sequences. Mutation rates of animal nuclear DNA (nDNA) are often an order of magnitude slower. Interestingly, cnidarian mtDNA evolves at a much slower rate than cnidarian nDNA. Thus, different genomes have the ability to provide different amounts of evolutionary information, and target genes can be tailored to one's specific research objectives. We can use the equation below to calculate the frequency of the R allele after a specific number of generations, assuming we know the starting allele frequency and the mutation rate.

$$(1 - \mu)^t = P_t/P_0$$

Where:

μ = mutation rate from the R allele to the r allele

t = time (e.g. number of generations)

P_t = frequency of the R allele after t generations

P_0 = initial frequency of the R allele

We can rearrange the problem to solve for pt:

$$P_t = P_0 * (1 - \mu)^t$$

Calculate the frequency of the R allele after 250 generations, assuming a mutation rate of R to r is 0.001 per generation and an initial frequency of 0.5.

Let's perform some simulations to observe how mutation alters allele frequencies. Setup a new run by first restoring the defaults (hit refresh on browser). To minimize the influence of drift, set the population size to 500. Enter 0.001 for the mutation rate from R to r and run the simulation for 250 generations. What do the results suggest? Would you say that the frequency of the R allele is changing dramatically? Explain your reasoning.

Next, go back to the settings window and enter a rate of 0.01 for the mutation rate from r to R, and rerun the simulation. How do these results differ from the previous simulation that assumed no mutation from r to R?

Finally, let's try to model a situation where mutation creates a new beneficial allele that is acted upon by natural selection. Theory predicts that the beneficial allele should become fixed in a population if the population size is large enough to minimize the effect of drift. In this example, we will assume that the population was originally monomorphic for the r allele (Initial "R" Allele Prop = 0), and mutation introduces a beneficial R allele. We will also assume that the mutation rate from R to r is the same as the mutation rate from r to R (0.001). We will use the following relative fitness values: wRR = 1.0 wRr = 0.7 wrr = 0.3

What is happening in your model?