**Hardy-Weinberg Guided-Inquiry Model Investigation**

**Learning Objectives**

* To use a data set that reflects a change in the genetic makeup of a populationover time and to apply mathematical methods and conceptual understandings toinvestigate the cause(s) and effect(s) of this change
* To apply mathematical methods to data from a real or simulated population topredict what will happen to the population in the future
* To evaluate data-based evidence that describes evolutionary changes in the geneticmakeup of a population over time
* To use data from mathematical models based on the Hardy-Weinberg equilibrium toanalyze genetic drift and the effect of selection in the evolution of specific populations
* To justify data from mathematical models based on the Hardy-Weinberg equilibriumto analyze genetic drift and the effects of selection in the evolution of specificpopulations
* To describe a model that represents evolution within a population
* To evaluate data sets that illustrate evolution as an ongoing process

**Step 1 Formulate the question.**

Think about a recessive Mendelian trait such as cystic fibrosis. Why do recessive alleleslike cystic fibrosis stay in the human population? Why don’t they gradually disappear?Now think about a dominant Mendelian trait such as polydactyly (more than fivefingers on a single hand or toes on a foot) in humans. Polydactyly isa dominant trait, but it is not a *common* trait in most human populations.Why not?How do inheritance patterns or allelefrequencies change in a population? Ourinvestigation begins with an exploration of answers to these simple questions.

**Step 2 Determine the basic ingredients.**

Let’s try to simplify the question *How do inheritance patterns or allele frequencies changein a population?* with some basic assumptions. For this model, assume that all theorganisms in our hypothetical population are diploid. Thisorganism has a gene locuswith two alleles — *A* and *B*. (We could use *A* and*a* to represent the alleles, but *A* and *B*are easier to work with in the spreadsheet you’ll be developing.) So far, this imaginarypopulation is much like any sexually reproducing population.How else can you simplify the question? Consider that the population has an infinitegene pool (all the alleles in the population at this particular locus). Gametes for the nextgeneration are selected totally at random. What does that mean? Focus on answeringthat question in your lab notebook for a moment — it is key to our model. For now let’sconsider that our model is going to look only at how allele frequencies might changefrom generation to generation. To do that we need to describe the system.

**Step 3 Qualitatively describe the biological system.**

To make this initial exploration into a model of inheritance patterns in a population,you need to make some important assumptions — all the gametes go into one infinitepool, and all have an equal chance of taking part in fertilization or formation of a zygote.For now, all zygotes live to be juveniles, all juveniles live to be adults, and no individualsenter or leave the population; there is also no mutation. Make sure to record theseassumptions in your notebook; later, you will need to explore how your model respondsas you change or modify these assumptions.

**Step 4 Quantitatively describe the biological system.**

Spreadsheets are valuable tools that allow us to ask *What if?* questions. They canrepeatedly make a calculation based on the results of another calculation. They can alsomodel the randomness of everyday events. Our goal is to model how allele frequencieschange through one life cycle of this imaginary population in the spreadsheet. Usethe diagram in Figure 1 as a guide to help you design the sequence and nature of yourspreadsheet calculation. The first step is to randomly draw gametes from the gene poolto form a number of zygotes that will make up the next generation.

To begin this model, let’s define a couple of variables.

*Let p* = the frequency of the *A* allele

and let *q* = the frequency of the *B* allele

Testing Your Mathematical Model

You now have a model with which you can explore how allele frequencies behave and change from generation to generation. Working with a partner, develop a plan to answer this general question: *How do inheritance patterns or allele frequencies change in a population over one generation?* As you work, think about the following more specific questions:

* What can you change in your model? If you change something, what does the change tell you about how alleles behave?
* Do alleles behave the same way if you make a particular variable more extreme? Less extreme?
* Do alleles behave the same way no matter what the population size is? To answer this question, you can insert rows of data somewhere between the first row of data and the last row and then copy the formulas down to fill in the space.

Try out different starting allele frequencies in the model. Look for and describe the patterns that you find as you try out different allele frequencies. Develop and use a pattern to select your values to test and organize your exploration. In particular, test your model with extreme values and intermediate values. In your lab notebook, describe your observations and conclusions about the population inheritance patterns you discover.

Try adding additional generations to your model to look at how allele frequencies change in multiple generations. To do this, use your newly recalculated *p* and *q* values to seed the next generation. Once you’ve included the second generation, you should be able to copy additional generations.

Try to create a graph of *p* values over several generations, for different-sized populations. See if you can detect a pattern of how population size affects the inheritance pattern. Be sure to try out both large and small populations of offspring.

This model relies on the RAND function to randomly select gametes from an infinite gene pool.

* What would happen if there were no randomness to this selection?
* What kind of pattern of genotypes would you expect in the next generation?

Analyzing Meaning of Data

Based on the calculations you made while testing your model, how would you answer the following questions?

* In the absence of random events (an infinitely large population), are the allele frequencies of the original population expected to change from generation to generation?
* How does this compare to a population that has random gamete selection but is small?
* What happens to allele frequencies in such a population? Is it predictable?

This mathematical model can predict allele frequencies from generation to generation. In fact, it is a *null* model. That is, in the absence of random events or other real-life factors that affect populations, the allele frequencies do not change from generation to generation. This is known as the Hardy-Weinberg equilibrium (H-W equilibrium). The H-W equilibrium is a valuable tool for population biologists because it serves as a baseline to measure changes in allele frequencies in a population. If a population is not in H-W equilibrium, then something else is happening that is making the allele frequencies change.

* What factors can cause allele frequencies to change in a population? (Hint: There are many.) How could you model these factors using your spreadsheet?

**Designing and Conducting Your Investigation**

By this point you’ve been able to use your model to explore how random chance affects the inheritance patterns of alleles in large and small populations. Perhaps you’ve also been able to find some interesting patterns in how alleles behave across generations. At the end of the last section you were asked what factors can cause allele frequencies to change in a population and how you would model them. Choose one of your answers, and try it out using your spreadsheet. This may involve adding multiple columns or rows along with a few extra operations. Keep the life cycle of your hypothetical population in mind as you develop additional strategies.

With your new spreadsheet model, generate your own questions regarding the evolution of allele frequencies in a population. From these questions (noted in your lab notebook), you need to develop hypotheses that you can test — those that allow you to easily manipulate the parameters of population size, number of generations, selection (fitness), mutation, migration, and genetic drift. Collect sufficient data by running your model repeatedly. Analyze your data. Formulate your conclusions and present a mini poster that supports your claim with sound reasoning and evidence to the class.