

Quantitative Genetics Simulation Demo

HORT / AGRO 4401

March 10, 2017

Introduction

The complexity of quantitative genetics often requires intensive computation tools for analysis. Researchers can also use computers to run simulations of breeding programs. These can be especially informative for hypothesis testing or for decision-making in breeding. We will use simulations to better-understand the concepts of population genetics, genetic variance, and response to selection.

Setup Instructions

If you have not yet installed the **shiny** package, do so using the following command:

```
install.packages("shiny")
```

Initiate the Simulation

1. First, load the **shiny** package by running the following command:

```
library(shiny)
```

2. Next, start the simulation by running this command:

```
runGitHub("neyhartj/qgdemo")
```

3. A window will pop up in your internet browser (or in a preview window if you are using R Studio).
You are ready to go!

Please read the instructions provided in the introduction, and then use the following questions to guide your interaction with the simulation.

Randomly-Mating Populations

Breeding programs rely on **breeding populations**. We learned in our population genetics section that breeding populations are defined by **genotype frequencies** and **allele frequencies**. We also learned that these frequencies can be predicted if populations are in Hardy-Weinberg Equilibrium.

For the first simulation, we will look at the change in allele frequency at a single gene in a randomly-mating breeding population.

Problem 1

Without changing parameters, click the “Simulate” button to run an iteration of the simulation. You will notice that the allele frequency is changing over generations. Are the populations in Hardy-Weinberg Equilibrium? Inspect the different parameters that you can change (i.e. allele frequency, population size, number of generations). What is the most obvious violation of the assumptions of Hardy-Weinberg Equilibrium?

Problem 2

Even though our populations are randomly-mating, allele frequencies can change over time simply by random chance. This is called **genetic drift**. Now, increase the breeding population size (try something large like **1,000** or **1,500** individuals) and re-run the simulation. What can you conclude about the effect of population size on the level of genetic drift (i.e. the random changes in allele frequency)?

Problem 3

Reset the population size to around **100** individuals and re-run the simulation. In any of the populations (defined in different colors), did the alleles reach **fixation** (i.e. $p = 0$ or 1)? You may re-run the simulation until this occurs.

Variation at a gene is lost when alleles reach fixation. This can happen simply as a result of **genetic drift**. Now, with the population size at around **100** individuals, change the starting allele frequency to be close to 0 (i.e. < 0.1) or close to 1 (i.e. > 0.9). Re-run the simulation. What can you conclude about the effect of starting allele frequency on the likelihood that an allele reaches fixation?

Problem 4

Keep the starting allele frequency close to 0 or 1. Adjust the population size to a large number of individuals. What can you conclude about the impact of starting allele frequency and population size on the likelihood that an allele reaches fixation?

Problem 5

Allele frequencies can change due to random chance, even in the absence of selection (i.e. from a breeder). Speculate on the importance to breeding of losing variation at a gene due to allele fixation. **Hint:** imagine this gene contributed to an important plant trait, such as yield.

Genetic Variance

Move on to the Genetic Variance tab. This simulation will assess the impact of **allele frequency** and **coded genotypic values** on the level of genetic variance at a single gene.

Problem 1

Adjust the starting allele frequency (the simulation will automatically update). What is the effect of starting allele frequency on the ratio of additive variance (V_A) to dominance variance (V_D)? What happens when the allele is **fixed** (i.e. $p = 0$ or 1)?

Problem 2

Set the starting allele frequency back to around **0.5**. Adjust the additive effect (a) and level of dominance (d). What can you conclude about impact of the ratio of a to d on the ratio of V_A to V_D ?

Problem 3

Set a to **0** and d to a value larger than 0 (say, **5**). Play with the starting allele frequency. Is it possible to have $V_A > 0$ if a is equal to 0? Why is this?

Response to Selection

Move on to the Response to Selection tab. The objective of a breeding program is to make selections to develop superior cultivars (or varieties) of plants. When breeding for a quantitative trait, such as yield, plant breeders aim to increase the **mean** of the population. The change in the population **mean** for a trait from one generation to the next is called the **response to selection**.

In this simulation, we will explore how different parameters impact the **response to selection** and other measurements for a trait influenced by **many genes**.

Without changing parameters, click the “Simulate” button to run an iteration of the simulation. The top graph shows the response to selection curve over generations. The lower-left graph shows the change in **genetic variance** across all of the genes impacting the trait. The lower-right graph shows the change in **allele frequency** for the value of p at all of the genes influencing the trait.

Problem 1

In this simulation, p measures the frequency of the allele with **positive additive effect** a . Change the starting allele frequency from **0.5** to close to 1 (i.e. > 0.9) and re-run the simulation. What do you notice about the steepness of the **response to selection** curve and the **genetic variance** curve?

Problem 2

Change the starting allele frequency to close to 0 (i.e. < 0.2) and re-run the simulation. What do you notice about the steepness of the **response to selection** curve and the **genetic variance** curve?

Problem 3

Inspect the change in **allele frequency**. Each line tracks the value of p at a different gene. Thicker lines indicate alleles with larger value of a and thinner lines indicate alleles with smaller values of a . You will notice that some alleles reach fixation. What can you conclude about the relationship between the value of a and the number of generations required for the allele to reach **fixation**?

Problem 4

Reset the simulation by refreshing the browser. Play with the value of **heritability** (h^2). What do you notice about the effect of heritability on the steepness of the **response to selection** curve?

Problem 5

Adjust the heritability to **0.5** and play with the value of the **selection intensity**. What do you notice about the effect of selection intensity on the steepness of the **response to selection** curve?

Problem 6

Return the selection intensity to 0.1 and increase the **number of breeding generations**. What do you notice about the trend in **response to selection**? What is the relationship between the **response to selection** and the **genetic variance**? Speculate on the reason for this relationship. **Hint:** observe the change in allele frequencies for each gene.

Optional Problem

Continue to adjust parameters and re-run simulations. What other observations and conclusions can you draw about the relationship between these parameters and the **response to selection**?