Exploring Genetic Drift via Manual Simulations

KEITH W. PECOR

ABSTRACT

Genetic drift is an important mechanism in microevolution, but it can be more challenging to understand than other mechanisms (e.g., natural selection). This group project allows students to simulate random changes in allelic frequencies over generational time using a few simple supplies and was well received when included in an introductory biology course at the collegiate level.

Key Words: Genetic drift; microevolution; active learning; in-class project.

Introduction

A search for “genetic drift” within the Next Generation Science Standards (NGSS Lead States, 2013) returns performance expectations ranging from elementary school through high school. In my own teaching at the undergraduate level, I have covered this topic in both introductory and intermediate courses. Genetic drift, or random fluctuations in allelic frequencies at a given locus, is commonly presented as one of the four major agents of microevolution, along with selection, migration (gene flow), and mutation. However, genetic drift is more abstract and less intuitive than the other three processes, which can make it a difficult concept to grasp for some students.

At its essence, genetic drift is an example of sampling error, which simply refers to the fact that a sample (e.g., alleles and genotypes in generation X + 1) may not mirror the population from which the sample was taken (e.g., alleles and genotypes in generation X). Generation-to-generation changes can result from matings that do not replicate the allelic frequencies of generation X in generation X + 1. In some instances, this is due in part to random survival of individuals in generation X, whereas other instances are influenced by the fact that not all combinations of alleles are realized (e.g., a pair of heterozygotes produce a homozygous offspring). If the alleles in question have no effect on fitness, then drift can lead to the fixation of one allele (and loss of all others) in small or medium-sized populations. A classic example is the study by Buri (1956), in which the frequencies of the two alleles for eye color were monitored in laboratory populations of fruit flies for 19 generations. All populations started at 50:50 for the two alleles, but one or the other allele became fixed in a number of populations over the course of the study.

A common presentation of genetic drift in collegiate textbooks (e.g., Freeman & Herron, 2007; Morris et al., 2012) is to show graphs of allelic frequencies over generational time from computer simulations that were coded to model random processes in populations of different sizes. There may also be some discussion of modeling random processes via the analogy of the “drunkard’s walk” (Futuyma, 2005). I have used these images and analogies in my courses but have never been completely satisfied with their efficacy. In searching for improvements, I found a variety of resources, both online (“Dueling Alleles” from Helms et al., 1997; “Genetic Drift Simulation” from University of Arizona, 1999) and in the pages of The American Biology Teacher (Western, 1976; Hammersmith & Mertens, 1990; Maret & Rissing, 1998; Brewer & Zabinski, 1999; Kephart et al., 2002; Young & Young, 2003; Robischon, 2015). These many projects approach drift from a variety of perspectives. Some are based on imaginary populations, whereas others draw on empirical studies. Some are focused on random survival, whereas others consider founder effects. However, none of these projects explicitly address the mechanism that the computer simulations in textbooks were meant to mimic. The project by Hammersmith and Mertens (1990) was focused on
mechanism, but a random numbers table was used to determine allelic frequencies rather than considering individual matings. Thus, I developed the following classroom activity that can be used in middle school, high school, and postsecondary settings.

**Project Description**

In the project, students conduct manual simulations using hypothetical organisms and alleles. In these simulations, there is a species of interest in which the population in each generation is four individuals and generations do not overlap. There is a single locus of interest, with two alleles: A₁ and A₂. Each group of four students is provided with a data sheet (Figure 1), a six-sided die, a coin, and instructions explaining the following parameters:

1. Each group member will represent a single member of the population (Individual 1, 2, 3, or 4) and will retain their assigned identity every generation.

2. The simulation will start with every individual being heterozygous (A₁ A₂).

3. In each generation, there will be four matings to yield four offspring for the next generation. The matings will be determined via four rolls of a standard six-sided die, where the result of each roll specifies a mating pair using the scheme shown in Table 1. Because this is a random process, it is possible that there will be more than one mating of a single type in a given generation.

4. Homozygous individuals will only have a single allele to pass on, whereas heterozygous individuals will need to determine the allele passed on in the gamete following meiosis via a coin flip: heads = A₁ and tails = A₂. Given that all individuals start as heterozygous in generation 1, the first iteration will require two coin flips for each of the four matings. In subsequent generations, matings may require two, one, or zero coin flips, depending on the genotypes of the individuals.

5. Starting in generation 2, the genotype of an individual is determined by the outcome of the corresponding mating from the previous generation. For example, individual 1 in generation 2 is the offspring of mating 1 from generation 1.

6. The matings continue until (a) one allele is fixed and the other is lost, (b) n generations have elapsed, or (c) a set time has elapsed. In an introductory undergraduate course for biology majors, I found that it took ~45 minutes for all groups to complete the project, with allele loss/fixation or 20 generations elapsed as the end point.

7. Upon completion, groups enter their allele frequencies for either A₁ or A₂ on a spreadsheet, from which a graph of allelic frequency over generational time among the simulations (Figure 2) can be produced.

**Reflections, Revisions & Extensions**

Our simulations were done in groups of four students, and we noted the following points. First, doing the first generation’s matings class-wide would improve efficiency. A number of students were uncertain and raised questions about the first iteration of the simulation but were confident for all subsequent iterations. Because of the random nature of genetic drift, the duration of a given simulation is unpredictable. We had one group with an allele

---

<table>
<thead>
<tr>
<th>Individual</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A₁ A₂</td>
<td>A₂ A₂</td>
<td>A₁ A₂</td>
<td>A₂ A₂</td>
<td>A₁ A₂</td>
<td>A₂ A₂</td>
<td>A₁ A₂</td>
<td>A₂ A₂</td>
<td>A₁ A₂</td>
<td>A₂ A₂</td>
</tr>
<tr>
<td>2</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
</tr>
<tr>
<td>3</td>
<td>A₁ A₂</td>
<td>A₂ A₂</td>
<td>A₂ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
</tr>
<tr>
<td>4</td>
<td>A₁ A₂</td>
<td>A₂ A₂</td>
<td>A₂ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
<td>A₁ A₂</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Matings</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3,4</td>
<td>2,3</td>
<td>2,3</td>
<td>2,4</td>
<td>3,4</td>
<td>2,3</td>
<td>1,4</td>
<td>2,4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2,3</td>
<td>2,4</td>
<td>1,2</td>
<td>1,2</td>
<td>1,3</td>
<td>1,2</td>
<td>1,4</td>
<td>2,3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3,4</td>
<td>1,2</td>
<td>1,4</td>
<td>2,3</td>
<td>1,2</td>
<td>2,4</td>
<td>3,4</td>
<td>1,4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1,4</td>
<td>1,3</td>
<td>3,4</td>
<td>3,4</td>
<td>1,4</td>
<td>1,3</td>
<td>3,4</td>
<td>1,4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>0.5</td>
<td>0.25</td>
<td>0.25</td>
<td>0.375</td>
<td>0.625</td>
<td>0.625</td>
<td>0.625</td>
<td>0.75</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>A₂</td>
<td>0.5</td>
<td>0.75</td>
<td>0.75</td>
<td>0.625</td>
<td>0.375</td>
<td>0.375</td>
<td>0.375</td>
<td>0.25</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

**Figure 1.** A sample data sheet. The items in bold are entered in the sheet prior to distribution to students. This sheet shows the results from one group simulation.
that went to fixation after six generations and another that lasted 16 generations (Figure 2). This degree of variation presented a challenge to keep all students engaged. After field testing the project, I recommend having additional data sheets (Figure 1) on hand and having groups that finish early engage in additional simulations to add to the data set.

After completion of the project, there are a number of ways to proceed. Before a class-wide discussion of the results from all groups, students could be asked to graph their group’s results and answer questions about their findings. For example, how would increasing the sample size impact the magnitude of generation-to-generation changes and/or the time to fixation of one allele? How would a more diverse gene pool (i.e., more than two alleles) impact the results? (Etc.)

In conclusion, this project – a relatively simple undertaking with modest supplies – helps illustrate genetic drift due to sampling error in a concrete, hands-on way. Completing this project (alone or in combination with extensions such as those described above) before a discussion of computer simulations of populations of various sizes would allow students to better understand both the mechanism of drift and the influence of population size on its impacts on populations.

Acknowledgments
I thank K. T. Elliott for helping test this project and providing feedback on it. M. A. Wund and anonymous reviewers provided feedback on the manuscript.

References