

## Class Exercise 1: Dynamic Evolution

BISC 111-01, Spring 2018

For this exercise, we'll use a set of mathematical models to investigate the effects of three key mechanisms of evolution – genetic drift, gene flow, and natural selection – on allele frequencies and phenotype frequencies within populations of organisms.

### Learning Objectives:

- 1) Understand that models are simplified versions of reality, but they are still useful for exploring mechanisms and their interactions.
- 2) Characterize the difference between *variables* and *parameters* in models.
- 3) Explain why the genetic drift model is *random*, or *probabilistic*.
- 4) Explain why the gene flow and natural selection models are *deterministic*.
- 5) Describe phenotype dynamics and allele frequency dynamics in models, and explain how the dynamics are altered by changes in the model *parameters*.

### VARIABLES VS. PARAMETERS

Variables and parameters are both components of equations. They play important different roles in equations and models, but their use is often confused (even by well-established scientists!). *Variables* are the quantities that change within an equation or a biological system. *Parameters* are values that help to define a relationship or equation, but remain constant during a calculation or simulation.

### GENETIC DRIFT MODEL

Genetic drift describes *random* variation in the alleles inherited by the next generation due to the fact that not every single member of the population produces offspring every generation. This is also called a *sampling error* problem, where in this case the *sample* is a smaller subset of the current generation that reproduces, and *error* describes the fact that this can't be perfectly predicted. Because of random variation, each simulation that you run for the genetic drift model is unique, and this is why multiple lines are plotted on the graph – they are showing different dynamics due to random variation with the same set of parameter values and initial conditions.

When a process is *random*, it doesn't mean that *literally anything* can happen, it means that the possible outcomes have a *probability* of happening. In random drift, the probability of the next generation inheriting allele **A** (as opposed to **a**, since **A** and **a** are

the only two possible allele outcomes) depends on the *frequency* of **A** alleles in the current generation. If there is a high frequency of **A** alleles in the current generation, there is a higher probability that there will be **A** alleles in the next generation.

Because we have two possible outcomes – **A** or **a** – we can think about the probability as connected to flipping a coin where you can get heads for **A** or tails for **a**. If the frequency of the **A** allele is 0.5 (it's present in half the population), then there's a 50% chance of passing either **A** or **a** to the next generation and the coin flip analogy works perfectly. But if the frequency of **A** is much larger than that of **a**, let's say **A** is 0.8, that's like flipping an unfair coin where we get heads 80% of the time, and tails only 20% of the time.

The frequency of the **A** allele in the next generation doesn't just depend on the frequency in the current generation; it also depends on the population size. If the current generation is small, there are only a few coin flips, and a few low probability results can have a disproportionate impact on the allele frequencies of the next generation. If the population size is large, there are lots of coin flip opportunities, and the allele frequencies of the next generation are more likely to reflect those of the present generation.

Within the Genetic Drift model on the app, you can simulate the random process of sampling the alleles passed on to the next generation. This works exactly through the coin flip analogy, where the number of coin flips depends on the population size and the probability depends on the allele frequency in the current generation.

## GENE FLOW MODEL

Gene flow describes the movement of reproducing organisms between populations. When populations are connected, organisms can move between populations and reproduce with organisms in the second population.

This gene flow model does not include the effects of genetic drift, so it makes the assumption that *all* members of the population reproduce each generation. Based on the current allele frequency of each island, the allele frequency of the next generation is calculated by the Hardy-Weinberg equation. Once the islands are connected after the number of generations that you select, a fixed fraction of members from the second population (determined by the "migration rate") are added to the other island in each generation. This is a *deterministic* model because it *does not include* any *random variables*; we get the same result for every simulation because all organisms are reproducing every generation and a fixed fraction are migrating each generation.

## NATURAL SELECTION

At this point, we've discussed many examples of natural selection in class. This model of natural selection we'll use here takes a simplified approach by scaling the Hardy-Weinberg equation by three parameters that represent fitness values for the AA, Aa, and aa phenotypes. In the typical form of the Hardy-Weinberg equation we have:

$$1 = A^2 + 2Aa + a^2$$

This diploid natural selection model scales the frequency of the Hardy-Weinberg results for the phenotypes by "fitness factors" that range from zero to one, with one being the highest fitness. The fitness factor for each phenotype (AA, Aa, and aa) is multiplied by its particular piece of the Hardy-Weinberg equation as:

$$\bar{w} = w_{AA}A^2 + w_{Aa}2Aa + w_{aa}a^2$$

where  $\bar{w}$  is the average fitness of the population,  $w_{AA}$  is the fitness factor (between 0 and 1) for the homozygous AA phenotype,  $w_{Aa}$  is the fitness factor for the heterozygous Aa phenotype, and  $w_{aa}$  is the fitness factor for the homozygous aa phenotype. This model does not include genetic drift nor gene flow, so it is also *deterministic*, because there are no random variables.

## ALL MECHANISMS

This simulation model includes all three mechanisms – genetic drift, gene flow, and natural selection – all happening at the same time. You'll use it to investigate questions once you've explored each of the mechanisms individually.

## ACCESSING THE EVOLUTION MECHANISMS APP

- When you're logged into the Wellesley WiFi<sup>1</sup>, open the evolution mechanisms app by clicking: <http://r.wellesley.edu:3838/apps/bisc111/evolution-mechanisms/>
- The app contains four models of evolutionary processes – Genetic Drift, Gene Flow, Natural Selection, and All Mechanisms – that you access by clicking the tabs at the top of the page.

---

<sup>1</sup> If you are using the app off-campus, you'll need to log into the campus VPN first, before opening the app in a web browser. Instructions for using the Wellesley VPN are here: <https://www.wellesley.edu/lts/techsupport/sslvpn>

### Class Exercise Questions

- The answers to the following Class Exercise questions are due on Monday, February 12 by 10pm via Sakai.
- Please answer the following questions in a separate document no longer than three single-spaced pages in length (it may be shorter!).
- Submit a PDF document titled "BISC111\_ClassEx1\_YOURBANNERID.pdf".
- The exercise is worth 25 total points, where each question is worth 2 points. (You get one "free" point for turning in your assignment.)

The Class Exercise questions are broken into different sections based on the four models within the app. In many of the questions, I ask you to describe the *dynamics* of the allele frequency, which means to describe how the allele frequencies are changing through time (*i.e.*, changing through generations), given the conditions that you set with the sliders and the models that are described above.

#### GENETIC DRIFT MODEL

1. Examine the dynamics of the A Allele Frequency when the population size is 100. Keeping the initial value of the Frequency of A allele the same (at 0.5), simulate the A Allele Frequency dynamics for a "founder" population of only 10 organisms that splits from the main population. *How* and *why* are the A allele frequency dynamics different for a population of 10 organisms compared to a population of 100 organisms, given the same initial frequency for A.
2. Move the population size back to 100, keep the Frequency of A allele at 0.5, and increase the number of generations to 200. What happens to the A allele frequencies across different starting populations (the colored lines) if we examine a longer time scale? Are the ending allele frequencies at generation 200 different or similar?
3. Keeping the population size at 100 and number of generations at 200, change the initial Frequency of A allele to 0.9. Describe the differences in the A Allele Frequency dynamics in comparison to your answer from Question 2.

#### GENE FLOW MODEL

4. With the default settings (Island 1 A frequency = 1, Island 2 = 0.2, Connect Populations = 0, Migration rate = 0.1), describe the dynamics of the A Allele Frequency as the generations approach 50.

5. With the same settings as Question 4, change the conditions to connect the populations on the two islands at generation 20. Do the dynamics of the A Allele Frequencies change? Does the end result at generation 50 change?
6. Try a smaller migration rate and a larger migration rate. How does this change the dynamics of the system? Does the end result at generation 50 change?

### NATURAL SELECTION

7. With the default conditions (initial A allele frequency = 0.9, AA fitness = 0.8, Aa fitness = 0.8, aa fitness = 1.0), describe the dynamics of changes in the phenotypes through many generations. What happens at the end of the simulation and why?
8. Leaving the other conditions from Question 7 the same, change the initial value of the A allele frequency to 0.5. Does this change the dynamics and/or the end result of the simulation? Why or why not?

### ALL MECHANISMS

The last four questions deal with scenarios that represent the mechanisms from all three models happening at once in each generation.

*Questions 9 & 10 use the default settings for All Mechanisms.*

9. Island 1 hosts a large population of chickens ( $n = 500$ ) that were released on the island by an ambitious chicken farmer. The initial frequency of the A allele is 0.9. Life on Chicken Island 1 is great for this population, with a slight advantage for chickens with the aa phenotype, chickens with a large beak, compared to chickens with the homozygous AA or heterozygous Aa phenotype, which have small beaks because they can scoop up more insects (chicken food) faster than the chickens with smaller beaks. Chicken Island 1 is connected to a much smaller island, Chicken Island 2, where chickens catch rides with the farmer in their boat, so the two populations mix. For Chicken Island 1 under these conditions, describe the phenotype dynamics across 100 generations.
10. When the farmer releases 500 chickens on Chicken Island 1, they also decide to put 50 chickens (with initial A allele frequency of 0.8) to Chicken Island 2, a smaller island off the coast of Chicken Island 1. On this island, chickens with small beaks (the Aa or Aa phenotypes) have an advantage because insects on this smaller island hide in tiny rock crevices. The populations on Chicken Islands 1

and 2 are connected by boat as described in Question 10. For Chicken Island 2 under these conditions, describe the phenotype dynamics across 100 generations.

11. On Chicken Island 2, a disease breaks out that reduces the fitness of the homozygous AA phenotype to 0.2. Leaving all other conditions the same, describe the phenotype frequency dynamics on Island 2 across 100 generations with these new conditions.
12. Do your own simulation (you can pick an organism other than chickens)! Think up a scenario for Island 1 and Island 2, describe what parameters suit your scenario, and describe what happens to the long-term phenotype and/or allele frequency dynamics for your scenario.