**BIO 101 Lab 10: Population Genetics**

Notification: If you have a disability that makes it difficult to complete this lab, please contact your instructor. Please provide your instructor a copy of the Memorandum of Accommodation (MOA) from NVCC Disability Support Services.

**Objectives:**

* Describe the Hardy-Weinberg equilibrium with respect to genotype and allele frequencies: ( p + q )2 = p2 + 2pq + q2 = 1
* Explain the conditions necessary to maintain the Hardy-Weinberg equilibrium
* Describe how the bottleneck effect and gene flow can alter the gene pool of a population

**Background:**

Population Genetics

Genetics can be applied on a small scale, as in the case of Gregor Mendel’s investigations of garden pea plants, but it can also be applied to whole populations. A **population** is a group of organisms in a specific area that have the opportunity to interbreed. These organisms share a common **gene pool**, all the alleles present in all individuals in a set population. The basic unit of evolution is the population, and small scale changes to the allele frequencies to a population’s gene pool is called **microevolution**. In short, populations can evolve, but individuals do not.

Two scientists, G. H. Hardy and W. Weinberg developed models of population genetics that showed gene pools did not evolve, or change allele frequencies, solely based on probability inherent during the process of heredity. The **Hardy-Weinberg** **Principle** states that allelic frequencies of an ideal population will not change from generation to generation. This principle holds true only under certain criteria:

1. The population is **large in number**
2. There is **random mating** between individuals in the population
3. There are **no mutations** that change one allele into a different allele
4. There is **no migration** of individuals into or out of the population
5. Individuals with each genotype are equally likely to survive and mate (**no natural selection**)

The Hardy-Weinberg Principle provides a testable framework to study the effects migration (gene flow), mutations, non-random mating, natural selection, population size etc. have on the evolution of populations. A major function of population genetics is to determine the allele and genotype frequencies in populations over generations. The Hardy-Weinberg Principle provides a mathematical formula to calculate expected allele and genotype frequencies in a population that is **not** undergoing microevolution.

The Hardy-Weinberg Equation is: (**p** + **q**)2 = **p2 + 2pq + q2 = 1**

|  |  |  |
| --- | --- | --- |
| **p2** ≡ Frequency of RR | **2pq** ≡ Frequency of Rr | **q2** ≡ Frequency of rr |

The Hardy-Weinberg equation is based on focusing on one trait in a population at a time, with two alleles contributing to the genotype for that trait. The proportion (frequency) of dominant alleles in the gene pool of the population is given the variable name **p** and the frequency of recessive alleles in the gene pool of the population is given the variable name **q**. Since **p** and **q** together represent all the alleles in the gene pool of the population, adding **p** and **q** will total 100%, or **p + q = 1**.

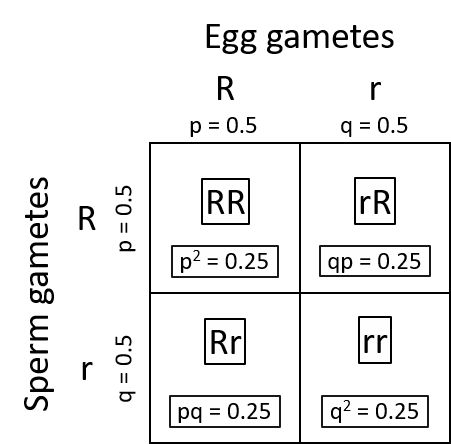
Assuming that the organisms in the population sexually reproduce and randomly mate with each other, alleles in the gene pool of the population will be randomly combined to result in the genotypes of the next generation of the population. In this case, if the dominant allele in the population is **R** and the recessive allele is **r**, the probability of having the genotype **RR** is equal to **p2**.

For example if the frequency of **R** in the gene pool is **p = 0.5** and the frequency of **r** in the gene pool is **q = 0.5**, then the probability of **RR** (homozygous dominant) in the next generation is **p** \* **p** = **p2** = **0.25**. Similarly, the probability of **rr** (homozygous recessive) is **q** \* **q** = **q2** = **0.25**. The probability of being **Rr** (heterozygous) is ( **p** \* **q** ) + ( **q** \* **p** ) = **2pq** = **0.5**. The **2** in the **2pq** term accounts for there being two ways an individual could be heterozygous, since either parent can provide the dominant or recessive allele.

**Example 1:**

* If the **R** (**p**) and **r** (**q**) alleles occur at equal frequencies in the initial population, then frequencies of alleles is **p = q = 0.5**
  + This means that half of all gametes in the gene pool will have the **R** allele and the other half of the gametes will have the **r** allele
* Again, assuming the population satisfies the Hardy-Weinberg Principles, the genotypic makeup of the next generation of the population can be determined using the Hardy-Weinberg equation:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ( **p** | + | **q** ) 2 | = | **p2** | + | **2pq** | + | **q2** | = | 1 |
|  |  |  |  |  |  |  |  |  |  |  |
| ( 0.5 | + | 0.5 ) 2 | = | 0.25 | + | 0.50 | + | 0.25 | = | 1 |
|  |  |  |  |  |  |  |  |  |  |  |
| Frequency of **R** allele |  | Frequency of **r** allele |  | Frequency of **RR** genotype |  | Frequency of **Rr** genotype |  | Frequency of **rr** genotype |  |  |



* If a population satisfies the Hardy-Weinberg Principle conditions, then the allele frequencies will not change from generation to generation. They will remain **p = q = 0.5**

**Example 2:**

* In nature the frequencies of dominant and recessive alleles are almost never equal. For example, in Mendel’s garden 4% of the pea plant flowers were white (a recessive trait). The genotype of those flowers would be homozygous recessive (**rr**). The frequency of that genotype in this population (if it satisfies the Hardy-Weinberg Principles) is **q2**. This means that the frequency of the recessive white allele (**q**) can be calculated by taking the square root of **q2**.

Therefore, if **q2** = 0.04, then **q** = √(0.04) = 0.2 (the frequency of the recessive allele).

* Since **p + q = 1**, if the frequency of one of the alleles in a population is known, the frequency of the other allele can be calculated.

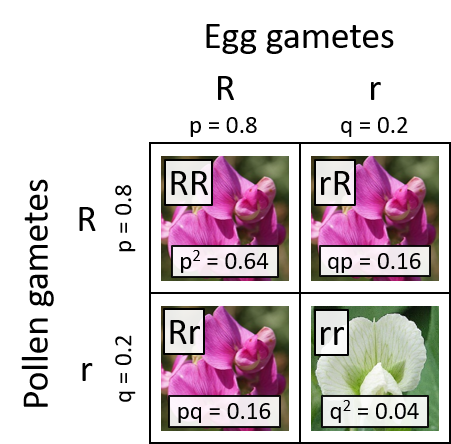
Therefore, if **q** = 0.2, then **p** + (0.2) = 1, so **p** = 0.8.

This means that 20% (0.2/1) of the alleles in the population are recessive and

80% (0.8/1) of the alleles are dominant.

* Again, assuming the population satisfies the Hardy-Weinberg Principles, the genotypic makeup of the next generation of the population can be determined using the Hardy-Weinberg equation:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ( **p** | + | **q** ) 2 | = | **p2** | + | **2pq** | + | **q2** | = | 1 |
|  |  |  |  |  |  |  |  |  |  |  |
| ( 0.8 | + | 0.2 ) 2 | = | 0.64 | + | 0.32 | + | 0.04 | = | 1 |
|  |  |  |  |  |  |  |  |  |  |  |
| Frequency of **R** allele |  | Frequency of **r** allele |  | Frequency of **RR** genotype |  | Frequency of **Rr** genotype |  | Frequency of **rr** genotype |  |  |
|  |  |  |  |  |  |  |  |  |  |  |



* Keep in mind that this method is only valid if all of the Hardy-Weinberg Principles are satisfied in a population.

If a population is in Hardy-Weinberg equilibrium then the allelic frequencies will not change from generation to generation, and the equation p2 + 2pq + q2 = 1 can be used to determine the expected genotype frequencies for future generations of the population. However, in nature the conditions for Hardy-Weinberg equilibrium are rarely met. In this lab exercise you will use the Hardy-Weinberg Principle as a model to predict the changes in a population’s gene pool based on events, such as genetic drift and gene flow.

**Materials:**

|  |  |
| --- | --- |
| * Internet | * Calculator |

**Safety:**

Follow all standard laboratory safety procedures.

**Procedure:**

**Population Genetics Activity 1: Testing Hardy-Weinberg Equilibrium**

This activity simulates how allele frequencies in a population will remain similar from generation to generation if the population satisfies the Hardy-Weinberg Equilibrium conditions.

1. In the laboratory, each group would obtain a plastic or paper bag that will hold the two colors of beans: black and brown (or red and white).
2. Each group would place 32 black (or red) beans and 48 brown (or white) beans into their group’s bag. The black (or red) beans represent the dominant allele **R** and the brown (or white) beans represent the recessive allele **r**.

Since the population has 32 + 48 = 80 alleles, this means the population will contain 40 individuals, with each individual represented by two beans. These two beans will represent the genotype of the individual.

1. Based on the Background information and Procedures described above, answer questions 1 – 2 in the Lab Worksheet.
2. Table 1 in the Lab Worksheet is partially completed for you. Fill in the rest of Table 1 by calculating the expected population genotype frequencies and number of individuals of each genotype by using the Hardy-Weinberg equation: p2 + 2pq + q2 = 1. This is the hypothesis (and prediction) that is made based on the assumption that the population satisfies the Hardy-Weinberg Equilibrium conditions.
   * Using the Hardy-Weinberg equation: **p2** + **2pq** + **q2** = 1, calculate the expected frequencies for the genotypes **RR**, **Rr**, and **rr** by using the allele frequencies for the parent population, where **p** is the frequency of allele **R** = p and **q** is the frequency of allele **r.**
   * Using the genotype frequencies, calculate the number of individuals with each genotype by multiplying the genotype frequency by the total population size.

For example: the allele frequency of the R allele is 0.4, so therefore p = 0.4. Since p = 0.4, then p2 = (0.4)2 = 0.16. This is the expected population genotype frequency of the RR genotype. Since there are 40 individuals in the population, the expected population number of RR individuals is 0.16 \* 40 = 6.4.

1. After completing Table 1, answer question 3 in the Lab Worksheet. This is where you will state your hypothesis and prediction.
2. Watch the following video that provides background on the Hardy-Weinberg Equation and the assumptions of the Hardy-Weinberg principle, as well as showing how beads representing alleles in a gene pool can be randomly chosen to simulate what the allele and genotype frequencies would be in the next generation of a population from time 1:30 to time 6:40.

<https://www.youtube.com/watch?v=8wyXM-URQSU>

Note that in the video above, red and yellow beads are used to represent the dominant allele “A” and the recessive allele “a” respectively, while in our lab experiments, we instead use black (or red) and brown (or white) beans to represent the alleles “R” and “r”, respectively. However, the process of randomly selecting the pairs of beads and keeping track of the genotypes of each pair selected is the same as what would occur in our lab experiment.

1. In the laboratory, without looking into the bag one student would remove two beans from the bag. These beans represent the genotype of an individual in the second generation. On a scrap sheet of paper keep a tally/count of the new generation’s genotypes (RR, Rr, and rr). This data would then be used to complete Table 2.
2. In the laboratory, once the genotype for an individual in the population is observed and counted, the beans would be placed back into the bag and step 7 would be repeated until a total of 40 individuals had been counted. This technique is called **sampling with replacement**.
3. Table 2 in the Lab Worksheet is partially completed for you. Fill in the rest of Table 2 by calculating the observed population genotype frequencies and allele frequencies by using the observed number of individuals with each genotype that are already filled in for Table 2.

For example: to calculate an observed population genotype frequency, if the number of RR individuals is 7 and the total population size is 40, then the genotype frequency would be 7/40 = 0.175 (which can be rounded to 0.18 in our experiment).

For example: to calculate an observed population allele frequency, if the number of Rr individuals is 18, then 18 “R” alleles will be contributed to the gene pool by these individuals, and if the number of RR individuals is 7, and each RR individual contributes 2 “R” alleles to the gene pool, that means that there are 7 \* 2 = 14 “R” alleles contributed by the RR individuals. Therefore, the total number of “R” alleles in the gene pool will be 18 + 14 = 34. Since there are 80 alleles in the gene pool in total, the allele frequency for “R” would be 34/80 = 0.425 (rounded to 0.43 in our experiment)

1. After completing Table 2, answer questions 4 – 6 in Lab Worksheet.

**Population Genetics Activity 2: Observing the effects of Genetic Drift (due to the Bottleneck Effect)**

This activity simulates how drastic decreases in the number of individuals in a population can result in changes in allele frequencies. These types of population decreases are known as **bottlenecks** and the changes in allele frequencies occur due to random chance as a result of **Genetic Drift** (a form of which is the **Bottleneck Effect**).

1. In the laboratory, each group would obtain a plastic or paper bag that will hold the two colors of beans: black and brown (or red and white).
2. Each group would place 40 black (or red) beans and 40 brown (or white) beans into their group’s bag. The black (or red) beans represent the dominant allele R and the brown (or white) beans represent the recessive allele r.

Since the population has 40 + 40 = 80 alleles, this means the population will contain 40 individuals, with each individual represented by two beans. These two beans will represent the genotype of the individual.

1. Without looking into the bag one student would then remove two beans from the bag. These beans represent the genotype of an individual that survived the catastrophe that substantially decreased the population size. On a scrap sheet of paper you would record the genotype (**RR**, **Rr**, and **rr**) of this individual. This data will be used for Generation 1 in Table 3. **Do not place these beans back into the bag**.
2. Step 3 would be repeated until 5 total individuals (5 pairs of beans) had been counted. Since you did not put the beans back into the bag, this technique is called **sampling without replacement.**
3. This data would be used to complete the row for Generation 1 in Table 3 by determining the genotype frequencies (**RR**, **Rr**, and **rr**) and the **new** allele frequencies for **R** (**p**) and **r** (**q**) – *your observed frequencies*. This is the gene pool of the surviving population.
   1. Count the number of **RR** individuals and divide by 5 to get the genotype frequency for **RR**. Repeat this process for the **Rr** individuals and the **rr** individuals to obtain the frequencies for the **Rr** and **rr** genotypes, respectively.
   2. Count the number of **R** alleles and divide by 10 to get the allele frequency for **R**, (**p**) then count the number of **r** alleles and divide by 10 to get the allele frequency for **r** (**q**).
4. The surviving individuals would now reproduce by randomly mating with each other, resulting in the next generation of organisms in the population.
   1. To simulate this, you would set all the beans aside and determine how many black (or red) and brown (or white) beans to put into the next generation’s gene pool by using the allele frequencies for **R** (**p**) and **r** (**q**) determined in step 3.

*Example: If p = 0.40 and q = 0.60 then your new population will have 4 black (or red) beans and 6 brown (or white) beans.*

Note that the allele frequencies for each generation may be different from the allele frequencies of the previous generation.

1. Again, in the laboratory, without looking into the bag, one student will remove two beans from the bag. These beans represent the genotype of an individual that survived the catastrophe that substantially decreased the population size. On a scrap sheet of paper you would record the genotype (**RR**, **Rr**, and **rr**) of this individual. This data will be used to complete Table 3. **Place these beans back into the bag**
2. Once the genotype of the individual had been observed and counted, the beans would be placed back into the bag, you would shake the bag, and you would repeat step 7 until you have counted 5 individuals. This technique is called **sampling with replacement**.
3. This data would be used to complete the rows for Generations 2-10 in Table 3 by determining the genotype frequencies (**RR**, **Rr**, and **rr**) and the **new** allele frequencies for **R** (**p**) and **r** (**q**) – *your observed frequencies* for each Generation. This is the gene pool of the surviving population.
   1. Count the number of **RR** individuals and divide by 5 to get the genotype frequency for **RR**. Repeat this process for the **Rr** individuals and the **rr** individuals to obtain the frequencies for the **Rr** and **rr** genotypes, respectively.
   2. Count the number of **R** alleles and divide by 10 to get the allele frequency for **R** (**p**), then count the number of **r** alleles and divide by 10 to get the allele frequency for **r** (**q**).
4. The surviving individuals now reproduce by randomly mating with each other, resulting in the next generation of organisms in the population.
   1. Set all the beans aside and determine how many black (or red) and brown (or white) beans to put into the next generation’s gene pool by using the allele frequencies for **R** (**p**) and **r** (**q**) determined in step 7.

Note that the allele frequencies for each generation may be different from the allele frequencies of the previous generation.

1. In the laboratory, you would repeat steps 7 – 10 for up to 10 generations. If at some point either the frequency of the **R** (**p**) allele or the frequency of the **r** (**q**) allele reaches 100% (or 1.0), you can stop. An allele reaching a frequency of 100% (or 1.0) in a population is called **fixation** and the population has become completely genetically uniform for that particular gene and trait as a result.
2. Table 3 in the Lab Worksheet has been completed for you.
3. Watch the following video that provides background on genetic drift and shows the results of the bottleneck effect on a population from time 0:00 to time 4:20.

<https://www.youtube.com/watch?v=W0TM4LQmoZY>

1. Answer question 7 in the Lab Worksheet by completing Table 4.
2. After completing Table 4, answer question 8 in the Lab Worksheet
3. Using the data in Table 3, answer questions 9 – 11 in the Lab Worksheet.
4. Based on the Background information and the videos you watched, answer questions 12 – 13 in the Mendelian and Population Genetics Lab Worksheet.

**BIO 101 Lab 11: Population Genetics Worksheet**

**Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Section: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Data Analysis and Synthesis Questions**

Population Genetics: Activity 1

1. Each bean represents a single allele. If the organisms in the population are diploid (they have two alleles for each gene), how many individuals are in the population?
2. Since each bean represents a single allele:
   1. What allele does the black (or red) bean represent?

* 1. What allele does the brown (or white) bean represent?
  2. What color combination of two beans represents a:
     1. homozygous dominant individual?
     2. Homozygous recessive individual?
     3. Heterozygous individual?

**Table 1: Expected Genotype and Allele Frequencies for the Second Generation**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Parent Population** | | | **Expected Population** | | | | | | |
| **Allele Frequency** | | | **Allele Frequency** | | **Genotype Frequency** | | | | |
| **R** | **r** | | **R** | **r** | **RR** | | **Rr** | | **rr** |
| 0.4 | 0.6 | | 0.4 | 0.6 | **0.16** | |  | |  |
|  |  | |  |  |  | |  | |  |
|  | |  |  | | **Number of individuals with each genotype** | | | | |
|  | |  |  | | **RR** | **Rr** | | **rr** | |
|  | |  |  | | **6.4** |  | |  | |

1. Based on the Hardy-Weinberg principle state your prediction for genotype frequencies of the future generations

**Table 2: Observed Genotype and Allele Frequencies for the Second Generation**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Parent Population** | | | **Observed Population** | | | | | | |
| **Allele Frequency** | | | **Allele Frequency** | | **Genotype Frequency** | | | | |
| **R** | **r** | | **R** | **r** | **RR** | | **Rr** | | **rr** |
| 0.4 | 0.6 | | **0.43** |  | **0.18** | |  | |  |
|  |  | |  |  |  | |  | |  |
|  | |  |  | | **Number of individuals with each genotype** | | | | |
|  | |  |  | | **RR** | **Rr** | | **rr** | |
|  | |  |  | | **7** | **18** | | **15** | |

1. Were the expected and observed results similar? Do your results match the prediction for a population in Hardy-Weinberg equilibrium? If not suggest an explanation.
2. What do you expect to happen to the genotype and allele frequencies of the population after running this simulation for 20 generations?
3. Explain how this simulation meets the conditions for a population in Hardy-Weinberg equilibrium.

**Table 3: Observed Genotype and Allele Frequencies: Bottleneck Effect**

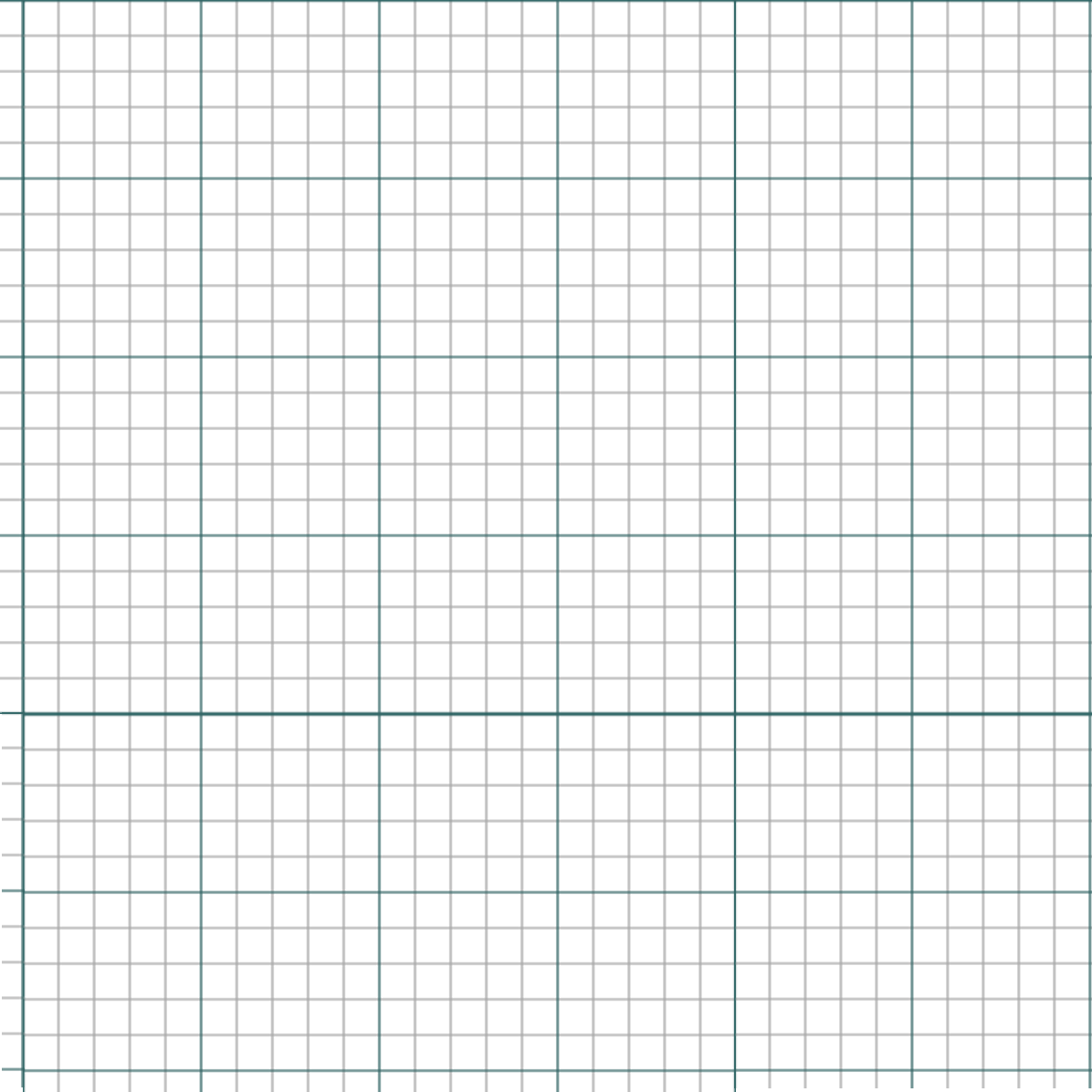
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Generation** | **Genotype Frequency Observed** | | | **Allele Frequency Observed** | | **Population Size** |
| **RR** | **Rr** | **rr** | **R (p)** | **r (q)** |
| **0** | 0.25 | 0.5 | 0.25 | 0.5 | 0.5 | 40 |
| **Bottleneck** | | | | | | |
| **1** | **0.2** | **0.2** | **0.6** | **0.3** | **0.7** | 5 |
| **2** | **0** | **0.6** | **0.4** | **0.3** | **0.7** | 5 |
| **3** | **0.2** | **0.6** | **0.2** | **0.5** | **0.5** | 5 |
| **4** | **0.4** | **0.4** | **0.2** | **0.6** | **0.4** | 5 |
| **5** | **0.6** | **0.4** | **0** | **0.8** | **0.2** | 5 |
| **6** | **0.4** | **0.4** | **0.2** | **0.6** | **0.4** | 5 |
| **7** | **0.6** | **0.2** | **0.2** | **0.7** | **0.3** | 5 |
| **8** | **0.6** | **0.4** | **0** | **0.8** | **0.2** | 5 |
| **9** | **0.6** | **0.2** | **0.2** | **0.7** | **0.3** | 5 |
| **10** | **0.8** | **0.2** | **0** | **0.9** | **0.1** | 5 |

1. The allele frequencies **p** and **q** for Generation 1 from Table 3 have been copied into Table 4 below. Use these allele frequencies to calculate the expected genotype frequencies (**p2**, **2pq**, **q2**) for Generation 1 and enter the results into Table 4.

**Table 4: Calculated Expected Genotype frequencies**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Generation** | **Allele Frequency**  **Observed** | | **Genotype Frequency**  **Expected** | | |
| **R (p)** | **r (q)** | **p2** | **2pq** | **q2** |
| **0** | 0.5 | 0.5 | 0.25 | 0.5 | 0.25 |
| **1** | **0.3** | **0.7** |  |  |  |

1. Compare the Expected Genotype Frequencies from Table 4 to the Observed Genotype Frequencies from Table 3 for Generation 1.
   1. Do your expected frequencies match your observed frequencies?
   2. Suggest a reason for why the frequencies do (or do not) match.
2. Construct a Line Graph showing the allele frequencies p and q in the population over all 10 Generations. You should have two lines, one for each allele. Don’t forget to give your graph a title and axes labels.



1. Did the allele frequencies **p** and **q** remain the same from generation to generation in the population?
2. Did the population undergo microevolution? Explain your reasoning.
3. Based on your observations, what is the result of drastic declines in population size (bottlenecks) on genetic diversity in a population?
4. Given your understanding of evolution by means of natural selection, is a change in the amount of genetic diversity in a population likely to impact the survival of this population over many generations? Explain your reasoning.