

PROCESS OF SCIENCE, NATURAL SELECTION, AND GENETIC DRIFT

OBJECTIVES

Process of science

- Understand the different steps in the scientific process: observations, hypotheses, experiments, results, conclusions
- Understand how these steps are related to each other

Natural selection

- Understand the following terms: population, gene, allele, genotype, phenotype, homozygous, heterozygous, recessive, dominant, and fixation
- Understand the 4 requirements for evolution by natural selection
- Understand how evolution by natural selection changes the allelic and genotypic proportions

Genetic drift

- Understand the differences between genetic drift and natural selection
- Understand how time affects evolution by genetic drift
- Understand how population size affects evolution by genetic drift
- Understand how starting allelic proportions affect evolution by genetic drift

TIMELINE

20 mins General Lab Introduction

70 mins Process of science

80 mins Evolutionary processes

45 mins Natural selection

35 mins Genetic drift

EXERCISE 1: THE PROCESS OF SCIENCE

The Process of Science is the method by which scientists observe and study phenomena as well as share their discoveries with other scientists and the public. Although scientists employ a variety of techniques to study the world in which they live, there is a unifying framework that all scientists use—the **scientific method**. The steps of the scientific method are shown in Figure 1.

As Figure 1 depicts, the scientific method is not a linear, unidirectional process, but rather a highly dynamic and multidirectional process. Each of the steps in the scientific method feed into the development and re-development of the other steps. Let's explore these primary steps in more detail.

Form a Question: A question should reflect an observation or set of observations. Observations can come from many different experiences, such as seeing something in nature, reading a scientific article, or talking with a colleague. The question should be stated so a scientist can answer it by designing a study and collecting data. For example, do individuals of the monarch butterfly (*Danaus plexippus*) look like their parents?

Form Hypotheses: A hypothesis is a possible answer to a question that is testable. Hypotheses are based on observations or intuition that occur before an investigator begins a study. Almost always scientists develop two or more hypotheses for a given question. A set of well-developed hypotheses include mutually exclusive answers to a question. For our monarch butterfly example, a possible set of hypotheses is simply yes they do or no they don't. Another possible set of hypotheses is they look like their father, they look like their mother, they look intermediate between the two, or they look more extreme than either parent. Both are well-formed sets of hypotheses, and allow an investigator to gain support for each hypothesis of a set by collecting data.

At this point you might be asking yourself, "How will a scientist determine whether the offspring and parents are similar or not?" To do this scientists develop another set of statistical hypotheses that are matched with their biological hypotheses. We will discuss statistical hypotheses in the following step of the scientific

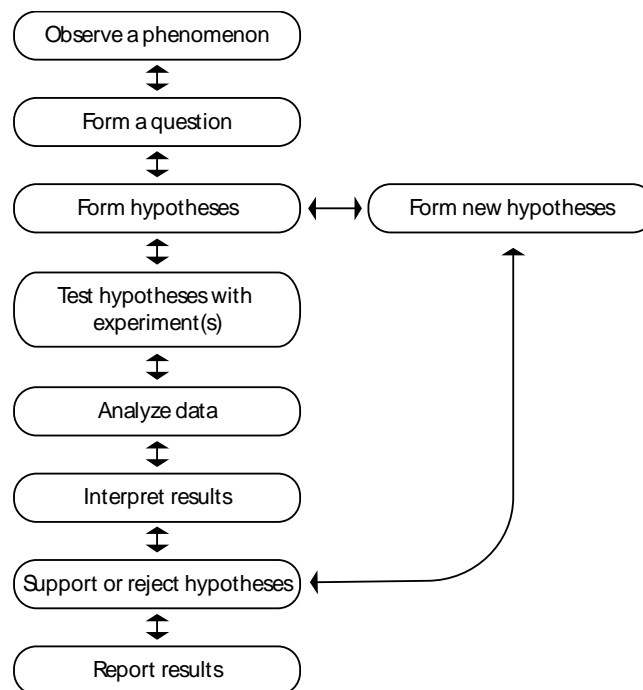


Figure 1. A flow chart depicting the scientific method.

process.

Test with Experiment: An experiment should be designed in a way such that the results will clearly support one hypothesis and will be able to eliminate (falsify) the others. The experiment should be clear for other scientists to follow and repeat. In our butterfly example, the investigator might be interested in antennae length. During a study, she will measure the length of antennae for parents and offspring of butterflies. The statistical hypotheses she poses might be 1) the average length of parents' antennae differs from the average length of offspring's antennae more than expected by chance, and 2) the average length of parents' antennae and the average length of offspring's antennae are the same or differ by an amount expected by chance. As you can see the important phrase is differs by chance, and is required because scientists almost always take a small sample from the population and there is chance involved in which individuals are chosen and measured.

Test yourself

Match up these two statistical hypotheses with the first set of hypotheses posed in the "Form Hypotheses" section, above.

Statistical hypothesis 1 matches with the hypothesis that parents and offspring are different
Statistical hypothesis 2 matches with the hypothesis that parents and offspring look the same

An experiment typically includes an **independent variable** (the variable that is changed in the experiment), and a **dependent variable** (the variable that you observe for changes in response to the independent variable). Generation is our independent variable, and antennae length is our dependent variable. The design of experiments is critical for clearly distinguishing among hypotheses and posing an answer to a question; scientists train for many years to design and execute experiments well. Even our seemingly simple butterfly example would require a great deal of thought about which individuals to measure, and when and how to measure them.

Analyze Data: Scientists typically analyze data from their experiments with graphs, statistics, and mathematical models. The results from these analyses are then used to draw conclusions about the hypotheses. Let's assume that there was a statistical difference between the average antennae length of parents and offspring. That is to say, statistical hypothesis 1 was supported.

Interpret Results: Scientists use the results of their experiments to distinguish between their hypotheses. This step requires critical thinking skills and is one of the most important steps in the scientific process. If the experiment was well designed and the statistical hypotheses matched well with the biological hypotheses, then this step is easy because all the hard work was done in the above steps. The outcome of the statistical test in our butterfly example supports the hypothesis that offspring of the monarch butterfly do not look like their parents. However, the results could be interpreted in more detail by also creating a graph of the mean antennae length of parents and offspring to determine whether parents have longer or shorter antennae on average compared to offspring.

Scientists learn from supporting and rejecting hypotheses. We come closer to the truth when we can narrow down the possible answers. This happens when there is strong evidence for a given hypotheses but no evidence against an alternative. It also happens when there is strong evidence against a hypothesis and no

evidence for an alternative. The latter case forces scientists to go back to the drawing board, and think more deeply about what might be causing the phenomenon they observed in nature. It is not uncommon that a well-reasoned, “common sense” hypothesis isn’t supported, but with the continued application of the scientific process to the problem, our understanding of the complexity of the natural world benefits in the long run.

Share Results and Conclusions: Scientific research is shared with the scientific community through peer reviewed papers and conferences. Results and conclusions can also be shared with the general public through a variety of media avenues.

The Scientific Process: In-Class Assignment

For this exercise, you will work with other students at your table. Your TA will assign each group to a biological discipline, and provide the class with a hypothetical scenario and question. Your TA will then pass out information about the research your group is doing and you will develop hypotheses for the research. After your group has a well-formed set of hypotheses, your TA will give your group the research results. Your group will then interpret the results, distinguish among your hypotheses, and develop a presentation. Once all the groups are finished, each group will present their findings to the class. Each group will take notes on what the other groups have found during their presentations. Each group will then use the new information from the other groups to re-evaluate their hypotheses and draw conclusions.

For this exercise you will need to rely on the steps of the scientific method to solve the problem:

- (1) **State the Question:** Based on the limited information provided, state in question form the problem you are trying to solve.
- (2) **Form Hypotheses:** Based on the information provided and any other previous knowledge. Form at least one primary hypothesis and one alternative hypothesis. These need to be mutually exclusive.
- (3) **Test with an Experiment & Analyze Results:** Each group will be provided with a set of results from experiments or field data with information about the outcome of analyses that need to be interpreted. In the space below briefly describe your data and results
- (4) **Interpret Results:** What conclusions can you draw from the results? Which hypotheses do the data support and reject?
- (5) **Share Results with the Class:** Each group will briefly (2-3 minutes) present their results and conclusions to the class in a mock scientific conference. Results presented by other groups and those from your own group will be used to re-evaluate the initial question and if necessary formulate alternative hypotheses to test.
 - (a) **Take notes on the results and conclusions presented by other groups.**
 - (b) **Do the results from the other groups support your original hypothesis? Why or why not?**
 - (c) **Do you think malaria is or will soon be endemic? What data support each hypothesis?**

EXERCISE 2: EVOLUTION BY NATURAL SELECTION AND GENETIC DRIFT

Evolution is the change in allele proportions in a population over time. There are five general ways in which evolution can occur.

- **Natural selection**
- **Genetic drift**
- **Gene flow**
- **Mutation**
- **Nonrandom mating**

Of these, only natural selection leads to adaptations. Mutations provide the source of genetic variation on which natural selection acts. The other processes lead to changes in allele frequencies that are not necessarily adaptive. During the remainder of the lab you will complete computer exercises to improve your understanding of natural selection and genetic drift. But first we need to review a few terms.

Is sex random?

When biologists learned that genes passed from parents to offspring cause the traits of an organism, they needed terms to distinguish between the information that codes for traits and the traits themselves—the terms genotype and phenotype were proposed early in the 20th century. A **genotype** is the genetic sequence of an organism and a **phenotype** is the sum of the traits an organism possesses. Genes are sequences of DNA or RNA that code for some product, typically a protein. An **allele** is a particular sequence of a gene that codes for a particular form of a given protein. Different alleles of a gene code for different forms of that same protein. These different alleles are responsible, in part¹, for the variation in traits observed among individuals of a given species.

Test yourself

Specify whether each of the following examples represent part of a phenotype or genotype: brown hair, shape of a red blood cell, an allele for blue eyes, the gene for a protein, the shape of a protein, and your height.

Phenotype, phenotype, genotype, genotype, phenotype, phenotype

A **population** is a group of organisms of the same species living in the same place at the same time. There can be many populations of a species, but only one species for a population. A **gene pool** refers to all of the alleles present within a population.

As you already know, offspring typically receive genes from their mothers and fathers². Commonly, organisms have two alleles for a given gene, receiving one allele from each parent, and are called **diploid** organisms. However, some organisms are haploid and have only one copy of each gene, and some organisms are polyploid and have more than two copies of each gene. Diploid organisms either have two copies of the same allele or two different alleles for a gene. Organisms with the same allele for each copy of a given gene are called **homozygotes** (the noun) or **homozygous** (the adjective) for that particular gene. Organisms with two different alleles for a given

¹ The environment also affects the traits of organisms.

² Some organisms only receive genetic information from one of the parents.

gene are called **heterozygotes** (the noun) or **heterozygous** (the adjective)³. Diploid organisms produce **haploid** gametes (sex cells, that is, sperm and eggs) with only one allele for each gene. All gametes of a homozygote have the same allele. Half of the gametes from a heterozygote have one of the two alleles, and the other half have the other allele. When sperm and egg merge during fertilization, they then create a new diploid zygote with the allele combinations from that sperm and egg. If we know the genotypes of all the individuals in a population in one generation and the genotypes of all of their offspring in the next generation, we can determine the allele proportions in each generation and whether or not these proportions are changing (that is, whether or not evolution is occurring). If mating and fertilization are random with respect to the gene in question, and none of the other mechanisms of evolution (see above) are occurring, then the proportions of the alleles for that gene in a population will stay the same generation after generation. If the allele proportions are changing generation to generation, then at least one of the evolutionary mechanisms must be occurring. The question that frequently confronts scientists is, “Which one?”

Now consider how an allele affects the phenotype of an organism. In the simplest case, a phenotypic difference is caused by different alleles of a single gene—in a set of classic experiments by Gregor Mendel, there was an allele that caused green peas and an allele that caused yellow peas. To begin with, let us consider one allele as **dominant** and the other as **recessive**. In the case of Mendel’s peas, the allele for green peas was dominant and the allele for yellow peas was recessive. An individual that received at least one copy of a dominant, green-pea allele produced green peas. An individual that received two copies of the recessive, yellow-pea allele produced yellow peas. Note that homozygotes with two copies of the dominant allele and heterozygotes have the same phenotype, green peas, even though they have different genotypes! When alleles have a dominant-recessive relationship, only two phenotypes can occur even though there are three different genotypes—homozygous dominant, homozygous recessive, and heterozygous. Not all alleles have a dominant-recessive relationship. Some alleles are incompletely dominant. When this occurs the phenotype of a heterozygote is different (and typically intermediate) from either of the homozygotes’ phenotypes. Incomplete dominance results in a different phenotype of each genotype.

Test yourself

Consider a population of 120 diploid goldfish in which there are two alleles for the gene that determines fin length, one for short fins (f) and one for long fins (F). The long-fin allele is dominant over the short-fin allele (capital letters typically denote the dominant allele), and there are 40 homozygous-dominant individuals (FF), 40 heterozygous individuals (Ff), and 40 homozygous-recessive individuals (ff).

How many alleles are there in total in the population for the fin-length gene? How many are the F allele? How many are the f allele? What is the proportion of the F allele in the population? What is the proportion of the f allele in the population?

240, 120, 120, 0.5, 0.5

³ “Homo” means alike and “hetero” means different in Greek. A zygote is the cell that results when an egg and sperm fuse.

How does evolution work?

There are four general requirements for evolution by natural selection to occur.

1. An organism's phenotype must be heritable.
2. An organism's traits must affect the number of offspring it produces.
3. There must be variation in the phenotypes of individuals within a population.
4. Organisms must produce more offspring than can survive to reproduce.

If these requirements are met, then parents who have more viable offspring (i.e., offspring that can reproduce) contribute more to the next generation, and the proportion of their alleles within the gene pool increase. For example, if a moth with cryptic coloration has less chance of being eaten by a bird, then, on average, cryptic moths will live to reproduce more than moths that are less camouflaged. The genes that code for cryptic coloration will then be more prevalent in subsequent generations of that population of moths.

Genetic drift also causes a shift in the proportion of alleles in a population. However, genetic drift does not lead to adaptation like natural selection. Genetic drift is the result of chance, or random, events. There is a chance that the number of offspring an individual produces is not the result of its traits. Take a simple example where there are two alleles for pea color in a population. Now let's say a large storm hits a field and kills half the pea plants. Just by chance, a greater proportion of individuals with one allele than the other dies in the storm (that is, pea color has no effect on storm survival). That means that more of the plants with one of the two alleles will survive and contribute more of their alleles to the next generation. Genetic drift can also occur through the chance combination of gametes. If we consider a single gene, then heterozygotes will produce two types of gametes—one with one of the two alleles and the other with the other allele. If this heterozygote is lucky enough to mate, then there is a chance either type of gamete will produce an offspring. Say this individual has four offspring. For simplicity, let's just focus on one parent for now. It is possible that two of the offspring get one allele from that parent and the other two offspring get the other allele, preserving the original allele proportions. However, it's quite possible that one offspring will get one allele, but three get the other, or that all four offspring get the same allele. Either of these latter cases would alter the allele proportions in the population from the parental generation, thus causing evolution.

The concepts of natural selection and genetic drift are critical to understand because they form the foundation of evolution, and of the science of biology itself. You are therefore going to spend the remainder of the lab working on computer exercises to improve your understanding of these critical biological concepts.

EVOBEAKER COMPUTER LAB—SICKLE-CELL ALLELES

Malaria is one of the world's most serious diseases, infecting upwards of 300 million people and killing 1.5 million people each year. It is most common in Africa but occurs in warmer climates worldwide. People are infected when bitten by mosquitoes carrying certain kinds of protozoa. The malarial protozoa are released as the mosquito's mouth parts pierce the skin of the unlucky victim. The protozoa then swim through the victim's blood until reaching the liver. There they reproduce and emerge to infect the host's red blood cells, after which another mosquito can suck them up and start the cycle over again.

Just about anything that would protect people from malaria would be beneficial for those who live in the malaria-prone areas of the world. And indeed, some people carry an allele for a gene that provides a defense.

Surprisingly, this anti-malaria allele was tracked down through studies of a seemingly completely unrelated disease: sickle-cell anemia. Sickle-cell anemia is every bit as nasty as malaria. Individuals with this disease have red blood cells that form a sickle shape instead of the normal circular shape. The sickle-shaped cells tend to get stuck in small blood vessels, blocking flow, and halting the supply of oxygen to downstream cells.

Unlike malaria, sickle-cell anemia is a genetic disease. Individuals inherit alleles that cause the disease from their parents. Sickle-cell anemia is associated with a gene that encodes part of the hemoglobin molecule (called the *Hb* gene). Hemoglobin is the protein in red blood cells that carries oxygen. The allele for the normal hemoglobin protein is called *HbA* and the allele for sickle cell anemia is called *HbS*. People who inherit the *HbS* allele from both parents (i.e., are homozygous for *HbS*) have a form of hemoglobin that makes their red blood cells highly prone to becoming sickle-shaped. People who inherit one sickle-cell allele and one normal hemoglobin allele (i.e., they are heterozygotes), often have such minor symptoms that they do not even realize that they carry the *HbS* allele.

Although people with sickle-cell anemia typically die from the disease before they are old enough to reproduce, it is relatively common in some parts of the world. Why doesn't natural selection eliminate the sickle-cell allele? The answer is that although the sickle-cell allele can kill you, it can also protect you from malaria. Having one copy of the *HbS* allele protects you from malaria. A heterozygote's red blood cells infected with the malaria protozoa will sickle, and their immune system will dispose of the infected cells. So anyone who is heterozygous with both the *HbS* and *HbA* alleles is protected from both malaria and sickle-cell anemia. This is an example of "heterozygote advantage". It is also called "balancing selection", because forces that are selecting for one allele are balanced by forces selecting for the other. Thus both alleles tend to stay in the population.

Test yourself

Complete the table below.

DESCRIPTION	GENOTYPE	SUSCEPTIBLE TO MALARIA?	HAS SICKLE-CELL ANEMIA?
Homozygous for normal hemoglobin allele	HbA/HbA		
Heterozygous			

Yes, no, HbA/HbS, no, no, homozygous for the sickle-cell allele, HbS/HbS, no, yes

For today's exercises, you will use a simulation model of individuals in a village. The population of this village hovers around 200 people, though that number rises and falls from year to year. Each year, each female in the population can have 1 offspring, which for simplicity, becomes an adult the following year (realism occasionally suffers when making models, though in this case the difference doesn't qualitatively affect the outcome; it just makes the calculations easier). Each offspring receives one allele for the *Hb* gene from each parent. The number of offspring in the population is limited by how close the population is to the carrying capacity of 200. Similarly, each year, each individual has a chance of dying that is independent of either disease, and that chance rises as the population increases in size. The malaria and sickle-cell death rates also add to the chance of dying for homozygous *HbA* and *HbS* individuals, respectively (remember that heterozygous individuals are immune from both diseases). The simulation is initialized in year 1 with 100 homozygotes for the *HbA* allele (50 females and 50 males), plus a certain number of heterozygotes ("carriers") that are all female (the default number of carriers is 50).

Exercise 2a: Natural selection

- ❖ Launch the Sickle-cell Alleles computer simulation. You should see a village in Africa depicted on the left side of the screen. Mosquitoes will hover in the blue sky above the village once you start the simulation. The graph entitled “Allele Frequencies” will plot the proportions⁴ of the normal hemoglobin allele (*HbA*) and the sickle-cell allele (*HbS*) in the village over time. The STATISTICS panel will track the number of deaths from malaria and sickle-cell anemia. The PARAMETERS panel will allow you to adjust the initial number of carriers of the sickle-cell allele in the population at the beginning of the simulation, as well as the death rate of villagers afflicted with sickle-cell anemia and malaria. In the bottom-left corner are buttons that start and stop the simulation.
- ❖ Select the STEP button at the bottom-left corner. In the blink of an eye you are seeing the village one year later. Now, select GO to start the simulation. After 150 years, several seconds for you, select STOP. You can select RESET to re-initialize the population and start over.

Questions

1. Recall that you initialize the population with 50 male and 50 female homozygotes for the *HbA* allele, and some number of carriers (which you can specify). The carriers are all female. How many *HbS/HbS* individuals do you expect in year one, why?

Use the graph of allele proportions to answer the following questions, making sure that you have selected “Very Wet” conditions in the settings. You can click the left mouse button on the graph to determine the exact values.

2. In each time step, what do the proportions of the two alleles add up to?
-

- ❖ Select GO again and let the simulation run for 300 hundred years, then select STOP. The proportions will bounce around but you should be able to estimate the average proportion for each allele.

Questions

3. Given death rates for malaria=0.2 and anemia=0.1 (you will have to change the anemia death rate from the default of 0.2), which allele should have a lower proportion? Why?
4. To check your answer, what is the average proportion for last 100 years of the sickle-cell allele and the normal allele (just eyeball it)?
5. Given the death rates and numbers of each genotype (look at the village), which illness should kill more people? Why?

⁴ In genetics, the term “frequency” and “proportion” are synonyms. However, in all other areas of biology and all other sciences they have different meanings. Frequencies are a count and restricted to integers. In contrast, proportions are a fraction restricted to values ≥ 0 and ≤ 1 . For clarity we will refer to the values in the graph as proportions, but note that this software and your textbook sticks to genetics conventions and refers to allele frequencies.

6. To check your answer, how many malaria deaths and anemia death were there in the past five years?

❖ It is now time for you to perform several experiments. Below are specific questions. Use the scientific method to answer them. Start by developing several hypotheses to answer the proposed question. Then collect data from the villagers and determine which hypothesis is best supported by your data. You will work in groups of 2-4. You will present your research to the class for only one of the questions below. However, you will not know which question until shortly before you present. So, make sure to manage your time so you can work on all the questions and keep careful notes on the hypotheses, methods, and results for each one. For your presentation, you will need to state the following:

- The scientific question you are addressing (hint: it is one of the questions below).
- A well-formed set of hypotheses, and the reasoning that led you to those hypotheses.
- Your methods – what did you do to test your hypotheses?
- Your results – what did the data show?
- Your conclusions – which of your hypotheses were best supported?

Scientific questions:

- a) How will the initial number of carriers (that is, heterozygotes) at beginning of the simulation, affect the proportions of alleles in the population after 300 years?
- b) What will happen to the proportions of each allele if sickle-cell anemia is cured with a drug? You can adjust the death rate from sickle-cell anemia in the textbox in PARAMETERS.
- c) In which region (e.g., very wet, wet, slightly wet, or dry) should sickle-cell anemia be most common?
- d) What will happen to the proportions of each allele as the death rate of sickle-cell anemia increases from 0.1 to 0.7 (this is a 700% increase in the death rate, such that 70% of the homozygous *HbS* individuals will die)?
- e) What will happen to the portions of each allele if the death rate of sickle-cell anemia is 0.1, and malaria cured? (Note, you can adjust the region of Africa to dry/no mosquitoes to manipulate the malaria death rate.)
- f) What will happen if both malaria and sickle-cell anemia are cured?

Questions

7. Under the initial conditions of this simulation, are the four general requirements for natural selection satisfied? List each and explain your answer.
 8. If malaria and sickle-cell anemia are both cured, which requirement for natural selection is no longer met?
 9. If malaria and sickle-cell anemia are both cured, which evolutionary process is still occurring? How can you tell?
-

Exercise 2b: Genetic Drift

- ❖ Select “Genetic Drift” from the EXPERIMENT pull-down menu at the bottom-right corner. You should see a newspaper article with headlines that indicates cures for both sickle-cell anemia and malaria were discovered. Below are PARAMETERS in which you can change the initial frequency of the *HbS* allele, and the number of people in a small, medium, and large village. To the right are three graphs. Each plots the proportion of the *HbS* allele through time for the small, medium, and large villages.
- ❖ As a group, answer the following questions using the scientific method. Start by developing several hypotheses, then collect data from the villagers and determine which hypothesis is best supported by your data. You will present your research to the class for only one of the questions below. However, you will not know which question until shortly before you present. So, make sure to manage your time so you can work on all the questions, and keep careful notes on your hypotheses, methods, results, and conclusions. Think about mechanistic explanations of the phenomena that you observe (that is, why do you think things happened the way they did?).

Scientific questions:

- a) Does the amount of genetic drift, that is the amount an allelic proportion might drift away from the initial proportion, increase with time?
- b) Are larger populations more affected by genetic drift than smaller populations?
- c) Are smaller populations more likely to lose the *HbS* allele through genetic drift than larger populations after 500 years?
- d) Does the chance of increase in the proportion of the *HbS* allele change with population size?
- e) Does the chance that the *HbS* allele becomes fixed (that is, proportion = 1) change with the initial proportion of the *HbS* allele?
- f) What starting allelic proportion will equal the average allelic proportions of many run after 100 years?

ASSIGNMENT

Turn in your answers for questions 1-9 above at the end of class. You should have one set of answers for your entire lab group.