

Evolutionary Forces: Generation X Simulation

To launch the GenX software:

1. *Right-click* 'My Computer'.
2. Click 'Map Network Drive'
3. Don't worry about what drive letter is assigned in the upper part of the dialogue box that pops up. In the lower part, enter: <\\hopper\labshare>. You can probably get this without typing by clicking the little triangle on the right of the field.
4. Using Windows Explorer, Click on the Labshare folder, then on Biology 303 folder, then on the GenX icon to start.

A. Recall from lecture:

Evolution is a change over time in allele (gene) frequencies within a population.

There are **4 main evolutionary forces**:

Mutation - new allele arises by physical change in structure of DNA

Genetic drift - random change in allele frequency by chance, important mainly in small populations (remember that variance \uparrow as sample size \downarrow)

Isolation - two populations that exchange members (even just 1 disperser per generation) do not diverge genetically. Isolated populations can diverge due to drift or natural selection

Natural selection - differential survival and reproduction of individuals with different phenotypes

Also recall that a locus is ***fixed*** for a single allele if any allele reaches a frequency of one. When a population holds only one allele, evolution cannot occur (until an alternative allele arrives by mutation or immigration).

B. Background for Generation X simulations:

Generation X is a computer model of evolution. It allows you to alter the four evolutionary forces, either one at a time or in combination, and see the effect on genotype and allele frequencies.

C. Basic Operation of GenX to simulate population genetics.

When you click run, GenX simulates changes in allele frequency through time (evolution) in two populations (one population on the left, the other population on the right). The default settings are to model 25 generations of time, and run 20 iterations of the entire simulation. You can change either of these settings (look at the bottom for the generations and iterations settings; these settings apply to both populations). By running many iterations each time, GenX lets you consider how variable the outcome of evolution is for the sort of situation you are modeling.

1. The main panel has controls that let you set the impact of (a) natural selection, (b) dispersal, (c) mutation, (d) initial genotype frequencies, (e) population size, and (f) number of generations, which is a measure of time. First, make sure all controls are set at their default values:

Natural selection inactive

Dispersal inactive

Mutation of 'A to a' off and 'a to A' off

Population size 500

Genotype frequencies $AA = 0.25$, $Aa = 0.50$, $aa = 0.25$

Generations = 25

If you've changed any numbers, you can quickly restore the defaults: click RENINITIALIZE ALL TO DEFAULT under the OPERATE menu.

2. Under OPERATE, click RUN. The arrow key near the top left does the same thing.

The **top display** shows how genotype frequencies change through time (25 generations), for 20 iterations of the simulation. Note that there is some random variation between runs. This is due to chance events that effect who breeds with whom, and which of an individual's two alleles is passed (by meiosis) into the sperm/egg that happens to create a descendant.

The **middle display** shows how allele frequencies change (it only displays values for the A allele to keep the display simple. The frequency of a would be a mirror image, because the frequency of $a = 1 - A$).

The **bottom display** is a frequency distribution for the final frequency of allele A. The X-axis is the frequency of A at the end of the run. The Y-axis shows how often each final allele frequency occurred.

Watch the top two panels while the simulation runs, because they are not cumulative; they disappear after each iteration. The bottom panel is cumulative, so you can concentrate on it after the 20 iterations are done.

D. Population size and the rate of genetic drift:

A common rule of thumb in conservation biology (Franklin 1980) is that a population size of at least 500 is needed to preserve genetic variability, so that the population can respond to future changes in environmental conditions. Recently, Lande (1996) argued that a population size of 5000 is a more realistic target.

To test these ideas, compare changes over time in two populations that are of different size, but are otherwise the same. How will changes in genotype frequencies through time differ between the two populations? Allele frequencies? How will the final frequency

distribution for allele frequencies differ for the two populations? Does either population actually lose genetic variability? Does the length of time that passes affect the likelihood that an allele will become fixed (that is, hit a frequency of 0 or 1)?

1. Double-click the GenX icon to start the program.
2. In the OPERATE menu, click the REINITIALIZE ALL TO DEFAULT
3. Set population size to 500 for population 1 and to 50 for population 2.
4. Make sure all the other buttons read “Inactive”
5. In the OPERATE menu, click RUN
6. RUN as often as needed to answer the above questions.
7. At the bottom, change GENERATIONS from 25 to 200, and RUN again. You’re now observing the changes over a period 8 times longer than before. Do the answers to any of the above questions change?
8. Now set population 1 to be 50 individuals, and population 2 to be 5 individuals.
9. Reset GENERATIONS to 25 and RUN. As before, the difference in population size is a factor of ten. Do the answers to any of the above questions change?

E. Effects of dispersal (or isolation) on drift within a population and divergence between two populations:

The last simulation showed that alleles often drift to fixation (all *A* or all *a*) by chance, even when no evolutionary force is operating, and that this happens more quickly in small populations. What effect does dispersal between two populations have on this process? Dispersal patterns can be complex. Sometimes individuals move between two populations in both directions, but sometimes movement is one-way, with one population as a ‘source’ and the other as a ‘sink’. Source-sink dynamics occur when one patch of habitat is better (more food, fewer predators, etc.) than nearby patches, allowing better reproduction and more dispersers. Rates of dispersal also vary, so that some populations are linked by dispersal but exchange few individuals, but other populations exchange many individuals.

To investigate how dispersal affects drift within a population and evolutionary divergence between populations, first set up a simulation with all evolutionary forces inactive:

1. Under OPERATE click REINITIALIZE ALL TO DEFAULT.
2. Set Population 1 to 200 individuals.
3. Set Population 2 to 10 individuals.

Use these population sizes for all of the dispersal simulations.

You should know how these two will differ, but if in doubt, RUN the simulation to check. How will the results change if there is dispersal between the populations? Does it matter if dispersal is two-way or one-way? If dispersal is one way, does it matter which population is the source? How does variation in the rate of dispersal change the outcome?

4. Set DISPERSAL ACTIVE for both populations.
5. Set DISPERSAL RATE = 0.01 for both populations. (This means that individuals have a 1 in 100 likelihood of dispersing.)
6. RUN

How does a low rate of two-way dispersal affect changes in allele frequency? Will the results differ if dispersal is one way?

7. Set DISPERSAL INACTIVE for population 1 and DISPERSAL ACTIVE for population 2.
8. RUN

The last simulation was set up for one-way dispersal from a small population to a large population, which is unlikely to occur in nature. Will the results change if we simulate more typical “source-to-sink” dispersal, from a large population to a small, marginal population?

9. Set DISPERSAL ACTIVE for population 1 and DISPERSAL INACTIVE for population 2.
10. RUN

What implications do the results of the dispersal simulations have for conserving genetic diversity in small, endangered populations?

How will the rate of dispersal affect these conclusions? To test your predictions:

11. With population 1 at 200 individuals and population 2 at 10, set DISPERSAL ACTIVE for population 1.
12. Set DISPERSAL RATE = 0.01 for population 1.
13. RUN, and note the results to compare with the next run.
14. Now set DISPERSAL RATE = 0.10 for population 1 (increasing the rate of dispersal from population 1 to population 2 by a factor of 10, with everything else held constant).
15. RUN
16. RUN several times with the DISPERSAL RATE for population 1 at different values between 0.10 and 0.01, to determine how much dispersal is enough to prevent allele frequencies in the small population from drifting.

Before you leave this part of the simulation, you should have a good *quantitative* idea of *how much* dispersal is sufficient to hold drift in check, in populations of varying size.

F. Natural Selection:

The effect that an allele has on an individual's fitness is not directly related to the allele's mode of inheritance (dominant/recessive/additive). An allele that increases fitness can be dominant or recessive. Similarly, an allele that lowers fitness can be dominant or recessive. Despite this, most harmful alleles are recessive. Why is this so?

Selection against a recessive harmful allele:

1. Under OPERATE, click REINITIALIZE ALL TO DEFAULT. This restores both populations to 500 individuals, at Hardy-Weinberg equilibrium (no evolution occurring).
2. Set NATURAL SELECTION to ACTIVE in population 1
3. Set the fitness of the *aa* genotype to 0.90

This simulation models selection against a harmful *recessive* allele that reduces fitness by 10% in population 1. Compare results to population 2, which has equal fitness for all genotypes (no selection).

How do allele frequencies differ between the populations?

Selection against a dominant harmful allele:

1. Under OPERATE, click REINITIALIZE ALL TO DEFAULT.
2. Set NATURAL SELECTION to ACTIVE in population 1.
3. Set the fitness of the *aa* genotype to 0.90. (*AA* and *Aa* remain at 1.0)
4. Set NATURAL SELECTION to ACTIVE in population 2.
5. Set the fitness of the *AA* genotype to 0.90.
6. Set the fitness of the *Aa* genotype to 0.90. (*aa* remains at 1.0)

This simulation compares selection against a harmful *recessive* allele that reduces fitness by 10% (population 1), to selection against an allele that has the same effect on fitness but is *dominant* (population 2).

How do allele frequencies differ between the populations?

Selection against an additive harmful allele:

1. Under OPERATE, click REINITIALIZE ALL TO DEFAULT.
2. Set NATURAL SELECTION to ACTIVE in population 1.
3. Set the fitness of the *aa* genotype to 0.90 (*AA* and *Aa* stay = 1)
4. Set NATURAL SELECTION to ACTIVE in population 2.
5. Set the fitness of the *aa* genotype to 0.90.
6. Set the fitness of the *Aa* genotype to 0.95. (*AA* stays = 1)

This simulation compares selection against a harmful *recessive* allele that reduces fitness by 10% (population 1), to selection against an allele that has the same effect on fitness but is *additive* (population 2). "Additive" means that the magnitude of the effect depends on how many copies of the allele an individual has. "Dosage-dependent" means the same thing.

How do allele frequencies differ between the populations?

Comparing the three selection simulations, why are most harmful alleles recessive?

OPTIONAL MATERIAL: UNDERSTANDING HARDY-WEINBERG EQUILIBRIUM

For example, fur color in foxes is (simplifying slightly) controlled by a single locus, with two alleles, A and a . Individuals with the AA genotype have typical 'red phase' coats. Individuals with the aa genotype have 'silver/black phase' coats. Individuals with the Aa genotype have 'cross phase' coats, which are similar to red phase, but with a blackish stripe down the back and across the shoulders, forming a cross. This is a case of incomplete dominance, because heterozygotes (Aa) have a phenotype intermediate between the two homozygotes (though more similar to AA than to aa). If A was completely dominant, then AA and Aa genotypes would produce the same coat color. If the alleles were completely additive (no dominance), then Aa would give a coat color equally similar to AA and aa .

In Yellowstone, surveys show that the fox population has the following proportions of coat colors:

Red	=	0.62
Cross	=	0.35
Black	=	0.03

So, in a population of 100 foxes with these proportions, there are $62 \times 2 = 124$ copies of the A allele in red foxes (genotype AA) and 35 copies of A in cross foxes (genotype Aa). There are 200 alleles total, so the frequency of the A allele is $124 + 35 / 200 = 0.795$. The frequency of the a allele is calculated the same way: 3 black individuals carry $3 \times 2 = 6$ a alleles, plus 35 a alleles in cross foxes gives 41 a alleles. $41 / 200 = 0.205$. To check that your calculations are correct, remember that the frequencies of A and a must total 1.

If a population is **NOT** evolving, it is said to be in Hardy-Weinberg equilibrium. To test if a population is in H-W equilibrium, you can use the Hardy-Weinberg equation, which uses current allele frequencies to calculate expected genotype frequencies for the next generation. *If the genotype frequencies in the next generation do not match H-W predictions, the population is currently evolving — one or more of the evolutionary forces (natural selection, genetic drift, mutation or dispersal) is altering allele or genotype frequencies.* That's the value of H-W equilibrium... it is simply a way of checking if a population is currently evolving (for a specific locus [gene]).

To use the Hardy Weinberg equation:

p is the frequency of the A allele

q is the frequency of the a allele.

IF the population is in HW equilibrium, the genotype frequencies in the next generation will be:

$$AA = p^2$$
$$Aa = 2pq$$
$$aa = q^2$$

note that $p^2 + 2pq + q^2 = 1$.

In Generation X you can test your understanding of the H-W equations by clicking on the button (bottom right) labeled USE PRACTICE EQUATIONS? so that it reads USING PRACTICE EQUATIONS. I don't recommend doing this during the class period: do it if you have time at the end, or out of class. After you run a simulation, a window will appear with H-W calculations to enter. After you type your answers, click the "Enter Answers" button, click the "Show Formulas" button, and then click the run button. The program will check your answers and show you which are correct. When you are doing these calculations, you must use the number of individuals for **both populations** (1 & 2) to determine the initial genotype frequencies and allele frequencies, or the program will show that your answers are wrong even if you used the correct logic.

Questions to consider:

1. How does the *length of time* over which drift occurs affect the likelihood of an allele drifting to fixation (frequency of 0 or 1)? How does *population size* affect the likelihood of an allele drifting to fixation? How does *dispersal among populations* affect the likelihood of an allele drifting to fixation?
2. When a new allele arises by mutation, it is equally likely to be dominant or recessive. Given that, why are almost all harmful alleles recessive, in genetic surveys of both humans and natural populations?
3. (Hardy-Weinberg) A survey of coat color in 200 foxes gave the following results:

<u>Phenotype</u>	<u>Genotype</u>	<u>Number of foxes</u>
Red	AA	96
Cross	Aa	76
Black	aa	28

What are the current allele frequencies for A and a in this population? What are the current genotype frequencies? What genotype frequencies are expected if the population is in HW equilibrium? Is the population evolving?