

Mathematical Modeling of Allele Frequency Changes over Time

Evolution occurs in populations of organisms as a result of the differential reproductive success among individuals that is a product of natural variation. One way to study evolution is to study how the frequency of alleles in a population changes from generation to generation. In other words, you can ask *“What are the inheritance patterns of alleles, not just from two parental organisms, but also in a population?”* You can then follow how allele frequencies change in populations and investigate how these changes might predict what will happen to a population in the future.

Mathematical models and computer simulations are tools used to explore the complexity of biological systems that might otherwise be difficult or impossible to study. In this activity, we'll use a spreadsheet to model how a hypothetical gene pool changes from one generation to the next. This model will let us explore parameters that affect allele frequencies, such as selection, mutation, and migration.

The real world is infinitely complicated. To penetrate that complexity using model building, we need to learn to make reasonable, simplifying assumptions about complex processes. For example, climate change models or weather forecasting models are simplifications of very complex processes - more than can be accounted for with even the best models. These models allow us to make predictions and test hypotheses about climate change and weather. Because any model is a simplification of the real world, we need to constantly evaluate the assumptions we make as we build a model, as well as evaluate the results of the model with a critical eye. This is actually one of the powerful benefits of a model - it forces us to think carefully about an idea.

The approach described by Otto and Day (2007) suggest the following steps:

Step 1 Formulate the question.

In this activity we will ask *“How do allele frequencies change in a population from one generation to the next?”*

Step 2 Determine the basic ingredients.

Let's try to simplify the question with some basic assumptions. For this model, assume that all the organisms in our hypothetical population are diploid. This organism has a gene locus with 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 we'll use.) So far, this imaginary population is much like any sexually reproducing population.

How else can we simplify the question? Let's assume that the population has an infinite gene pool which includes all alleles and that the gametes for the next generation are selected totally at random.

Q. 1 What implications do these two assumptions have for our model?

Our model is going to look at how allele frequencies might change from generation to generation.

Step 3 Qualitatively describe the biological system.

Q. 2 Describe how gametes would be produced by our sample organism and how alleles would be assorted into those gametes.

Q. 3 What assumptions will we have to make in building the model?

Step 4 Quantitatively describe the biological system.

Spreadsheets are valuable tools that allow us to ask *What if?* questions. They can repeatedly make a calculation based on the results of another calculation. They can also model the randomness of everyday events. Our goal is to model how allele frequencies change from one generation to the next in an imaginary population.

Q. 4 How do you think we can use a spreadsheet to answer the question “*How do allele frequencies change in a population from one generation to the next?*”

Step 5. Analyze the equations.

This step has already been done.

Step 6. Perform checks and balances.

The spreadsheet model must be tested to make sure it behaves as we would expect.

Q. 5 How might we accomplish this?

Q. 6 Explain what it means for a population to be in Hardy-Weinberg equilibrium.

Q. 7 Does the frequency of any allele ever reach zero?

Q. 8 a) How can we use the spreadsheet to model the effect of genetic drift?

b) Why are many endangered or threatened species at risk due to genetic drift ?

Q. 9 a) How can we change the model to investigate the effect of different population sizes?

b) What do you predict about the allele frequencies when we increase the population size?

Q. 10 Perform a chi-squared analysis for each of the populations to determine if the allele frequency is as you would expect after 10 generations.

Q. 11 How can we change the model to account for variables like immigration, emigration, death, *etc*?

Q. 12 Suggest some additional questions which could be investigated using this model.