Principles of Biology

Principles of Biology

BIOLOGY 211, 212, AND 213

LISA BARTEE; WALTER SHRINER; AND CATHERINE CREECH

Open Oregon Educational Resources

Contents

Principles of Biology: The Principles of Biology sequence (BI 211, 212 and 213) introduces biology as a scientific discipline for students planning to major in biology and other science disciplines. Laboratories and classroom activities introduce techniques used to study biological processes and provide opportunities for students to develop their ability to conduct research.

- BI211 focuses on how structure defines function in organisms and the pathways and transformation of energy in living systems.
- BI212 uses genetics as a model system to understand information flow in living organisms.
- BI213 focuses on the interactions of living systems and the ecology and evolution of biodiversity.

THE PROCESS OF SCIENCE

Learning Objectives

Course Outcomes for this section:

Apply the scientific method to biological questions by designing experiments and using the resulting data to form a conclusion.

1. Design a controlled experiment to answer a biological question.
2. Predict the outcome of an experiment.
3. Collect, manipulate, and analyze quantitative and qualitative data
4. Answer a biological question using data.

Select, evaluate, and utilize discipline-specific information and literature to research a biological topic.

1. Differentiate between questions that can and cannot be answered using science.
2. Identify appropriate credible sources of

information to research a topic.

3. Evaluate sources of information for their strengths and weaknesses.
4. Differentiate between popular and scholarly sources.

Like geology, physics, and chemistry, **biology** is a science that gathers knowledge about the natural world. Specifically, biology is the study of life. The discoveries of biology are made by a community of researchers who work individually and together using agreed-on methods. In this sense, biology, like all sciences is a social enterprise like politics or the arts. The methods of science include careful observation, record keeping, logical and mathematical reasoning, experimentation, and submitting conclusions to the scrutiny of others. Science also requires considerable imagination and creativity; a well-designed experiment is commonly described as elegant, or beautiful. Like politics, science has considerable practical implications and some science is dedicated to practical applications, such as the prevention of disease (see Figure 1). Other science proceeds largely motivated by curiosity. Whatever its goal, there is no doubt that science, including biology, has transformed human existence and will continue to do so.

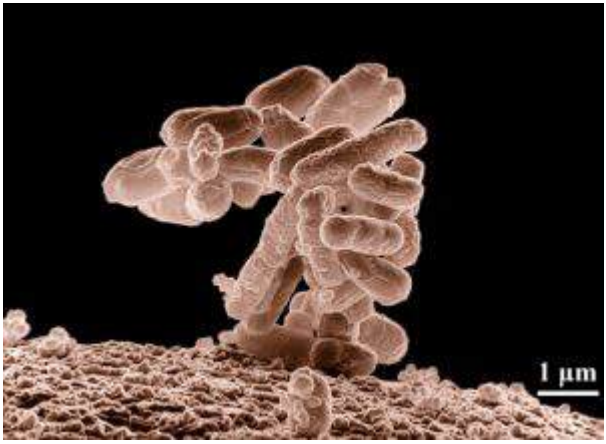


Figure 1 Escherichia coli (E. coli) bacteria, seen in this scanning electron micrograph, are normal residents of our digestive tracts that aid in the absorption of vitamin K and other nutrients. However, virulent strains are sometimes responsible for disease outbreaks. (credit: Eric Erbe, digital colorization by Christopher Pooley, both of USDA, ARS, EMU)

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:RD6ERYiU@5/The-Process-of-Science.

The Nature of Science

Biology is a science, but what exactly is science? What does the study of biology share with other scientific disciplines? **Science** (from the Latin scientia, meaning "knowledge") can be defined as knowledge about the natural world. Science is a very specific way of learning, or knowing, about the world. The history of the past 500 years demonstrates that science is a very powerful way of knowing about the world; it is largely responsible for the technological revolutions that have taken place during this time. There are however, areas of knowledge and human experience that the methods of science cannot be applied to. These include such things as answering purely moral questions, aesthetic questions, or what can be generally categorized as spiritual questions. Science can not investigate these areas because they are outside the realm of material phenomena, the phenomena of matter and energy, and can not be observed and measured.

The **scientific method** is a method of research with defined steps that include experiments and careful observation. One of the most important aspects of this method is the testing of hypotheses. A **hypothesis** is a suggested explanation for an event, which can be tested. Hypotheses, or tentative explanations, are generally produced within the context of a scientific theory. A **scientific theory** is a generally accepted, thoroughly

tested and confirmed explanation for a set of observations or phenomena. Scientific theory is the foundation of scientific knowledge. In addition, in many scientific disciplines (less so in biology) there are **scientific laws**, often expressed in mathematical formulas, which describe how elements of nature will behave under certain specific conditions. There is not an evolution of hypotheses through theories to laws as if they represented some increase in certainty about the world. Hypotheses are the day-to-day material that scientists work with and they are developed within the context of theories. Laws are concise descriptions of parts of the world that are amenable to formulaic or mathematical description.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=42#h5p-7

The scientific community has been debating for the last few decades about the value of different types of science. Is it valuable to pursue science for the sake of simply gaining knowledge, or does scientific knowledge only have worth if we can apply it to solving a specific problem or bettering our lives? This question focuses on the differences between two types of science: basic science and applied science.

- Basic science or "pure" science seeks to expand knowledge regardless of the short-term application of that knowledge. It is not focused on developing a product or a service of immediate public or commercial value. The immediate goal of basic

science is knowledge for knowledge's sake, though this does not mean that in the end it may not result in an application.

- In contrast, applied science or "technology," aims to use science to solve real-world problems, making it possible, for example, to improve a crop yield, find a cure for a particular disease, or save animals threatened by a natural disaster. In applied science, the problem is usually defined for the researcher.

Some individuals may perceive applied science as "useful" and basic science as "useless." A question these people might pose to a scientist advocating knowledge acquisition would be, "What for?" A careful look at the history of science, however, reveals that basic knowledge has resulted in many remarkable applications of great value. Many scientists think that a basic understanding of science is necessary before an application is developed; therefore, applied science relies on the results generated through basic science. Other scientists think that it is time to move on from basic science and instead to find solutions to actual problems. Both approaches are valid. It is true that there are problems that demand immediate attention; however, few solutions would be found without the help of the knowledge generated through basic science.

One example of how basic and applied science can work together to solve practical problems occurred after the discovery of DNA structure led to an understanding of the molecular mechanisms governing DNA replication. Strands of DNA, unique in every human, are found in our cells, where they provide the instructions necessary for life. During DNA replication, new copies of DNA are made, shortly before a cell divides to form new cells. Understanding the mechanisms of DNA replication enabled scientists to develop laboratory techniques that are now used to identify genetic diseases,

pinpoint individuals who were at a crime scene, and determine paternity. Without basic science, it is unlikely that applied science would exist.

Another example of the link between basic and applied research is the Human Genome Project, a study in which each human chromosome was analyzed and mapped to determine the precise sequence of DNA subunits and the exact location of each gene. (The gene is the basic unit of heredity; an individual's complete collection of genes is his or her genome.) Other organisms have also been studied as part of this project to gain a better understanding of human chromosomes. The Human Genome Project (**Figure 1**) relied on basic research carried out with non-human organisms and, later, with the human genome. An important end goal eventually became using the data for applied research seeking cures for genetically related diseases.

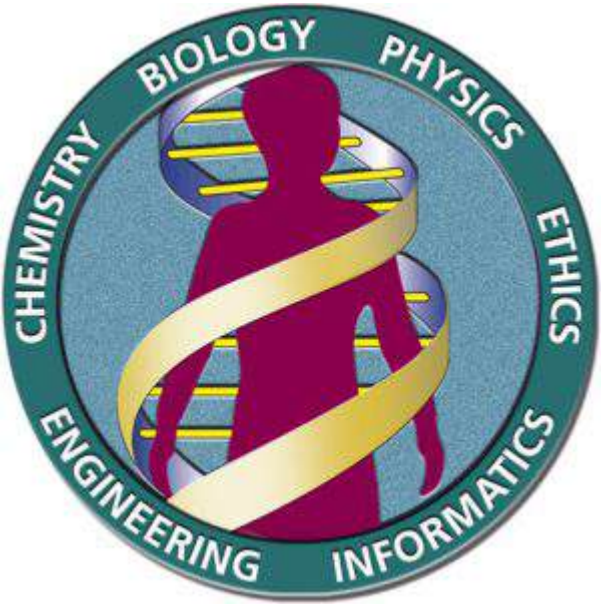


Figure 1 The Human Genome Project was a 13-year collaborative effort among researchers working in several different fields of science. The project, which sequenced the entire human genome, was completed in 2003. (credit: the U.S. Department of Energy Genome Programs (http://genomics.energy.gov))

While research efforts in both basic science and applied science are usually carefully planned, it is important to note that some discoveries are made by serendipity, that is, by means of a fortunate accident or a lucky surprise. Penicillin was discovered when biologist Alexander Fleming accidentally left a petri dish of Staphylococcus bacteria open. An unwanted mold grew, killing the bacteria. The mold turned out to be Penicillium, and a new antibiotic was discovered. Even in the highly organized world of science, luck—when combined with an observant, curious mind—can lead to unexpected breakthroughs.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:RD6ERYiU@5/The-Process-of-Science.

The Scientific Process

Biologists study the living world by posing questions about it and seeking science-based responses. This approach is common to other sciences as well and is often referred to as the **scientific method**. The scientific process was used even in ancient times, but it was first documented by England's Sir Francis Bacon (1561–1626) (**Figure 1**), who set up inductive methods for scientific inquiry. The scientific method is not exclusively used by biologists but can be applied to almost anything as a logical problem solving method.

Figure 1 Sir Francis Bacon (1561–1626) is credited with being the first to define the scientific method. (credit: Paul van Somer)

QUESTION

The scientific process typically starts with an **observation** (often a problem to be solved) that leads to a **question.** Remember that science is very good at answering questions having to do with observations about the natural world, but is very bad at answering questions having to do with morals, ethics, or personal opinions.

Questions that can be answered using science	Questions that cannot be answered using science
• What is the optimum temperature for the growth of E. coli bacteria?	• How tall is Santa Claus?
• Do birds prefer bird feeders of a specific color?	• Do angels exist?
• What is the cause of this disease?	• Which is better: classical music or rock and roll?
• How effective is this drug in treating this disease?	• What are the ethical implications of human cloning?

Let's think about a simple problem that starts with an observation and apply the scientific method to solve the problem. Imagine that one morning when you wake up and flip a the switch to turn on your bedside lamp, the light won't turn on. That is an observation that also describes a problem: the lights won't turn on. Of course, you would next ask the question: "Why won't the light turn on?"

HYPOTHESIS

Recall that a **hypothesis** is a suggested explanation that can be tested. *A hypothesis is NOT the question you are trying to answer – it is what you think the answer to the question will be **and why**.* To solve a problem, several hypotheses may be proposed. For example, one hypothesis might be, "The light won't turn on because the bulb is burned out." But there could be other answers to the question, and therefore other hypotheses may be proposed. A second hypothesis might be, "The light won't turn on because the lamp is unplugged" or "The light won't turn on because the power is out." A hypothesis should be based on credible background information. A hypothesis is NOT just a guess (not even an educated one), although it can be based on your prior experience (such as in the example where the light won't turn on). In general, hypotheses in biology should be based on a credible, referenced source of information.

A hypothesis must be **testable** to ensure that it is valid. For example, a hypothesis that depends on what a dog thinks is not testable, because we can't tell what a dog thinks. It should also be **falsifiable,** meaning that it can be disproven by experimental results. An example of an unfalsifiable hypothesis is "Red is a better color than blue." There is no experiment that might show this statement to be false. To test a hypothesis, a researcher will conduct one or more experiments designed to eliminate one or more of the hypotheses. This is important: *a hypothesis can be disproven, or eliminated, but it can never be proven.* Science does not deal in proofs like mathematics. If an experiment fails to disprove a hypothesis, then that explanation (the hypothesis) is supported as the answer to the question. However, that doesn't mean that later on, we won't find a better explanation or design a better experiment that will be found to falsify the first hypothesis and lead to a better one.

VARIABLES

A variable is any part of the experiment that can vary or change during the experiment. Typically, an experiment only tests one variable and all the other conditions in the experiment are held constant.

- The variable that is tested is known as the **independent variable**.
- The **dependent variable** is the thing (or things) that you are measuring as the outcome of your experiment.
- A **constant** is a condition that is the same between all of the tested groups.
- A **confounding variable** is a condition that is not held constant that could affect the experimental

results.

A hypothesis often has the format "If [I change the independent variable in this way] then [I will observe that the dependent variable does this] because [of some reason]." For example, the first hypothesis might be, "If you change the light bulb, then the light will turn on because the bulb is burned out." In this experiment, the independent variable (the thing that you are testing) would be changing the light bulb and the dependent variable is whether or not the light turns on. It would be important to hold all the other aspects of the environment constant, for example not messing with the lamp cord or trying to turn the lamp on using a different light switch. If the entire house had lost power during the experiment because a car hit the power pole, that would be a confounding variable.

You may have learned that a hypothesis can be phrased as an "If..then..." statement. Simple hypotheses can be phrased that way (but they must also include a "because"), but more complicated hypotheses may require several sentences. It is also very easy to get confused by trying to put your hypothesis into this format. Hypotheses do not have to be phrased as "if..then.." statements, it is just sometimes a useful format.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=46#h5p-8

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=46#h5p-17

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=46#h5p-18

RESULTS

The **results** of your experiment are the data that you collect as the outcome. In the light experiment, your results are either that the light turns on or the light doesn't turn on. Based on your results, you can make a conclusion. Your **conclusion** uses the results to answer your original question.

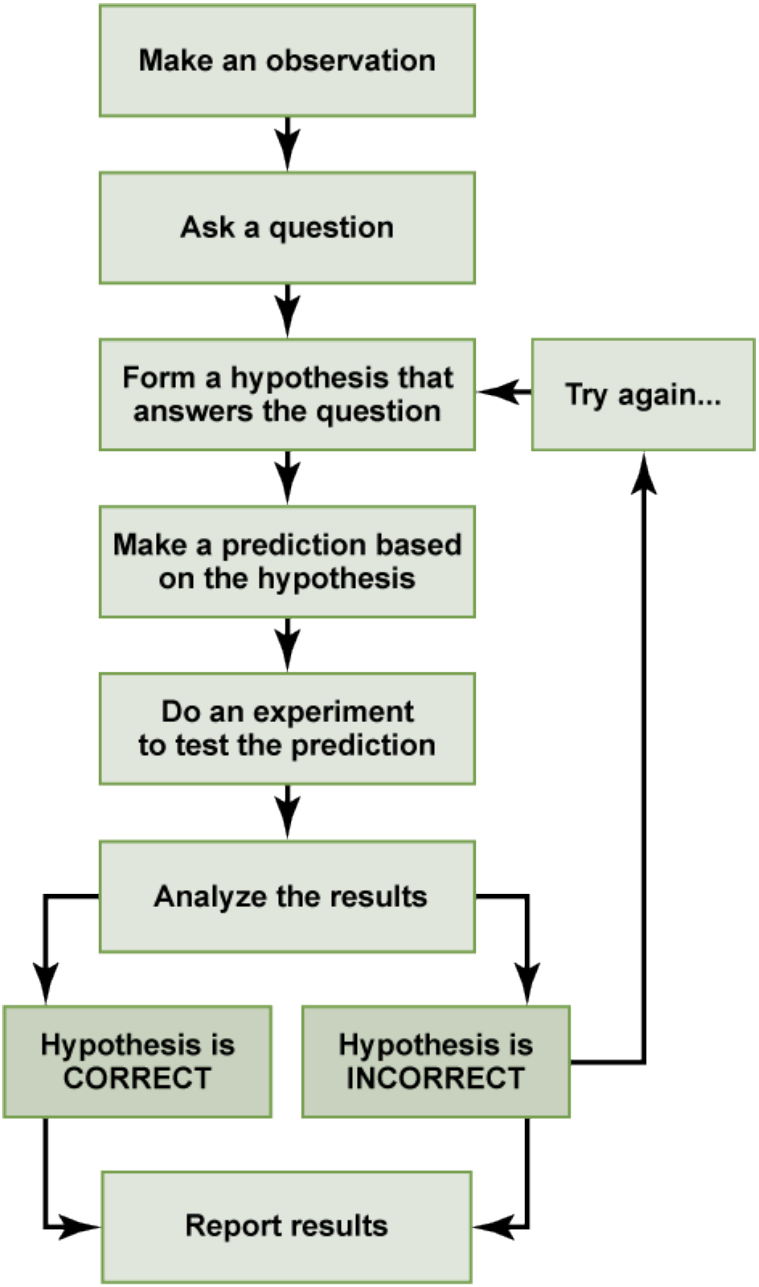

Make an observation
Ask a question
Form a hypothesis that answers the question
Try again...
Make a prediction based on the hypothesis
Do an experiment to test the prediction
Analyze the results
Hypothesis is CORRECT
Hypothesis is INCORRECT
Report results

Figure 4 The basic process of the scientific method. This is what science looks like in a simplified world.

We can put the experiment with the light that won't go in into the figure above:

1. Observation: the light won't turn on.
2. Question: why won't the light turn on?
3. Hypothesis: the lightbulb is burned out.
4. Prediction: if I change the lightbulb (independent variable), then the light will turn on (dependent variable).
5. Experiment: change the lightbulb while leaving all other variables the same.
6. Analyze the results: the light didn't turn on.
7. Conclusion: The lightbulb isn't burned out. The results do not support the hypothesis, time to develop a new one!
8. Hypothesis 2: the lamp is unplugged.
9. Prediction 2: if I plug in the lamp, then the light will turn on.
10. Experiment: plug in the lamp
11. Analyze the results: the light turned on!
12. Conclusion: The light wouldn't turn on because the lamp was unplugged. The results support the hypothesis, it's time to move on to the next experiment!

In practice, the scientific method is not as rigid and structured as it might at first appear. Sometimes an experiment leads to conclusions that favor a change in approach; often, an experiment brings entirely new scientific

questions to the puzzle. Many times, science does not operate in a linear fashion; instead, scientists continually draw inferences and make generalizations, finding patterns as their research proceeds. Scientific reasoning is more complex than the scientific method alone suggests.

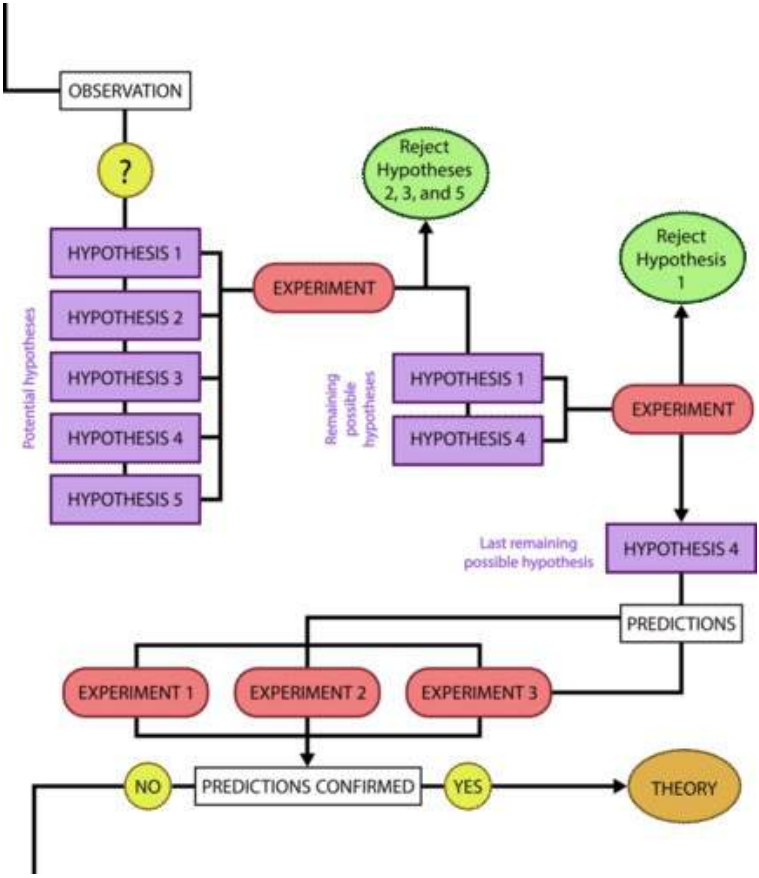


Figure 5 The actual process of using the scientific method. "The general process of scientific investigations" by Laura Guerin, CK-12 Foundation is licensed under CC BY-NC 3.0

CONTROL GROUPS

Another important aspect of designing an experiment is the presence of one or more control groups. A **control group** allows you to make a comparison that is important for interpreting your results. Control groups are samples that help you to determine that differences between your experimental groups are due to your treatment rather than a different variable – they eliminate alternate explanations for your results (including experimental error and experimenter bias). They increase reliability, often through the comparison of control measurements and measurements of the experimental groups. Often, the control group is a sample that is not treated with the independent variable, but is otherwise treated the same way as your experimental sample. This type of control group contains every feature of the experimental group except it is not given the manipulation that is hypothesized about (it does not get treated with the independent variable). Therefore, if the results of the experimental group differ from the control group, the difference must be due to the hypothesized manipulation, rather than some outside factor. It is common in complex experiments (such as those published in scientific journals) to have more control groups than experimental groups.

Example 1

Question: Which fertilizer will produce the greatest number of tomatoes when applied to the plants?

Prediction and Hypothesis: If I apply different brands of fertilizer to tomato plants, the most tomatoes will be produced from plants watered with Brand A because Brand A advertises

that it produces twice as many tomatoes as other leading brands.

Experiment: Purchase 10 tomato plants of the same type from the same nursery. Pick plants that are similar in size and age. Divide the plants into two groups of 5. Apply Brand A to the first group and Brand B to the second group according to the instructions on the packages. After 10 weeks, count the number of tomatoes on each plant.

Independent Variable: Brand of fertilizer.

Dependent Variable: Number of tomatoes.

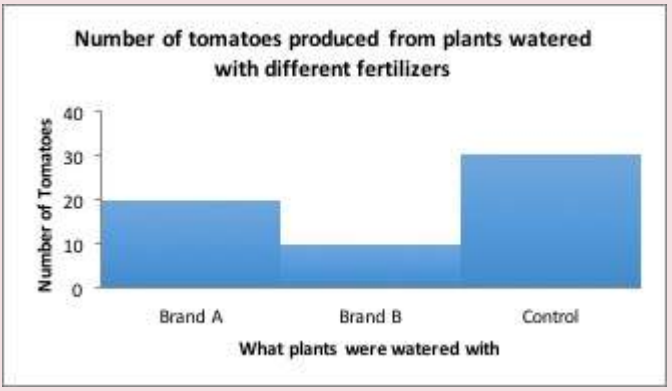
The number of tomatoes produced depends on the brand of fertilizer applied to the plants.

Constants: amount of water, type of soil, size of pot, amount of light, type of tomato plant, length of time plants were grown.

Confounding variables: any of the above that are not held constant, plant health, diseases present in the soil or plant before it was purchased.

Results: Tomatoes fertilized with Brand A produced an average of 20 tomatoes per plant, while tomatoes fertilized with Brand B produced an average of 10 tomatoes per plant.

You'd want to use Brand A next time you grow tomatoes, right? But what if I told you that plants grown without fertilizer produced an average of 30 tomatoes per plant! Now what will you use on your tomatoes?



Results including control group: Tomatoes which received no fertilizer produced more tomatoes than either brand of fertilizer.

Conclusion: Although Brand A fertilizer produced more tomatoes than Brand B, neither fertilizer should be used because plants grown without fertilizer produced the most tomatoes!

Positive control groups are often used to show that the experiment is valid and that everything has worked correctly. You can think of a positive control group as being a group where you should be able to observe the thing that you are measuring ("the thing" should happen). The conditions in a positive control group should guarantee a positive result. If the positive control group doesn't work, there may be something wrong with the experimental procedure.

Negative control groups are used to show whether a treatment had any effect. If your treated sample is the same as your negative control group, your treatment had no effect. You can also think of a negative control group as being a group where you should NOT be able to observe the thing that you are measuring ("the thing" shouldn't happen), or

where you should not observe any change in the thing that you are measuring (there is no difference between the treated and control group). The conditions in a negative control group should guarantee a negative result. A placebo group is an example of a negative control group.

As a general rule, you need a positive control to validate a negative result, and a negative control to validate a positive result.

- **You read an article in the NY Times that says some spinach is contaminated with Salmonella.** You want to test the spinach you have at home in your fridge, so you wet a sterile swab and wipe it on the spinach, then wipe the swab on a nutrient plate (petri plate).
    - *You observe growth*. Does this mean that your spinach is really contaminated? Consider an alternate explanation for growth: the swab, the water, or the plate is contaminated with bacteria. You could use a negative control to determine which explanation is true. If a swab is wet and wiped on a nutrient plate, do bacteria grow?
    - *You don't observe growth.* Does this mean that your spinach is really safe? Consider an alternate explanation for no growth: Salmonella isn't able to grow on the type of nutrient you used in your plates. You could use a positive control to determine which explanation is true. If you wipe a known sample of Salmonella bacteria on the plate, do bacteria grow?

- **In a drug trial, one group of subjects are given a new drug, while a second group is given a placebo drug** (a sugar pill; something which appears like the drug, but doesn't contain the active ingredient). Reduction in disease symptoms are measured. The second group receiving the placebo is a negative control group. You might expect a reduction in disease symptoms purely because the person knows they are taking a drug so they should be getting better. If the group treated with the real drug does not show more a reduction in disease symptoms than the placebo group, the drug doesn't really work. The placebo group sets a baseline against which the experimental group (treated with the drug) can be compared. A positive control group is not required for this experiment.
- In an experiment measuring the **preference of birds for various types of food**, a negative control group would be a "placebo feeder". This would be the same type of feeder, but with no food in it. Birds might visit a feeder just because they are interested in it; an empty feeder would give a baseline level for bird visits. A positive control group might be a food that squirrels are known to like. This would be useful because if no squirrels visited any of the feeders, you couldn't tell if this was because there were no squirrels around or because they didn't like any of your food offerings!
- To test **the effect of pH on the function of an enzyme**, you would want a positive control group where you knew the enzyme would function (pH not changed) and a negative control group where you knew the enzyme would not function (no enzyme added). You need the positive control

group so you know your enzyme is working: if you didn't see a reaction in any of the tubes with the pH adjusted, you wouldn't know if it was because the enzyme wasn't working at all or because the enzyme just didn't work at any of your tested pH. You need the negative control group so you can ensure that there is no reaction taking place in the absence of enzyme: if the reaction proceeds without the enzyme, your results are meaningless.

REFERENCES

Text adapted from: OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:RD6ERYiU@5/The-Process-of-Science.

Presenting Data - Graphs and Tables

TYPES OF DATA

There are different types of data that can be collected in an experiment. Typically, we try to design experiments that collect objective, quantitative data.

Objective data is fact-based, measurable, and observable. This means that if two people made the same measurement with the same tool, they would get the same answer. The measurement is determined by the object that is being measured. The length of a worm measured with a ruler is an objective measurement. The observation that a chemical reaction in a test tube changed color is an objective measurement. Both of these are observable facts.

Subjective data is based on opinions, points of view, or emotional judgment. Subjective data might give two different answers when collected by two different people. The measurement is determined by the subject who is doing the measuring. Surveying people about which of two chemicals smells worse is a subjective measurement. Grading the quality of a presentation is a subjective measurement. Rating your relative happiness on a scale of 1-5 is a subjective measurement. All of these depend on the person who is

making the observation – someone else might make these measurements differently.

Quantitative measurements gather numerical data. For example, measuring a worm as being 5cm in length is a quantitative measurement.

Qualitative measurements describe a quality, rather than a numerical value. Saying that one worm is longer than another worm is a qualitative measurement.

	Quantitative	Qualitative
Objective	The chemical reaction has produced 5cm of bubbles.	The chemical reaction has produced a lot of bubbles.
Subjective	I give the amount of bubbles a score of 7 on a scale of 1-10.	I think the bubbles are pretty.

After you have collected data in an experiment, you need to figure out the best way to present that data in a meaningful way. Depending on the type of data, and the story that you are trying to tell using that data, you may present your data in different ways.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=31#h5p-19

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=31#h5p-20

DATA TABLES

The easiest way to organize data is by putting it into a data table. In most data tables, the independent variable (the variable that you are testing or changing on purpose) will be in the column to the left and the dependent variable(s) will be across the top of the table.

Be sure to:

- Label each row and column so that the table can be interpreted
- Include the units that are being used
- Add a descriptive caption for the table

Example

You are evaluating the effect of different types of fertilizers on plant growth. You plant 12 tomato plants and divide them into three groups, where each group contains four plants. To the first group, you do not add fertilizer and the plants are watered with plain water. The second and third groups are watered with two different brands of fertilizer. After three

weeks, you measure the growth of each plant in centimeters and calculate the average growth for each type of fertilizer.

The effect of different brands of fertilizer on tomato plant growth over three weeks

Treatment	Plant Number				
	1	2	3	4	Av
No treatment	10	12	8	9	9.7
Brand A	15	16	14	12	14
Brand B	22	25	21	27	23

Scientific Method Review: Can you identify the key parts of the scientific method from this experiment?

- Independent variable – Type of treatment (brand of fertilizer)
- Dependent variable – plant growth in cm
- Control group(s) – Plants treated with no fertilizer
- Experimental group(s) – Plants treated with different brands of fertilizer

GRAPHING DATA

Graphs are used to display data because it is easier to see trends in the data when it is displayed visually compared to when it is displayed numerically in a table. Complicated data can often be displayed and interpreted more easily in a graph format than in a data table.

In a graph, the X-axis runs horizontally (side to side) and the Y-axis runs vertically (up and down). Typically, the

independent variable will be shown on the X axis and the dependent variable will be shown on the Y axis (just like you learned in math class!).

LINE GRAPH

Line graphs are the best type of graph to use when you are displaying a change in something over a continuous range. For example, you could use a line graph to display a change in temperature over time. Time is a continuous variable because it can have any value between two given measurements. It is measured along a continuum. Between 1 minute and 2 minutes are an infinite number of values, such as 1.1 minute or 1.93456 minutes.

Changes in several different samples can be shown on the same graph by using lines that differ in color, symbol, etc.

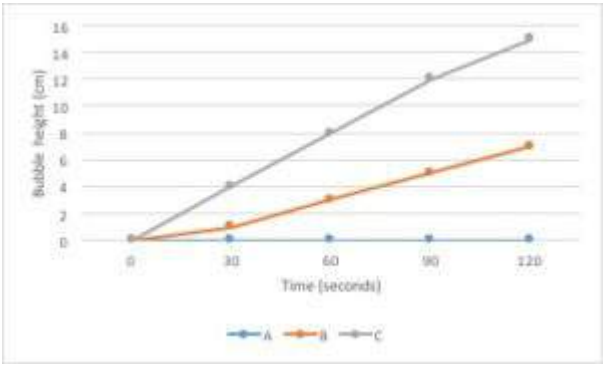


Figure 1: Change in bubble height in centimeters over 120 seconds for three samples containing different amounts of enzyme. Sample A contained no enzyme, sample B contained 1mL of enzyme, sample C contained 2 mL of enzyme.

BAR GRAPH

Bar graphs are used to compare measurements between

different groups. Bar graphs should be used when your data is not continuous, but rather is divided into different categories. If you counted the number of birds of different species, each species of bird would be its own category. There is no value between "robin" and "eagle", so this data is not continuous.

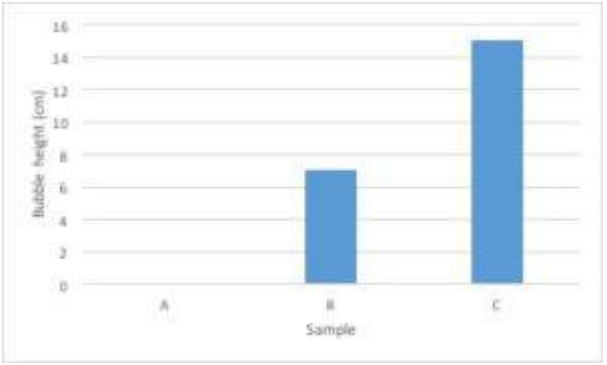


Figure 2: Final bubble height after 120 seconds for three samples containing different combinations of ingredients. Sample A contained enzyme but no substrate, sample B contained substrate but no enzyme, sample C contained substrate and enzyme.

SCATTER PLOT

Scatter Plots are used to evaluate the relationship between two different continuous variables. These graphs compare changes in two different variables at once. For example, you could look at the relationship between height and weight. Both height and weight are continuous variables. You could not use a scatter plot to look at the relationship between number of children in a family and weight of each child because the number of children in a family is not a continuous variable: you can't have 2.3 children in a family.

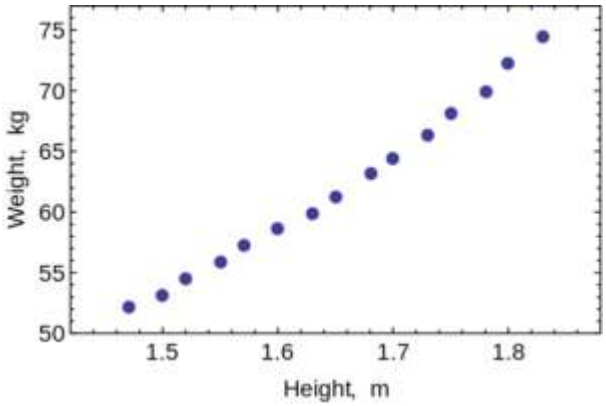


Figure 3: The relationship between height (in meters) and weight (in kilograms) of members of the girls softball team. "OLS example weight vs height scatterplot" by Stpasha is in the Public Domain

AN INTERACTIVE H5P ELEMENT HAS BEEN EXCLUDED FROM THIS VERSION OF THE TEXT. YOU CAN VIEW IT ONLINE HERE: HTTPS://OPENOREGON.PRESSBOOKS.PUB/MHCCMAJORSBIO/?P=31#H5P-11

HOW TO MAKE A GRAPH

1. Identify your independent and dependent variables.
2. Choose the correct type of graph by determining whether each variable is continuous or not.
3. Determine the values that are going to go on the X and Y axis. If the values are continuous, they need to be evenly spaced based on the value.
4. Label the X and Y axis, including units.
5. Graph your data.
6. Add a descriptive caption to your graph. Note that

data tables are titled above the figure and graphs are captioned below the figure.

Example

Let's go back to the data from our fertilizer experiment and use it to make a graph. I've decided to graph only the average growth for the four plants because that is the most important piece of data. Including every single data point would make the graph very confusing.

1. The independent variable is type of treatment and the dependent variable is plant growth (in cm).
2. Type of treatment is not a continuous variable. There is no midpoint value between fertilizer brands (Brand A 1/2 doesn't make sense). Plant growth is a continuous variable. It makes sense to sub-divide centimeters into smaller values. Since the independent variable is categorical and the dependent variable is continuous, this graph should be a bar graph.
3. Plant growth (the dependent variable) should go on the Y axis and type of treatment (the independent variable) should go on the X axis.
4. Notice that the values on the Y axis are continuous and evenly spaced. Each line represents an increase of 5cm.
5. Notice that both the X and the Y axis have labels that include units (when required).
6. Notice that the graph has a descriptive caption that allows the figure to stand alone without additional information given from the procedure: you know

that this graph shows the average of the measurements taken from four tomato plants.

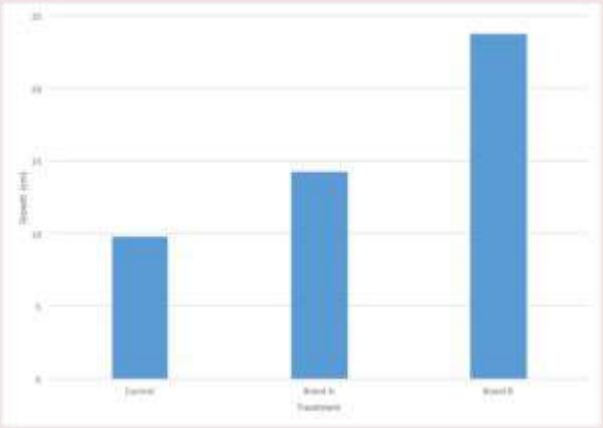


Figure 4: Average growth (in cm) of tomato plants when treated with different brands of fertilizer. There were four tomato plants in each group (n = 4).

DESCRIPTIVE CAPTIONS

All figures that present data should stand alone – this means that you should be able to interpret the information contained in the figure without referring to anything else (such as the methods section of the paper). This means that all figures should have a descriptive caption that gives information about the independent and dependent variable. Another way to state this is that the caption should describe what you are testing and what you are measuring. A good starting point to developing a caption is "the effect of [the independent variable] on the [dependent variable]."

Here are some examples of good caption for figures:

- The effect of exercise on heart rate
- Growth rates of E. coli at different temperatures
- The relationship between heat shock time and transformation efficiency

Here are a few less effective captions:

- Heart rate and exercise
- Graph of E. coli temperature growth
- Table for experiment 1

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=31#h5p-12

Writing for Science

Whether scientific research is basic science or applied science, scientists must share their findings for other researchers to expand and build upon their discoveries. Communication and collaboration within and between sub disciplines of science are key to the advancement of knowledge in science. For this reason, an important aspect of a scientist's work is disseminating results and communicating with peers. Scientists can share results by presenting them at a scientific meeting or conference, but this approach can reach only the limited few who are present. Instead, most scientists present their results in peer-reviewed articles that are published in scientific journals. **Peer-reviewed articles** are scientific papers that are reviewed, usually anonymously by a scientist's colleagues, or peers. These colleagues are qualified individuals, often experts in the same research area, who judge whether or not the scientist's work is suitable for publication. The process of peer review helps to ensure that the research described in a scientific paper or grant proposal is original, significant, logical, and thorough. Grant proposals, which are requests for research funding, are also subject to peer review. Scientists publish their work so other scientists can reproduce their experiments under similar or different conditions to expand on the findings. The experimental results must be consistent with the findings of other scientists.

There are many journals and the popular press that do not use a peer-review system. A large number of online open-access journals, journals with articles available without cost, are now available many of which use rigorous peer-review systems, but some of which do not. Results of any studies published in these forums without peer review are not reliable and should not form the basis for other scientific work. In one exception, journals may allow a researcher to cite a personal communication from another researcher about unpublished results with the cited author's permission.

Scientific articles are not literary works. Instead, they are meant to transmit information effectively and concisely. The need for clarity and brevity is especially important for other forms of science communication such as posters where the audience must be able to understand the significance of your research in just a few minutes, but the need is there for all forms of scientific communication.

There is an explicit format that scientific papers follow, with relatively small variations in style among journals. Papers are broken down into the following sections: title, abstract, introduction, methods, results, and discussion. Every section, except the title, should be labeled as such. Generally the section name is centered and underlined (or bold-faced) over the text. Although posters follow the same format as a paper, each section is abbreviated (once again, clarity is critical).

TITLE

The title should give the reader a concise, informative description of the content and scope of the paper.

ABSTRACT

The abstract is a concise summary of the major findings of the study. It should be no longer than 9-10 sentences. It should summarize every subsequent section of the paper. It should state the purposes of the study, and then briefly summarize the methods, results, and conclusions of the study. The abstract should be able to stand-alone. Do not refer to any figures or tables, or cite any references. Because the abstract is a distillation of the paper, it is often written last. It is typically the hardest part of the paper to write.

INTRODUCTION

In many journals, the introduction is also unlabeled, and simply starts after the abstract.

The introduction gives the rationale for the research. It answers the question "Why should anyone be interested in this work?" It usually includes background information, including the work of others, and a description of your objectives. If you are studying a particular species, give both the scientific (Latin) name and the common name the first time you mention your study animal. The scientific name is always underlined or italicized, and the genus name is capitalized while the species name is not. Cite only references pertinent to your study. Direct quotations are rarely used in scientific writing; instead state the findings of others in your own words. Furthermore, footnotes are rarely used in a scientific paper. Instead cite the author by last name, and the year that the source was published.

Smith (1987) found that male mice prefer the odor of non-pregnant female mice to that of pregnant female mice. Male mice prefer the odor of non- pregnant female mice to that of pregnant female mice (Smith, 1987).

When two people co-author a paper, both are cited:

For instance:

When more than two people co-author a paper, cite only the first author, and refer to the other authors with the Latin phrase, “et al.”, indicating “and others”:

Undergraduate students who came to lectures were more likely to receive a high grade on the exams (Thatcher, et al., 2000).

Harrett and Garrett (1999) found no differences between male and female elephants in their response to the tape of a female vocalization.

The full reference for each work must be given in the literature cited section at the end of the paper. For references, select work from the primary literature: that is, work that is published by the same people who did it. In general, citing an encyclopedia or textbook is not appropriate for a scientific paper.

When organizing your introduction, begin with a general description of the topic, and then become more specific. For example, in a study of the olfaction in the reproductive behavior of mice, the skeleton of the introduction might be:

For reproduction to be successful, animals must be able to correctly assess the reproductive condition of a potential partner. Many different signals have evolved in animals to facilitate such assessment.Olfactory signals seem to be particularly important in mammals.

Mice are particularly suited for studying the role of olfaction in reproductive behavior. Odor cues are involved in several aspects

of mouse reproductive behavior, including... The aim of this study was...

Each of these sentences would be a good topic sentence of a different paragraph in the introduction. References should be cited where appropriate.

In sum, an introduction should convey your overall purpose in conducting the experiment as well as your specific objectives.

METHODS

This section is also often called Materials and Methods. This section is a very concise summary of the subjects, equipment, and procedures used. This section should contain enough information so that someone else could replicate your work. It is NOT a list, but a narrative description. Because it is a narrative, it should not include a list of your materials. Rather, they should be described in the narrative as required. For example, you could say: "We measured 5mL of enzyme solution into a test tube and heated it on a hot plate until it boiled." From this, it is obvious that you used some sort of tool to accurately measure 5mL of solution, as well as a test tube, and a hot plate.

Only include information that is relevant to your experiment: do not include information that any scientist should know to do or that won't affect the results (label the tubes, clean up afterwards, make a graph). If you are following the methods of another paper or a lab manual,

simply cite the source. Then, you can concentrate on describing any changes that you made to that procedure. **A common mistake is to let results creep into this section.**

RESULTS

The results includes presentations of your data and the results of statistical analysis of your data. First, state the overall trend of the data. Did the majority of the data statistically support or contradict the null hypothesis?

Address each statistical test separately, often in separate paragraphs. For each type of data analyzed say whether your results are statistically significant, and in parentheses give the statistical test used, the value of the test statistic, and the probability level for that computed value. For example, "Male mice visited non-pregnant females significantly more often than pregnant females (chi square = 4.69; $p < 0.05$)."

Do not present your raw data. Instead, present data in an easy to read form. You will probably use a figure or a table to present your results. Refer to each table by a number (Table 1, Table 2, etc.) It should have a concise heading at the top. Graphs and diagrams are both called figures and are numbered consecutively (Fig. 1, Fig. 2, etc.) They have headings at the bottom. Axes on graphs should be clearly labeled. See the section on Presenting Your Data for more information.

You must refer to every table and figure at least once in the text. Often this can be done parenthetically: "Male mice visited non-pregnant females significantly more often than pregnant females (chi square = 4.69; $p < 0.05$; Fig. 2)."

Do not use the word "significant" unless it can be supported by statistical evidence.**A common mistake is to discuss the implications of your findings.** Save that for the discussion section.

DISCUSSION

Here you are to give a reader the "take home" message of the study. Begin by briefly summarizing the major findings of your study. Then discuss each finding one at a time (usually in separate paragraphs).

Interpret your results in light of the biology you are studying. Your discussion section should parallel your introduction: if you discussed the role of reproductive biology of the mouse at the beginning of your study, come back to it again here. The paper should come full circle.

Use references throughout your discussion to support your points. Compare your findings with those of similar studies.

Do not make statements that cannot be supported by the data, and be sure none of your conclusions are contradicted by the data. **Discuss unexpected results or possible errors in the experiment, but don't focus on "what didn't work". We all know this was a classroom research project!**

LITERATURE CITED

Each academic discipline uses a different format to cite the references they use. These differences can be dramatic (English vs. Science, for example) or small (Psychology vs. Biology), but they are based on what information is seen as important. In this course, we follow the format of the most biology journals by using CSE format. See the section on Citing Your Sources for more specific information.

GENERAL HINTS

For stylistic hints, browse one of the many books in the library on scientific writing. Remember, being a good writer in

English "121" doesn't mean your skills will translate to science writing without work (though you have a great start!).

Outline your paper. Use **topic sentences** for every paragraph. You should be able to go back and underline each topic sentence after you are finished.

Keep your report as short as you can, consistent with clarity and completeness. Do not "pad" with a lot of irrelevant information just to show you know a lot.

A note on Plagiarism: Plagiarism is a serious academic offense. However, most instances of plagiarism are the result of a lack of care and effort, and not intentional misbehavior. Here is a general rule to follow: **Don't Cut and Paste!** Accidental or not, any occurrence of academic dishonesty will be treated seriously. Ignorance is no defense.

Be sure to proofread for typographical errors, poor grammar, or unclear sentence structure.

Try to start paragraphs with a topic sentence or a summary statement. Then follow it with supporting statements. This technique makes your writing clearer and easier to follow. Ideally, someone could read the first sentence of each paragraph and still understand the gist of our paper.

PLEASE avoid dull scientific writing, particularly the use of the passive voice. As much as possible, use an active voice. Passive writing takes up more space and is dull, dull, dull. Look at the example here; see how this is more exciting and can lead to an interesting ecological observation about the importance of the predator – prey relationship involved?

BAD: Mussels are eaten by sea stars.GOOD: Sea stars eat mussels. BETTER STILL: Sea stars are voracious predators of mussels.

Make sure the object to which words such as "this" or "it" refer is clear.

Combine sentences with low information content into one

sentence. This will make your writing more streamlined and less repetitive. But don't write run-on sentences either!

Always refer to work people have done in the past in the past tense. Refer to species attributes or other on going, continuing states in the present tense.

The word "data" is plural. Say either "these data are…" or "this datum is…"

BAD: Wentworth (1985) studied vegetation in Arizona. He found that tree species distributions followed gradients.GOOD: In the Huachuca Mountains of Arizona, both elevation and the amount of light influenced tree species distributions (Wentworth 1985).

Scientific names of animals and plants are underlined or italicized (as are most Latin words), such as Homo sapiens or *Homo sapiens* (genera and all higher taxa are capitalized, species names are lowercase).

Do not anthropomorphize. A honeybee or a dandelion does not have the same consciousness or emotional life as your roommate. In extreme forms, this type of writing is appropriate for the tabloids in supermarket checkout lines…

Try varying the length of your sentences, and keep in mind that a sentence with 4 words is probably too short, and one with 20 too long.

BAD: Kudzu, an Asian super weed, intends to dominate and conquer the entire southeastern United States.GOOD: Kudzu is a noxious weed introduced from Asia that has quickly spread from its point of introduction throughout the southeastern United States.

Avoid using too many clauses in one sentence. If you see that you have a lot of commas, that is a clue that you've overdone the number of clauses in the sentence.

Try reading your work out loud. Anything that is written poorly will be difficult to read. This technique will alert you to problem areas in your writing.

BE PREPARED TO WRITE SEVERAL DRAFTS! Good, hard

editing will turn you from a mediocre to a good writer. **And with good writing, you are able to show your GREAT thinking!**

SOURCES:

These instructions are adapted by Walter Shriner. Originally from ***Jakob, E.*** *1995. Laboratory manual for animal behavior. Bowling Green University and* ***Muller, K.*** *1991. Ornithology laboratory. University of California, Davis.*

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:RD6ERYiU@5/The-Process-of-Science.

Using Credible Sources

Learning Objectives

Course Objective for this section: Select, evaluate, and utilize discipline-specific information and literature to explore topics.

- Differentiate between questions that can and cannot be answered using science.
- Identify appropriate credible sources of information to research a topic.
- Evaluate sources of information for their strengths and weaknesses.

Science is a very specific way of learning, or knowing, about the world. Humans have used the process of science to learn a huge amount about the way the natural world works. Science is responsible for amazing innovations in medicine, hygiene, and technology. There are however, areas of knowledge and human experience that the methods of science cannot be applied to. These include such things as answering purely moral questions, aesthetic questions, or what can be generally categorized as spiritual questions.

Science has cannot investigate these areas because they are outside the realm of material phenomena, the phenomena of matter and energy, and cannot be observed and measured.

Questions that can be answered using science	Questions that cannot be answered using science
· What is the optimum temperature for the growth of E. coli bacteria? · Do birds prefer bird feeders of a specific color? · What is the cause of this disease? · How effective is this drug in treating this disease?	· How tall is Santa Claus? · Do angels exist? · Which is better: classical music or rock and roll? · What are the ethical implications of human cloning?

Since this is a biology class, we will be focusing on questions that can be answered scientifically. Remember that in the scientific process, observations lead to questions. A scientific question is one that can be answered by using the process of science (testing hypotheses, making observations about the natural world, designing experiments).

Sometimes you will directly make observations yourself about the natural world that lead you to ask scientific questions, other times you might hear or read something that leads you to ask a question. Regardless of how you make your initial observation, you will want to do research about your topic before you start setting up an experiment. When you're learning about a topic, it's important to use credible sources of information.

TYPES OF SOURCES

Whether conducting research in the social sciences, humanities (especially history), arts, or natural sciences, the ability to distinguish between **primary** and **secondary source material** is essential. Basically, this distinction illustrates the degree to which the author of a piece is removed from the actual event being described. This means

whether the author is reporting information *first hand* (or is first to record these immediately following an event), or conveying the experiences and opinions of others—that is, *second hand*. In biology, the distinction would be between the person (or people) who conducted the research and someone who didn't actually do the research, but is merely reporting on it.

PRIMARY SOURCES

These are contemporary accounts of an event, written by someone who experienced or witnessed the event in question. In general, these original documents (i.e., they are not about another document or account) are often diaries, letters, memoirs, journals, speeches, manuscripts, interviews, photographs, audio or video recordings, or original literary or theatrical works.

In science, a "primary source" or the "primary literature" refers to the original publication of a scientist's new data, results, and conclusions. These articles are written for other experts in a specific scientific field.

You've probably done a writing assignment or other project during which you have participated in a **peer review** process. During this process, your project was critiqued and evaluated by people of similar competence to yourself (your peers). This gave you feedback on which to improve your work. Scientific articles typically go through a peer review process before they are published in an academic journal. In this case, the peers who are reviewing the article are other experts in the specific field about which the paper is written. This allows other scientists to critique experimental design, data, and conclusions before that information is published in an academic journal. Often, the scientists who did the experiment and who are trying to publish it are required to do additional work or edit their

paper before it is published. The goal of the scientific peer review process is to ensure that published primary articles contain the best possible science.

SECONDARY SOURCES

The function of a secondary source is to interpret the **primary source.** A secondary source can be described as at least one step removed from the event or phenomenon under review. Secondary source materials interpret, assign value to, conjecture upon, and draw conclusions about the events reported in primary sources. These are usually in the form of published works such as magazine articles or books, but may include radio or television documentaries, or conference proceedings.

POPULAR VS. SCHOLARLY SOURCES

POPULAR	SCHOLARLY
Broad range of topics, presented in shorter articles	Specific, narrowly focused topics in lengthy, in-depth articles
Articles offer overview of subject matter; interpretation, rather than original research; sometimes contain feature articles and reports on current social issues and public opinion	Articles often contain previously unpublished research and detail new developments in field
Intended to attract a general readership without any particular expertise or advanced education	Intended for specialist readership of researchers, academics, students and professionals
Written by staff (not always attributed) or freelance writers using general, popular language	Written by identified specialists and researchers in subject area, usually employing technical, subject-specific language and jargon
Edited and approved for publication in-house (not peer-reviewed)	Critically evaluated by peers (fellow scholars) in field for content, scholarly soundness, and academic value
Articles rarely contain references or footnotes and follow no specific format	Well-researched, documented articles nearly always follow standard format: abstract, introduction, literature review, methodology, results, conclusion, bibliography/references
Designed to attract eye of potential newsstand customers: usually filled with photographs or illustrations, printed on glossier paper	Sober design: mostly text with some tables or graphs accompanying articles; usually little or no photography; negligible, if any, advertising; rarely printed on high-gloss paper
Each issue begins with page number '1'	Page numbers of issues *within a volume (year)* are usually consecutive (i.e., first page of succeeding issue is number following last page number of previous issue)
Presented to entertain, promote point of view, and/or sell products	Intended to present researchers' opinions and findings based on original research
Examples: *Newsweek, Rolling Stone, Vogue*	Examples: *Science*, Nature, *Journal of Microbial and Biochemical Technology*

In science, it is often extremely difficult to read and understand primary articles unless you are an expert in that specific scientific field. Secondary sources are typically easier to read and can give you the important information from

a primary source, but only if the secondary source has interpreted the information correctly! It is always better to go to the primary source if possible because otherwise you are relying on someone else's interpretation of the information. However, it is always better to use a source that you can read and understand rather than a source that you can't. For this reason, it is very important to be able to identify credible secondary sources.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=33#h5p-14

EVALUATING CREDIBILITY

When you write a scientific paper (or any paper, really), you want to back up your statements with credible sources. You will need to identify credible sources to help you research scientific topics to help you develop interesting scientific questions. You will also need sources to help you form a well-educated hypothesis that is not just based on your guess about what will happen. A credible source is one that is trustworthy from which the information can be believed. Credible sources are written by people who are experts in the field (or at least are very knowledgeable) about the subject that they are commenting on.

We will be using a variation of the CRAAP test to help you determine whether or not sources that you find are credible

or not. The CRAAP Test was created by Sarah Blakeslee, of the University of California at Chico's Meriam Library. It is adapted below. When evaluating the credibility of sources using this method, if it's CRAAP, it's good!

You can use the table below to help you evaluate the credibility of your sources.

Credibility Table

Factors to consider	*Least reliable (0 points)*	*Possibly reliable (1 point)*	*Most reliable (2 points)*
Currency	No date of publication or revision given	Outdated for this particular topic	Recently published or revised
Reliable source	Unreliable website, no additional info available	Possibly reliable	Official government or organization, institutional sites, academic journals
Author	No author is given / the author is not qualified to write about this topic	Author is educated on topic or is staff of an organization assumed to be knowledgeable on this specific topic	Specifically identified expert in this field with degrees / credentials in this subject
Accuracy	No review process and information is not supported by evidence from cited sources	The information may have been reviewed or edited by someone knowledgeable in the field. It mentions but does not directly cite other sources	The information has been peer reviewed and is supported by evidence from cited credible sources
Purpose	Obviously biased or trying to sell you something	Sponsored source; may present unbalanced information	Balanced, neutral, presents all sides of the issue fully

In general, do not use a source if it doesn't pass the CRAAP test! For our purposes, do not use any sources that score less than 6 points using the credibility table.

Several examples are given below for sources that you might come across if you were researching the topic of vaccine safety.

Example 1:

CDC (Centers for Disease Control and Prevention). Aug 28,

2015. Vaccine Safety [Internet]. [cited May 12, 2016]. Available from: http://www.cdc.gov/vaccinesafety/index.html

	Score	Discussion – why did you give that score?
Currency	2	Aug 28 2015 is recent and shows that this information is updated frequently.
Reliable source	2	I looked at the "about this organization" and learned that the CDC is a major government organization that works to protect Americans from health, safety, and security threats. They are a division of the US department of health and human services.
Author	1	A specific author was not identified, but the page states that the content is from the CDC, which suggests that it was written by a knowledgeable staff member.
Accuracy	1.5	No information is given about the review process, but it was probably edited by staff at the CDC. There is a list of citations and links to primary scientific articles supporting the information.
Purpose	2	The point of view does not appear to be biased because it seems to be presenting factual information. Admittedly, it only presents the pro-vaccine side of the argument. There are no ads on the page or other information trying to change the reader's viewpoint.
Credibility Score	8.5/ 10	This seems like an excellent source to use for research. It's readable and I could look at the primary articles if I wanted to check them out.

Example 2:

Stop Mandatory Vaccination. N.d.. The Dangers of Vaccines and Vaccinations [Internet]. [cited May 12, 2016]. Available from: http://www.stopmandatoryvaccination.com/vaccine-dangers/

	Score	Discussion – why did you give that score?
Currency	1	The copyright is given as 2015, but there is no date for this specific article. It does reference something that took place in 2015, so it is likely written after that.
Reliable source	0	The "About" page states that the organization was started by Larry Cook using a GoFundMe platform
Author	0	Larry Cook has been devoted to the natural lifestyle for 25 years, but doesn't appear to have any degrees or specific expertise on this topic. Other contributing authors include Landee Martin, who has a Bachelor's of Science in Psychology (which isn't related to vaccine safety), and Brittney Kara, who is a mother who has studied holistic living for the last 17 years. None of the individuals specifically identified on the website appear to be experts in the field.
Accuracy	0.5	It seems unlikely that there is any review process. There are links to several sources, but none of them appear to be primary scientific articles. Several are links to interviews.
Purpose	0	This source is extremely biased. Even the name of the website is biased. There is a link to donate to the webpage. There are at least 10 ads for anti-vaccine books and websites.
Credibility Score	1.5/ 10	I would not want to use this source to research this topic. It's extremely biased and doesn't seem to offer much evidence for its assertions.

CITING YOUR SOURCES

One of the goals for any class is to help students become better scholars. And, one of the important skills of scholarship is proper citation of resources used. Citations demonstrate your "credentials" as a scholar, and provide a resource to your readers of good reference material.

WHY DO YOU HAVE TO CITE YOUR SOURCES?

No research paper is complete without a list of the sources that you used in your writing. Scholars are very careful to keep accurate records of the resources they've used, and of the ideas and concepts they've quoted or used from others. This record keeping is generally presented in the form of citations.

A citation is a description of a book, article, URL, etc. that provides enough information so that others can locate the source you used themselves. It allows you to credit the authors of the sources you use and clarify which ideas belong to you and which belong to other sources. And providing a citation or reference will allow others to find and use these sources as well. Most research papers have a list of citations or cited references and there are special formatting guidelines for different types of research.

However, there are many "proper" formats because each discipline has its own rules. In general we ask only that you use one of the "official" formats and that you use it consistently. To understand what we mean by "consistent", compare the citations in two scientific journals. You will notice that each journal has its own rules for whether an article title is in quotes, bold, underlined, etc., but within each journal the rule applies to all reference citations. Below is a condensed guide to the general format used in science (CSE). For more detailed information consult one of the online citation guides and generators.

PLAGIARISM

Plagiarism is presenting the words or ideas of someone else as your own without proper acknowledgment of the source. When you work on a research paper you will probably find supporting material for your paper from works by others. It's okay to quote people and use their ideas, but you do need to correctly credit them. Even when you summarize or paraphrase information found in books, articles, or Web pages, you must acknowledge the original author. To avoid plagiarism, include a reference to any material you use that provides a fact not commonly known, or whenever you use information from another author. In short, if you didn't collect the data or reach the conclusion on your own, cite it!

These are all examples of plagiarism:

- Buying or using a term paper written by someone else.
- Cutting and pasting passages from the Web, a book, or an article and insert them into your paper without citing them. **Warning!** It is now easy for your instructors to search and identify passages that you have copied from the Web.
- Using the words or ideas of another person without citing them.
- Paraphrasing that person's words without citing them.

Tips for Avoiding Plagiarism:

- First, use your own ideas—it should be your paper and your ideas should be the focus.
- Use the ideas of others sparingly—only to support or reinforce your own argument.
- When taking notes, include complete citation information for each item you use.
- Use quotation marks when directly stating another person's words. Quotes are not frequently used in scientific writing unless you are directly quoting someone's spoken words.

CITING SOURCES IN CSE FORMAT

The Council of Science Editors (CSE) citation format is commonly used in scientific writing. CSE format emphasizes the information that is important when writing scientifically: who wrote the information and when they wrote it. In

different fields, there is an emphasis on different types of information. In the humanities, MLA format is commonly used. This style emphasizes the author's name and the page number. This information allows a reader to track down the exact quotes that are being discussed. Another commonly used format, APA, emphasizes the author's name and the year the information was published.

The standard format for citing a source in science writing is the Name-year format. In this format, the first author's last name is followed by the date. For example: *Not all populations of alligators in the everglades are at risk from habitat loss (Nicholson, 2002).*

If you are not familiar with the CSE citation style, you can get additional information and examples at http://writing.wisc.edu/Handbook/DocCSE_NameYear.html

Beware of computerized "citation creators." While they can get you part way to a correct citation, they rarely are 100% correct. For example, they often fail to put the last name first.

Citing a scientific journal article

Author's last name first initial, next author's last name first initial. Date published. Title of Article. Journal Name. Volume (issue): pages.

Please note that you need to cite the JOURNAL, not the DATABASE that you got it from. Citing the database in which you found a scientific journal article is like citing Google for an internet resource that you are using.

Flores-Cruz Z, Allen C. 2011. Necessity of OxyR for the hydrogen peroxide stress response and full virulence

in *Ralstonia solanacearum*. Appl Environ Microbiol. 77(18):6426-6432.

Werling BP, Lowenstein DM, Straub CS, Gratton C. 2012. Multi-predator effects produced by functionally distinct species vary with prey density. J Insect Sci; 12(30): 346-378.

Shriner, W.M. 1998. Yellow-bellied marmot and golden-mantled ground squirrel responses to heterospecific alarm calls. Animal Behaviour 55:529-536.

Citing an internet resource

Author's last name, first initial. Date published. Title of Website [Internet]. Publisher information. [cited on date that you accessed the information]. Available from: URL where you accessed the source.

Williamson RC. 2004. Deciduous tree galls [Internet]. Madison (WI): University of Wisconsin-Madison; [cited 2013 Sep 12]. Available from http://labs.russell.wisc.edu/pddc/files/Fact_Sheets/FC_PDF/Deciduous_Tree_Galls.pdf

[BP] The Biology Project. 2003. The chemistry of amino acids [Internet]. University of Arizona; [cited 2004 Mar 17]. Available from: http://www.biology.arizona.edu/biochemistry/problem_sets/aa/aa.html

Hilton-Taylor C, compiler. 2000. 2000 IUCN red list of threatened species [Internet]. Gland (Switzerland) and Cambridge (UK): IUCN; [cited 2002 Feb 12]. Available from: http://www.redlist.org/.

AN INTERACTIVE H5P ELEMENT HAS BEEN EXCLUDED FROM THIS VERSION OF THE TEXT. YOU CAN VIEW IT ONLINE HERE: HTTPS://OPENOREGON.PRESSBOOKS.PUB/

MHCCMAJORSBIO/?P=33#H5P-15

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=33#h5p-16

CITING SOURCES WITHIN TEXT

We will be using the CSE Name-year format for citations. When you want to provide a citation reference for a statement that you are making, you should end the sentence with (First author's last name, year). If the article was written by an organization and not a specific author, you can use the name of the organization (or an abbreviation for the name).

Example: Sickle cell anemia is caused by abnormally-shaped haemoglobin proteins (NIH, 2012).

In the References Cited section (a.k.a. Literature cited…) list all the sources you cited in your paper, but do not include any items that you did not specifically cite within the body of your paper or project, even if you read them! Except in rare instances, do not cite a reference that you have not personally read.

You should then list all your references in the Literature Cited section alphabetically by author's last name.

For more information and lots of examples of what to do in specific instances, please visit http://writing.wisc.edu/Handbook/DocCSE_NameYear.html

SOURCES:

CHEMISTRY FOR BIOLOGY

Learning Outcomes

Course Outcomes for this section:

- Describe the structure of biologically-important molecules (carbohydrates, lipids, proteins, nucleic acids, water) and how their structure leads to their function.

Living things are highly organized and structured, following a hierarchy that can be examined on a scale from small to large. The examination of the smallest parts involves a knowledge of chemistry. We can put the levels of organization of living things in order from smallest to largest.

- The **atom** is the smallest and most fundamental unit of matter. It consists of a nucleus surrounded by electrons.
- Atoms form **molecules.** A molecule is a chemical structure consisting of at least two atoms held together by one or more chemical bonds. Many

molecules that are biologically important are **macromolecules,** large molecules that are typically formed by polymerization (a polymer is a large molecule that is made by combining smaller units called monomers, which are simpler than macromolecules).

These first 2 levels (or 3, depending on how you categorize macromolecules) are typically studied in chemistry or biochemistry courses. However, a working knowledge of atoms and molecules is required to understand how these small pieces work to make larger, living organisms.

- Some cells contain aggregates of macromolecules surrounded by membranes; these are called **organelles.** Organelles are small structures that exist within cells.
- All living things are made of **cells;** the cell itself is the smallest fundamental unit of structure and function in living organisms.
- In larger multicellular organisms, cells combine to make **tissues,** which are groups of similar cells carrying out similar or related functions.
- **Organs** are collections of tissues grouped together performing a common function. An organ system is a higher level of organization that consists of functionally related organs.
- Mammals have many **organ systems**. For instance, the circulatory system transports blood through the body and to and from the lungs; it includes organs such as the heart and blood vessels.
- **Organisms** are individual living entities. For example, each tree in a forest is an organism. Single-celled prokaryotes and single-celled

eukaryotes are also considered organisms and are typically referred to as microorganisms.

Once we move beyond one single organism, we have reached the study of ecology. You'll look at these topics in BI213.

- All the individuals of a species living within a specific area are collectively called a **population.** For example, a forest may include many pine trees. All of these pine trees represent the population of pine trees in this forest. Different populations may live in the same specific area. For example, the forest with the pine trees includes populations of flowering plants and also insects and microbial populations.
- A **community** is the sum of populations inhabiting a particular area. For instance, all of the trees, flowers, insects, and other populations in a forest form the forest's community. The forest itself is an ecosystem.
- An **ecosystem** consists of all the living things in a particular area together with the abiotic, non-living parts of that environment such as nitrogen in the soil or rain water.
- At the highest level of organization, the **biosphere** is the collection of all ecosystems, and it represents the zones of life on earth. It includes land, water, and even the atmosphere to a certain extent.

A flow chart shows the hierarchy of living organisms. From smallest to largest, this hierarchy includes: (1) Organelles, such as nuclei, that exist inside cells. (2) Cells, such as a red blood cell. (3) Tissues, such as human skin tissue. (4) Organs such as the stomach make up the human digestive system, an example of an organ system. (5) Organisms, populations, and communities. In a forest, each pine tree is an organism. Together, all the pine trees make up a population. All the plant and animal species in the forest comprise a community. (6) Ecosystems: the coastal ecosystem in the Southeastern United States includes living organisms and the environment in which they live. (7) The biosphere: encompasses all the ecosystems on Earth.

Figure 1 *The biological levels of organization of living things are shown. From a single organelle to the entire biosphere, living organisms are parts of a highly structured hierarchy. (credit "organelles": modification of work by Umberto Salvagnin; credit "cells": modification of work by Bruce Wetzel, Harry Schaefer/ National Cancer Institute; credit "tissues": modification of work by Kilbad; Fama Clamosa; Mikael Häggström; credit "organs": modification of work by Mariana Ruiz Villareal; credit "organisms": modification of work by "Crystal"/Flickr; credit "ecosystems": modification of work by US Fish and Wildlife Service Headquarters; credit "biosphere": modification of work by NASA)*

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 25, 2017 https://cnx.org/contents/GFy_h8cu@10.99:gNLp76vu@13/Themes-and-Concepts-of-Biology

Atoms

An **atom** is the smallest component of an element that retains all of the chemical properties of that element. For example, one hydrogen atom has all of the properties of the element hydrogen, such as it exists as a gas at room temperature, and it bonds with oxygen to create a water molecule. Hydrogen atoms cannot be broken down into anything smaller while still retaining the properties of hydrogen. If a hydrogen atom were broken down into subatomic particles, it would no longer have the properties of hydrogen.

All atoms contain **protons, electrons,** and **neutrons** (Figure 1). The only exception is hydrogen (H), which is made of one proton and one electron. A proton is a positively charged particle that resides in the nucleus (the core of the atom) of an atom and has a mass of 1 and a charge of +1. An electron is a negatively charged particle that travels in the space around the nucleus. In other words, it resides outside of the nucleus. It has a negligible mass and has a charge of –1.

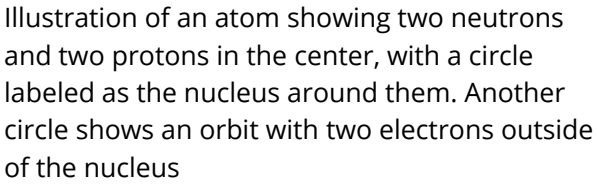


Figure 1 Atoms are made up of protons and neutrons located within the nucleus, and electrons surrounding the nucleus.

Neutrons, like protons, reside in the nucleus of an atom. They have a mass of 1 and no charge. The positive (protons) and negative (electrons) charges balance each other in a neutral atom, which has a net zero charge.

Because protons and neutrons each have a mass of 1, the mass of an atom is equal to the number of protons and neutrons of that atom. The number of electrons does not factor into the overall mass, because their mass is so small.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=56#h5p-68

At the most basic level, all organisms are made of a combination of **elements.** An element is a substance whose atoms all have the same number of protons. They contain atoms that combine together to form molecules. In multicellular organisms, such as animals, molecules can interact to form cells that combine to form tissues, which make up organs. These combinations continue until entire multicellular organisms are formed.

Each element has its own unique properties. Each contains a different number of protons and neutrons, giving it its own atomic number and mass number. The atomic number of an element is equal to the number of protons that element contains. The mass number, or atomic mass, is the number of protons plus the number of neutrons of that element. Therefore, it is possible to determine the number of neutrons by subtracting the atomic number from the mass number.

These numbers provide information about the elements and how they will react when combined. Different elements have different melting and boiling points, and are in different states (liquid, solid, or gas) at room temperature. They also combine in different ways. Some form specific types of bonds, whereas others do not. How they combine is based on the number of electrons present. Because of these characteristics, the elements are arranged into the periodic table of elements, a chart of the elements that includes the atomic number and relative atomic mass of each element. The periodic table also provides key information about the properties of elements (Figure 2) —often indicated by color-coding. The arrangement of the table also shows how the electrons in each element are organized and provides important details about how atoms will react with each other to form molecules.

Isotopes are different forms of the same element that have the same number of protons, but a different number of neutrons. Some elements, such as carbon, potassium, and uranium, have naturally occurring isotopes. Carbon-12, the most common isotope of carbon, contains six protons and six neutrons. Therefore, it has a mass number of 12 (six protons and six neutrons) and an atomic number of 6 (which makes it carbon). Carbon-14 contains six protons and eight neutrons. Therefore, it has a mass number of 14 (six protons and eight neutrons) and an atomic number of 6, meaning it is still the element carbon. These two alternate forms of

carbon are isotopes. Some isotopes are unstable and will lose protons, other subatomic particles, or energy to form more stable elements. These are called radioactive isotopes or radioisotopes.

Periodic table of elements.

Figure 2 Arranged in columns and rows based on the characteristics of the elements, the periodic table provides key information about the elements and how they might interact with each other to form molecules. Most periodic tables provide a key or legend to the information they contain.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=56#h5p-69

Evolution in Action

Carbon Dating: Carbon-14 (^{14}C) is a naturally occurring

radioisotope that is created in the atmosphere by cosmic rays. This is a continuous process, so more ^{14}C is

always being created. As a living organism develops, the relative level of ^{14}C in its body is equal to the concentration of ^{14}C in the

atmosphere. When an organism dies, it is no longer ingesting ^{14}C, so the ratio will decline. ^{14}C decays to ^{14}N by a process

called beta decay; it gives off energy in this slow process.

After approximately 5,730 years, only one-half

of the starting concentration of ^{14}C will have been converted to ^{14}N. The time it takes for half of the original concentration

of an isotope to decay to its more stable form is called its half-life. Because the half-life of ^{14}C is long, it is used to age

formerly living objects, such as fossils. Using the ratio of the ^{14}C concentration found in an object to the amount of ^{14}C

detected in the atmosphere, the amount of the isotope that has not yet decayed can be determined. Based on this

amount, the age of the fossil can be calculated to about 50,000 years (Figure 3). Isotopes with longer half-lives, such

as potassium-40, are used to calculate the ages of older fossils. Through the use of carbon dating,

scientists can reconstruct the ecology and biogeography of organisms living within the past 50,000 years.

Photograph shows scientists digging pygmy mammoth skeleton fossils from the ground.

Figure 3 The age of remains that contain carbon and are less than about 50,000 years old, such as this pygmy mammoth, can be determined

using carbon dating. (credit: Bill Faulkner/ NPS)

An

interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=56#h5p-70

REFERENCES

Unless otherwise noted, images on this page are licensed under CC-BY 4.0 by OpenStax.

OpenStax, Concepts of Biology.

OpenStax CNX. March 22, 2017 https://cnx.org/contents/s8Hh0oOc@9.21:IBRqRY3C@8/The-Building-Blocks-of-Molecules

Chemical Bonds

How elements interact with one another depends on how their electrons are arranged and how many openings for electrons exist at the outermost region where electrons are present in an atom. Electrons exist at energy levels that form shells around the nucleus. The closest shell can hold up to two electrons. The closest shell to the nucleus is always filled first, before any other shell can be filled. Hydrogen has one electron; therefore, it has only one spot occupied within the lowest shell. Helium has two electrons; therefore, it can completely fill the lowest shell with its two electrons. If you look at the periodic table, you will see that hydrogen and helium are the only two elements in the first row. This is because they only have electrons in their first shell. Hydrogen and helium are the only two elements that have the lowest shell and no other shells.

The second and third energy levels can hold up to eight electrons. The eight electrons are arranged in four pairs and one position in each pair is filled with an electron before any pairs are completed.

Looking at the periodic table again (Figure 1), you will notice that there are seven rows. These rows correspond to the number of shells that the elements within that row have. The elements within a particular row have increasing numbers of electrons as the columns proceed from left to right. Although each element has the same number of shells,

not all of the shells are completely filled with electrons. If you look at the second row of the periodic table, you will find lithium (Li), beryllium (Be), boron (B), carbon (C), nitrogen (N), oxygen (O), fluorine (F), and neon (Ne). These all have electrons that occupy only the first and second shells. Lithium has only one electron in its outermost shell, beryllium has two electrons, boron has three, and so on, until the entire shell is filled with eight electrons, as is the case with neon.

Not all elements have enough electrons to fill their outermost shells, but an atom is at its most stable when all of the electron positions in the outermost shell are filled. Because of these vacancies in the outermost shells, we see the formation of chemical bonds, or interactions between two or more of the same or different elements that result in the formation of molecules. To achieve greater stability, atoms will tend to completely fill their outer shells and will bond with other elements to accomplish this goal by sharing electrons, accepting electrons from another atom, or donating electrons to another atom. Because the outermost shells of the elements with low atomic numbers (up to calcium, with atomic number 20) can hold eight electrons, this is referred to as the octet rule. An element can donate, accept, or share electrons with other elements to fill its outer shell and satisfy the octet rule.

When an atom does not contain equal numbers of protons and electrons, it is called an **ion.** Because the number of electrons does not equal the number of protons, each ion has a net charge. Positive ions are formed by losing electrons and are called cations. Negative ions are formed by gaining electrons and are called anions.

For example, sodium only has one electron in its outermost shell. It takes less energy for sodium to donate that one electron than it does to accept seven more electrons to fill the outer shell. If sodium loses an electron, it now has

11 protons and only 10 electrons, leaving it with an overall charge of +1. It is now called a sodium ion.

The chlorine atom has seven electrons in its outer shell. Again, it is more energy-efficient for chlorine to gain one electron than to lose seven. Therefore, it tends to gain an electron to create an ion with 17 protons and 18 electrons, giving it a net negative (–1) charge. It is now called a chloride ion. This movement of electrons from one element to another is referred to as electron transfer. As Figure 1 illustrates, a sodium atom (Na) only has one electron in its outermost shell, whereas a chlorine atom (Cl) has seven electrons in its outermost shell. A sodium atom will donate its one electron to empty its shell, and a chlorine atom will accept that electron to fill its shell, becoming chloride. Both ions now satisfy the octet rule and have complete outermost shells. Because the number of electrons is no longer equal to the number of protons, each is now an ion and has a +1 (sodium) or –1 (chloride) charge.

Diagram shows electron transfer between elements.

Figure 1 Elements tend to fill their outermost shells with electrons. To do this, they can either donate or accept electrons from other elements.

IONIC BONDS

There are four types of bonds or interactions: ionic, covalent, hydrogen bonds, and van der Waals interactions. Ionic and covalent bonds are strong interactions that require a larger energy input to break apart. When an element donates an electron from its outer shell, as in the sodium atom example above, a positive ion is formed (Figure 2). The element accepting the electron is now negatively charged. Because positive and negative charges attract, these ions stay together and form an **ionic bond**, or a bond between ions.

The elements bond together with the electron from one element staying predominantly with the other element. When Na^+ and Cl^- ions combine to produce NaCl, an electron from a sodium atom stays with the other seven from the chlorine atom, and the sodium and chloride ions attract each other in a lattice of ions with a net zero charge.

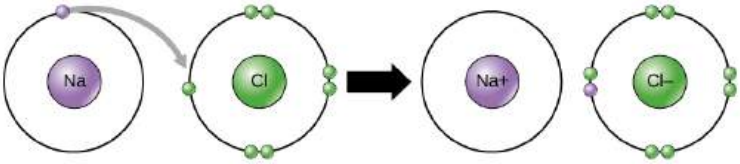


Figure 2 In the formation of an ionic compound, metals lose electrons and nonmetals gain electrons to achieve an octet.

COVALENT BONDS

Another type of strong chemical bond between two or more atoms is a **covalent bond**. These bonds form when an electron is shared between two elements and are the strongest and most common form of chemical bond in living organisms. Covalent bonds form between the elements that make up the biological molecules in our cells. Unlike ionic bonds, covalent bonds do not dissociate in water.

Interestingly, chemists and biologists measure bond strength in different ways. Chemists measure the absolute strength of a bond (the theoretical strength) while biologists are more interested in how the bond behaves in a biological system, which is usually **aqueous** (water-based). In water, ionic bonds come apart much more readily than covalent bonds, so biologists would say that they are weaker than covalent

bonds. If you look in a chemistry textbook, you'll see something different. This is a great example of how the same information can lead to different answers depending on the perspective that you're viewing it from.

The hydrogen and oxygen atoms that combine to form water molecules are bound together by covalent bonds. The electron from the hydrogen atom divides its time between the outer shell of the hydrogen atom and the incomplete outer shell of the oxygen atom. To completely fill the outer shell of an oxygen atom, two electrons from two hydrogen atoms are needed, hence the subscript "2" in H_2O. The electrons are shared between the atoms, dividing their time between them to "fill" the outer shell of each. This sharing is a lower energy state for all of the atoms involved than if they existed without their outer shells filled.

There are two types of covalent bonds: polar and nonpolar. **Nonpolar covalent bonds** form between two atoms of the same element or between different elements that share the electrons equally. For example, an oxygen atom can bond with another oxygen atom to fill their outer shells. This association is **nonpolar** because the electrons will be equally distributed between each oxygen atom. Two covalent

bonds form between the two oxygen atoms because oxygen requires two shared electrons to fill its outermost shell. Nitrogen atoms will form three covalent bonds (also called triple covalent) between two atoms of nitrogen because each nitrogen atom needs three electrons to fill its outermost shell. Another example of a nonpolar covalent bond is found in the methane (CH_4) molecule. The carbon atom has four electrons in its outermost shell and needs four more to fill it. It gets these four from four hydrogen atoms, each atom providing one. These elements all share the electrons equally, creating four nonpolar covalent bonds (Figure 3).

In a **polar covalent bond**, the electrons shared by the atoms spend more time closer to one nucleus than to the other nucleus. Because of the unequal distribution of electrons between the different nuclei, a slightly positive ($\delta+$) or slightly negative ($\delta-$) charge develops. The covalent bonds between hydrogen and oxygen atoms in water are polar covalent bonds. The shared electrons spend more time near the oxygen nucleus, giving it a small negative charge, than they spend near the hydrogen nuclei, giving these molecules a small positive charge.

Diagram depicting polar and nonpolar covalent bonds

Figure 3 The water molecule (left) depicts a polar bond with a slightly positive charge on the hydrogen atoms and a slightly negative charge on the oxygen. Examples of nonpolar bonds include methane (middle) and oxygen (right).

HYDROGEN BONDS

Ionic and covalent bonds are strong bonds that require considerable energy to break. However, not all bonds between elements are ionic or covalent bonds. Weaker bonds can also form. These are attractions that occur between positive and negative charges that do not require much energy to break. Two weak bonds that occur frequently are hydrogen bonds and van der Waals interactions. These bonds give rise to the unique properties of water and the unique structures of DNA and proteins.

When polar covalent bonds containing a hydrogen atom form, the hydrogen atom in that bond has a slightly positive charge. This is because the shared electron is pulled more strongly toward the other element and away from the hydrogen nucleus. Because the hydrogen atom is slightly positive (δ+), it will be attracted to neighboring negative partial charges (δ–). When this happens, a weak interaction occurs between the δ+ charge of the hydrogen atom of one molecule and the δ– charge of the other molecule. This interaction is called a hydrogen bond. This type of bond is common; for example, the liquid nature of water is caused by the hydrogen bonds between water molecules (Figure 4). Hydrogen bonds give water the unique properties that sustain life. If it were not for hydrogen bonding, water would be a gas rather than a liquid at room temperature.

Diagram showing hydrogen bonds formed between adjacent water molecules.

Figure 4 *Hydrogen bonds form between slightly positive (δ+) and slightly negative (δ–) charges of polar covalent molecules, such as water.*

Hydrogen bonds can form between different molecules and they do not always have to include a water molecule. Hydrogen atoms in polar bonds within any molecule can form bonds with other adjacent molecules. For example, hydrogen bonds hold together two long strands of DNA to give the DNA molecule its characteristic double-stranded structure. Hydrogen bonds are also responsible for some of the three-dimensional structure of proteins.

VAN DER WAALS INTERACTIONS

Like hydrogen bonds, van der Waals interactions are weak attractions or interactions between molecules. They occur between polar, covalently bound, atoms in different molecules. Some of these weak attractions are caused by temporary partial charges formed when electrons move around a nucleus. These weak interactions between molecules are important in biological systems.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=61#h5p-72

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=61#h5p-71

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. March 22, 2017 https://cnx.org/contents/s8Hh0oOc@9.21:IBRqRY3C@8/The-Building-Blocks-of-Molecul

Water

Figure 1 *Water: without it, life wouldn't exist. Photo credit ronymichaud; CC0 license; https://pixabay.com/en/users/ronymichaud-647623/*

Do you ever wonder why scientists spend time looking for water on other planets? It is because water is essential to life; even minute traces of it on another planet can indicate that life could or did exist on that planet. Water is one of the more abundant molecules in living cells and the one most critical to life as we know it. Approximately 60–70 percent of your body is made up of water. Without it, life simply would not exist.

WATER IS POLAR

The hydrogen and oxygen atoms within water molecules form **polar covalent bonds**. The shared electrons spend more time associated with the oxygen atom than they do with hydrogen atoms. There is no overall charge to a water molecule, but there is a slight positive charge on each hydrogen atom and a slight negative charge on the oxygen atom. Because of these charges, the slightly positive hydrogen atoms repel each other and form the unique shape. Each water molecule attracts other water molecules because of the positive and negative charges in the different parts of the molecule.

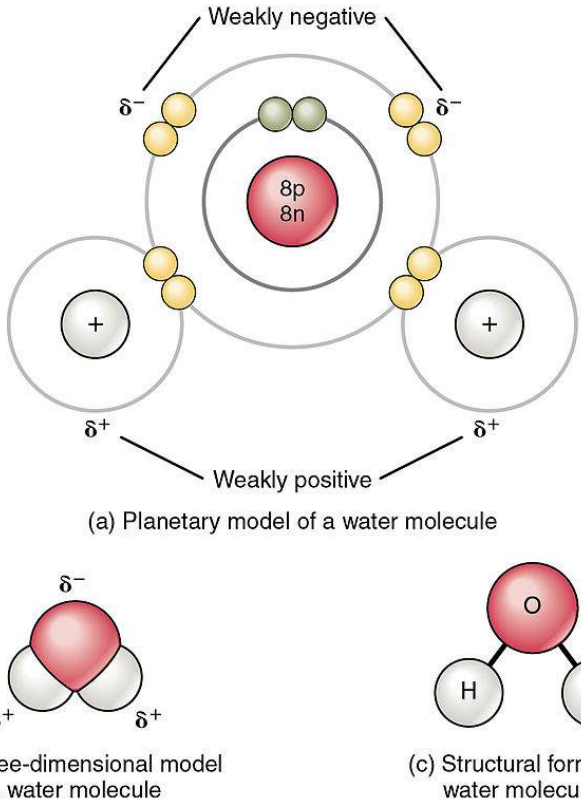


Figure 2 The electrons in the covalent bond connecting the two hydrogens to the atom of oxygen in a water molecule spend more time on the oxygen atom. This gives the oxygen atom a slightly negative charge (since electrons are negatively charged). Credit Anatomy & Physiology, Connexions Web site. http://cnx.org/content/col11496/1.6/, Jun 19, 2013.

Water also attracts other polar molecules (such as sugars), forming hydrogen bonds. When a substance readily forms hydrogen bonds with water, it can dissolve in water and is referred to as **hydrophilic** ("water-loving"). Hydrogen bonds are not readily formed with nonpolar substances like oils and

fats (Figure 3). These nonpolar compounds are **hydrophobic** ("water-fearing") and will not dissolve in water.

Picture of oil in water.

Figure 3 As this macroscopic image of oil and water show, oil is a nonpolar compound and, hence, will not dissolve in water. Oil and water do not mix. (credit: Gautam Dogra)

WATER STABILIZES TEMPERATURE

The hydrogen bonds in water allow it to absorb and release heat energy more slowly than many other substances. Temperature is a measure of the motion (kinetic energy) of molecules. As the motion increases, energy is higher and thus temperature is higher. Water absorbs a great deal of energy before its temperature rises. Increased energy disrupts the hydrogen bonds between water molecules. Because these bonds can be created and disrupted rapidly, water absorbs an increase in energy and temperature changes only minimally. This means that water moderates temperature changes within organisms and in their environments. As energy input continues, the balance between hydrogen-bond formation and destruction swings toward the destruction side. More bonds are broken than are formed. This process results in the release of individual water molecules at the surface of the liquid (such as a body of water, the leaves of a plant, or the skin of an organism) in a process called evaporation. Evaporation of sweat, which is 90 percent water, allows for cooling of an organism, because breaking hydrogen bonds requires an input of energy and takes heat away from the body.

Conversely, as molecular motion decreases and temperatures drop, less energy is present to break the hydrogen bonds between water molecules. These bonds

remain intact and begin to form a rigid, lattice-like structure (e.g., ice) (Figure 4**a**). When frozen, ice is less dense than liquid water (the molecules are farther apart). This means that ice floats on the surface of a body of water (Figure4**b**). In lakes, ponds, and oceans, ice will form on the surface of the water, creating an insulating barrier to protect the animal and plant life beneath from freezing in the water. If this did not happen, plants and animals living in water would freeze in a block of ice and could not move freely, making life in cold temperatures difficult or impossible.

Part A shows the lattice-like molecular structure of ice. Part B is a photo of ice on water.

Figure 4 (a) The lattice structure of ice makes it less dense than the freely flowing molecules of liquid water. Ice's lower density enables it to (b) float on water. (credit a: modification of work by Jane Whitney; credit b: modification of work by Carlos Ponte)

WATER IS AN EXCELLENT SOLVENT

Because water is polar, with slight positive and negative charges, ionic compounds and polar molecules can readily dissolve in it. Water is, therefore, what is referred to as a solvent—a substance capable of dissolving another substance. The charged particles will form hydrogen bonds with a surrounding layer of water molecules. This is referred to as a sphere of hydration and serves to keep the particles separated or dispersed in the water. In the case of table salt (NaCl) mixed in water (Figure , the sodium and chloride ions separate, or dissociate, in the water, and spheres of hydration are formed around the ions. A positively charged sodium ion is surrounded by the partially negative charges of oxygen atoms in water molecules. A negatively charged chloride ion is surrounded by the partially positive charges

of hydrogen atoms in water molecules. These spheres of hydration are also referred to as hydration shells. The polarity of the water molecule makes it an effective solvent and is important in its many roles in living systems.

Illustration of spheres of hydration around sodium and chlorine ions.

Figure 5 When table salt (NaCl) is mixed in water, spheres of hydration form around the ions.

WATER IS COHESIVE

Have you ever filled up a glass of water to the very top and then slowly added a few more drops? Before it overflows, the water actually forms a dome-like shape above the rim of the glass. This water can stay above the glass because of the property of cohesion. In cohesion, water molecules are attracted to each other (because of hydrogen bonding), keeping the molecules together at the liquid-air (gas) interface, although there is no more room in the glass. Cohesion gives rise to surface tension, the capacity of a substance to withstand rupture when placed under tension or stress. When you drop a small scrap of paper onto a droplet of water, the paper floats on top of the water droplet, although the object is denser (heavier) than the water. This occurs because of the surface tension that is created by the water molecules. Cohesion and surface tension keep the water molecules intact and the item floating on the top. It is even possible to "float" a steel needle on top of a glass of water if you place it gently, without breaking the surface tension (Figure 6).

Picture of a needle floating on top of water because of cohesion and surface tension.

Figure 6 The weight of a needle on top of water pulls the surface tension downward; at the same time, the surface tension of the water is pulling it up, suspending the needle on the surface of the water and keeping it from sinking. Notice the indentation in the water around the needle. (credit: Cory Zanker)

These cohesive forces are also related to the water's property of adhesion, or the attraction between water molecules and other molecules. This is observed when water "climbs" up a straw placed in a glass of water. You will notice that the water appears to be higher on the sides of the straw than in the middle. This is because the water molecules are attracted to the straw and therefore adhere to it.

Cohesive and adhesive forces are important for sustaining life. For example, because of these forces, water can flow up from the roots to the tops of plants to feed the plant.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=67#h5p-73

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. March 22,

2017 https://cnx.org/contents/s8Hh0oOc@9.21:t90BfSb7@4/Water

Buffers, pH, Acids, and Bases

The **pH** of a solution is a measure of its **acidity** or **alkalinity.** You have probably used litmus paper, paper that has been treated with a natural water-soluble dye so it can be used as a pH indicator, to test how much acid or base (alkalinity) exists in a solution. You might have even used some to make sure the water in an outdoor swimming pool is properly treated. In both cases, this pH test measures the amount of hydrogen ions that exists in a given solution. High concentrations of hydrogen ions yield a low pH, whereas low levels of hydrogen ions result in a high pH. The overall concentration of hydrogen ions is inversely related to its pH and can be measured on the pH scale (Figure 1). Therefore, the more hydrogen ions present, the lower the pH; conversely, the fewer hydrogen ions, the higher the pH.
The pH scale ranges from 0 to 14. A change of one unit on the pH scale represents a change in the concentration of hydrogen ions by a factor of 10, a change in two units represents a change in the concentration of hydrogen ions by a factor of 100. Thus, small changes in pH represent large changes in the concentrations of hydrogen ions. Pure water is **neutral.** It is neither acidic nor basic, and has a pH of 7.0. Anything below 7.0 (ranging from 0.0 to 6.9) is **acidic,** and anything above 7.0 (from 7.1 to 14.0) is **alkaline (basic).** The blood in your veins is slightly alkaline (pH = 7.4). The environment in your stomach is highly acidic (pH = 1 to 2).

Orange juice is mildly acidic (pH = approximately 3.5), whereas baking soda is basic (pH = 9.0).

The pH scale with representative substances and their pHs.

Figure 1 The pH scale measures the amount of hydrogen ions (H+) in a substance. (credit: modification of work by Edward Stevens)

Acids are substances that provide hydrogen ions (H^+) and lower pH, whereas bases provide hydroxide ions (OH^-) and raise pH. The stronger the acid, the more readily it donates H^+. For example, hydrochloric acid and lemon juice are very acidic and readily give up H^+ when added to water. Conversely, bases are those substances that readily donate OH^-. The OH^- ions combine with H^+ to produce water, which raises a substance's pH. Sodium hydroxide and many household cleaners are very alkaline and give up OH^- rapidly when placed in water, thereby raising the pH.

Most cells in our bodies operate within a very narrow window of the pH scale, typically ranging only from 7.2 to 7.6. If the pH of the body is outside of this range, the respiratory system malfunctions, as do other organs in the body. Cells no longer function properly, and proteins will break down. Deviation outside of the pH range can induce coma or even cause death.

So how is it that we can ingest or inhale acidic or basic substances and not die? Buffers are the key. **Buffers** readily absorb excess H^+ or OH^-, keeping the pH of the body carefully maintained in the aforementioned narrow range. Carbon dioxide is part of a prominent buffer system in the human body; it keeps the pH within the proper range. This buffer system involves carbonic acid (H_2CO_3) and bicarbonate (HCO_3^-) anion. If too much H^+ enters the body, bicarbonate will combine with the H^+ to create carbonic acid and limit the decrease in pH. Likewise, if too much OH^- is

introduced into the system, carbonic acid will rapidly dissociate into bicarbonate and H^+ ions. The H^+ ions can combine with the OH^- ions, limiting the increase in pH. While carbonic acid is an important product in this reaction, its presence is fleeting because the carbonic acid is released from the body as carbon dioxide gas each time we breathe. Without this buffer system, the pH in our bodies would fluctuate too much and we would fail to survive.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=75#h5p-75

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=75#h5p-74

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. March 22,

2017 https://cnx.org/contents/s8Hh0oOc@9.21:t90BfSb7@4/Water

Absolutely Necessary Chemistry Summary

Matter

- Matter is anything that occupies space and has mass.
- Matter is made up of atoms of different elements.
- All of the 92 elements that occur naturally have unique qualities that allow them to combine in various ways to create compounds or molecules.
- Atoms consist of protons, neutrons, and electrons.
- Atoms are the smallest units of an element that retain all of the properties of that element.

Chemical Bonds

- Electrons can be donated or shared between atoms to create bonds.
    - Ionic bonds form between a positively and a negatively charged atom. They are fairly strong bonds.
    - Covalent bonds form when atoms share one or more electrons. They are very

strong bonds.

- Hydrogen bonds form between partially charged atoms. They are weak bonds.
- van der Waals interactions form between polar, covalently bound atoms. They are weak attractions that are often temporary.

Water

- is POLAR, allowing for the formation of hydrogen bonds,
- is an excellent SOLVENT: because water is polar, it allows ions and other polar molecules to dissolve.
- STABILIZES TEMPERATURE: the hydrogen bonds between water molecules give water the ability to hold heat better than many other substances. As the temperature rises, the hydrogen bonds between water continually break and reform, allowing for the overall temperature to remain stable, although increased energy is added to the system.
- is COHESIVE: hydrogen bonds allow for the property of surface tension.

pH, Acids, Bases, and Buffers

- The pH of a solution is a measure of the concentration of hydrogen ions in the solution. The pH scale ranges from 0 to 14.
    - A solution with an equal number of hydrogen ions and hydroxide ions is neutral and has a pH of 7.

    - A solution with a high number of hydrogen ions is acidic and has a low pH value (below 7).
    - A solution with a high number of hydroxide ions is basic and has a high pH value (above 7).
- Buffers are solutions that moderate pH changes when an acid or base is added to the buffer system. Buffers are important in biological systems because of their ability to maintain constant pH conditions.

BIOLOGICAL MOLECULES

Learning Outcomes

- Describe the structure of biologically-important molecules (carbohydrates, lipids, proteins, nucleic acids, water) and how their structure leads to their function.

Food provides an organism with nutrients—the matter it needs to survive. Many of these critical nutrients come in the form of **biological macromolecules**, or large molecules necessary for life. These macromolecules are built from different combinations of smaller organic molecules. What specific types of biological macromolecules do living things require? How are these molecules formed? What functions do they serve? In this chapter, we will explore these questions.

There are four major classes of biological macromolecules (carbohydrates, lipids, proteins, and nucleic acids), and each is an important component of the cell and performs a wide array of functions. Combined, these molecules make up the

majority of a cell's mass. Biological macromolecules are organic, meaning that they contain carbon atoms. In addition, they may contain atoms of hydrogen, oxygen, nitrogen, phosphorus, sulfur, and additional minor elements.

These molecules are made up of subunits called monomers. Each type of biological molecule is made up of different monomers. The monomers are connected together into a chain by strong covalent bonds. It is important that covalent bonds connect the monomers. If they were connected by hydrogen bonds the monomers would easily separate from each other and the biological molecule would come apart. If ionic bonds connected the monomers, the biological molecule would be likely to fall apart if it came into contact with water.

Figure 1 The structure of a macromolecule can be compared to a necklace: both are larger structures that are built out of small pieces connected together into a chain. The "string" in a macromolecule would be strong covalent bonds connecting the individual subunits together. ("Beads on a string" by Daniel is licensed under CC BY-NC-ND 2.0)

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=54#h5p-76

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:QhGQhr4x@6/Biological-Molecules

Carbohydrates

Carbohydrates are macromolecules with which most consumers are somewhat familiar. To lose weight, some individuals adhere to “low-carb” diets. Athletes, in contrast, often “carb-load” before important competitions to ensure that they have sufficient energy to compete at a high level. Carbohydrates are, in fact, an essential part of our diet; grains, fruits, and vegetables are all natural sources of carbohydrates. Carbohydrates provide energy to the body, particularly through glucose, a simple sugar. Carbohydrates also have other important functions in humans, animals, and plants.

Figure 1 Bread, pasta, and sugar all contain high levels of carbohydrates. ("Wheat products" by US Department of Agriculture is in the Public Domain)

Carbohydrates can be represented by the stoichiometric formula $(CH_2O)_n$, where n is the number of carbons in the molecule. In other words, the ratio of carbon to hydrogen to oxygen is 1:2:1 in carbohydrate molecules. This formula also explains the origin of the term "carbohydrate": the components are carbon ("carbo") and the components of water (hence, "hydrate"). Carbohydrates are classified into three subtypes: monosaccharides, disaccharides, and polysaccharides.

MONOSACCHARIDES

Monosaccharides (mono- = "one"; sacchar- = "sweet") are simple sugars, the most common of which is glucose. In monosaccharides, the number of carbons usually ranges

from three to seven. Most monosaccharide names end with the suffix -ose.

The chemical formula for glucose is $C_6H_{12}O_6$. In humans, glucose is an important source of energy. During cellular respiration, energy is released from glucose, and that energy is used to help make adenosine triphosphate (ATP). Plants synthesize glucose using carbon dioxide and water, and glucose in turn is used for energy requirements for the plant. Excess glucose is often stored as starch that is catabolized (the breakdown of larger molecules by cells) by humans and other animals that feed on plants.

Galactose (part of lactose, or milk sugar) and fructose (found in sucrose, in fruit) are other common monosaccharides. Although glucose, galactose, and fructose all have the same chemical formula ($C_6H_{12}O_6$), they differ structurally and chemically (and are known as isomers) because of the different arrangement of functional groups around the asymmetric carbon; all of these monosaccharides have more than one asymmetric carbon. Within one monosaccharide, all of the atoms are connected to each other with strong covalent bonds.

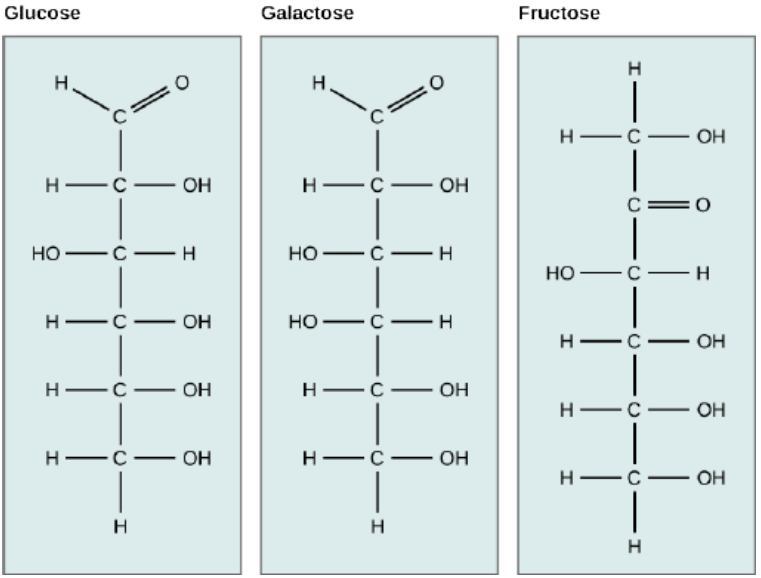


Figure 2 Glucose, galactose, and fructose are all hexoses. They are structural isomers, meaning they have the same chemical formula ($C_6H_{12}O_6$) but a different arrangement of atoms. The lines between atoms represent covalent bonds.

DISACCHARIDES

Disaccharides (di- = "two") form when two monosaccharides undergo a dehydration reaction (also known as a condensation reaction or dehydration synthesis). During this process, the hydroxyl (OH) group of one monosaccharide combines with the hydrogen of another monosaccharide, releasing a molecule of water and forming a covalent bond which joins the two monosaccharides together.

Common disaccharides include lactose, maltose, and sucrose (Figure 3). Lactose is a disaccharide consisting of the monomers glucose and galactose. It is formed by a dehydration reaction between the glucose and the galactose molecules, which removes a water molecule and forms a

covalent bond. connected by a covalent bond. It is found naturally in milk. Maltose, or malt sugar, is a disaccharide composed of two glucose molecules connected by a covalent bond. The most common disaccharide is sucrose, or table sugar, which is composed of the monomers glucose and fructose, also connected by a covalent bond.

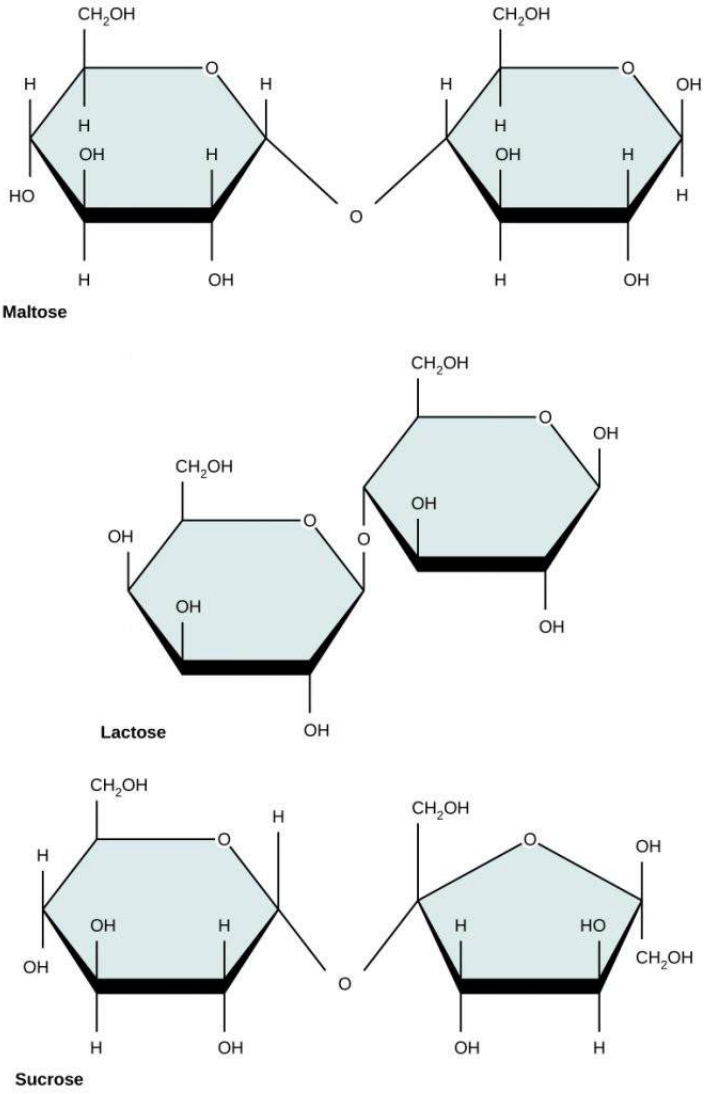


Figure 3 Common disaccharides include maltose (grain sugar), lactose (milk sugar), and sucrose (table sugar).

POLYSACCHARIDES

A long chain of monosaccharides linked by glycosidic bonds is known as a polysaccharide (poly- = "many"). The chain may be branched or unbranched, and it may contain different types of monosaccharides. All of the monosaccharides are connected together by covalent bonds. The molecular weight may be 100,000 daltons or more depending on the number of monomers joined. Starch, glycogen, cellulose, and chitin are primary examples of polysaccharides.

Starch is the stored form of sugars in plants and is made up of a mixture of amylose and amylopectin (both polymers of glucose). Basically, starch is a long chain of glucose monomers. Plants are able to synthesize glucose, and the excess glucose, beyond the plant's immediate energy needs, is stored as starch in different plant parts, including roots and seeds. The starch in the seeds provides food for the embryo as it germinates and can also act as a source of food for humans and animals. The starch that is consumed by humans is broken down by enzymes, such as salivary amylases, into smaller molecules, such as maltose and glucose. The cells can then absorb the glucose.

Glycogen is the storage form of glucose in humans and other vertebrates and is made up of monomers of glucose. Glycogen is the animal equivalent of starch and is a highly branched molecule usually stored in liver and muscle cells. Whenever blood glucose levels decrease, glycogen is broken down to release glucose in a process known as glycogenolysis.

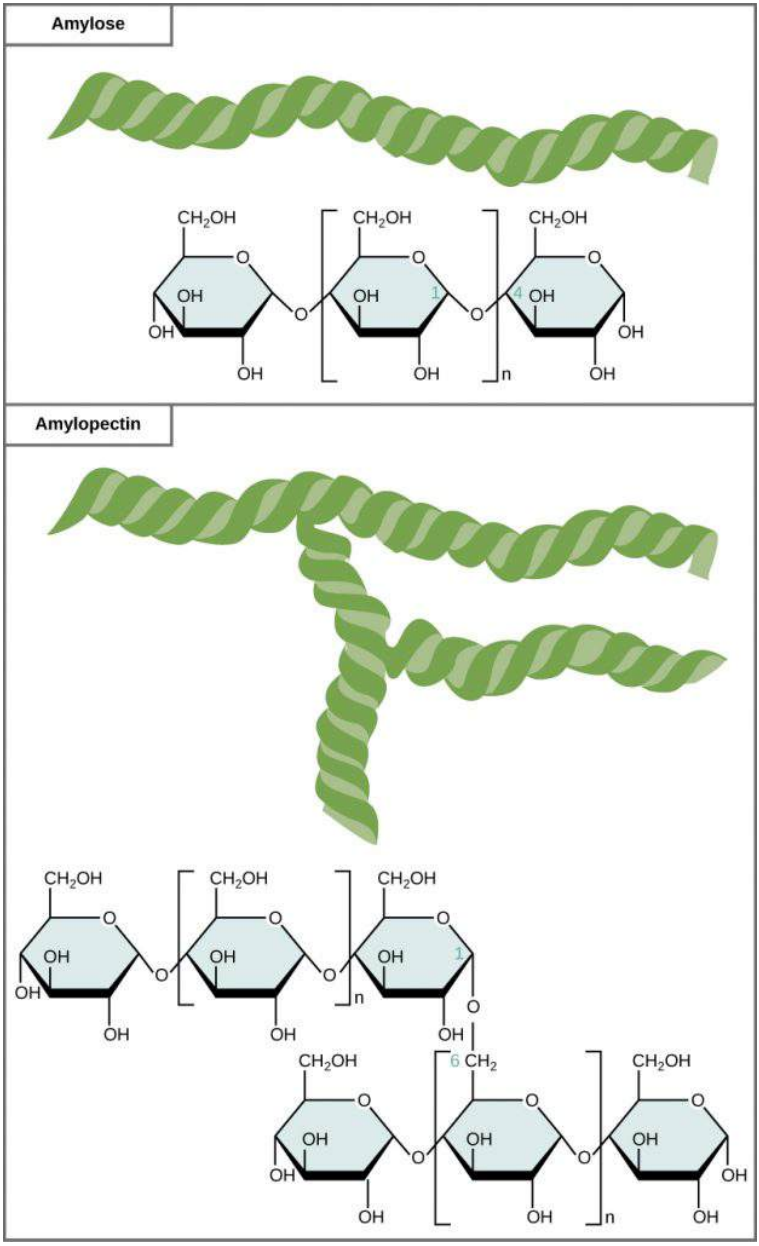


Figure 4 *Amylose and amylopectin are two different forms of starch.*

Amylose is composed of unbranched chains of glucose monomers. Amylopectin is composed of branched chains of glucose monomers. Because of the way the subunits are joined, the glucose chains have a helical structure. Glycogen (not shown) is similar in structure to amylopectin but more highly branched.

Cellulose is the most abundant natural biopolymer. The cell wall of plants is mostly made of cellulose; this provides structural support to the cell. Wood and paper are mostly cellulosic in nature. Cellulose is made up of glucose monomers (Figure 5).

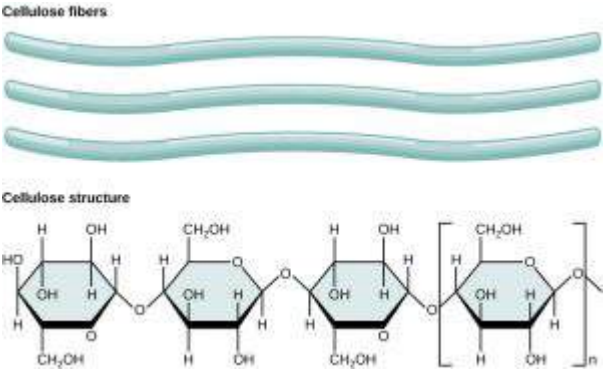


Figure 5 In cellulose, glucose monomers are linked in unbranched chains. Because of the way the glucose subunits are joined, every glucose monomer is flipped relative to the next one resulting in a linear, fibrous structure.

Carbohydrates serve various functions in different animals. Arthropods (insects, crustaceans, and others) have an outer skeleton, called the exoskeleton, which protects their internal body parts (as seen in the bee in Figure 6). This exoskeleton is made of the biological macromolecule chitin, which is a polysaccharide-containing nitrogen. It is made of repeating units of N-acetyl-β-d-glucosamine, a modified sugar. Chitin is also a major component of fungal cell walls; fungi are neither

animals nor plants and form a kingdom of their own in the domain Eukarya.

Figure 6 Insects have a hard outer exoskeleton made of chitin, a type of polysaccharide. (credit: Louise Docker)

HOW DOES CARBOHYDRATE STRUCTURE RELATE TO FUNCTION?

Energy can be stored within the bonds of a molecule. Bonds connecting two carbon atoms or connecting a carbon atom to a hydrogen atom are high energy bonds. Breaking these bonds releases energy. This is why our cells can get energy from a molecule of glucose ($C_6H_{12}O_6$).

Polysaccharides form long, fibrous chains which are able to build strong structures such as cell walls.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:QhGQhr4x@6/Biological-Molecules

Lipids

Lipids are a diverse group of compounds that are united by a common feature. **Lipids** are hydrophobic ("water-fearing"), or insoluble in water. Lipids perform many different functions in a cell. Cells store energy for long-term use in the form of lipids called fats. Lipids also provide insulation from the environment for plants and animals. For example, they help keep aquatic birds and mammals dry because of their water-repelling nature. Lipids are also the building blocks of many hormones and are an important constituent of the plasma membrane. Lipids include fats, oils, waxes, phospholipids, and steroids.

FATS AND OILS

A fat molecule consists of two main components—glycerol and fatty acids. Glycerol is an organic compound (an alcohol) that contains three carbons, five hydrogens, and three hydroxyl (OH) groups. Fatty acids have a long chain of hydrocarbons to which a carboxyl group is attached, hence the name "fatty acid." The number of carbons in the fatty acid may range from 4 to 36; most common are those containing 12–18 carbons. In a fat molecule, the fatty acids are attached to each of the three carbons of the glycerol molecule with a covalent bond. This molecule is called a triglyceride.

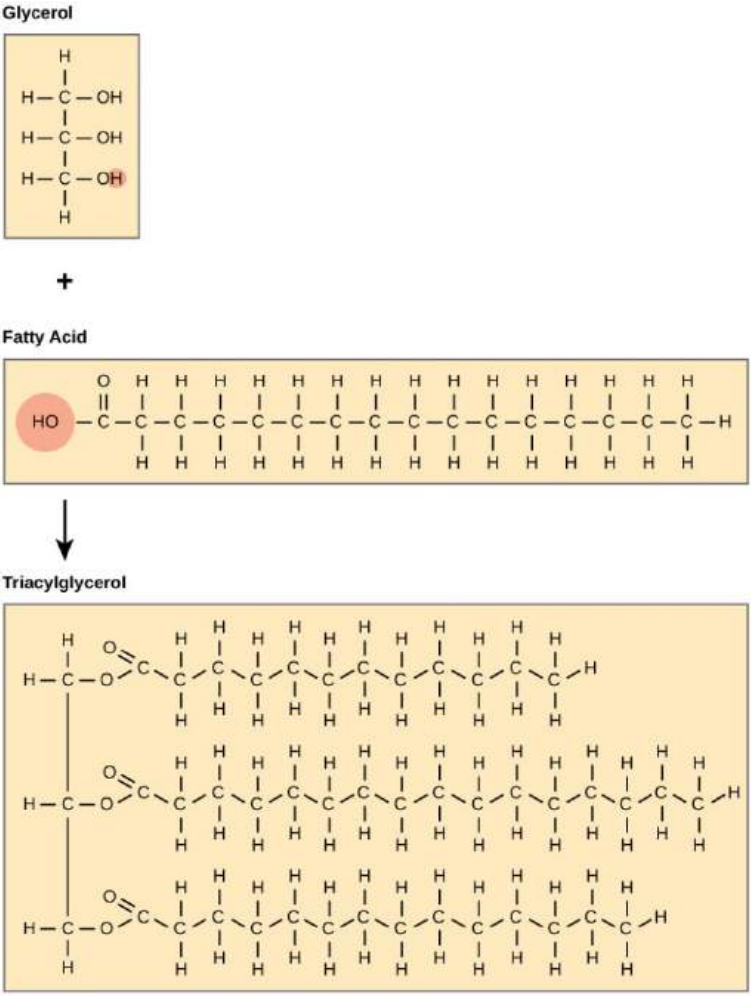


Figure 4 *Triacylglycerol is formed by the joining of three fatty acids to a glycerol backbone in a dehydration reaction (remember this removes a water molecule and forms a covalent bond). Three molecules of water are released in the process.*

WAXES

Wax covers the feathers of some aquatic birds and the leaf surfaces of some plants. Because of the hydrophobic nature of waxes, they prevent water from sticking on the surface (Figure 5). Waxes are made up of long fatty acid chains covalently bonded to long-chain alcohols.

Figure 5 Waxy coverings on some leaves are made of lipids. (credit: Roger Griffith)

PHOSPHOLIPIDS

Phospholipids are major constituents of the plasma membrane, the outermost layer of animal cells. Like fats, they are composed of fatty acid chains covalently bonded to a glycerol or sphingosine backbone. Instead of three fatty acids attached as in triglycerides, however, there are two fatty acids forming diacylglycerol, and the third carbon of the glycerol backbone is occupied by a modified phosphate group (Figure 6). Phosphatidylcholine and

phosphatidylserine are two important phospholipids that are found in plasma membranes.

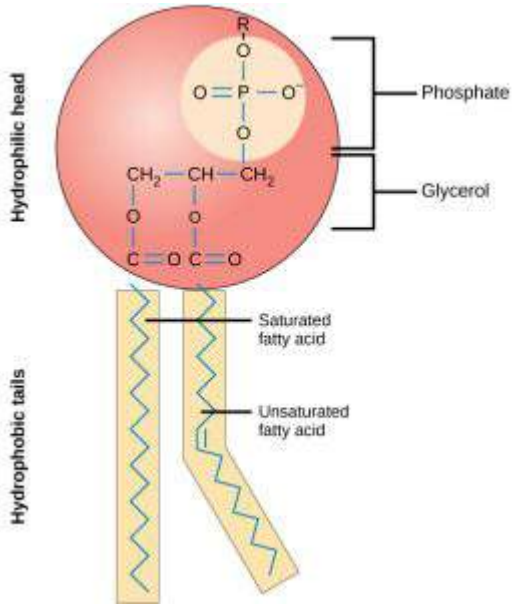


Figure 6 A phospholipid is a molecule with two fatty acids and a modified phosphate group attached to a glycerol backbone. The phosphate may be modified by the addition of charged or polar chemical groups. Two chemical groups that may modify the phosphate, choline and serine, are shown here. Both choline and serine attach to the phosphate group at the position labeled R via the hydroxyl group indicated in green.

A phospholipid is an **amphipathic** molecule, meaning it has a hydrophobic and a hydrophilic part. The fatty acid chains are hydrophobic and cannot interact with water, whereas the phosphate-containing group is hydrophilic and interacts with water (Figure 7). The head is the hydrophilic part, and the tail contains the hydrophobic fatty acids. In a membrane, a bilayer of phospholipids forms the matrix of the structure, the fatty acid tails of phospholipids face inside, away from water, whereas the phosphate group faces the outside,

aqueous side. This forms a hydrophobic layer on the inside of the bilayer, where the tails are located.

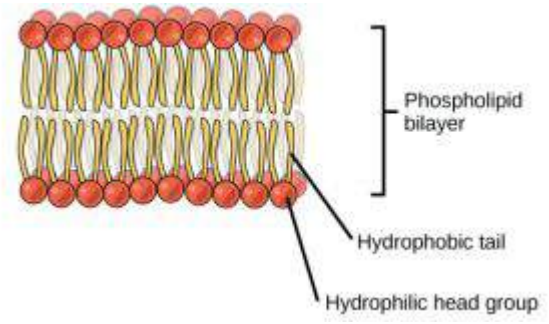


Figure 7 The phospholipid bilayer is the major component of all cellular membranes. The hydrophilic head groups of the phospholipids face the aqueous solution. The hydrophobic tails are sequestered in the middle of the bilayer.

Phospholipids are responsible for the dynamic nature of the plasma membrane. If a drop of phospholipids is placed in water, it spontaneously forms a structure known as a micelle, where the hydrophilic phosphate heads face the outside and the fatty acids face the interior of this structure (Figure 8).

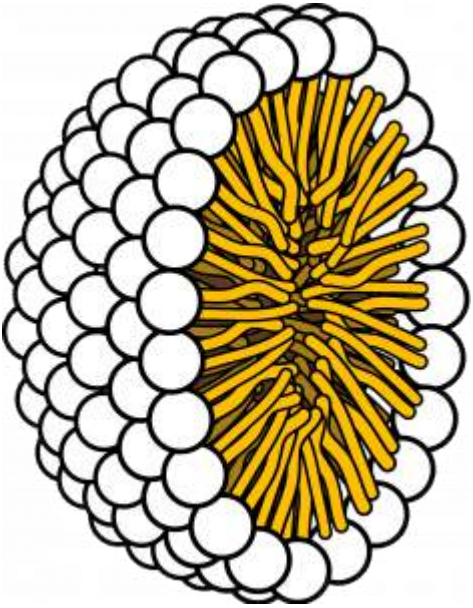
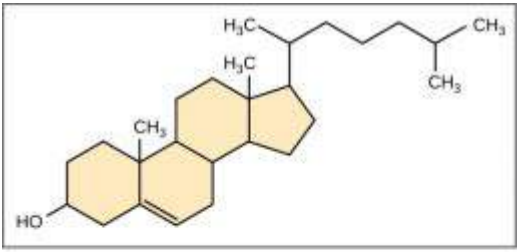


Figure 8 A micelle may be the very early precursor of a cell. It is a single layer of phospholipids that form spontaneously. Credit AmitWo, Wikimedia; https://commons.wikimedia.org/wiki/File:Micelle.svg

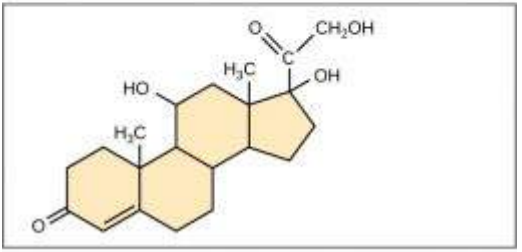
STEROIDS

Unlike the phospholipids and fats discussed earlier, steroids have a fused ring structure. Although they do not resemble the other lipids, they are grouped with them because they are also hydrophobic and insoluble in water. All steroids have four linked carbon rings and several of them, like cholesterol, have a short tail (Figure 9). Many steroids also have the –OH functional group, which puts them in the alcohol classification (sterols). Remember that each line in these diagrams of chemical structures represents a covalent bond. The points where the lines connect to each other show the location of carbon atoms – these carbon atoms are not

labeled, but their existence is implied in the chemical structure.



Cholesterol



Cortisol

Figure 9 Steroids such as cholesterol and cortisol are composed of four fused hydrocarbon rings.

Cholesterol is the most common steroid. Cholesterol is mainly synthesized in the liver and is the precursor to many steroid hormones such as testosterone and estradiol, which are secreted by the gonads and endocrine glands. It is also the precursor to Vitamin D. Cholesterol is also the precursor of bile salts, which help in the emulsification of fats and their subsequent absorption by cells. Although cholesterol is often spoken of in negative terms by lay people, it is necessary for proper functioning of the body. It is a component of the plasma membrane of animal cells and is found within the phospholipid bilayer. Being the outermost structure in animal cells, the plasma membrane is responsible for the transport of materials and cellular recognition and it is involved in cell-to-cell communication.

HOW DOES LIPID STRUCTURE RELATE TO FUNCTION?

Fats (triglycerides) are made up of three fatty acid hydrocarbon chains connected to a glycerol. Fatty acid chains contain large numbers of carbon-carbon and carbon-hydrogen bonds – they are typically made up of between 4 and 28 carbons connected together in a chain. Just like the carbon-carbon and carbon-hydrogen bonds in glucose allow that molecule to store energy, the bonds in fatty acids allow triglycerides to store energy. In fact, triglycerides can store much more energy than carbohydrates because they contain so many more bonds! This is why fats contain more calories (a measure of energy) than sugars do.

Waxes function to provide a waterproof coating on a surface. Because they are hydrophobic, they can form a coating that repels water.

The structure of phospholipids is very important to their function. Because they are amphipathic (have a hydrophobic and a hydrophilic portion), they self-assemble into structures where the hydrophobic tails are hidden away from the watery environment. This gives the cell membrane a structure that prevents many molecules from moving through it.

Cholesterol is also amphipathic. It can insert into cell membranes in a manner similar to phospholipids. The presence of cholesterol within a membrane prevents the phospholipid tails from packing together tightly. This allows the membrane to remain fluid at lower temperatures.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:QhGQhr4x@6/Biological-Molecules

Proteins

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective; they may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each cell in a living system may contain thousands of different proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, polymers of amino acids, arranged in a linear sequence and connected together by covalent bonds.

Amino acids are the monomers that makeup proteins (Figure 1). Each amino acid has the same fundamental structure, which consists of a central carbon atom, also known as the alpha (α) carbon, bonded to an amino group (NH_2), a carboxyl group (COOH), and to a hydrogen atom. Every amino acid also has another atom or group of atoms bonded to the central atom known as the R group.

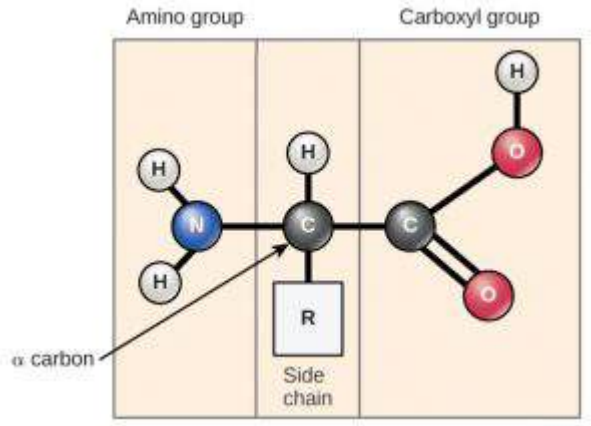


Figure 1 Amino acids have a central asymmetric carbon to which an amino group, a carboxyl group, a hydrogen atom, and a side chain (R group) are attached.

Function	Examples	Description
Defense	Immunoglobulins	Antibodies bind to specific foreign particles, such as viruses and bacteria, to help protect the body.
Enzyme	Digestive enzymes such as amylase, lipase, pepsin, trypsin	Enzymes carry out almost all of the thousands of chemical reactions that take place in cells. They also assist with the formation of new molecules by reading the genetic information stored in DNA.
Messenger	Insulin, thyroxine	Messenger proteins, such as some types of hormones, transmit signals to coordinate biological processes between different cells, tissues, and organs.
Structural component	Actin, tubulin, keratin	These proteins provide structure and support for cells. On a larger scale, they also allow the body to move.
Transport/ storage	Hemoglobin, albumin, Legume storage proteins, egg white (albumin)	These proteins bind and carry atoms and small molecules within cells and throughout the body. Some provide nourishment in early development of the embryo and the seedling
Contractile	Actin, myosin	Affect muscle contraction.

You may have noticed that “source of energy” was not listed among the function of proteins. This is because proteins in our diet are typically broken back down into individual amino

acids that our cells then assemble into our own proteins. Humans are actually unable to build some amino acids inside our own cells – we require them in our diet (these are the so-called "essential" amino acids). Our cells can digest proteins to release energy, but will usually only do so when carbohydrates or lipids are not available.

Figure 2 Examples of foods that contain high levels of protein. ("Protein" by National Cancer Institute is in the Public Domain)

The functions of proteins can be very diverse because they are made up of are 20 different chemically distinct amino acids that form long chains, and the amino acids can be in any order. The function of the protein is dependent on the protein's shape. The shape of a protein is determined by the order of the amino acids. Proteins are often hundreds of amino acids long and they can have very complex shapes because there are so many different possible orders for the 20 amino acids (Figure 3)!

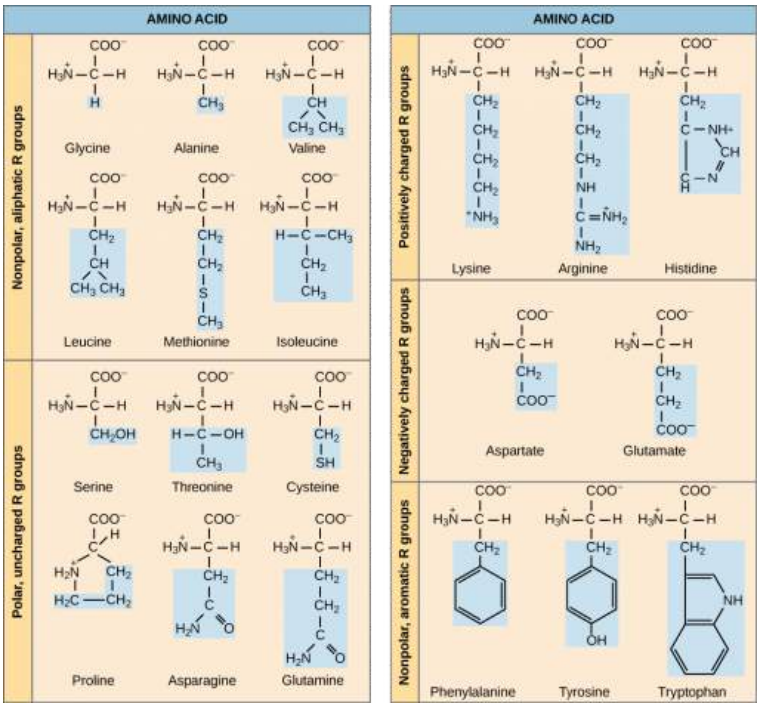


Figure 3 There are 20 common amino acids commonly found in proteins, each with a different R group (variant group) that determines its chemical nature.

The chemical nature of the side chain determines the nature of the amino acid (that is, whether it is acidic, basic, polar, or nonpolar). For example, the amino acid glycine has a hydrogen atom as the R group. Amino acids such as valine, methionine, and alanine are nonpolar or hydrophobic in nature, while amino acids such as serine, threonine, and cysteine are polar and have hydrophilic side chains. The side chains of lysine and arginine are positively charged, and therefore these amino acids are also known as basic amino acids. Proline has an R group that is linked to the amino group, forming a ring-like structure. Proline is an exception to the standard structure of an amino acid since its amino

group is not separate from the side chain (Figure 3). Amino acids are represented by a single upper case letter as well as a three-letter abbreviation. For example, valine is known by the letter V or the three-letter symbol val.

Just as some fatty acids are essential to a diet, some amino acids are necessary as well. They are known as essential amino acids, and in humans they include isoleucine, leucine, and cysteine. Essential amino acids refer to those necessary for construction of proteins in the body, although not produced by the body; which amino acids are essential varies from organism to organism.

The sequence and the number of amino acids ultimately determine the protein's shape, size, and function. Each amino acid is attached to another amino acid by a covalent bond, known as a peptide bond, which is formed by a dehydration reaction. The carboxyl group of one amino acid and the amino group of the incoming amino acid combine, releasing a molecule of water. The resulting bond is the peptide bond (Figure 4).

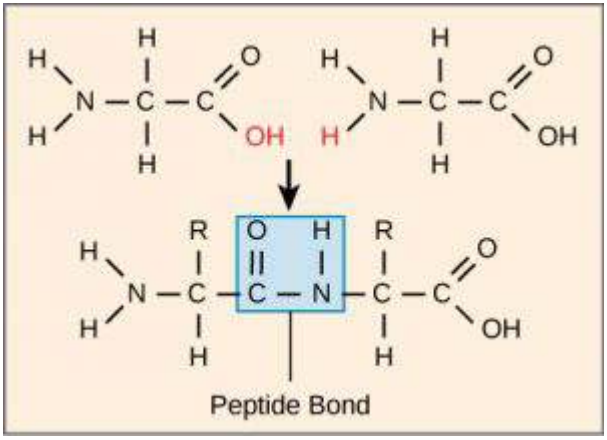


Figure 4 Peptide bond formation is a dehydration synthesis reaction. The carboxyl group of one amino acid is linked to the amino group of the incoming amino acid. In the process, a molecule of water is released.

PROTEIN STRUCTURE

As discussed earlier, the shape of a protein is critical to its function. For example, an enzyme can bind to a specific substrate at a site known as the active site. If this active site is altered because of local changes or changes in overall protein structure, the enzyme may be unable to bind to the substrate. To understand how the protein gets its final shape or conformation, we need to understand the four levels of protein structure: primary, secondary, tertiary, and quaternary (Figure 5).

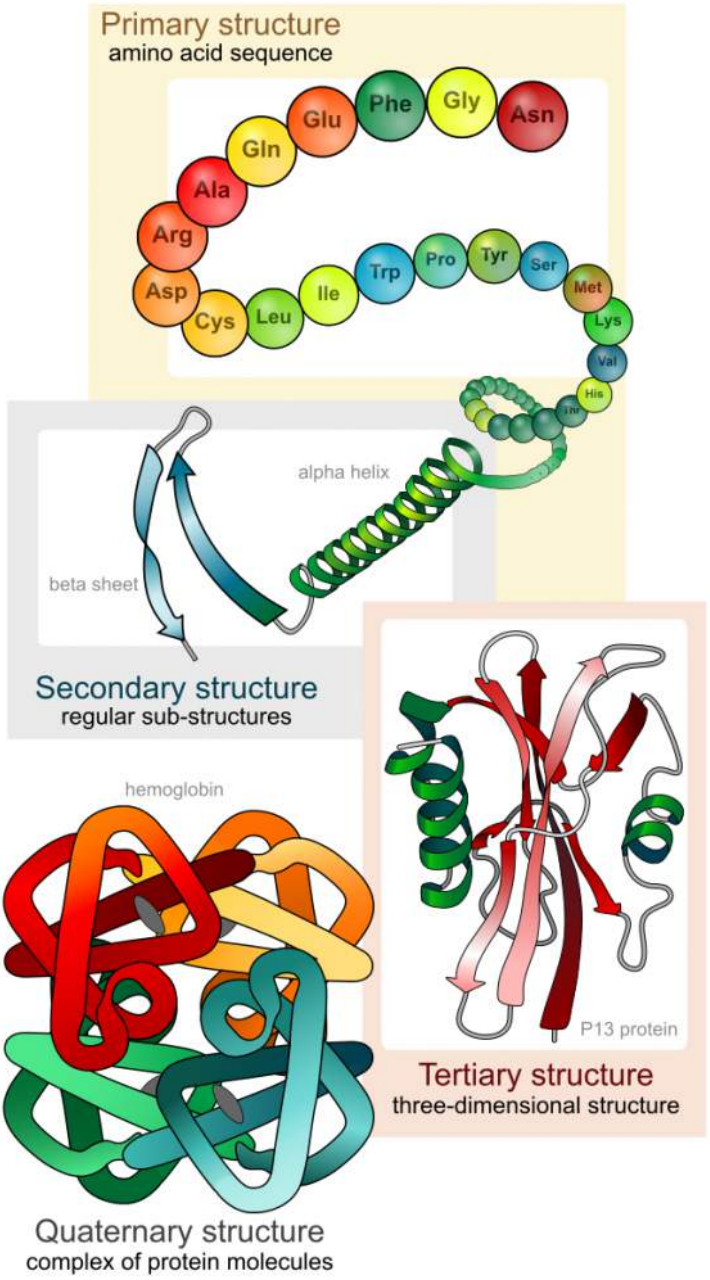

Primary structure
amino acid sequence
Phe
Gly
Asn
Glu
Gln
Ala
Arg
Asp
Cys
Leu
Ile
Trp
Pro
Tyr
Ser
Met
Lys
Val
His
alpha helix
beta sheet
Secondary structure
regular sub-structures
P13 protein
Tertiary structure
three-dimensional structure
hemoglobin
Quaternary structure
complex of protein molecules

Figure 5 Main levels of protein structure. ("Main protein structure levels en" by LadyofHats is in the Public Domain)

PRIMARY STRUCTURE

The unique sequence of amino acids in a polypeptide chain is its primary structure. For example, the pancreatic hormone insulin has two polypeptide chains, A and B, and they are linked together by disulfide bonds. The N terminal amino acid of the A chain is glycine, whereas the C terminal amino acid is asparagine (). The sequences of amino acids in the A and B chains are unique to insulin.

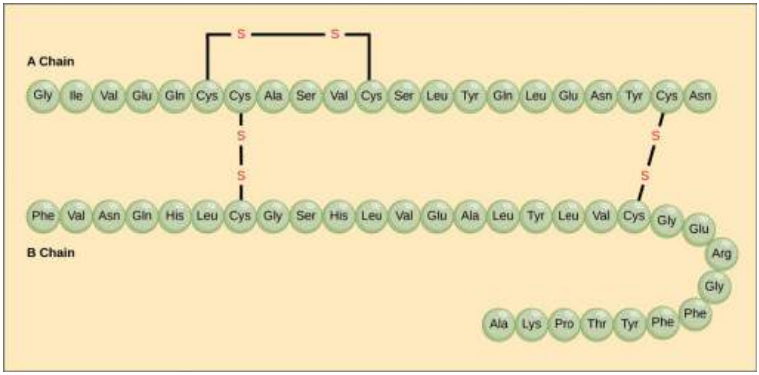


Figure 6 Bovine serum insulin is a protein hormone made of two peptide chains, A (21 amino acids long) and B (30 amino acids long). In each chain, primary structure is indicated by three-letter abbreviations that represent the names of the amino acids in the order they are present. The amino acid cysteine (cys) has a sulfhydryl (SH) group as a side chain. Two sulfhydryl groups can react in the presence of oxygen to form a disulfide (S-S) bond. Two disulfide bonds connect the A and B chains together, and a third helps the A chain fold into the correct shape. Note that all disulfide bonds are the same length, but are drawn different sizes for clarity.

SECONDARY STRUCTURE

The local folding of the polypeptide in some regions gives rise to the secondary structure of the protein. The most common are the *α*-helix and *β*-pleated sheet structures (Figure 7). Both structures are the *α*-helix structure—the helix held in shape by hydrogen bonds. The hydrogen bonds form between the oxygen atom in the carbonyl group in one amino acid and another amino acid that is four amino acids farther along the chain.

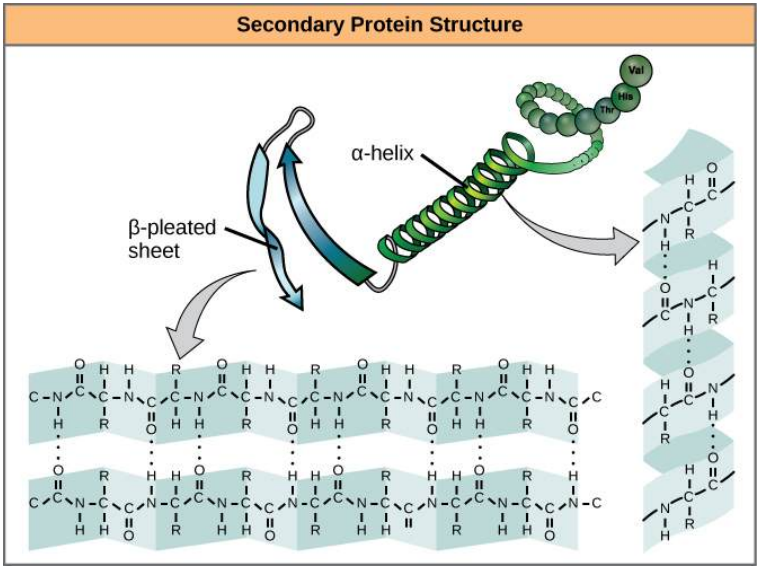


Figure 7 *The α-helix and β-pleated sheet are secondary structures of proteins that form because of hydrogen bonding between carbonyl and amino groups in the peptide backbone. Certain amino acids have a propensity to form an α-helix, while others have a propensity to form a β-pleated sheet.*

TERTIARY STRUCTURE

The unique three-dimensional structure of a polypeptide is

its tertiary structure (Figure 8). This structure is in part due to chemical interactions at work on the polypeptide chain. Primarily, the interactions among R groups (the variable part of the amino acid) creates the complex three-dimensional tertiary structure of a protein. The nature of the R groups found in the amino acids involved can counteract the formation of the hydrogen bonds described for standard secondary structures. For example, R groups with like charges are repelled by each other and those with unlike charges are attracted to each other (ionic bonds). When protein folding takes place, the hydrophobic R groups of nonpolar amino acids lay in the interior of the protein, whereas the hydrophilic R groups lay on the outside. The former types of interactions are also known as hydrophobic interactions. Interaction between cysteine side chains forms disulfide linkages in the presence of oxygen, the only covalent bond forming during protein folding.

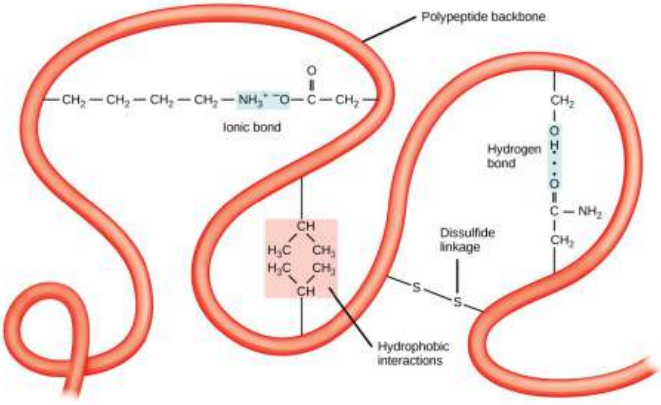


Figure 8 The tertiary structure of proteins is determined by a variety of chemical interactions. These include hydrophobic interactions, ionic bonding, hydrogen bonding and disulfide linkages.

QUATERNARY STRUCTURE

In nature, some proteins are formed from several polypeptides, also known as subunits, and the interaction of these subunits forms the quaternary structure. Weak interactions between the subunits help to stabilize the overall structure. For example, insulin (a globular protein) has a combination of hydrogen bonds and disulfide bonds that cause it to be mostly clumped into a ball shape. Insulin starts out as a single polypeptide and loses some internal sequences in the presence of post-translational modification after the formation of the disulfide linkages that hold the remaining chains together. Silk (a fibrous protein), however, has a β-pleated sheet structure that is the result of hydrogen bonding between different chains.

The four levels of protein structure (primary, secondary, tertiary, and quaternary) are illustrated in Figure 9.

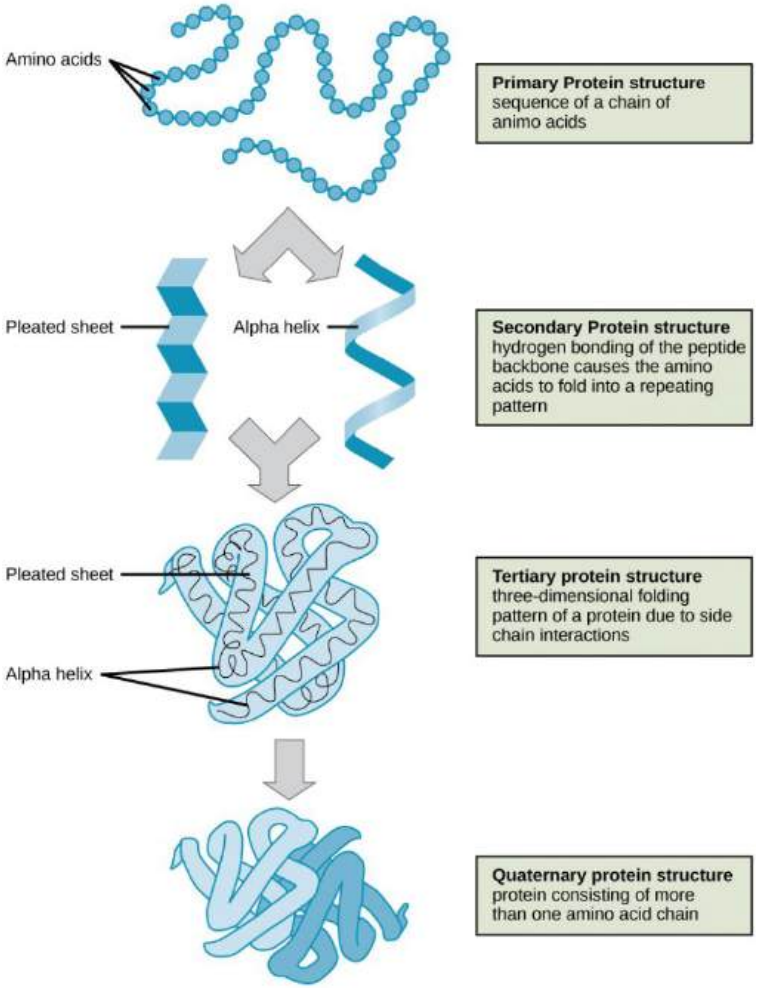


Figure 9 The four levels of protein structure can be observed in these illustrations. (credit: modification of work by National Human Genome Research Institute)

The unique shape for every protein is ultimately determined by the gene that encodes the protein. Any change in the gene sequence may lead to a different amino acid being added to the polypeptide chain, causing a change in protein

structure and function. Individuals who are affected by sickle cell anemia can have a variety of serious health problems, such as breathlessness, dizziness, headaches, and abdominal pain. In this disease, the hemoglobin β chain has a single amino acid substitution, causing a change in both the structure (shape) and function (job) of the protein. What is most remarkable to consider is that a hemoglobin molecule is made up of about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule is a single amino acid of the 600 (Figure 10).

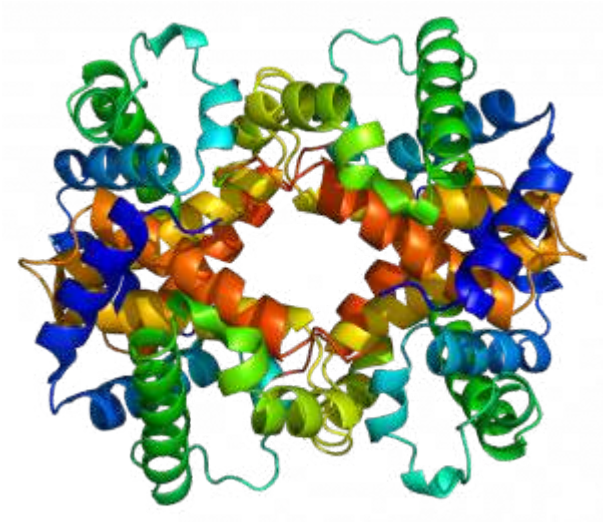


Figure 10 The unique shape of the normal hemoglobin protein. ("Structure of hemoglobin Gower 2" by Emw is licensed under CC BY-SA 3.0)

DENATURATION AND PROTEIN FOLDING

Each protein has its own unique sequence and shape that are held together by chemical interactions. If the protein is subject to changes in temperature, pH, or exposure to

chemicals, the protein structure may change, losing its shape without losing its primary sequence in what is known as denaturation. Denaturation is often reversible because the primary structure of the polypeptide is conserved in the process if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to loss of function. One example of irreversible protein denaturation is when an egg is fried. The albumin protein in the liquid egg white is denatured when placed in a hot pan. Not all proteins are denatured at high temperatures; for instance, bacteria that survive in hot springs have proteins that function at temperatures close to boiling. The stomach is also very acidic, has a low pH, and denatures proteins as part of the digestion process; however, the digestive enzymes of the stomach retain their activity under these conditions.

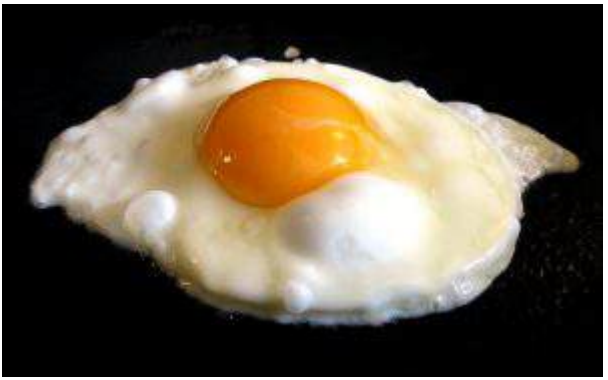


Figure 11 The reason an egg white turns white as you cook it is because the albumin in the white denatures and then reconnects in an abnormal fashion. Credit Matthew Murdock; https://www.flickr.com/photos/54423233@N05/13916201522

Protein folding is critical to its function. It was originally thought that the proteins themselves were responsible for the folding process. Only recently was it found that often they receive assistance in the folding process from protein

helpers known as chaperones (or chaperonins) that associate with the target protein during the folding process. They act by preventing aggregation of polypeptides that make up the complete protein structure, and they disassociate from the protein once the target protein is folded.

HOW DOES PROTEIN STRUCTURE RELATE TO FUNCTION?

Recall that a protein is built from a long chain of amino acids connected together in a specific order. The specific order of amino acids determines how they will interact together to form the 3-D shape of the protein. The shape of a protein determines its function. Therefore, the order of the amino acids determines the protein's shape, which determines its function.

Because there are 20 different amino acids, they can be combined together in a practically infinite number of ways. This means that there is a huge number of different protein shapes that can be assumed based on the amino acid order. This is very important since proteins fulfill so many different functions within cells.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:QhGQhr4x@6/Biological-Molecules

Nucleic Acids

Nucleic acids are key macromolecules in the continuity of life. They carry the genetic blueprint of a cell and carry instructions for the functioning of the cell. The two main types of **nucleic acids** are **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. DNA is the genetic material found in all living organisms, ranging from single-celled bacteria to multicellular mammals. The other type of nucleic acid, RNA, is mostly involved in protein synthesis. The DNA molecules never leave the nucleus, but instead use an RNA intermediary to communicate with the rest of the cell. Other types of RNA are also involved in protein synthesis and its regulation. We will be going into more detail about nucleic acids in a later section.

DNA and RNA are made up of monomers known as **nucleotides** connected together in a chain with covalent bonds. Each nucleotide is made up of three components: a nitrogenous base, five-carbon sugar, and a phosphate group (**Figure 1**). The nitrogenous base in one nucleotide is attached to the sugar molecule, which is attached to the phosphate group.

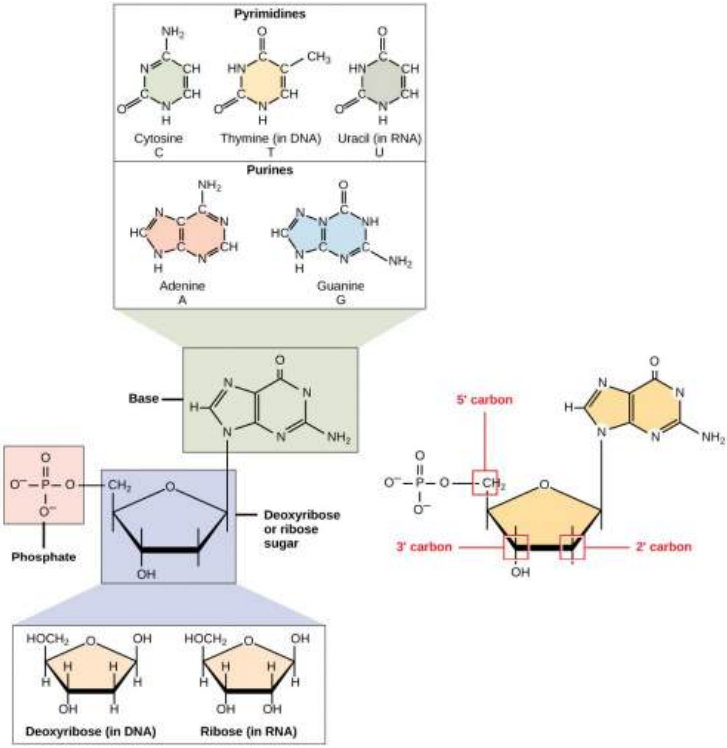


Figure 1 A nucleotide is made up of three components: a nitrogenous base, a pentose sugar, and one or more phosphate groups.

The nitrogenous bases, important components of nucleotides, are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus, decreases the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T). RNA contains the base uracil (U) instead of thymine. The order of the bases in a nucleic acid determines the information that the molecule of DNA or RNA carries. This is because the order of the bases in a DNA gene

determines the order that amino acids will be assembled together to form a protein.

The pentose sugar in DNA is deoxyribose, and in RNA, the sugar is ribose (Figure 1). The difference between the sugars is the presence of the hydroxyl group on the second carbon of the ribose and hydrogen on the second carbon of the deoxyribose. The carbon atoms of the sugar molecule are numbered as 1', 2', 3', 4', and 5' (1' is read as “one prime”). The phosphate residue is attached to the hydroxyl group of the 5' carbon of one sugar and the hydroxyl group of the 3' carbon of the sugar of the next nucleotide, which forms a 5'–3' phosphodiester linkage (a specific type of covalent bond). A polynucleotide may have thousands of such phosphodiester linkages.

DNA DOUBLE-HELICAL STRUCTURE

DNA has a double-helical structure (**Figure 2**). It is composed of two strands, or chains, of nucleotides. The double helix of DNA is often compared to a twisted ladder. The strands (the outside parts of the ladder) are formed by linking the phosphates and sugars of adjacent nucleotides with strong chemical bonds, called **covalent bonds**. The rungs of the twisted ladder are made up of the two bases attached together with a weak chemical bond, called **a hydrogen bonds**. Two bases hydrogen bonded together is called a **base pair**. The ladder twists along its length, hence the “double helix” description, which means a double spiral.

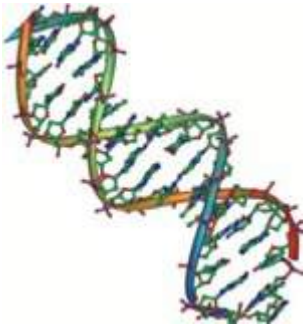


Figure 2 The double-helix model shows DNA as two parallel strands of intertwining molecules. (credit: Jerome Walker, Dennis Myts).

The alternating sugar and phosphate groups lie on the outside of each strand, forming the backbone of the DNA. The nitrogenous bases are stacked in the interior, like the steps of a staircase, and these basespair; the pairs are bound to each other by hydrogen bonds. The bases pair in such a way that the distance between the backbones of the two strands is the same all along the molecule.

HOW DOES NUCLEIC ACID STRUCTURE DETERMINE FUNCTION?

The major function of both DNA and RNA is to store and carry genetic information. The specific order of nucleotides in the molecule of DNA or RNA is what determines the genetic information it carries. You can think of it like letters in a book – if the order of the letters were changed, the book would no longer contain the same (or correct) information.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016. https://cnx.org/contents/GFy_h8cu@10.120:U7tPDRxK@9/DNA-Structure-and-Sequencing

Biological Macromolecule Practice Questions

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-87

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-84

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-83

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-79

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-78

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-82

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-80

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-77

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-85

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1267#h5p-86

WHAT IS LIFE?

What is biology? In simple terms, biology is the study of living organisms and their interactions with one another and their environments. This is a very broad definition because the scope of biology is vast. Biologists may study anything from the microscopic or submicroscopic view of a cell to ecosystems and the whole living planet. Listening to the daily news, you will quickly realize how many aspects of biology are discussed every day. For example, recent news topics include Zika virus and using a new technology called CRISPR to specifically target and edit human genes. Other subjects include efforts toward finding a cure for AIDS, Alzheimer's disease, and cancer. On a global scale, many researchers are committed to finding ways to protect the planet, solve environmental issues, and reduce the effects of climate change. All of these diverse endeavors are related to different facets of the discipline of biology.

What makes something living?

All living organisms share several key characteristics or functions: order, sensitivity or response to the environment, reproduction, adaptation, growth and development, homeostasis, energy processing, and evolution. When viewed together, these characteristics serve to define life. Different sources may use slightly different terms to describe these characteristics, but the basic ideas are always present.

ORDER

A photo shows a light-colored toad covered in bright green spots.

Figure 1 *A toad represents a highly organized structure consisting of cells, tissues, organs, and organ systems. (credit: "Ivengo"/Wikimedia Commons)*

Organisms are highly organized, coordinated structures that consist of one or more cells. Even very simple, single-celled organisms are remarkably complex: inside each cell, atoms make up molecules; these in turn make up cell organelles and other cellular inclusions. In multicellular organisms, such as the toad seen in Figure 1, similar cells form tissues. Tissues, in turn, collaborate to create organs (body structures

with a distinct function). Organs work together to form organ systems.

In this class, we will be focusing on how cells function, so we will be concentrating on biological molecules, how they make up cells, and how those cells function.

SENSITIVITY OR RESPONSE TO STIMULI

A photograph of the Mimosa pudica shows a plant with many tiny leaves connected to a central stem. Four of these stems connect together.

Figure 2 The leaves of this sensitive plant (Mimosa pudica) will instantly droop and fold when touched. After a few minutes, the plant returns to normal. (credit: Alex Lomas)

Organisms respond to diverse stimuli. For example, plants can bend toward a source of light, climb on fences and walls, or respond to touch (Figure 2). Even tiny bacteria can move toward or away from chemicals (a process called *chemotaxis*) or light (*phototaxis*).

REPRODUCTION

Single-celled organisms reproduce by first duplicating their DNA, and then dividing it equally as the cell prepares to divide to form two new cells. Multicellular organisms often produce specialized reproductive germline (reproductive) cells that will form new individuals. When reproduction occurs, DNA is passed from the organism to that organism's offspring. DNA contains the instructions to produce all the physical traits for the organism. This means that because parents and offspring share DNA ensures that the offspring will belong to the same species and will have similar characteristics, such as size and shape.

GROWTH AND DEVELOPMENT

All living things increase in size and/or change over their lifespan. For example, a human grows from a baby into an adult and goes through developmental processes such as puberty. Organisms grow and develop following specific instructions coded for by their genes (DNA). These genes provide instructions that will direct cellular growth and development, ensuring that a species' young will grow up to exhibit many of the same characteristics as its parents, like the kittens seen in Figure 3.

A photograph depicts a mother cat nursing three kittens: one has an orange and white tabby coat, another is black with a white foot, while the third has a black and white tabby coat.

Figure 3 *Although no two look alike, these kittens have inherited genes from both parents and share many of the same characteristics. (credit: Rocky Mountain Feline Rescue)*

HOMEOSTASIS AND REGULATION

A photos shows a white, furry polar bear.

Figure 4 *Polar bears (Ursus maritimus) and other mammals living in ice-covered regions maintain their body temperature by generating heat and reducing heat loss through thick fur and a dense layer of fat under their skin. (credit: "longhorndave"/Flickr)*

In order to function properly, cells need to have appropriate conditions such as proper temperature, pH, and appropriate concentration of diverse chemicals. These conditions may, however, change from one moment to the next. Organisms are able to maintain internal conditions within a narrow range almost constantly, despite environmental changes, through homeostasis (literally, "steady state")—the ability of

an organism to maintain constant internal conditions. For example, an organism needs to regulate body temperature through a process known as thermoregulation. Organisms that live in cold climates, such as the polar bear (Figure 4), have body structures that help them withstand low temperatures and conserve body heat. Structures that aid in this type of insulation include fur, feathers, blubber, and fat. In hot climates, organisms have methods (such as perspiration in humans or panting in dogs) that help them to shed excess body heat.

Even the smallest organisms are complex and require multiple regulatory mechanisms to coordinate internal functions, respond to stimuli, and cope with environmental stresses. Two examples of internal functions regulated in an organism are nutrient transport and blood flow. Organs (groups of tissues working together) perform specific functions, such as carrying oxygen throughout the body, removing wastes, delivering nutrients to every cell, and cooling the body.

ENERGY PROCESSING

Figure 5 *The California condor (Gymnogyps californianus) uses chemical energy derived from food to power flight. California condors are an endangered species; this bird has a wing tag that helps biologists identify the individual. (credit: Pacific Southwest Region U.S. Fish and Wildlife Service)*

All organisms use a source of energy for their metabolic activities. Some organisms capture energy from the sun and convert it into chemical energy in food (such as grass and bacteria that can perform photosynthesis); others use

chemical energy in molecules they take in as food (such as the condor seen in Figure 5).

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=241#h5p-21

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=241#h5p-22

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=241#h5p-23

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 25, 2017 https://cnx.org/contents/GFy_h8cu@10.99:gNLp76vu@13/Themes-and-Concepts-of-Biology

Levels of Organization

Living things are highly organized and structured, following a hierarchy that can be examined on a scale from small to large. The **atom** is the smallest and most fundamental unit of matter. It consists of a nucleus surrounded by electrons. Atoms form **molecules.** A molecule is a chemical structure consisting of at least two atoms held together by one or more chemical bonds. Many molecules that are biologically important are **macromolecules,** large molecules that are typically formed by polymerization (a polymer is a large molecule that is made by combining smaller units called monomers, which are simpler than macromolecules). An example of a macromolecule is deoxyribonucleic acid (DNA) (Figure 1), which contains the instructions for the structure and functioning of all living organisms. See the section of your textbook about the chemistry of biological molecules for more information.

Molecular model depicts a DNA molecule, showing its double helix structure.

Figure 1 All molecules, including this DNA molecule, are composed of atoms. (credit: "brian0918"/Wikimedia Commons)

Some cells contain aggregates of macromolecules surrounded by membranes; these are called **organelles.** Organelles are small structures that exist within cells. Examples of organelles include mitochondria and

chloroplasts, which carry out indispensable functions: mitochondria produce energy to power the **cell,** while chloroplasts enable green plants to utilize the energy in sunlight to make sugars. All living things are made of cells; the cell itself is the smallest fundamental unit of structure and function in living organisms. This requirement is one of the reasons why viruses are not considered living: they are not made of cells. To make new viruses, they have to invade and hijack the reproductive mechanism of a living cell; only then can they obtain the materials they need to reproduce. Some organisms consist of a single cell and others are multicellular. Cells are classified as prokaryotic or eukaryotic. Prokaryotes are single-celled or colonial organisms that do not have membrane-bound nuclei; in contrast, the cells of eukaryotes do have membrane-bound organelles and a membrane-bound nucleus.

In larger organisms, cells combine to make **tissues,** which are groups of similar cells carrying out similar or related functions. **Organs** are collections of tissues grouped together performing a common function. Organs are present not only in animals but also in plants. An organ system is a higher level of organization that consists of functionally related organs. Mammals have many **organ systems**. For instance, the circulatory system transports blood through the body and to and from the lungs; it includes organs such as the heart and blood vessels. **Organisms** are individual living entities. For example, each tree in a forest is an organism. Single-celled prokaryotes and single-celled eukaryotes are also considered organisms and are typically referred to as microorganisms.

All the individuals of a species living within a specific area are collectively called a **population.** For example, a forest may include many pine trees. All of these pine trees represent the population of pine trees in this forest. Different populations may live in the same specific area. For example,

the forest with the pine trees includes populations of flowering plants and also insects and microbial populations. A **community** is the sum of populations inhabiting a particular area. For instance, all of the trees, flowers, insects, and other populations in a forest form the forest's community. The forest itself is an ecosystem. An **ecosystem** consists of all the living things in a particular area together with the abiotic, non-living parts of that environment such as nitrogen in the soil or rain water. At the highest level of organization, the biosphere is the collection of all ecosystems, and it represents the zones of life on earth. It includes land, water, and even the atmosphere to a certain extent.

A flow chart shows the hierarchy of living organisms. From smallest to largest, this hierarchy includes: (1) Organelles, such as nuclei, that exist inside cells. (2) Cells, such as a red blood cell. (3) Tissues, such as human skin tissue. (4) Organs such as the stomach make up the human digestive system, an example of an organ system. (5) Organisms, populations, and communities. In a forest, each pine tree is an organism. Together, all the pine trees make up a population. All the plant and animal species in the forest comprise a community. (6) Ecosystems: the coastal ecosystem in the Southeastern United States includes living organisms and the environment in which they live. (7) The biosphere: encompasses all the ecosystems on Earth.

Figure 1 *The biological levels of organization of living things are shown. From a single organelle to the entire biosphere, living organisms are parts of a highly structured hierarchy. (credit "organelles": modification of work by Umberto Salvagnin; credit "cells": modification of work by Bruce Wetzel, Harry Schaefer/ National Cancer Institute; credit "tissues": modification of work by Kilbad; Fama Clamosa; Mikael Häggström; credit "organs": modification of work by Mariana Ruiz Villareal; credit "organisms": modification of work by "Crystal"/Flickr; credit "ecosystems": modification of work by US Fish and Wildlife Service Headquarters; credit "biosphere": modification of work by NASA)*

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=245#h5p-24

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 25, 2017 https://cnx.org/contents/GFy_h8cu@10.99:gNLp76vu@13/Themes-and-Concepts-of-Biology

The Diversity of Life

The fact that biology, as a science, has such a broad scope has to do with the tremendous diversity of life on earth. The source of this diversity is evolution, the process of gradual change during which new species arise from older species. Evolutionary biologists study the evolution of living things in everything from the microscopic world to ecosystems.

The evolution of various life forms on Earth can be summarized in a phylogenetic tree (Figure 1). A phylogenetic tree is a diagram showing the evolutionary relationships among biological species based on similarities and differences in genetic or physical traits or both. A phylogenetic tree is composed of branches (the lines) and nodes (places where two lines diverge). The internal nodes represent ancestors and are points in evolution when, based on scientific evidence, an ancestor is thought to have diverged to form two new species. The length of each branch is proportional to the time elapsed since the split.

This phylogenetic tree shows that the three domains of life, bacteria, archaea and eukarya, all arose from a common ancestor.

Figure 1 This phylogenetic tree was constructed by microbiologist Carl Woese using data obtained from sequencing ribosomal RNA genes. The tree shows the separation of living organisms into three domains: Bacteria, Archaea, and Eukarya. Bacteria and Archaea are prokaryotes, single-celled organisms lacking intracellular organelles. (credit: Eric Gaba; NASA Astrobiology Institute)

While this is the most common way that is used to group organisms, other divisions have been proposed.

- Some scientists believe that organisms should be divided into two groups: Prokaryota (or Monera) and Eukaryota. In this method, Archae is typically included in Prokaryota. This view has become less popular due to scientific advancements, specifically genetic analysis of various organisms.
- Another two-group division groups Archae with Eukaryotes. This is often called the "Eocyte hypothesis". This hypothesis has become more popular as the genomes of more Archaeic organisms are sequenced.

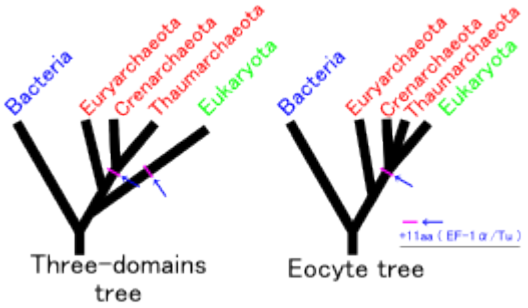


Figure 2 The relationship between Archae (in red) and Eukaryotes (green) may be closer than you think. Figure credit: Crion, Wikimedia. https://commons.wikimedia.org/wiki/File:Eocyte_hypothesis.png

Viruses

None of the three systems currently

include non-cellular life. As of 2011 there is talk about Nucleocytoplas mic large DNA viruses possibly being a fourth branch domain of life, a view supported by

researchers in 2012.

Stefan Luketa in 2012 proposed a five-domain system, adding Prionobiota (acellular and without nucleic acid) and Virusobiota

(acellular but with nucleic acid) to the traditional three domains.

Evolution Connection

Carl Woese and the Phylogenetic Tree

In the past, biologists grouped living organisms into five kingdoms: animals, plants, fungi, protists, and bacteria. The organizational scheme was based mainly on physical features, as opposed to physiology, biochemistry, or molecular biology, all of which are used by modern systematics. The pioneering work of American microbiologist Carl Woese in the early 1970s has shown, however, that life on Earth has evolved along three lineages, now called domains—Bacteria, Archaea, and Eukarya. The first two are prokaryotic cells with microbes that lack membrane-enclosed nuclei and organelles. The third domain contains the eukaryotes and includes unicellular microorganisms together with the four original kingdoms (excluding bacteria). Woese defined Archaea as a new domain, and this resulted in a new taxonomic tree (Figure 1). Many organisms belonging to the Archaea domain live under extreme conditions and are called extremophiles. To construct his tree,

Woese used genetic relationships rather than similarities based on morphology (shape).

Woese's tree was constructed from comparative sequencing of the genes that are universally distributed, present in every organism, and conserved (meaning that these genes have remained essentially unchanged throughout evolution). Woese's approach was revolutionary because comparisons of physical features are insufficient to differentiate between the prokaryotes that appear fairly similar in spite of their tremendous biochemical diversity and genetic variability (Figure 3). The comparison of homologous DNA and RNA sequences provided Woese with a sensitive device that revealed the extensive variability of prokaryotes, and which justified the separation of the prokaryotes into two domains: bacteria and archaea.

Photo depict: A: bacterial cells. Photo depict: B: a natural hot vent. Photo depict: C: a sunflower. Photo depict: D: a lion.

Figure 4 *These images represent different domains. The (a) bacteria in this micrograph belong to Domain Bacteria, while the (b) extremophiles (not visible) living in this hot vent belong to Domain Archaea. Both the (c) sunflower and (d) lion are part of Domain Eukarya. (credit a: modification of work by Drew March; credit b: modification of work by Steve Jurvetson; credit c: modification of work by Michael Arrighi; credit d: modification of work by Leszek Leszcynski)*

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=248#h5p-25

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=248#h5p-26

REFERENCES

Text adapted from:

OpenStax, Concepts of Biology. OpenStax CNX. May 25, 2017 https://cnx.org/contents/GFy_h8cu@10.99:gNLp76vu@13/Themes-and-Concepts-of-Biology

Eocyte Hypothesis, Wikipedia. May 25, 2017. https://en.wikipedia.org/wiki/Eocyte_hypothesis

Domain (biology), Wikipedia. May 25, 2017. https://en.wikipedia.org/wiki/Domain_(biology)

CELL STRUCTURE AND FUNCTION

Learning Objectives

Course Objectives for this section:

1. **Explain how basic units of cellular structure define the function of all living things.**

- Explain how various cell structures participate in the function of a cell and/or organism.
- Discuss the role of evolution in shaping cellular structure and function.

Close your eyes and picture a brick wall. What is the basic building block of that wall? It is a single brick, of course. Like a brick wall, your body is composed of basic building blocks, and the building blocks of your body are cells (**Figure 1a-c**).

Your body has many kinds of cells, each specialized for a specific purpose. Just as a home is made from a variety of building materials, the human body is constructed from many cell types. For example, epithelial cells protect the surface of the body and cover the organs and body cavities

within. Bone cells help to support and protect the body. Cells of the immune system fight invading bacteria. Additionally, red blood cells carry oxygen throughout the body. Each of these cell types plays a vital role during the growth, development, and day-to-day maintenance of the body. In spite of their enormous variety, however, all cells share certain fundamental characteristics.

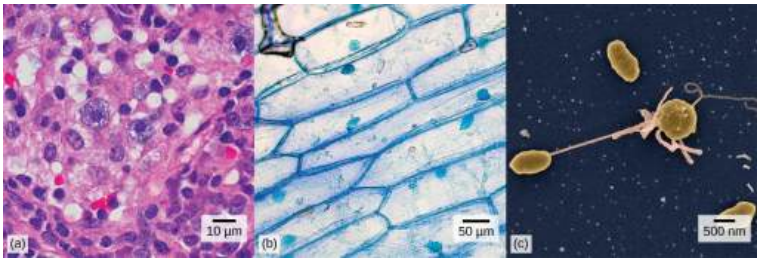


Figure 1 *(a) Nasal sinus cells (viewed with a light microscope), (b) onion cells (viewed with a light microscope), and (c) Vibrio tasmaniensis bacterial cells (viewed using a scanning electron microscope) are from very different organisms, yet all share certain characteristics of basic cell structure. (credit a: modification of work by Ed Uthman, MD; credit b: modification of work by Umberto Salvagnin; credit c: modification of work by Anthony D'Onofrio; scale-bar data from Matt Russell)*

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

How Cells Are Studied

A cell is the smallest unit of a living thing. A living thing, like you, is called an organism. Thus, cells are the basic building blocks of all organisms.

In multicellular organisms, **cells** of one particular cell type interconnect with each other and perform shared functions to form **tissues** (for example, muscle tissue, connective tissue, and nervous tissue), several tissues combine to form an **organ** (for example, stomach, heart, or brain), and several organs make up an **organ system** (such as the digestive system, circulatory system, or nervous system). Several systems functioning together form an **organism** (such as an elephant, for example).

There are many types of cells, and all are grouped into one of two broad categories: **prokaryotic** and **eukaryotic**. Animal cells, plant cells, fungal cells, and protist cells are classified as eukaryotic, whereas bacteria and archaea cells are classified as prokaryotic. Before discussing the criteria for determining whether a cell is prokaryotic or eukaryotic, let us first examine how biologists study cells.

MICROSCOPY

Cells vary in size. With few exceptions, individual cells are too small to be seen with the naked eye, so scientists use

microscopes to study them. A **microscope** is an instrument that magnifies an object. Most images of cells are taken with a microscope and are called micrographs.

LIGHT MICROSCOPES

To give you a sense of the size of a cell, a typical human red blood cell is about eight millionths of a meter or eight micrometers (abbreviated as μm) in diameter; the head of a pin is about two thousandths of a meter (millimeters, or mm) in diameter. That means that approximately 250 red blood cells could fit on the head of a pin.

The optics of the lenses of a light microscope changes the orientation of the image. A specimen that is right-side up and facing right on the microscope slide will appear upside-down and facing left when viewed through a microscope, and vice versa. Similarly, if the slide is moved left while looking through the microscope, it will appear to move right, and if moved down, it will seem to move up. This occurs because microscopes use two sets of lenses to magnify the image. Due to the manner in which light travels through the lenses, this system of lenses produces an inverted image (binoculars and a dissecting microscope work in a similar manner, but include an additional magnification system that makes the final image appear to be upright).

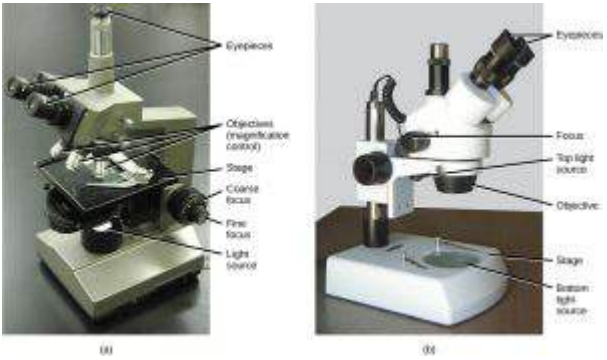


Figure 1 (a) Most light microscopes used in a college biology lab can magnify cells up to approximately 400 times. (b) Dissecting microscopes have a lower magnification than light microscopes and are used to examine larger objects, such as tissues.

Most student microscopes are classified as light microscopes (**Figure 1a**). Visible light both passes through and is bent by the lens system to enable the user to see the specimen. Light microscopes are advantageous for viewing living organisms, but since individual cells are generally transparent, their components are not distinguishable unless they are colored with special stains. Staining, however, usually kills the cells.

Light microscopes commonly used in the undergraduate college laboratory magnify up to approximately 400 times. Two parameters that are important in microscopy are magnification and resolving power. **Magnification** is the degree of enlargement of an object. **Resolving power** is the ability of a microscope to allow the eye to distinguish two adjacent structures as separate; the higher the resolution, the closer those two objects can be, and the better the clarity and detail of the image. When oil immersion lenses are used, magnification is usually increased to 1,000 times for the study of smaller cells, like most prokaryotic cells. Because light entering a specimen from below is focused onto the eye of an observer, the specimen can be viewed using light

microscopy. For this reason, for light to pass through a specimen, the sample must be thin or translucent.

A second type of microscope used in laboratories is the dissecting microscope (**Figure 1b**). These microscopes have a lower magnification (20 to 80 times the object size) than light microscopes and can provide a three-dimensional view of the specimen. Thick objects can be examined with many components in focus at the same time. These microscopes are designed to give a magnified and clear view of tissue structure as well as the anatomy of the whole organism.

Like light microscopes, most modern dissecting microscopes are also binocular, meaning that they have two separate lens systems, one for each eye. The lens systems are separated by a certain distance, and therefore provide a sense of depth in the view of their subject to make manipulations by hand easier. Dissecting microscopes also have optics that correct the image so that it appears as if being seen by the naked eye and not as an inverted image. The light illuminating a sample under a dissecting microscope typically comes from above the sample, but may also be directed from below.

ELECTRON MICROSCOPES

In contrast to light microscopes, electron microscopes use a beam of electrons instead of a beam of light (**Figure 2**). Not only does this allow for higher magnification and, thus, more detail, it also provides higher resolving power. Preparation of a specimen for viewing under an electron microscope will kill it; therefore, live cells cannot be viewed using this type of microscopy. In addition, the electron beam moves best in a vacuum, making it impossible to view living materials. There are two major types of electron microscopes which differ in the images they provide:

- In a scanning electron microscope (SEM) (**Figure 3**), a beam of electrons moves back and forth across a cell's surface, rendering the details of cell surface characteristics by reflection. Cells and other structures are usually coated with a metal like gold.
- In a transmission electron microscope (TEM), the electron beam is transmitted through the cell and provides details of a cell's internal structures. As you might imagine, electron microscopes are significantly more bulky and expensive than are light microscopes.

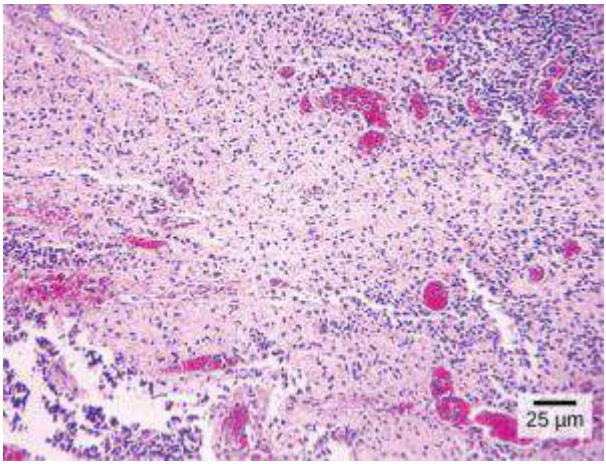


Figure 2 Salmonella bacteria are viewed with a light microscope. (credit: credit a: modification of work by CDC, Armed Forces Institute of Pathology, Charles N. Farmer)

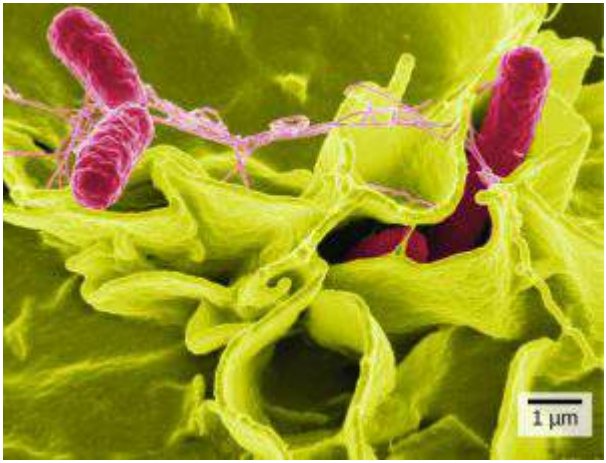


Figure 3 This scanning electron micrograph (SEM) shows Salmonella bacteria (in red) invading human cells. (credit: modification of work by Rocky Mountain Laboratories, NIAID, NIH; scale-bar data from Matt Russell)

CELL THEORY

The microscopes we use today are far more complex than those used in the 1600s by Antony van Leeuwenhoek, a Dutch shopkeeper who had great skill in crafting lenses. Despite the limitations of his now-ancient lenses, van Leeuwenhoek observed the movements of protists (a type of single-celled organism) and sperm, which he collectively termed "animalcules."

In a 1665 publication called *Micrographia*, experimental scientist Robert Hooke coined the term "cell" (from the Latin *cella*, meaning "small room") for the box-like structures he observed when viewing cork tissue through a lens. In the 1670s, van Leeuwenhoek discovered bacteria and protozoa. Later advances in lenses and microscope construction

enabled other scientists to see different components inside cells.

By the late 1830s, botanist Matthias Schleiden and zoologist Theodor Schwann were studying tissues and proposed the **unified cell theory**. This theory has three principles which still stand today. They are:

1. *All living things are composed of one or more cells.*
2. *The cell is the basic unit of life.*
3. *All new cells arise from existing cells*.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=93#h5p-27

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Comparing Prokaryotic and Eukaryotic Cells

Cells fall into one of two broad categories: **prokaryotic** and **eukaryotic.** The predominantly single-celled organisms of the domains Bacteria and Archaea are classified as prokaryotes (*pro*– = before; –*karyon*– = nucleus). Animal cells, plant cells, fungi, and protists are eukaryotes (*eu*– = true).

All cells share four common components: 1) a plasma membrane, an outer covering that separates the cell's interior from its surrounding environment; 2) cytoplasm, consisting of a gel-like region within the cell in which other cellular components are found; 3) DNA, the genetic material of the cell; and 4) ribosomes, particles that synthesize proteins.

COMPONENTS OF PROKARYOTIC CELLS

Prokaryotes differ from eukaryotic cells in several important ways. A **prokaryotic cell** is a simple, single-celled (unicellular) organism that lacks a nucleus, or any other membrane-bound organelle. We will shortly come to see that this is significantly different in eukaryotes. Prokaryotic DNA is found in the central part of the cell: a darkened region called the nucleoid (**Figure 1**).

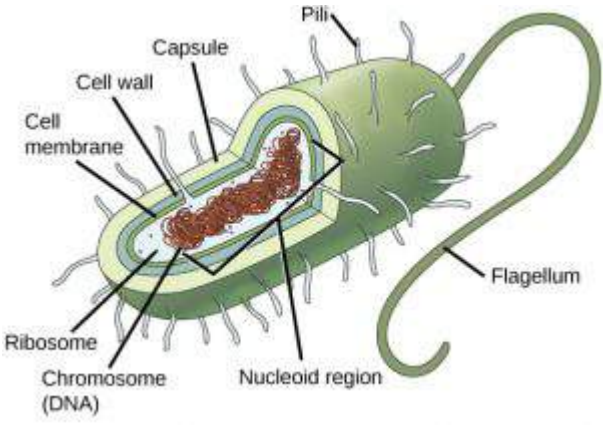


Figure 1 This figure shows the generalized structure of a prokaryotic cell.

Unlike Archaea and eukaryotes, bacteria have a cell wall made of peptidoglycan, comprised of sugars and amino acids, and many have a polysaccharide (carbohydrate) capsule (**Figure 1**). The cell wall acts as an extra layer of protection, helps the cell maintain its shape, and prevents dehydration. The capsule enables the cell to attach to surfaces in its environment. Some prokaryotes have flagella, pili, or fimbriae. Flagella are used for locomotion, while most pili are used to exchange genetic material during a type of reproduction called conjugation.

COMPONENTS OF EUKARYOTIC CELLS

In nature, the relationship between form and function is apparent at all levels, including the level of the cell, and this will become clear as we explore eukaryotic cells. The principle “form follows function” is found in many contexts. For example, birds and fish have streamlined bodies that allow them to move quickly through the medium in which they live, be it air or water. It means that, in general, one

can deduce the function of a structure by looking at its form, because the two are matched.

A **eukaryotic cell** is a cell that has a membrane-bound nucleus and other membrane-bound compartments or sacs, called **organelles**, which have specialized functions. The rest of this chapter will discuss functions of the various organelles. The word eukaryotic means "true kernel" or "true nucleus," alluding to the presence of the membrane-bound nucleus in these cells. The word "organelle" means "little organ," and, as already mentioned, organelles have specialized cellular functions, just as the organs of your body have specialized functions.

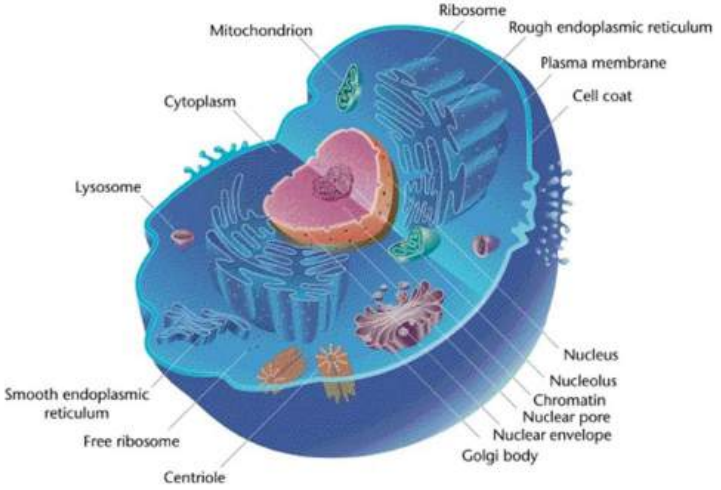
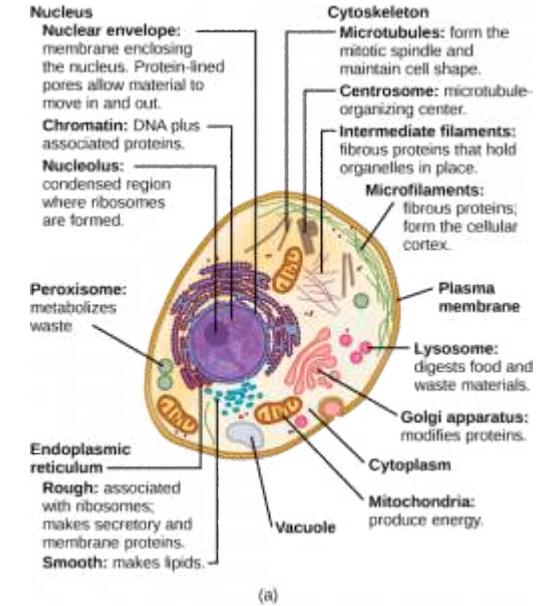
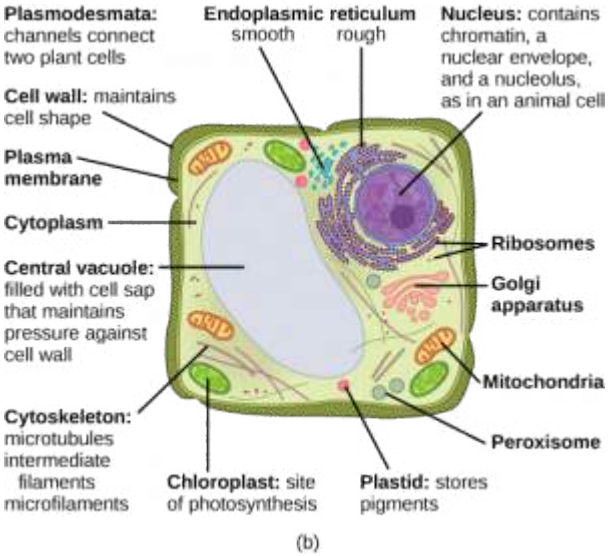


Figure 2 A generalized eukaryotic cell showing some of the organelles. (Photo credit: Mediran, Wikimedia. 14 Aug 2002)

Both animals and plants are eukaryotes. Despite their fundamental similarities, there are some striking differences between animal and plant cells. Animal cells have centrioles, centrosomes (discussed under the cytoskeleton), and lysosomes, whereas plant cells do not. Plant cells have a cell wall, chloroplasts, plasmodesmata, and plastids used for

storage, and a large central vacuole, whereas animal cells do not.





An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=96#h5p-28

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=96#h5p-29

CELL SIZE

At 0.1–5.0 μm in diameter, prokaryotic cells are significantly smaller than eukaryotic cells, which have diameters ranging from 10–100 μm (**Figure 3**). The small size of prokaryotes allows ions and organic molecules that enter them to quickly spread to other parts of the cell. Similarly, any wastes produced within a prokaryotic cell can quickly move out. However, larger eukaryotic cells have evolved different structural adaptations to enhance cellular transport. Indeed, the large size of these cells would not be possible without these adaptations. In general, cell size is limited because volume increases much more quickly than does cell surface area. As a cell becomes larger, it becomes more and more

difficult for the cell to acquire sufficient materials to support the processes inside the cell, because the relative size of the surface area across which materials must be transported declines.

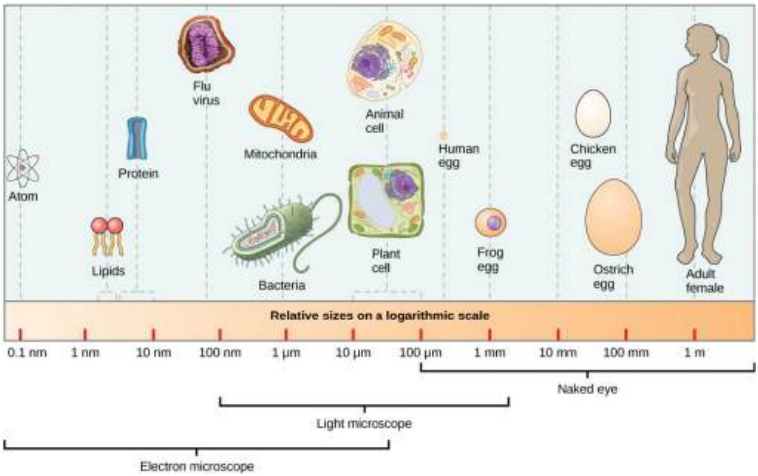


Figure 3 This figure shows the relative sizes of different kinds of cells and cellular components. An adult human is shown for comparison.

Small size, in general, is necessary for all cells, whether prokaryotic or eukaryotic. Let's examine why that is so. First, we'll consider the area and volume of a typical cell. Not all cells are spherical in shape, but most tend to approximate a sphere. You may remember from your geometry course that the formula for the surface area of a sphere is $4\pi r^2$, while the formula for its volume is $4\pi r^3/3$. Thus, as the radius of a cell increases, its surface area increases as the square of its radius, but its volume increases as the cube of its radius (much more rapidly). Therefore, as a cell increases in size, its surface area-to-volume ratio decreases. This same principle would apply if the cell had the shape of a cube (Figure 4). If the cell grows too large, the plasma membrane will not have sufficient surface area to support the rate of diffusion required for the increased volume. In other words, as a cell

grows, it becomes less efficient. One way to become more efficient is to divide; another way is to develop organelles that perform specific tasks. These adaptations lead to the development of more sophisticated cells called eukaryotic cells.

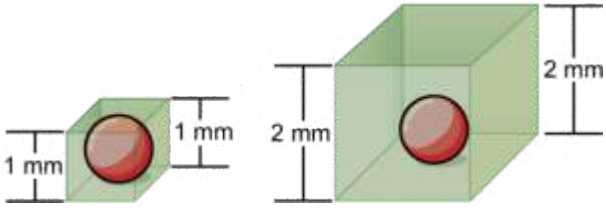


Figure 4 Volume increases faster than surface area. The surface area of the small cell is 1mm x 1mm x 6 sides = 6mm^2. The volume of the small cell is 1mm x 1mm x 1mm = 1mm^3. This gives a surface area to volume ratio of 6:1. The surface area of the larger cell is 2mm x 2mm x 6 sides = 24mm^2. The volume of the large cell is 2mm x 2mm x 2mm = 8mm^3. This gives a surface area to volume ratio of 3:1 (24:8 reduces to 3:1).

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=96#h5p-30

REFERENCES

Text adapted from: OpenStax, Concepts of Biology.

OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Plasma Membrane and The Cytoplasm

At this point, it should be clear that eukaryotic cells have a more complex structure than do prokaryotic cells. Organelles allow for various functions to occur in the cell at the same time. Before discussing the functions of organelles within a eukaryotic cell, let us first examine two important components of all cells (prokaryotic and eukaryotic): the plasma membrane and the cytoplasm.

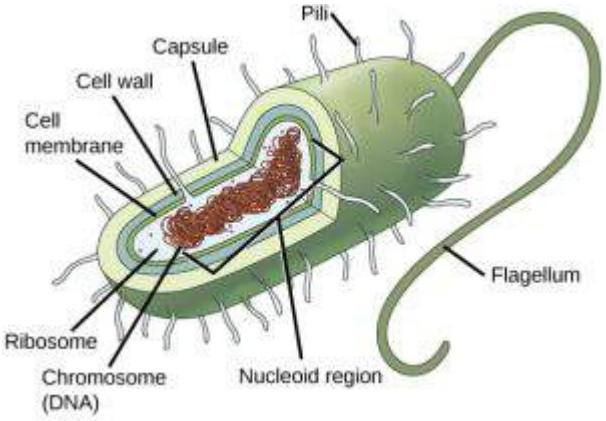


Figure 1 *A prokaryotic cell. The cytoplasm is not labeled, but is the light blue area inside the cell membrane. The ribosome label is pointing to one of the small brown dots representing the ribosome.*

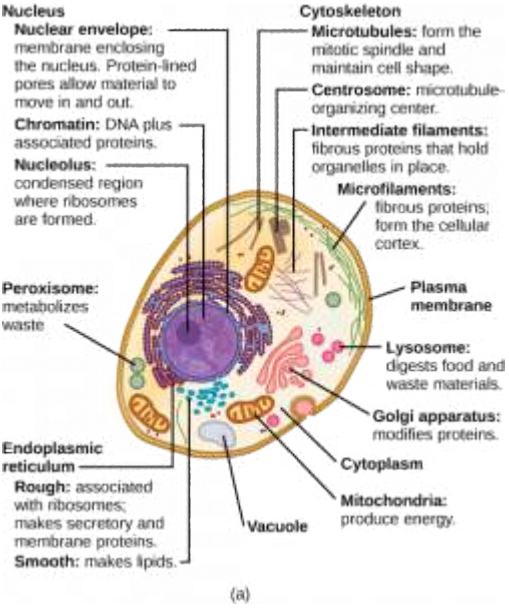


Figure 2 This figure shows a typical animal cell

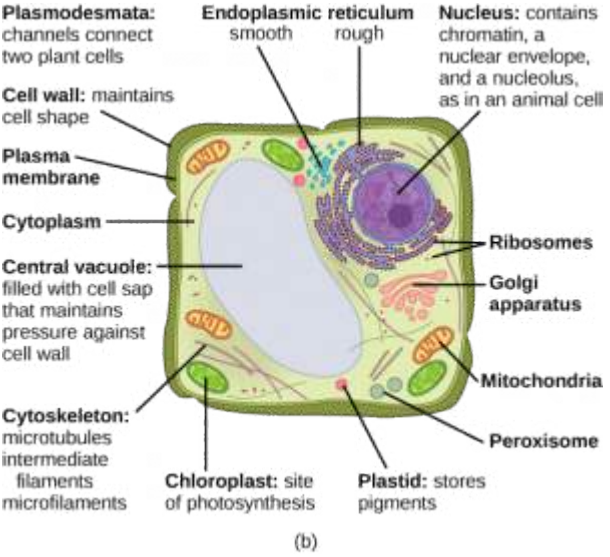


Figure 3 This figure shows a typical plant cell.

THE PLASMA MEMBRANE

Like prokaryotes, eukaryotic cells have a **plasma membrane** (Find it in **Figures 1-3**, then look at the detailed structure in **Figure 4**) made up of a phospholipid bilayer with embedded proteins that separates the internal contents of the cell from its surrounding environment. A phospholipid is a lipid molecule composed of two fatty acid chains, a glycerol backbone, and a phosphate group. The plasma membrane regulates the passage of some substances, such as organic molecules, ions, and water, preventing the passage of some to maintain internal conditions, while actively bringing in or removing others. Other compounds move passively across the membrane.

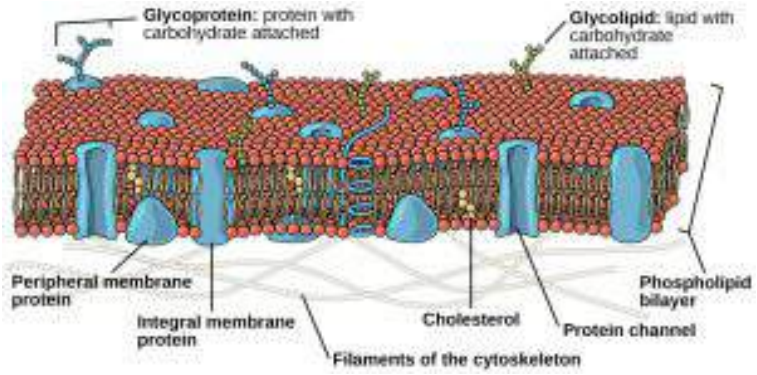


Figure 4 The plasma membrane is a phospholipid bilayer with embedded proteins. There are other components, such as cholesterol and carbohydrates, which can be found in the membrane in addition to phospholipids and protein.

The plasma membranes of cells that specialize in absorption are folded into fingerlike projections called **microvilli** (singular = microvillus). This folding increases the surface area of the plasma membrane. Such cells are typically found lining the small intestine, the organ that absorbs nutrients

from digested food (Figure 5). This is an excellent example of form matching the function of a structure.

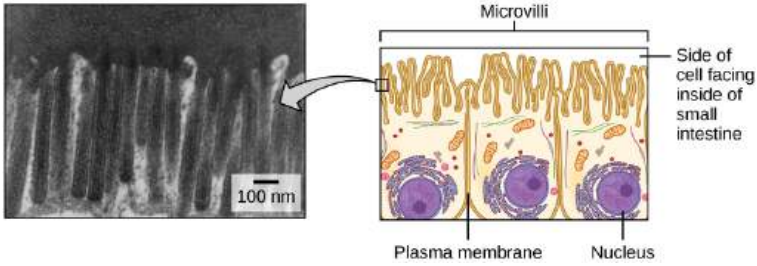


Figure 5 Microvilli, shown here as they appear on cells lining the small intestine, increase the surface area available for absorption. These microvilli are only found on the area of the plasma membrane that faces the cavity from which substances will be absorbed. (credit "micrograph": modification of work by Louisa Howard)

THE CYTOPLASM

The **cytoplasm** comprises the contents of a cell between the plasma membrane and the nuclear envelope (a structure to be discussed shortly). It is made up of organelles suspended in the gel-like **cytosol**, the cytoskeleton, and various chemicals (Find it in **Figures 1-3**). Even though the cytoplasm consists of 70 to 80 percent water, it has a semi-solid consistency, which comes from the proteins within it. However, proteins are not the only organic molecules found in the cytoplasm. Glucose and other simple sugars, polysaccharides, amino acids, nucleic acids, fatty acids, and derivatives of glycerol are found there too. Ions of sodium, potassium, calcium, and many other elements are also dissolved in the cytoplasm. Many metabolic reactions, including protein synthesis, take place in the cytoplasm. Take note that the cytoplasm is not "empty" or "filler" – it is a vitally

important component of cells that allows chemical reactions to take place!

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Ribosomes

Ribosomes are the cellular structures responsible for protein synthesis. The word "synthesis" means "to combine things to produce something else." In this context, protein synthesis means combining different amino acids together to form a protein. Ribosomes join amino acids together in a chain to form a protein (**Figure 1**). This amino acid chain then folds into a complex 3-dimensional structure. The shape of a protein is what gives the protein its specific function.

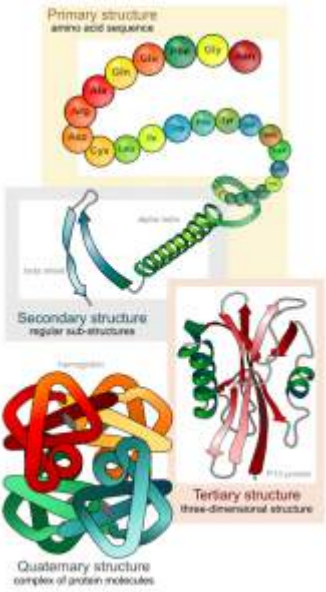


Figure 1 Protein structure. The colored balls at the top of this diagram represent different amino acids. Amino acids are the subunits that are joined together by the ribosome to form a protein. This chain of amino acids then folds to form a complex 3D structure. (Credit: Lady of Hats from Wikipedia; public domain)

Helpful Hint: Proteins are not typically used as a source of energy for the body. Protein from your diet is broken down into individual amino acids which are reassembled by your ribosomes into proteins that your cells need. Ribosomes do not produce energy.

When viewed through an electron microscope, free

ribosomes appear as either clusters or single tiny dots floating freely in the cytoplasm. Ribosomes may be attached to either the cytoplasmic side of the plasma membrane or the cytoplasmic side of the rough endoplasmic reticulum (Figure 2).

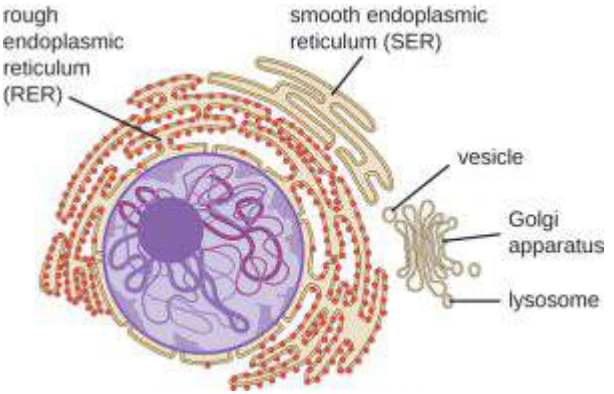


Figure 2 Ribosomes can be found free in the cytoplasm (not shown in this diagram), or attached to the outer membrane of the nucleus and the rough endoplasmic reticulum (RER). Credit CFCF; Wikimedia; CC license.

Because protein synthesis is essential for all cells, ribosomes are found in practically every cell, although they are smaller in prokaryotic cells. They are particularly abundant in immature red blood cells for the synthesis of hemoglobin, which functions in the transport of oxygen throughout the body. Electron microscopy has shown us that ribosomes, which are large complexes of protein and RNA, consist of two subunits, aptly called large and small (**Figure 3**). Ribosomes receive their "orders" for protein synthesis from the nucleus where the DNA is transcribed into messenger RNA (mRNA). The mRNA travels to the ribosomes, which translate the code provided by the sequence of the nitrogenous bases in the mRNA into a specific order of amino acids in a protein. Amino acids are the building blocks of proteins.

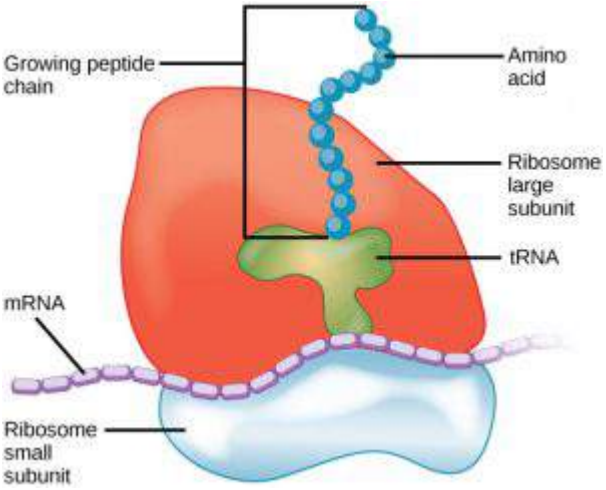


Figure 3 Ribosomes are made up of a large subunit (top) and a small subunit (bottom). During protein synthesis, ribosomes assemble amino acids into proteins.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Cytoskeleton

If you were to remove all the organelles from a cell, would the plasma membrane and the cytoplasm be the only components left? No. Within the cytoplasm, there would still be ions and organic molecules, plus a network of protein fibers known as the **cytoskeleton**.

Both prokaryotes and eukaryotes have a cytoskeleton. Both types of organisms use their cytoskeleton for cell division, protection, and shape determination.

In addition, in eukaryotes the cytoskeleton also functions to secure certain organelles in specific positions, and to allow cytoplasm and vesicles to move within the cell. It also enables unicellular organisms to move independently. There are three types of fibers within the cytoskeleton: microfilaments, also known as actin filaments, intermediate filaments, and microtubules (**Figure 1**).

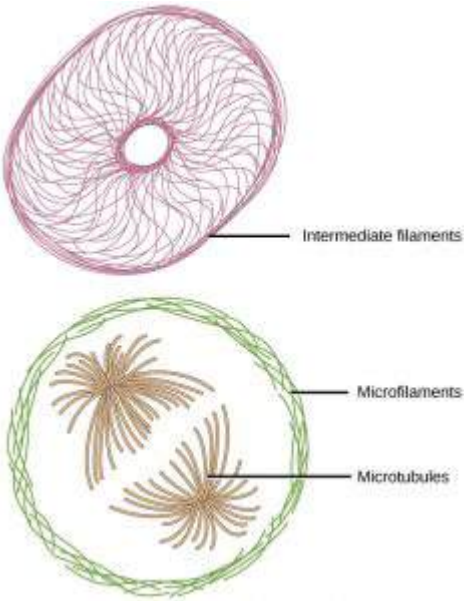


Figure 1 Microfilaments, intermediate filaments, and microtubules compose a cell's cytoskeleton.

MICROFILAMENTS

Of the three types of protein fibers in the cytoskeleton, **microfilaments** are the narrowest. They function in cellular movement, have a diameter of about 7 nm, and are made of two intertwined strands of a globular protein called actin. For this reason, microfilaments are also known as actin filaments.

ATP is required for actin proteins to assemble into long filaments. These long actin filaments serve as a track for the movement of a motor protein called myosin. Actin and myosin are plentiful in muscle cells. When your actin and myosin filaments slide past each other, your muscles contract. Actin also enables your cells to engage in cellular

events requiring motion, such as cell division in animal cells and cytoplasmic streaming, which is the circular movement of the cell cytoplasm in plant cells.

Microfilaments also provide some rigidity and shape to the cell. They can depolymerize (disassemble) and reform quickly, thus enabling a cell to change its shape and move. White blood cells (your body's infection-fighting cells) make good use of this ability. They can move to the site of an infection and phagocytize the pathogen.

INTERMEDIATE FILAMENTS

Intermediate filaments are made of several strands of fibrous proteins that are wound together. These elements of the cytoskeleton get their name from the fact that their diameter, 8 to 10 nm, is between those of microfilaments and microtubules.

Intermediate filaments have no role in cell movement. Their function is purely structural. They bear tension, thus maintaining the shape of the cell, and anchor the nucleus and other organelles in place. Figure 1 shows how intermediate filaments create a supportive scaffolding inside the cell.

The intermediate filaments are the most diverse group of cytoskeletal elements. Several types of fibrous proteins are found in the intermediate filaments. You are probably most familiar with keratin, the fibrous protein that strengthens your hair, nails, and the epidermis of the skin.

MICROTUBULES

As their name implies, microtubules are small hollow tubes. The walls of the microtubule are made of polymerized dimers of α-tubulin and β-tubulin, two globular proteins.

With a diameter of about 25 nm, microtubules are the widest components of the cytoskeleton. They help the cell resist compression, provide a track along which vesicles move through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. Like microfilaments, microtubules can dissolve and reform quickly.

Microtubules are also the structural elements of flagella, cilia, and centrioles (the latter are the two perpendicular bodies of the centrosome). In fact, in animal cells, the centrosome is the microtubule-organizing center. In eukaryotic cells, flagella and cilia are quite different structurally from their counterparts in prokaryotes, as discussed below.

The centrosome replicates itself before a cell divides, and the centrioles play a role in pulling the duplicated chromosomes to opposite ends of the dividing cell. However, the exact function of the centrioles in cell division is not clear, since cells that have the centrioles removed can still divide, and plant cells, which lack centrioles, are capable of cell division.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Flagella and Cilia

Flagella (singular = flagellum) are long, hair-like structures that extend from the plasma membrane and are used to move an entire cell, (for example, sperm, *Euglena*). When present, the cell has just one flagellum or a few flagella. Prokaryotes sometimes have flagella, but they are structurally very different from eukaryotic flagella. Prokaryotes can have more than one flagella. They serve the same function in both prokaryotes and eukaryotes (to move an entire cell).

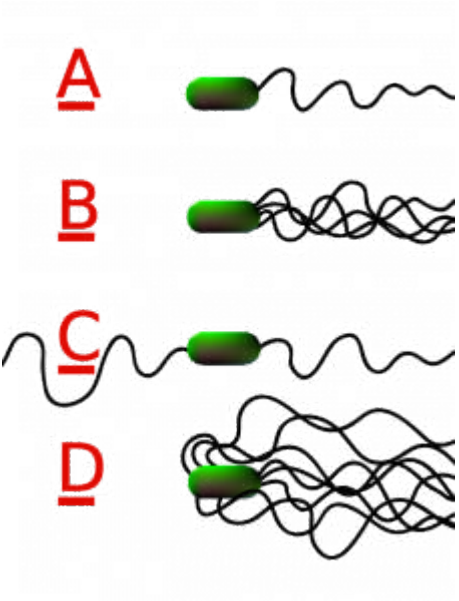


Figure 1 Examples of bacterial flagella arrangement schemes. Credit Adenosine; Wikimedia.

When **cilia** (singular = cilium) are present, however, they are many in number and extend along the entire surface of the plasma membrane. They are short, hair-like structures that are used to move entire cells (such as paramecium) or move substances along the outer surface of the cell (for example, the cilia of cells lining the fallopian tubes that move the ovum toward the uterus, or cilia lining the cells of the respiratory tract that move particulate matter toward the throat that mucus has trapped). Cilia are not found on prokaryotes.

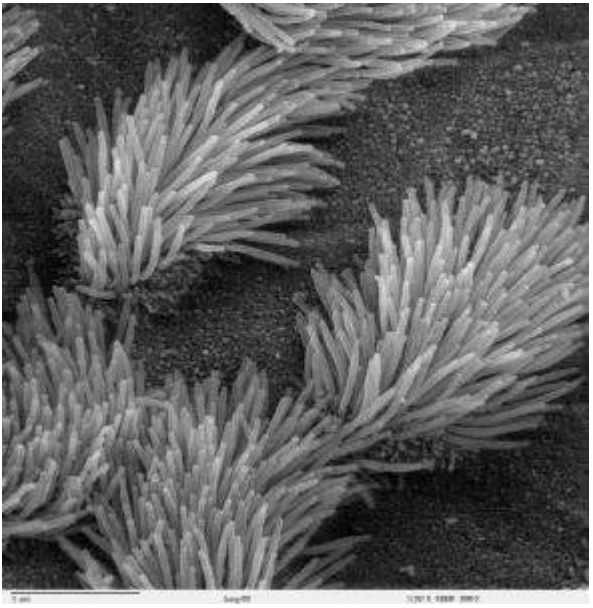


Figure 2 Scanning electron microscope image of lung trachea epithelium. There are both ciliated and non-ciliated cells in this epithelium. Note the difference in size between the cilia and the microvilli (on the non-ciliated cell surface). Photo credit Charles Daghlian; Wikimedia; public domain.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Endomembrane System

The **endomembrane system** (*endo* = within) is a group of membranes and organelles (see **Figure 1**) in eukaryotic cells that work together to modify, package, and transport lipids and proteins. It includes the nuclear envelope, lysosomes, and vesicles, the endoplasmic reticulum and Golgi apparatus, which we will cover shortly. Although not technically *within* the cell, the plasma membrane is included in the endomembrane system because, as you will see, it interacts with the other endomembranous organelles. None of the organelles that make up the endomembrane system are found in prokaryotes with the exception of the plasma membrane.

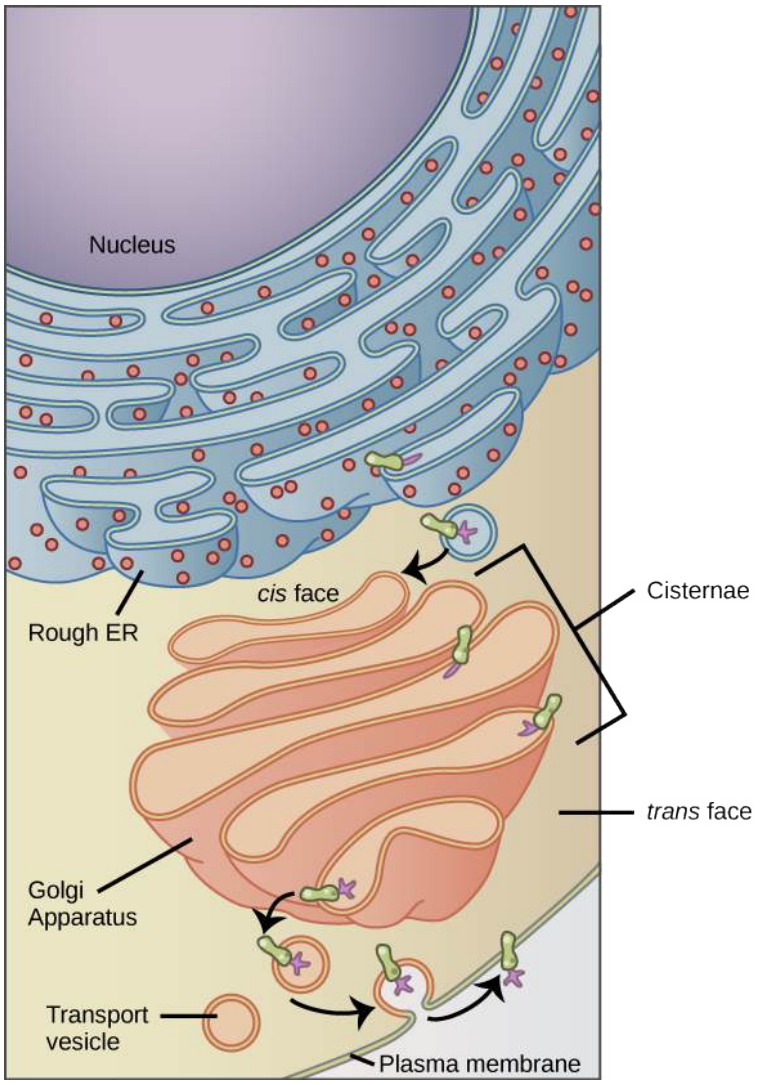


Figure 1 Membrane and secretory proteins are synthesized in the rough endoplasmic reticulum (RER). The RER also sometimes modifies proteins. In this illustration, a (green) integral membrane protein in the ER is modified by attachment of a (purple) carbohydrate. Vesicles with the integral protein bud from the ER and fuse with the cis face of the Golgi apparatus. As the protein passes along the Golgi's cisternae, it is

further modified by the addition of more carbohydrates. After its synthesis is complete, it exits as integral membrane protein of the vesicle that bud from the Golgi's trans face and when the vesicle fuses with the cell membrane the protein becomes integral portion of that cell membrane. (credit: modification of work by Magnus Manske)

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Nucleus

Typically, the nucleus is the most prominent organelle in a cell. The **nucleus** (plural = nuclei) houses the cell's DNA in the form of chromatin and directs the synthesis of ribosomes and proteins. Let us look at it in more detail (**Figure 1**).

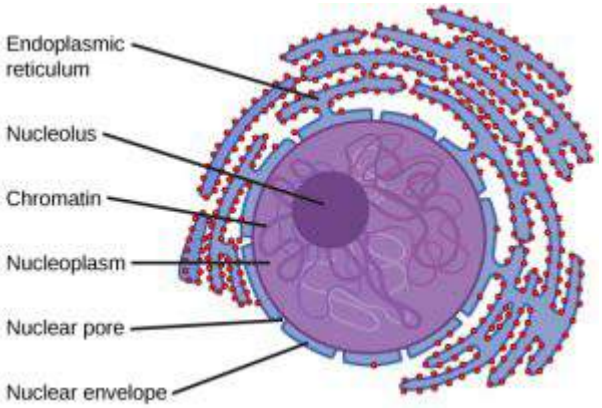


Figure 1 The outermost boundary of the nucleus is the nuclear envelope. Notice that the nuclear envelope consists of two phospholipid bilayers (membranes)—an outer membrane and an inner membrane—in contrast to the plasma membrane, which consists of only one phospholipid bilayer. (credit: modification of work by NIGMS, NIH)

The **nuclear envelope** is a double-membrane structure that constitutes the outermost portion of the nucleus (**Figure 2**).

Both the inner and outer membranes of the nuclear envelope are phospholipid bilayers.

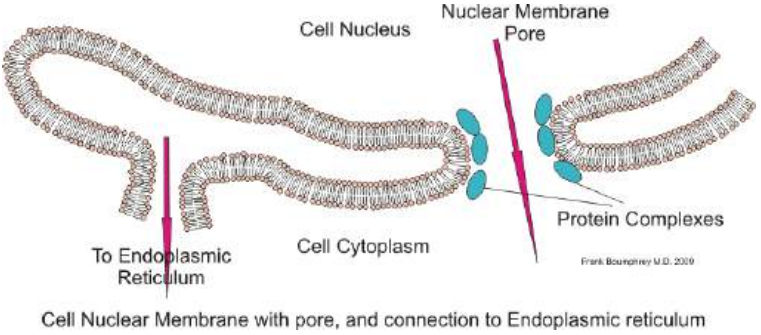


Figure 2 This illustration shows the double membrane structure surrounding the nucleus. Notice that both membranes are composed of a phospholipid bilayer. Credit Boumphreyfr; Wikimedia

CHROMATIN

The nuclear envelope is punctuated with pores that control the passage of ions, molecules, and RNA between the nucleoplasm and the cytoplasm (**Figure 2**). The nucleoplasm is the semi-solid fluid inside the nucleus, where we find the chromatin and the nucleolus.

You may remember that in prokaryotes, DNA is organized into a single circular chromosome. In eukaryotes, chromosomes are linear structures. In eukaryotes, chromosomes are structures within the nucleus that are made up of DNA, the hereditary material, and proteins. This combination of DNA and proteins is called chromatin. Every species has a specific number of chromosomes in the nucleus of its body cells. For example, in humans, the chromosome number is 46, whereas in fruit flies, the chromosome number is eight.

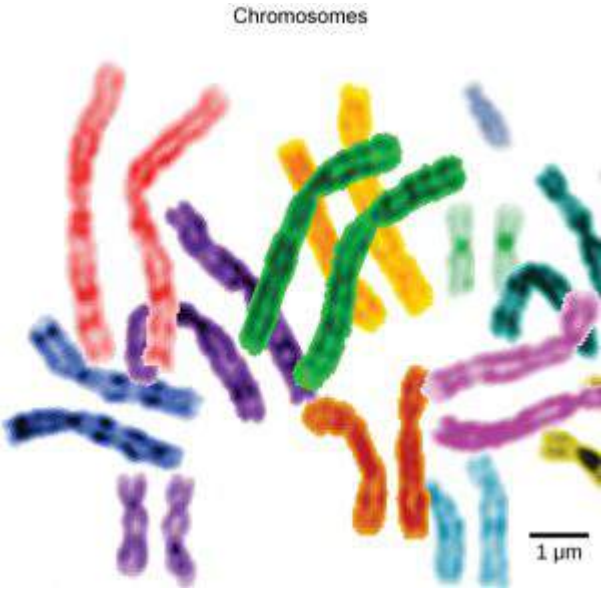


Figure 3 This image shows paired chromosomes. Each pair of chromosomes is shown in a different color. In reality, chromosomes are not colorful and typically look grayish. (Credit: modification of work by NIH; scale-bar data from Matt Russell)

Chromosomes are only visible and distinguishable from one another when the cell is getting ready to divide. When the cell is in the growth and maintenance phases of its life cycle, the chromosomes resemble an unwound, jumbled bunch of threads. These unwound protein-chromosome complexes are called chromatin (Figure 4); chromatin describes the material that makes up the chromosomes both when condensed and decondensed.

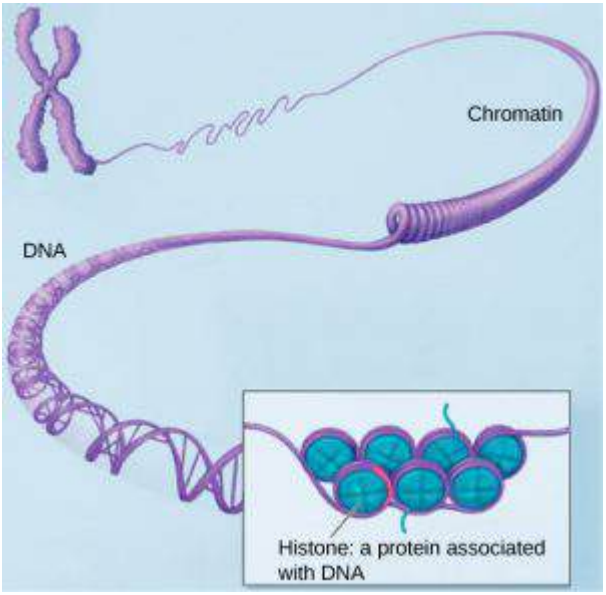


Figure 4 This image shows various levels of the organization of chromatin (DNA and protein).

NUCLEOLUS

We already know that the nucleus directs the synthesis of ribosomes, but how does it do this? Some chromosomes have sections of DNA that encode ribosomal RNA. A darkly staining area within the nucleus, called the **nucleolus** (plural = nucleoli) (See Figure 1), aggregates the ribosomal RNA with associated proteins to assemble the ribosomal subunits that are then transported through the nuclear pores into the cytoplasm.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Endoplasmic Reticulum

The **endoplasmic reticulum (ER)** is a series of interconnected membranous tubules that collectively modify proteins and synthesize lipids. However, these two functions are performed in separate areas of the endoplasmic reticulum: the **rough endoplasmic reticulum** and the **smooth endoplasmic reticulum**, respectively.

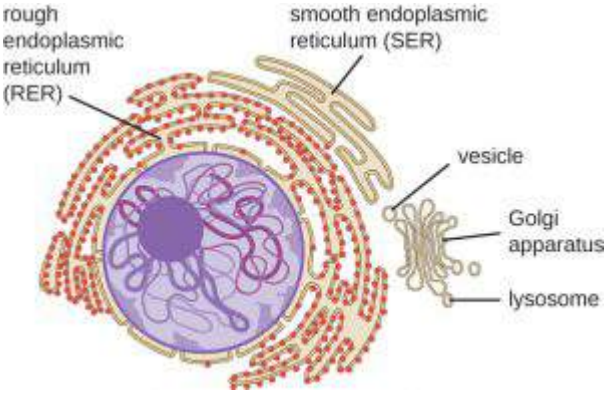


Figure 1 The rough and smooth endoplasmic reticulum are part of the endomembrane system.

The hollow portion of the ER tubules is called the lumen or cisternal space. The membrane of the ER, which is a phospholipid bilayer embedded with proteins, is continuous with the nuclear envelope (**Figure 1**).

The **rough endoplasmic reticulum (RER)** is so named

because the ribosomes attached to its cytoplasmic surface give it a studded appearance when viewed through an electron microscope (**Figure 2**). The ribosomes synthesize proteins while attached to the ER, resulting in transfer of their newly synthesized proteins into the lumen of the RER where they undergo modifications such as folding or addition of sugars. The RER also makes phospholipids for cell membranes.

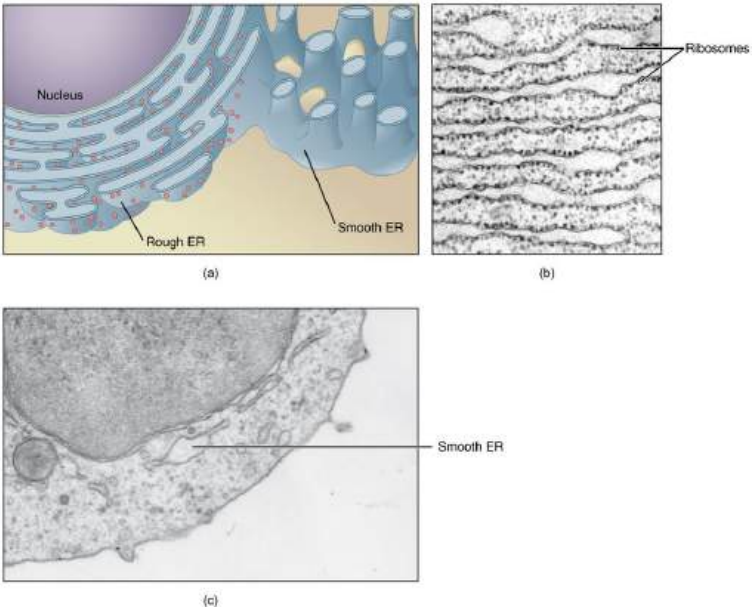


*Figure 2 (a) The ER is a winding network of thin membranous sacs found in close association with the cell nucleus. The smooth and rough endoplasmic reticula are very different in appearance and function (source: mouse tissue). **(b)** Rough ER is studded with numerous ribosomes, which are sites of protein synthesis (source: mouse tissue). EM × 110,000. **(c)** Smooth ER synthesizes phospholipids, steroid hormones, regulates the concentration of cellular Ca++, metabolizes some carbohydrates, and breaks down certain toxins (source: mouse tissue). EM × 110,510. (Micrographs provided by the Regents of University of Michigan Medical School © 2012). Figure from The Cytoplasm and Cellular Organelles; OpenStax.*

If the phospholipids or modified proteins are not destined to stay in the RER, they will be packaged within vesicles and transported from the RER by budding from the membrane (**Figure 3**). Since the RER is engaged in modifying proteins that will be secreted from the cell, it is abundant in cells that secrete proteins, such as the liver.

The **smooth endoplasmic reticulum (SER)** is continuous

with the RER but has few or no ribosomes on its cytoplasmic surface (see **Figures 1-3**). The SER's functions include synthesis of carbohydrates, lipids, and steroid hormones; detoxification of medications and poisons; alcohol metabolism; and storage of calcium ions.

In muscle cells, a specialized SER called the sarcoplasmic reticulum is responsible for storage of the calcium ions that are needed to trigger the coordinated contractions of the muscle cells.

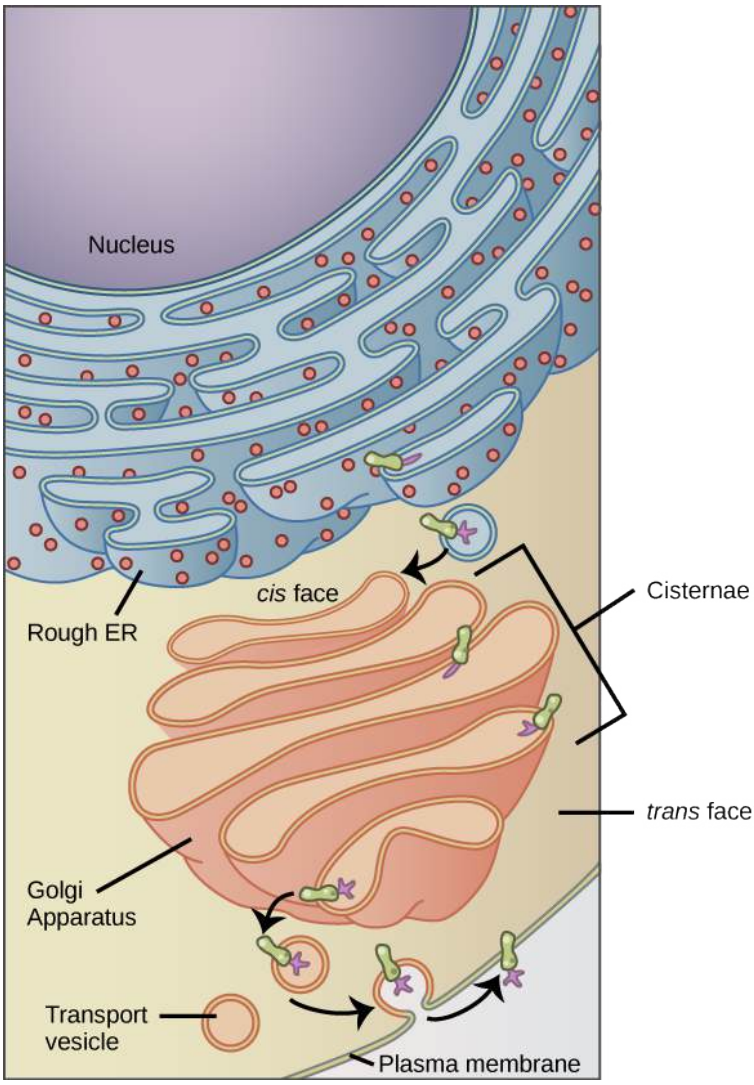


Figure 3 The endomembrane system works to modify, package, and transport lipids and proteins. (credit: modification of work by Magnus Manske)

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Golgi Apparatus

We have already mentioned that vesicles can bud from the ER, but where do the vesicles go? Before reaching their final destination, the lipids or proteins within the transport vesicles need to be sorted, packaged, and tagged so that they wind up in the right place. The sorting, tagging, packaging, and distribution of lipids and proteins take place in the **Golgi apparatus** (also called the Golgi body), a series of flattened membranous sacs (**Figure 1**).

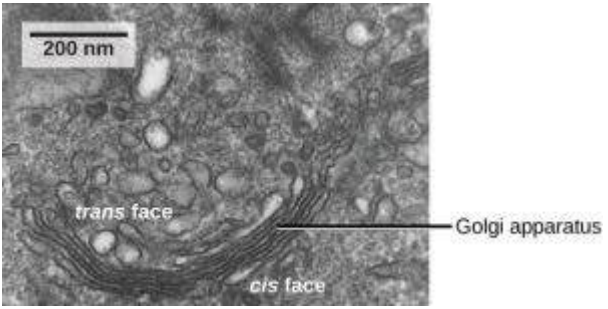


Figure 1 *The Golgi apparatus in this transmission electron micrograph of a white blood cell is visible as a stack of semicircular flattened rings in the lower portion of this image. Several vesicles can be seen near the Golgi apparatus.(credit: modification of work by Louisa Howard; scale-bar data from Matt Russell)*

The Golgi apparatus has a receiving face near the endoplasmic reticulum (the *cis* face) and a releasing face on the side away from the ER, toward the cell membrane (the

trans face) (**Figure 2**). The transport vesicles that form from the ER travel to the receiving face, fuse with it, and empty their contents into the lumen (empty space inside) of the Golgi apparatus. As the proteins and lipids travel through the Golgi, they undergo further modifications. The most frequent modification is the addition of short chains of sugar molecules. The newly modified proteins and lipids are then tagged with small molecular groups to enable them to be routed to their proper destinations.

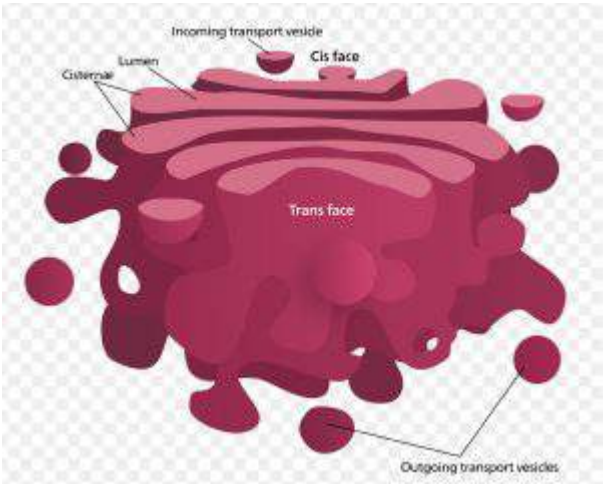


Figure 2 Diagram of the Golgi apparatus showing the cis and trans faces. The cis face would be near the nucleus while the trans face would be facing the cell membrane. Credit Kelvinsong; Wikimedia

Finally, the modified and tagged proteins are packaged into vesicles that bud from the opposite face of the Golgi. While some of these vesicles, transport vesicles, deposit their contents into other parts of the cell where they will be used, others, secretory vesicles, fuse with the plasma membrane and release their contents outside the cell.

The amount of Golgi in different cell types again illustrates that form follows function within cells. Cells that engage in a great deal of secretory activity (such as cells of the salivary

glands that secrete digestive enzymes or cells of the immune system that secrete antibodies) have an abundant number of Golgi.

In plant cells, the Golgi has an additional role of synthesizing polysaccharides, some of which are incorporated into the cell wall and some of which are used in other parts of the cell.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Vesicles and Vacuoles, Lysosomes, and Peroxisomes

VESICLES AND VACUOLES

Vesicles and **vacuoles** are membrane-bound sacs that function in storage and transport. Vacuoles are somewhat larger than vesicles, and the membrane of a vacuole does not fuse with the membranes of other cellular components. Vesicles can fuse with other membranes within the cell system (**Figure 1**). Additionally, enzymes within plant vacuoles can break down macromolecules.

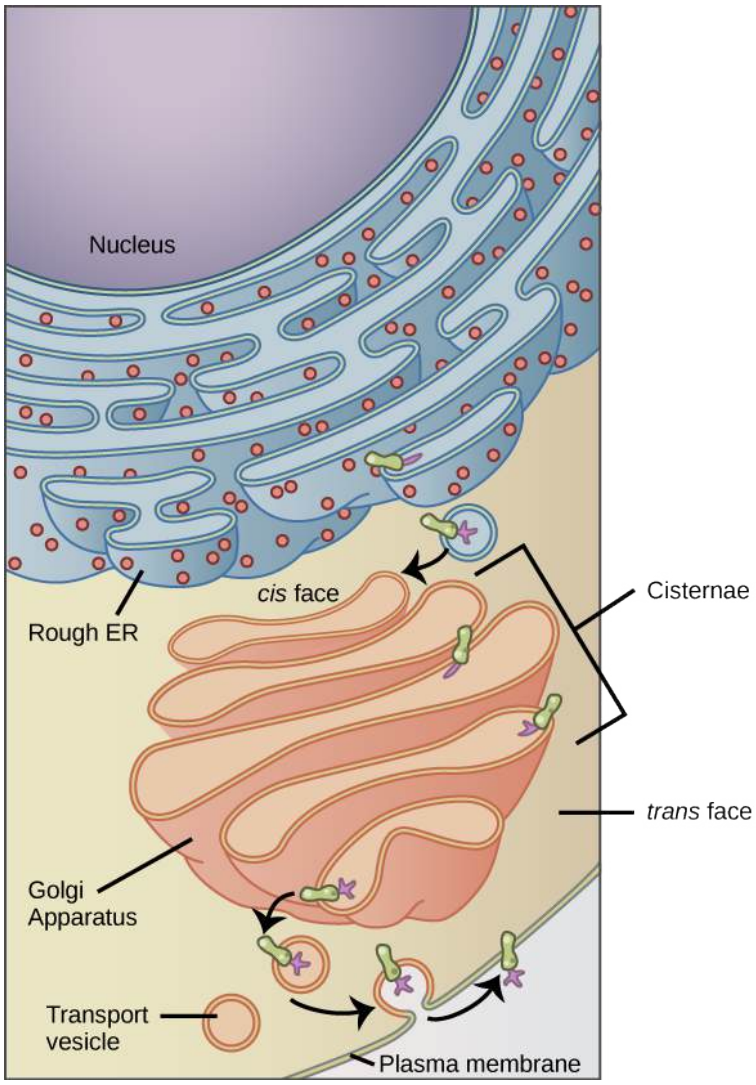


Figure 1 The endomembrane system works to modify, package, and transport lipids and proteins. (credit: modification of work by Magnus Manske)

THE CENTRAL VACUOLE (PLANTS)

Previously, we mentioned vacuoles as essential components of plant cells. If you look at **Figure 2**, you will see that plant cells each have a large, central vacuole that occupies most of the cell.

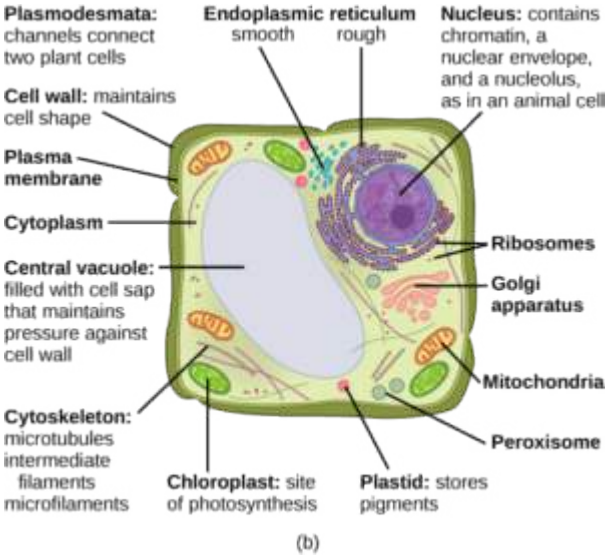


Figure 2 Diagram of a plant cell.

The **central vacuole** plays a key role in regulating the cell's concentration of water in changing environmental conditions. In plant cells, the liquid inside the central vacuole provides turgor pressure, which is the outward pressure caused by the fluid inside the cell. Have you ever noticed that if you forget to water a plant for a few days, it wilts? That is because as the water concentration in the soil becomes lower than the water concentration in the plant, water moves out of the central vacuoles and cytoplasm and into the soil. As the central vacuole shrinks, it leaves the cell wall unsupported. This loss of support to the cell walls of a plant

results in the wilted appearance. Additionally, this fluid has a very bitter taste, which discourages consumption by insects and animals. The central vacuole also functions to store proteins in developing seed cells.

LYSOSOME

In animal cells, the **lysosomes** are the cell's "garbage disposal." Digestive enzymes within the lysosomes aid the breakdown of proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles. In single-celled eukaryotes, lysosomes are important for digestion of the food they ingest and the recycling of organelles. These enzymes are active at a much lower pH (more acidic) than those located in the cytoplasm. Many reactions that take place in the cytoplasm could not occur at a low pH, thus the advantage of compartmentalizing the eukaryotic cell into organelles is apparent.

Lysosomes also use their hydrolytic enzymes to destroy disease-causing organisms that might enter the cell. A good example of this occurs in a group of white blood cells called macrophages, which are part of your body's immune system. In a process known as phagocytosis, a section of the plasma membrane of the macrophage invaginates (folds in) and engulfs a pathogen. The invaginated section, with the pathogen inside, then pinches itself off from the plasma membrane and becomes a vesicle. The vesicle fuses with a lysosome. The lysosome's hydrolytic enzymes then destroy the pathogen (**Figure 3**).

Lysosomes are basically small bags of membrane containing enzymes, so they look structurally similar to a small vacuole.

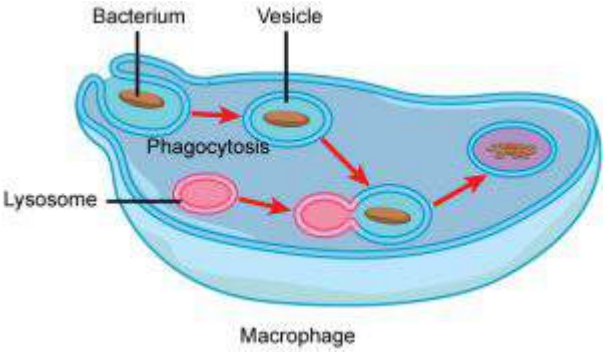


Figure 3 A macrophage has phagocytized a potentially pathogenic bacterium into a vesicle, which then fuses with a lysosome within the cell so that the pathogen can be destroyed. Other organelles are present in the cell, but for simplicity, are not shown.

PEROXISOMES

Peroxisomes are small, round organelles enclosed by single membranes (so again, they look similar to small vacuoles). They carry out oxidation reactions that break down fatty acids and amino acids. They also detoxify many poisons that may enter the body. Alcohol is detoxified by peroxisomes in liver cells. A byproduct of these oxidation reactions is hydrogen peroxide, H_2O_2, which is contained within the peroxisomes to prevent the chemical from causing damage to cellular components outside of the organelle. Hydrogen peroxide is safely broken down by peroxisomal enzymes into water and oxygen.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Mitochondria and Chloroplasts

MITOCHONDRIA

Mitochondria (singular = mitochondrion) are often called the "powerhouses" or "energy factories" of a cell because they are responsible for making adenosine triphosphate (ATP), the cell's main energy-carrying molecule. The formation of ATP from the breakdown of glucose is known as cellular respiration. Mitochondria are oval-shaped, double-membrane organelles (**Figure 1**) that have their own ribosomes and DNA. Each membrane is a phospholipid bilayer embedded with proteins. The inner layer has folds called cristae, which increase the surface area of the inner membrane. The area surrounded by the folds is called the mitochondrial matrix. The cristae and the matrix have different roles in cellular respiration.

In keeping with our theme of form following function, it is important to point out that muscle cells have a very high concentration of mitochondria because muscle cells need a lot of energy to contract.

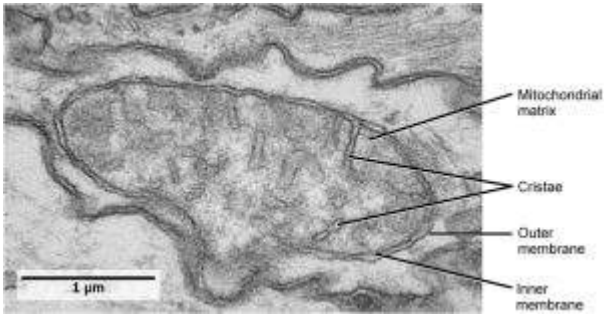


Figure 1 This transmission electron micrograph shows a mitochondrion as viewed with an electron microscope. Notice the inner and outer membranes, the cristae, and the mitochondrial matrix. (credit: modification of work by Matthew Britton; scale-bar data from Matt Russell)

Like mitochondria, chloroplasts also have their own DNA and ribosomes. **Chloroplasts** function in photosynthesis and can be found in eukaryotic cells such as plants and algae. Carbon dioxide (CO_2), water, and light energy are used to make glucose and oxygen in photosynthesis. This is the major difference between plants and animals: Plants (autotrophs) are able to make their own food, like glucose, whereas animals (heterotrophs) must rely on other organisms for their organic compounds or food source.

Like mitochondria, chloroplasts have outer and inner membranes, but within the space enclosed by a chloroplast's inner membrane is a set of interconnected and stacked, fluid-filled membrane sacs called thylakoids (**Figure 2**). Each stack of thylakoids is called a granum (plural = grana). The fluid enclosed by the inner membrane and surrounding the grana is called the stroma.

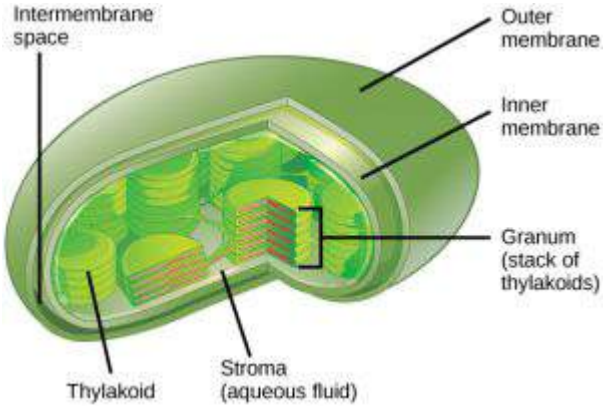


Figure 2 This simplified diagram of a chloroplast shows the outer membrane, inner membrane, thylakoids, grana, and stroma.

The chloroplasts contain a green pigment called **chlorophyll**, which captures the energy of sunlight for photosynthesis. Like plant cells, photosynthetic protists also have chloroplasts. Some bacteria also perform photosynthesis, but they do not have chloroplasts. Their photosynthetic pigments are located in the thylakoid membrane within the cell itself.

Theory of Endosymbiosis

We have mentioned that both mitochondria and chloroplasts contain DNA and ribosomes. Have you wondered why? Strong evidence points to endosymbiosis as the explanation.

Symbiosis is a relationship in which organisms from two separate species live in close association and typically exhibit specific

adaptations to each other. Endosymbiosis (endo-= within) is a relationship in which one organism lives inside the other. Endosymbiotic relationships abound in nature. Microbes that produce vitamin K live inside the human gut. This relationship is beneficial for us because we are unable to synthesize vitamin K. It is also beneficial for the microbes because they are protected from other organisms and are provided a stable habitat and abundant food by living within the large intestine.

Scientists have long noticed that bacteria, mitochondria, and chloroplasts are similar in size. We also know that mitochondria and chloroplasts have DNA and ribosomes, just as bacteria do. Scientists believe that host cells and bacteria formed a mutually beneficial endosymbiotic relationship when the host cells ingested aerobic bacteria and cyanobacteria but did not destroy them. Through evolution, these ingested bacteria became more specialized in their functions, with the aerobic bacteria becoming mitochondria and the photosynthetic bacteria becoming chloroplasts.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Cell Wall

The **cell wall** is a rigid covering that protects the cell, provides structural support, and gives shape to the cell. Cell walls are found in both prokaryotes and eukaryotes, although not all cells have cell walls. In Figure 1, the diagram of a plant cell, you see a structure external to the plasma membrane which is the cell wall. The cell wall is the reason why vegetables such as celery crunch when you bite into them.

Fungal and protist cells also have cell walls, but they are structurally different from those found in plants..

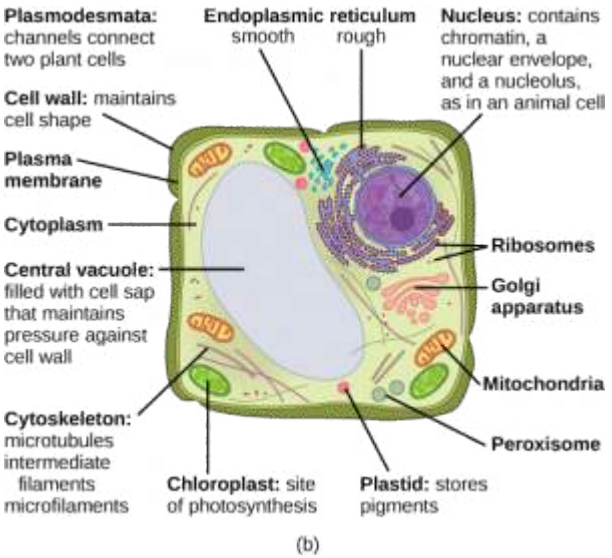


Figure 1 Note that the cell wall is located outside the cell membrane.

While the chief component of prokaryotic cell walls is peptidoglycan, the major organic molecule in the plant cell wall is cellulose, a polysaccharide made up of long, straight chains of glucose units. When nutritional information refers to dietary fiber, it is referring to the cellulose content of food. Fungal cell walls are made up of a molecule called chitin.

Animal cells do not have cell walls. Steak does not crunch when you bite it.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Extracellular matrix and intercellular junctions

EXTRACELLULAR MATRIX OF ANIMAL CELLS

Most animal cells release materials into the extracellular space. The primary components of these materials are glycoproteins and the protein collagen. Collectively, these materials are called the **extracellular matrix** (**Figure 1**). Not only does the extracellular matrix hold the cells together to form a tissue, but it also allows the cells within the tissue to communicate with each other.

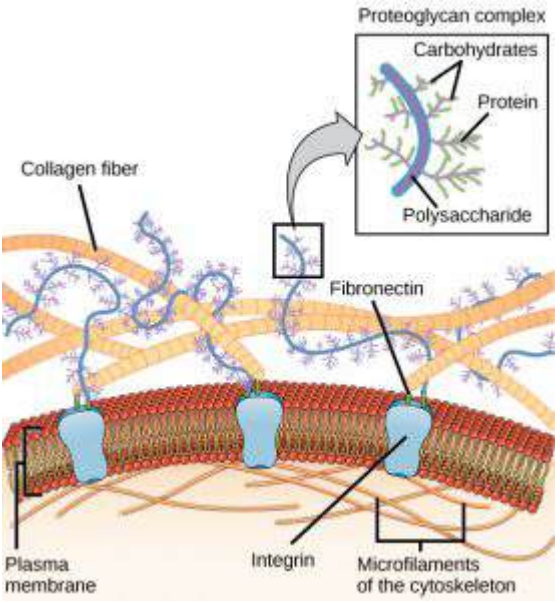


Figure 1 The extracellular matrix consists of a network of substances secreted by cells.

Blood clotting provides an example of the role of the extracellular matrix in cell communication.

When the cells lining a blood vessel are damaged, they display a protein receptor called tissue factor. When tissue factor binds with another factor in the extracellular matrix, it causes platelets to adhere to the wall of the damaged blood vessel, stimulates adjacent smooth muscle cells in the blood vessel to contract (thus constricting the blood vessel), and initiates a series of steps that stimulate the platelets to produce clotting factors.

INTERCELLULAR JUNCTIONS

Cells can also communicate with each other by direct contact, referred to as intercellular junctions. There are some

differences in the ways that plant and animal cells do this. **Plasmodesmata** (singular = plasmodesma) are junctions between plant cells, whereas animal cell contacts include tight and gap junctions, and desmosomes.

In general, long stretches of the plasma membranes of neighboring plant cells cannot touch one another because they are separated by the cell walls surrounding each cell. Plasmodesmata are numerous channels that pass between the cell walls of adjacent plant cells, connecting their cytoplasm and enabling signal molecules and nutrients to be transported from cell to cell (**Figure 2a**).

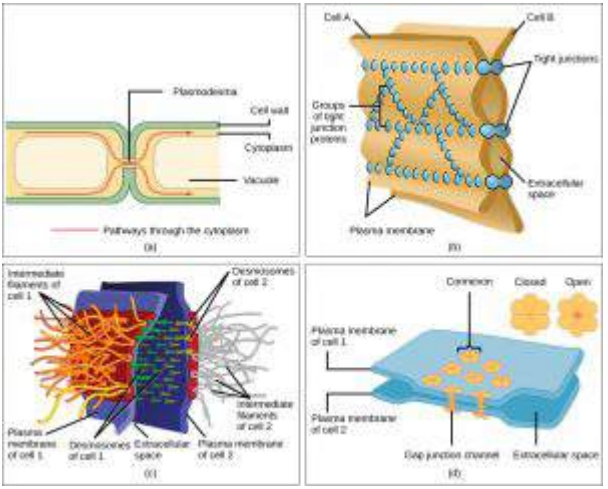


Figure 2 There are four kinds of connections between cells. (a) A plasmodesma is a channel between the cell walls of two adjacent plant cells. (b) Tight junctions join adjacent animal cells. (c) Desmosomes join two animal cells together. (d) Gap junctions act as channels between animal cells. (credit b, c, d: modification of work by Mariana Ruiz Villareal)

A **tight junction** is a watertight seal between two adjacent animal cells (**Figure 2b**). Proteins hold the cells tightly against each other. This tight adhesion prevents materials from leaking between the cells. Tight junctions are typically found

in the epithelial tissue that lines internal organs and cavities, and composes most of the skin. For example, the tight junctions of the epithelial cells lining the urinary bladder prevent urine from leaking into the extracellular space.

Also found only in animal cells are **desmosomes**, which act like spot welds between adjacent epithelial cells (**Figure 2c**). They keep cells together in a sheet-like formation in organs and tissues that stretch, like the skin, heart, and muscles.

Gap junctions in animal cells are like plasmodesmata in plant cells in that they are channels between adjacent cells that allow for the transport of ions, nutrients, and other substances that enable cells to communicate (**Figure 2d**). Structurally, however, gap junctions and plasmodesmata differ.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Production of a Protein

Proteins are one of the most abundant organic molecules in living systems and have an incredibly diverse range of functions. Proteins are used to:

- Build structures within the cell (such as the cytoskeleton)
- Regulate the production of other proteins by controlling protein synthesis
- Slide along the cytoskeleton to cause muscle contraction
- Transport molecules across the cell membrane
- Speed up chemical reactions (enzymes)
- Act as toxins

Each cell in a living system may contain thousands of different proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, polymers of amino acids, arranged in a linear sequence (**Figure 1**).

The functions of proteins are very diverse because they are made up of are 20 different chemically distinct amino acids that form long chains, and the amino acids can be in any order. The function of the protein is dependent on the protein's shape. The shape of a protein is determined by the

order of the amino acids. Proteins are often hundreds of amino acids long and they can have very complex shapes because there are so many different possible orders for the 20 amino acids!

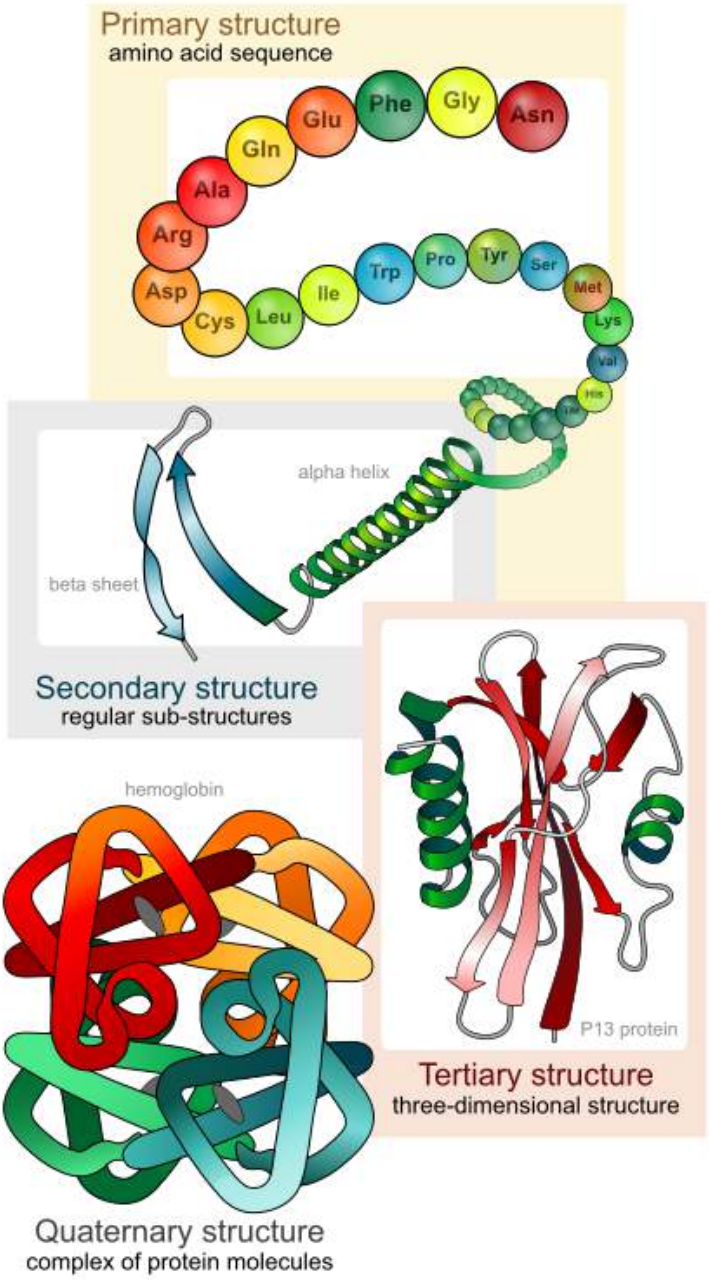

Primary structure
amino acid sequence
Phe
Gly
Asn
Glu
Gln
Ala
Arg
Asp
Cys
Leu
Ile
Trp
Pro
Tyr
Ser
Met
Lys
Val
His
alpha helix
beta sheet
Secondary structure
regular sub-structures
P13 protein
Tertiary structure
three-dimensional structure
hemoglobin
Quaternary structure
complex of protein molecules

Figure 1 Protein structure. The colored balls at the top of this diagram represent different amino acids. Amino acids are the subunits that are joined together by the ribosome to form a protein. This chain of amino acids then folds to form a complex 3D structure. (Credit: Lady of Hats from Wikipedia; public domain)

Contrary to what you may believe, proteins are not typically used as a source of energy by cells. Protein from your diet is broken down into individual amino acids which are reassembled by your ribosomes into proteins that your cells need. Ribosomes do not produce energy.

Figure 2 Examples of foods that contain high levels of protein. ("Protein" by National Cancer Institute is in the Public Domain)

The information to produce a protein is encoded in the cell's DNA. When a protein is produced, a copy of the DNA is made (called mRNA) and this copy is transported to a ribosome. Ribosomes read the information in the mRNA and use that information to assemble amino acids into a protein. If the protein is going to be used within the cytoplasm of the cell, the ribosome creating the protein will be free-floating in the cytoplasm. If the protein is going to be targeted to the lysosome, become a component of the plasma membrane, or be secreted outside of the cell, the protein will be

synthesized by a ribosome located on the rough endoplasmic reticulum (RER). After being synthesized, the protein will be carried in a vesicle from the RER to the *cis* face of the Golgi (the side facing the inside of the cell). As the protein moves through the Golgi, it can be modified. Once the final modified protein has been completed, it exits the Golgi in a vesicle that buds from the *trans* face. From there, the vesicle can be targeted to a lysosome or targeted to the plasma membrane. If the vesicle fuses with the plasma membrane, the protein will become part of the membrane or be ejected from the cell.

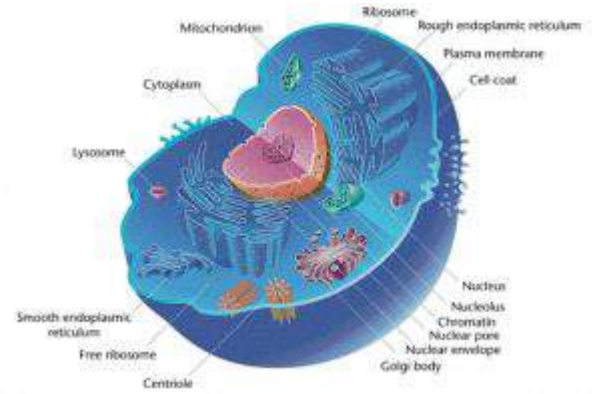


Figure 3 Diagram of a eukaryotic cell. (Photo credit: Mediran, Wikimedia. 14 Aug 2002)

Insulin

Insulin is a protein hormone that is made by specific cells inside the pancreas called beta cells. When the beta cells sense that glucose (sugar) levels in the bloodstream are high, they produce insulin protein and secrete it outside of the cells into the bloodstream. Insulin signals cells to absorb sugar from the bloodstream. Cells can't absorb sugar without insulin. Insulin

protein is first produced as an immature, inactive chain of amino acids (preproinsulin – See Figure 4). It contains a signal sequence that targets the immature protein to the rough endoplasmic reticulum, where it folds into the correct shape. The targeting sequence is then cut off of the amino acid chain to form proinsulin. This trimmed, folded protein is then shipped to the Golgi inside a vesicle. In the Golgi, more amino acids (chain C) are trimmed off of the protein to produce the final mature insulin. Mature insulin is stored inside special vesicles until a signal is received for it to be released into the bloodstream.

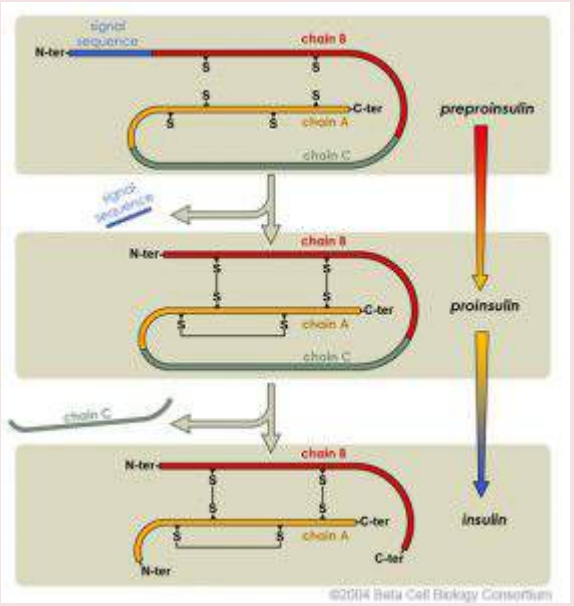


Figure 4 Insulin maturation. (Photo credit: Beta Cell Biology Consortium, Wikimedia. 2004. This picture is in the public domain.

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=126#h5p-31

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Summary Table of Prokaryotic and Eukaryotic Cells and Functions

Components of Prokaryotic and Eukaryotic Cells and Functions

Cell Component	Function	Present in Prokaryotes	Present in Animal Cells	Present in Plant Cells
Plasma Membrane	Separates cell from external environment; controls passage of organic molecules, ions, water, oxygen, and wastes into and out of the cell	Yes	Yes	Yes
Cytoplasm	Provides structure to cell; site of many metabolic reactions; medium in which organelles are found	Yes	Yes	Yes
Nucleoid	Location of DNA	Yes	No	No
Nucleus	Cell organelle that houses DNA and directs synthesis of ribosomes and proteins	No	Yes	Yes
Ribosomes	Protein synthesis	Yes	Yes	Yes
Mitochondria	ATP production/cellular respiration	No	Yes	Yes
Peroxisomes	Oxidizes and breaks down fatty acids and amino acids, and detoxifies poisons	No	Yes	Yes
Vesicles and vacuoles	Storage and transport; digestive function in plant cells	No	Yes	Yes
Centrosome	Unspecified role in cell division in animal cells; organizing center of microtubules in animal cells	No	Yes	No
Lysosomes	Digestion of macromolecules; recycling of worn-out organelles	No	Yes	No
Cell wall	Protection, structural support and maintenance of cell shape	Yes, primarily peptidoglycan in bacteria but not Archaea	No	Yes, primarily cellulose
Chloroplasts	Photosynthesis	No	No	Yes
Endoplasmic reticulum	Modifies proteins and synthesizes lipids	No	Yes	Yes
Golgi apparatus	Modifies, sorts, tags, packages, and distributes lipids and proteins	No	Yes	Yes

Cell Component	Function	Present in Prokaryotes	Present in Animal Cells	Present in Plant Cells
Cytoskeleton	Maintains cell's shape, secures organelles in specific positions, allows cytoplasm and vesicles to move within the cell, and enables unicellular organisms to move independently	Yes	Yes	Yes
Flagella	Cellular locomotion	Some	Some	No, except for some plant sperm.
Cilia	Cellular locomotion, movement of particles along extracellular surface of plasma membrane, and filtration	No	Some	No

Table 1 The components of prokaryotic and eukaryotic cells and their respective functions.

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=127#h5p-35

An interactive H5P element has been excluded from this version of the

text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=127#h5p-34

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=127#h5p-33

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=127#h5p-32

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

THE CELL MEMBRANE AND TRANSPORT

Learning Objectives

By the end of this section, you will be able to:

- Explain how the structure of cell membranes leads to its various functions including selective permeability and transport, and cell signaling.

The plasma membrane, which is also called the cell membrane, has many functions, but the most basic one is to define the borders of the cell and keep the cell functional. The plasma membrane is selectively permeable. This means that the membrane allows some materials to freely enter or leave the cell, while other materials cannot move freely, but require the use of a specialized structure, and occasionally, even energy investment for crossing.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 https://cnx.org/contents/GFy_h8cu@10.117:oaLwOnAf@2/Introduction

The Plasma Membrane

Cells closely control the exchange of substances in and out of the cell. Some substances are excluded, others are taken in, and still others are excreted – all in controlled quantities. Although the **plasma membrane** encloses the cell's borders, it is far from being a static barrier; it is dynamic and constantly in flux. The plasma membrane must be sufficiently flexible to allow certain cells, such as red blood cells and white blood cells, to change shape as they pass through narrow capillaries. In addition to these more obvious functions, the surface of the plasma membrane carries markers which allow cells to recognize one another. This is vital as these markers play a role in the "self" versus "non-self" distinction of the immune response.

FLUID MOSAIC MODEL

In 1972, S. J. Singer and Garth L. Nicolson proposed a new model of the plasma membrane. This theory, compared to earlier theories, best explains both microscopic observations and the function of the plasma membrane. This theory is called the **fluid mosaic model**. The model has evolved somewhat over time, but still best accounts for the structure and functions of the plasma membrane as we now understand them. The fluid mosaic model describes the

structure of the plasma membrane as comprised of diverse components—including phospholipids, cholesterol, proteins, and carbohydrates—that are able to flow and change position, while maintaining the basic integrity of the membrane. Both phospholipid molecules and embedded proteins are able to move laterally in the membrane. The fluidity of the plasma membrane is necessary for the activities of certain enzymes and transport molecules within the membrane.

Plasma membranes range from 5–10 nm thick. As a comparison, human red blood cells, visible via light microscopy, are approximately 8 μm thick, or approximately 1,000 times thicker than a plasma membrane.

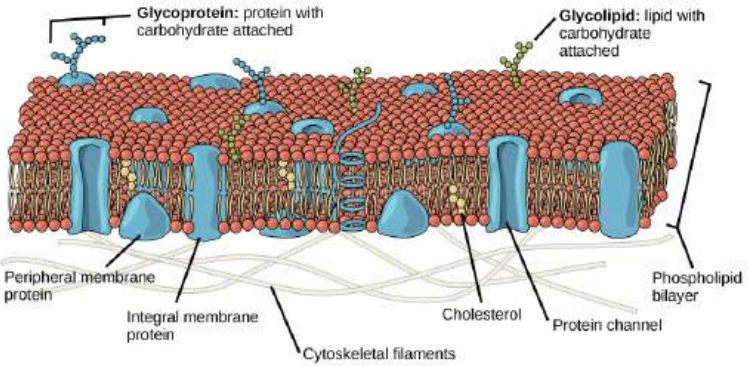


Figure 1 The fluid mosaic model of the plasma membrane structure describes the plasma membrane as a fluid combination of phospholipids, cholesterol, proteins, and carbohydrates.

COMPONENTS OF THE PLASMA MEMBRANE

The plasma membrane is made up primarily of a bilayer of phospholipids with embedded proteins, carbohydrates,

glycolipids, and glycoproteins, and, in animal cells, cholesterol (**Figure 1**).

PHOSPHOLIPIDS

The main fabric of the membrane is composed of two layers of phospholipid molecules, and the polar ends of these molecules (which look like a collection of balls in an artist's rendition of the model) (**Figure 2**) are in contact with aqueous fluid both inside and outside the cell. Thus, both surfaces of the plasma membrane are **hydrophilic** ("water loving"). In contrast, the interior of the membrane, between its two surfaces, is a **hydrophobic** ("water fearing") or nonpolar region because of the fatty acid tails. This region has no attraction for water or other polar molecules.

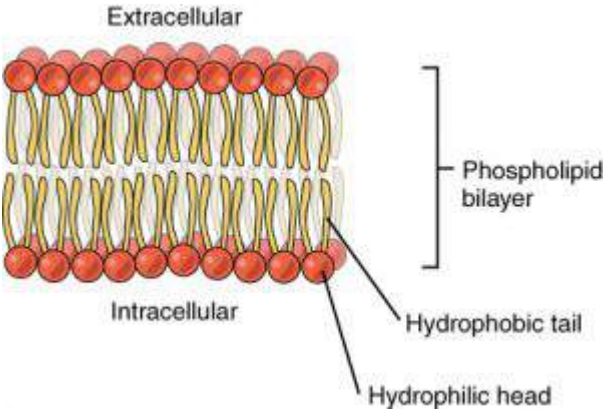


Figure 2 Phospholipid bilayer. "Extracellular" = outside the cell; "Intracellular" = inside the cell. Photo credit: OpenStax Anatomy and Physiology.

A phospholipid molecule (**Figure 3**) consists of a three-carbon glycerol backbone with two fatty acid molecules attached to carbons 1 and 2, and a phosphate-containing group attached to the third carbon. This arrangement gives the overall molecule an area described as its head (the

phosphate-containing group), which has a polar character or negative charge, and an area called the tail (the fatty acids), which has no charge. The head can form hydrogen bonds, but the tail cannot.

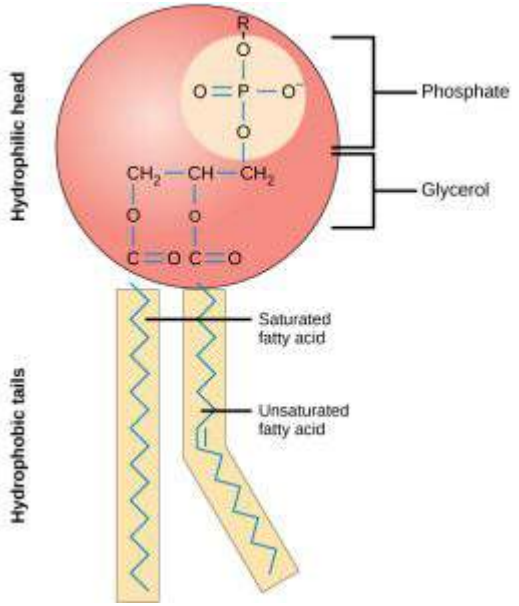


Figure 3 This phospholipid molecule is composed of a hydrophilic head and two hydrophobic tails. The hydrophilic head group consists of a phosphate-containing group attached to a glycerol molecule. The hydrophobic tails, each containing either a saturated or an unsaturated fatty acid, are long hydrocarbon chains.

PROTEINS

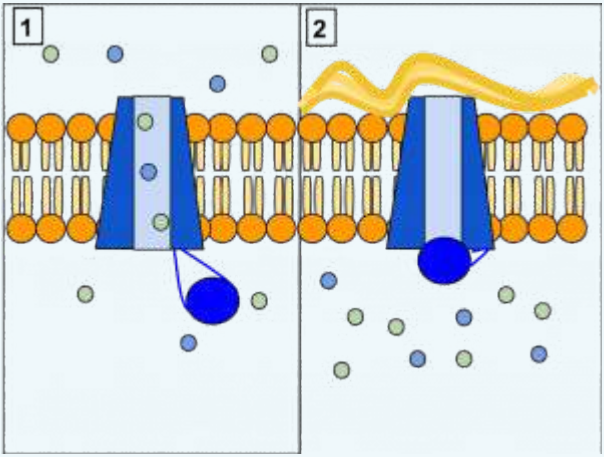
Proteins make up the second major chemical component of plasma membranes (see **Figure 1**). **Integral proteins** are embedded in the plasma membrane and may span all or part of the membrane (**Figure 1**). Integral proteins may serve as channels or pumps to move materials into or out of the cell. **Peripheral proteins** are found on the exterior or interior

surfaces of membranes, attached either to integral proteins or to phospholipid molecules (**Figure 1**). Both integral and peripheral proteins may serve as enzymes, as structural attachments for the fibers of the cytoskeleton, or as part of the cell's recognition sites.

The recognition sites on the plasma membrane are called **receptors,** which are attachment sites for substances that interact with the cell. Each receptor is structured to bind with a specific substance. The binding of a specific substance to its receptor on the plasma membrane can activate processes within the interior of the cell – such as activating enzymes involved in metabolic pathways. These metabolic pathways might be vital for providing the cell with energy, making substances for the cell, or breaking down cellular waste or toxins for disposal. Likewise, extracellular hormones and neurotransmitters bind to plasma membrane receptors which transmit a signal into the cell to intracellular molecules. Some recognition sites are used by viruses as attachment points. Although they are highly specific, pathogens like viruses may evolve to exploit receptors to gain entry to a cell by mimicking the specific substance that the receptor is meant to bind. This specificity helps to explain why human immunodeficiency virus (HIV) or any of the five types of hepatitis viruses invade only specific cells.

Cystic Fibrosis is caused by a defect in an integral protein in the cell membrane which acts as a channel. The CFTR protein moves ions from one side of the membrane to another. When it is not

functioning correctly, this causes very thick mucus to build up in the lungs and digestive tract.



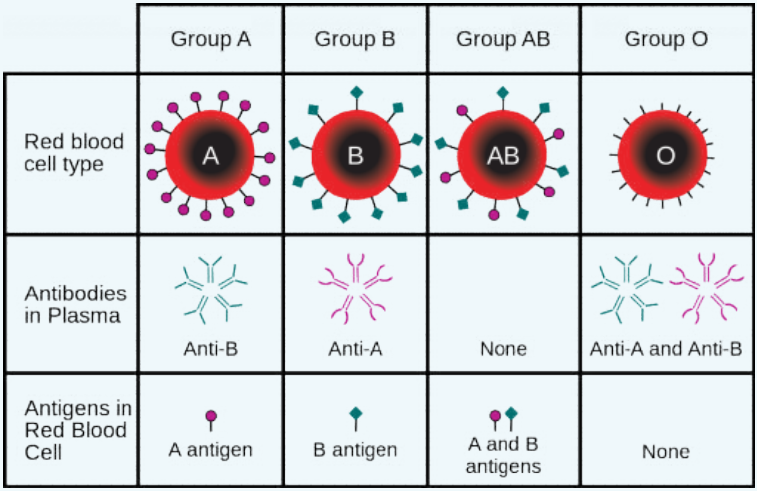
When the CFTR channel protein is functioning correctly (1), ions (small balls) are able to pass through the membrane. When it is not functioning correctly (2), ions are unable to cross the membrane. Photo credit: LBudd14, May, 2013. Wikimedia.

CARBOHYDRATES

Carbohydrates are the third major component of plasma membranes. They are always found on the exterior surface of cells and are bound either to proteins (forming **glycoproteins**) or to lipids (forming **glycolipids**). These carbohydrate chains may consist of 2–60 monosaccharide units and may be either straight or branched. Along with peripheral proteins, carbohydrates form specialized sites on

the cell surface that allow cells to recognize each other. These sites have unique patterns that allow the cell to be recognized, much the way that the facial features unique to each person allow him or her to be recognized. This recognition function is very important to cells, as it allows the immune system to differentiate between body cells (called "self") and foreign cells or tissues (called "non-self"). Similar types of glycoproteins and glycolipids are found on the surfaces of viruses and may change frequently, preventing immune cells from recognizing and attacking them.

The carbohydrates that make up glycoproteins are responsible for determining human A, B, O blood types. These glycoproteins are recognized by the immune system, which leads to incompatibility in blood types.



ABO Blood types. In this figure, the membrane carbohydrate is represented by the "lollipops". They are termed "antigens". Photo credit: InvictaHOG, 2006. Wikimedia.

MEMBRANE FLUIDITY

The mosaic characteristic of the membrane, described in the fluid mosaic model, helps to illustrate its nature. The proteins and other components that exist in the membrane can move with respect to each other, rather like boats floating on a lake. The membrane is not like a balloon, however, that can expand and contract; rather, it is fairly rigid and can burst if penetrated or if a cell takes in too much water. However, because of its mosaic nature, a very fine needle can easily

penetrate a plasma membrane without causing it to burst, and the membrane will flow and self-seal when the needle is extracted.

The mosaic characteristics of the membrane explain some but not all of its fluidity. There are two other factors that help maintain this fluid characteristic. One factor is the nature of the phospholipids themselves. The structure of the fatty acid tails in each phospholipid can make the membrane more dense and rigid, or less dense and flexible. The relative fluidity of the membrane is particularly important in a cold environment. A cold environment tends to make membranes less fluid and more susceptible to rupturing. Many organisms (fish are one example) are capable of adapting to cold environments by changing the proportion of different types of fatty acids in their membranes in response to the lowering of the temperature.

Animals have an additional membrane constituent that assists in maintaining fluidity. **Cholesterol,** which lies alongside the phospholipids in the membrane, tends to dampen the effects of temperature on the membrane. Thus, this lipid functions as a buffer, preventing lower temperatures from inhibiting fluidity and preventing increased temperatures from increasing fluidity too much. Thus, cholesterol extends, in both directions, the range of temperature in which the membrane is appropriately fluid and consequently functional. Cholesterol also serves other functions, such as organizing clusters of transmembrane proteins into lipid rafts.

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=130#h5p-36

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Transport Across Membranes

Plasma membranes act not only as a barrier, but also as a gatekeeper. It must allow needed substances to enter and cell products to leave the cell, while preventing entrance of harmful material and exit of essential material. In other words, plasma membranes are **selectively permeable**—they allow some substances through but not others. If the membrane were to lose this selectivity, the cell would no longer be able to maintain homeostasis, or to sustain itself, and it would be destroyed. Some cells require larger amounts of specific substances than other cells; they must have a way of obtaining these materials from the extracellular fluids.

This may happen passively, as certain materials move back and forth, or the cell may have special mechanisms that ensure transport. Most cells expend most of their energy, in the form of adenosine triphosphate (ATP), to create and maintain an uneven distribution of ions on the opposite sides of their membranes. The structure of the plasma membrane contributes to these functions.

SELECTIVE PERMEABILITY

Plasma membranes are asymmetric, meaning that despite the mirror image formed by the phospholipids, the interior

of the membrane is not identical to the exterior of the membrane. Integral proteins that act as channels or pumps work in one direction. Carbohydrates, attached to lipids or proteins, are also found on the exterior surface of the plasma membrane.

These carbohydrate complexes help the cell bind substances in the extracellular fluid that the cell needs. This adds considerably to the selective nature of plasma membranes.

Recall that plasma membranes have hydrophilic and hydrophobic regions. This characteristic helps the movement of certain materials through the membrane and hinders the movement of others. Lipid-soluble material can easily slip through the hydrophobic lipid core of the membrane. Substances such as the fat-soluble vitamins A, D, E, and K readily pass through the plasma membranes in the digestive tract and other tissues. Fat-soluble drugs also gain easy entry into cells and are readily transported into the body's tissues and organs. Molecules of oxygen and carbon dioxide have no charge and pass through by simple diffusion.

Polar substances, with the exception of water, present problems for the membrane. While some polar molecules connect easily with the outside of a cell, they cannot readily pass through the lipid core of the plasma membrane. Additionally, whereas small ions could easily slip through the spaces in the mosaic of the membrane, their charge prevents them from doing so. Ions such as sodium, potassium, calcium, and chloride must have a special means of penetrating plasma membranes. Simple sugars and amino acids also need help with transport across plasma membranes.

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=131#h5p-37

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Passive Transport: Diffusion

The most direct forms of membrane transport are passive. **Passive transport** is a naturally occurring phenomenon and does not require the cell to expend energy to accomplish the movement. In passive transport, substances move from an area of higher concentration to an area of lower concentration in a process called **diffusion**. A physical space in which there is a different concentration of a single substance is said to have a **concentration gradient**.

DIFFUSION

Diffusion is a passive process of transport. A single substance tends to move from an area of high concentration to an area of low concentration until the concentration is equal across the space. You are familiar with diffusion of substances through the air. For example, think about someone opening a bottle of perfume in a room filled with people. The perfume is at its highest concentration in the bottle and is at its lowest at the edges of the room. The perfume vapor will diffuse, or spread away, from the bottle, and gradually, more and more people will smell the perfume as it spreads. Materials move within the cell's cytosol by diffusion, and certain materials move through the plasma membrane by diffusion (**Figure 1**). Diffusion expends no

energy. Rather the different concentrations of materials in different areas are a form of potential energy, and diffusion is the dissipation of that potential energy as materials move down their concentration gradients, from high to low.

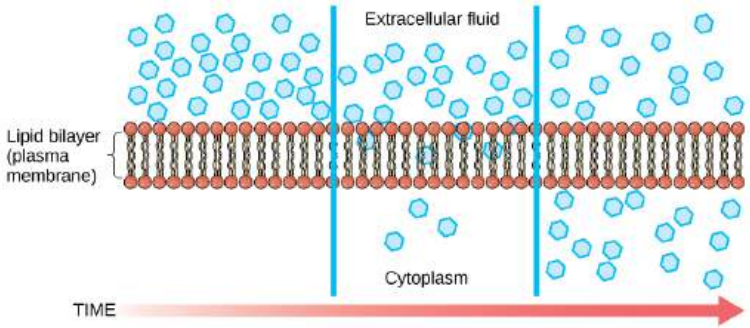


Figure 1 Diffusion through a permeable membrane follows the concentration gradient of a substance, moving the substance from an area of high concentration to one of low concentration. (credit: modification of work by Mariana Ruiz Villarreal)

Each separate substance in a medium, such as the extracellular fluid, has its own concentration gradient, independent of the concentration gradients of other materials. Additionally, each substance will diffuse according to that gradient.

Several factors affect the rate of diffusion:

- Extent of the concentration gradient: The greater the difference in concentration, the more rapid the diffusion. The closer the distribution of the material gets to equilibrium, the slower the rate of diffusion becomes.
- Mass of the molecules diffusing: More massive molecules move more slowly, because it is more difficult for them to move between the molecules of the substance they are moving through; therefore, they diffuse more slowly.

- Temperature: Higher temperatures increase the energy and therefore the movement of the molecules, increasing the rate of diffusion.
- Solvent density: As the density of the solvent increases, the rate of diffusion decreases. The molecules slow down because they have a more difficult time getting through the denser medium.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=133#h5p-38

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Passive Transport: Facilitated Transport

In **facilitated transport**, also called facilitated diffusion, material moves across the plasma membrane with the assistance of transmembrane proteins down a concentration gradient (from high to low concentration) *without* the expenditure of cellular energy. However, the substances that undergo facilitated transport would otherwise not diffuse easily or quickly across the plasma membrane. The solution to moving polar substances and other substances across the plasma membrane rests in the proteins that span its surface. The material being transported is first attached to protein or glycoprotein receptors on the exterior surface of the plasma membrane. This allows the material that is needed by the cell to be removed from the extracellular fluid. The substances are then passed to specific integral proteins that facilitate their passage, because they form channels or pores that allow certain substances to pass through the membrane. The integral proteins involved in facilitated transport are collectively referred to as transport proteins, and they function as either channels for the material or carriers.

CHANNELS

The integral proteins involved in facilitated transport are collectively referred to as transport proteins, and they function as either channels for the material or carriers. In both cases, they are transmembrane proteins (they span across the membrane). Channels are specific for the substance that is being transported. Channel proteins have hydrophilic domains exposed to the intracellular and extracellular fluids; they additionally have a hydrophilic channel through their core that provides a hydrated opening through the membrane layers (**Figure 1**). Passage through the channel allows polar compounds to avoid the nonpolar central layer of the plasma membrane that would otherwise slow or prevent their entry into the cell. Aquaporins are channel proteins that allow water to pass through the membrane at a very high rate.

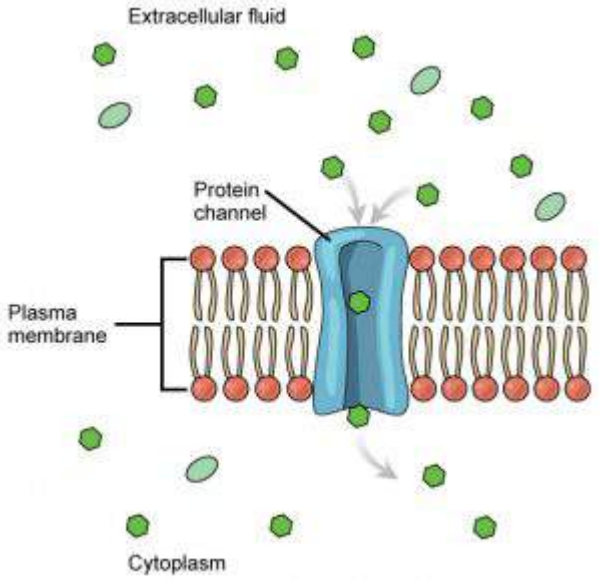
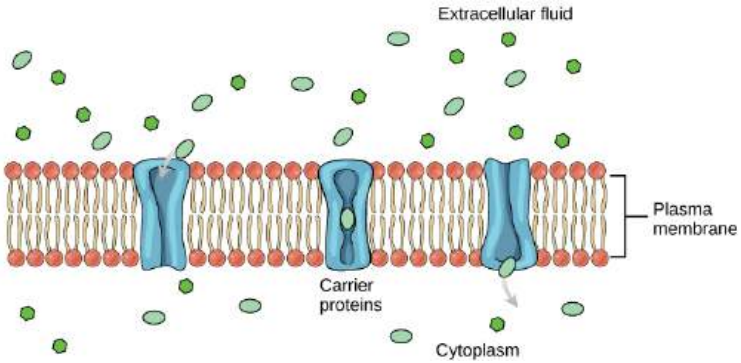


Figure 1 Facilitated transport moves substances down their concentration gradients. They may cross the plasma membrane with the aid of channel proteins. (credit: modification of work by Mariana Ruiz Villareal)

CARRIER PROTEINS

Another type of protein embedded in the plasma membrane is a carrier protein. This aptly named protein binds a substance and, in doing so, triggers a change of its own shape, moving the bound molecule from the outside of the cell to its interior (**Figure 2**); depending on the gradient, the material may move in the opposite direction. Carrier proteins are typically specific for a single substance. This selectivity adds to the overall selectivity of the plasma membrane. The exact mechanism for the change of shape is poorly understood. Proteins can change shape when their hydrogen

bonds are affected, but this may not fully explain this mechanism. Each carrier protein is specific to one substance, and there are a finite number of these proteins in any membrane. This can cause problems in transporting enough of the material for the cell to function properly. When all of the proteins are bound to their ligands, they are saturated and the rate of transport is at its maximum. Increasing the concentration gradient at this point will not result in an increased rate of transport.



An example of this process occurs in the kidney. Glucose, water, salts, ions, and amino acids needed by the body are filtered in one part of the kidney. This filtrate, which includes glucose, is then reabsorbed in another part of the kidney. Because there are only a finite number of carrier proteins for glucose, if more glucose is present than the proteins can handle, the excess is not transported and it is excreted from the body in the urine. In a diabetic individual, this is described as "spilling glucose into the urine." A different group of carrier proteins called glucose transport proteins, or GLUTs, are involved in transporting glucose and other hexose sugars through plasma membranes within the body.

Channel and carrier proteins transport material at different rates. Channel proteins transport much more quickly than do carrier proteins. Channel proteins facilitate diffusion at a rate of tens of millions of molecules per second,

whereas carrier proteins work at a rate of a thousand to a million molecules per second.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Passive Transport: Osmosis

Osmosis is the diffusion of water through a semipermeable membrane according to the concentration gradient of water across the membrane. Whereas diffusion transports material across membranes and within cells, osmosis transports *only water* across a membrane and the membrane limits the diffusion of solutes in the water. Osmosis is a special case of diffusion. Water, like other substances, moves from an area of higher concentration to one of lower concentration. Imagine a beaker with a semipermeable membrane, separating the two sides or halves **(Figure 3**). On both sides of the membrane, the water level is the same, but there are different concentrations on each side of a dissolved substance, or **solute**, that cannot cross the membrane. If the volume of the water is the same, but the concentrations of solute are different, then there are also different concentrations of water, the **solvent,** on either side of the membrane.

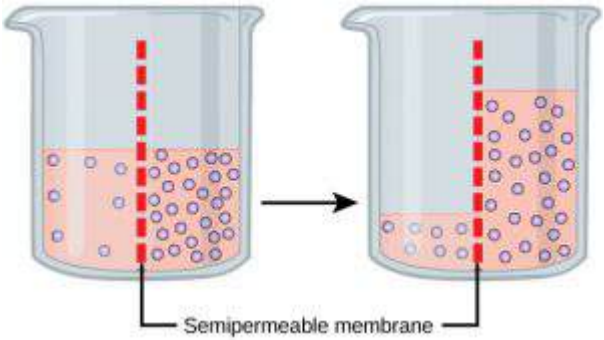


Figure 3 In osmosis, water always moves from an area of higher concentration (of water) to one of lower concentration (of water). In this system, the solute cannot pass through the selectively permeable membrane.

A principle of diffusion is that the molecules move around and will spread evenly throughout the medium if they can. However, only the material capable of getting through the membrane will diffuse through it. In this example, the solute cannot diffuse through the membrane, but the water can. Water has a concentration gradient in this system. Therefore, water will diffuse down its concentration gradient, crossing the membrane to the side where it is less concentrated. This diffusion of water through the membrane— **osmosis** —will continue until the concentration gradient of water goes to zero. Osmosis proceeds constantly in living systems.

TONICITY

Tonicity describes the amount of solute in a solution. The measure of the tonicity of a solution, or the total amount of solutes dissolved in a specific amount of solution, is called its **osmolarity**. Three terms—hypotonic, isotonic, and hypertonic—are used to relate the osmolarity of a cell to the osmolarity of the extracellular fluid that contains the cells. All

three of these terms are a *comparison* between two different solutions (for example, inside a cell compared to outside the cell).

In a **hypotonic** solution, such as tap water, the extracellular fluid has a lower concentration of solutes than the fluid inside the cell, and water enters the cell. (In living systems, the point of reference is always the cytoplasm, so the prefix *hypo–* means that the extracellular fluid has a lower concentration of solutes, or a lower osmolarity, than the cell cytoplasm.) It also means that the extracellular fluid has a higher concentration of water than does the cell. In this situation, water will follow its concentration gradient and enter the cell. This may cause an animal cell to burst, or **lyse.**

In a **hypertonic** solution (the prefix *hyper–* refers to the extracellular fluid having a higher concentration of solutes than the cell's cytoplasm), the fluid contains less water than the cell does, such as seawater. Because the cell has a lower concentration of solutes, the water will leave the cell. In effect, the solute is drawing the water out of the cell. This may cause an animal cell to shrivel, or **crenate.**

In an **isotonic** solution, the extracellular fluid has the same osmolarity as the cell. If the concentration of solutes of the cell matches that of the extracellular fluid, there will be no net movement of water into or out of the cell. The cell will retain its "normal" appearance. Blood cells in hypertonic, isotonic, and hypotonic solutions take on characteristic appearances (**Figure 4**).

Remember that all three of these terms are *comparisons* between two solutions (i.e. inside and outside the cell). A solution can't be hypotonic, that would be like saying that Bob is taller. That doesn't make sense – you need to say that Bob is taller than Mike. You can say that the solution inside the cell is hypotonic to the solution outside the cell. That also means that the solution outside is hypertonic to the solution inside (just like Mike would be shorter than Bob).

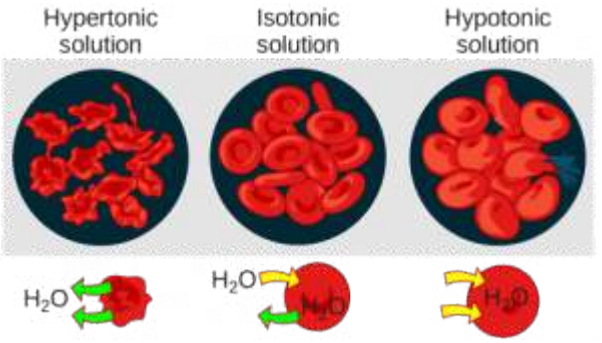


Figure 4 Osmotic pressure changes the shape of red blood cells in hypertonic, isotonic, and hypotonic solutions. (credit: modification of work by Mariana Ruiz Villarreal)

Some organisms, such as plants, fungi, bacteria, and some protists, have **cell walls** that surround the plasma membrane and prevent cell lysis. The plasma membrane can only expand to the limit of the cell wall, so the cell will not lyse. In fact, the cytoplasm in plants is always slightly hypertonic compared to the cellular environment, and water will always enter the plant cell if water is available. This influx of water produces **turgor pressure**, which stiffens the cell walls of the plant (**Figure 5**). In nonwoody plants, turgor pressure supports the plant. If the plant cells become hypertonic, as occurs in drought or if a plant is not watered adequately, water will leave the cell. Plants lose turgor pressure in this condition and wilt.

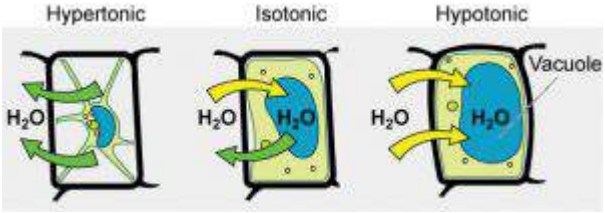


Figure 5 The turgor pressure within a plant cell depends on the tonicity of the solution that it is bathed in. (credit: modification of work by Mariana Ruiz Villarreal)

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=138#h5p-39

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=138#h5p-40

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=138#h5p-41

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Active Transport

Active transport mechanisms require the use of the cell's energy, usually in the form of adenosine triphosphate (ATP). If a substance must move into the cell against its concentration gradient, that is, if the concentration of the substance inside the cell must be greater than its concentration in the extracellular fluid, the cell must use energy to move the substance. Some active transport mechanisms move small-molecular weight material, such as ions, through the membrane.

In addition to moving small ions and molecules through the membrane, cells also need to remove and take in larger molecules and particles. Some cells are even capable of engulfing entire unicellular microorganisms. You might have correctly hypothesized that the uptake and release of large particles by the cell requires energy. A large particle, however, cannot pass through the membrane, even with energy supplied by the cell.

ELECTROCHEMICAL GRADIENT

We have discussed simple concentration gradients—differential concentrations of a substance across a space or a membrane. However, in living systems gradients are more complex. Cells contain many proteins, most of

which are negatively charged. Due to these negatively charged proteins, coupled with the movement of ions into and out of cells, there is an electrical gradient (a difference of charge) across the plasma membrane. The interior of living cells is electrically negative as compared to the extracellular fluid in which cells are bathed; at the same time, cells contain higher concentrations of potassium (K+) and lower concentrations of sodium (Na+) than does the extracellular fluid. Thus, in a living cell, the concentration gradient and electrical gradient of Na+ promotes diffusion of the ion into the cell, and the electrical gradient of Na+ (a positive ion) tends to drive it inward to the negatively charged interior. The situation is more complex, however, for other elements such as potassium. The electrical gradient of K+ promotes diffusion of the ion *into* the cell, but the concentration gradient of K+ promotes diffusion *out* of the cell (**Figure 5**). The combined gradient that affects an ion is called its **electrochemical gradient**, and it is especially important to muscle and nerve cells.

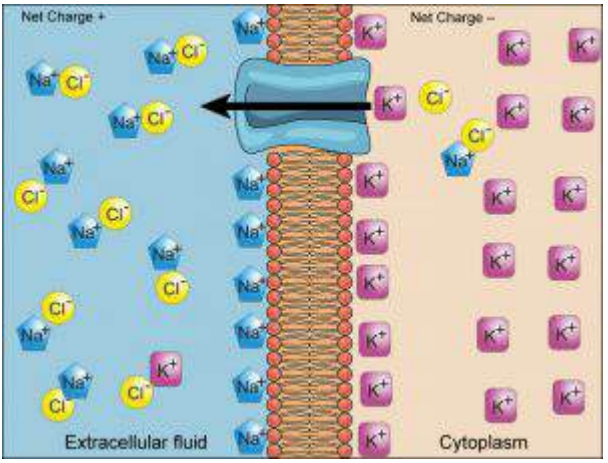


Figure 5 Electrochemical gradients arise from the combined effects of concentration gradients and electrical gradients. (credit: modification of work by "Synaptitude"/Wikimedia Commons)

MOVING AGAINST A GRADIENT

To move substances against a concentration or an electrochemical gradient, the cell must use energy. This energy is harvested from ATP that is generated through cellular metabolism. Active transport mechanisms, collectively called pumps or carrier proteins, work against electrochemical gradients. With the exception of ions, small substances constantly pass through plasma membranes. Active transport maintains concentrations of ions and other substances needed by living cells in the face of these passive changes. Much of a cell's supply of metabolic energy may be spent maintaining these processes. As active transport mechanisms depend on cellular metabolism for energy, they are sensitive to many metabolic poisons that interfere with the supply of ATP.

Two mechanisms exist for the transport of small-molecular weight material and macromolecules. **Primary active transport** moves ions across a membrane and creates a difference in charge across that membrane. The primary active transport system uses ATP to move a substance, such as an ion, into the cell, and often at the same time, a second substance is moved out of the cell. The sodium-potassium pump, an important pump in animal cells, expends energy to move potassium ions into the cell and a different number of sodium ions out of the cell (**Figure 6**). The action of this pump results in a concentration and charge difference across the membrane.

Secondary active transport describes the movement of material using the energy of the electrochemical gradient established by primary active transport. Using the energy of the electrochemical gradient created by the primary active transport system, other substances such as amino acids and glucose can be brought into the cell through membrane channels. ATP itself is formed through secondary active

transport using a hydrogen ion gradient in the mitochondrion.

ENDOCYTOSIS

Endocytosis is a type of active transport that moves particles, such as large molecules, parts of cells, and even whole cells, into a cell. There are different variations of endocytosis, but all share a common characteristic: The plasma membrane of the cell invaginates, forming a pocket around the target particle. The pocket pinches off, resulting in the particle being contained in a newly created vacuole that is formed from the plasma membrane.

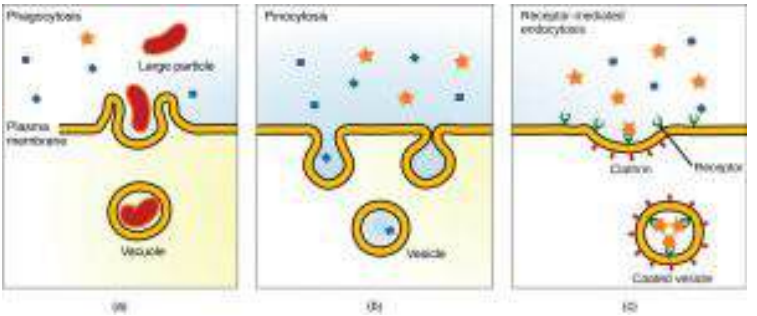


Figure 7 Three variations of endocytosis are shown. (a) In one form of endocytosis, phagocytosis, the cell membrane surrounds the particle and pinches off to form an intracellular vacuole. (b) In another type of endocytosis, pinocytosis, the cell membrane surrounds a small volume of fluid and pinches off, forming a vesicle. (c) In receptor-mediated endocytosis, uptake of substances by the cell is targeted to a single type of substance that binds at the receptor on the external cell membrane. (credit: modification of work by Mariana Ruiz Villarreal)

Phagocytosis is the process by which large particles, such as cells, are taken in by a cell. For example, when microorganisms invade the human body, a type of white blood cell called a neutrophil removes the invader through

this process, surrounding and engulfing the microorganism, which is then destroyed by the neutrophil (**Figure 7**a).

A variation of endocytosis is called **pinocytosis**. This literally means "cell drinking" and was named at a time when the assumption was that the cell was purposefully taking in extracellular fluid. In reality, this process takes in solutes that the cell needs from the extracellular fluid (**Figure 7**b).

A targeted variation of endocytosis employs binding proteins in the plasma membrane that are specific for certain substances (**Figure 7**c). The particles bind to the proteins and the plasma membrane invaginates, bringing the substance and the proteins into the cell. If passage across the membrane of the target of **receptor-mediated endocytosis** is ineffective, it will not be removed from the tissue fluids or blood. Instead, it will stay in those fluids and increase in concentration.

Some human diseases are caused by a failure of receptor-mediated endocytosis. For example, the form of cholesterol termed low-density lipoprotein or LDL (also referred to as "bad" cholesterol) is removed from the blood by receptor mediated endocytosis. In the human genetic disease familial hypercholesterolemia, the LDL receptors are defective or missing entirely. People with this condition have life-threatening levels of cholesterol in their blood, because their cells cannot clear the chemical from their blood.

EXOCYTOSIS

In contrast to these methods of moving material into a cell is the process of exocytosis. **Exocytosis** is the opposite of the processes discussed above in that its purpose is to expel material from the cell into the extracellular fluid. A particle enveloped in membrane fuses with the interior of the plasma membrane. This fusion opens the membranous envelope to

the exterior of the cell, and the particle is expelled into the extracellular space (**Figure 8**).

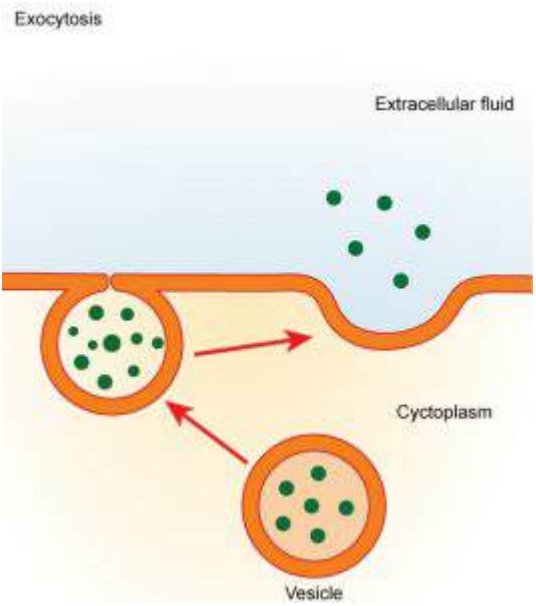


Figure 8 In exocytosis, a vesicle migrates to the plasma membrane, binds, and releases its contents to the outside of the cell. (credit: modification of work by Mariana Ruiz Villarreal)

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=143#h5p-43

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=143#h5p-42

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

CELL COMMUNICATION

Learning Objectives

Course Outcomes for this section:

Explain how basic units of cellular structure define the function of all living things.

d. Explain how the structure of cell membranes leads to its various functions including selective permeability and transport, and cell signaling.

Imagine what life would be like if you and the people around you could not communicate. You would not be able to express your wishes to others, nor could you ask questions to find out more about your environment. Social organization is dependent on communication between the individuals that comprise that society; without communication, society would fall apart.

As with people, it is vital for individual cells to be able to interact with their environment. This is true whether a cell is growing by itself in a pond or is one of many cells

that form a larger organism. In order to properly respond to external stimuli, cells have developed complex mechanisms of communication that can receive a message, transfer the information across the plasma membrane, and then produce changes within the cell in response to the message.

In multicellular organisms, cells send and receive chemical messages constantly to coordinate the actions of distant organs, tissues, and cells. The ability to send messages quickly and efficiently enables cells to coordinate and fine-tune their functions.

While the necessity for cellular communication in larger organisms seems obvious, even single-celled organisms communicate with each other. Yeast cells signal each other to aid mating. Some forms of bacteria coordinate their actions in order to form large complexes called biofilms or to organize the production of toxins to remove competing organisms. The ability of cells to communicate through chemical signals originated in single cells and was essential for the evolution of multicellular organisms. The efficient and error-free function of communication systems is vital for all life as we know it.

REFERENCES

OpenStax, Biology. OpenStax CNX. October 13, 2017. https://cnx.org/contents/GFy_h8cu@10.118:1e9l33C7@2/Introduction

A Summary of Cell Communication

Receptors are protein molecules inside the target cell or on its surface that receive a chemical signal. Chemical signals are released by signaling cells in the form of small, usually volatile or soluble molecules called **ligands.** A ligand is a molecule that binds another specific molecule, in some cases, delivering a signal in the process. Ligands can thus be thought of as signaling molecules. Ligands and receptors exist in several varieties; however, a specific ligand will have a specific receptor that typically binds only that ligand.

There are two basic types of **receptors:** internal receptors and cell surface receptors.

- **Internal receptors** are found in the cytoplasm of the cell and respond to ligands that cross the cell membrane into the cell. These receptors can have a direct effect on protein production by binding directly to the DNA.
- **Cell-surface receptors** are found on the cell membrane. They bind to ligands that do not cross the cell membrane. After the ligand binds, the receptor responds in some way. One response is to open a channel to allow ions to pass through the membrane. A second response is to activate an

enzyme that sets off a response inside the cell. A third response is to activate a protein which is not an enzyme, but which can affect other cell components.

There are several different types of **ligands.**

- Small hydrophobic ligands can pass directly through the cell membrane. They typically interact with internal receptors. Steroid hormones are an example.
- Water soluble hydrophilic ligands can not pass directly through the cell membrane. They typically interact with cell-surface receptors. Peptide (protein) hormones are an example.
- There are a variety of other ligands such as nitric oxide (NO) gas. Nitroglycerin and Viagra affect the NO pathway.

Once a ligand binds to a receptor, the signal is transmitted through the membrane and into the cytoplasm. Continuation of a signal in this manner is called **signal transduction**. Signal transduction only occurs with cell-surface receptors because internal receptors are able to interact directly with DNA in the nucleus to initiate protein synthesis.

Signal transduction pathways can be extremely complicated and involve large numbers of enzymes and other proteins. These pathways can help amplify a signal received by one receptor. There can also be different effects from the same ligand in different cell types due to different proteins present in different types of cells.

- **Kinases** are a type of enzyme that adds a phosphate group to another molecule (including other proteins). This is called **phosphorylation.** Phosphorylation can activate or

deactivate other proteins.

- **Second messengers** are small molecules that help to spread a signal through the cytoplasm after a ligand binds to a receptor. They do this by altering the behavior of certain cellular proteins. Some examples of second messengers are cAMP (a modified version of AMP, which is related to ATP but only contains one phosphate) and calcium ions.

There are several categories of cellular responses to signals.

- **Changes in gene expression**: an increase or decrease in the production of a protein produced by a specific gene.
- **An increase in cellular metabolism**: the conversion of glucose to glycogen (and back) can be regulated depending on the energy needs of the cell.
- **Cell growth**: cells do not normally divide unless they are stimulated by signals from other cells.
- **Cell death**: apoptosis is controlled cell death; cells can be stimulated die if they are abnormal, infected with a bacteria or virus, or during specific parts of development (for example, to separate the fingers).

Stopping cell signaling pathways at the right time is just as important as starting them correctly. Tumors often display abnormal responses to cell signaling pathways.

Types of Receptors

A cell within a multicellular organism may need to signal to other cells that are at various distances from the original cell (**Figure 1**). Not all cells are affected by the same signals. Different types of signaling are used for different purposes.

The illustration shows four forms of chemical signaling. In autocrine signaling, a cell targets itself. In signaling across a gap junction, a cell targets a cell connected via gap junctions. In paracrine signaling, a cell targets a nearby cell. In endocrine signaling, a cell targets a distant cell via the bloodstream

Figure 1 In chemical signaling, a cell may target itself (autocrine signaling), a cell connected by gap junctions, a nearby cell (paracrine signaling), or a distant cell (endocrine signaling). Paracrine signaling acts on nearby cells, endocrine signaling uses the circulatory system to transport ligands, and autocrine signaling acts on the signaling cell. Signaling via gap junctions involves signaling molecules moving directly between adjacent cells.

Receptors are protein molecules inside the target cell or on its surface that receive a chemical signal. Chemical signals are released by signaling cells in the form of small, usually volatile or soluble molecules called **ligands.** A ligand is a molecule that binds another specific molecule, in some cases, delivering a signal in the process. Ligands can thus be thought of as signaling molecules. Ligands and receptors

exist in several varieties; however, a specific ligand will have a specific receptor that typically binds only that ligand.

INTERNAL RECEPTORS

Internal receptors, also known as intracellular or cytoplasmic receptors, are found in the cytoplasm of the cell and respond to hydrophobic ligand molecules that are able to travel across the plasma membrane. Once inside the cell, many of these molecules bind to proteins that act as regulators of mRNA synthesis. Recall that mRNA carries genetic information from the DNA in a cell's nucleus out to the ribosome, where the protein is assembled. When the ligand binds to the internal receptor, a change in shape is triggered that exposes a DNA-binding site on the receptor protein. The ligand-receptor complex moves into the nucleus, then binds to specific regions of the DNA and promotes the production of mRNA from specific genes (**Figure 2**). Internal receptors can directly influence gene expression (how much of a specific protein is produced from a gene) without having to pass the signal on to other receptors or messengers.

This illustration shows a hydrophobic signaling molecule that diffuses across the plasma membrane and binds an intracellular receptor in the cytoplasm. The intracellular receptor-signaling molecule complex then travels to the nucleus and binds DNA.

Figure 2 Hydrophobic signaling molecules typically diffuse across the plasma membrane and interact with intracellular receptors in the cytoplasm. Many intracellular receptors are transcription factors that interact with DNA in the nucleus and regulate gene expression.

CELL-SURFACE RECEPTORS

Cell-surface receptors, also known as **transmembrane receptors**, are proteins that are found attached to the cell membrane. These receptors bind to external ligand molecules (ligands that do not travel across the cell membrane). This type of receptor spans the plasma membrane and performs **signal transduction**, in which an extracellular signal is converted into an intercellular signal. Ligands that interact with cell-surface receptors do not have to enter the cell that they affect. Cell-surface receptors are also called cell-specific proteins or markers because they are specific to individual cell types.

Each cell-surface receptor has three main components: an external ligand-binding domain, a hydrophobic membrane-spanning region, and an intracellular domain inside the cell. The size and extent of each of these domains vary widely, depending on the type of receptor.

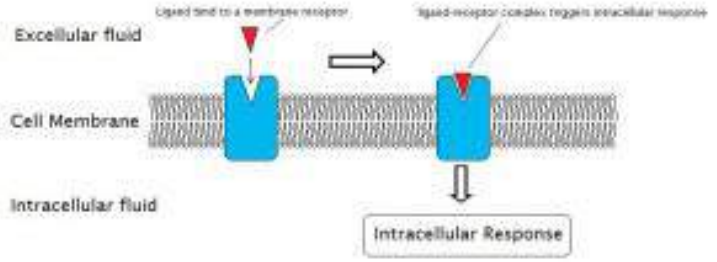


Figure 3 Cell-surface receptors function by transmitting a signal through the cell membrane. The ligand does not directly enter the cell. Photo credit Laozhengzz; Wikimedia commons.

Cell-surface receptors are involved in most of the signaling in multicellular organisms. There are three general categories of cell-surface receptors: ion channel-linked receptors, G-protein-linked receptors, and enzyme-linked receptors.

Ion channel-linked receptors

Ion channel-linked receptors bind a ligand and open a channel through the membrane that allows specific ions to pass through. To form a channel, this type of cell-surface receptor has an extensive membrane-spanning region. When a ligand binds to the extracellular region of the channel, there is a conformational change in the proteins structure that allows ions such as sodium, calcium, magnesium, and hydrogen to pass through (**Figure 4)**.

This illustration shows a gated ion channel that is closed in the absence of a signaling molecule. When a signaling molecule binds, a pore in the middle of the channel opens, allowing ions to enter the cell.

Figure 4 Gated ion channels form a pore through the plasma membrane that opens when the signaling molecule binds. The open pore then allows ions to flow into or out of the cell.

G-protein-coupled receptors

G-protein-coupled receptors bind a ligand and activate a membrane protein called a G-protein. The activated G-protein then interacts with either an ion channel or an enzyme in the membrane (**Figure 5**). Before the ligand binds, the inactive G-protein can bind to a site on a specific receptor. Once the G-protein binds to the receptor, the G-protein changes shape, becomes active, and splits into two different subunits. One or both of these subunits may be able to activate other proteins as a result.

This illustration shows the activation pathway for a heterotrimeric G-protein, which has three subunits: alpha beta, and gamma, all associated with the inside of the plasma membrane. When a signaling molecule binds to a G-protein-coupled receptor in the plasma membrane, a GDP molecule associated with the alpha subunit is exchanged for GTP. The alpha subunit dissociates from the beta and gamma subunits and triggers a cellular response. Hydrolysis of GTP to GDP terminates the signal.

Figure 5 When a signaling molecule binds to a G-protein-coupled receptor in the plasma membrane, a GDP molecule associated with the G-protein is exchanged for GTP. The subunits come apart from each other, and a cellular response is triggered either by one or both of the subunits. Hydrolysis of GTP to GDP terminates the signal.

Enzyme-linked receptors

Enzyme-linked receptors are cell-surface receptors with intracellular domains that are associated with an enzyme. In some cases, the intracellular domain of the receptor itself is an enzyme. Other enzyme-linked receptors have a small intracellular domain that interacts directly with an enzyme. When a ligand binds to the extracellular domain, a signal is transferred through the membrane, activating the enzyme. Activation of the enzyme sets off a chain of events within the cell that eventually leads to a response.

How Viruses Recognize a Host

Unlike living cells, many viruses do not have a plasma membrane or any of the structures necessary to sustain life. Some viruses are simply composed of an inert protein shell containing DNA or RNA. To reproduce, viruses must invade a

living cell, which serves as a host, and then take over the hosts cellular apparatus. But how does a virus recognize its host?

Viruses often bind to cell-surface receptors on the host cell. For example, the virus that causes human influenza (flu) binds specifically to receptors on membranes of cells of the respiratory system. Chemical differences in the cell-surface receptors among hosts mean that a virus that infects a specific species (for example, humans) cannot infect another species (for example, chickens).

However, viruses have very small amounts of DNA or RNA compared to humans, and, as a result, viral reproduction can occur rapidly. Viral reproduction invariably produces errors that can lead to changes in newly produced viruses; these changes mean that the viral proteins that interact with cell-surface receptors may evolve in such a way that they can bind to receptors in a new host. Such changes happen randomly and quite often in the reproductive cycle of a virus, but the changes only matter if a virus with new binding properties comes into contact with a suitable host. In the case of influenza, this situation can occur in settings where animals and people are in close contact, such as poultry and swine farms (Sigalov, 2010). Once a virus jumps to a new host, it can spread quickly. Scientists watch newly appearing viruses (called emerging viruses) closely in the hope that such monitoring can reduce the likelihood of global viral epidemics.

REFERENCES

Text adapted from: OpenStax, Biology. OpenStax CNX. October 13, 2017. https://cnx.org/contents/GFy_h8cu@10.118:H4oMpCSi@8/Signaling-Molecules-and-Cellul#footnote1

A. B. Sigalov, The School of Nature. IV. Learning from Viruses, *Self/Nonself* 1, no. 4 (2010): 282-298. Y. Cao, X. Koh, L. Dong, X. Du, A. Wu, X. Ding, H. Deng, Y. Shu, J. Chen, T. Jiang, Rapid Estimation of Binding Activity of Influenza Virus Hemagglutinin to Human and Avian Receptors, *PLoS One* 6, no. 4 (2011): e18664.

Types of signaling molecules

Ligands are produced by signaling cells and act as chemical signals that travel to target cells to coordinate responses. The types of molecules that serve as ligands are incredibly varied and range from small proteins to small ions like calcium (Ca^{2+}).

SMALL HYDROPHOBIC LIGANDS

Small hydrophobic ligands can directly diffuse through the plasma membrane and interact with internal receptors. Important members of this class of ligands are the **steroid hormones**. Steroids are lipids that have a hydrocarbon skeleton with four fused rings; different steroids have different functional groups attached to the carbon skeleton. Steroid hormones include the female sex hormone, estradiol, which is a type of estrogen; the male sex hormone, testosterone; and cholesterol, which is an important structural component of biological membranes and a precursor of steriod hormones (**Figure 1**). Other hydrophobic hormones include thyroid hormones and vitamin D.

The molecular structures of estradiol, testosterone, and cholesterol are shown. All three molecules share a four-ring structure but differ in the types of functional groups attached to it.

Figure 1 Steroid hormones have similar chemical structures to their precursor, cholesterol. Because these molecules are small and hydrophobic, they can diffuse directly across the plasma membrane into the cell, where they interact with internal receptors.

WATER-SOLUBLE LIGANDS

Water-soluble ligands are polar and therefore cannot pass through the plasma membrane unaided; sometimes, they are too large to pass through the membrane at all. Instead, most water-soluble ligands bind to the portion of a cell-surface receptor which is on the outside of the cell. This group of ligands is quite diverse and includes small molecules, peptides (short chains of amino acids), and proteins.

OTHER LIGANDS

Nitric oxide (NO) is a gas that also acts as a ligand. It is able to diffuse directly across the plasma membrane, and one of its roles is to interact with receptors in smooth muscle and induce relaxation of the tissue. NO has a very short half-life and therefore only functions over short distances. Nitroglycerin, a treatment for heart disease, acts by triggering the release of NO, which causes blood vessels to dilate (expand), thus restoring blood flow to the heart. NO has become better known recently because the pathway that it affects is targeted by prescription medications for erectile

dysfunction, such as Viagra (erection involves dilated blood vessels).

REFERENCES

OpenStax, Biology. OpenStax CNX. October 13, 2017. https://cnx.org/contents/GFy_h8cu@10.118:H4oMpCSi@8/Signaling-Molecules-and-Cellul#footnote1

Propagation of the signal

Once a ligand binds to a receptor, the signal is transmitted through the membrane and into the cytoplasm. Continuation of a signal in this manner is called **signal transduction**. Signal transduction only occurs with cell-surface receptors because internal receptors are able to interact directly with DNA in the nucleus to initiate protein synthesis.

BINDING INITIATES A SIGNALING PATHWAY

After the ligand binds to the cell-surface receptor, the activation of the receptor's intracellular components sets off a chain of events that is called a **signaling pathway** or a **signaling cascade**. Signaling pathways can get very complicated very quickly because most cellular proteins can affect different downstream events, depending on the conditions within the cell. A single pathway can branch off toward different endpoints based on the interplay between two or more signaling pathways, and the same ligands are often used to initiate different signals in different cell types. This variation in response is due to differences in protein expression in different cell types. Another complicating element is signal integration of the pathways, in which signals from two or more different cell-surface receptors merge to activate the same response in the cell. This process

can ensure that multiple external requirements are met before a cell commits to a specific response.

The effects of extracellular signals can also be amplified by enzymatic cascades. At the initiation of the signal, a single ligand binds to a single receptor. However, activation of a receptor-linked enzyme can activate many copies of a component of the signaling cascade, which amplifies the signal.

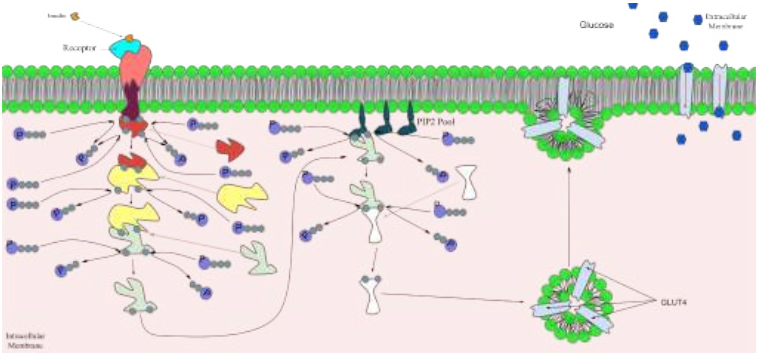


Figure 1 Example of a signal transduction cascade. In this example, insulin serves as the ligand and activates a cascade that leads to activation of the GLUT4 glucose transporter on the cell membrane. Photo credit Luuis12321; Wikimedia commons

METHODS OF INTRACELLULAR SIGNALING

The activation of a signaling pathway depends on the modification of a cellular component by an enzyme. There are numerous types of enzymatic modifications that can occur, and they are recognized in turn by the next component downstream. The following are some of the more common events in intracellular signaling.

Phosphorylation

One of the most common chemical modifications that occurs in signaling pathways is the addition of a phosphate group (PO_4^{-3}) to a molecule such as a protein in a process called **phosphorylation.** The transfer of the phosphate is catalyzed by an enzyme called a **kinase.** Various kinases are named for the substrate they phosphorylate. Phosphorylation can create a binding site that interacts with downstream components in the signaling cascade. Phosphorylation may activate or inactivate enzymes, and the reversal of phosphorylation, dephosphorylation by a phosphatase, will reverse the effect.

Second Messengers

Second messengers are small molecules that help to spread a signal through the cytoplasm after a ligand binds to a receptor. They do this by altering the behavior of certain cellular proteins. Some examples of second messengers are cAMP (a modified version of AMP, which is related to ATP but only contains one phosphate) and calcium ions.

REFERENCES

OpenStax, Biology. OpenStax CNX. October 13, 2017. https://cnx.org/contents/GFy_h8cu@10.118:sTkSwdf7@5/Propagation-of-the-Signal

Response to the signal

Ligands which can enter the cell and bind to internal receptors are able to directly affect the cell's DNA and protein-producing machinery. Ligands which can not enter the cell bind to receptors in the plasma membrane and use signal transduction pathways to produce a variety of effects. The results of signaling pathways are extremely varied and depend on the type of cell involved as well as the external and internal conditions. A small sampling of responses is described below.

GENE EXPRESSION

Some signal transduction pathways regulate the production of mRNA. Others regulate the synthesis of proteins from mRNA by ribosomes. Typically these pathways increase gene expression so that more of a specific protein is produced from a gene, but some pathways do decrease gene expression.

An example of a protein that regulates translation in the nucleus is the MAP kinase ERK. ERK is activated in a phosphorylation cascade when epidermal growth factor (EGF) binds the EGF receptor. Upon phosphorylation, ERK enters the nucleus and activates a protein kinase that, in turn, regulates protein translation (Figure 1).

This illustration shows the pathway by which ERK, a MAP kinase, activates protein synthesis. Phosphorylated ERK phosphorylates MNK1, which in turn phosphorylates eIF-4E, which is associated with mRNA. When eIF-4E is phosphorylated, the mRNA unfolds and protein synthesis begins.

Figure 1 ERK is a MAP kinase that activates translation when it is phosphorylated. ERK phosphorylates MNK1, which in turn phosphorylates eIF-4E, an elongation initiation factor that, with other initiation factors, is associated with mRNA. When eIF-4E becomes phosphorylated, the mRNA unfolds, allowing protein synthesis in the nucleus to begin.

INCREASE IN CELLULAR METABOLISM

The result of another signaling pathway affects muscle cells. The activation of β-adrenergic receptors in muscle cells by adrenaline leads to an increase in cyclic AMP (cAMP) inside the cell. Also known as epinephrine, adrenaline is a hormone (produced by the adrenal gland attached to the kidney) that readies the body for short-term emergencies. Cyclic AMP activates PKA (protein kinase A), which in turn phosphorylates two enzymes. The first enzyme promotes the degradation of glycogen by activating intermediate glycogen phosphorylase kinase (GPK) that in turn activates glycogen phosphorylase (GP) that catabolizes glycogen into glucose. (Recall that your body converts excess glucose to glycogen for short-term storage. When energy is needed, glycogen is quickly reconverted to glucose.) Phosphorylation of the second enzyme, glycogen synthase (GS), inhibits its ability to form glycogen from glucose. In this manner, a muscle cell obtains a ready pool of glucose by activating its formation via glycogen degradation and by inhibiting the use of glucose to form glycogen, thus preventing a futile cycle of glycogen

degradation and synthesis. The glucose is then available for use by the muscle cell in response to a sudden surge of adrenaline—the "fight or flight" reflex.

CELL GROWTH

Cell signaling pathways also play a major role in cell division. Cells do not normally divide unless they are stimulated by signals from other cells. The ligands that promote cell growth are called growth factors. Most growth factors bind to cell-surface receptors that are linked to tyrosine kinases. These cell-surface receptors are called receptor tyrosine kinases (RTKs). Activation of RTKs initiates a signaling pathway that includes a G-protein called RAS, which activates the MAP kinase pathway described earlier. The enzyme MAP kinase then stimulates the expression of proteins that interact with other cellular components to initiate cell division.

CELL DEATH

When a cell is damaged, superfluous, or potentially dangerous to an organism, a cell can initiate a mechanism to trigger programmed cell death, or **apoptosis.** Apoptosis allows a cell to die in a controlled manner that prevents the release of potentially damaging molecules from inside the cell. There are many internal checkpoints that monitor a cell's health; if abnormalities are observed, a cell can spontaneously initiate the process of apoptosis. However, in some cases, such as a viral infection or uncontrolled cell division due to cancer, the cell's normal checks and balances fail. External signaling can also initiate apoptosis. For example, most normal animal cells have receptors that interact with the extracellular matrix, a network of glycoproteins that provides structural support for cells in an

organism. The binding of cellular receptors to the extracellular matrix initiates a signaling cascade within the cell. However, if the cell moves away from the extracellular matrix, the signaling ceases, and the cell undergoes apoptosis. This system keeps cells from traveling through the body and proliferating out of control, as happens with tumor cells that metastasize.

Another example of external signaling that leads to apoptosis occurs in T-cell development. T-cells are immune cells that bind to foreign macromolecules and particles, and target them for destruction by the immune system. Normally, T-cells do not target "self" proteins (those of their own organism), a process that can lead to autoimmune diseases. In order to develop the ability to discriminate between self and non-self, immature T-cells undergo screening to determine whether they bind to so-called self proteins. If the T-cell receptor binds to self proteins, the cell initiates apoptosis to remove the potentially dangerous cell.

Apoptosis is also essential for normal embryological development. In vertebrates, for example, early stages of development include the formation of web-like tissue between individual fingers and toes (Figure 2). During the course of normal development, these unneeded cells must be eliminated, enabling fully separated fingers and toes to form. A cell signaling mechanism triggers apoptosis, which destroys the cells between the developing digits.

This photo shows a histological section of a foot of a 15-day-old mouse embryo. Tissue connects the space between the toes.

Figure 2 The histological section of a foot of a 15-day-old mouse embryo, visualized using light microscopy, reveals areas of tissue between the toes, which apoptosis will eliminate before the mouse reaches its full gestational age at 27 days. (credit: modification of work by Michal Mañas)

TERMINATION OF THE SIGNAL CASCADE

The aberrant signaling often seen in tumor cells is proof that the termination of a signal at the appropriate time can be just as important as the initiation of a signal. One method of stopping a specific signal is to degrade the ligand or remove it so that it can no longer access its receptor. One reason that hydrophobic hormones like estrogen and testosterone trigger long-lasting events is because they bind carrier proteins. These proteins allow the insoluble molecules to be soluble in blood, but they also protect the hormones from degradation by circulating enzymes.

Inside the cell, many different enzymes reverse the cellular modifications that result from signaling cascades. For example, phosphatases are enzymes that remove the phosphate group attached to proteins by kinases in a process called dephosphorylation. Cyclic AMP (cAMP) is degraded into AMP by phosphodiesterase, and the release of calcium stores is reversed by the Ca^{2+} pumps that are located in the external and internal membranes of the cell.

REFERENCES

OpenStax, Biology. OpenStax CNX. October 13, 2017. https://cnx.org/contents/GFy_h8cu@10.118:yQ4MP9JZ@7/Response-to-the-Signal

ENZYME-CATALYZED REACTIONS

Learning Objectives

By the end of this section, you will be able to:

- Explain the role of enzyme-catalyzed reactions in cellular metabolism.

Figure 1 A hummingbird needs energy to maintain prolonged flight. The bird obtains its energy from taking in food and transforming the energy contained in food molecules into forms of energy to power its flight through a series of biochemical reactions. (credit: modification of work by Cory Zanker)

Virtually every task performed by living organisms requires **energy.** Energy is needed to perform heavy labor and exercise, but humans also use energy while thinking, and even during sleep. In fact, the living cells of every organism constantly use energy. Nutrients and other molecules are imported into the cell have many different potential paths: metabolized (broken down) and used for energy, synthesized into new molecules, modified if needed, transported around the cell, and even distributed to the entire organism. For example, the large proteins that make up muscles are built from smaller molecules imported from dietary amino acids. Complex carbohydrates are broken down into simple sugars that the cell uses for energy. Just as energy is required to both build and demolish a building, energy is required for the synthesis and breakdown of molecules as well as the transport of molecules into and out of cells. In addition, processes such as ingesting and breaking down pathogenic bacteria and viruses, exporting wastes and toxins, and movement of the cell require energy.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Energy

Scientists use the term **bioenergetics** to describe the concept of energy flow (**Figure 2**) through living systems, such as cells. Cellular processes such as the building and breaking down of complex molecules occur through stepwise chemical reactions. Some of these chemical reactions are spontaneous and release energy, whereas others require energy to proceed.

Just as living things must continually consume food to replenish their energy supplies, cells must continually produce more energy to replenish that used by the many energy-requiring chemical reactions that constantly take place. Together, all of the chemical reactions that take place inside cells, including those that consume or generate energy, are referred to as the cell's **metabolism**.

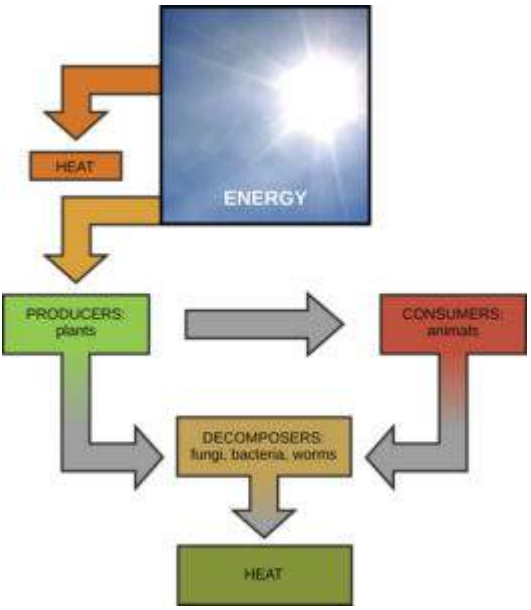


Figure 2 Ultimately, most life forms get their energy from the sun. Plants use photosynthesis to capture sunlight, and herbivores eat the plants to obtain energy. Carnivores eat the herbivores, and eventual decomposition of plant and animal material contributes to the nutrient pool.

ENERGY

Thermodynamics refers to the study of energy and energy transfer involving physical matter. The matter relevant to a particular case of energy transfer is called a system, and everything outside of that matter is called the surroundings. For instance, when heating a pot of water on the stove, the system includes the stove, the pot, and the water. Energy is transferred within the system (between the stove, pot, and water). There are two types of systems: open and closed. In an **open system**, energy can be exchanged with its surroundings. The stovetop system is open because heat can

be lost to the air. A **closed system** cannot exchange energy with its surroundings.

Biological organisms are open systems. Energy is exchanged between them and their surroundings as they use energy from the sun to perform photosynthesis or consume energy-storing molecules and release energy to the environment by doing work and releasing heat. Like all things in the physical world, energy is subject to physical laws. The laws of thermodynamics govern the transfer of energy in and among all systems in the universe.

In general, **energy** is defined as the ability to do work, or to create some kind of change. Energy exists in different forms. For example, electrical energy, light energy, and heat energy are all different types of energy. To appreciate the way energy flows into and out of biological systems, it is important to understand two of the physical laws that govern energy.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Metabolic Pathways

Consider the metabolism of sugar. This is a classic example of one of the many cellular processes that use and produce energy. Living things consume sugars as a major energy source, because sugar molecules have a great deal of energy stored within their bonds. For the most part, photosynthesizing organisms like plants produce these sugars. During photosynthesis, plants use energy (originally from sunlight) to convert carbon dioxide gas (CO_2) into sugar molecules (like glucose: $C_6H_{12}O_6$). They consume carbon dioxide and produce oxygen as a waste product. This reaction is summarized as:

$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$

Because this process involves synthesizing an energy-storing molecule, it requires energy input to proceed. During the light reactions of photosynthesis, energy is provided by a molecule called **adenosine triphosphate (ATP)**, which is the primary energy currency of all cells. Just as the dollar is used as currency to buy goods, cells use molecules of ATP as energy currency to perform immediate work. In contrast, energy-storage molecules such as glucose are consumed only to be broken down to use their energy. The reaction that harvests the energy of a sugar molecule in cells requiring oxygen to survive can be summarized by the reverse reaction to photosynthesis. In this reaction, oxygen is consumed and

carbon dioxide is released as a waste product. The reaction is summarized as:

$C_6H_{12}O_6 + 6O_2 \rightarrow 6H_2O + 6CO_2$

Both of these reactions involve many steps.

The processes of making and breaking down sugar molecules illustrate two examples of metabolic pathways. A **metabolic pathway** is a series of chemical reactions that takes a starting molecule and modifies it, step-by-step, through a series of metabolic intermediates, eventually yielding a final product. In the example of sugar metabolism, the first metabolic pathway synthesized sugar from smaller molecules, and the other pathway broke sugar down into smaller molecules. These two opposite processes—the first requiring energy and the second producing energy—are referred to as **anabolic** pathways (building polymers) and **catabolic** pathways (breaking down polymers into their monomers), respectively. Consequently, metabolism is composed of synthesis (anabolism) and degradation (catabolism) (**Figure 3**).

It is important to know that the chemical reactions of metabolic pathways do not take place on their own. Each reaction step is facilitated, or catalyzed, by a protein called an **enzyme**. Enzymes are important for catalyzing all types of biological reactions—those that require energy as well as those that release energy.

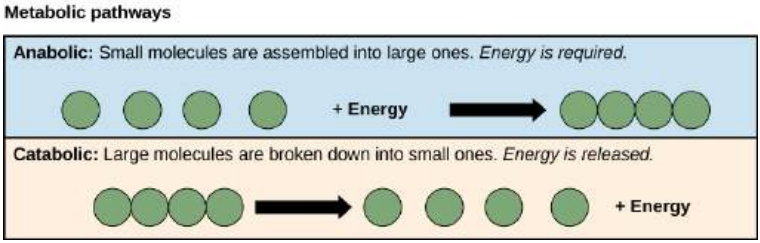


Figure 3 Catabolic pathways are those that generate energy by breaking down larger molecules. Anabolic pathways are those that require energy to synthesize larger

molecules. Both types of pathways are required for maintaining the cell's energy balance.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=148#h5p-46

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Thermodynamics

The first law of thermodynamics states that the total amount of energy in the universe is constant and conserved. In other words, there has always been, and always will be, exactly the same amount of energy in the universe. Energy exists in many different forms. According to the first law of thermodynamics, energy may be transferred from place to place or transformed into different forms, but it cannot be created or destroyed. The transfers and transformations of energy take place around us all the time. Light bulbs transform electrical energy into light and heat energy. Gas stoves transform chemical energy from natural gas into heat energy. Plants perform one of the most biologically useful energy transformations on earth: that of converting the energy of sunlight to chemical energy stored within organic molecules (**Figure 2**). Some examples of energy transformations are shown in **Figure 4**.

The challenge for all living organisms is to obtain energy from their surroundings in forms that they can transfer or transform into usable energy to do work. Living cells have evolved to meet this challenge. Chemical energy stored within organic molecules such as sugars and fats is transferred and transformed through a series of cellular chemical reactions into energy within molecules of **ATP** (adenosine triphosphate). Energy in ATP molecules is easily accessible to do work. Examples of the types of work that

cells need to do include building complex molecules, transporting materials, powering the motion of cilia or flagella, and contracting muscle fibers to create movement.

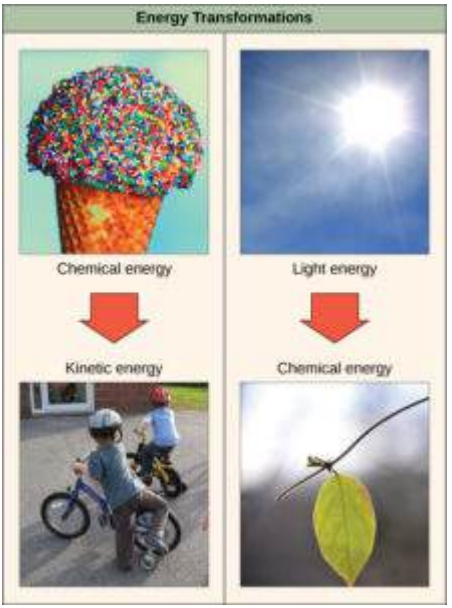


Figure 4 Shown are some examples of energy transferred and transformed from one system to another and from one form to another. (credit "ice cream": modification of work by D. Sharon Pruitt; credit "kids": modification of work by Max from Providence; credit "leaf": modification of work by Cory Zanker)

A living cell's primary tasks of obtaining, transforming, and using energy to do work may seem simple. However, the second law of thermodynamics explains why these tasks are harder than they appear. All energy transfers and transformations are never completely efficient. In every energy transfer, some amount of energy is lost in a form that is unusable. In most cases, this form is heat energy. Thermodynamically, **heat energy** is defined as the energy transferred from one system to another that is not work. For example, when a light bulb is turned on, some of the energy

being converted from electrical energy into light energy is lost as heat energy. Likewise, some energy is lost as heat energy during cellular metabolic reactions.

An important concept in physical systems is that of order and disorder. The more energy that is lost by a system to its surroundings, the less ordered and more random the system is. Scientists refer to the measure of randomness or disorder within a system as entropy. High entropy means high disorder and low energy. Molecules and chemical reactions have varying entropy as well. For example, entropy increases as molecules at a high concentration in one place diffuse and spread out. The second law of thermodynamics says that energy will always be lost as heat in energy transfers or transformations. Living things are highly ordered, requiring constant energy input to be maintained in a state of low entropy.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=150#h5p-45

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Potential and Kinetic Energy

When an object is in motion, there is energy associated with that object. Think of a wrecking ball. Even a slow-moving wrecking ball can do a great deal of damage to other objects. Energy associated with objects in motion is called **kinetic energy** (**Figure 5**). A speeding bullet, a walking person, and the rapid movement of molecules in the air (which produces heat) all have kinetic energy.

Now what if that same motionless wrecking ball is lifted two stories above ground with a crane? If the suspended wrecking ball is unmoving, is there energy associated with it? The answer is yes. The energy that was required to lift the wrecking ball did not disappear, but is now stored in the wrecking ball by virtue of its position and the force of gravity acting on it. This type of energy is called **potential energy** (**Figure 5**). If the ball were to fall, the potential energy would be transformed into kinetic energy until all of the potential energy was exhausted when the ball rested on the ground. Wrecking balls also swing like a pendulum; through the swing, there is a constant change of potential energy (highest at the top of the swing) to kinetic energy (highest at the bottom of the swing). Other examples of potential energy include the energy of water held behind a dam or a person about to skydive out of an airplane.

Figure 5 Still water has potential energy; moving water, such as in a waterfall or a rapidly flowing river, has kinetic energy. (credit "dam": modification of work by "Pascal"/Flickr; credit "waterfall": modification of work by Frank Gualtieri)

Potential energy is not only associated with the location of matter, but also with the *structure* of matter. A spring on the ground has potential energy if it is compressed; so does a rubber band that is pulled taut. On a molecular level, the bonds that hold the atoms of molecules together exist in a particular structure that has potential energy. Remember that anabolic cellular pathways *require* energy to synthesize complex molecules from simpler ones and catabolic pathways *release* energy when complex molecules are broken down. The fact that energy can be released by the breakdown of certain chemical bonds implies that those bonds have potential energy. In fact, there is potential energy stored within the bonds of all the food molecules we eat, which is eventually harnessed for use. This is because these bonds can release energy when broken. The type of potential energy that exists within chemical bonds, and is released when those bonds are broken, is called **chemical energy**. Chemical energy is responsible for providing living cells with energy from food. The release of energy occurs when the molecular bonds within food molecules are broken.

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=152#h5p-44

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Free and Activation Energy

After learning that chemical reactions release energy when energy-storing bonds are broken, an important next question is the following: How is the energy associated with these chemical reactions quantified and expressed? How can the energy released from one reaction be compared to that of another reaction? A measurement of **free energy** is used to quantify these energy transfers. Recall that according to the second law of thermodynamics, all energy transfers involve the loss of some amount of energy in an unusable form such as heat. Free energy specifically refers to the energy associated with a chemical reaction that is available after the losses are accounted for. In other words, *free energy is usable energy, or energy that is available to do work.* Looking at this concept in a biological sense, free energy is the energy within a molecule that can be used to perform work. Glucose has a lot of free energy because there is a lot of energy stored within the bonds of the glucose molecule. Carbon dioxide has a much lower free energy because there is much less energy stored in its bonds.

If energy is released during a chemical reaction, then the change in free energy from the conversion of the reactants to the products, signified as ΔG (delta G) will be a negative number. A negative change in free energy also means that the products of the reaction have less free energy than the reactants, because they release some free energy during the

reaction. Reactions that have a negative change in free energy and consequently release free energy are called **exergonic reactions**. Think: *ex*ergonic means energy is *ex*iting the system. These reactions are also referred to as spontaneous reactions, and their products have less stored energy than the reactants. An important distinction must be drawn between the term spontaneous and the idea of a chemical reaction occurring immediately. Contrary to the everyday use of the term, a spontaneous reaction is not one that suddenly or quickly occurs. The rusting of iron is an example of a spontaneous reaction that occurs slowly, little by little, over time.

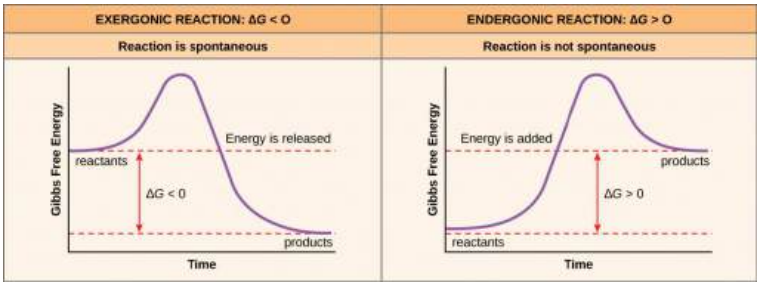


Figure 1 Free energy of endergonic and exergonic reactions. In an exergonic reaction, the reactants have more free energy than the products. Therefore, energy is released as the reaction proceeds. In an endergonic reaction, the reactants have more less energy than the products. Therefore, energy must be added to make the reaction move take place.

If a chemical reaction absorbs energy rather than releases energy on balance, then the ΔG for that reaction will be a positive value. In this case, the products have more free energy than the reactants. Thus, the products of these reactions can be thought of as energy-storing molecules. These chemical reactions are called **endergonic reactions** and they are nonspontaneous.

An endergonic reaction will **not** take place on its own without the addition of free energy.

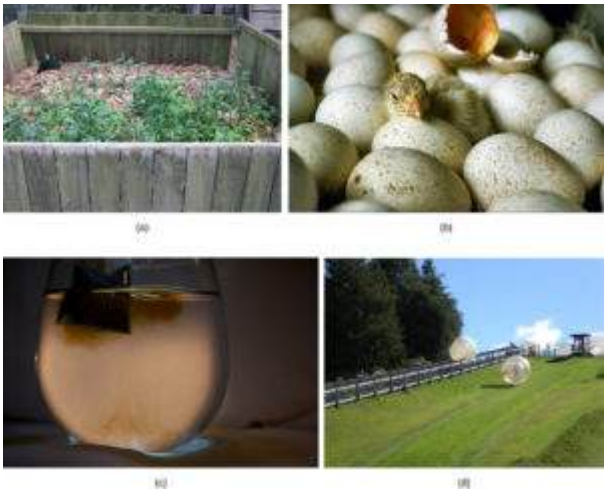


Figure 2 Shown are some examples of endergonic processes (ones that require energy) and exergonic processes (ones that release energy). (credit a: modification of work by Natalie Maynor; credit b: modification of work by USDA; credit c: modification of work by Cory Zanker; credit d: modification of work by Harry Malsch)

There is another important concept that must be considered regarding endergonic and exergonic reactions. Exergonic reactions require a small amount of energy input to get going, before they can proceed with their energy-releasing steps.

These reactions have a net release of energy, but still require some energy input in the beginning. This small amount of energy input necessary for all chemical reactions to occur is called the **activation energy (Figure 3)**.

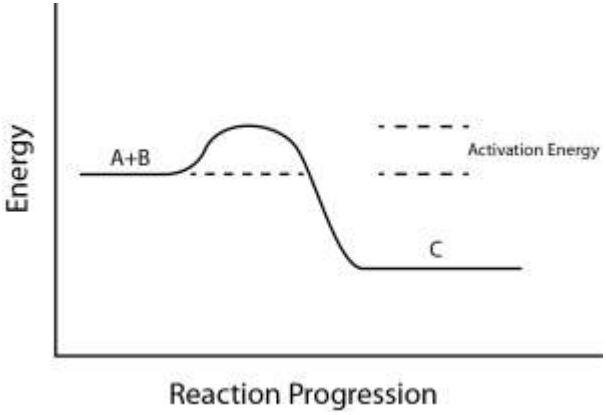


Figure 3 Activation energy is the small amount of energy that must be put into a system in order for the reaction to take place. Photo credit Brazosport College; Wikimedia.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Enzymes

A substance that helps a chemical reaction to occur is called a *catalyst*, and the molecules that catalyze biochemical reactions are called **enzymes**. Most enzymes are proteins and perform the critical task of lowering the activation energies of chemical reactions inside the cell. Most of the reactions critical to a living cell happen too slowly at normal temperatures to be of any use to the cell. Without enzymes to speed up these reactions, life could not persist. Enzymes do this by binding to the reactant molecules and holding them in such a way as to make the chemical bond-breaking and -forming processes take place more easily. It is important to remember that enzymes do not change whether a reaction is exergonic (spontaneous) or endergonic. This is because they do not change the free energy of the reactants or products. They only reduce the activation energy required for the reaction to go forward (**Figure 1**). In addition, an enzyme itself is unchanged by the reaction it catalyzes. Once one reaction has been catalyzed, the enzyme is able to participate in other reactions.

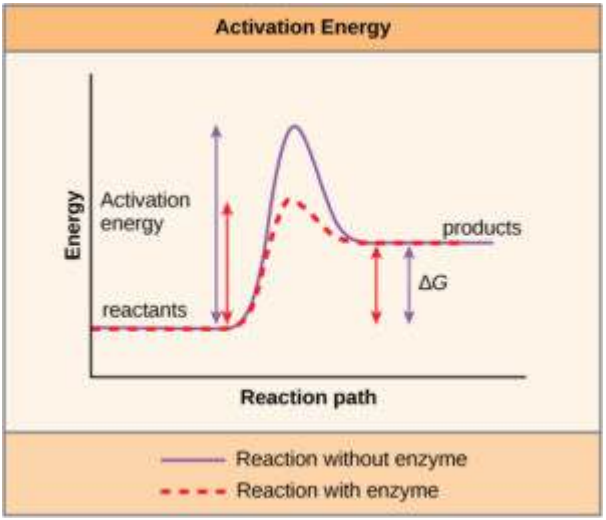


Figure 1 Enzymes lower the activation energy of the reaction but do not change the free energy of the reaction.

The chemical reactants to which an enzyme binds are called the enzyme's **substrates**. There may be one or more substrates, depending on the particular chemical reaction. In some reactions, a single reactant substrate is broken down into multiple products. In others, two substrates may come together to create one larger molecule. Two reactants might also enter a reaction and both become modified, but they leave the reaction as two products. The location within the enzyme where the substrate binds is called the enzyme's **active site**. The active site is where the "action" happens. Since enzymes are proteins, there is a unique combination of amino acid side chains within the active site. Each side chain is characterized by different properties. They can be large or small, weakly acidic or basic, hydrophilic or hydrophobic, positively or negatively charged, or neutral. The unique combination of side chains creates a very specific chemical environment within the active site. This specific environment

is suited to bind to one specific chemical substrate (or substrates).

Active sites are subject to influences of the local environment. Increasing the environmental temperature generally increases reaction rates, enzyme-catalyzed or otherwise. However, temperatures outside of an optimal range reduce the rate at which an enzyme catalyzes a reaction. Hot temperatures will eventually cause enzymes to **denature,** an irreversible change in the three-dimensional shape and therefore the function of the enzyme (**Figure 8**). Enzymes are also suited to function best within a certain pH and salt concentration range, and, as with temperature, extreme pH, and salt concentrations can cause enzymes to denature.

Figure 2 Heat applied to an egg during cooking irreversibly denatures the proteins. (credit: "K-Wall"/Flickr)

Typically, enzymes function optimally in the environment where they are typically found and used. For example, the enzyme amylase is found in saliva, where it functions to break down starch (a polysaccharide – carbohydrate chain) into smaller sugars. Note that in this example, amylase is the enzyme, starch is the substrate, and smaller sugars are the product. The pH of saliva is typically between 6.2 and

7.6, with roughly 6.7 being the average. The optimum pH of amylase is between 6.7 and 7.0, which is close to neutral (Figure 3). The optimum temperature for amylase is close to 37ºC (which is human body temperature).

graphs showing effect of temperature and pH on amylase. each graph has a bell-curve shape above the optimum pH

Figure 3 The effect of pH and temperature on the activity of an enzyme. Amylase is shown in blue in both graphs. (top) Amylase (blue) has an optimum pH of about 7. The green enzyme, which has an optimum pH of about 2.3, might function in the stomach where it is very acidic. (bottom) Amylase (blue) has an optimum temperature of about 37 degrees C. The orange enzyme, which has an optimum temperature of about 15 degrees C (about 60F) might function in a plant found outdoors.

For many years, scientists thought that enzyme-substrate binding took place in a simple "lock and key" fashion. This model asserted that the enzyme and substrate fit together perfectly in one instantaneous step. However, current research supports a model called **induced fit** (**Figure 9**). The induced-fit model expands on the lock-and-key model by describing a more dynamic binding between enzyme and substrate. As the enzyme and substrate come together, their interaction causes a mild shift in the enzyme's structure that forms an ideal binding arrangement between enzyme and substrate.

When an enzyme binds its substrate, an enzyme-substrate complex is formed. This complex lowers the activation energy of the reaction and promotes its rapid progression in one of multiple possible ways.

- On a basic level, enzymes promote chemical reactions that involve more than one substrate by bringing the substrates together in an optimal orientation for reaction.

- Enzymes promote the reaction of their substrates is by creating an optimal environment within the active site for the reaction to occur. The chemical properties that emerge from the particular arrangement of amino acid R groups (side chains) within an active site create the perfect environment for an enzyme's specific substrates to react.
- The enzyme-substrate complex can also lower activation energy by compromising the bond structure so that it is easier to break.
- Finally, enzymes can also lower activation energies by taking part in the chemical reaction itself. In these cases, it is important to remember that the enzyme will always return to its original state by the completion of the reaction.

One of the hallmark properties of enzymes is that they remain ultimately unchanged by the reactions they catalyze. After an enzyme has catalyzed a reaction, it releases its product(s) and can catalyze a new reaction.

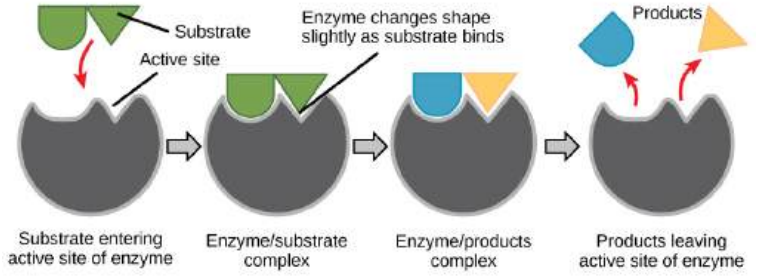


Figure 9 The induced-fit model is an adjustment to the lock-and-key model and explains how enzymes and substrates undergo dynamic modifications during the transition state to increase the affinity of the substrate for the active site.

It would seem ideal to have a scenario in which all of an

organism's enzymes existed in abundant supply and functioned optimally under all cellular conditions, in all cells, at all times. However, a variety of mechanisms ensures that this does not happen. Cellular needs and conditions constantly vary from cell to cell, and change within individual cells over time. The required enzymes of stomach cells differ from those of fat storage cells, skin cells, blood cells, and nerve cells. Furthermore, a digestive organ cell works much harder to process and break down nutrients during the time that closely follows a meal compared with many hours after a meal. As these cellular demands and conditions vary, so must the amounts and functionality of different enzymes.

Since the rates of biochemical reactions are controlled by activation energy, and enzymes lower and determine activation energies for chemical reactions, the relative amounts and functioning of the variety of enzymes within a cell ultimately determine which reactions will proceed and at what rates. This determination is tightly controlled in cells. In certain cellular environments, enzyme activity is partly controlled by environmental factors like pH, temperature, salt concentration, and, in some cases, cofactors or coenzymes.

Enzymes can also be regulated in ways that either promote or reduce enzyme activity. There are many kinds of molecules that inhibit or promote enzyme function, and various mechanisms by which they do so. In some cases of enzyme inhibition, an inhibitor molecule is similar enough to a substrate that it can bind to the active site and simply block the substrate from binding. When this happens, the enzyme is inhibited through **competitive inhibition**, because an inhibitor molecule competes with the substrate for binding to the active site.

On the other hand, in **noncompetitive inhibition**, an inhibitor molecule binds to the enzyme in a location other than the active site, called an allosteric site, but still manages

to block substrate binding to the active site. Some inhibitor molecules bind to enzymes in a location where their binding induces a conformational change that reduces the affinity of the enzyme for its substrate. This type of inhibition is called **allosteric inhibition** (**Figure 10**). Most allosterically regulated enzymes are made up of more than one polypeptide, meaning that they have more than one protein subunit. When an allosteric inhibitor binds to a region on an enzyme, all active sites on the protein subunits are changed slightly such that they bind their substrates with less efficiency. There are allosteric activators as well as inhibitors. Allosteric activators bind to locations on an enzyme away from the active site, inducing a conformational change that increases the affinity of the enzyme's active site(s) for its substrate(s) (**Figure 10**).

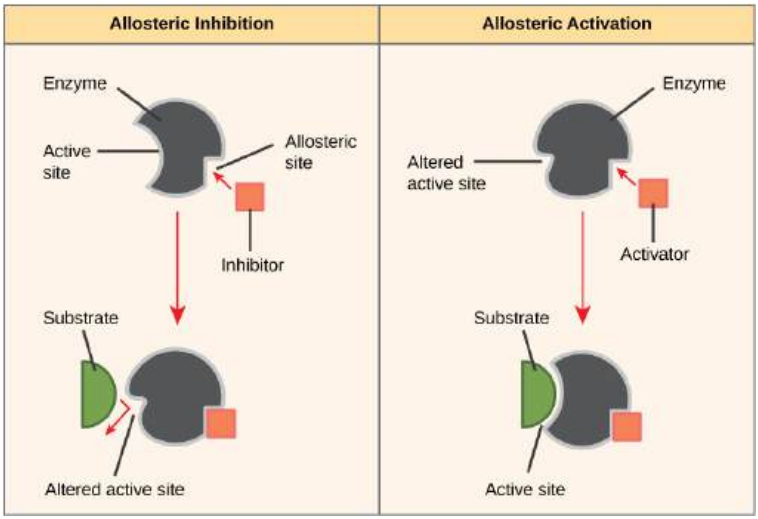


Figure 10 Allosteric inhibition works by indirectly inducing a conformational change to the active site such that the substrate no longer fits. In contrast, in allosteric activation, the activator molecule modifies the shape of the active site to allow a better fit of the substrate.

Many enzymes do not work optimally, or even at all, unless bound to other specific non-protein helper molecules. They may bond either temporarily through ionic or hydrogen bonds, or permanently through stronger covalent bonds. Binding to these molecules promotes optimal shape and function of their respective enzymes. Two examples of these types of helper molecules are *cofactors* and *coenzymes*. Cofactors are inorganic ions such as ions of iron and magnesium. Coenzymes are organic helper molecules, those with a basic atomic structure made up of carbon and hydrogen. Like enzymes, these molecules participate in reactions without being changed themselves and are ultimately recycled and reused. Vitamins are the source of coenzymes. Some vitamins are the precursors of coenzymes and others act directly as coenzymes. Vitamin C is a direct coenzyme for multiple enzymes that take part in building the important connective tissue, collagen. Therefore, enzyme function is, in part, regulated by the abundance of various cofactors and coenzymes, which may be supplied by an organism's diet or, in some cases, produced by the organism.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=159#h5p-47

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=159#h5p-48

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=159#h5p-49

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=159#h5p-50

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=159#h5p-51

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Feedback Inhibition in Metabolic Pathways

Molecules can regulate enzyme function in many ways. The major question remains, however: What are these molecules and where do they come from? Some are cofactors and coenzymes, as you have learned. What other molecules in the cell provide enzymatic regulation such as allosteric modulation, and competitive and non-competitive inhibition? Perhaps the most relevant sources of regulatory molecules, with respect to enzymatic cellular metabolism, are the products of the cellular metabolic reactions themselves. In a most efficient and elegant way, cells have evolved to use the products of their own reactions for feedback inhibition of enzyme activity. **Feedback inhibition** involves the use of a reaction product to regulate its own further production (**Figure 11**). The cell responds to an abundance of the products by slowing down production during anabolic or catabolic reactions. Such reaction products may inhibit the enzymes that catalyzed their production through the mechanisms described above.

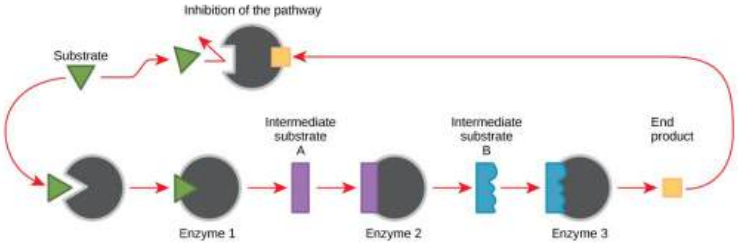


Figure 11 Metabolic pathways are a series of reactions catalyzed by multiple enzymes. Feedback inhibition, where the end product of the pathway inhibits an upstream process, is an important regulatory mechanism in cells.

The production of both amino acids and nucleotides is controlled through feedback inhibition. Additionally, ATP is an allosteric regulator of some of the enzymes involved in the catabolic breakdown of sugar, the process that creates ATP. In this way, when ATP is in abundant supply, the cell can prevent the production of ATP. On the other hand, ADP serves as a positive allosteric regulator (an allosteric activator) for some of the same enzymes that are inhibited by ATP. Thus, when relative levels of ADP are high compared to ATP, the cell is triggered to produce more ATP through sugar catabolism.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

HOW CELLS OBTAIN ENERGY

Learning Objectives

By the end of this section, you will begin to be able to:

- Compare energy-generating processes within different types of cells.

Energy in Living Systems

All living organisms require energy to perform their life processes. Energy, as you learned earlier in the chapter about enzymes, is the ability to do work or to create some kind of change. You are familiar with or have learned about many processes that can require energy:

- Movement
- Reproduction
- Maintaining homeostasis of many different conditions
- Acquiring and digesting food
- Producing proteins

Just as living things must continually consume food to replenish their energy supplies, cells must continually produce more energy to replenish that used by the many energy-requiring chemical reactions that constantly take place. Together, all of the chemical reactions that take place inside cells, including those that consume or generate energy, are referred to as the cell's **metabolism**.

A living cell cannot store significant amounts of free energy. Free energy is energy that is not stored in molecules. Excess free energy would result in an increase of heat in the cell, which would denature enzymes and other proteins,

and destroy the cell. Instead, a cell must be able to store energy safely and release it for use only as needed. Living cells accomplish this using ATP, which can be used to fill any energy need of the cell. How? It functions like a rechargeable battery.

When ATP is broken down, energy is released. This energy is used by the cell to do work. For example, in the mechanical work of muscle contraction, ATP supplies energy to move the contractile muscle proteins.

ATP STRUCTURE AND FUNCTION

ATP is a complex-looking molecule, but for our purposes you can think of it as a rechargeable battery. ATP, the fully charged form of our battery, is made up of three phosphates (the "TP" part of ATP means "tri phosphate") attached to a sugar and an adenine (the "A" part of ATP) (**Figure 1**). When the last phosphate is broken off of the ATP, energy is released. The result is a single phosphate and a molecule called ADP ("D" stands for "di" which means two).

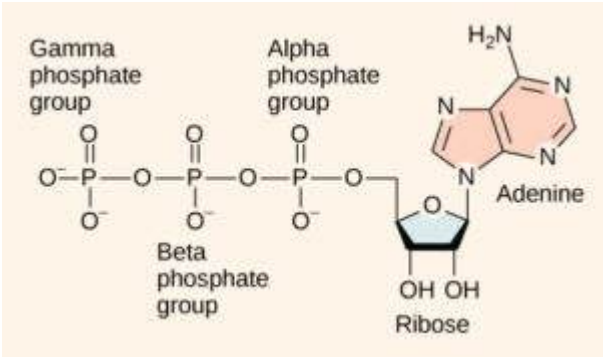


Figure 1 The structure of ATP shows the basic components of a two-ring adenine, five-carbon ribose sugar, and three phosphate groups.

A large amount of energy is required in order to recharge a

molecule of ADP into ATP. This energy is stored in the bond between the second and third phosphates. When this bond is broken, the energy is released in a way that the cell can use it.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=164#h5p-52

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

From Mouth to Molecule: Digestion

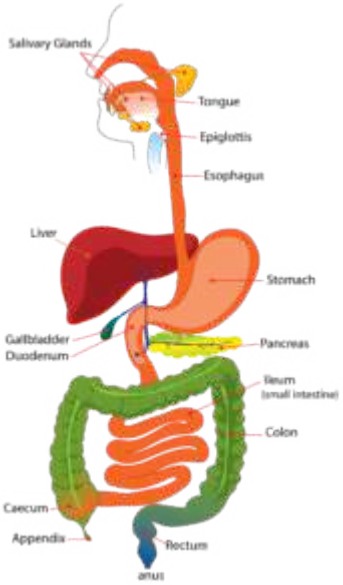
While plants can produce their own energy using the process of photosynthesis, animals (and other organisms that can't do photosynthesis) must eat to get energy from food molecules. Just like energy can be stored in the chemical bond between the second and third phosphate of an ATP molecule, energy can also be stored in the chemical bonds that make up food molecules. Most of the energy that we use comes from molecules of glucose, a simple sugar.

Food energy is chemical energy that animals (including humans) derive from their food through the process of cellular respiration. Cellular respiration involves either joining oxygen from air with the molecules of food (aerobic respiration) or reorganizing the atoms within the molecules in the absence of oxygen (anaerobic respiration).

Humans and other animals need a minimum intake of food energy to sustain their metabolism and to drive their muscles. Foods are composed chiefly of carbohydrates, fats, proteins, water, vitamins, and minerals. Carbohydrates, fats, proteins, and water represent virtually all the weight of food, with vitamins and minerals making up only a small percentage of the weight. In fact, carbohydrates, fats, and proteins comprise ninety percent of the dry weight of foods. Organisms derive food

energy mainly from carbohydrates and fats present in the diet, and to a smaller extent proteins and other organic molecules. Some diet components that provide little or no food energy, such as water, minerals, vitamins, cholesterol, and fiber, may still be necessary to health and survival for other reasons. Water, minerals, vitamins, and cholesterol are not broken down; they are used by the body in the form in which they are taken in, so they cannot be used for energy. Fiber, a type of carbohydrate, cannot be completely digested by the human body so energy is not released from fiber when it is digested. Instead, it moves mostly intact through the digestive system.

After you put food into your mouth, you begin to break it down mechanically using your teeth. Enzymes in your saliva begin breaking the food molecules down as well. After you swallow your food, it is further broken down by additional enzymes in the stomach, followed by the small intestine. In the small intestine, the fully broken-down food is absorbed into the blood. The majority of the nutrients (about 95%) are absorbed in the small intestine. Water is reabsorbed from the remaining material in the colon. Then the residual waste is eliminated during defecation.



The human digestive system. (Credit: Leysi24, from Wikimedia. Creative Commons Attribution-Share Alike 3.0 Unported)

Once in the bloodstream, nutrients enter individual cells. Glucose is too large to diffuse through the cell membrane and is typically transported inside cells by proteins. After molecules enter a cell, the breakdown process to produce energy in the form of ATP can be completed.

REFERENCES

Wikipedia. Creative Commons Attribution-ShareAlike License.

Metabolism

An organism's metabolism is the sum total of all the chemical reactions that occur within the organism. These chemical reactions fall into two basic categories:

- Anabolism: building polymers (large molecules that the cell needs).
- Catabolism: breaking down polymers to release energy.

This means that metabolism is composed of synthesis (anabolism) and degradation (catabolism) (**Figure 1**).

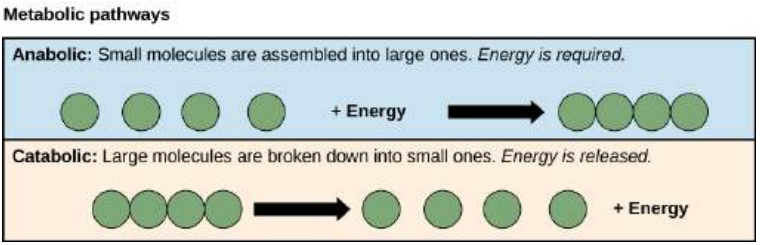


Figure 1 Catabolic pathways are those that generate energy by breaking down larger molecules. Anabolic pathways are those that require energy to synthesize larger molecules. Both types of pathways are required for maintaining the cell's energy balance.

It is important to know that the chemical reactions of metabolic pathways do not take place on their own. Each

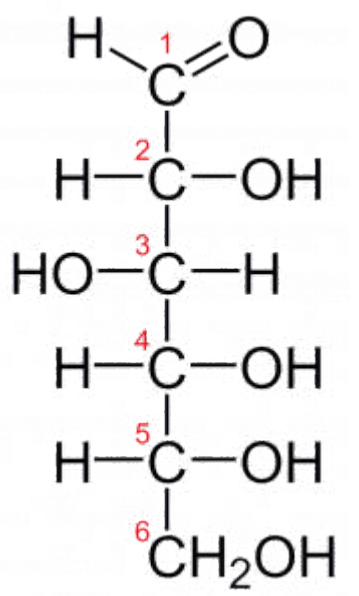
reaction step is facilitated, or catalyzed, by a protein called an **enzyme**. Enzymes are important for catalyzing all types of biological reactions—those that require energy as well as those that release energy. Refer back to the chapter on enzymes if you need a reminder about this topic.

Consider the metabolism of sugar (a carbohydrate). This is a classic example of one of the many cellular processes that use and produce energy. Living things consume sugars as a major energy source, because sugar molecules have a great deal of energy stored within their bonds. For the most part, photosynthesizing organisms like plants produce these sugars. During photosynthesis, plants use energy (originally from sunlight) to convert carbon dioxide gas (CO_2) into sugar molecules (like glucose: $C_6H_{12}O_6$). They consume carbon dioxide and produce oxygen as a waste product. This reaction is summarized as:

$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$

Recall from chemistry that the abbreviation "CO_2" means "one carbon atom covalently bonded to two oxygen atoms." Water, "H_2O" is two hydrogen atoms covalently bonded to one oxygen atom. And "$C_6H_{12}O_6$" has 6 carbon atoms, 12 hydrogen atoms, and 6 oxygen atoms that are covalently bonded together.

Carbon dioxide (CO2) contains one carbon atom covalently bonded to two oxygen atoms. Credit: wikimedia



Glucose contains 6 carbons, 6 oxygens, and 12 hydrogen atoms. Credit: Ben, 2006. Wikimedia. Public domain.

The process of producing glucose from carbon dioxide and water requires an energy input to proceed because glucose contains more energy in its molecular bonds than carbon dioxide does.

In contrast, energy-storage molecules such as glucose are consumed to be broken down to use their energy. The reaction that harvests the energy of a sugar molecule in cells requiring oxygen to survive can be summarized by the reverse reaction to photosynthesis. In this reaction, oxygen is consumed and carbon dioxide is released as a waste product. The reaction is summarized as:

$C_6H_{12}O_6 + 6O_2 \rightarrow 6H_2O + 6CO_2$

Both of these reactions involve many steps.

The processes of making and breaking down sugar molecules illustrate two examples of metabolic pathways. A **metabolic pathway** is a series of chemical reactions that

takes a starting molecule and modifies it, step-by-step, through a series of metabolic intermediates, eventually yielding a final product. In the example of sugar metabolism, the first metabolic pathway synthesized sugar from smaller molecules, and the other pathway broke sugar down into smaller molecules.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=167#h5p-53

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

An overview of Cellular Respiration

Glucose and other molecules from food are broken down to release energy in a complex series of chemical reactions that together are called **cellular respiration**.

Cellular respiration is a set of metabolic reactions and processes that take place in the cells of organisms to convert biochemical energy from nutrients into ATP, and then release waste products. The reactions involved in respiration are catabolic reactions, which break large molecules into smaller ones, releasing energy in the process. These processes require a large number of enzymes which each perform one specific chemical reaction.

AEROBIC RESPIRATION

Aerobic respiration requires oxygen. This is the reason why we breathe oxygen in from the air. This type of respiration releases a large amount of energy from glucose that can be stored as ATP. Aerobic respiration happens all the time in animals and plants, where most of the reactions occur in the mitochondria. Even some prokaryotes can perform aerobic respiration (although since prokaryotes don't contain mitochondria, the reactions are slightly

different). The overall chemical formula for aerobic respiration can be written as:

$C_6H_{12}O_2 + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O$ + (approximately) 38 ATP

Translating that formula into English: One molecule of glucose can be broken down in the presence of oxygen gas to produce waste products of carbon dioxide (which we breathe out) and water. This process has an overall release of energy which is captured and stored in 38 molecules of ATP.

Aerobic respiration is a complex process that can be divided into three basic stages: glycolysis, the citric acid cycle, and oxidative phosphorylation. The next several sections in the textbook address the details of these stages.

ANAEROBIC RESPIRATION

Anaerobic respiration occurs in the absence of oxygen. It releases a much smaller amount of energy than aerobic respiration. Anaerobic respiration does not release enough energy to power human cells for long – think about how long a person can live if they are not able to breathe. Anaerobic respiration occurs in muscle cells during hard exercise (after the oxygen has been used up). It also occurs in yeast when brewing beer. Many prokaryotes perform anaerobic respiration.

There are several different types of anaerobic respiration, which will be discussed in more detail later. All the types of anaerobic respiration involve glycolysis, and none of them go through the citric acid cycle or oxidative phosphorylation. Instead, various other methods are used to regenerate the molecules needed for glycolysis, For now, we will summarize them all using this chemical formula:

$C_6H_{12}O_2$ NAD+ $\rightarrow$ various waste products + NADH + 2 ATP

NAD+ and NADH are two states of a molecule that will carry energy during this process. It will be addressed further in

a later section. For right now, just know that NADH carries energy (similar to ATP) and NAD+ is the form that carries less energy (similar to ADP)

AEROBIC VS ANAEROBIC RESPIRATION

	Aerobic	Anaerobic
Requires oxygen?	Yes	No
Glucose breakdown	Complete	Incomplete
End products	CO_2 and H_2O	Animal cells: lactic acid Plant cells and yeast: carbon dioxide and ethanol
ATP produced	About 38	2

Aerobic respiration is much more efficient than anaerobic respiration. One molecule of glucose can generate up to 38 molecules of ATP if aerobic respiration is used. In contrast, only 2 molecules of ATP are generated in anaerobic respiration.

To put it another way, a cellular process which requires 100 molecules of ATP:

- Will require about 2.5 molecules of glucose to be broken down using aerobic respiration (100 / 38 = 2.63)
- Will require 50 molecules of glucose to be broken down using anaerobic respiration (100 / 2 = 50)

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=168#h5p-54

REFERENCES

Wikipedia.

Aerobic Respiration, Part 1: Glycolysis

You have read that nearly all of the energy used by living things comes to them in the bonds of the sugar, glucose. **Glycolysis** is the first step in the breakdown of glucose to extract energy for cell metabolism. Many living organisms carry out glycolysis as part of their metabolism. Glycolysis takes place in the cytoplasm of most prokaryotic and all eukaryotic cells.

Glycolysis begins with a molecule of **glucose** ($C_6H_{12}O_6$). Various enzymes are used to break glucose down into two molecules of **pyruvate** ($C_3H_4O_3$, basically a glucose molecule broken in half) (**Figure 1**). This process releases a small amount of energy.

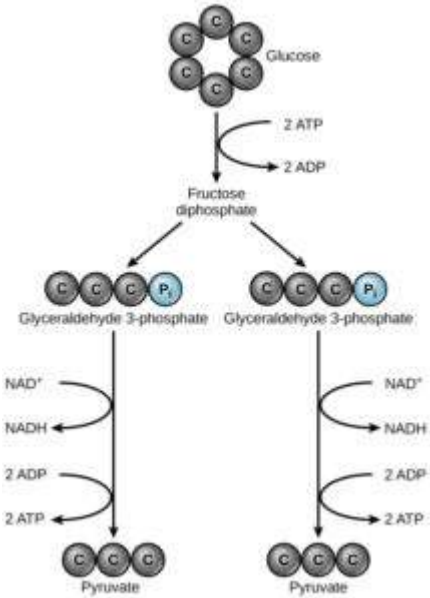


Figure 1 An overview of glycolysis. In glycolysis, a glucose molecule is converted into two pyruvate molecules.

Glycolysis consists of two distinct phases: energy-requiring, and energy-producing.

ENERGY-REQUIRING STEPS

The first part of the glycolysis pathway requires an input of energy to begin. The first step in glycolysis is catalyzed by **hexokinase,** an enzyme with broad specificity that catalyzes the phosphorylation of six-carbon sugars. Hexokinase **phosphorylates** (adds a phosphate to) glucose using ATP as the source of the phosphate (**Figure 2**). This produces glucose-6-phosphate, a more chemically reactive form of glucose. This phosphorylated glucose molecule can no longer leave the cell because the negatively charged phosphate will not allow it to cross the hydrophobic interior of the plasma membrane.

Several additional enzymatic reactions occur (**Figure 2**), one of which requires an additional ATP molecule. At the end of the energy-requiring steps, the original glucose has been split into two three-carbon molecules, and two ATPs have been used as sources of energy for this process.

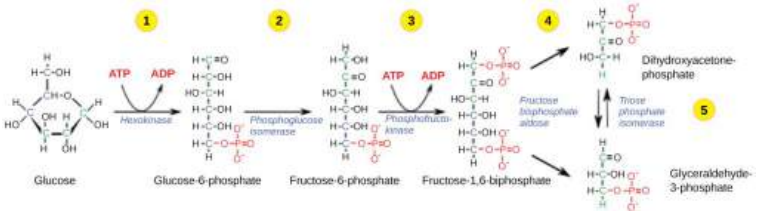


Figure 2 The first half of glycolysis uses two ATP molecules in the phosphorylation of glucose, which is then split into two three-carbon molecules.

ENERGY-PRODUCING STEPS

So far, glycolysis has cost the cell two ATP molecules and produced two small, three-carbon sugar molecules. Both of these molecules will proceed through the second half of the pathway, and sufficient energy will be extracted to pay back the two ATP molecules used as an initial investment and produce a profit for the cell of two additional ATP molecules and two even higher-energy **NADH** molecules (**Figure 3**).

During the energy-producing steps, additional enzymes continue to catalyze the breakdown of glucose (**Figure 3**). The end result of these reactions is two 3-carbon molecules of **pyruvate.**

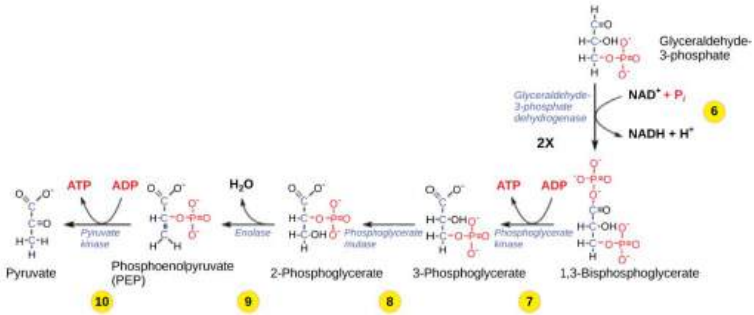


Figure 3 The second half of glycolysis involves phosphorylation without ATP investment (step 6) and produces two NADH and four ATP molecules per glucose.

An important rate-limiting step occurs at step 6 in glycolysis. If you look at Figure 3, you will notice that during step 6, **NAD^+** is converted into NADH. NADH contains more energy than NAD^+, and is therefore a desired product from this reaction. However, the continuation of the reaction depends upon the availability NAD^+. Thus, NADH must be continuously converted back into NAD^+ in order to keep this step going. If NAD^+ is not available, the second half of glycolysis slows down or stops.

If oxygen is available in the system, the NADH will be converted readily back into NAD^+ by the later processes in aerobic cellular respiration. However, if there is no oxygen available, NADH is not converted back into NAD^+. Without NAD^+, the reaction in step 6 cannot proceed and glycolysis slows or stops. In an environment without oxygen, an alternate pathway (fermentation) can provide the oxidation of NADH to NAD^+.

OUTCOMES OF GLYCOLYSIS

Glycolysis starts with glucose and ends with two pyruvate molecules, a total of four ATP molecules and two molecules

of NADH. Two ATP molecules were used in the first half of the pathway to prepare the six-carbon ring for cleavage, so the cell has a net gain of two ATP molecules and 2 NADH molecules for its use. If the cell cannot catabolize (break down) the pyruvate molecules further, it will harvest only two ATP molecules from one molecule of glucose. Mature mammalian red blood cells are not capable of aerobic respiration—the process in which organisms convert energy in the presence of oxygen—and glycolysis is their sole source of ATP. If glycolysis is interrupted, these cells lose their ability to maintain their sodium-potassium pumps, and eventually, they die.

SECTION SUMMARY

Glycolysis is the first pathway used in the breakdown of glucose to extract energy. It was probably one of the earliest metabolic pathways to evolve and is used by nearly all of the organisms on earth. Glycolysis consists of two parts: The first part prepares the six-carbon ring of glucose for cleavage into two three-carbon sugars. ATP is invested in the process during this half to energize the separation. The second half of glycolysis extracts ATP and high-energy electrons from hydrogen atoms and attaches them to NAD^+. Two ATP molecules are invested in the first half and four ATP molecules are formed by substrate phosphorylation during the second half. This produces a net gain of two ATP and two NADH molecules for the cell.

What was produced (per molecule of glucose)?

- 2 pyruvate (3 carbon molecules), 2 NADH, net gain of 2 ATP

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=170#h5p-55

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

OpenStax, Biology. OpenStax CNX. September 16, 2017 https://cnx.org/contents/GFy_h8cu@10.118:tYtpI6rX@6/Glycolysis

Aerobic Respiration, Part 2: Oxidation of Pyruvate and The Citric Acid Cycle

If oxygen is available, aerobic respiration will go forward. In eukaryotic cells, the pyruvate molecules produced at the end of glycolysis are transported into mitochondria (**Figure 1**), which are the sites of cellular respiration. In order for pyruvate, the product of glycolysis, to enter the next pathway, it must undergo several changes. The conversion is a three-step process.

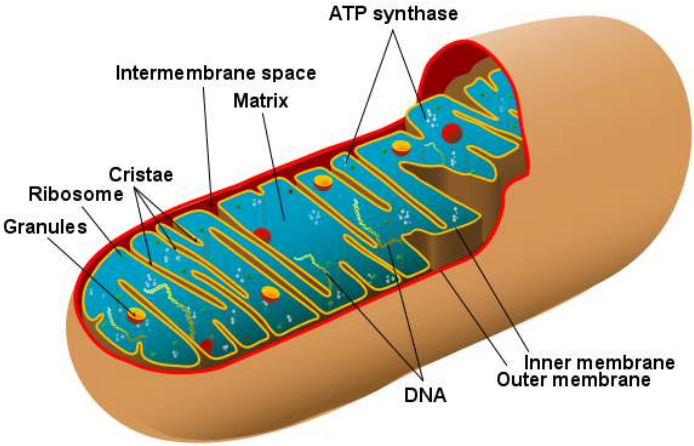


Figure 1 Diagram of a human mitochondrion. Recall that mitochondria have two membranes: an inner and an outer membrane. Between the two membranes is a region known as the intermembrane space. The mitochondrial matrix is located inside the inner membrane. Photo credit PsChemp, Wikimedia.

OXIDATION OF PYRUVATE

In eukaryotic cells, the pyruvate molecules produced at the end of glycolysis are transported into the **mitochondrial matrix** (the middle region of the mitochondria) (**Figure 1**). In the mitochondrial matrix, pyruvate will be transformed into a two-carbon acetyl group by removing a molecule of carbon dioxide. This also produces NADH. The acetyl group is picked up by a carrier compound called coenzyme A (CoA), which is made from vitamin B5. The resulting compound is called **acetyl CoA** (**Figure 2**). Acetyl CoA can be used in a variety of ways by the cell, but its major function is to deliver the acetyl

group derived from pyruvate to the next pathway in glucose catabolism.

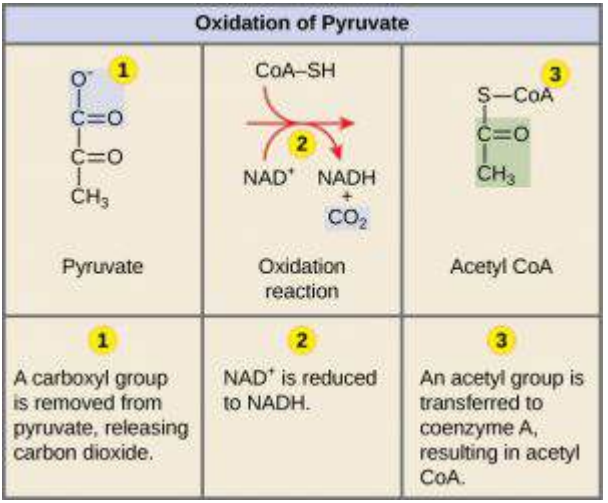


Figure 2 Upon entering the mitochondrial matrix, a multi-enzyme complex converts pyruvate into acetyl CoA. In the process, carbon dioxide is released and one molecule of NADH is formed.

ACETYL COA TO CO_2

In the presence of oxygen, acetyl CoA delivers its acetyl group to a four-carbon molecule, oxaloacetate, to form citrate, a six-carbon molecule with three carboxyl groups; this pathway will harvest the remainder of the extractable energy from what began as a glucose molecule. This single pathway is called by different names: the **citric acid cycle** (for the first intermediate formed—citric acid, or citrate—when acetate joins to the oxaloacetate), the **TCA cycle** (since citric acid or citrate and isocitrate are tricarboxylic acids), and the **Krebs cycle**, after Hans Krebs, who first identified the steps in the pathway in the 1930s in pigeon flight muscles.

Like the conversion of pyruvate to acetyl CoA, the **citric acid cycle** in eukaryotic cells also takes place in the matrix

of the mitochondria (**Figure 1**). Unlike glycolysis, the citric acid cycle is a closed loop: the last part of the pathway regenerates the compound used in the first step. The eight steps of the cycle are a series of chemical reactions that produces the following from *each* of the two molecules of pyruvate produced per molecule of glucose that originally went into glycolysis (**Figure 3**):

- 2 carbon dioxide molecules
- 1 ATP molecule (or an equivalent)
- 3 NADH and 1 $FADH_2$, which carry energy to the last part of the aerobic respiration pathway.

Part of this is considered an **aerobic** pathway (oxygen-requiring) because the NADH and $FADH_2$ produced must transfer their electrons to the next pathway in the system, which will use oxygen. If oxygen is not present, this transfer does not occur. *The citric acid cycle does NOT occur in anaerobic respiration.*

Two carbon atoms come into the citric acid cycle from each acetyl group. Two carbon dioxide molecules are released on each turn of the cycle; however, these do not contain the same carbon atoms contributed by the acetyl group on that turn of the pathway. The two acetyl-carbon atoms will eventually be released on later turns of the cycle; in this way, all six carbon atoms from the original glucose molecule will be eventually released as carbon dioxide. Carbon dioxide is a waste product in most animal cells and will be released outside the organism. It takes two turns of the cycle to process the equivalent of one glucose molecule. Each turn of the cycle forms three high-energy NADH molecules and one high-energy $FADH_2$ molecule. These high-energy carriers will connect with the last portion of aerobic respiration to produce ATP molecules. One ATP (or an equivalent) is also made in each cycle. Several of the intermediate compounds

in the citric acid cycle can be used in synthesizing non-essential amino acids; therefore, the cycle is both anabolic and catabolic.

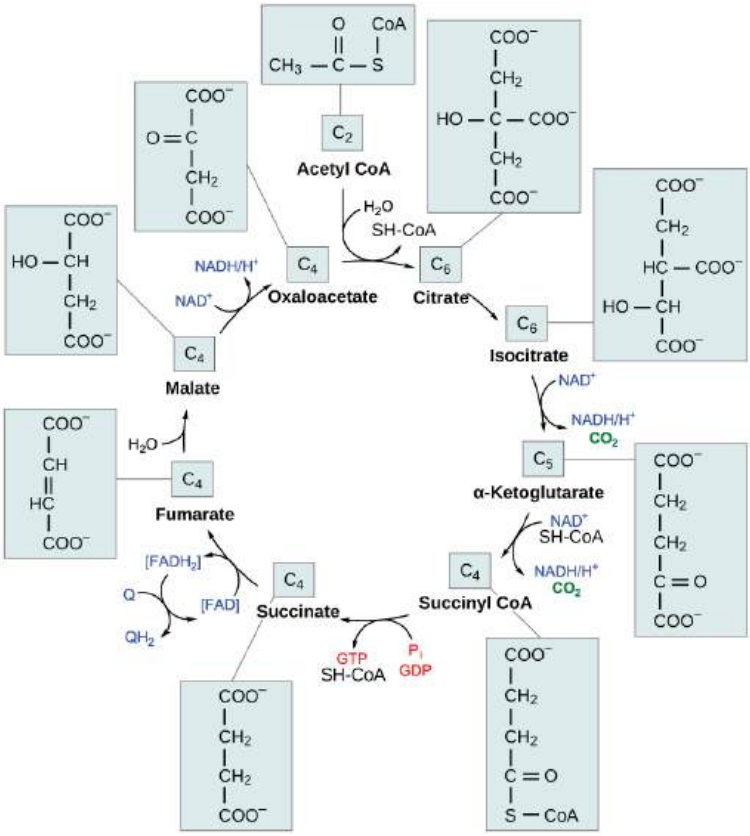


Figure 3 In the citric acid cycle, the acetyl group from acetyl CoA is attached to a four-carbon oxaloacetate molecule to form a six-carbon citrate molecule. Through a series of steps, citrate is oxidized, releasing two carbon dioxide molecules for each acetyl group fed into the cycle. In the process, three NAD+ molecules are reduced to NADH, one FAD molecule is reduced to FADH2, and one ATP or GTP (depending on the cell type) is produced (by substrate-level phosphorylation). Because the final product of the citric acid cycle is also the first reactant, the cycle runs continuously in the presence of sufficient reactants. (credit: modification of work by "Yikrazuul"/Wikimedia Commons)

SECTION SUMMARY

In the presence of oxygen, 3-carbon pyruvate is converted into a 2-carbon acetyl group, which is attached to a carrier molecule of coenzyme A. The resulting acetyl CoA can enter several pathways, but most often, the acetyl group is delivered to the citric acid cycle for further catabolism (breakdown). During the conversion of pyruvate into the acetyl group, a molecule of carbon dioxide and two high-energy electrons are removed. Because two pyruvate were produced from each molecule of glucose during glycolysis, the production of two carbon dioxide molecules (which are released as waste) accounts for two of the six carbons of the original glucose molecule. The other four carbons are released as carbon dioxide during two turns of the citric acid cycle. The electrons are picked up by NAD^+, and the NADH carries the electrons to a later pathway for ATP production. At this point, the glucose molecule that originally entered cellular respiration has been completely broken down. Chemical potential energy stored within the glucose molecule has been transferred to electron carriers or has been used to synthesize a few ATPs.

What was produced (per molecule of glucose)?

- Oxidation of pyruvate: 2 CO_2, 2 NADH, 2 acetyl (2 carbon molecule)
- Products of the citric acid cycle: 4 CO_2, 6 NADH, 2 $FADH_2$, 2 ATP

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=172#h5p-56

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

OpenStax, Biology. OpenStax CNX. September 16, 2017. https://cnx.org/contents/GFy_h8cu@10.118:weAHBat1@7/Oxidation-of-Pyruvate-and-the-

Aerobic Respiration, Part 3: Oxidative Phosphorylation

You have just read about two pathways in glucose catabolism—glycolysis and the citric acid cycle—that generate ATP. Most of the ATP generated during the aerobic catabolism of glucose, however, is not generated directly from these pathways. Rather, it derives from a process that begins with passing electrons through a series of chemical reactions to a final electron acceptor, oxygen. This is the only place in aerobic respiration where O_2 is actually required. These reactions take place in specialized protein complexes located in the inner membrane of the mitochondria of eukaryotic organisms and on the inner part of the cell membrane of prokaryotic organisms. The energy of the electrons is used to generate ATP. The entirety of this process is called **oxidative phosphorylation**.

During oxidative phosphorylation:

- The energy from NADH and $FADH_2$ is used up.
- Oxygen gas is converted into water.
- 30-36 ATP are recharged from ADP

ELECTRON TRANSPORT CHAIN

The electron transport chain (**Figure 1**) is the last component of aerobic respiration and is the only part of metabolism that uses atmospheric oxygen. Oxygen continuously diffuses into plants for this purpose. In animals, oxygen enters the body through the respiratory system. Electron transport is a series of chemical reactions that resembles a bucket brigade in that electrons are passed rapidly from one component to the next, to the endpoint of the chain where oxygen is the final electron acceptor and water is produced. There are four complexes composed of proteins, labeled I through IV in **Figure 1**, and the aggregation of these four complexes, together with associated mobile, accessory electron carriers, is called the **electron transport chain**. The electron transport chain is present in multiple copies in the **inner mitochondrial membrane** of eukaryotes and in the plasma membrane of prokaryotes. In each transfer of an electron through the electron transport chain, the electron loses energy, but with some transfers, the energy is stored as potential energy by using it to pump hydrogen ions (H^+, protons) across the inner mitochondrial membrane into the **intermembrane space**, creating an **electrochemical gradient**. An electrochemical gradient consists of two parts: a difference in solute concentration across the membrane combined with a difference in charge across the membrane. Here, the electrochemical gradient is made up of a higher concentration of H+ in the inner membrane space compared to the mitochondrial matrix.

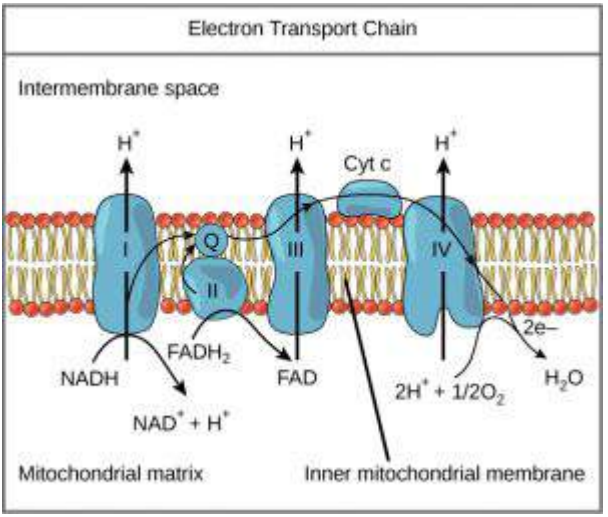


Figure 1 The electron transport chain is a series of electron transporters embedded in the inner mitochondrial membrane that shuttles electrons from NADH and $FADH_2$ to molecular oxygen. In the process, protons are pumped from the mitochondrial matrix to the intermembrane space, and oxygen is reduced to form water.

Electrons from NADH and $FADH_2$ are passed to protein complexes in the electron transport chain. As they are passed from one complex to another (there are a total of four), the electrons lose energy, and some of that energy is used to pump hydrogen ions from the mitochondrial matrix into the intermembrane space. In the fourth protein complex, the electrons are accepted by oxygen, the terminal acceptor. The oxygen with its extra electrons then combines with two hydrogen ions, further enhancing the electrochemical gradient, to form water. If there were no oxygen present in the mitochondrion, the electrons could not be removed from the system, and the entire electron transport chain would back up and stop. The mitochondria would be unable to generate new ATP in this way, and the cell would ultimately die from lack of energy. This is the

reason we must breathe to draw in new oxygen. This is the only place where oxygen is required during the processes of aerobic respiration.

In the electron transport chain, the free energy from the series of reactions just described is used to pump hydrogen ions across the membrane. The uneven distribution of H+ ions across the membrane establishes an electrochemical gradient, owing to the H+ ions' positive charge and their higher concentration on one side of the membrane.

Hydrogen ions diffuse from the intermembrane space through the inner membrane into the mitochondrial matrix through an integral membrane protein called **ATP synthase** (**Figure 2**). This complex protein acts as a tiny generator, turned by the force of the hydrogen ions diffusing through it, down their electrochemical gradient from the intermembrane space, where there are many mutually repelling hydrogen ions to the matrix, where there are few. The turning of the parts of this molecular machine regenerate ATP from ADP. This flow of hydrogen ions across the membrane through ATP synthase is called **chemiosmosis**.

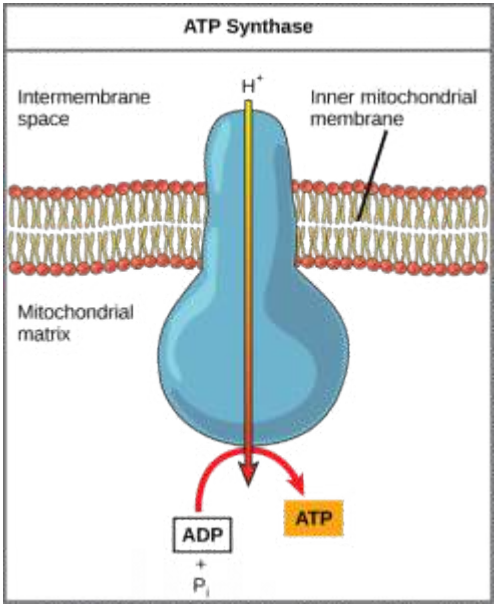


Figure 2 ATP synthase is a complex, molecular machine that uses a proton (H+) gradient to form ATP from ADP and inorganic phosphate (Pi). (Credit: modification of work by Klaus Hoffmeier)

Chemiosmosis (**Figure 2**) is used to generate 90 percent of the ATP made during aerobic glucose catabolism. The result of the reactions is the production of ATP from the energy of the electrons removed from hydrogen atoms. These atoms were originally part of a glucose molecule. At the end of the electron transport system, the electrons are used to reduce an oxygen molecule to oxygen ions. The extra electrons on the oxygen ions attract hydrogen ions (protons) from the surrounding medium, and water is formed. The electron transport chain and the production of ATP through chemiosmosis are collectively called **oxidative phosphorylation (Figure 3)**.

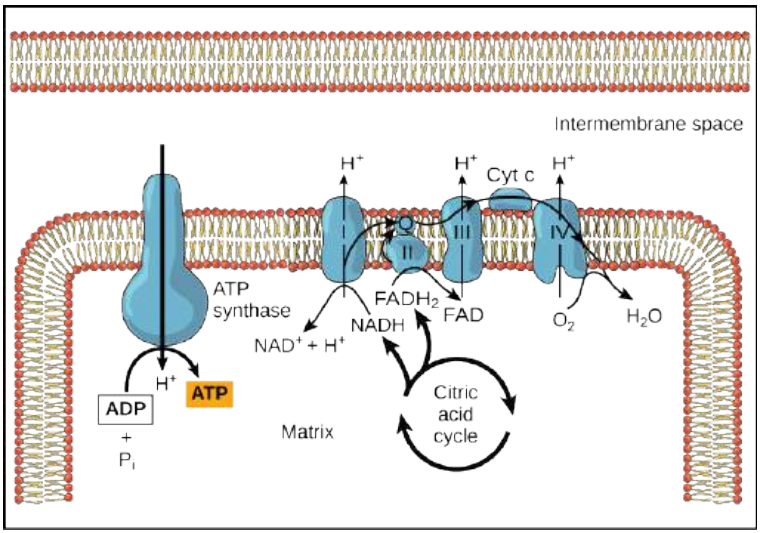


Figure 3 In oxidative phosphorylation, the pH gradient formed by the electron transport chain is used by ATP synthase to form ATP.

ATP YIELD

The number of ATP molecules generated from the catabolism of glucose varies. For example, the number of hydrogen ions that the electron transport chain complexes can pump through the membrane varies between species. Another source of variance stems from the shuttle of electrons across the membranes of the mitochondria because the NADH generated from glycolysis cannot easily enter mitochondria. Thus, electrons are picked up on the inside of mitochondria by either NAD^+ or FAD^+. As you have learned earlier, these FAD^+molecules can transport fewer ions; consequently, fewer ATP molecules are generated when FAD^+ acts as a carrier. NAD^+ is used as the electron transporter in the liver and FAD^+ acts in the brain.

Another factor that affects the yield of ATP molecules

generated from glucose is the fact that intermediate compounds in these pathways are used for other purposes. Glucose catabolism connects with the pathways that build or break down all other biochemical compounds in cells, and the result is somewhat messier than the ideal situations described thus far. For example, sugars other than glucose are fed into the glycolytic pathway for energy extraction. Moreover, the five-carbon sugars that form nucleic acids are made from intermediates in glycolysis. Certain nonessential amino acids can be made from intermediates of both glycolysis and the citric acid cycle. Lipids, such as cholesterol and triglycerides, are also made from intermediates in these pathways, and both amino acids and triglycerides are broken down for energy through these pathways. Overall, in living systems, these pathways of glucose catabolism extract about 34 percent of the energy contained in glucose.

SECTION SUMMARY

The electron transport chain is the portion of aerobic respiration that uses free oxygen as the final electron acceptor of the electrons removed from the intermediate compounds in glucose catabolism. The electron transport chain is composed of four large, multiprotein complexes embedded in the inner mitochondrial membrane and two small diffusible electron carriers shuttling electrons between them. The electrons are passed through a series of reactions, with a small amount of free energy used at three points to transport hydrogen ions across a membrane. This process contributes to the gradient used in chemiosmosis. The electrons passing through the electron transport chain gradually lose energy until eventually they are donated to oxygen gas which accepts two protons (H+) and is converted into water. The end products of the electron transport chain

are water and roughly 30-34 molecules of ATP. A number of intermediate compounds of the citric acid cycle can be diverted into the anabolism of other biochemical molecules, such as nonessential amino acids, sugars, and lipids. These same molecules can serve as energy sources for the glucose pathways.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=174#h5p-57

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=174#h5p-58

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

OpenStax, Biology. OpenStax CNX. September 16, 2017 https://cnx.org/contents/GFy_h8cu@10.118:7oTVAgrZ@7/Oxidative-Phosphorylation

Metabolism without Oxygen: Fermentation

In aerobic respiration, the final electron acceptor for the electron transport chain is an oxygen molecule, O_2. If aerobic respiration occurs, then approximately 30 molecules of ATP will be produced during the electron transport chain and chemiosmosis using the energy of the high-energy electrons carried by NADH or $FADH_2$ to the electron transport chain. When NADH or $FADH_2$ give their high energy electrons to the electron transport chain, NAD^+ and FAD are regenerated. These low energy molecules cycle back to glycolysis and/or the citric acid cycle, where they pick up more high energy electrons and allow the process to continue.

Glycolysis and the citric acid cycle can not occur if there is not NAD^+ present to pick up electrons as the reactions proceed. When oxygen is present, this isn't a problem – all of the NADH and $FADH_2$ that were produced during glycolysis and the citric acid cycle are converted back into NAD^+ and FAD after the electron transport chain. When no oxygen is present, the electron transport chain can't run because there is no oxygen to act as the final electron acceptor. This means that the ETC will not be accepting electrons from NADH as its source of power, so NAD^+ will not be regenerated. Both glycolysis and the citric acid cycle require NAD^+ to accept electrons during their chemical reactions. In order for the cell to continue to

generate *any* ATP, NADH must be converted back to NAD^+ for use as an electron carrier. Anaerobic processes use different mechanisms, but all function to convert NAD^+ back into NADH.

How is this done?

- Processes that use an organic molecule to regenerate NAD^+ from NADH are collectively referred to as **fermentation**.
- In contrast, some living systems use an inorganic molecule (such as nitrate or sulfur) to regenerate NAD^+.

Both of these methods are called **anaerobic cellular respiration**. They do not require oxygen to achieve NAD^+ regeneration and enable organisms to convert energy for their use in the absence of oxygen.

During anaerobic respiration, only glycolysis occurs. The 2 molecules of NADH that are generated during glycolysis are then converted back into NAD+ during anaerobic respiration so that glycolysis can continue. Since glycolysis only produces 2 ATP, anaerobic respiration is much less efficient than aerobic respiration (2 ATP molecules compared to 36-ish ATP molecules). However, 2 ATP molecules is much better for a cell than 0 ATP molecules. In anaerobic situations, the cell needs to continue performing glycolysis to generate 2 ATP per glucose because if a cell is not generating any ATP, it will die.

Note that the only part of aerobic respiration that physically uses oxygen is the electron transport chain. However, the citric acid cycle can not occur in the absence of oxygen because there is no way to regenerate the NAD+ used during this process.

LACTIC ACID FERMENTATION

The fermentation method used by animals and some bacteria like those in yogurt is lactic acid fermentation (**Figure 1**). This occurs routinely in mammalian red blood cells and in skeletal muscle that does not have enough oxygen to allow aerobic respiration to continue (such as in muscles after hard exercise). The chemical reaction of lactic acid fermentation is the following:

Pyruvic acid + NADH ↔ lactic acid + NAD^+

The build-up of lactic acid causes muscle stiffness and fatigue. In muscles, lactic acid produced by fermentation must be removed by the blood circulation and brought to the liver for further metabolism. Once the lactic acid has been removed from the muscle and is circulated to the liver, it can be converted back to pyruvic acid and further catabolized (broken down) for energy.

Note that the purpose of this process is not to produce lactic acid (which is a waste product and is excreted from the body). The purpose is to convert NADH back into NAD^+ so that glycolysis can continue so that the cell can produce 2 ATP per glucose.

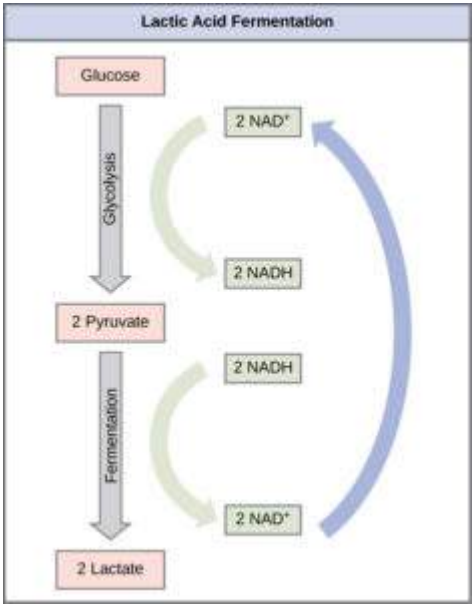


Figure 1 Lactic acid fermentation is common in muscles that have become exhausted by use.

ALCOHOL FERMENTATION

Another familiar fermentation process is alcohol fermentation (**Figure 2**), which produces ethanol, an alcohol. The alcohol fermentation reaction is the following:

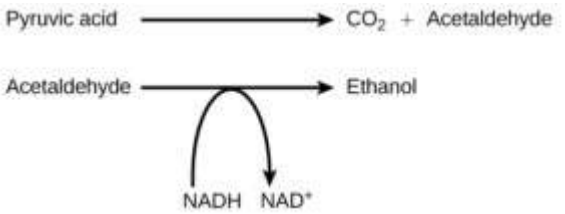


Figure 2 The reaction resulting in alcohol fermentation is shown.

The fermentation of pyruvic acid by yeast produces the

ethanol found in alcoholic beverages (**Figure 3**). If the carbon dioxide produced by the reaction is not vented from the fermentation chamber, for example in beer and sparkling wines, it remains dissolved in the medium until the pressure is released. Ethanol above 12 percent is toxic to yeast, so natural levels of alcohol in wine occur at a maximum of 12 percent.

Figure 3 Fermentation of grape juice to make wine produces CO_2 as a byproduct. Fermentation tanks have valves so that pressure inside the tanks can be released.

Again, the purpose of this process is not to produce ethanol, but rather to convert NADH back into NAD^+ so that glycolysis can continue.

An interactive H5P element has been excluded from this version of the

text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=178#h5p-59

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=178#h5p-61

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Metabolism of molecules other than glucose

You have learned about the catabolism of glucose, which provides energy to living cells. But living things consume more than just glucose for food. How does a turkey sandwich, which contains various carbohydrates, lipids, and protein, provide energy to your cells?

Basically, all of these molecules from food are converted into molecules that can enter the cellular respiration pathway somewhere. Some molecules enter at glycolysis, while others enter at the citric acid cycle. This means that all of the catabolic pathways for carbohydrates, proteins, and lipids eventually connect into glycolysis and the citric acid cycle pathways. Metabolic pathways should be thought of as porous—that is, substances enter from other pathways, and other substances leave for other pathways. These pathways are not closed systems. Many of the products in a particular pathway are reactants in other pathways.

CARBOHYDRATES

So far, we have discussed the carbohydrate from which organisms derive the majority of their energy: glucose. Many carbohydrate molecules can be broken down into

glucose or otherwise processed into glucose by the body. **Glycogen,** a polymer of glucose, is a short-term energy storage molecule in animals (**Figure 1**). When there is plenty of ATP present, the extra glucose is converted into glycogen for storage. Glycogen is made and stored in the liver and muscle. Glycogen will be taken out of storage if blood sugar levels drop. The presence of glycogen in muscle cells as a source of glucose allows ATP to be produced for a longer time during exercise.

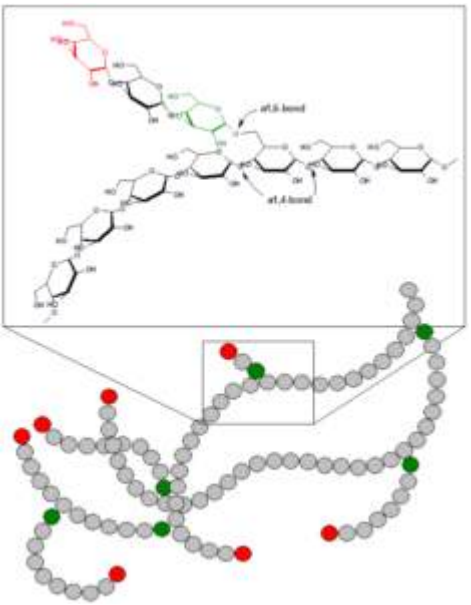


Figure 1 Glycogen is made of many molecules of glucose attached together into branching chains. Each of the balls in the bottom diagram represents one molecule of glucose. (Credit: Glycogen by BorisTM. This work has been released into the public domain)

Most other carbohydrates enter the cellular respiration pathway during glycolysis. For example, **sucrose** is a disaccharide made from glucose and fructose bonded together. Sucrose is broken down in the small intestine. The glucose enters the beginning of glycolysis as previously

discussed, while fructose can be slightly modified and enter glycolysis at the third step. **Lactose,** the disaccharide sugar found in milk, can be broken down by lactase enzyme into two smaller sugars: galactose and glucose. Like fructose, galactose can be slightly modified to enter glycolysis.

Because these carbohydrates enter near the beginning of glycolysis, their catabolism (breakdown) produces the same number of ATP molecules as glucose.

PROTEINS

Proteins are broken down by a variety of enzymes in cells. Most of the time, amino acids are recycled into new proteins and not used as a source of energy. This is because it is more energy efficient to reuse amino acids rather than making new ones from scratch. The body will use protein as a source of energy if:

- There are excess amino acids (you consume a lot of protein)
- The body is in a state of famine (you are starving and have no other source of energy available)

When proteins are used in the cellular respiration pathway, they are first broken down into individual amino acids. The amino group from each amino acid is removed (deaminated) and is converted into ammonia. In mammals, the liver synthesizes urea from two ammonia molecules and a carbon dioxide molecule. Thus, urea is the principal waste product in mammals from the nitrogen originating in amino acids, and it leaves the body in urine.

Once the amino acid has been deaminated, its chemical properties determine which intermediate of the cellular respiration pathway it will be converted into. These

intermediates enter cellular respiration at various places in the Citric Acid Cycle (**Figure 2**).

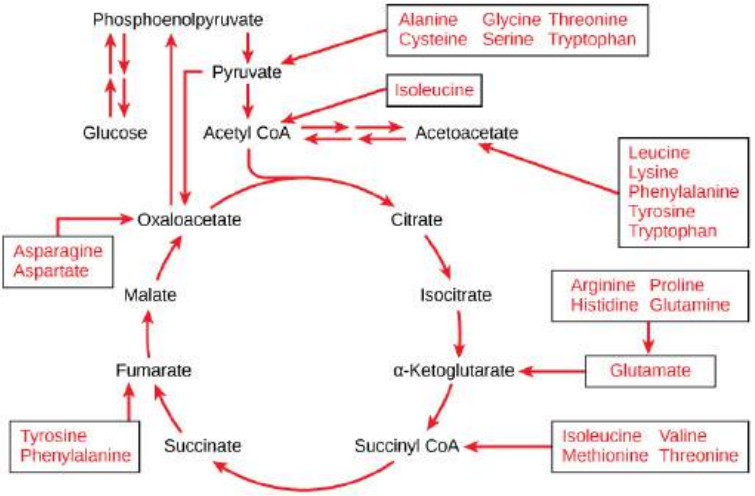


Figure 2 The carbon skeletons of certain amino acids (indicated in boxes) derived from proteins can feed into the citric acid cycle. (credit: modification of work by Mikael Häggström)

LIPIDS

Triglycerides (fats) are a form of long-term energy storage in animals. Triglycerides store about twice as much energy as carbohydrates. Triglycerides are made of glycerol and three fatty acids. Glycerol can enter glycolysis. Fatty acids are broken into two-carbon units that enter the citric acid cycle (**Figure 3**).

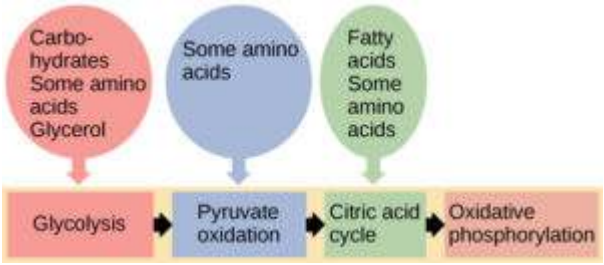


Figure 3 Glycogen from the liver and muscles, together with fats, can feed into the catabolic pathways for carbohydrates.

Remember that if oxygen is not available, glycolysis can occur but not the citric acid cycle or oxidative phosphorylation. Since fatty acids enter the pathway at the citric acid cycle, they can not be broken down in the absence of oxygen. This means that if cells are not performing aerobic cellular respiration, the body can not burn fat for energy. This is why posters about the "Fat Burning Zone" in a gym specify that you need to have a lower heart rate / breathing rate to burn more fat – cells that are not doing aerobic respiration can't burn fat for fuel!

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=181#h5p-60

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Anaerobic Cellular Respiration in Prokaryotes

Certain prokaryotes, including some species of bacteria and Archaea, use anaerobic respiration. For example, the group of Archaea called methanogens reduces carbon dioxide to methane to oxidize NADH. These microorganisms are found in soil and in the digestive tracts of ruminants, such as cows and sheep. Similarly, sulfate-reducing bacteria and Archaea, most of which are anaerobic (**Figure 8**), reduce sulfate to hydrogen sulfide to regenerate NAD^+ from NADH.

Figure 8 The green color seen in these coastal waters is from an eruption of hydrogen sulfide. Anaerobic, sulfate-reducing bacteria release hydrogen sulfide gas as they decompose algae in the water. (credit: NASA image courtesy Jeff Schmaltz, MODIS Land Rapid Response Team at NASA GSFC)

Other fermentation methods occur in bacteria. Many prokaryotes are facultatively anaerobic. This means that they can switch between aerobic respiration and fermentation, depending on the availability of oxygen. Certain prokaryotes, like *Clostridia* bacteria, are obligate anaerobes. Obligate anaerobes live and grow in the absence of molecular oxygen. Oxygen is a poison to these microorganisms and kills them upon exposure. It should be noted that all forms of fermentation, except lactic acid fermentation, produce gas. The production of particular types of gas is used as an indicator of the fermentation of specific carbohydrates, which plays a role in the laboratory identification of the bacteria. The various methods of fermentation are used by different organisms to ensure an adequate supply of NAD^+ for the sixth step in glycolysis. Without these pathways, that step would not occur, and no ATP would be harvested from the breakdown of glucose.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

PHOTOSYNTHESIS

Learning Objectives

By the end of this section, you will be able to:

- Compare energy-generating processes within different types of cells

All living organisms on earth consist of one or more cells. Each cell runs on the chemical energy found mainly in carbohydrate molecules (food), and the majority of these molecules are produced by one process: **photosynthesis.** Through photosynthesis, certain organisms convert solar energy (sunlight) into chemical energy, which is then used to build carbohydrate molecules. The energy used to hold these molecules together is released when an organism breaks down food. Cells then use this energy to perform work, such as cellular respiration.

The energy that is harnessed from photosynthesis enters the ecosystems of our planet continuously and is transferred from one organism to another. Therefore, directly or

indirectly, the process of photosynthesis provides most of the energy required by living things on earth.

Photosynthesis also results in the release of oxygen into the atmosphere. In short, to eat and breathe, **humans depend almost entirely on the organisms that carry out photosynthesis.**

Putting photosynthesis into context

All living things require energy. Carbohydrates are storage molecules for energy. Living things access energy by breaking down carbohydrate molecules during the process of cellular respiration. Plants produce carbohydrates during photosynthesis. So if plants make carbohydrate molecules during photosynthesis, do they also perform cellular respiration? The answer is yes, they do. Although energy can be stored in molecules like ATP, carbohydrates (and lipids, which can also enter cellular respiration as a source of energy) are much more stable and efficient reservoirs for chemical energy. Photosynthetic organisms also carry out the reactions of respiration to harvest the energy that they have stored in carbohydrates during photosynthesis. Plants have mitochondria in addition to chloroplasts.

The overall reaction for photosynthesis:

$$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$$

is the reverse of the overall reaction for cellular respiration:

$$6O_2 + C_6H_{12}O_6 \rightarrow 6CO_2 + 6H_2O$$

Photosynthesis produces oxygen as a byproduct, and respiration produces carbon dioxide as a byproduct. In nature, there is no such thing as waste. Every single atom of matter is conserved, recycling indefinitely. Substances change form or move from one type of molecule to another, but never disappear (**Figure 1**).

CO_2 is no more a form of waste produced by respiration than oxygen is a waste product of photosynthesis. Both are byproducts of reactions that move on to other reactions. Photosynthesis absorbs energy from sunlight to build carbohydrates in the chloroplasts, and aerobic cellular respiration releases that stored energy by using oxygen to break down carbohydrates. Both organelles use electron transport chains to generate the energy necessary to drive other reactions. Photosynthesis and cellular respiration function in a biological cycle, allowing organisms to access life-sustaining energy that originates millions of miles away in a star.

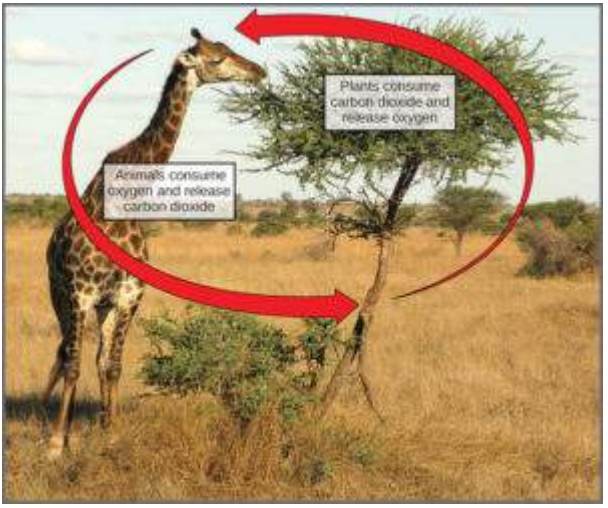


Figure 1 *In the carbon cycle, the reactions of photosynthesis and cellular respiration share reciprocal reactants and products. (credit: modification of work by Stuart Bassil)*

There are two basic parts of photosynthesis: the light dependent reactions and the light independent reactions (also known as the Calvin cycle). During the light reactions, the energy from sunlight is stored in energy carrier molecules. These energy carrier molecules are then used to power the reactions of the Calvin cycle, where CO_2 molecules are joined together to produce carbohydrates such as glucose.

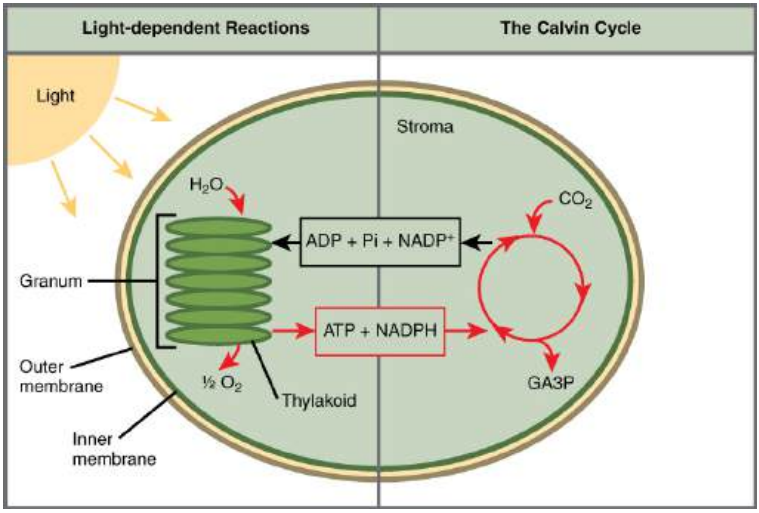


Figure 2 Overview of the process of photosynthesis. During the light-dependent reactions, the energy from sunlight is used by the chloroplast to create energy molecules: ATP and NADPH. These energy molecules power the Calvin cycle, which creates carbohydrates (G3P) from CO2.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=205#h5p-62

REFERENCES

Text adapted from: OpenStax, Biology. OpenStax CNX. November 11, 2017. https://cnx.org/contents/GFy_h8cu@10.118:W7ctJeSI@8/Overview-of-Photosynthesis

The structure of the chloroplast

In plants, photosynthesis takes place primarily in leaves, which consist of many layers of cells and have differentiated top and bottom sides. The process of photosynthesis occurs not on the surface layers of the leaf, but rather in a middle layer called the mesophyll (**Figure 1**).

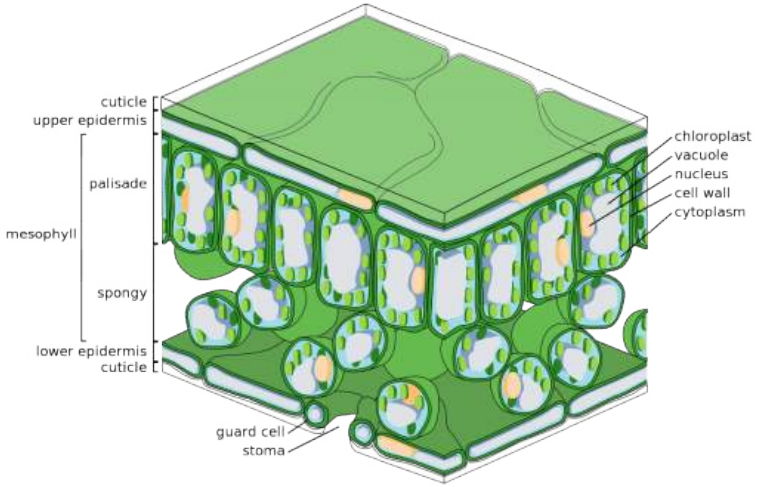


Figure 1 Not all cells of a leaf carry out photosynthesis. Cells within the middle layer of a leaf have chloroplasts, which contain the photosynthetic apparatus. (credit Zephyris; wikimedia)

The gas exchange of carbon dioxide and oxygen occurs through small, regulated openings called **stomata**.

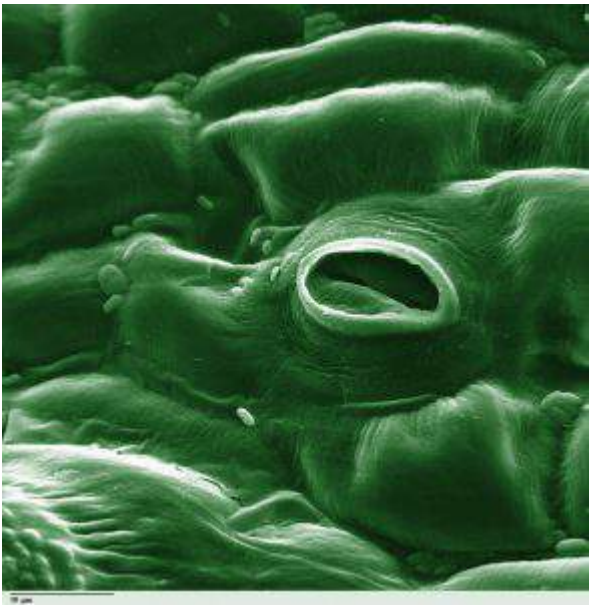


Figure 2 Tomato leaf stomate (singular of stomata). Photo credit: Photohound; Wikimedia; Public Domain.

In eukaryotes, photosynthesis takes place inside an organelle called a **chloroplast**. Some prokaryotes can perform photosynthesis, but they do not contain chloroplasts (or other membrane-bound organelles). In plants, chloroplast-containing cells exist in the mesophyll. Chloroplasts are surrounded by a double membrane similar to the double membrane found within a mitochondrion. Within the chloroplast is a third membrane that forms stacked, disc-shaped structures called **thylakoids**. Embedded in the thylakoid membrane are molecules of **chlorophyll**, a pigment (a molecule that absorbs light) through which the entire process of photosynthesis begins. Chlorophyll is responsible for the green color of plants. The thylakoid membrane encloses an internal space called the thylakoid

lumen or space. Other types of pigments are also involved in photosynthesis, but chlorophyll is by far the most important. As shown in **Figure 3**, a stack of thylakoids is called a **granum**, and the space surrounding the granum is called **stroma** (not to be confused with stomata, the openings on the leaves).

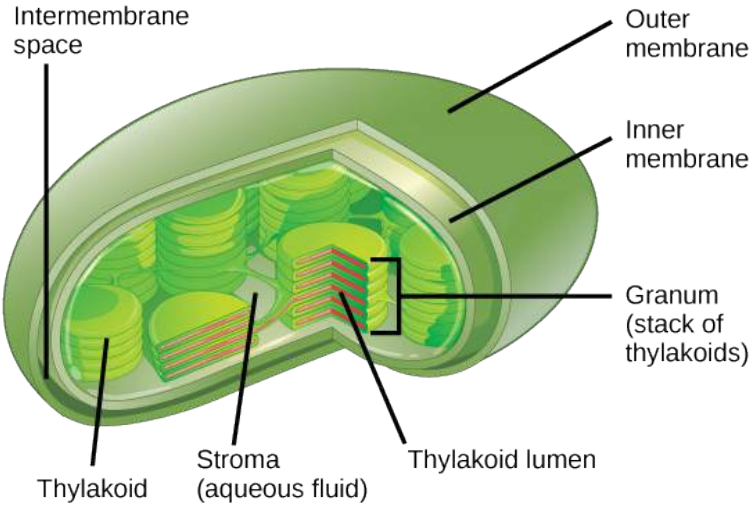


Figure 3 structure of the chloroplast. Note that the chloroplast is surrounded by a double membrane, but also contains a third set of membranes, which enclose the thylakoids.

Just like the structure of the mitochondria was important to its ability to perform aerobic cellular respiration, the structure of the chloroplast allows the process of photosynthesis to take place. Both the light-dependent reactions and the Calvin cycle take place inside of the chloroplast.

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Light and Pigments

How can light be used to make food? It is easy to think of light as something that exists and allows living organisms, such as humans, to see, but light is a form of energy. Like all energy, light can travel, change form, and be harnessed to do work. In the case of photosynthesis, light energy is transformed into chemical energy, which autotrophs use to build carbohydrate molecules. However, autotrophs only use a specific component of sunlight (**Figure 1**).

Figure 1 Autotrophs can capture light energy from the sun, converting it into chemical energy used to build food molecules. (credit: modification of work by Gerry Atwell, U.S. Fish and Wildlife Service)

WHAT IS LIGHT ENERGY?

The sun emits an enormous amount of electromagnetic radiation (solar energy). Humans can see only a fraction of this energy, which is referred to as "visible light." The manner in which solar energy travels can be described and measured as waves. Scientists can determine the amount of energy of a wave by measuring its **wavelength**, the distance between two consecutive, similar points in a series of waves, such as from crest to crest or trough to trough (**Figure 2**).

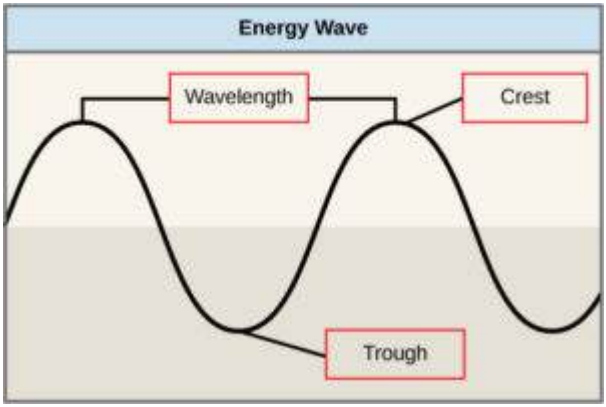


Figure 2 The wavelength of a single wave is the distance between two consecutive points along the wave.

Visible light constitutes only one of many types of electromagnetic radiation emitted from the sun. The **electromagnetic spectrum** is the range of all possible wavelengths of radiation (**Figure 3**). Each wavelength corresponds to a different amount of energy carried.

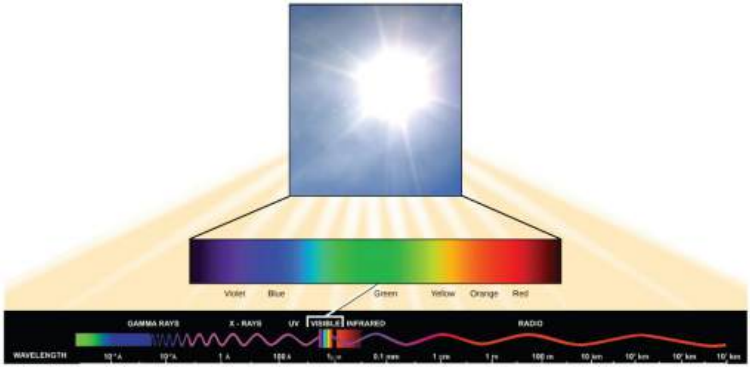


Figure 3 The sun emits energy in the form of electromagnetic radiation. This radiation exists in different wavelengths, each of which has its own characteristic energy. Visible light is one type of energy emitted from the sun.

Each type of electromagnetic radiation has a characteristic

range of wavelengths. The longer the wavelength (or the more stretched out it appears), the less energy is carried. Short, tight waves carry the most energy. This may seem illogical, but think of it in terms of a piece of moving rope. It takes little effort by a person to move a rope in long, wide waves. To make a rope move in short, tight waves, a person would need to apply significantly more energy.

The sun emits a broad range of electromagnetic radiation, including X-rays and ultraviolet (UV) rays (**Figure 3**) . The higher-energy waves are dangerous to living things; for example, X-rays and UV rays can be harmful to humans.

ABSORPTION OF LIGHT

Light energy enters the process of photosynthesis when pigments absorb the light. In plants, pigment molecules absorb only visible light for photosynthesis. The visible light seen by humans as white light actually exists in a rainbow of colors. Certain objects, such as a prism or a drop of water, disperse white light to reveal these colors to the human eye. The visible light portion of the electromagnetic spectrum is perceived by the human eye as a rainbow of colors, with violet and blue having shorter wavelengths and, therefore, higher energy. At the other end of the spectrum toward red, the wavelengths are longer and have lower energy.

The wavelengths of light that are reflected from an object and bounce off are detected by our eyes. The wavelengths of light that are absorbed by an object do not make it to our eyes. This means that the color an object appears is due to the wavelengths that are reflected and not those that are absorbed. For example, the apple in **Figure 4** appears red (assuming you are not color-blind). This is because the red wavelengths of light are reflected off the apple and the other

wavelengths (yellow, green, blue, purple) are absorbed by the apple.

Figure 4 This apple appears red because it is reflecting the red wavelengths of light. Other wavelengths are absorbed by the apple.

UNDERSTANDING PIGMENTS

Different kinds of pigments exist, and each absorbs only certain wavelengths (colors) of visible light. Pigments reflect the color of the wavelengths that they cannot absorb. All photosynthetic organisms contain a pigment called **chlorophyll** *a*, which humans see as the common green color associated with plants. Chlorophyll *a* absorbs wavelengths from either end of the visible spectrum (blue and red), but not from green. Because green is reflected, chlorophyll appears green.

Other pigment types include **chlorophyll** *b* (which absorbs

blue and red-orange light) and the carotenoids. Each type of pigment can be identified by the specific pattern of wavelengths it absorbs from visible light, which is its absorption spectrum.

Many photosynthetic organisms have a mixture of pigments; between them, the organism can absorb energy from a wider range of visible-light wavelengths. Not all photosynthetic organisms have full access to sunlight. Some organisms grow underwater where light intensity decreases with depth, and certain wavelengths are absorbed by the water. Other organisms grow in competition for light. Plants on the rainforest floor must be able to absorb any bit of light that comes through, because the taller trees block most of the sunlight (**Figure 5**).

Figure 5 Plants that commonly grow in the shade benefit from having a variety of light-absorbing pigments. Each pigment can absorb different wavelengths of light, which allows the plant to absorb any light that passes through the taller trees. (credit: Jason Hollinger)

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=576#h5p-64

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=576#h5p-63

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Light-Dependent Reactions

Photosynthesis takes place in two stages: the light-dependent reactions and the Calvin cycle. In the **light-dependent reactions**, which take place at the thylakoid membrane, chlorophyll absorbs energy from sunlight and then converts it into chemical energy with the use of water. The light-dependent reactions release oxygen as a byproduct as water is broken apart. In the Calvin cycle, which takes place in the stroma, the chemical energy derived from the light-dependent reactions drives both the capture of carbon in carbon dioxide molecules and the subsequent assembly of sugar molecules.

The two reactions use carrier molecules to transport the energy from one to the other. The carriers that move energy from the light-dependent reactions to the Calvin cycle reactions can be thought of as "full" because they bring energy. After the energy is released, the "empty" energy carriers return to the light-dependent reactions to obtain more energy. You should be familiar with the energy carrier molecules used during cellular respiration: NADH and $FADH_2$. Photosynthesis uses a different energy carrier, **NADPH,** but it functions in a comparable way. The lower energy form, **NADP+**, picks up a high energy electron and a

proton and is converted to NADPH. When NADPH gives up its electron, it is converted back to NADP+.

HOW THE LIGHT-DEPENDENT REACTIONS WORK

The overall purpose of the light-dependent reactions is to convert solar energy into chemical energy in the form of NADPH and ATP. This chemical energy will be used by the Calvin cycle to fuel the assembly of sugar molecules.

The light-dependent reactions begin in a grouping of pigment molecules and proteins called a **photosystem**. There are two photosystems (Photosystem I and II), which exist in the membranes of thylakoids. Both photosystems have the same basic structure: a number of antenna proteins to which chlorophyll molecules are bound surround the reaction center where the photochemistry takes place. Each photosystem is serviced by the light-harvesting complex, which passes energy from sunlight to the reaction center. It consists of multiple antenna proteins that contain a mixture of 300–400 chlorophyll *a* and *b* molecules as well as other pigments like carotenoids. A photon of light energy travels until it reaches a molecule of chlorophyll pigment. The photon causes an electron in the chlorophyll to become "excited." The energy given to the electron allows it to break free from an atom of the chlorophyll molecule. Chlorophyll is therefore said to "donate" an electron (**Figure 1**).The absorption of a single **photon** or distinct quantity or "packet" of light by any of the chlorophylls pushes that molecule into an excited state. In short, the light energy has now been captured by biological molecules but is not stored in any useful form yet. The energy is transferred from chlorophyll to chlorophyll until eventually (after about a millionth of a second), it is delivered to the reaction center. Up to this point,

only energy has been transferred between molecules, not electrons.

To replace the electron in the chlorophyll, a molecule of water is split. This splitting releases two electrons and results in the formation of oxygen (O_2) and 2 hydrogen ions (H^+) in the thylakoid space. The replacement of the electron enables chlorophyll to respond to another photon. The oxygen molecules produced as byproducts exit the leaf through the stomata and find their way to the surrounding environment. The hydrogen ions play critical roles in the remainder of the light-dependent reactions.

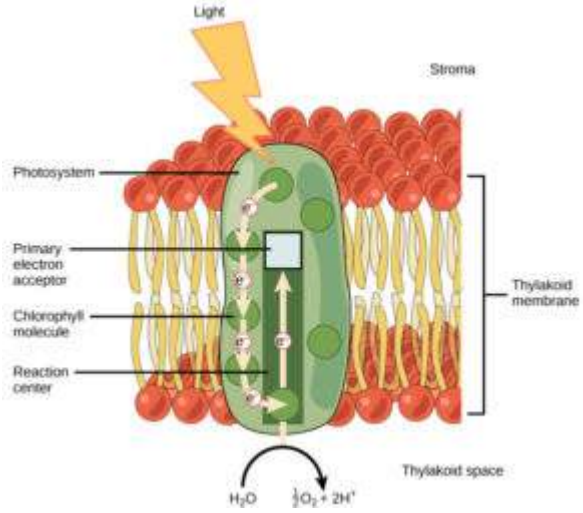


Figure 1 Light energy is absorbed by a chlorophyll molecule and is passed along a pathway to other chlorophyll molecules. The energy culminates in a molecule of chlorophyll found in the reaction center. The energy "excites" one of its electrons enough to leave the molecule and be transferred to a nearby primary electron acceptor. A molecule of water splits to release an electron, which is needed to replace the one donated. Oxygen and hydrogen ions are also formed from the splitting of water.

Keep in mind that the purpose of the light-dependent

reactions is to convert solar energy into chemical carriers (NADPH and ATP) that will be used in the Calvin cycle. In eukaryotes and some prokaryotes, two photosystems exist. The first is called **photosystem II** (PSII), which was named for the order of its discovery rather than for the order of the function. After a photon hits the photosystem II (PSII) reaction center, energy from sunlight is used to extract electrons from water. The electrons travel through the **chloroplast electron transport chain** to photosystem I (PSI), which reduces NADP+ to NADPH (**Figure 3**). As the electron passes along the electron transport chain, energy from the electron fuels proton pumps in the membrane that actively move hydrogen ions against their concentration gradient from the stroma into the thylakoid space. The electron transport chain moves protons across the thylakoid membrane into the lumen (the space inside the thylakoid disk). At the same time, splitting of water adds additional protons into the lumen, and reduction of NADPH removes protons from the stroma (the space outside the thylakoids). The net result is a high concentration of protons (H+) in the thylakoid lumen, and a low concentration of protons in the stroma. ATP synthase uses this electrochemical gradient to make ATP, just like it did in cellular respiration. Note that a high concentration of protons = an acidic pH, so the thylakoid lumen has a much more acidic (lower) pH than the stroma.

This whole process is quite analogous to the process that occurs during cellular respiration in the mitochondria. Recall that during CR, the energy carried by NADH and FADH2 is used to pump protons across the inner mitochondrial membrane and into the intermembrane space, creating an electrochemical proton gradient. This gradient is used to power oxidative phosphorylation by ATP synthase to create ATP.

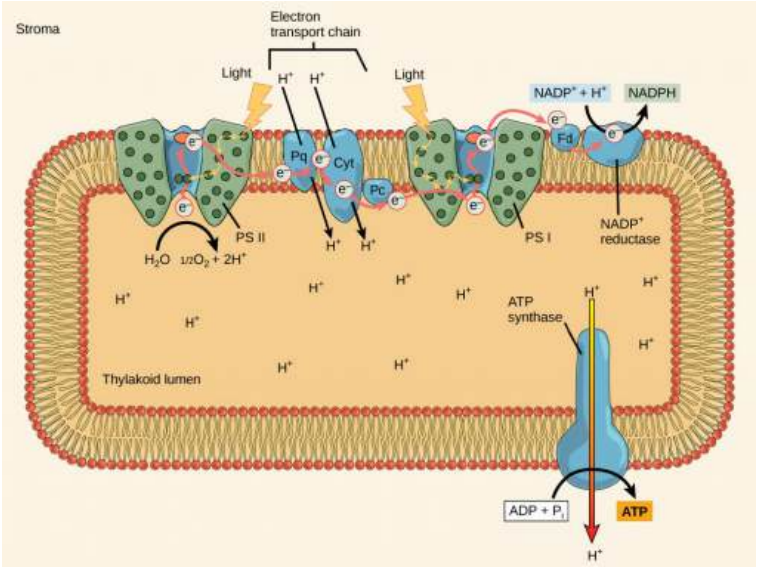


Figure 3 Energy from light is used by the chloroplast electron transport chain to pump protons across the thylakoid membrane into the lumen of the thylakoid. This creates a proton gradient that is used as a source of energy by ATP synthase.

GENERATING AN ENERGY MOLECULE: ATP

In the light-dependent reactions, energy absorbed by sunlight is stored by two types of energy-carrier molecules: ATP and NADPH. The energy that these molecules carry is stored in a bond that holds a single atom to the molecule. For ATP, it is a phosphate atom, and for NADPH, it is a hydrogen atom. Recall that NADH was a similar molecule that carried energy in the mitochondrion from the citric acid cycle to the electron transport chain. When these molecules release energy into the Calvin cycle, they each lose atoms to become the lower-energy molecules ADP and $NADP^+$.

The buildup of hydrogen ions in the thylakoid space forms

an electrochemical gradient because of the difference in the concentration of protons (H^+) and the difference in the charge across the membrane that they create. This potential energy is harvested and stored as chemical energy in ATP through chemiosmosis, the movement of hydrogen ions down their electrochemical gradient through the transmembrane enzyme ATP synthase, just as in the mitochondrion.

The hydrogen ions are allowed to pass through the thylakoid membrane through an embedded protein complex called ATP synthase. This same protein generated ATP from ADP in the mitochondrion. The energy generated by the hydrogen ion stream allows ATP synthase to attach a third phosphate to ADP, which forms a molecule of ATP in a process called photophosphorylation. The flow of hydrogen ions through ATP synthase is called chemiosmosis (just like in cellular respiration), because the ions move from an area of high to low concentration through a semi-permeable structure.

GENERATING ANOTHER ENERGY CARRIER: NADPH

The remaining function of the light-dependent reaction is to generate the other energy-carrier molecule, NADPH. As the electron from the electron transport chain arrives at photosystem I, it is re-energized with another photon captured by chlorophyll. The energy from this electron drives the formation of NADPH from $NADP^+$ and a hydrogen ion (H^+). Now that the solar energy is stored in energy carriers, it can be used to make a sugar molecule.

SECTION SUMMARY

The pigments of the first part of photosynthesis, the light-dependent reactions, absorb energy from sunlight. A photon strikes the antenna pigments of photosystem II to initiate photosynthesis. The energy travels to the reaction center that contains chlorophyll *a* to the electron transport chain, which pumps hydrogen ions into the thylakoid interior (the lumen). This action builds up a high concentration of hydrogen ions. The ions flow through ATP synthase via chemiosmosis to form molecules of ATP, which are used for the formation of sugar molecules in the second stage of photosynthesis. Photosystem I absorbs a second photon, which results in the formation of an NADPH molecule, another energy and reducing power carrier for the light-independent reactions.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=202#h5p-65

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

The Light Independent Reactions (aka the Calvin Cycle)

After the energy from the sun is converted and packaged into ATP and NADPH, the cell has the fuel needed to build carbohydrate molecules. The carbohydrate molecules made will have a backbone of carbon atoms. Where does the carbon come from? The carbon atoms used to build carbohydrate molecules comes from carbon dioxide, which diffuses into the leaves through the stomata. The **Calvin cycle** is the term used for the reactions of photosynthesis that use the energy stored by the light-dependent reactions to form glucose and other carbohydrate molecules (**Figure 1**).

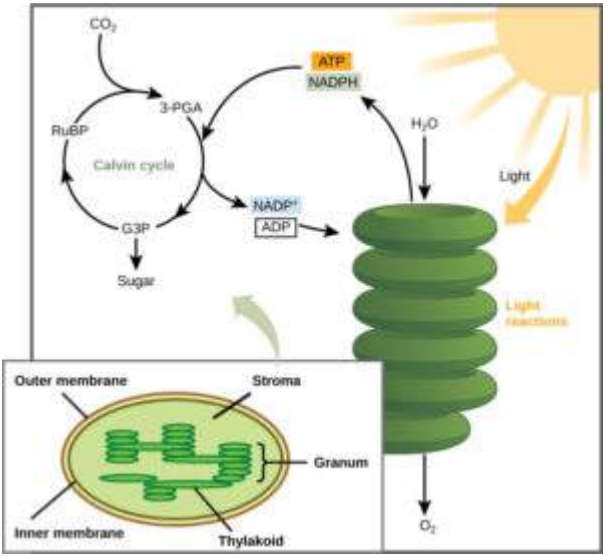


Figure 1 The light-dependent reactions harness energy from the sun to produce ATP and NADPH. These energy-carrying molecules travel into the stroma where the Calvin cycle reactions take place.

THE INTERWORKINGS OF THE CALVIN CYCLE

In plants, carbon dioxide (CO_2) enters the chloroplast through the stomata and diffuses into the stroma of the chloroplast—the site of the Calvin cycle reactions where sugar is synthesized. The reactions are named after the scientist who discovered them, and reference the fact that the reactions function as a cycle. Others call it the Calvin-Benson cycle to include the name of another scientist involved in its discovery.

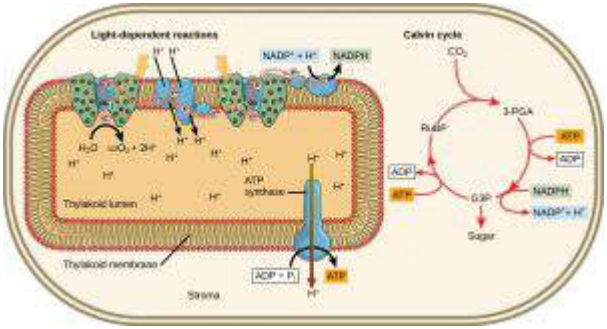


Figure 2 Light reactions harness energy from the sun to produce chemical bonds, ATP, and NADPH. These energy-carrying molecules are made in the stroma where carbon fixation takes place.

The Calvin cycle reactions (**Figure 2**) can be organized into three basic stages: fixation, reduction, and regeneration. In the stroma, in addition to CO_2, two other molecules are present to initiate the Calvin cycle: an enzyme abbreviated RuBisCO (which stands for ribulose-1,5-bisphosphate carboxylase/oxygenase, in case you're interested), and the molecule ribulose bisphosphate (RuBP). RuBP has five atoms of carbon and a phosphate group on each end.

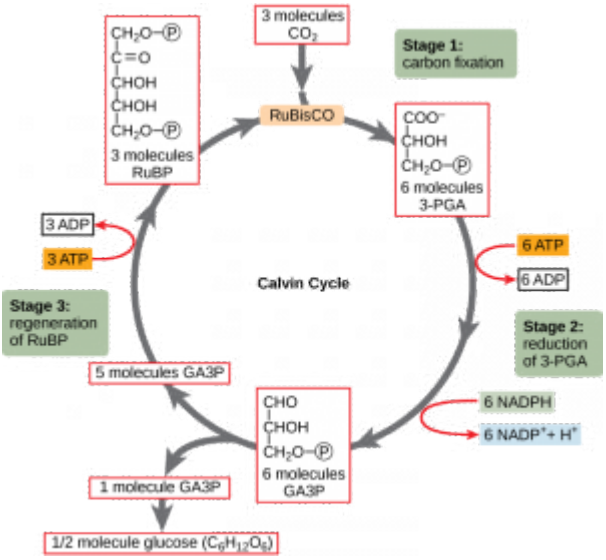


Figure 3 The Calvin cycle has three stages. In stage 1, the enzyme RuBisCO incorporates carbon dioxide into an organic molecule, 3-PGA. In stage 2, the organic molecule is reduced using electrons supplied by NADPH. In stage 3, RuBP, the molecule that starts the cycle, is regenerated so that the cycle can continue. Only one carbon dioxide molecule is incorporated at a time, so the cycle must be completed three times to produce a single three-carbon GA3P molecule, and six times to produce a six-carbon glucose molecule.

RuBisCO catalyzes a reaction between CO_2 and RuBP, which forms a six-carbon compound that is immediately converted into two three-carbon compounds. This process is called **carbon fixation**, because CO_2 is "fixed" from its inorganic form into organic molecules. You can think this as the carbon being converted from the "broken" form in CO_2 (which organisms are not able to directly use) into a "fixed" form, which organisms are able to utilize. Because of this very important role in photosynthesis, RuBisCO is probably the most abundant enzyme on earth.

ATP and NADPH use their stored energy to convert the

three-carbon compound, 3-PGA, into another three-carbon compound called **G3P.** This type of reaction is called a reduction reaction, because it involves the gain of electrons. A reduction is the gain of an electron by an atom or molecule. The molecules of ADP and NAD^+, resulting from the reduction reaction, return to the light-dependent reactions to be re-energized.

One of the G3P molecules leaves the Calvin cycle to contribute to the formation of the carbohydrate molecule, which is commonly glucose ($C_6H_{12}O_6$). Because the carbohydrate molecule has six carbon atoms, it takes six turns of the Calvin cycle to make one carbohydrate molecule (one for each carbon dioxide molecule fixed). The remaining G3P molecules regenerate RuBP, which enables the system to prepare for the carbon-fixation step. ATP is also used in the regeneration of RuBP.

In summary, it takes six turns of the Calvin cycle to fix six carbon atoms from CO_2. These six turns require energy input from 12 ATP molecules and 12 NADPH molecules in the reduction step and 6 ATP molecules in the regeneration step.

Evolution Connection

Photosynthesis in desert plants has evolved adaptations that conserve water. In the harsh dry heat, every drop of water must be used to survive. Because stomata must open to allow for the uptake of CO_2, water escapes from the leaf during active photosynthesis. Desert plants have evolved processes to conserve water and deal with harsh conditions. A more efficient use of CO_2 allows plants to adapt to living with less water. Some plants such as cacti can prepare materials for photosynthesis during the night by a temporary carbon fixation/storage process, because opening the stomata at this time conserves water due to cooler temperatures. In addition, cacti have

evolved the ability to carry out low levels of photosynthesis without opening stomata at all, an extreme mechanism to face extremely dry periods.

SECTION SUMMARY

Using the energy carriers formed in the first steps of photosynthesis, the light-independent reactions, or the Calvin cycle, take in CO_2 from the environment. An enzyme, RuBisCO, catalyzes a reaction with CO_2 and another molecule, RuBP. After three cycles, a three-carbon molecule of G3P leaves the cycle to become part of a carbohydrate molecule. The remaining G3P molecules stay in the cycle to be regenerated into RuBP, which is then ready to react with more CO_2. Photosynthesis forms an energy cycle with the process of cellular respiration. Plants need both photosynthesis and respiration for their ability to function in both the light and dark, and to be able to interconvert essential metabolites. Therefore, plants contain both chloroplasts and mitochondria.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=187#h5p-66

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Photosynthesis in Prokaryotes

The two parts of photosynthesis—the light-dependent reactions and the Calvin cycle—have been described, as they take place in chloroplasts. However, prokaryotes, such as cyanobacteria, lack membrane-bound organelles (including chloroplasts). Prokaryotic photosynthetic organisms have infoldings of the plasma membrane for chlorophyll attachment and photosynthesis (**Figure 1**). It is here that organisms like cyanobacteria can carry out photosynthesis.

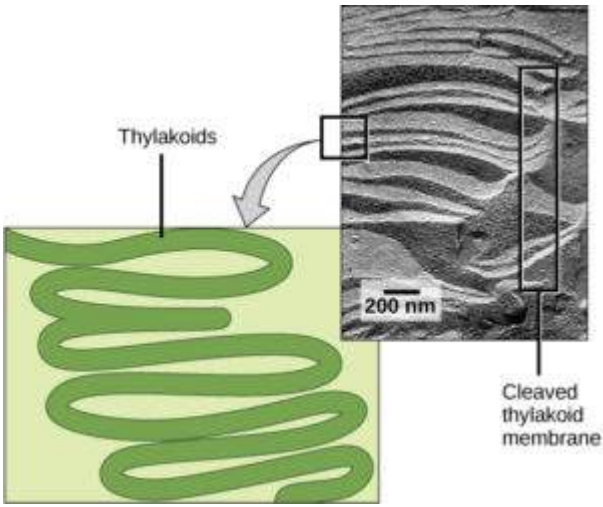


Figure 1 A photosynthetic prokaryote has infolded regions of the plasma membrane that function like thylakoids. Although these are not contained in an organelle, such as a chloroplast, all of the necessary components are present to carry out photosynthesis. (credit: scale-bar data from Matt Russell)

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=189#h5p-67

REFERENCES

Text adapted from: OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

CELL DIVISION - BINARY FISSION AND MITOSIS

Learning Objectives

Describe the processes used for cell division. Compare the process and consequences of binary fission, mitosis, and meiosis.

Since all living things are made up of one or more cells, all living things have to undergo some type of cell division. Cell division serves several basic purposes: reproduction, repair, and growth. Typically during cell division, the DNA of the organism is copied, then divided into the new cells using one or more divisions. We will therefore start our discussion of cell division with a brief overview of the process of DNA replication (copying DNA), followed by the process by which various types of cells divide that DNA into new cells.

How DNA is arranged in a cell

DNA is a working molecule; it must be replicated (copied) when a cell is ready to divide, and it must be "read" to produce the molecules, such as proteins, to carry out the functions of the cell. For this reason, the DNA is protected and packaged in very specific ways. Because they must carry so much information, DNA molecules can be very long. Stretched end-to-end, the DNA molecules in a single human cell would come to a length of about 2 meters (roughly 6 feet). Thus, the DNA for a cell must be packaged in a very ordered way to fit and function within a structure (the cell) that is not visible to the naked eye.

A cell's complete complement of DNA is called its **genome**. In prokaryotes (bacteria), the genome is composed of a single, double-stranded DNA molecule in the form of a loop or circle. The region in the cell containing this genetic material is called a **nucleoid.** Some prokaryotes also have smaller loops of DNA called **plasmids** that are not essential for normal growth.

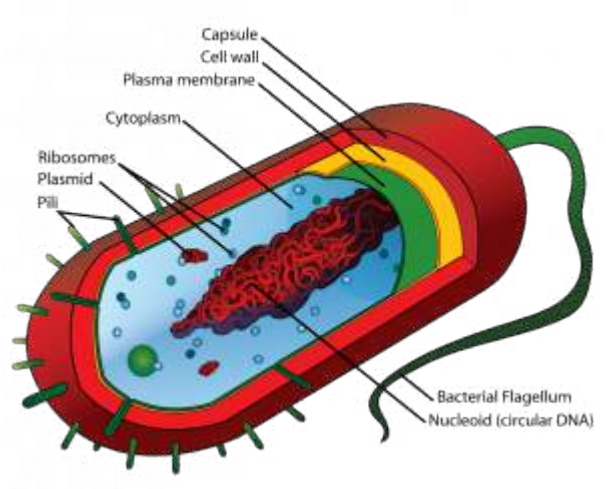


Figure 1 An average prokaryotic cell. Note that the DNA is not surrounded by a membrane to create a nucleus. Photo credit Lady of Hats; Wikipedia.

The size of the genome in one of the most well-studied prokaryotes, *Escherichia coli,* is 4.6 million base pairs. This would extend a distance of about 1.6 mm if stretched out. Compare that to the length of an E. coli cell, which is approximately 1-2μm long. 1.6mm = 1600μm: so how does all this DNA fit inside a tiny cell? The DNA is twisted beyond the double helix in what is known as supercoiling. Some proteins are known to be involved in the supercoiling; other proteins and enzymes help in maintaining the supercoiled structure.

Eukaryotes, such as animals and plants, have **chromosomes** that consist of linear DNA molecules. Chromosomes can be seen as thread-like structures located inside the nucleus of eukaryotic cells. Each chromosome is made of protein and a single linear double-helix of DNA (**Figure 2**). The term chromosome comes from the Greek words for color (chroma) and body (soma). Scientists gave this name to chromosomes because they are cell structures,

or bodies, that are strongly stained by some colorful dyes used in research.

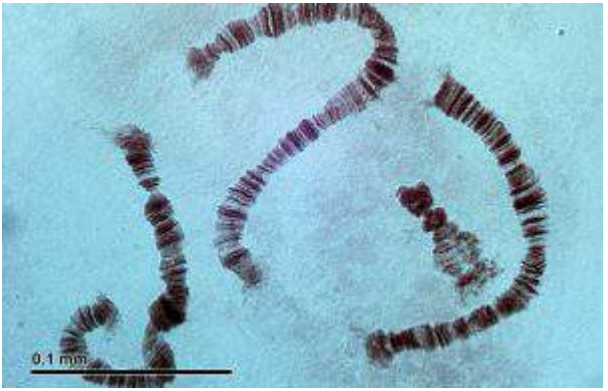


Figure 2 Linear chromosomes from the salivary glands of nonbiting midge larvae. Photo credit Joseph Resichig; Wikimedia.

Eukaryotes typically have much more DNA than prokaryotes: the human genome is roughly 3 *billion* base pairs while the E. coli genome is roughly 4 *million.* For this reason, eukaryotes employ a different type of packing strategy to fit their DNA inside the **nucleus** (**Figure 4**). At the most basic level, DNA is wrapped around proteins known as **histones**. The DNA wrapped around histones wraps and stacks through several additional levels of complexity. These thicker more compact structures are what you have seen before in pictures labeled "chromosomes".

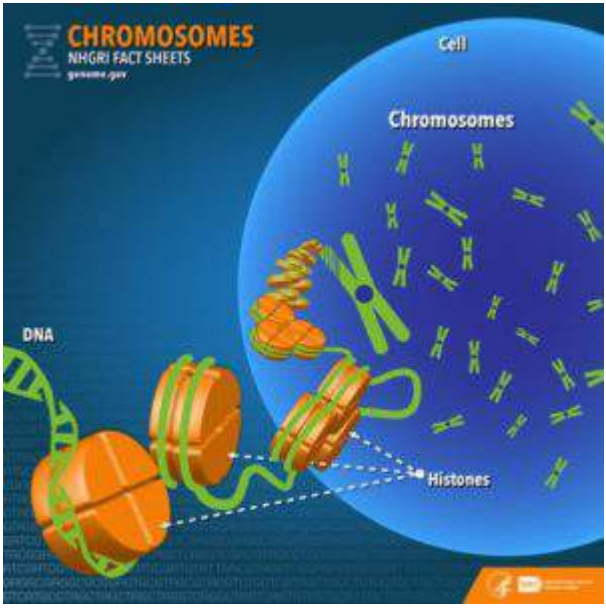


Figure 4: The basic structure of eukaryotic chromosomes inside the nucleus of a cell ("Chromosomes" by National Human Genome Research Institute is in the Public Domain)

To summarize:

- Prokaryotes have relatively small amounts of DNA (millions of basepairs) found in one circular genome, which is located in the cytoplasm in the nucleoid.
- Eukaryotes have larger amounts of DNA (billions of basepairs) found in several linear chromosomes, which are located inside the nucleus.

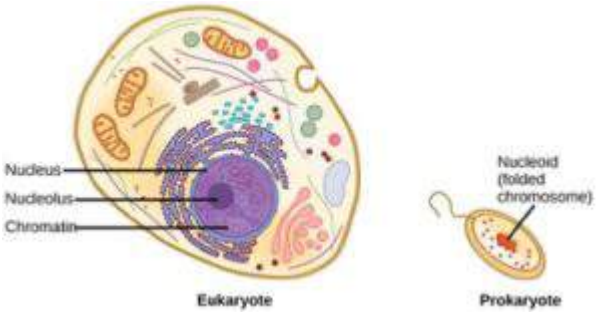


Figure 5: A eukaryote contains a well-defined nucleus, whereas in prokaryotes, the chromosome lies in the cytoplasm in an area called the nucleoid.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8v2Xzdco@5/The-Structure-of-DNA

"Genetics Home Reference: Help Me Understand Genetics" by , National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

An Overview of DNA Replication

When a cell divides, it is important that each daughter cell receives an identical copy of the DNA. This is accomplished by the process of **DNA replication**. The replication of DNA occurs before the cell begins to divide into two separate cells.

The discovery and characterization of the structure of the double helix provided a hint as to how DNA is copied. Recall that adenine nucleotides pair with thymine nucleotides, and cytosine with guanine, and that DNA is double stranded. This means that the two strands are **complementary** to each other. For example, a strand of DNA with a nucleotide sequence of AGTCATGA will have a complementary strand with the sequence TCAGTACT (**Figure 1**).

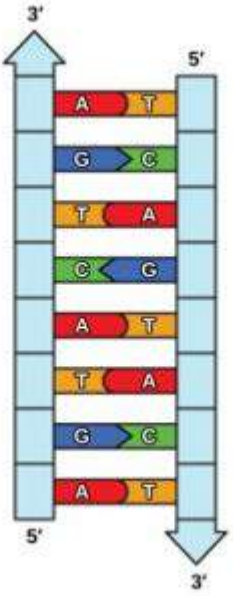


Figure 1: The two strands of DNA are complementary, meaning the sequence of bases in one strand can be used to create the correct sequence of bases in the other strand.

Because of the complementarity of the two strands, having one strand means that it is possible to recreate the other strand. This model for replication suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied (**Figure 2**).

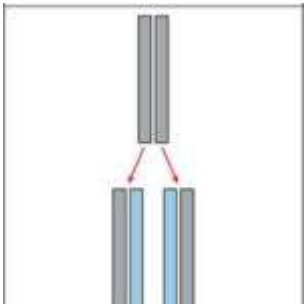


Figure 2: The semiconservative model of DNA replication is shown. Gray indicates the original DNA strands, and blue indicates newly synthesized DNA.

During DNA replication, each of the two strands that make up the double helix serves as a template from which new strands are copied. The new strand will be complementary to the parental or "old" strand. Each new double strand consists of one parental strand and one new daughter strand. This is known as **semiconservative replication**. When two DNA copies are formed, they have an identical sequence of nucleotide bases and are divided equally into two daughter cells.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:2ousESf0@5/DNA-Replication

Prokaryotic Cell Division

The cell division process used by prokaryotes (such as E. coli bacteria) and some unicellular eukaryotes is called **binary fission**. For unicellular organisms, cell division is the only method to produce new individuals. The outcome of this type of cell reproduction is a pair of **daughter cells** that are genetically identical to the original **parent cell**. In unicellular organisms, daughter cells are whole individual organisms. This is a less complicated and much quicker process than cell division in eukaryotes. Because of the speed of bacterial cell division, populations of bacteria can grow very rapidly.

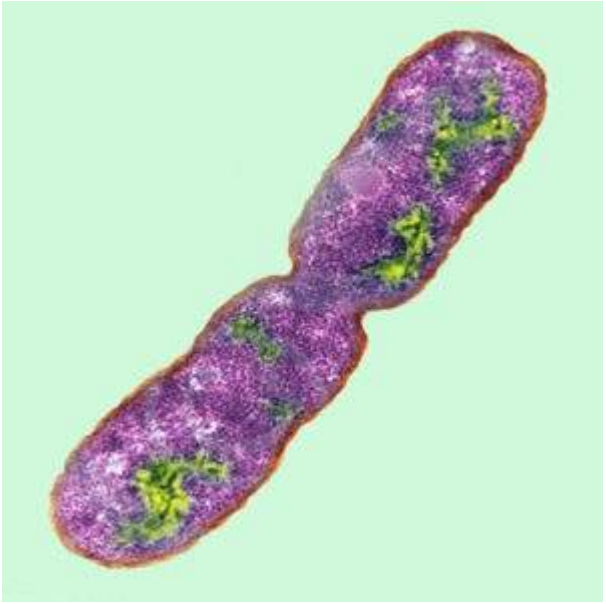


Figure 1: An E. coli bacteria dividing into two identical daughter cells

To achieve the outcome of identical daughter cells, there are some essential steps. The genomic DNA must be replicated (using DNA replication) to produce two identical copies of the entire genome. Then, one copy must be moved into each of the daughter cells. The cytoplasmic contents must also be divided to give both new cells the machinery to sustain life. Since bacterial cells have a genome that consists of a single, circular DNA chromosome, the process of cell division is very simple.

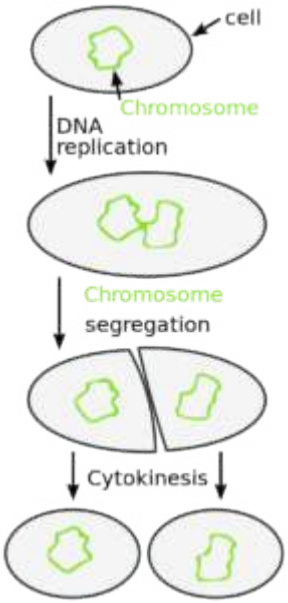


Figure 2: Prokaryotic cell division occurs via a process called binary fission.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Vbi92lHB@9/The-Cell-Cycle

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:LlKfCy5H@4/Prokaryotic-Cell-Division

The Eukaryotic Cell Cycle

Eukaryotes have two major types of cell division: mitosis and meiosis. Mitosis is used to produce new body cells for growth and healing, while meiosis is used to produce sex cells (eggs and sperm). Meiosis will be discussed in a later chapter.

The **cell cycle** is an ordered series of events involving cell growth and cell division that produces two new daughter cells via mitosis. The length of the cell cycle is highly variable even within the cells of an individual organism. In humans, the frequency of cell turnover ranges from a few hours in early embryonic development to an average of two to five days for epithelial cells, or to an entire human lifetime spent without dividing in specialized cells such as cortical neurons or cardiac muscle cells. There is also variation in the time that a cell spends in each phase of the cell cycle. When fast-dividing mammalian cells are grown in culture (outside the body under optimal growing conditions), the length of the cycle is approximately 24 hours. The timing of events in the cell cycle is controlled by mechanisms that are both internal and external to the cell.

Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages of growth, DNA replication, and division that produce two genetically identical cells. The cell cycle has two major phases: interphase and the mitotic phase (**Figure 1**). During **interphase**, the cell grows and DNA is replicated. During the

mitotic phase, the replicated DNA and cytoplasmic contents are separated and the cell divides.

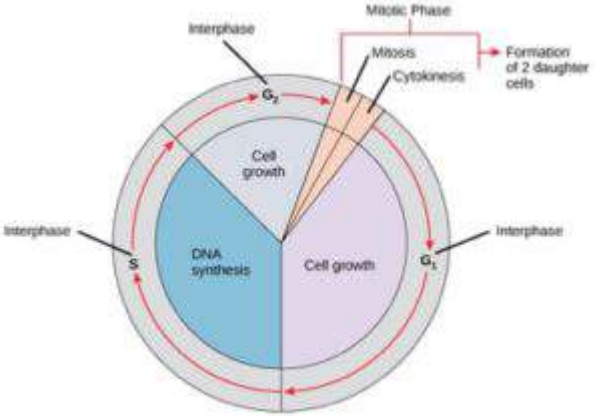


Figure 1: A cell moves through a series of phases in an orderly manner. During interphase, G1 involves cell growth and protein synthesis, the S phase involves DNA replication and the replication of the centrosome, and G2 involves further growth and protein synthesis. The mitotic phase follows interphase. Mitosis is nuclear division during which duplicated chromosomes are segregated and distributed into daughter nuclei. Usually the cell will divide after mitosis in a process called cytokinesis in which the cytoplasm is divided and two daughter cells are formed.

INTERPHASE

During interphase, the cell undergoes normal processes while also preparing for cell division. For a cell to move from interphase to the mitotic phase, many internal and external conditions must be met. The three stages of interphase are called G_1, S, and G_2.

G_1 PHASE (FIRST GAP)

The first stage of interphase is called the G_1 phase (first gap) because, from a microscopic aspect, little change is visible. However, during the G_1 stage, the cell is quite active at the biochemical level. The cell is accumulating the building blocks of chromosomal DNA and the associated proteins as well as accumulating sufficient energy reserves to complete the task of replicating each chromosome in the nucleus.

S PHASE (SYNTHESIS OF DNA)

Throughout interphase, nuclear DNA remains in a semi-condensed chromatin configuration. In the S phase, DNA replication can proceed through the mechanisms that result in the formation of identical pairs of DNA molecules—**sister chromatids**—that are firmly attached to the centromeric region (**Figure 2**).

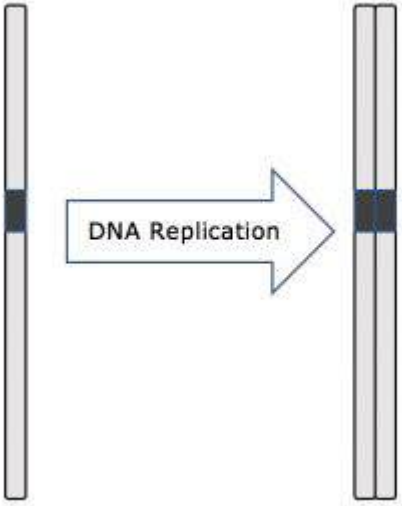


Figure 2 DNA replication during S phase copies each linear chromosome. The chromosomes remain attached together at a region called the centromere. Photo credit: Lisa Bartee

The **centrosome** is also duplicated during the S phase. The two centrosomes will give rise to the mitotic spindle, the apparatus that orchestrates the movement of chromosomes during mitosis. At the center of each animal cell, the centrosomes of animal cells are associated with a pair of rod-like objects, the **centrioles,** which are at right angles to each other. Centrioles help organize cell division. Centrioles are not present in the centrosomes of other eukaryotic species, such as plants and most fungi.

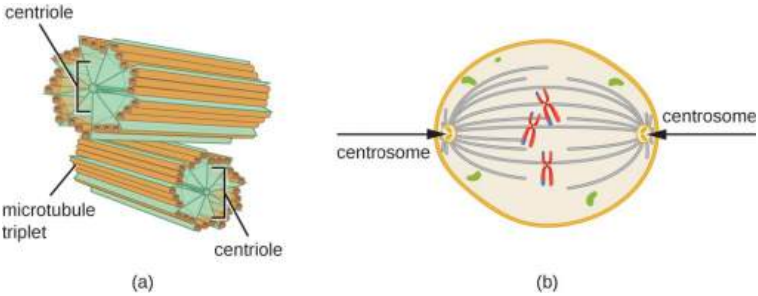


Figure 3 (a) Structure of the centrioles making up the centrosome. (b) Centrioles give rise to the mitotic spindle (grey threadlike structures). Photo credit: CNX OpenStax Microbiology.

G_2 PHASE (SECOND GAP)

In the G_2 phase, the cell replenishes its energy stores and synthesizes proteins necessary for chromosome manipulation. Some cell organelles are duplicated, and the cytoskeleton is dismantled to provide resources for the mitotic phase. There may be additional cell growth during G_2. The final preparations for the mitotic phase must be completed before the cell is able to enter the first stage of mitosis.

THE MITOTIC PHASE

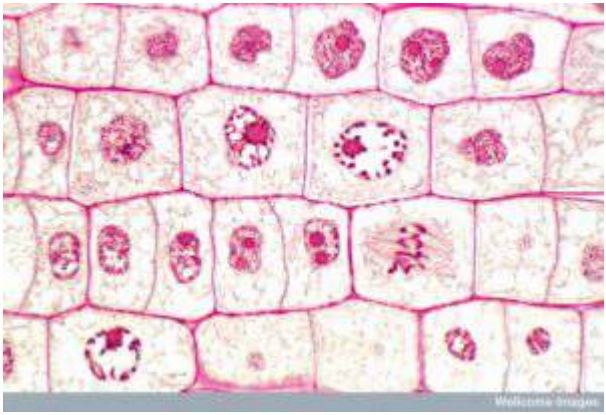


Figure 4: Mitosis in onion root cells. The cells in this image are in various stages of mitosis. (Credit: Spike Walker. Wellcome Images images@wellcome.ac.uk)

To make two daughter cells, the contents of the nucleus and the cytoplasm must be divided. The mitotic phase is a multistep process during which the duplicated chromosomes are aligned, separated, and moved to opposite poles of the cell, and then the cell is divided into two new identical daughter cells. The first portion of the mitotic phase, **mitosis**, is composed of five stages, which accomplish nuclear division (**Figure 5**). The second portion of the mitotic phase, called **cytokinesis,** is the physical separation of the cytoplasmic components into two daughter cells. Although the stages of mitosis are similar for most eukaryotes, the process of cytokinesis is quite different for eukaryotes that have cell walls, such as plant cells.

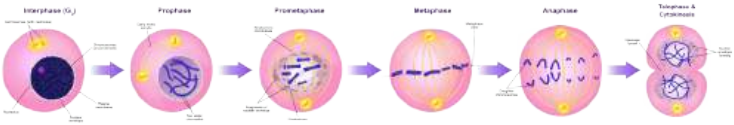


Figure 5 Summary of the process of mitosis. Photo credit Oganesson007, Wikimedia.

PROPHASE

During prophase, the "first phase," the nuclear envelope starts to dissociate into small vesicles, and the membranous organelles (such as the Golgi apparatus and endoplasmic reticulum), fragment and disperse toward the edges of the cell. The nucleolus disappears. The centrosomes begin to move to opposite poles of the cell. Microtubules that will form the mitotic spindle extend between the centrosomes, pushing them farther apart as the microtubule fibers lengthen. The sister chromatids begin to coil more tightly with the aid of condensin proteins and become visible under a light microscope.

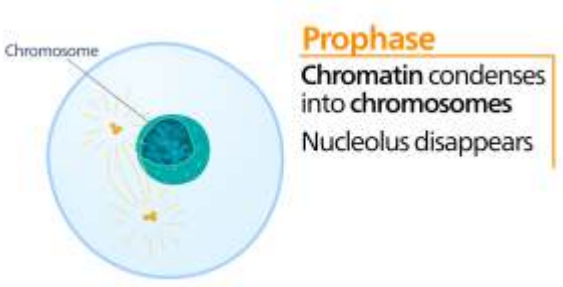


Figure 6 Prophase. Photo credit Kelvin13; Wikimedia.

PROMETAPHASE

During **prometaphase,** the "first change phase," many

processes that were begun in prophase continue to advance. The remnants of the nuclear envelope fragment. The mitotic spindle continues to develop as more microtubules assemble and stretch across the length of the former nuclear area. Chromosomes become more condensed and discrete. Each sister chromatid develops a protein structure called a **kinetochore** in the centromeric region.

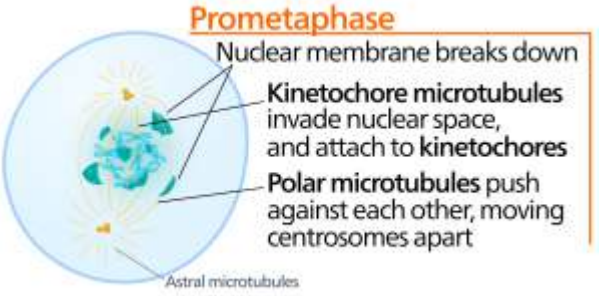


Figure 7 Prometaphase. Photo credit Kelvin13; Wikimedia.

The proteins of the kinetochore attract and bind mitotic spindle microtubules. As the spindle microtubules extend from the centrosomes, some of these microtubules come into contact with and firmly bind to the kinetochores. Once a mitotic fiber attaches to a chromosome, the chromosome will be oriented until the kinetochores of sister chromatids face the opposite poles. Eventually, all the sister chromatids will be attached via their kinetochores to microtubules from opposing poles. Spindle microtubules that do not engage the chromosomes are called polar microtubules. These microtubules overlap each other midway between the two poles and contribute to cell elongation. Astral microtubules are located near the poles, aid in spindle orientation, and are required for the regulation of mitosis.

This illustration shows two sister chromatids. Each has a kinetochore at the centromere, and mitotic spindle microtubules radiate from the kinetochore.

Figure 8 During prometaphase, mitotic spindle microtubules from opposite poles attach to each sister chromatid at the kinetochore. In anaphase, the connection between the sister chromatids breaks down, and the microtubules pull the chromosomes toward opposite poles.

METAPHASE

During **metaphase,** the "change phase," all the chromosomes are aligned in a plane called the metaphase plate, or the equatorial plane, midway between the two poles of the cell. The sister chromatids are still tightly attached to each other by cohesin proteins. At this time, the chromosomes are maximally condensed.

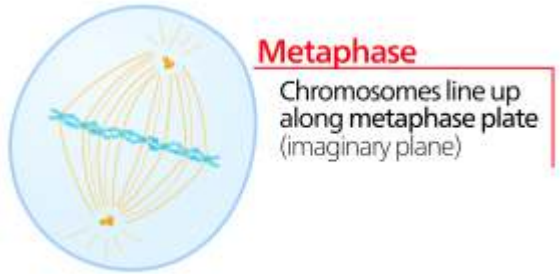


Figure 9 Metaphase. Photo credit Kelvin13; Wikimedia.

ANAPHASE

During anaphase, the "upward phase," the cohesin proteins degrade, and the sister chromatids separate at the centromere. Each chromatid, now called a chromosome, is pulled rapidly toward the centrosome to which its microtubule is attached. The cell becomes visibly elongated

(oval shaped) as the polar microtubules slide against each other at the metaphase plate where they overlap.

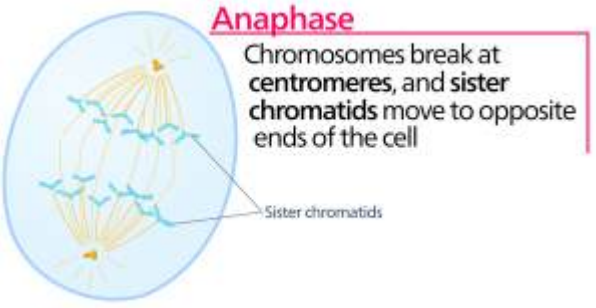


Figure 10 Anaphase. Photo credit Kelvin13; Wikimedia.

TELOPHASE

During telophase, the "distance phase," the chromosomes reach the opposite poles and begin to decondense (unravel), relaxing into a chromatin configuration. The mitotic spindles are depolymerized into tubulin monomers that will be used to assemble cytoskeletal components for each daughter cell. Nuclear envelopes form around the chromosomes, and nucleosomes appear within the nuclear area.

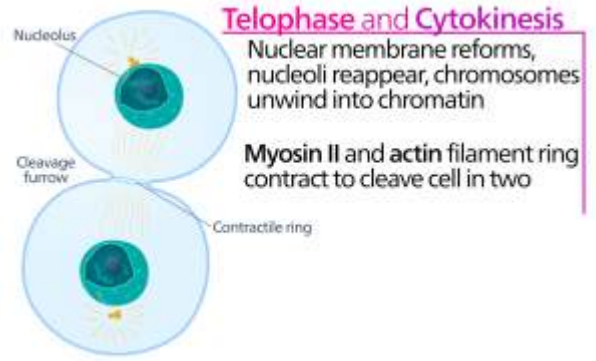


Figure 11 Telophase. Photo credit Kelvin13; Wikimedia.

CYTOKINESIS

Cytokinesis, or "cell motion," is the second main stage of the mitotic phase during which cell division is completed via the physical separation of the cytoplasmic components into two daughter cells. Division is not complete until the cell components have been divided and completely separated into the two daughter cells. Although the stages of mitosis are similar for most eukaryotes, the process of cytokinesis is quite different for eukaryotes that have cell walls, such as plant cells.

In cells such as animal cells that lack cell walls, cytokinesis follows the onset of anaphase. A contractile ring composed of actin filaments forms just inside the plasma membrane at the former metaphase plate (**Figure 12**). The actin filaments pull the equator of the cell inward, forming a fissure. This fissure, or "crack," is called the cleavage furrow. The furrow deepens as the actin ring contracts, and eventually the membrane is cleaved in two.

In plant cells, a new cell wall must form between the daughter cells. During interphase, the Golgi apparatus accumulates enzymes, structural proteins, and glucose molecules prior to breaking into vesicles and dispersing throughout the dividing cell (**Figure 12**). During telophase, these Golgi vesicles are transported on microtubules to form a phragmoplast (a vesicular structure) at the metaphase plate. There, the vesicles fuse and coalesce from the center toward the cell walls; this structure is called a cell plate. As more vesicles fuse, the cell plate enlarges until it merges with the cell walls at the periphery of the cell. Enzymes use the glucose that has accumulated between the membrane layers to build a new cell wall. The Golgi membranes become parts of the plasma membrane on either side of the new cell wall.

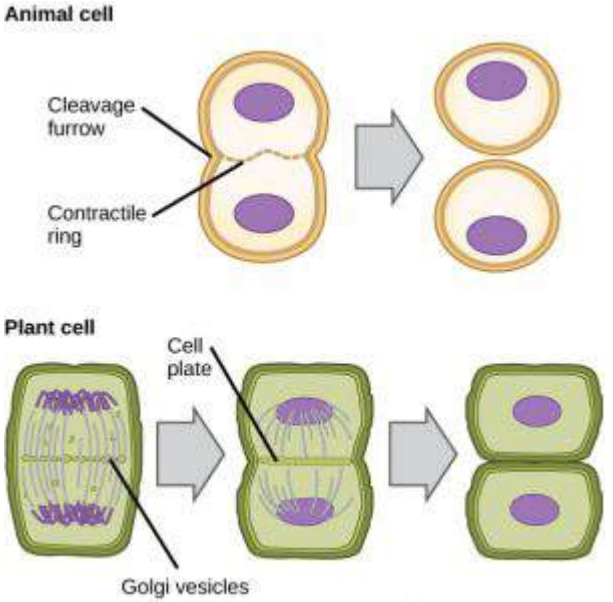


Figure 12 During cytokinesis in animal cells, a ring of actin filaments forms at the metaphase plate. The ring contracts, forming a cleavage furrow, which divides the cell in two. In plant cells, Golgi vesicles coalesce at the former metaphase plate, forming a phragmoplast. A cell plate formed by the fusion of the vesicles of the phragmoplast grows from the center toward the cell walls, and the membranes of the vesicles fuse to form a plasma membrane that divides the cell in two.

SUMMARY OF MITOSIS AND CYTOKINESIS

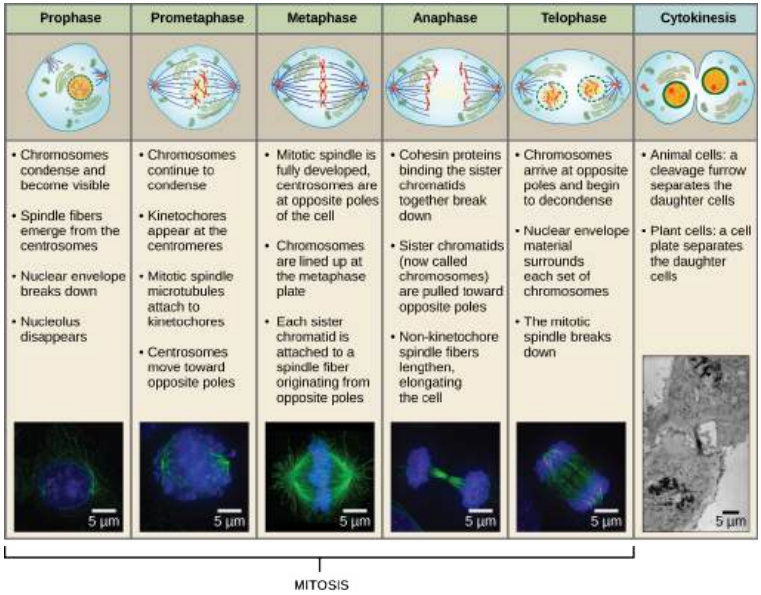


Figure 13 Mitosis is divided into five stages—prophase, prometaphase, metaphase, anaphase, and telophase. The pictures at the bottom were taken by fluorescence microscopy of cells artificially stained by fluorescent dyes: blue fluorescence indicates DNA (chromosomes) and green fluorescence indicates microtubules (spindle apparatus). (credit "mitosis drawings": modification of work by Mariana Ruiz Villareal; credit "micrographs": modification of work by Roy van Heesbeen; credit "cytokinesis micrograph": Wadsworth Center/New York State Department of Health; scale-bar data from Matt Russell)

G0 PHASE

Not all cells adhere to the classic cell-cycle pattern in which a newly formed daughter cell immediately enters interphase,

closely followed by the mitotic phase. Cells in the **G0 phase** are not actively preparing to divide. The cell is in a quiescent (inactive) stage, having exited the cell cycle. Some cells enter G0 temporarily until an external signal triggers the onset of G1. Other cells that never or rarely divide, such as mature cardiac muscle and nerve cells, remain in G0 permanently).

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Vbi92lHB@9/The-Cell-Cycle

Control of the Cell Cycle

It is essential that daughter cells be exact duplicates of the parent cell. Mistakes in the duplication or distribution of the chromosomes lead to mutations that may be passed forward to every new cell produced from the abnormal cell. To prevent a compromised cell from continuing to divide, there are internal control mechanisms that operate at three main **cell cycle checkpoints** at which the cell cycle can be stopped until conditions are favorable.

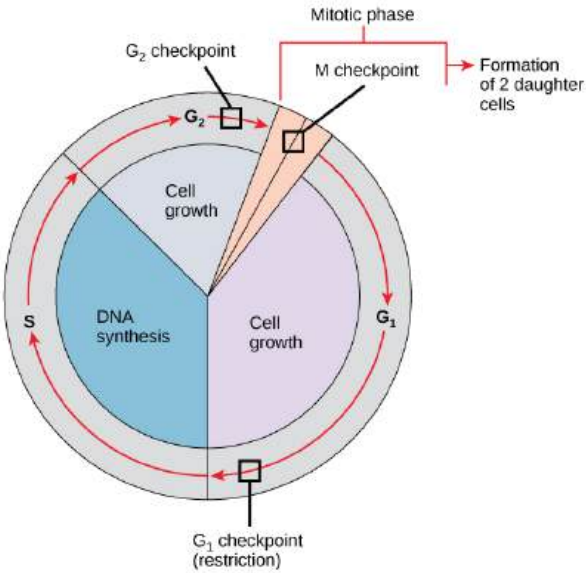


Figure 1 The cell cycle is controlled at three checkpoints. Integrity of the DNA is assessed at the G1 checkpoint. Proper chromosome duplication is assessed at the G2 checkpoint. Attachment of each kinetochore to a spindle fiber is assessed at the M checkpoint.

The first checkpoint (G_1) determines whether all conditions are favorable for cell division to proceed. This checkpoint is the point at which the cell irreversibly commits to the cell-division process. In addition to adequate reserves and cell size, there is a check for damage to the genomic DNA. A cell that does not meet all the requirements will not be released into the S phase.

The second checkpoint (G_2) bars the entry to the mitotic phase if certain conditions are not met. The most important role of this checkpoint is to ensure that all of the chromosomes have been replicated and that the replicated DNA is not damaged.

The final checkpoint (M) occurs in the middle of mitosis. This checkpoint determines if all of the copied chromosomes are arranged appropriately to be separated to opposite sides

of the cell. If this doesn't happen correctly, incorrect numbers of chromosomes can be partitioned into each of the daughter cells, which would likely cause them to die.

REGULATOR MOLECULES OF THE CELL CYCLE

In addition to the internally controlled checkpoints, there are two groups of intracellular molecules that regulate the cell cycle. These regulatory molecules either promote progress of the cell to the next phase (positive regulation) or halt the cycle (negative regulation). Regulator molecules may act individually, or they can influence the activity or production of other regulatory proteins. Therefore, it is possible that the failure of a single regulator may have almost no effect on the cell cycle, especially if more than one mechanism controls the same event. It is also possible that the effect of a deficient or non-functioning regulator can be wide-ranging and possibly fatal to the cell if multiple processes are affected.

POSITIVE REGULATION OF THE CELL CYCLE

Two groups of proteins, called **cyclins** and **cyclin-dependent kinases (Cdks)**, are responsible for the progress of the cell through the various checkpoints. The levels of the four cyclin proteins fluctuate throughout the cell cycle in a predictable pattern (**Figure 2**). Increases in the concentration of cyclin proteins are triggered by both external and internal signals. After the cell moves to the next stage of the cell cycle, the cyclins that were active in the previous stage are degraded.

This graph shows the concentrations of different cyclin proteins during various phases of the cell cycle. Cyclin D concentrations increase in G_{1} and decrease at the end of mitosis. Cyclin E levels rise during G_{1} and fall during S phase. Cyclin A levels rise during S phase and fall during mitosis. Cyclin B levels rise in S phase and fall during mitosis.

Figure 2 The concentrations of cyclin proteins change throughout the cell cycle. There is a direct correlation between cyclin accumulation and the three major cell cycle checkpoints. Also note the sharp decline of cyclin levels following each checkpoint (the transition between phases of the cell cycle), as cyclin is degraded by cytoplasmic enzymes. (credit: modification of work by "WikiMiMa"/Wikimedia Commons)

Cyclins regulate the cell cycle only when they are tightly bound to Cdks. To be fully active, the Cdk/cyclin complex must also be phosphorylated in specific locations. Like all kinases, Cdks are enzymes (kinases) that phosphorylate other proteins. Phosphorylation activates the protein by changing its shape. The proteins phosphorylated by Cdks are involved in advancing the cell to the next phase (**Figure 3**). The levels of Cdk proteins are relatively stable throughout the cell cycle; however, the concentrations of cyclin fluctuate and determine when Cdk/cyclin complexes form. The different cyclins and Cdks bind at specific points in the cell cycle and thus regulate different checkpoints.

This illustration shows a cyclin protein binding to a Cdk. The cyclin/Cdk complex is activated when a kinase phosphorylates it. The cyclin/Cdk complex, in turn, phosphorylates other proteins, thus advancing the cell cycle.

Figure 3 Cyclin-dependent kinases (Cdks) are protein kinases that, when fully activated, can phosphorylate and thus activate other proteins that advance the cell cycle past a checkpoint. To become fully activated, a Cdk must bind to a cyclin protein and then be phosphorylated by another kinase.

Since the cyclic fluctuations of cyclin levels are based on the timing of the cell cycle and not on specific events, regulation of the cell cycle usually occurs by either the Cdk molecules alone or the Cdk/cyclin complexes. Without a specific concentration of fully activated cyclin/Cdk complexes, the cell cycle cannot proceed through the checkpoints.

NEGATIVE REGULATION OF THE CELL CYCLE

The second group of cell cycle regulatory molecules are negative regulators. In positive regulation, active molecules such as CDK/cyclin complexes cause the cell cycle to progress. In negative regulation, active molecules halt the cell cycle.

The best understood negative regulatory molecules are retinoblastoma protein (Rb), p53, and p21. Much of what is known about cell cycle regulation comes from research conducted with cells that have lost regulatory control. All three of these regulatory proteins were discovered to be damaged or non-functional in cells that had begun to replicate uncontrollably (became cancerous). In each case, the main cause of the unchecked progress through the cell cycle was a faulty copy of the regulatory protein. For this reason, Rb and other proteins that negatively regulate the cell cycle are sometimes called **tumor suppressors**.

Rb, p53, and p21 act primarily at the G_1 checkpoint. **p53** is a multi-functional protein that has a major impact on the commitment of a cell to division because it acts when there is damaged DNA in cells that are undergoing the preparatory processes during G_1. If damaged DNA is detected, p53 halts the cell cycle and recruits enzymes to repair the DNA. If the DNA cannot be repaired, p53 can trigger apoptosis, or cell suicide, to prevent the duplication of damaged chromosomes. As p53 levels rise, the production of **p21** is triggered. p21 enforces the halt in the cycle dictated by p53

by binding to and inhibiting the activity of the Cdk/cyclin complexes. As a cell is exposed to more stress, higher levels of p53 and p21 accumulate, making it less likely that the cell will move into the S phase.

Rb exerts its regulatory influence on other positive regulator proteins. Chiefly, Rb monitors cell size. In the active, dephosphorylated state, Rb binds to proteins called transcription factors (**Figure 4**). Transcription factors "turn on" specific genes, allowing the production of proteins encoded by that gene. When Rb is bound to transcription factors, production of proteins necessary for the G_1/S transition is blocked. As the cell increases in size, Rb is slowly phosphorylated until it becomes inactivated. Rb releases the transcription factors, which can now turn on the gene that produces the transition protein, and this particular block is removed. For the cell to move past each of the checkpoints, all positive regulators must be "turned on," and all negative regulators must be "turned off."

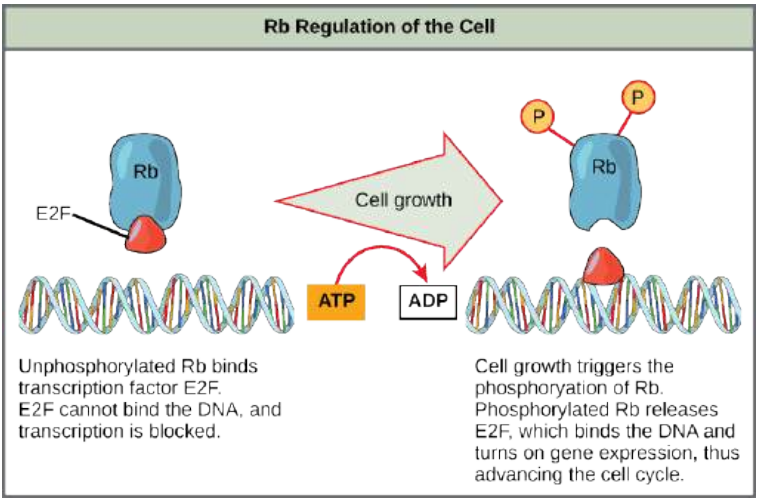


Figure 4 Rb halts the cell cycle and releases its hold in response to cell growth.

REFERENCES

Unless otherwise noted, images on this page are licensed under CC-BY 4.0 by OpenStax.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Vbi92lHB@9/The-Cell-Cycle

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:LlKfCy5H@4/Prokaryotic-Cell-Division

Cancer and the cell cycle

Cancer comprises many different diseases caused by a common mechanism: uncontrolled cell growth. Despite the redundancy and overlapping levels of cell cycle control, errors do occur. One of the critical processes monitored by the cell cycle checkpoint surveillance mechanism is the proper replication of DNA during the S phase. Even when all of the cell cycle controls are fully functional, a small percentage of replication errors (mutations) will be passed on to the daughter cells. If changes to the DNA nucleotide sequence of a gene are not corrected, a gene mutation results. All cancers start when a gene mutation causes a change in the order of the amino acids that make up a protein that plays a key role in cell reproduction. Changes in the amino acid sequence can change the shape of the protein. Since the shape of the protein is changed, its function may be changed as well. The change in the cell that results from the misshaped protein may be minor: perhaps a slight delay in the binding of Cdk to cyclin or an Rb protein that detaches from its target DNA while still phosphorylated. Even minor mistakes, however, may allow subsequent mistakes to occur more readily. Over and over, small uncorrected errors are passed from the parent cell to the daughter cells and amplified as each generation produces more non-functional proteins from uncorrected DNA damage. Eventually, the pace of the cell cycle speeds up

as the effectiveness of the control and repair mechanisms decreases. Uncontrolled growth of the mutated cells outpaces the growth of normal cells in the area, and a tumor ("-oma") can result.

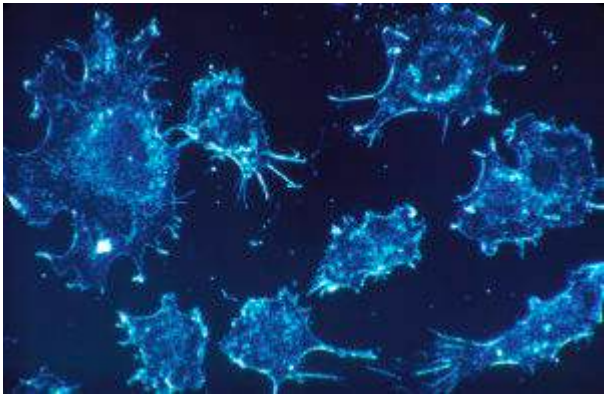


Figure 1 Cancer cells in culture from human connective tissue, illuminated by darkfield amplified contrast, at a magnification of 500x.

PROTO-ONCOGENES

The genes that code for the positive cell cycle regulators are called **proto-oncogenes**. Proto-oncogenes are normal genes that, when mutated in certain ways, become **oncogenes,** genes that cause a cell to become cancerous. Consider what might happen to the cell cycle in a cell with a recently acquired oncogene. In most instances, a mutation in the DNA sequence of a gene will result in a less functional or non-functional protein. This result is detrimental to the cell and will likely prevent the cell from completing the cell cycle, which means that this cell cannot create daughter cells. In this case, the organism is not harmed because the mutation will not be carried forward and the damage is minimal.

Occasionally, however, a gene mutation causes a change that increases the activity of a positive regulator. For

example, a mutation that allows Cdk to be activated without being partnered with cyclin could push the cell cycle past a checkpoint before all of the required conditions are met. If the resulting daughter cells are too damaged to undergo further cell divisions, the mutation would not be propagated and no harm would come to the organism. However, if the atypical daughter cells are able to undergo further cell divisions, subsequent generations of cells will probably accumulate even more mutations, some possibly in additional genes that regulate the cell cycle.

The Cdk gene in the above example is only one of many genes that are considered proto-oncogenes. In addition to the cell cycle regulatory proteins, any protein that influences the cycle can be altered in such a way as to override cell cycle checkpoints. An oncogene is any gene that, when altered, leads to an increase in the rate of cell cycle progression.

TUMOR SUPPRESSOR GENES

Like proto-oncogenes, many of the negative cell cycle regulatory proteins were discovered in cells that had become cancerous. **Tumor suppressor genes** are segments of DNA that code for negative regulator proteins. Activated negative regulator proteins prevent the cell from undergoing uncontrolled division. The collective function of the best-understood tumor suppressor gene proteins, Rb, p53, and p21, is to put up a roadblock to cell cycle progression until certain events are completed. A cell that carries a mutated form of a negative regulator might not be able to halt the cell cycle if there is a problem. Tumor suppressors are similar to brakes in a vehicle: Malfunctioning brakes can contribute to a car crash.

Mutated p53 genes have been identified in more than one-half of all human tumor cells. This discovery is not surprising

in light of the multiple roles that the p53 protein plays at the G_1 checkpoint. A cell with a faulty p53 may fail to detect errors present in the genomic DNA (**Figure 2**). Even if a partially functional p53 does identify the mutations, it may no longer be able to signal the necessary DNA repair enzymes. Either way, damaged DNA will remain uncorrected. At this point, a functional p53 will deem the cell unsalvageable and trigger programmed cell death (apoptosis). The damaged version of p53 found in cancer cells, however, cannot trigger apoptosis.

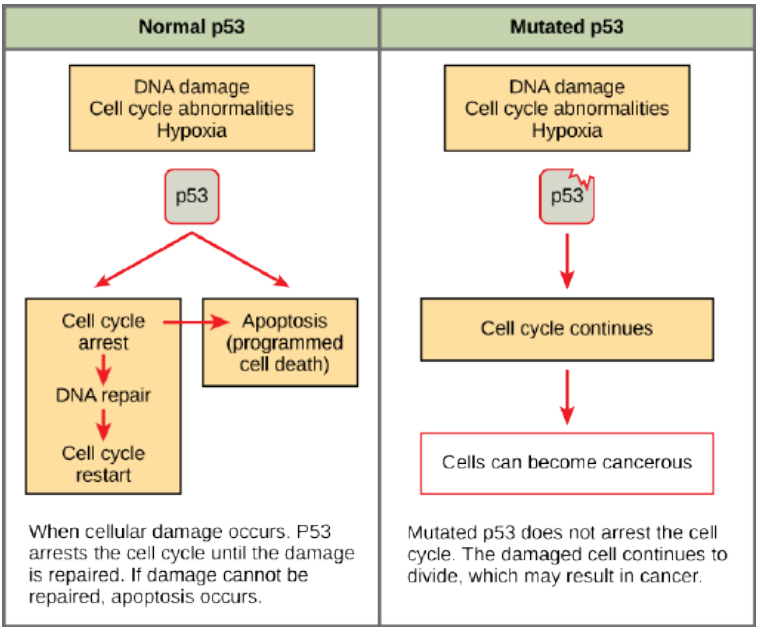


Figure 2 The role of normal p53 is to monitor DNA and the supply of oxygen (hypoxia is a condition of reduced oxygen supply). If damage is detected, p53 triggers repair mechanisms. If repairs are unsuccessful, p53 signals apoptosis. A cell with an abnormal p53 protein cannot repair damaged DNA and thus cannot signal apoptosis. Cells with abnormal p53 can become cancerous. (credit: modification of work by Thierry Soussi)

REFERENCES

Unless otherwise noted, images on this page are licensed under CC-BY 4.0 by OpenStax.

OpenStax, Biology. OpenStax CNX. November 11, 2017 https://cnx.org/contents/GFy_h8cu@10.118:CYZpmedR@7/Cancer-and-the-Cell-Cycle

BIOLOGY 212 - GENETICS

The Principles of Biology sequence (BI 211, 212, & 213) introduces biology as a scientific discipline for students planning to major in biology and other science disciplines. Laboratories and classroom activities introduce techniques used to study biological processes and provide opportunities for students to develop their ability to conduct research. BI212 uses genetics as a model system to understand information flow in living organisms.

Course Outcomes: Upon successful completion of this course, students should be able to

1. Apply the scientific method to biological questions by designing experiments and using the resulting data to form and communicate a conclusion.
    1. Design a controlled experiment to answer a biological question.
    2. Predict the outcome of an experiment.
    3. Collect, manipulate, and analyze quantitative and qualitative data.
    4. Answer a biological question using data.
    5. Use an appropriate written format to

present scientific information.

6. Use appropriate biological terminology to answer written and oral questions.
7. Orally present results from an experiment.

2. Assess the strengths and weaknesses of the design of and conclusions drawn from scientific studies.
    1. Differentiate between questions that can and cannot be answered using science.
    2. Identify appropriate credible sources of information to research a topic.
    3. Evaluate sources of information for their strengths and weaknesses.
3. Select, evaluate and utilize discipline-specific information and literature to research a biological topic.
    1. Evaluate the strengths and weaknesses of their own as well as published experiments.
4. Use evidence to develop informed opinions on contemporary biological issues while considering cultural and ethical implications.
    1. Research current ethical issues in genetics and biotechnology.
    2. Form opinions based on published scientific research.
5. Apply biological theories and concepts to solve problems related to classical and molecular genetics.
    1. Describe the molecular basis of inheritance.

2. Determine the outcome in crosses involving various types of inheritance (e.g. simple dominance, codominance, incomplete dominance, sex-linkage).
3. Present and decipher information about trait information using a pedigree.
4. Discuss the possible evolutionary consequences of various types of inheritance.

6. Discuss the potential implications of mutations at the cellular, organismal, and evolutionary levels.
    1. Describe the structure of DNA and the process of DNA replication.
    2. Summarize the processes involved in protein synthesis.
    3. Describe how mutations affect the process of protein synthesis and its products.
    4. Discuss the possible evolutionary consequences of mutations.
7. Describe the purpose of the regulation of gene expression and the mechanisms by which gene expression is regulated.
    1. Describe processes through which gene expression can be regulated.
    2. Differentiate between gene regulation processes used by prokaryotes and eukaryotes.
    3. Discuss the possible evolutionary consequences of changes in gene expression.

DNA AND CHROMOSOME STRUCTURE

Learning Outcomes

- Discuss the potential implications of mutations at cellular, organismal, and evolutionary levels
    - Describe the structure of DNA and the process of DNA replication.

DNA is a nucleic acid, which is one of the four biological macromolecules that you began learning about in BI211. Recall that nucleic acids are made up of monomers called nucleotides joined together by strong covalent bonds.

All living things store genetic information using nucleic acids: either DNA or a related molecule called RNA. Different organisms have different ways of packaging and storing DNA inside their cells.

DNA Structure

In the 1950s, Francis Crick and James Watson worked together to determine the structure of DNA at the University of Cambridge, England. Other scientists like Linus Pauling and Maurice Wilkins were also actively exploring this field. Pauling had discovered the secondary structure of proteins using X-ray crystallography. In Wilkins' lab, researcher Rosalind Franklin was using X-ray diffraction methods to understand the structure of DNA. Watson and Crick were able to piece together the puzzle of the DNA molecule on the basis of Franklin's data because Crick had also studied X-ray diffraction (Figure 1).

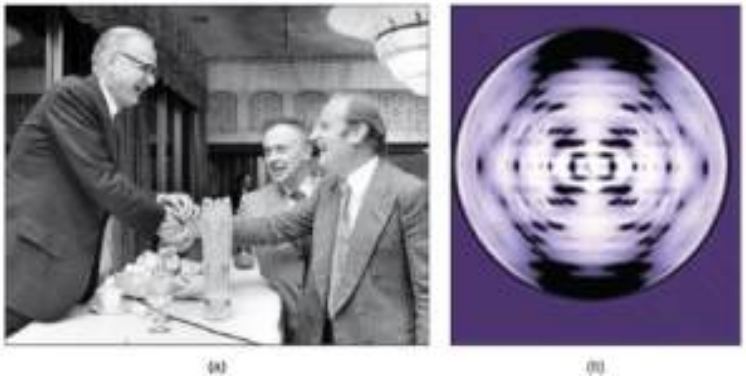


Figure 1 The work of pioneering scientists (a) James Watson, Francis Crick, and Maclyn McCarty led to our present day understanding of DNA. Scientist Rosalind Franklin discovered (b) the X-ray diffraction pattern of DNA, which helped to elucidate its double helix structure. (credit a: modification of work by Marjorie McCarty, Public Library of Science)

Unfortunately, Watson and Crick gained access to Franklin's data without her knowledge or approval. In 1962, James Watson, Francis Crick, and Maurice Wilkins were awarded the Nobel Prize in Medicine. Unfortunately, by then Franklin had died (of ovarian cancer, likely caused by exposure to X-rays), and Nobel prizes are not awarded posthumously (after death). This is actually a really interesting story of "sexism in the sciences" – there's a movie called "The Secret of Photo 51" that you can find on YouTube if you're interested.

Based on Rosalind Franklin's X-ray diffraction photograph, and work by other scientists, Watson and Crick proposed that DNA is made up of two strands of **nucleotides** that are twisted around each other to form a right-handed helix. The nucleotides are joined together in a chain by covalent bonds known as **phosphodiester** bonds. Scientists already knew that nucleotides contain the same three important

components: a **nitrogenous base**, a **deoxyribose** (5-carbon sugar), and a **phosphate group** (Figure 2). The nucleotide is named depending on the nitrogenous base: adenine (A), thymine (T), cytosine (C), and guanine (G). Adenine and guanine are both **purines**, while cytosine and thymine are **pyrimidines.** The purines have a double ring structure with a six-membered ring fused to a five-membered ring. Pyrimidines are smaller in size; they have a single six-membered ring structure. One good way to remember this is that cytosine, thymine, and pyrimidine all contain the letter "y".

Watson and Crick's model proposed that the two strands of nucleotides interact through base pairing between the nucleotides: A pairs with T and G pairs with C. Adenine and thymine are **complementary base pairs**, and cytosine and guanine are also complementary base pairs. The base pairs are stabilized by **hydrogen bonds** (a weak type of bond that forms between partially positive and partially negative atoms). Adenine and thymine form two hydrogen bonds and cytosine and guanine form three hydrogen bonds. Since a purine is "2 rings" across and a pyrimidine is "1 ring" across (Figure 2), the diameter of **the DNA double helix** remains constant at "3 rings" (Figure 3).

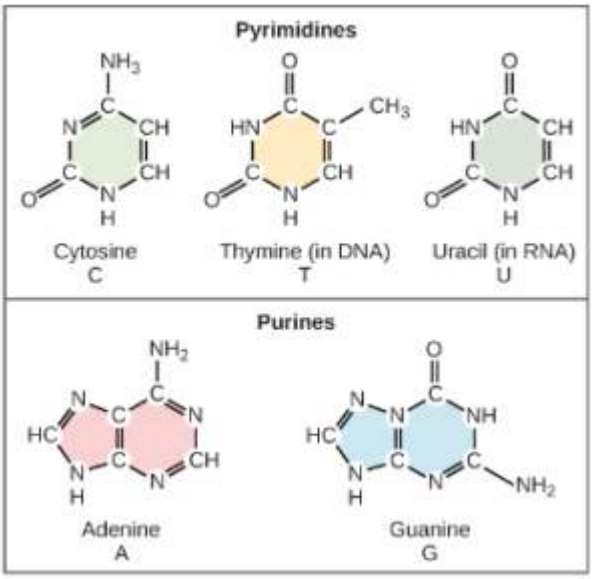


Figure 2: Each nucleotide is made up of a sugar, a phosphate group, and a nitrogenous base. The sugar is deoxyribose in DNA and ribose in RNA.

The carbon atoms of the five-carbon sugar are numbered in order starting from the carbon connected to the nitrogenous base: 1', 2', 3', 4', and 5' (1' is read as "one prime"). We don't particularly care about the 1', 2', or 4' positions – you'll never hear them mentioned again. At the 3' position, there is always a hydroxyl (OH) group that is a part of the sugar. The 5' carbon is attached to a phosphate group. When nucleotides are joined together into a chain, the 5' phosphate of one nucleotide is attached to the 3' hydroxyl group of the next nucleotide, thereby forming a 5'-3' **phosphodiester bond.** What this means is that when nucleotides are joined together in a chain, there will always be a free 3' OH group (from the sugar) at one end of the chain and a free 5' phosphate at the other end (Figure 3).

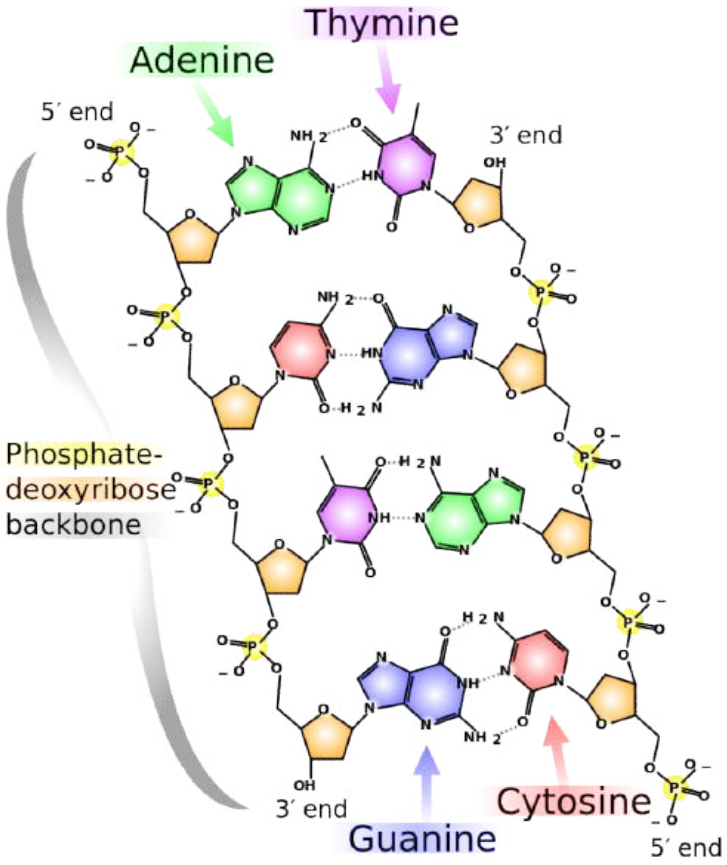


Figure 3 Structure of DNA. Notice that adenine (a purine) and thymine (a pyrimidine) are connected together with 2 hydrogen bonds, while guanine (a purine) and cytosine (a pyrimidine) are connected by three hydrogen bonds. There is a 5' and 3' end to both chains of nucleotides, which are antiparallel to each other. Photo credit Madeline Price Ball; Wikimedia.

The two strands are **anti-parallel** in nature; that is, the 3' end of one strand points in one direction, while the 5' end of the other strand points in that direction (Figure 3). The sugar and phosphate of the nucleotides form the backbone of the structure, while the nitrogenous bases are stacked

inside. Each base pair is separated from the other base pair by a distance of 0.34 nm (nanometer: 1 x 10^{-9} meters), and each turn of the helix measures 3.4 nm. Therefore, ten base pairs are present per turn of the helix. The diameter of the DNA double helix is 2 nm, and it is uniform throughout. Only the pairing between a purine and pyrimidine can explain the uniform diameter (3 "rings" across). The twisting of the two strands around each other results in the formation of uniformly spaced major and minor grooves (Figure 4). The major and minor grooves are very important for protein binding to DNA, but we will not be discussing this more specifically.

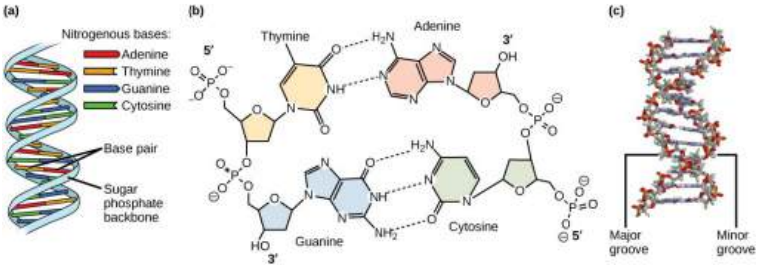


Figure 4 DNA has (a) a double helix structure and (b) phosphodiester bonds. The (c) major and minor grooves are binding sites for DNA binding proteins during processes such as transcription (the copying of RNA from DNA) and replication.

REFERENCES

OpenStax, Biology. OpenStax CNX. December 21, 2017 http://cnx.org/contents/s8Hh0oOc@9.10:QhGQhr4x@6/Biological-Molecules

DNA organization inside a cell

DNA ORGANIZATION IN PROKARYOTES

A cell's DNA, packaged as a double-stranded DNA molecule, is called its **genome.** In prokaryotes, the genome is composed of a single, double-stranded DNA molecule in the form of a loop or circle (Figure 1). The region in the cell containing this genetic material is called a **nucleoid** (remember that prokaryotes do not have a separate membrane-bound nucleus). Some prokaryotes also have smaller loops of DNA called **plasmids** that are not essential for normal growth. Bacteria can exchange these plasmids with other bacteria, sometimes receiving beneficial new genes that the recipient can add to their chromosomal DNA. Antibiotic resistance is one trait that often spreads through a bacterial colony through plasmid exchange.

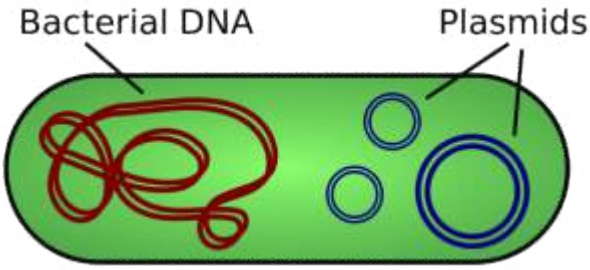


Figure 1 Bacterial DNA and plasmids are both circular. Photo credit Spaully; Wikipedia.

The size of the genome in one of the most well-studied prokaryotes, *E.coli,* is 4.6 million base pairs (which would be approximately 1.1 mm in length, if cut and stretched out). So how does this fit inside a small bacterial cell? The DNA is twisted by what is known as supercoiling. Supercoiled DNA is coiled more tightly than would be typically be found in a cell (more than 10 nucleotides per twist of the helix). If you visualize twisting a rope until it twists back on itself, you have a pretty good visual of supercoiled DNA. This process allows the DNA to be compacted into the small space inside a bacteria.

DNA ORGANIZATION IN EUKARYOTES

Eukaryotes have much more DNA than prokaryotes. For example, an *E. coli* bacteria contains roughly 3 million base pairs of DNA, while a human contains roughly 3 *billion.* In eukaryotes such as humans and other animals, the genome consists of several double-stranded linear DNA molecules (Figure 2), which are located inside a membrane-bound nucleus. Each species of eukaryotes has a characteristic number of chromosomes in the nuclei (plural of nucleus) of its cells. A normal human **gamete** (sperm or egg) contains 23 chromosomes. A normal human body cell, or **somatic**

cell, contains 46 chromosomes (one set of 23 from the egg and one set of 23 from the sperm; Figure 2). The letter *n* is used to represent a single set of chromosomes; therefore, a gamete (sperm or egg) is designated *1n*, and is called a **haploid** cell. Somatic cells (body cells) are designated 2*n* and are called **diploid** cells.

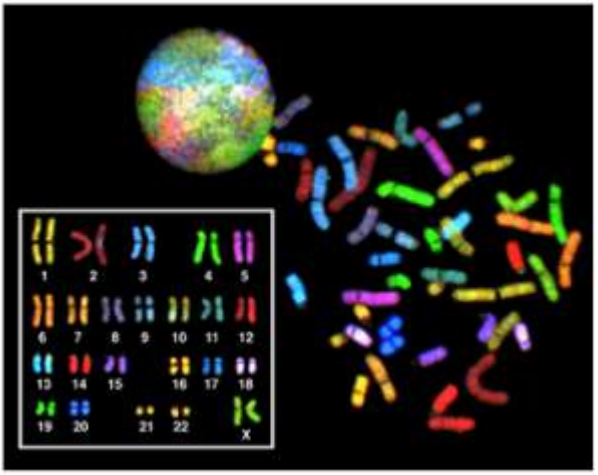


Figure 2 There are 23 pairs of homologous chromosomes in a female human somatic cell. The condensed chromosomes are viewed within the nucleus (top), removed from a cell in mitosis and spread out on a slide (right), and artificially arranged according to length (left); an arrangement like this is called a karyotype. In this image, the chromosomes were exposed to fluorescent stains for differentiation of the different chromosomes. A method of staining called "chromosome painting" employs fluorescent dyes that highlight chromosomes in different colors. (credit: National Human Genome Project/NIH)

Matched pairs of chromosomes in a diploid organism are called **homologous** ("same knowledge") chromosomes. Of a pair of homologous chromosomes, one came from the egg and the second came from the sperm. Homologous chromosomes are the same length and have specific nucleotide segments called **genes** in exactly the same

location, or **locus.** Genes, the functional units of chromosomes, determine specific characteristics by coding for specific proteins. Traits are the variations of those characteristics. For example, hair color is a characteristic with traits that are blonde, brown, or black.

Each copy of a homologous pair of chromosomes originates from a different parent; therefore, the sequence of DNA present in the two genes on a pair of homologous chromosomes is not necessarily identical, despite the same genes being present in the same locations on the chromosome. These different versions of a gene that contain different sequences of DNA are called **alleles.**

If you look at Figure 3, you can see a pair of homologous chromosomes. The chromosome shown in the figure is chromosome 15. The HERC2 gene is located on this chromosome. This gene is one of at least three genes that helps determine eye color. Each person inherits two copies of the HERC2 gene: one from the egg and one from the sperm. However, the alleles of the HERC2 gene that they inherit can be different. In the figure, the cell containing this homologous pair of chromosomes contains one blue allele and one brown allele.

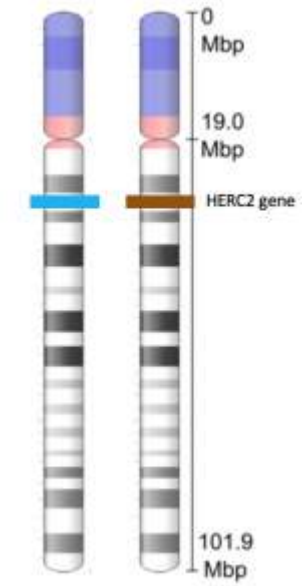


Figure 3 Homologous pair of chromosome 15s, showing the location of HERC2 gene. Two different alleles of this gene are shown in either blue or brown.

The variation of individuals within a species is due to the specific combination of the genes inherited from both parents. Even a slightly altered sequence of nucleotides within a gene can result in an alternative trait. For example, there are three possible gene sequences (alleles) on the human chromosome that code for blood type: sequence A, sequence B, and sequence O. Because all diploid human cells have two copies of the chromosome that determines blood type, the blood type (the trait) is determined by which two versions of the marker gene are inherited. It is possible to have two copies of the same allele on both homologous chromosomes, with one on each (for example, AA, BB, or OO), or two different alleles, such as AB.

Minor variations of traits, such as blood type, eye color, and handedness, contribute to the natural variation found

within a species. However, if the entire DNA sequence from any pair of human homologous chromosomes is compared, the difference is less than one percent. The sex chromosomes, X and Y, are the single exception to the rule of homologous chromosome uniformity: Other than a small amount of homology that is necessary to accurately produce gametes, the genes found on the X and Y chromosomes are different.

EUKARYOTIC CHROMOSOMAL STRUCTURE AND COMPACTION

If the DNA from all 46 chromosomes in a human cell nucleus was laid out end to end, it would measure approximately two meters; however, its diameter would be only 2 nm. Considering that the size of a typical human cell is about 10 μm (100,000 cells lined up to equal one meter), DNA must be tightly packaged to fit in the cell's nucleus. At the same time, it must also be readily accessible for the genes to be expressed. During some stages of the cell cycle, the long strands of DNA are condensed into compact chromosomes.

Eukaryotes, whose chromosomes each consist of a linear DNA molecule, employ a complex type of packing strategy to fit their DNA inside the nucleus (Figure 4). At the most basic level, DNA is wrapped around proteins known as **histones** to form structures called **nucleosomes**. The histones are evolutionarily conserved proteins that form an octamer of eight histone proteins attached together. DNA, which is negatively charged because of the phosphate groups, is wrapped tightly around the histone core, which has an overall positive charge. This nucleosome is linked to the next one with the help of a linker DNA. This is also known as the "beads on a string" structure. This is further compacted into a 30 nm fiber, which is the diameter of the structure. At the metaphase stage, the chromosomes are at their most

compact, are approximately 700 nm in width, and are found in association with scaffold proteins.

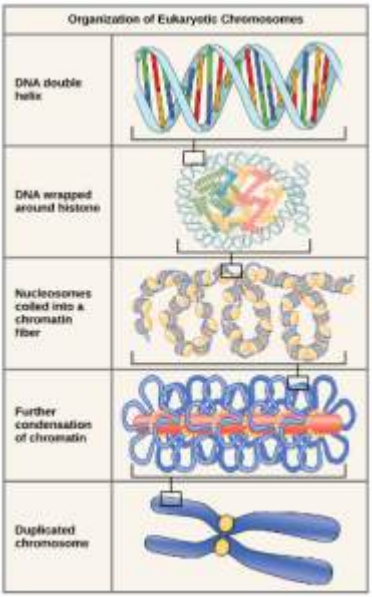


Figure 4 Double-stranded DNA wraps around histone proteins to form nucleosomes that have the appearance of "beads on a string." The nucleosomes are coiled into a 30-nm chromatin fiber. When a cell undergoes mitosis, the duplicated chromosomes condense even further.

DNA replicates in the S phase of interphase. After replication, the chromosomes are composed of two linked **sister chromatids (Figure 5)**. This means that the only time chromosomes look like an "X" is after DNA replication has taken place and the chromosomes have condensed. During the majority of the cell's life, chromosomes are composed of only one copy and they are not tightly compacted into chromosomes. When fully compact, the pairs of identically packed chromosomes are bound to each other by cohesin proteins. The connection between the sister chromatids is closest in a region called the **centromere.** The conjoined

sister chromatids, with a diameter of about 1 μm, are visible under a light microscope. The centromeric region is highly condensed and thus will appear as a constricted area. In Figure 4, it is shown as an oval because it is easier to draw that way.

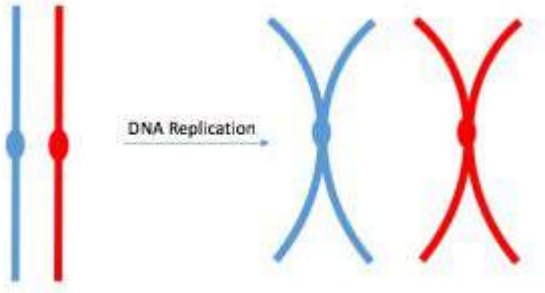


Figure 5 Result of DNA replication in eukaryotes. In body cells, there are two copies of each chromosome (one from each parent). After replication, two identical sister chromatids remain connected at the centromere (shown as an oval).

REFERENCES

OpenStax, Biology. OpenStax CNX. December 21, 2017 https://cnx.org/contents/GFy_h8cu@10.120:U7tPDRxK@9/DNA-Structure-and-Sequencing

Homologous Chromosomes and Sexual Reproduction

It's very important to keep chromosome number in context as we discuss genetics and inheritance. As discussed in the last section, the somatic cells of a human are diploid (2n; contain two copies of each chromosome, creating homologous pairs). In Figure 1, both the adult humans are diploid for chromosome 15, along the other 22 pairs of chromosomes in humans, which are not shown for simplicity. When gametes (eggs or sperm) are created, those cells are haploid (1n) and contain only one copy of each chromosome. During fertilization, the male and female gametes fuse together creating a fertilized egg, or **zygote**, which again contains two copies of each chromosome and is therefore diploid. The zygote goes through millions of cell divisions to eventually become a baby and then an adult.

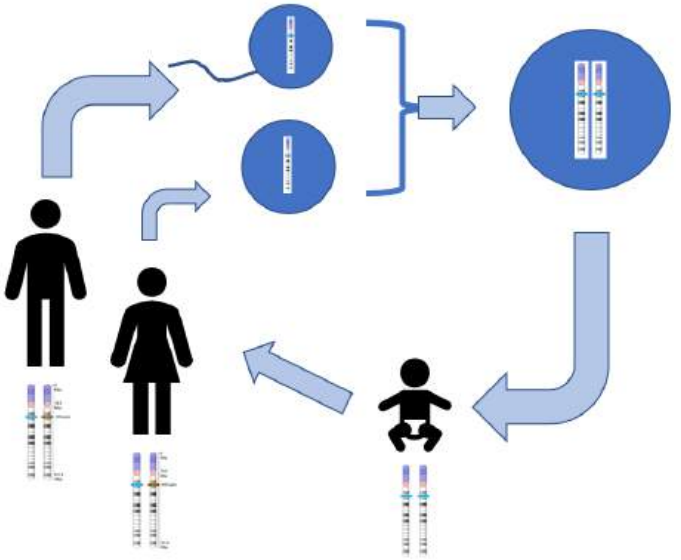


Figure 1 Sexual life cycle of humans.

DNA REPLICATION

Learning Outcomes

1. Discuss the potential implications of mutations at cellular, organismal, and evolutionary levels
1. Describe the structure of DNA and the process of DNA replication.

When a cell divides, it is important that each daughter cell receives an identical copy of the DNA. This is accomplished by the process of **DNA replication**. The replication of DNA occurs before the cell begins to divide into two separate cells.

The discovery and characterization of the structure of the double helix provided a hint as to how DNA is copied. Recall that adenine nucleotides pair with thymine nucleotides, and cytosine with guanine. This means that the two strands are **complementary** to each other. For example, a strand of DNA with a nucleotide sequence of 3'-AGTCATGA-5' will have a complementary strand with the sequence 5'-TCAGTACT-3' (**Figure 1**). Recall that the two strands of the double helix are

anti-parallel as you look at the order of bases in these two sequences.

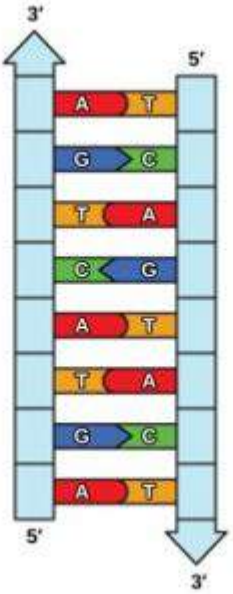


Figure 1: The two strands of DNA are complementary, meaning the sequence of bases in one strand can be used to create the correct sequence of bases in the other strand.

Because of the complementarity of the two strands, having one strand means that it is possible to recreate the other strand. During DNA replication, the two strands are separated and the missing bases are filled in on each side (**Figure 2**). Each strand of DNA serves as a template to produce the missing strand. The new strand will be complementary to the parental or "old" strand. Each new double helix consists of one parental strand and one new daughter strand. This is known as **semiconservative replication**. When two DNA copies are formed, they have an identical sequence of nucleotide bases and are divided equally into two daughter cells.

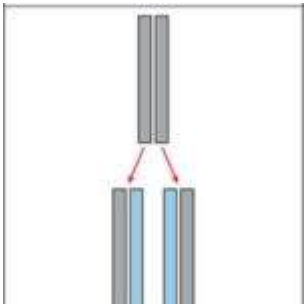


Figure 2: The semiconservative model of DNA replication is shown. Gray indicates the original DNA strands, and blue indicates newly synthesized DNA.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:2ousESf0@5/DNA-Replication

DNA Replication in Prokaryotes

The prokaryotic chromosome is a circular molecule with a less extensive coiling structure than eukaryotic chromosomes. The eukaryotic chromosome is linear and highly coiled around proteins. While there are many similarities in the DNA replication process, these structural differences necessitate some differences in the DNA replication process in these two life forms. DNA replication in prokaryotes has been extensively studied, so we will learn the basic process of prokaryotic DNA replication, then focus on the differences between prokaryotes and eukaryotes.

How does the replication machinery know where to start? It turns out that there are specific nucleotide sequences called **origins of replication** where replication begins. *E. coli* has a single origin of replication on its one chromosome, as do most prokaryotes (**Figure 1**). The origin of replication is approximately 245 base pairs long and is rich in AT sequences. This sequence of base pairs is recognized by certain proteins that bind to this site. An enzyme called **helicase** unwinds the DNA by breaking the hydrogen bonds between the nitrogenous base pairs. ATP hydrolysis is required for this process because it requires energy. As the DNA opens up, Y-shaped structures called **replication forks** are formed (**Figure 1**). Two replication forks are formed at

the origin of replication and these get extended bi-directionally as replication proceeds. **Single-strand binding proteins** (Figure 2) coat the single strands of DNA near the replication fork to prevent the single-stranded DNA from winding back into a double helix.

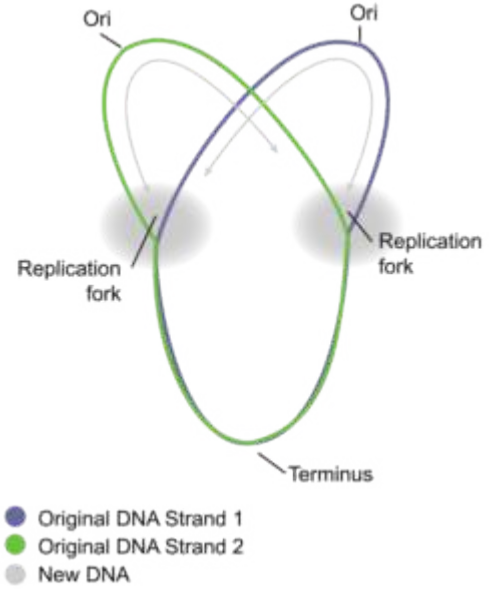


Figure 1: DNA replication in prokaryotes, which have one circular chromosome.

The next important enzyme is **DNA polymerase III**, also known as DNA pol III, which adds nucleotides one by one to the growing DNA chain (Figure 2). The addition of nucleotides requires energy; this energy is obtained from the nucleotides that have three phosphates attached to them. ATP structurally is an adenine nucleotide which has three phosphate groups attached; breaking off the third phosphate releases energy. In addition to ATP, there are also TTP, CTP, and GTP. Each of these is made up of the corresponding nucleotide with three phosphates attached. When the bond between the phosphates is broken, the

energy released is used to form the phosphodiester bond between the incoming nucleotide and the existing chain.

In prokaryotes, three main types of polymerases are known: DNA pol I, DNA pol II, and DNA pol III. DNA pol III is the enzyme required for DNA synthesis; DNA pol I is used later in the process and DNA pol II is used primarily required for repair (this is another irritating example of naming that was done based on the order of discovery rather than an order that makes sense).

DNA polymerase is able to add nucleotides only in the 5' to 3' direction (a new DNA strand can be only extended in this direction). It requires a free 3'-OH group (located on the sugar) to which it can add the next nucleotide by forming a phosphodiester bond between the 3'-OH end and the 5' phosphate of the next nucleotide. This essentially means that it cannot add nucleotides if a free 3'-OH group is not available. Then how does it add the first nucleotide? The problem is solved with the help of a **primer** that provides the free 3'-OH end. Another enzyme, **RNA primase**, synthesizes an RNA primer that is about five to ten nucleotides long and complementary to the DNA. RNA primase does not require a free 3'-OH group. Because this sequence primes the DNA synthesis, it is appropriately called the primer. DNA polymerase can now extend this RNA primer, adding nucleotides one by one that are complementary to the template strand (**Figure 2**).

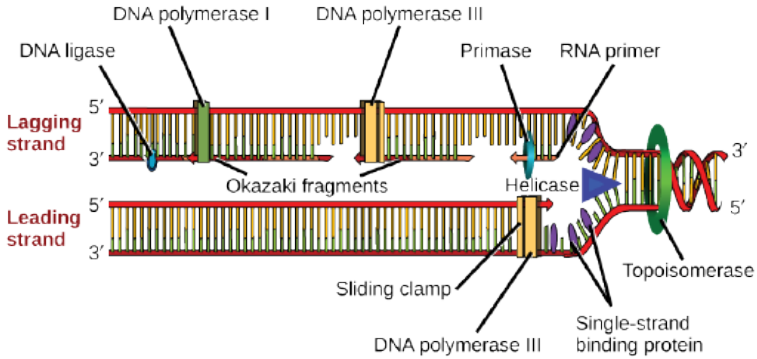


Figure 2 A replication fork is formed when helicase separates the DNA strands at the origin of replication. The DNA tends to become more highly coiled ahead of the replication fork. Topoisomerase breaks and reforms DNA's phosphate backbone ahead of the replication fork, thereby relieving the pressure that results from this supercoiling. Single-strand binding proteins bind to the single-stranded DNA to prevent the helix from re-forming. Primase synthesizes an RNA primer. DNA polymerase III uses this primer to synthesize the daughter DNA strand. On the leading strand, DNA is synthesized continuously, whereas on the lagging strand, DNA is synthesized in short stretches called Okazaki fragments. DNA polymerase I replaces the RNA primer with DNA. DNA ligase seals the gaps between the Okazaki fragments, joining the fragments into a single DNA molecule. (credit: modification of work by Mariana Ruiz Villareal)

The replication fork moves at the rate of 1000 nucleotides per second. DNA polymerase can only extend in the 5' to 3' direction, which poses a slight problem at the replication fork. As we know, the DNA double helix is anti-parallel; that is, one strand is in the 5' to 3' direction and the other is oriented in the 3' to 5' direction. One strand, which is complementary to the 3' to 5' parental DNA strand, is synthesized continuously towards the replication fork because the polymerase can add nucleotides in this direction. This continuously synthesized strand is known as the **leading strand**. The other strand, complementary to the 5' to 3'

parental DNA, is extended away from the replication fork, in small fragments known as **Okazaki fragments**, each requiring a primer to start the synthesis. Okazaki fragments are named after the Japanese scientist who first discovered them. The strand with the Okazaki fragments is known as the **lagging strand**.

The leading strand can be extended by one primer alone, whereas the lagging strand needs a new primer for each of the short Okazaki fragments. The overall direction of the lagging strand will be 3' to 5', and that of the leading strand 5' to 3'. A protein called the **sliding clamp** holds the DNA polymerase in place as it continues to add nucleotides. The sliding clamp is a ring-shaped protein that binds to the DNA and holds the polymerase in place. **Topoisomerase** prevents the over-winding of the DNA double helix ahead of the replication fork as the DNA is opening up; it does so by causing temporary nicks in the DNA helix and then resealing it. As synthesis proceeds, the RNA primers are replaced by DNA pol I, which breaks down the RNA and fills the gaps with DNA nucleotides. The nicks that remain between the newly synthesized DNA (that replaced the RNA primer) and the previously synthesized DNA are sealed by the enzyme **DNA ligase** that catalyzes the formation of phosphodiester linkage between the 3'-OH end of one nucleotide and the 5' phosphate end of the other fragment.

(Lisa's note: I think this process is almost impossible to visualize from reading text. I strongly recommend that you watch a couple of animations / videos like the one available here. There are additional links in Blackboard)

Once the chromosome has been completely replicated, the two DNA copies move into two different cells during cell division. The process of DNA replication can be summarized as follows:

1. DNA unwinds at the origin of replication.

2. Helicase opens up the DNA-forming replication forks; these are extended in both directions.
3. Single-strand binding proteins coat the DNA around the replication fork to prevent rewinding of the DNA.
4. Topoisomerase binds at the region ahead of the replication fork to prevent supercoiling (over-winding).
5. Primase synthesizes RNA primers complementary to the DNA strand.
6. DNA polymerase III starts adding nucleotides to the 3'-OH (sugar) end of the primer.
7. Elongation of both the lagging and the leading strand continues.
8. RNA primers are removed and gaps are filled with DNA by DNA pol I.
9. The gaps between the DNA fragments are sealed by DNA ligase.

Table 1: The enzymes involved in prokaryotic DNA replication and the functions of each.

Prokaryotic DNA Replication: Enzymes and Their Function

Enzyme/protein	Specific Function
DNA pol I	Exonuclease activity removes RNA primer and replaces with newly synthesized DNA
DNA pol II	Repair function
DNA pol III	Main enzyme that adds nucleotides in the 5'-3' direction
Helicase	Opens the DNA helix by breaking hydrogen bonds between the nitrogenous bases
Ligase	Seals the gaps between the Okazaki fragments to create one continuous DNA strand
Primase	Synthesizes RNA primers needed to start replication
Sliding Clamp	Helps to hold the DNA polymerase in place when nucleotides are being added
Topoisomerase	Helps relieve the stress on DNA when unwinding by causing breaks and then resealing the DNA
Single-strand binding proteins (SSB)	Binds to single-stranded DNA to avoid DNA rewinding back.

DNA replication has been extremely well-studied in prokaryotes, primarily because of the small size of the genome and large number of variants available. *Escherichia coli* has 4.6 million base pairs in a single circular chromosome, and all of it gets replicated in approximately 42 minutes, starting from a single origin of replication and proceeding around the chromosome in both directions. This means that approximately 1000 nucleotides are added per second. The process is much more rapid than in eukaryotes.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:2ousESf0@5/DNA-Replication

DNA Replication in Eukaryotes

The essential steps of replication are the same as in prokaryotes. Starting replication is more complex in eukaryotes. At the origin of replication, a pre-replication complex is made with other initiator proteins. Other proteins are then recruited to start the replication process. The overall process is the same, although differently named enzymes fulfill the same function. For example, DNA pol III is used for the majority of replication in prokaryotes, while in eukaryotes the leading strand is continuously synthesized by the enzyme pol *δ*, the lagging strand is synthesized by pol *ε*. We are focusing on the enzymes used in prokaryotic replication, so don't worry about these name differences.

Here are the important differences between prokaryotic and eukaryotic replication:

Table 1: Differences between prokaryotic and eukaryotic replication

Property	Prokaryotes	Eukaryotes
Origin of replication	Single	Multiple
Rate of replication	1000 nucleotides/s	50 to 100 nucleotides/s
DNA polymerase types	5	14
DNA packaging	supercoiling	wound around histones
Telomerase	Not present	Present

ORIGINS AND RATE OF REPLICATION

Eukaryotic genomes are much more complex and larger in size than prokaryotic genomes. The human genome has three billion base pairs per haploid set of chromosomes, and 6 billion base pairs are replicated during the S phase of the cell cycle. This means that there must be multiple origins of replication on the eukaryotic chromosome in order for all the DNA to be replicated in a timely manner; humans can have up to 100,000 origins of replication. The rate of replication is approximately 100 nucleotides per second, much slower than prokaryotic replication.

DNA POLYMERASE TYPES

The number of DNA polymerases in eukaryotes is much more than prokaryotes: 14 are known, of which five are known to have major roles during replication and have been well studied. They are known as pol α, pol β, pol γ, pol δ, and pol ε. I won't ever ask you the names of these polymerases – learn the names of the prokaryotic polymerases.

DNA PACKAGING

Eukaryotic DNA is wound around proteins known as histones to form structures called nucleosomes. The DNA must be made accessible in order for DNA replication to proceed.

The chromatin (the complex between DNA and proteins) may undergo some chemical modifications, so that the DNA may be able to slide off the histones or otherwise be accessible to the enzymes of the DNA replication machinery. Prokaryotes do not packaged their DNA by wrapping it around histones.

TELOMERE REPLICATION

Unlike prokaryotic chromosomes, eukaryotic chromosomes are linear. As you've learned, the enzyme DNA pol can add nucleotides only in the 5' to 3' direction. In the leading strand, synthesis continues until the end of the chromosome is reached. On the lagging strand, DNA is synthesized in short stretches, each of which is initiated by a separate primer. When the replication fork reaches the end of the linear chromosome, there is no place for a primer to be made for the DNA fragment to be copied at the end of the chromosome. These ends thus remain unpaired, and over time these ends may get progressively shorter as cells continue to divide.

The ends of the linear chromosomes are known as **telomeres,** which have repetitive sequences that do not code for a particular gene. These telomeres protect the genes that are located on the chromosome from getting deleted as cells continue to divide. In humans, a six base pair sequence, TTAGGG, is repeated 100 to 1000 times. The discovery of the enzyme **telomerase** (**Figure 1**) helped in the understanding of how chromosome ends are maintained. The telomerase enzyme contains a catalytic part and a built-in RNA template. It attaches to the end of the chromosome, and complementary bases to the RNA template are added on the 3' end of the DNA strand. Once the 3' end of the lagging strand template is sufficiently elongated, DNA polymerase can add the nucleotides complementary to the ends of the

chromosomes. Thus, the ends of the chromosomes are replicated.

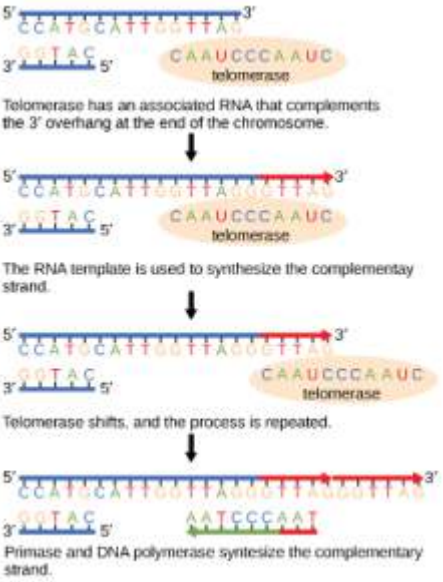


Figure 1 The ends of linear chromosomes are maintained by the action of the telomerase enzyme.

Telomerase is typically active in germ cells and adult stem cells. It is not active in adult somatic cells. For her discovery of telomerase and its action, Elizabeth Blackburn (**Figure 2**) received the Nobel Prize for Medicine and Physiology in 2009.

Figure 2: Elizabeth Blackburn, 2009 Nobel Laureate, is the scientist who discovered how telomerase works. (credit: US Embassy Sweden)

Telomerase and Aging

Cells that undergo cell division continue to have their telomeres shortened because most somatic cells do not make telomerase. This essentially means that telomere shortening is associated with aging. With the advent of modern medicine, preventative health care, and healthier lifestyles, the human life span has increased, and there is an increasing demand for people to look younger and have a better quality of life as they grow older.

In 2010, scientists found that telomerase can reverse some age-related conditions in mice. This may have potential in regenerative medicine (Jaskelioff, 2011). Telomerase-deficient mice were used in these studies; these mice have tissue atrophy, stem cell depletion, organ system failure, and impaired tissue injury responses. Telomerase reactivation in these mice caused extension of telomeres, reduced DNA damage, reversed neurodegeneration, and improved the function of the testes, spleen, and intestines. Thus, telomere

reactivation may have potential for treating age-related diseases in humans.

Cancer is characterized by uncontrolled cell division of abnormal cells. The cells accumulate mutations, proliferate uncontrollably, and can migrate to different parts of the body through a process called metastasis. Scientists have observed that cancerous cells have considerably shortened telomeres and that telomerase is active in these cells. Interestingly, only after the telomeres were shortened in the cancer cells did the telomerase become active. If the action of telomerase in these cells can be inhibited by drugs during cancer therapy, then the cancerous cells could potentially be stopped from further division.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:2ousESf0@5/DNA-Replication

Jaskelioff et al., 2011 Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice. Nature 469: 102-7.

MUTATIONS

Learning Objectives

By the end of this section, you will be able to:

- Describe how mutations affect protein synthesis and its products.

In both prokaryotes and eukaryotes, the major purpose of DNA is to provide the information needed to construct the **proteins** necessary for the cell can perform all of its functions. Proteins are large, complex molecules that play many critical roles in the body. They do most of the work in cells and are required for the structure, function, and regulation of the body's tissues and organs.

The information to make proteins is stored in an organism's DNA. Each protein is coded for by a specific section of DNA called a **gene.** A gene is the section of DNA required to produce one protein. Genes are typically hundreds or thousands of base pairs in length because they code for proteins made of hundreds or thousands of amino acids.

A gene **mutation** is a permanent alteration in the DNA sequence that makes up a gene, such that the sequence

differs from what is found in most people. Mutations range in size; they can affect anywhere from a single DNA building block (base pair) to a large segment of a chromosome that includes multiple genes.

Since the DNA sequence found within a gene controls protein synthesis. If the DNA sequence is altered, this can alter the amino acid sequence within a protein. This can have a variety of potential effects, which will be discussed in this chapter.

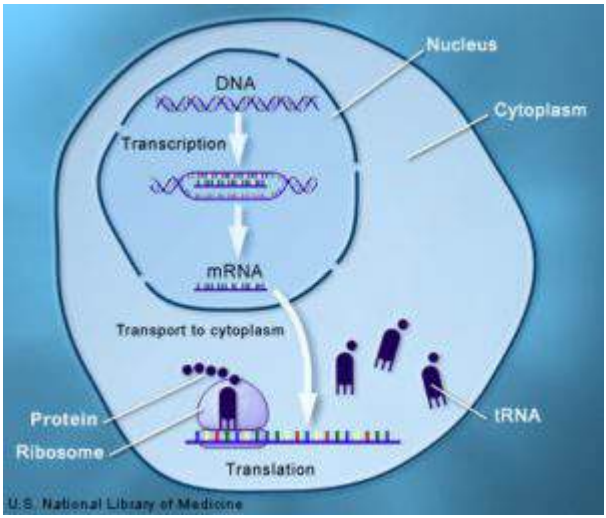


Figure 1 The process of protein synthesis first creates an mRNA copy of a DNA sequence during the process of transcription. This mRNA is translated into a sequence of amino acids by the ribosome. In this way, the information encoded in the sequence of bases in the DNA making up a gene is used to produce a protein.

How do genes direct the production of proteins?

In order to understand the potential effect of mutations, we must first have a background in how the information in DNA is used to produce a protein. Each protein is coded for by a gene, which is typically hundreds or thousands of base pairs in length. The information in the gene specifies the order in which the amino acids will be assembled into the protein.

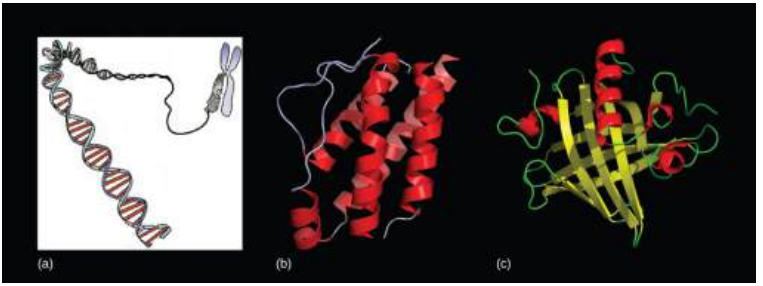


Figure 1 Genes, which are carried on (a) chromosomes, are linearly organized instructions for making the RNA and protein molecules that are necessary for all of processes of life. The (b) interleukin-2 protein and (c) alpha-2u-globulin protein are just two examples of the array of different molecular structures that are encoded by genes. (credit "chromosome: National Human Genome Research Institute; credit "interleukin-2": Ramin Herati/Created from PDB 1M47 and rendered with Pymol; credit "alpha-2u-globulin": Darren Logan/rendered with AISMIG)

The journey from gene to protein is complex and tightly controlled within each cell. It consists of two major steps: **transcription** and **translation.** Together, transcription and translation are known as **gene expression**.

TRANSCRIPTION

During the process of **transcription,** the information stored in a gene's DNA is used as a blueprint to produce a similar molecule called RNA (ribonucleic acid) in the cell nucleus. Both RNA and DNA are made up of a chain of nucleotide bases, but they have slightly different chemical properties (Figure 2).

- Both RNA and DNA contain a 5-carbon sugar, but the sugar differs: it is deoxyribose in DNA and ribose in RNA (DNA stands for deoxyribonucleic acid; RNA stands for ribonucleic acid).
- DNA and RNA also differ in the nitrogenous bases they contain. DNA contains A, T, C, and G. RNA contains A, C, and G, but no thymine. Instead it contains a base called uracil (U).
- DNA is almost always double-stranded (a double helix), while RNA is typically single stranded.

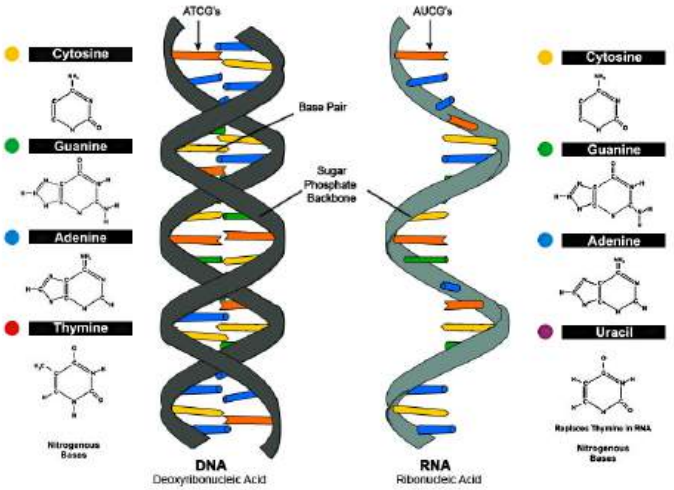


Figure 2 DNA vs RNA. Photo credit Zappys Technology Solution; Flickr.

The type of RNA that contains the information for making a protein is called messenger RNA (mRNA) because it carries the information, or message, from the DNA out of the nucleus into the cytoplasm. During transcription, this mRNA copy is made from a DNA molecule. This is possible because of the base-pairing rules: A with T (or U) and C with G. The hydrogen bonds connecting the base pairs in a DNA molecule are broken, and an enzyme creates a chain of RNA nucleotides that correspond to the DNA sequence.

In eukaryotes, transcription occurs in the nucleus (because that's where the DNA is). In prokaryotes, transcription occurs in the cytoplasm because there is no nucleus.

RNA PROCESSING

After prokaryotes produce an mRNA, it can be immediately translated since both processes occur in the cytoplasm. In fact, transcription and translation can occur at the same time

– as an mRNA is being transcribed, it can also begin to be translated.

Eukaryotes require a more complex process since the mRNA must move from the nucleus to the cytoplasm. Additionally, eukaryotic mRNAs are typically modified in several different ways: portions of the mRNA that do not code for amino acids are removed ("spliced" out), and the 5' and 3' ends are modified to help with recognition and mRNA stability. After these modifications are made, the mature mRNA is transported to the cytoplasm.

TRANSLATION

Translation, the second step in getting from a gene to a protein, takes place in the cytoplasm. The mRNA interacts with a specialized complex called a **ribosome,** which "reads" the sequence of mRNA bases. In conjunction with a type of RNA called transfer RNA (tRNA), the protein is assembled according to the instructions in the mRNA molecule. Each sequence of three bases in the mRNA, called a **codon,** usually codes for one particular amino acid. Remember that amino acids are the building blocks of proteins. Protein assembly continues until the ribosome encounters a "stop" codon (a sequence of three bases that does not code for an amino acid).

Recall that ribosomes are located in two different places in eukaryotic cells: free-floating in the cytoplasm and attached to the rough endoplasmic reticulum. The final destination of the protein determines where it will be synthesized.

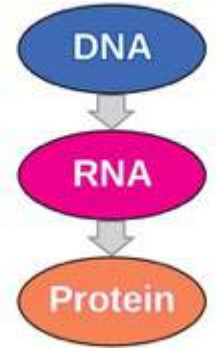


Figure 3: The Central Dogma – DNA is used to make RNA is used to make protein

The flow of information from DNA to RNA to proteins is one of the fundamental principles of molecular biology. It is so important that it is sometimes called the "central dogma" (Figures 3 and 4).

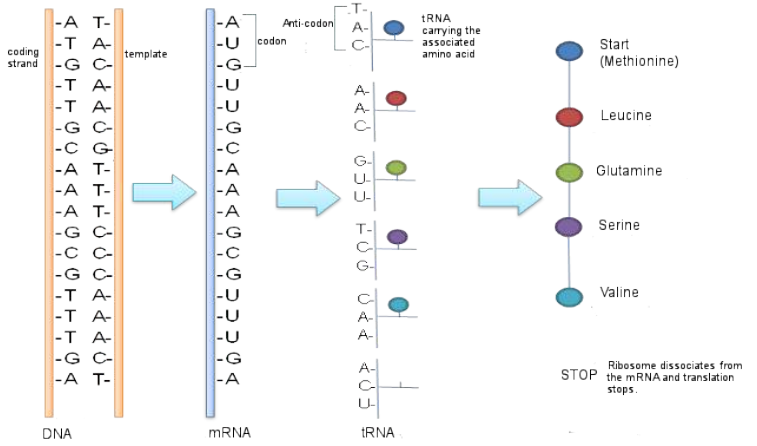


Figure 4: More detail on the central dogma. ("Overview of Protein Synthesis" by Becky Boone is licensed under CC BY-SA 2.0)

REFERENCES

How Gene Mutations Occur

Gene mutations can be classified in two major ways:

- **Hereditary mutations** are inherited from a parent and are present throughout a person's life in virtually every cell in the body. These mutations are also called germline mutations because they are present in the parent's egg or sperm cells, which are also called germ cells. When an egg and a sperm cell unite, the resulting fertilized egg cell receives DNA from both parents. If this DNA has a mutation, the child that grows from the fertilized egg will have the mutation in each of his or her cells.
- **Acquired (or somatic) mutations** occur at some time during a person's life and are present only in certain cells, not in every cell in the body. These changes can be caused by environmental factors such as ultraviolet radiation from the sun, or can occur if a mistake is made as DNA copies itself during cell division. Acquired mutations in somatic cells (cells other than sperm and egg cells) cannot be passed on to the next generation.

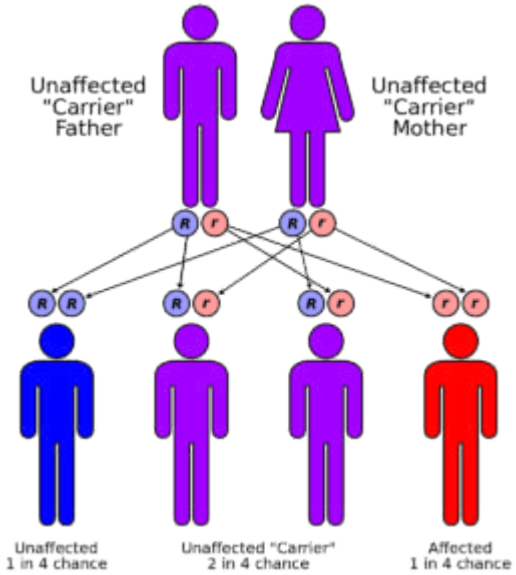


Figure 2 The red individual has inherited two mutated alleles of a gene from their parents. This is an example of a hereditary mutation.

Figure: The color variation seen in this tulip is caused by a somatic mutation – one which occurred early in the development of this individual flower.

Genetic changes that are described as ***de novo*** **(new) mutations** can be either hereditary or somatic. In some cases, the mutation occurs in a person's egg or sperm cell but is not present in any of the person's other cells. In other cases, the mutation occurs in the fertilized egg shortly after the egg and sperm cells unite. It is often impossible to tell exactly when a de novo mutation happened. As the fertilized egg divides, each resulting cell in the growing embryo will have the mutation. *De novo* mutations may explain genetic disorders in which an affected child has a mutation in every cell in the body but the parents do not, and there is no family history of the disorder.

Somatic mutations that happen in a single cell early in embryonic development can lead to a situation called mosaicism. These genetic changes are not present in a parent's egg or sperm cells, or in the fertilized egg, but

happen a bit later when the embryo includes several cells. As all the cells divide during growth and development, cells that arise from the cell with the altered gene will have the mutation, while other cells will not. Depending on the mutation and how many cells are affected, mosaicism may or may not cause health problems.

Most disease-causing gene mutations are uncommon in the general population. However, other genetic changes occur more frequently. Genetic alterations that occur in more than 1 percent of the population are called **polymorphisms**. They are common enough to be considered a normal variation in the DNA. Polymorphisms are responsible for many of the normal differences between people such as eye color, hair color, and blood type. Although many polymorphisms have no negative effects on a person's health, some of these variations may influence the risk of developing certain disorders.

REFERENCES

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Introduction to Genetic Disorders

To function correctly, each cell depends on thousands of proteins to do their jobs in the right places at the right times. Sometimes, gene mutations prevent one or more of these proteins from working properly. By changing a gene's instructions for making a protein, a mutation can cause the protein to malfunction or to be missing entirely. When a mutation alters a protein that plays a critical role in the body, it can disrupt normal development or cause a medical condition. A condition caused by mutations in one or more genes is called a **genetic disorder**.

In some cases, gene mutations are so severe that they prevent an embryo from surviving until birth. These changes occur in genes that are essential for development, and often disrupt the development of an embryo in its earliest stages. Because these mutations have very serious effects, they are incompatible with life.

It is important to note that genes themselves do not cause disease—genetic disorders are caused by mutations that make a gene function improperly. For example, when people say that someone has "the cystic fibrosis gene," they are usually referring to a mutated version of the *CFTR* gene, which causes the disease. All people, including those without cystic fibrosis, have a version of the *CFTR* gene.

REFERENCES

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Do all mutations affect health and development?

No; only a small percentage of mutations cause genetic disorders—most have no impact on health or development. For example, some mutations alter a gene's DNA sequence but do not change the function of the protein made by the gene.

Often, gene mutations that could cause a genetic disorder are repaired by certain enzymes before the gene is expressed and an altered protein is produced. Each cell has a number of pathways through which enzymes recognize and repair mistakes in DNA. Because DNA can be damaged or mutated in many ways, DNA repair is an important process by which the body protects itself from disease.

A very small percentage of all mutations actually have a positive effect. These mutations lead to new versions of proteins that help an individual better adapt to changes in his or her environment. For example, a beneficial mutation could result in a protein that protects an individual and future generations from a new strain of bacteria.

Because a person's genetic code can have a large number of mutations with no effect on health, diagnosing genetic conditions can be difficult. Sometimes, genes thought to be related to a particular genetic condition have mutations, but whether these changes are involved in development of the condition has not been determined; these genetic changes are known as variants of unknown significance (VOUS). Sometimes, no mutations are found in suspected disease-related genes, but mutations are found in other genes whose relationship to a particular genetic condition is unknown. It is difficult to know whether these variants are involved in the disease.

Figure: This lobster contains a mutation that causes it to be blue. This is estimated to occur in roughly one in two million lobsters.

REFERENCES

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Types of Mutations

The DNA sequence of a gene can be altered in a number of ways. Gene mutations have varying effects on health, depending on where they occur and whether they alter the function of essential proteins. The types of mutations include:

- **Silent mutation:** Silent mutations cause a change in the sequence of bases in a DNA molecule, but do not result in a change in the amino acid sequence of a protein (Figure 1).
- **Missense mutation:** This type of mutation is a change in one DNA base pair that results in the substitution of one amino acid for another in the protein made by a gene (Figure 1).
- **Nonsense mutation:** A nonsense mutation is also a change in one DNA base pair. Instead of substituting one amino acid for another, however, the altered DNA sequence prematurely signals the cell to stop building a protein (Figure 1). This type of mutation results in a shortened protein that may function improperly or not at all.

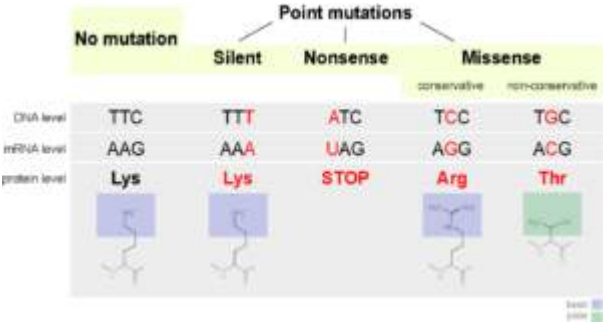


Figure: Some mutations do not change the sequence of amino acids in a protein. Some swap one amino acid for another. Others introduce an early stop codon into the sequence causing the protein to be truncated.

- **Insertion or Deletion:** An insertion changes the number of DNA bases in a gene by adding a piece of DNA. A deletion removes a piece of DNA. Insertions or deletions may be small (one or a few base pairs within a gene) or large (an entire gene, several genes, or a large section of a chromosome). In any of these cases, the protein made by the gene may not function properly.
- **Duplication:** A duplication consists of a piece of DNA that is abnormally copied one or more times. This type of mutation may alter the function of the resulting protein.
- **Frameshift mutation:** This type of mutation occurs when the addition or loss of DNA bases changes a gene's reading frame. A reading frame consists of groups of 3 bases that each code for one amino acid. A frameshift mutation shifts the grouping of these bases and changes the code for amino acids. The resulting protein is usually nonfunctional.

Insertions, deletions, and duplications can all be frameshift mutations.

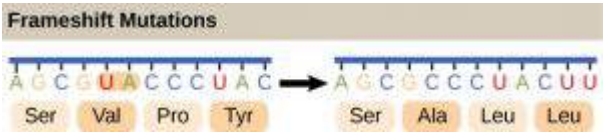


Figure 2 A frameshift mutation adds or deletes 1 or 2 bases. This results in a shift of the "reading frame" for the ribosome causing a drastic change in amino acid sequence. Photo Credit: OpenStax Biology.

- **Repeat expansion:** Nucleotide repeats are short DNA sequences that are repeated a number of times in a row. For example, a trinucleotide repeat is made up of 3-base- pair sequences, and a tetranucleotide repeat is made up of 4-base-pair sequences. A repeat expansion is a mutation that increases the number of times that the short DNA sequence is repeated. This type of mutation can cause the resulting protein to function in a completely different way than it would have originally.

REFERENCES

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Multifactorial Disorders and Genetic Predispositions

MULTIFACTORIAL DISORDERS

Researchers are learning that nearly all conditions and diseases have a genetic component. Some disorders, such as sickle cell disease and cystic fibrosis, are caused by mutations in a single gene. The causes of many other disorders, however, are much more complex. Common medical problems such as heart disease, diabetes, and obesity do not have a single genetic cause—they are likely associated with the effects of multiple genes in combination with lifestyle and environmental factors. Conditions caused by many contributing factors are called complex or **multifactorial disorders**.

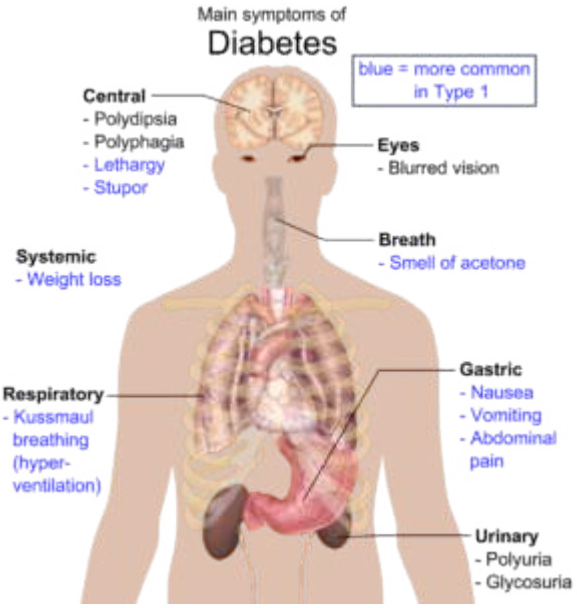


Figure 1 The main symptoms of diabetes, a multifactorial disorder

Although complex disorders often cluster in families, they do not have a clear- cut pattern of inheritance. This makes it difficult to determine a person's risk of inheriting or passing on these disorders. Complex disorders are also difficult to study and treat because the specific factors that cause most of these disorders have not yet been identified. Researchers continue to look for major contributing genes for many common complex disorders.

GENETIC PREDISPOSITIONS

A **genetic predisposition** (sometimes also called genetic susceptibility) is an increased likelihood of developing a particular disease based on a person's genetic makeup. A genetic predisposition results from specific genetic variations

that are often inherited from a parent. These genetic changes contribute to the development of a disease but do not directly cause it. Some people with a predisposing genetic variation will never get the disease while others will, even within the same family.

Genetic variations can have large or small effects on the likelihood of developing a particular disease. For example, certain mutations in the *BRCA1* or *BRCA2* genes greatly increase a person's risk of developing breast cancer and ovarian cancer. Variations in other genes, such as *BARD1* and *BRIP1*, also increase breast cancer risk, but the contribution of these genetic changes to a person's overall risk appears to be much smaller.

Current research is focused on identifying genetic changes that have a small effect on disease risk but are common in the general population. Although each of these variations only slightly increases a person's risk, having changes in several different genes may combine to increase disease risk significantly. Changes in many genes, each with a small effect, may underlie susceptibility to many common diseases, including cancer, obesity, diabetes, heart disease, and mental illness.

In people with a genetic predisposition, the risk of disease can depend on multiple factors in addition to an identified genetic change. These include other genetic factors (sometimes called modifiers) as well as lifestyle and environmental factors. Although a person's genetic makeup cannot be altered, some lifestyle and environmental modifications (such as having more frequent disease screenings and maintaining a healthy weight) may be able to reduce disease risk in people with a genetic predisposition.

REFERENCES

DNA Repair

DNA replication is a highly accurate process, but mistakes can occasionally occur, such as a DNA polymerase inserting a wrong base. Uncorrected mistakes may sometimes lead to serious consequences, such as cancer. Repair mechanisms correct the mistakes. In rare cases, mistakes are not corrected, leading to mutations; in other cases, repair enzymes are themselves mutated or defective.

Most of the mistakes during DNA replication are promptly corrected by DNA polymerase by proofreading the base that has been just added (Figure 1). In proofreading, the DNA pol reads the newly added base before adding the next one, so a correction can be made. The polymerase checks whether the newly added base has paired correctly with the base in the template strand. If it is the right base, the next nucleotide is added. If an incorrect base has been added, the enzyme makes a cut at the phosphodiester bond and releases the wrong nucleotide. This is performed by the exonuclease action of DNA pol III. Once the incorrect nucleotide has been removed, a new one will be added again (**Figure 1**).

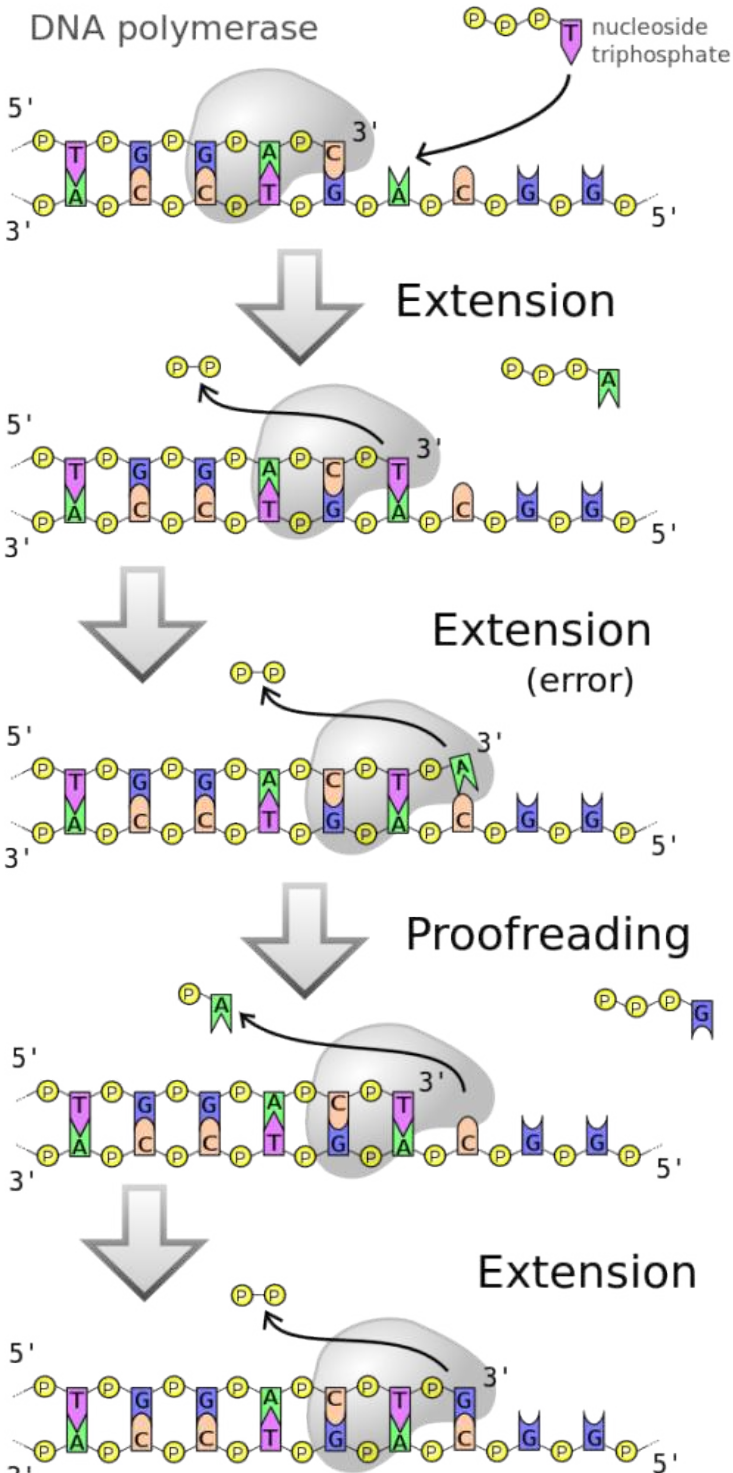

DNA polymerase
nucleoside triphosphate
5'
3'
Extension
Extension
(error)
Proofreading
Extension

Figure 1 Proofreading by DNA polymerase corrects errors during replication. Photo credit Madeline Price Ball; Wikimedia.

Some errors are not corrected during replication, but are instead corrected after replication is completed; this type of repair is known as **mismatch repair** (**Figure 2**). The enzymes recognize the incorrectly added nucleotide and excise it; this is then replaced by the correct base. If this remains uncorrected, it may lead to more permanent damage. How do mismatch repair enzymes recognize which of the two bases is the incorrect one? In *E. coli*, after replication, the nitrogenous base adenine acquires a methyl group (CH_3); the parental DNA strand will have methyl groups, whereas the newly synthesized strand lacks them. Thus, DNA polymerase is able to remove the wrongly incorporated bases from the newly synthesized, non-methylated strand. In eukaryotes, the mechanism is not very well understood, but it is believed to involve recognition of unsealed nicks in the new strand, as well as a short-term continuing association of some of the replication proteins with the new daughter strand after replication has completed.

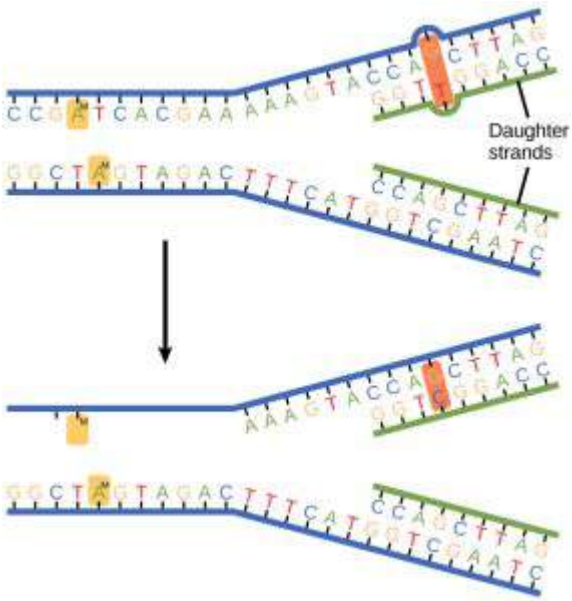


Figure 2 In mismatch repair, the incorrectly added base is detected after replication. The mismatch repair proteins detect this base and remove it from the newly synthesized strand by nuclease action. The gap is now filled with the correctly paired base.

In another type of repair mechanism, **nucleotide excision repair**, enzymes replace incorrect bases by making a cut on both the 3' and 5' ends of the incorrect base (**Figure 3**). The segment of DNA is removed and replaced with the correctly paired nucleotides by the action of DNA pol. Once the bases are filled in, the remaining gap is sealed with a phosphodiester linkage catalyzed by DNA ligase. This repair mechanism is often employed when UV exposure causes the formation of thymine-thymine dimers (the small – connecting the two Ts in Figure 3).

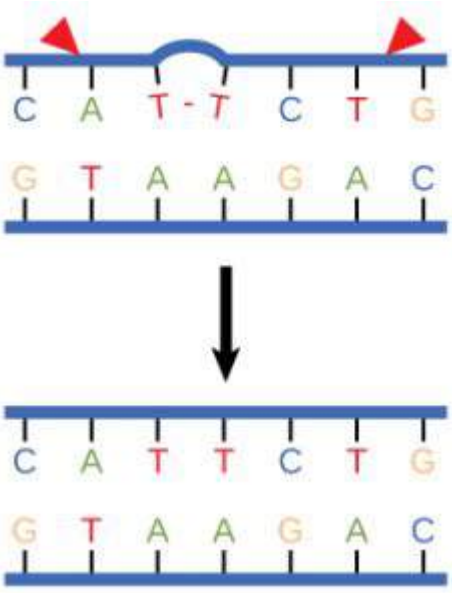


Figure 3 Nucleotide excision repairs thymine dimers. When exposed to UV, thymines lying adjacent to each other can form thymine dimers. In normal cells, they are excised and replaced.

A well-studied example of mistakes not being corrected is seen in people suffering from xeroderma pigmentosa (**Figure 4**). Affected individuals have skin that is highly sensitive to UV rays from the sun. When individuals are exposed to UV, pyrimidine dimers, especially those of thymine, are formed; people with xeroderma pigmentosa are not able to repair the damage. These are not repaired because of a defect in the nucleotide excision repair enzymes, whereas in normal individuals, the thymine dimers are excised and the defect is corrected. The thymine dimers distort the structure of the DNA double helix, and this may cause problems during DNA replication. People with xeroderma pigmentosa have a higher risk of contracting skin cancer than those who don't have the condition.

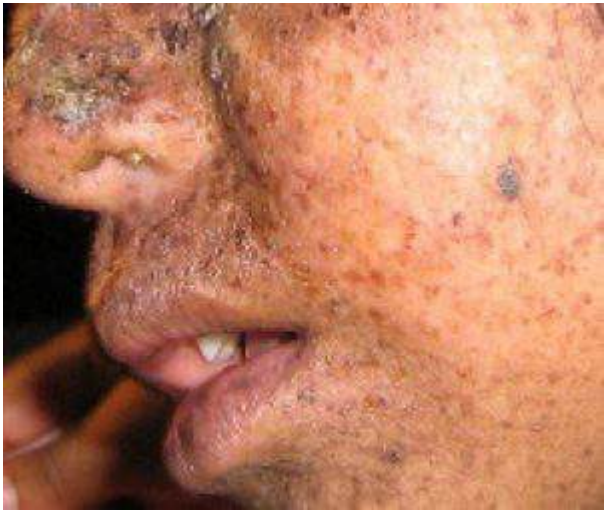


Figure 4 Xeroderma pigmentosa is a condition in which thymine dimerization from exposure to UV is not repaired. Exposure to sunlight results in skin lesions. (credit: James Halpern et al.)

Errors during DNA replication are not the only reason why mutations arise in DNA. Mutations, variations in the nucleotide sequence of a genome, can also occur because of damage to DNA. Such mutations may be of two types: induced or spontaneous. **Induced mutations** are those that result from an exposure to a **mutagen:** chemicals, UV rays, x-rays, or some other environmental agent. **Spontaneous mutations** occur without any exposure to any environmental agent; they are a result of natural reactions taking place within the body.

Mutations may have a wide range of effects. Some mutations are not expressed; these are known as **silent mutations**. Other mutations can have serious effects on the organism (such as the mutation that causes xeroderma pigmentosa.

Mutations in repair genes have been known to cause cancer. Many mutated repair genes have been implicated

in certain forms of pancreatic cancer, colon cancer, and colorectal cancer. Mutations can affect either somatic cells or gametes. If many mutations accumulate in a somatic cell, they may lead to problems such as the uncontrolled cell division observed in cancer. If a mutation takes place in a gamete, the mutation can be passed on to the next generation.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. December 21, 2017. https://cnx.org/contents/GFy_h8cu@10.120:MPy5UMKm@7/DNA-Repair

Changes in Numbers of Genes or Chromosomes

CHANGES IN NUMBERS OF GENES

People have two copies of most genes, one copy inherited from each parent. In some cases, however, the number of copies varies—meaning that a person can be born with one, three, or more copies of particular genes. Less commonly, one or more genes may be entirely missing. This type of genetic difference is known as copy number variation (CNV).

Copy number variation results from insertions, deletions, and duplications of large segments of DNA. These segments are big enough to include whole genes. Variation in gene copy number can influence the activity of genes and ultimately affect many body functions.

Researchers were surprised to learn that copy number variation accounts for a significant amount of genetic difference between people. More than 10 percent of human DNA appears to contain these differences in gene copy number. While much of this variation does not affect health or development, some differences likely influence a person's risk of disease and response to certain drugs. Future research will focus on the consequences of copy number variation in different parts of the genome and study the contribution of these variations to many types of disease.

Human cells normally contain 23 pairs of chromosomes, for a total of 46 chromosomes in each cell. A change in the number of chromosomes can cause problems with growth, development, and function of the body's systems. These changes can occur during the formation of reproductive cells (eggs and sperm), in early fetal development, or in any cell

after birth. A gain or loss of chromosomes from the normal 46 is called **aneuploidy**.

A common form of aneuploidy is **trisomy,** or the presence of an extra chromosome in cells. "Tri-" is Greek for "three"; people with trisomy have three copies of a particular chromosome in cells instead of the normal two copies. Down syndrome is an example of a condition caused by trisomy. People with Down syndrome typically have three copies of chromosome 21 in each cell, for a total of 47 chromosomes per cell.

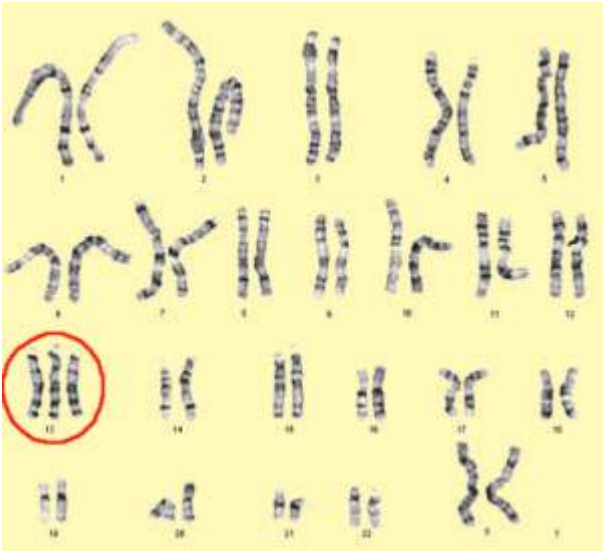


Figure 1 This karyotype, which is a picture of all the chromosomes from one individual, is from a person who has Trisomy 13.

Monosomy, or the loss of one chromosome in cells, is another kind of aneuploidy. "Mono-" is Greek for "one"; people with monosomy have one copy of a particular chromosome in cells instead of the normal two copies. Turner syndrome is a condition caused by monosomy. Women with Turner syndrome usually have only one copy

of the X chromosome in every cell, for a total of 45 chromosomes per cell.

Rarely, some cells end up with complete extra sets of chromosomes. Cells with one additional set of chromosomes, for a total of 69 chromosomes, are called **triploid.** Cells with two additional sets of chromosomes, for a total of 92 chromosomes, are called tetraploid. A condition in which every cell in the body has an extra set of chromosomes is not compatible with life.

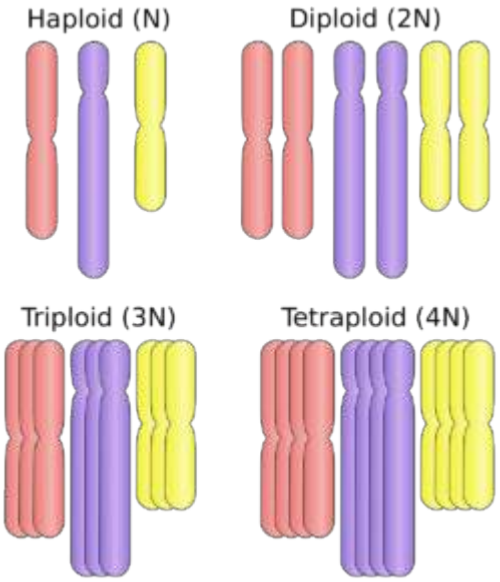


Figure 2 "Ploid" refers to the number of copies of each chromosome found in a somatic cell.

Figure 3 Human and other animal cells do not develop if they have an entire extra set of chromosomes. In contrast, plants often have entire copied sets of chromosomes. This strawberry is an example of a plant that is tetraploid.

In some cases, a change in the number of chromosomes occurs only in certain cells. When an individual has two or more cell populations with a different chromosomal makeup, this situation is called **chromosomal mosaicism**. Chromosomal mosaicism occurs from an error in cell division in cells other than eggs and sperm. Most commonly, some cells end up with one extra or missing chromosome (for a total of 45 or 47 chromosomes per cell), while other cells have the usual 46 chromosomes. Mosaic Turner syndrome is one example of chromosomal mosaicism. In females with this condition, some cells have 45 chromosomes because they are missing one copy of the X chromosome, while other cells have the usual number of chromosomes.

Many cancer cells also have changes in their number of chromosomes. These changes are not inherited; they occur in somatic cells (cells other than eggs or sperm) during the formation or progression of a cancerous tumor.

REFERENCES

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

REFERENCES

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

PROTEIN SYNTHESIS

Learning Outcomes

- Discuss the potential implications of mutations at cellular, organismal, and evolutionary levels
  - Summarize the processes involved in protein synthesis.
  - Describe how mutations affect the process of protein synthesis and its products.

In both prokaryotes and eukaryotes, the major purpose of DNA is to provide the information needed to construct the **proteins** necessary for the cell can perform all of its functions. Proteins are large, complex molecules that play many critical roles in the body. They do most of the work in cells and are required for the structure, function, and regulation of the body's tissues and organs.

Recall that proteins are made up of hundreds or thousands of smaller units called amino acids, which are attached to one another in long chains. There are 20 different types of

amino acids that can be combined to make a protein. The sequence of amino acids determines each protein's unique 3-dimensional structure and its specific function.

Table 1: Some of the functions of proteins in cells, listed in alphabetical order:

Function	Description
Antibody	Antibodies bind to specific foreign particles, such as viruses and bacteria, to help protect the body.
Enzyme	Enzymes carry out almost all of the thousands of chemical reactions that take place in cells. They also assist with the formation of new molecules by reading the genetic information stored in DNA.
Messenger	Messenger proteins, such as some types of hormones, transmit signals to coordinate biological processes between different cells, tissues, and organs.
Structural component	These proteins provide structure and support for cells. On a larger scale, they also allow the body to move.
Transport/ storage	These proteins bind and carry atoms and small molecules within cells and throughout the body.

The information to make proteins is stored in an organism's DNA. Each protein is coded for by a specific section of DNA called a **gene.** A gene is the section of DNA required to produce one protein. Genes are typically hundreds or thousands of base pairs in length because they code for proteins made of hundreds or thousands of amino acids.

Remember that DNA in eukaryotes is found as long linear molecules called chromosomes (Figure 1). Chromosomes are millions of base pairs in length and each contain many, many genes (Table 2). An organisms complete set of DNA (including all its genes) is referred to as its **genome.**

Table 2 Size and number of genes of several human chromosomes.

Chromosome	Size (in base pairs)	# of genes
1	248,956,422	2058
10	133,797,422	733
22	50818468	488

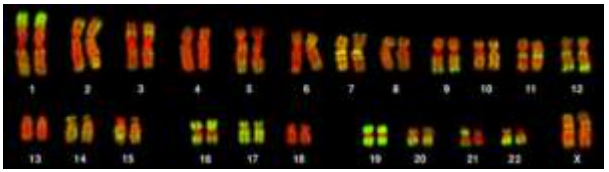


Figure 1: A karyotype showing the sizes of all the human chromosomes. Notice that they decrease in size.

To summarize: many base pairs make up one gene, many genes are found on one chromosome, and many chromosomes can be found in one genome.

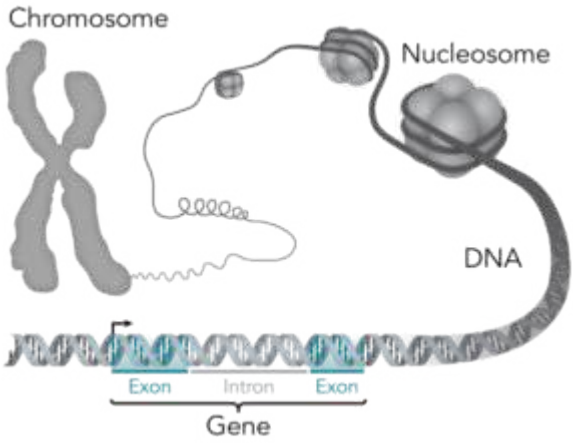


Figure 2: The arrangement of DNA into chromosomes. Photo credit: Thomas Splettstoesser (www.scistyle.com)

REFERENCES

The Genetic Code

THE CENTRAL DOGMA: DNA ENCODES RNA; RNA ENCODES PROTEIN

To summarize what we know to this point, the cellular process of transcription generates messenger RNA (mRNA), a mobile molecular copy of one or more genes with an alphabet of A, C, G, and uracil (U). Translation of the mRNA template converts nucleotide-based genetic information into a protein product. This flow of genetic information in cells from DNA to mRNA to protein is described by the Central Dogma (Figure 1), which states that genes specify the sequence of mRNAs, which in turn specify the sequence of proteins. The decoding of one molecule to another is performed by specific proteins and RNAs. Because the information stored in DNA is so central to cellular function, it makes intuitive sense that the cell would make mRNA copies of this information for protein synthesis, while keeping the DNA itself intact and protected.

It turns out that the central dogma is not always true. We will not discuss the exceptions here, however.

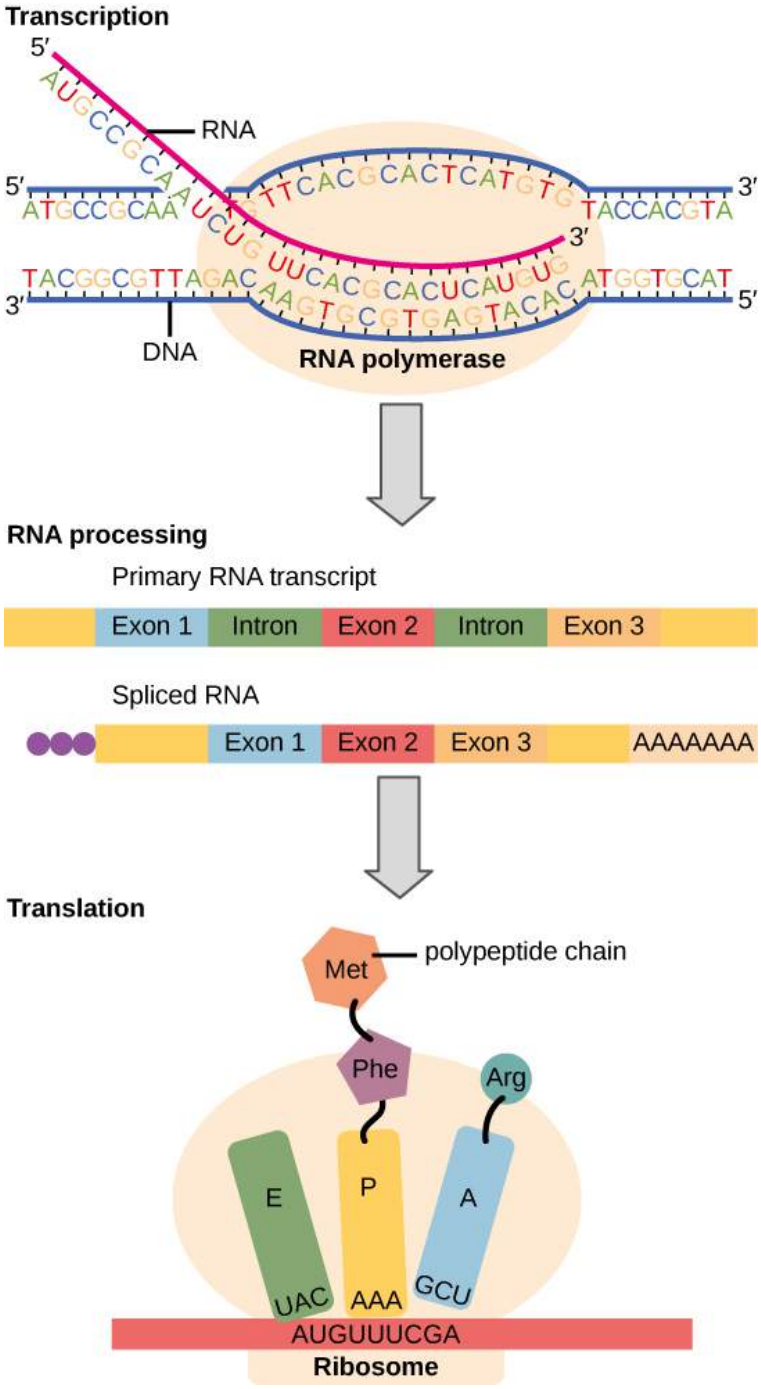

Transcription
5′
RNA
5′
3′
ATGCCGCAA
TACCACGTA
3′
TACGGCGTTAGAC
ATGGTGCAT
3′
5′
DNA
RNA polymerase
RNA processing
Primary RNA transcript
Exon 1
Intron
Exon 2
Intron
Exon 3
Spliced RNA
Exon 1
Exon 2
Exon 3
AAAAAAA
Translation
Met
polypeptide chain
Phe
Arg
E
P
A
UAC
AAA
GCU
AUGUUUCGA
Ribosome

Figure 1 Instructions on DNA are transcribed onto messenger RNA. Ribosomes are able to read the genetic information inscribed on a strand of messenger RNA and use this information to string amino acids together into a protein.

AMINO ACID STRUCTURE

Protein sequences consist of 20 commonly occurring amino acids (Figure 2); therefore, it can be said that the protein alphabet consists of 20 letters. Different amino acids have different chemistries (such as acidic versus basic, or polar and non-polar) and different structural constraints. Variation in amino acid sequence gives rise to enormous variation in protein structure and function.

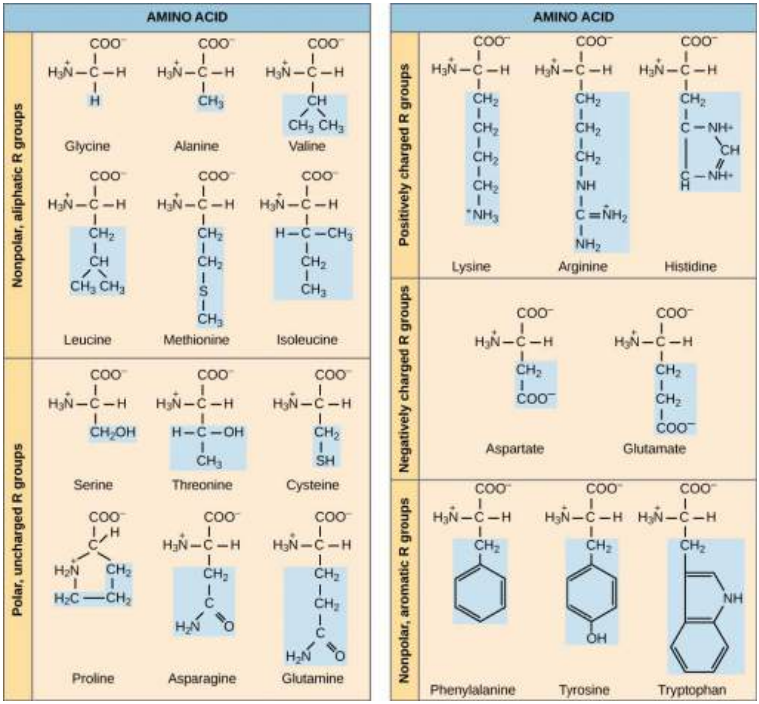


Figure 2 Structures of the 20 amino acids found in proteins are shown. Each amino acid is composed of an amino group (NH+3), a carboxyl group (COO-), and a side chain (blue). The side chain may be nonpolar, polar, or charged, as well as large or small. It is the variety of amino acid side chains that gives rise to the incredible variation of protein structure and function.

GENETIC CODE

Each amino acid is defined by a three-nucleotide sequence called the triplet **codon**. The relationship between a nucleotide codon and its corresponding amino acid is called the **genetic code**. Given the different numbers of "letters" in the mRNA (4 – A, U, C, G) and protein "alphabets" (20 different amino acids) one nucleotide could not correspond to one amino acid. Nucleotide doublets would also not be

sufficient to specify every amino acid because there are only 16 possible two-nucleotide combinations (4^2). In contrast, there are 64 possible nucleotide triplets (4^3), which is far more than the number of amino acids. Scientists theorized that amino acids were encoded by nucleotide triplets and that the genetic code was degenerate. In other words, a given amino acid could be encoded by more than one nucleotide triplet. (**Figure 2**). These nucleotide triplets are called codons.

The same codon will always specify the insertion of one specific amino acid. The chart seen in Figure 2 can be used to translate an mRNA sequence into an amino acid sequence. For example, the codon UUU will always cause the insertion of the amino acid phenylalanine (Phe), while the codon UUA will cause the insertion of leucine (Leu).

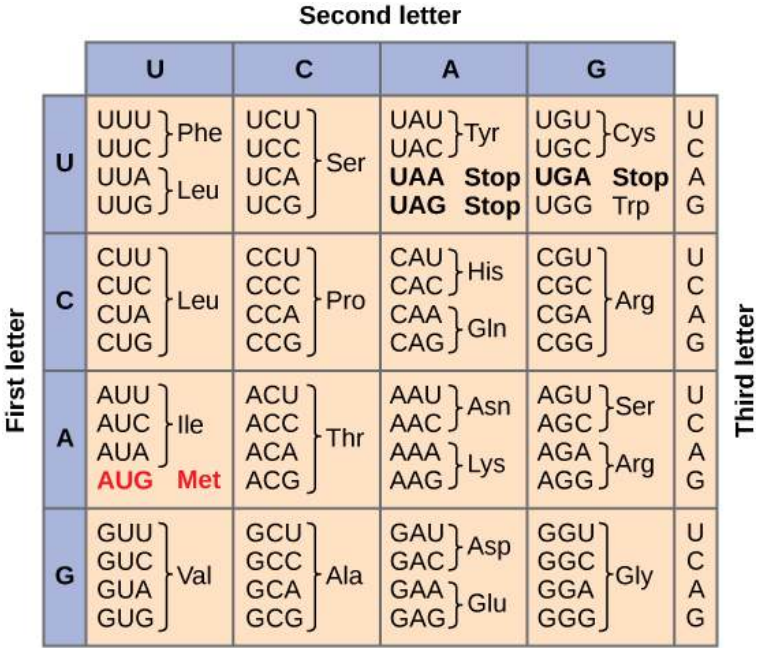


Figure 3 This figure shows the genetic code for translating each nucleotide triplet in mRNA into an amino acid or a termination signal in a nascent protein. (credit: modification of work by NIH)

Each set of three bases (one codon) causes the insertion of one specific amino acid into the growing protein. This means that the insertion of one or two nucleotides can completely change the triplet "reading frame", thereby altering the message for every subsequent amino acid (Figure 4). Though insertion of three nucleotides caused an extra amino acid to be inserted during translation, the integrity of the rest of the protein was maintained.

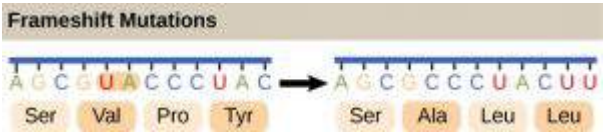


Figure 4 The deletion of two nucleotides shifts the reading frame of an mRNA and changes the entire protein message, creating a nonfunctional protein or terminating protein synthesis altogether.

Three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called **stop codons**. Another codon, AUG, also has a special function. In addition to specifying the amino acid methionine, it also serves as the **start codon** to initiate translation. The reading frame for translation is set by the AUG start codon near the 5' end of the mRNA. The genetic code is universal. With a few exceptions, virtually all species use the same genetic code for protein synthesis, which is powerful evidence that all life on Earth shares a common origin.

REFERENCES

OpenStax, Biology. OpenStax CNX. January 2, 2017 https://cnx.org/contents/

GFy_h8cu@10.120:QEibhJMi@9/The-Genetic-Code#fig-ch15_01_05

Prokaryotic Transcription

Both prokaryotes and eukaryotes perform fundamentally the same process of transcription, with the important difference of the membrane-bound nucleus in eukaryotes. With the genes bound in the nucleus, transcription occurs in the nucleus of the cell and the mRNA transcript must be transported to the cytoplasm. In prokaryotes, which lack membrane-bound nuclei and other organelles, transcription occurs in the cytoplasm of the cell.

RNA POLYMERASE

RNA Polymerase is the enzyme that produces the mRNA molecule (just like DNA polymerase produced a new DNA molecule during DNA replication). Prokaryotes use the same RNA polymerase to transcribe all of their genes. In *E. coli*, the polymerase is composed of five polypeptide subunits. These subunits assemble every time a gene is transcribed, and they disassemble once transcription is complete. Each subunit has a unique role (which you do not need to memorize). The polymerase comprised of all five subunits is called the **holoenzyme.**

INITIATION

Transcription in prokaryotes (and in eukaryotes) requires the DNA double helix to partially unwind in the region of mRNA synthesis. The region of unwinding is called a transcription bubble. The DNA sequence onto which the proteins and enzymes involved in transcription bind to initiate the process is called a **promoter**. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all of the time, some of the time, or hardly at all. The structure and function of a prokaryotic promoter is relatively simple (Figure 1). One important sequence in the prokaryotic promoter is located 10 bases before the transcription start site (-10) and is commonly called the TATA box.

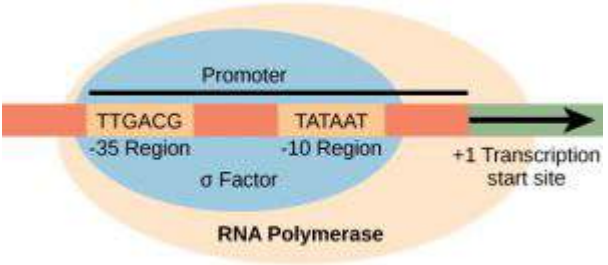


Figure 1 The general structure of a prokaryotic promoter.

To begin transcription, the RNA polymerase holoenzyme assembles at the promoter. The dissociation of σ allows the core enzyme to proceed along the DNA template, synthesizing mRNA by adding RNA nucleotides according to the base pairing rules, similar to the way a new DNA molecule is produced during DNA replication. Only one of the two DNA strands is transcribed. The transcribed strand of DNA is called the **template strand** because it is the template for mRNA production. The mRNA product is complementary to

the template strand and is almost identical to the other DNA strand, called the **non-template strand**, with the exception that RNA contains a uracil (U) in place of the thymine (T) found in DNA. Like DNA polymerase, RNA polymerase adds new nucleotides onto the 3'-OH group of the previous nucleotide. This means that the growing mRNA strand is being synthesized in the 5' to 3' direction. Because DNA is anti-parallel, this means that the RNA polymerase is moving in the 3' to 5' direction down the template strand (Figure 2).

ELONGATION

As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme as the hydrogen bonds that connect the complementary base pairs in the DNA double helix are broken (Figure 2). The DNA is rewound behind the core enzyme as the hydrogen bonds are reformed. The base pairing between DNA and RNA is not stable enough to maintain the stability of the mRNA synthesis components. Instead, the RNA polymerase acts as a stable linker between the DNA template and the newly forming RNA strand to ensure that elongation is not interrupted prematurely.

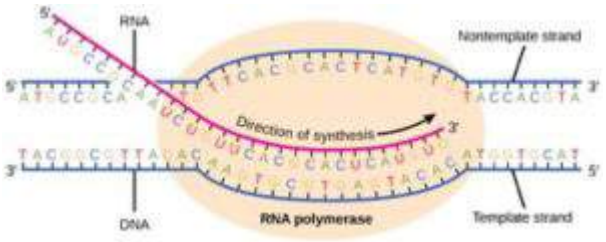


Figure 2 During elongation, RNA polymerase tracks along the DNA template, synthesizes mRNA in the 5' to 3' direction, and unwinds then rewinds the DNA as it is read.

TERMINATION

Once a gene is transcribed, the RNA polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA. Depending on the gene being transcribed, there are two kinds of termination signals. One is protein-based and the other is RNA-based. Both termination signals rely on specific sequences of DNA near the end of the gene that cause the polymerase to release the mRNA.

In a prokaryotic cell, by the time transcription ends, the transcript would already have been used to begin making copies of the encoded protein because the processes of transcription and translation can occur at the same time since both occur in the cytoplasm (**Figure 3**). In contrast, transcription and translation cannot occur simultaneously in eukaryotic cells since transcription occurs inside the nucleus and translation occurs outside in the cytoplasm.

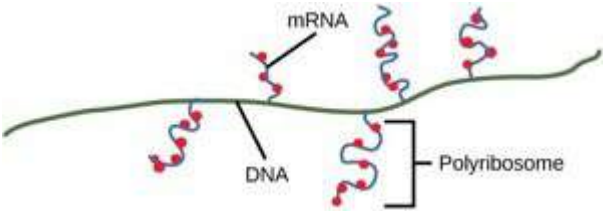


Figure 3: Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

REFERENCES

OpenStax, Biology. OpenStax CNX. December 21, 2017 https://cnx.org/contents/GFy_h8cu@10.120:SPzsALhh@7/Prokaryotic-Transcription

Eukaryotic Transcription

Prokaryotes and eukaryotes perform fundamentally the same process of transcription, with a few key differences. The most important difference between prokaryotes and eukaryotes is the latter's membrane-bound nucleus and organelles. With the genes enclosed in a nucleus, the eukaryotic cell must be able to transport its mRNA to the cytoplasm and must protect its mRNA from degrading before it is translated. Eukaryotes also employ three different polymerases that each transcribe a different subset of genes.

INITIATION

The eukaryotic promoters that we are most interested in are similar to prokaryotic promoters in that they contain a TATA box (Figure 1). However, initiation of transcription is much more complex in eukaryotes compared to prokaryotes. Unlike the prokaryotic RNA polymerase that can bind to a DNA template on its own, eukaryotes require several other proteins, called **transcription factors**, to first bind to the promoter region and then help recruit the appropriate polymerase.

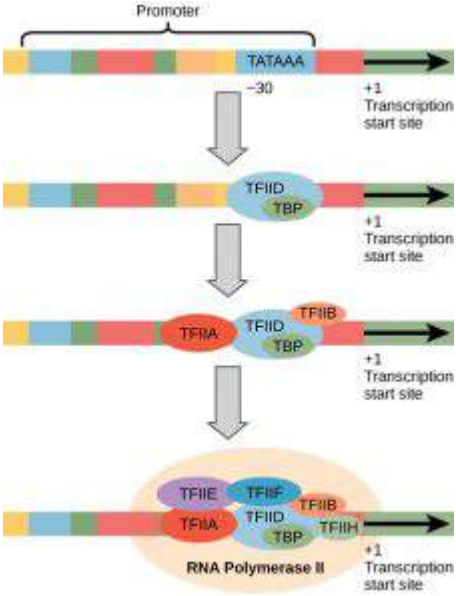


Figure 1 The generalized structure of a eukaryotic promoter and transcription factors.

In addition, there are three different RNA polymerases in eukaryotes, each of which is made up of 10 subunits or more. Each eukaryotic RNA polymerase also requires a distinct set of transcription factors to bring it to the DNA template.

RNA polymerase I is located in the nucleolus, a specialized nuclear substructure in which ribosomal RNA (rRNA) is transcribed, processed, and assembled into ribosomes. The rRNA molecules are considered structural RNAs because they have a cellular role but are not translated into protein. The rRNAs are components of the ribosome and are essential to the process of translation. RNA polymerase I synthesizes most of the rRNAs.

RNA polymerase II is located in the nucleus and synthesizes all protein-coding nuclear pre-mRNAs. Eukaryotic pre-mRNAs undergo extensive processing after transcription but before translation. For clarity, the term “mRNA” will only

be used to describe the mature, processed molecules that are ready to be translated. RNA polymerase II is responsible for transcribing the overwhelming majority of eukaryotic genes.

RNA polymerase III is also located in the nucleus. This polymerase transcribes a variety of structural RNAs including transfer pre-RNAs (pre-tRNAs), and small nuclear pre-RNAs. The tRNAs have a critical role in translation; they serve as the adaptor molecules between the mRNA template and the growing polypeptide chain. Small nuclear RNAs have a variety of functions, including "splicing" pre-mRNAs and regulating transcription factors.

Each of the types of RNA polymerase recognizes a different promoter sequence and requires different transcription factors.

ELONGATION

Following the formation of the preinitiation complex, the polymerase is released from the other transcription factors, and elongation is allowed to proceed as it does in prokaryotes with the RNA polymerase synthesizing pre-mRNA in the 5' to 3' direction. As discussed previously, RNA polymerase II transcribes the major share of eukaryotic genes, so this section will focus on how this polymerase accomplishes elongation and termination.

Although the enzymatic process of elongation is essentially the same in eukaryotes and prokaryotes, the DNA template is more complex. When eukaryotic cells are not dividing, their genes exist as a diffuse mass of DNA and proteins called chromatin. The DNA is tightly packaged around charged histone proteins at repeated intervals. These DNA–histone complexes, collectively called nucleosomes, are regularly

spaced and include 146 nucleotides of DNA wound around eight histones like thread around a spool.

For RNA synthesis to occur, the transcription machinery needs to move histones out of the way every time it encounters a nucleosome. This is accomplished by a special protein complex called FACT, which stands for "facilitates chromatin transcription." This complex pulls histones away from the DNA template as the polymerase moves along it. Once the pre-mRNA is synthesized, the FACT complex replaces the histones to recreate the nucleosomes.

TERMINATION

The termination of transcription is different for the different polymerases. Unlike in prokaryotes, elongation by RNA polymerase II in eukaryotes takes place 1,000–2,000 nucleotides beyond the end of the gene being transcribed. This pre-mRNA tail is removed during mRNA processing. RNA polymerases I and III require termination signals. Genes transcribed by RNA polymerase I contain a specific 18-nucleotide sequence that is recognized by a termination protein. The process of termination in RNA polymerase III involves an mRNA hairpin that causes the mRNA to be released.

REFERENCES

OpenStax, Biology. OpenStax CNX. December 21, 2017. https://cnx.org/contents/GFy_h8cu@10.120:6l70P9u6@5/Eukaryotic-Transcription

Eukaryotic RNA Processing

Eukaryotic mRNAs must undergo several processing steps before they can be transferred from the nucleus to the cytoplasm and translated into a protein. The additional steps involved in eukaryotic mRNA maturation create a molecule that is much more stable than a prokaryotic mRNA. Eukaryotic mRNAs typically last for several hours, whereas the typical prokaryotic mRNA lasts no more than five seconds.

The mRNA transcript is coated in RNA-stabilizing proteins to prevent it from degrading while it is processed and exported out of the nucleus. The three most important steps of pre-mRNA processing are the addition of stabilizing and signaling factors at the 5' and 3' ends of the molecule, and the removal of intervening sequences that do not specify the appropriate amino acids. In rare cases, the mRNA transcript can be "edited" after it is transcribed.

5' CAPPING

While the pre-mRNA is still being synthesized, a 7-methylguanosine cap is added to the 5' end of the growing transcript by a phosphate linkage. This **5' cap** protects the nascent mRNA from degradation. In addition, factors involved in protein synthesis recognize the cap to help initiate translation by ribosomes. Structurally,

7-methylguanosine looks like a guanine nucleotide, but with an added methyl group (Figure 1).

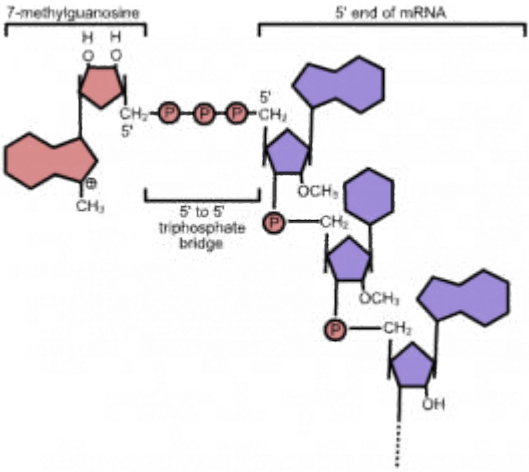


Figure 1 The structure of the 5' cap. Notice that the 7-methylguanosine is shaped like a purine nucleotide, but has an extra methyl group (CH3). It is attached to the mRNA upside-down compared to the other nucleotides. Photo credit Zephyris; Wikipedia.

3' POLY-A TAIL

Once elongation is complete, the pre-mRNA is cleaved by an endonuclease between an AAUAAA consensus sequence and a GU-rich sequence, leaving the AAUAAA sequence on the pre-mRNA. An enzyme called poly-A polymerase then adds a string of approximately 200 adenine nucleotides, called the **poly-A tail**. This modification further protects the pre-mRNA from degradation and signals the export of the cellular factors that the transcript needs to the cytoplasm.

PRE-MRNA SPLICING

Eukaryotic genes are composed of **exons,** which correspond to protein-coding sequences (*ex*-on signifies that they are *ex*pressed), and *int*ervening sequences called **introns** (*int*-ron denotes their *int*ervening role), which may be involved in gene regulation but are removed from the pre-mRNA during processing (Figure 2). Intron sequences in mRNA do not encode functional proteins.

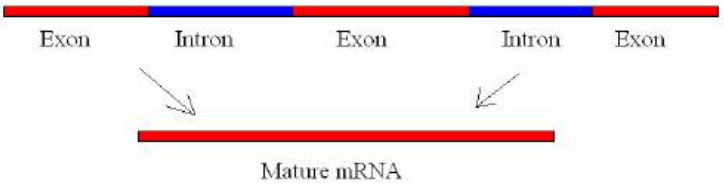


Figure 2 Eukaryotic mRNA contains introns that must be spliced out. A 5' cap and 3' tail are also added. Photo credit Kazulanth; Wikimedia. This work has been released into the public domain.

The discovery of introns came as a surprise to researchers in the 1970s who expected that pre-mRNAs would specify protein sequences without further processing, as they had observed in prokaryotes. The genes of higher eukaryotes very often contain one or more introns. These regions may correspond to regulatory sequences; however, the biological significance of having many introns or having very long introns in a gene is unclear. It is possible that introns slow down gene expression because it takes longer to transcribe pre-mRNAs with lots of introns. Alternatively, introns may be nonfunctional sequence remnants left over from the fusion of ancient genes throughout evolution. This is supported by the fact that separate exons often encode separate protein subunits or domains. For the most part, the sequences of introns can be mutated without ultimately affecting the protein product.

All of a pre-mRNA's introns must be completely and precisely removed before protein synthesis. If the process errs by even a single nucleotide, the reading frame of the rejoined exons would shift, and the resulting protein would be dysfunctional. The process of removing introns and reconnecting exons is called **splicing** (Figure 2 and 3). Introns are removed and degraded while the pre-mRNA is still in the nucleus. Splicing occurs by a sequence-specific mechanism that ensures introns will be removed and exons rejoined with the accuracy and precision of a single nucleotide. The splicing of pre-mRNAs is conducted by complexes of proteins and RNA molecules called **spliceosomes** (Figure 3).

Note that more than 70 individual introns can be present, and each has to undergo the process of splicing—in addition to 5' capping and the addition of a poly-A tail—just to generate a single, translatable mRNA molecule.

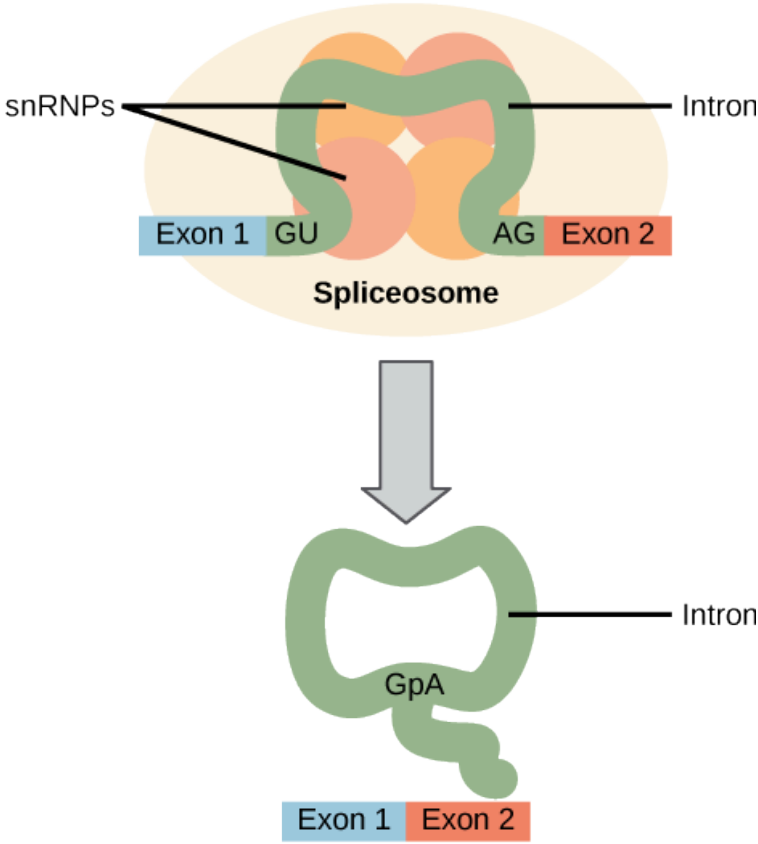


Figure 3 Pre-mRNA splicing involves the precise removal of introns from the primary RNA transcript. The splicing process is catalyzed by protein complexes called spliceosomes that are composed of proteins and RNA molecules called snRNAs. Spliceosomes recognize sequences at the 5' and 3' end of the intron.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TkuNUJis@3/Transcription

Translation

The synthesis of proteins consumes more of a cell's energy than any other metabolic process. In turn, proteins account for more mass than any other component of living organisms (with the exception of water), and proteins perform virtually every function of a cell. The process of translation, or **protein synthesis,** involves the decoding of an mRNA message into a polypeptide (protein) product. Amino acids are covalently strung together by peptide bonds in lengths ranging from approximately 50 amino acid residues to more than 1,000. Each individual amino acid has an amino group (NH_2) and a carboxyl (COOH) group. Polypeptides are formed when the amino group of one amino acid forms an amide (i.e., peptide) bond with the carboxyl group of another amino acid (Figure 1). This reaction is catalyzed by ribosomes and generates one water molecule.

Remember that there are 20 different amino acids that are commonly used. In Figure 1, R' represents the part of the amino acid which is different between these 20 structures.

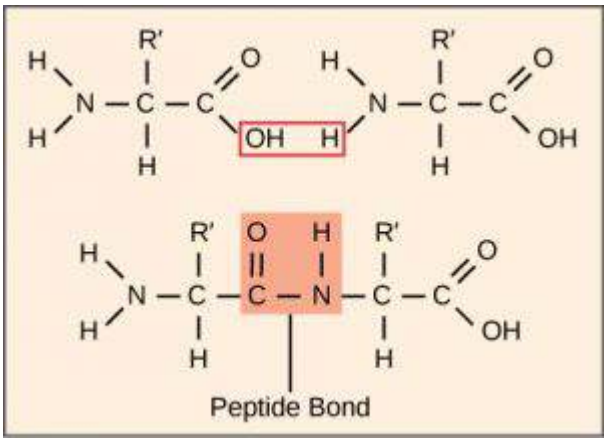


Figure 1 A peptide bond links the carboxyl end of one amino acid with the amino end of another, expelling one water molecule. For simplicity in this image, only the functional groups involved in the peptide bond are shown. The R and R' designations refer to the rest of each amino acid structure.

THE PROTEIN SYNTHESIS MACHINERY

In addition to the mRNA template, many other molecules contribute to the process of translation and the general structures and functions of the protein synthesis machinery are comparable from bacteria to human cells. Translation requires the input of an mRNA template, **ribosomes, tRNAs,** and various enzymatic factors (**Figure 1**).

RIBOSOMES

Even before an mRNA is translated, a cell must invest energy to build each of its ribosomes. **Ribosomes** are the part of the cell which reads the information in the mRNA molecule and joins amino acids together in the correct order. A ribosome is a very large, complex macromolecule composed of structural

and catalytic rRNAs, and many distinct polypeptides. In eukaryotes, the nucleolus is completely specialized for the synthesis and assembly of rRNAs (the RNA component that makes up ribosomes).

Ribosomes are made up of two subunits that come together for translation, rather like a hamburger bun comes together around the meat (the mRNA). The small subunit is responsible for binding the mRNA template, whereas the large subunit sequentially binds **tRNAs**, a type of RNA molecule that brings amino acids to the growing chain of the polypeptide. Each mRNA molecule can be simultaneously translated by many ribosomes, all synthesizing protein in the same direction: reading the mRNA from 5' to 3' and synthesizing the polypeptide from the N terminus to the C terminus (refer to Figure 1 – the N terminus is the end of the amino acid with the Nitrogen; the C terminus is the end with the Carbon).

Ribosomes exist in the cytoplasm in prokaryotes and in the cytoplasm and rough endoplasmic reticulum in eukaryotes. Mitochondria and chloroplasts also have their own ribosomes in the matrix and stroma, which look more similar to prokaryotic ribosomes (and have similar drug sensitivities) than the ribosomes just outside their outer membranes in the cytoplasm.

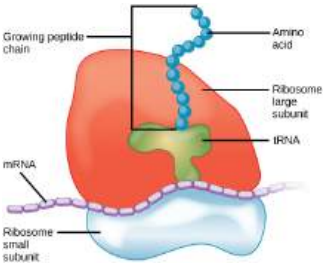


Figure 2 The protein synthesis machinery includes the large and small subunits of the ribosome, mRNA, and tRNA.

TRNA

Depending on the species, 40 to 60 types of **tRNA** exist in the cytoplasm. Serving as adaptors, specific tRNAs bind to sequences on the mRNA template and add the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually "translate" the language of RNA into the language of proteins.

Each tRNA is made up of a linear RNA molecule that is folded into a complex shape (Figure 3). At one end of the tRNA is an anticodon, which recognizes and base pairs with one of the mRNA codons. At the other end, a specific amino acid is attached. Of the 64 possible mRNA codons—or triplet combinations of A, U, G, and C—three specify the termination of protein synthesis and 61 specify the addition of amino acids to the polypeptide chain. Of these 61, one codon (AUG) also encodes the initiation of translation. Each tRNA anticodon can base pair with one of the mRNA codons and add an amino acid or terminate translation, according to the genetic code. For instance, if the sequence CUA occurred on an mRNA template in the proper reading frame, it would bind

a tRNA expressing the complementary sequence, GAU, which would be linked to the amino acid leucine.

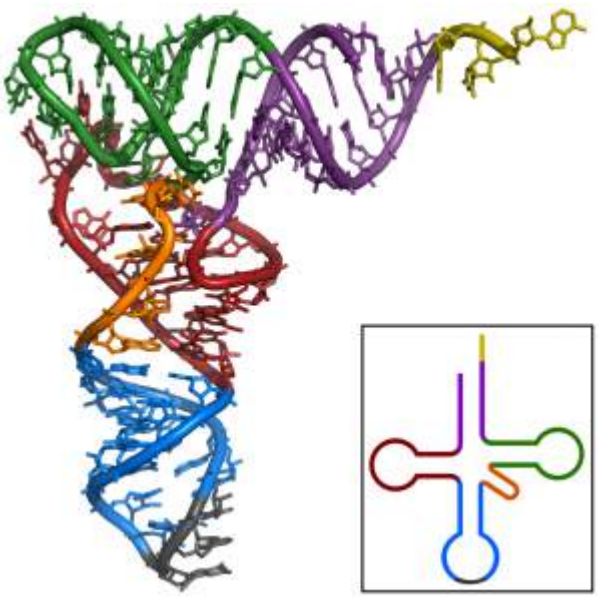


Figure 3 The RNA molecule that makes up a tRNA folds into the complex 3-D structure seen here. In this figure, the anticodon is the grey section at the bottom of the structure. The amino acid would be attached to the yellow part at the top right. Photo credit Yikrazuul; Wikimedia.

Aminoacyl tRNA Synthetases

For each tRNA to function, it must have its specific amino acid bonded to it. In the process of tRNA "charging," each tRNA molecule is bonded to its correct amino acid by a group of enzymes called **aminoacyl tRNA synthetases**. At least one type of aminoacyl tRNA synthetase exists for each of the 20 amino acids; the exact number of aminoacyl tRNA synthetases varies by species. These enzymes utilize the energy from ATP to energize a specific amino acid, which is

then transferred to the tRNA. In this way, tRNA molecules can be used over and over again, but each tRNA always carries the same amino acid because of the specificity of the aminoacyl tRNA synthetase enzymes.

THE MECHANISM OF PROTEIN SYNTHESIS

As with mRNA synthesis, protein synthesis can be divided into three phases: initiation, elongation, and termination. The process of translation is similar in prokaryotes and eukaryotes. Here we'll explore how translation occurs in *E. coli*, a representative prokaryote, and specify any differences between prokaryotic and eukaryotic translation.

INITIATION OF TRANSLATION

Protein synthesis begins with the formation of an initiation complex. In *E. coli*, this complex involves the small ribosomal subunit, the mRNA template, three initiation factors (IF-1, IF-2, and IF-3), and a special initiator tRNA, called $tRNA_F^{Met}$. The initiator tRNA interacts with the start codon AUG, links to a formylated methionine amino acid called fMet, and can also bind IF-2. Formylated methionine is inserted by fMet-$tRNA_F^{Met}$ at the beginning of every polypeptide chain synthesized by *E. coli*, but it is usually clipped off after translation is complete. When an in-frame AUG is encountered during translation elongation, a non-formylated methionine is inserted by a regular Met-$tRNA^{Met}$.

In *E. coli* mRNA, a sequence upstream of the first AUG codon, called the Shine-Dalgarno sequence (AGGAGG), interacts with the rRNA molecules that compose the ribosome. This interaction anchors the small ribosomal subunit at the correct location on the mRNA template. Guanosine triphosphate (GTP), which is a purine nucleotide

triphosphate, acts as an energy source during translation—both at the start of elongation and during the ribosome's translocation.

In eukaryotes, a similar initiation complex forms, comprising mRNA, the small ribosomal subunit, IFs, and nucleoside triphosphates (GTP and ATP). The charged initiator tRNA, called Met-tRNA$_i$, does not bind fMet in eukaryotes, but is distinct from other Met-tRNAs in that it can bind IFs. Like in E. coli, a "normal" methionine amino acid is inserted when the ribosome encounters in-frame AUG codons.

Instead of depositing at the Shine-Dalgarno sequence, the eukaryotic initiation complex recognizes the 7-methylguanosine cap at the 5' end of the mRNA. A cap-binding protein (CBP) and several other IFs assist the movement of the ribosome to the 5' cap. Once at the cap, the initiation complex tracks along the mRNA in the 5' to 3' direction, searching for the AUG start codon. Many eukaryotic mRNAs are translated from the first AUG, but this is not always the case. The sequence of bases around each AUG helps determine if that AUG will be used as the start codon.

Once the appropriate AUG is identified, the large ribosomal subunit binds to the complex of Met-tRNA$_i$, mRNA, and the small ribosomal subunit. This step completes the initiation of translation in eukaryotes.

Summary: In both prokaryotes and eukaryotes, the small ribosomal subunit binds to the special initiator methionine tRNA. With the help of several other factors, this complex identifies the start codon (AUG) based on the sequence of nucleotides nearby (Figure 4, top diagram). Then, the large ribosomal subunit binds (Figure 4, middle diagram).

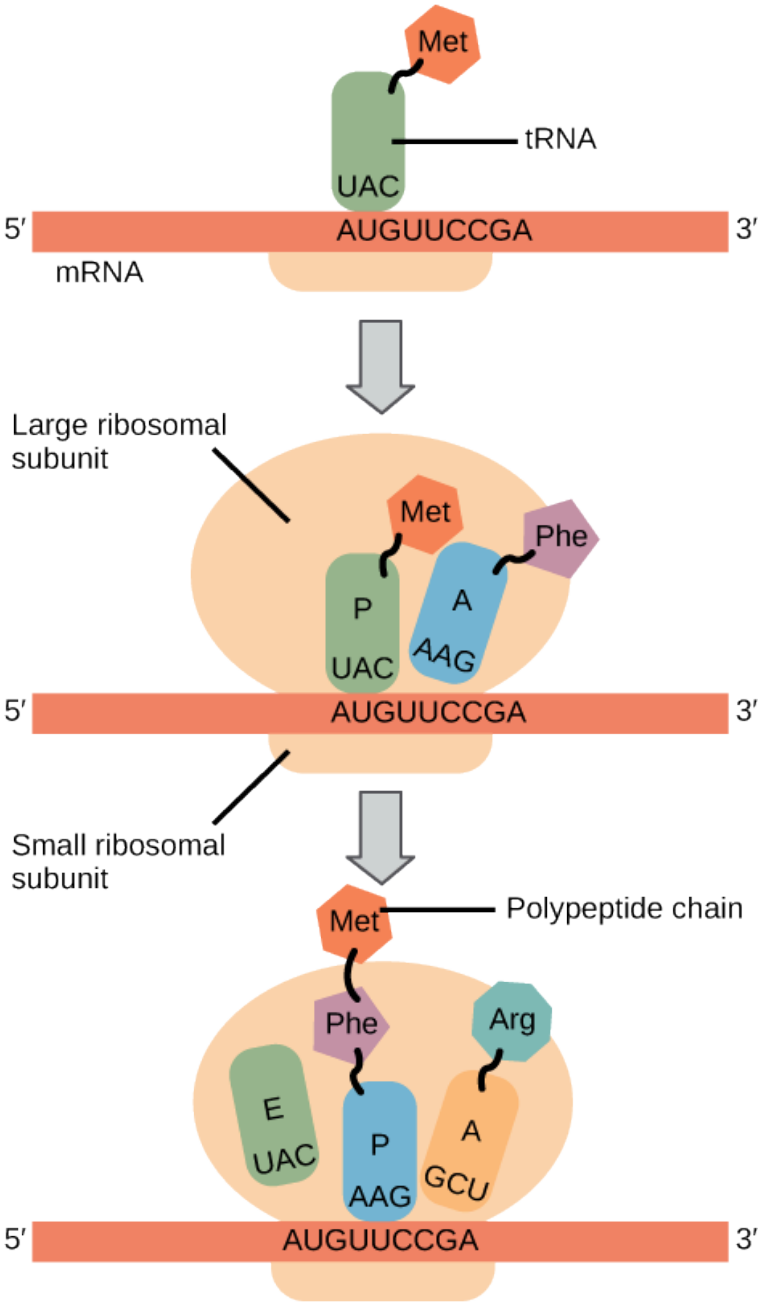

Met
tRNA
UAC
5′
AUGUUCCGA
3′
mRNA
Large ribosomal subunit
Met
Phe
P
A
UAC
AAG
5′
AUGUUCCGA
3′
Small ribosomal subunit
Met
Polypeptide chain
Phe
Arg
E
P
A
UAC
AAG
GCU
5′
AUGUUCCGA
3′

Figure 4 Translation begins when a tRNA anticodon recognizes a codon on the mRNA. The large ribosomal subunit joins the small subunit, and a second tRNA is recruited. As the mRNA moves relative to the ribosome, the polypeptide chain is formed. Entry of a release factor into the A site terminates translation and the components dissociate.

TRANSLATION, ELONGATION, AND TERMINATION

In prokaryotes and eukaryotes, the basics of elongation are the same. The large ribosomal subunit of consists of three compartments: the **A (aminoacyl) site** binds incoming charged aminoacyl tRNAs. The **P (peptidyl) site** binds charged tRNAs carrying amino acids that have formed peptide bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA. The **E (exit) site** releases dissociated tRNAs so that they can be recharged with free amino acids. There is one exception to this assembly line of tRNAs: in *E. coli*, fMet-tRNA$_F^{Met}$ is capable of entering the P site directly without first entering the A site. Similarly, the eukaryotic Met-tRNA$_i$, with help from other proteins of the initiation complex, binds directly to the P site. In both cases, this creates an initiation complex with a free A site ready to accept the tRNA corresponding to the first codon after the AUG.

During translation elongation, the mRNA template provides specificity. As the ribosome moves along the mRNA, each mRNA codon comes into register, and specific binding with the corresponding charged tRNA anticodon is ensured. If mRNA were not present in the elongation complex, the ribosome would bind tRNAs nonspecifically and a nonsense protein would be produced.

Elongation proceeds with charged tRNAs entering the A site and then shifting to the P site followed by the E site with each single-codon “step” of the ribosome (Figure 4, bottom diagram). Ribosomal steps are induced by conformational

changes that advance the ribosome by three bases in the 3' direction. The energy for each step of the ribosome is donated by an elongation factor that hydrolyzes GTP. Peptide bonds form between the amino group of the amino acid attached to the A-site tRNA and the carboxyl group of the amino acid attached to the P-site tRNA. The formation of each peptide bond is catalyzed by peptidyl transferase, an RNA-based enzyme that is integrated into the large ribosomal subunit. The energy for each peptide bond formation is derived from GTP hydrolysis, which is catalyzed by a separate elongation factor. The amino acid bound to the P-site tRNA is also linked to the growing polypeptide chain. As the ribosome steps across the mRNA, the former P-site tRNA enters the E site, detaches from the amino acid, and is expelled (Figure 4). Amazingly, the *E. coli* translation apparatus takes only 0.05 seconds to add each amino acid, meaning that a 200-amino acid protein can be translated in just 10 seconds.

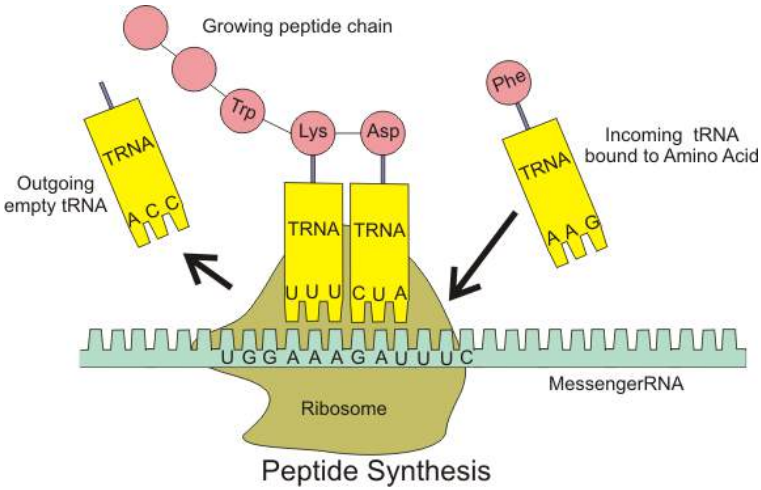


Figure 5 The movement of the tRNA molecules through the ribosome during protein synthesis. Note that the ribosome is moving from 5' to 3' along the mRNA, and the tRNAs are coming in from the front (the 3' direction) and exiting at the back (the 5' direction). Photo credit Boumphreyfr; Wikimedia.

REFERENCES

OpenStax, Biology. OpenStax CNX. January 2, 2017 http://cnx.org/contents/s8Hh0oOc@9.10:FUH9XUkW@6/Translation

Optional Section - Micropigs

Micropigs are tiny, genetically-edited pigs that have recently been developed by a Chinese genomics institute (Li, 2014). The Chinese scientists used a technique called TALENs to edit the genome of pig cells (Figure 1). Each cell inside a pig contains two copies of the growth hormone receptor gene: one from each of its parents. The TALENs technique was used to delete one of the two copies of this gene.

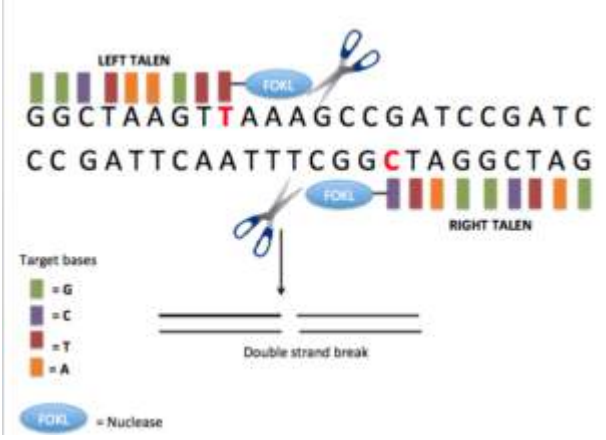


Figure 1 General overview of the TALEN process. The left and right TALEN bind to a specific sequence of genomic DNA inside the nucleus of a cell. When they are correctly bound, nuclease enzymes (represented by scissors) cut the genomic DNA. The TALEN sequences can be edited by scientists to target different DNA sequences in the genome. (Photo credit: Ogletreerd, Wikimedia.

Growth hormone (GH) stimulates the growth of essentially all tissues within the body. GH is a 191 amino acid peptide (protein) hormone which is produced from the GH gene. Cells sense the presence of GH protein hormone with the growth hormone receptor (GHR) protein on the outside of the cell. GHR protein is produced from the GHR gene and is found on the cell membrane on the outside of cells.
The GHR protein has three major parts:

- An extracellular region that sticks out from the outside surface of the cell
- A transmembrane region that goes through the cell membrane and anchors the receptor to the membrane
- An intracellular region on the inside of the cell membrane that transmits signals to the interior of the cell.

The extracellular region binds (attaches) to GH, fitting together like a lock and its key. The binding of growth hormone transmits signals through the cell membrane to the intracellular region of the receptor (Figure 2). These signals "turn on" genes involved in growth and metabolism so that those genes are made into proteins. These proteins stimulate the growth and division of other cells in the organism.

If growth hormone is not present, the organism will not grow to full size. In humans, severe GH deficiency can lead to an adult height of only 4 feet tall. If growth hormone receptor is not present, the "grow" signal from the GH will not be transmitted inside of cells, so growth will not be stimulated (Figure 3).

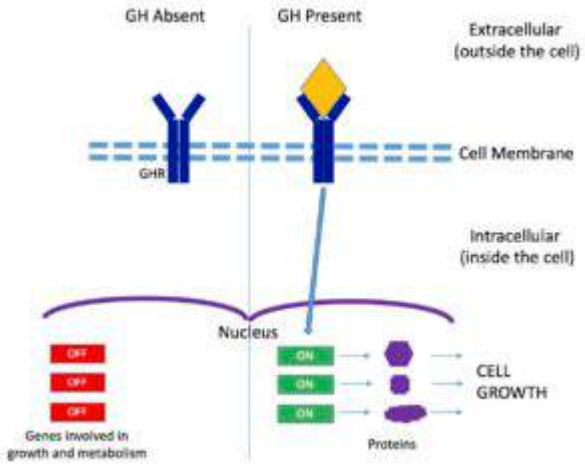


Figure 2 Growth hormone signaling pathway. When GH (growth hormone) binds to GHR (growth hormone receptor), a signal is sent through the cell membrane and into the nucleus of the cell. This signal turns on genes involved in cell metabolism and growth. (Photo credit: Lisa Bartee, 2017)

Figure 3: Micro pigs next to a tea cup. Image from: https://www.dailymail.co.uk/news/article-1218472/The-700-teacup-sized-pigs-latest-celebrity-pet.html

REFERENCES

Li f, Li Y, Liu H, Zhang H, Liu C, Zhang X, Dou H, Yang W, Du Y. 2014 Sep. Production of GHR double-allelic knockout Bama pig by TALENs and handmade cloning. Yi Chuan: 36(9):903-11.

GENE REGULATION

Learning Objectives

By the end of this section, you will be able to:

- Describe processes through which gene expression can be regulated.
- Differentiate between gene regulation processes used by prokaryotes and eukaryotes.
- Discuss the possible evolutionary consequences of changes in gene expression.

Each cell expresses, or turns on, only a fraction of its genes. "Expresses" or "turns on" means that protein is being produced from that gene. The rest of the genes are repressed, or turned off (no protein is being produced from those genes). The process of turning genes on and off is known as **gene regulation**. Gene regulation is an important part of normal development. Genes are turned on and off in different patterns during development to make a brain cell look and act different from a liver cell or a muscle cell, for

example. Gene regulation also allows cells to react quickly to changes in their environments. Although we know that the regulation of genes is critical for life, this complex process is not yet fully understood.

For a cell to function properly, necessary proteins must be synthesized at the proper time. All organisms and cells control or regulate the transcription and translation of their DNA into protein. The process of turning on a gene to produce RNA and protein is called **gene expression**. Whether in a simple unicellular organism or in a complex multicellular organism, each cell controls when and how its genes are expressed. For this to occur, there must be a mechanism to control when a gene is expressed to make RNA and protein, how much of the protein is made, and when it is time to stop making that protein because it is no longer needed.

Cells in multicellular organisms are specialized; cells in different tissues look very different and perform different functions. For example, a muscle cell is very different from a liver cell, which is very different from a skin cell. These differences are a consequence of the expression of different sets of genes in each of these cells. All cells have certain basic functions they must perform for themselves, such as converting the energy in sugar molecules into energy in ATP. Therefore, there is a set of "housekeeping" genes that are expressed in all cells. Each type of cell also has many genes that are not expressed because the cell does not need to perform those functions. Specific cells also express many genes that are not expressed by other cells so that they can carry out their specialized functions. In addition, cells will turn on or off certain genes at different times in response to changes in the environment or at different times during the development of the organism. Unicellular organisms, both eukaryotic and prokaryotic, also turn on and off genes in

response to the demands of their environment so that they can respond to special conditions.

Figure 1 The unique color pattern of this cat's fur is caused by either the orange or the black allele of a gene being randomly silenced (turned off).

The control of gene expression is extremely complex. Malfunctions in this process are detrimental to the cell and can lead to the development of many diseases, including cancer.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Prokaryotic versus Eukaryotic Gene Expression

To understand how gene expression is regulated, we must first understand how a gene becomes a functional protein in a cell. The process occurs in both prokaryotic and eukaryotic cells, just in slightly different fashions.

Because prokaryotic organisms lack a cell nucleus, the processes of transcription and translation occur almost simultaneously. When the protein is no longer needed, transcription stops. When there is no mRNA present, no protein can be made. As a result, the primary method to control what type and how much protein is expressed in a prokaryotic cell is through the regulation of DNA transcription into RNA. All the subsequent steps happen automatically. When more protein is required, more transcription occurs. Therefore, in prokaryotic cells, the control of gene expression is almost entirely at the transcriptional level.

Eukaryotic cells, in contrast, have intracellular organelles and are much more complex. Recall that in eukaryotic cells, the DNA is contained inside the cell's nucleus and that is where it is transcribed to produce mRNA. The newly synthesized mRNA is transported out of the nucleus into the cytoplasm, where ribosomes translate the mRNA to produce protein. The processes of transcription and translation are physically separated by the nuclear membrane; transcription occurs only within the nucleus, and translation only occurs outside the nucleus in the cytoplasm. The regulation of gene expression can occur at any stage of the process (**Figure 1**):

- **Epigenetic level**: regulates how tightly the DNA is wound around histone proteins to package it into chromosomes
- **Transcriptional level**: regulates how much transcription takes place
- **Post-transcriptional level**: regulates aspects of RNA processing (such as splicing) and transport out

of the nucleus

- **Translational level**: regulates how much of the RNA is translated into protein
- **Post-translational level**: regulates how long the protein lasts after it has been made and whether the protein is processed into an active form

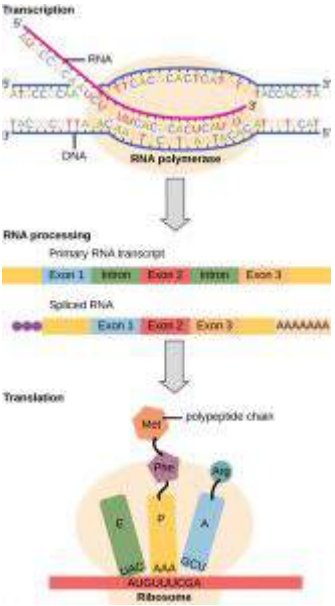


Figure 1 Eukaryotic gene expression is regulated during transcription and RNA processing, which take place in the nucleus, as well as during protein translation, which takes place in the cytoplasm. Further regulation may occur through post-translational modifications of proteins.

The differences in the regulation of gene expression between prokaryotes and eukaryotes are summarized in **Table 1**.

Table 1: Differences in the Regulation of Gene Expression of Prokaryotic and Eukaryotic Organisms

Prokaryotic organisms	Eukaryotic organisms
Lack nucleus	Contain nucleus
RNA transcription and protein translation occur almost simultaneously	RNA transcription occurs prior to protein translation, and it takes place in the nucleus. RNA translation to protein occurs in the cytoplasm. RNA post-processing includes addition of a 5' cap, poly-A tail, and excision of introns and splicing of exons.
Gene expression is regulated primarily at the transcriptional level	Gene expression is regulated at many levels (epigenetic, transcriptional, post-transcriptional, translational, and posttranslational)

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Eukaryotic epigenetic regulation

DNA modifications that do not change the DNA sequence can affect gene activity. Chemical compounds that are added to single genes can regulate their activity; these modifications are known as **epigenetic** changes. The **epigenome** comprises all of the chemical compounds that have been added to the entirety of one's DNA (genome) as a way to regulate the activity (expression) of all the genes within the genome. The chemical compounds of the epigenome are not part of the DNA sequence, but are on or attached to DNA ("epi-" means above in Greek). Epigenomic modifications remain as cells divide and in some cases can be inherited through the generations. Environmental influences, such as a person's diet and exposure to pollutants, can also impact the epigenome.

Epigenetic changes can help determine whether genes are turned on or off and can influence the production of proteins in certain cells, ensuring that only necessary proteins are produced. For example, proteins that promote bone growth are not produced in muscle cells. Patterns of epigenome modification vary among individuals, different tissues within an individual, and even different cells.

REGULATING ACCESS TO GENES

The human genome encodes over 20,000 genes, which means that each of the 23 pairs of human chromosomes contains thousands of genes. The DNA in the nucleus of each cell is precisely wound, folded, and compacted into chromosomes so that it will fit inside the nuclear membrane. It is also organized so that specific segments can be accessed as needed by a specific cell type.

The first level of organization, or packing, is the winding of DNA strands around **histone** proteins. Histones package and order DNA into structural units called **nucleosome**

complexes, which can control the access of proteins to the DNA regions (Figure 1**a**). Under the electron microscope, this winding of DNA around histone proteins to form nucleosomes looks like small beads on a string (Figure 1**b**). These beads (histone proteins) can move along the string (DNA) and change the structure of the molecule.

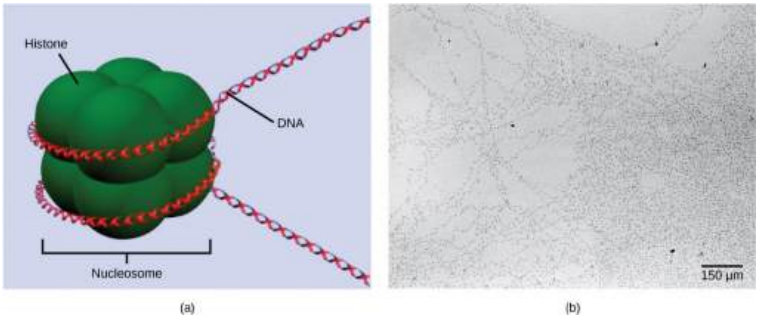


Figure 1 DNA is folded around histone proteins to create (a) nucleosome complexes. These nucleosomes control the access of proteins to the underlying DNA. When viewed through an electron microscope (b), the nucleosomes look like beads on a string. (credit "micrograph": modification of work by Chris Woodcock)

If DNA encoding a specific gene is to be transcribed into RNA, the nucleosomes surrounding that region of DNA can slide down the DNA to open that specific chromosomal region and allow for the transcriptional machinery (RNA polymerase) to initiate transcription. Nucleosomes can move to open the chromosome structure to expose a segment of DNA, but do so in a very controlled manner.

Active open regions of chromatin are called **euchromatin (Figure 2).** Regions of the genome that are transcriptionally active are typically euchromatic. Tightly wound regions of chromatin are called **heterochromatin.** Heterochromatic regions of the genome are typically silenced and transcriptionally inactive.

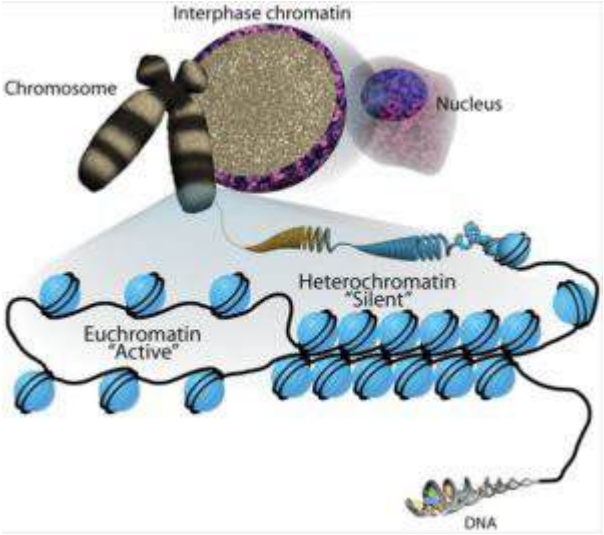


Figure 2 The difference in chromatin packaging between an active (euchromatic) and inactive (heterochromatic) region of DNA.

MODIFICATIONS TO DNA AND HISTONES

How the histone proteins move, and whether the DNA is wrapped loosely or tightly around them, is dependent on signals found on both the histone proteins and on the DNA. These signals are chemical tags added to histone proteins and DNA that tell the histones if a chromosomal region should be open or closed. These tags are not permanent, but may be added or removed as needed. They are chemical modifications (phosphate, methyl, or acetyl groups) that are attached to specific amino acids in the protein or to the nucleotides of the DNA. The tags do not alter the DNA base sequence, but they do alter how tightly wound the DNA is around the histone proteins.

This type of gene regulation is called **epigenetic regulation**. Epigenetic means "around or above genetics."

The changes that occur to the histone proteins and DNA do not alter the nucleotide sequence and are not permanent. Instead, these changes are temporary, although they can and often do persist through multiple rounds of cell division. They alter the chromosomal structure (open euchromatin or closed heterochromatin) as needed, but do not change the sequence of bases within the DNA.

A gene can be turned on or off depending upon the location and modifications to the histone proteins and DNA. If a gene is to be transcribed, the histone proteins and DNA are modified surrounding the chromosomal region encoding that gene. This opens the chromosomal region (it becomes euchromatic) to allow access for RNA polymerase and other proteins, called transcription factors, to bind to the promoter region, located just upstream of the gene, and initiate transcription. If a gene is to remain turned off, or silenced, the histone proteins and DNA have different modifications that signal a closed chromosomal configuration. In this closed configuration (heterochromatin), the RNA polymerase and transcription factors do not have access to the DNA and transcription cannot occur (Figure 2).

DNA METHYLATION

A common type of epigenomic modification is called **methylation.** Methylation involves attaching small molecules called methyl groups, each consisting of one carbon atom and three hydrogen atoms, to DNA nucleotides or the amino acids that make up the histone proteins.

When DNA is methylated, the methyl group is typically added to cytosine nucleotides. This occurs within very specific regions called CpG islands. These are stretches with a high frequency of cytosine and guanine dinucleotide DNA pairs (CG) found in the promoter regions of genes. When this configuration exists, the cytosine member of the pair can

be methylated (a methyl group is added). This modification changes how the DNA interacts with proteins, including the histone proteins that control access to the region. When methyl groups are added to a particular gene, that gene is turned off or silenced, and no protein is produced from that gene (Figure 3).

"HISTONE CODE" HYPOTHESIS

The **histone code hypothesis** is the hypothesis that transcription of a gene is in part regulated by modifications made to histone proteins, primarily on their somewhat floppy ends (their "tails"). Many of the histone tail modifications correlate very well to chromatin structure and both histone modification state and chromatin structure correlates well to gene expression levels. The most important concept in the histone code hypothesis is that the histone modifications serve to recruit other proteins by specific recognition of the modified histone, rather than through simply stabilizing or destabilizing the interaction between histone and the underlying DNA. These recruited proteins then act to alter chromatin structure actively or to promote transcription.

The histone code has the potential to be massively complex. There are at least 20 modifications that are made to histone tails that have been relatively well characterized, and there is the potential for many more that we have not discovered. Each histone can be modified on multiple amino acids, with multiple different chemical modifications. The information that can be stored in the histone code dwarfs the amount that is stored in the order of the bases in the human genome.

HISTONE METHLYATION

A portion of the histone protein known as the histone tail can have methyl groups (CH_3) added to it. This is the same modification that is made to cytosine nucleotides in DNA. The specific amino acid in the histone tail that gets methylated is very important for determining whether it will tighten or loosen chromatin structure. Modification to several amino acids in the tail is correlated with euchromatin and active transcription, while modification to other amino acids is correlated with heterochromatin and gene silencing. You should know that histones can be methylated, but we can't use **histone methylation** as a predictor for euchromatin or heterochromatin.

HISTONE ACETYLATION

Histone tails can also be modified by the addition of an acetyl group (this process is known as **histone acetylation**). If you remember from cellular respiration, an acetyl group (such as that found in acetyl-CoA) is a 2-carbon molecule. When histone tails are acetylated, this typically causes the tails to loosen from around the DNA, allowing the chromatin to loosen (Figure 3).

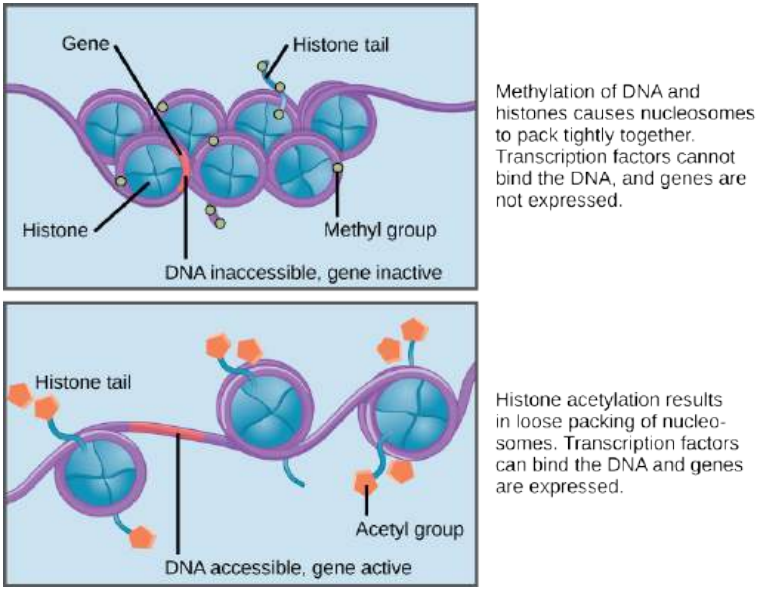


Figure 3 Nucleosomes can slide along DNA. When nucleosomes are spaced closely together (top), transcription factors cannot bind and gene expression is turned off. When the nucleosomes are spaced far apart (bottom), the DNA is exposed. Transcription factors can bind, allowing gene expression to occur. Modifications to the histones and DNA affect nucleosome spacing.

OTHER MODIFICATIONS

There are many other modifications that can be made to histone proteins in addition to methylation and acetylation. Histone tails can be phosphorylated or ubiquitinated (where a small protein called ubiquitin is attached). Histone phosphorylation seems to be related to DNA repair. Ubiquitination has been shown to be associated with both transcriptional activation or inactivation, depending on the specific location.

EPIGENETIC CHANGES

Because errors in the epigenetic process, such as modifying the wrong gene or failing to add a compound to a gene, can lead to abnormal gene activity or inactivity, they can cause genetic disorders. Conditions including cancers, metabolic disorders, and degenerative disorders have all been found to be related to epigenetic errors.

Cancerous cells often have regions of DNA that show different levels of methylation compared to normal cells. Some genes are methylated and silenced in cancerous cells, while they are unmethylated and active in normal cells. Other genes are active in cancerous cells, but inactive in normal cells. Each specific cancer in each specific individual can show different patterns of methylation, although there are similarities between many different types of cancer.

Scientists continue to explore the relationship between the genome and the chemical compounds that modify it. In particular, they are studying what effect the modifications have on gene function, protein production, and human health.

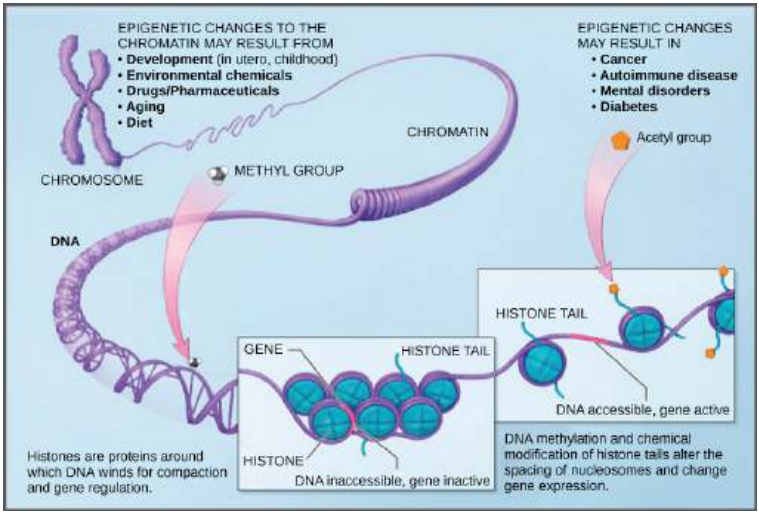


Figure 4 Histone proteins and DNA nucleotides can be modified chemically. Modifications affect nucleosome spacing and gene expression. (credit: modification of work by NIH)

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Eukaryotic Transcriptional Regulation

Like prokaryotic cells, the transcription of genes in eukaryotes requires the actions of an RNA polymerase to bind to a sequence upstream of a gene to initiate transcription. However, unlike prokaryotic cells, the eukaryotic RNA polymerase requires other proteins, or transcription factors, to facilitate transcription initiation. **Transcription factors** are proteins that bind to the promoter sequence and other regulatory sequences to control the transcription of the target gene. RNA polymerase by itself cannot initiate transcription in eukaryotic cells. Transcription factors must bind to the promoter region first and recruit RNA polymerase to the site for transcription to be established.

THE PROMOTER AND THE TRANSCRIPTION MACHINERY

Genes are organized to make the control of gene expression easier. The **promoter region** is immediately upstream of the coding sequence. This region can be short (only a few nucleotides in length) or quite long (hundreds of nucleotides long). The longer the promoter, the more available space

for proteins to bind. This also adds more control to the transcription process. The length of the promoter is gene-specific and can differ dramatically between genes. Consequently, the level of control of gene expression can also differ quite dramatically between genes. The purpose of the promoter is to bind transcription factors that control the initiation of transcription.

The interaction between parts of the promoter and transcription factors is quite complex (Figure 1) and differs between specific genes. In addition to the general transcription factors, there are gene-specific transcription factors. There are hundreds of transcription factors in a cell that each bind specifically to a particular DNA sequence motif. Transcription factors respond to environmental stimuli that cause the proteins to find their binding sites and initiate transcription of the gene that is needed.

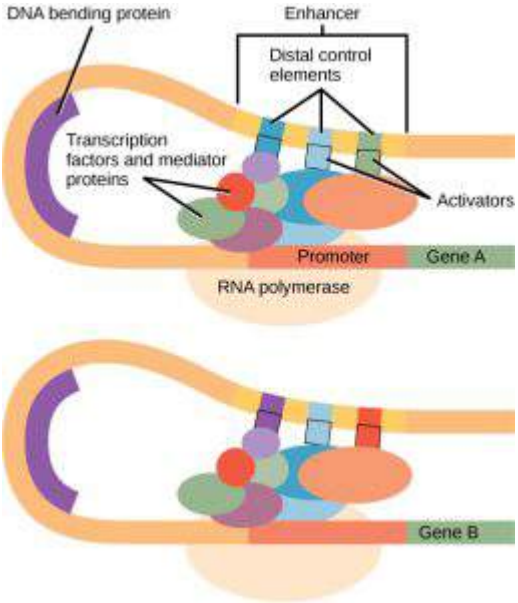


Figure 1 An enhancer is a DNA sequence that promotes transcription. Each enhancer is made up of short DNA sequences called distal control elements. Activators bound to the distal control elements interact with mediator proteins and transcription factors. Two different genes may have the same promoter but different distal control elements, enabling differential gene expression.

ENHANCERS AND TRANSCRIPTION

In some eukaryotic genes, there are regions that help increase or enhance transcription. These regions, called **enhancers**, are not necessarily close to the genes they enhance. They can be located upstream of a gene, within the coding region of the gene, downstream of a gene, or may be thousands of nucleotides away.

Enhancer regions are binding sequences, or sites, for transcription factors. When a DNA-bending protein binds, the shape of the DNA changes (Figure 1). This shape change

allows for the interaction of the activators bound to the enhancers with the transcription factors bound to the promoter region and the RNA polymerase. Whereas DNA is generally depicted as a straight line in two dimensions, it is actually a three-dimensional object. Therefore, a nucleotide sequence thousands of nucleotides away can fold over and interact with a specific promoter.

TURNING GENES OFF: TRANSCRIPTIONAL REPRESSORS

Like prokaryotic cells, eukaryotic cells also have mechanisms to prevent transcription. **Transcriptional repressors** can bind to promoter or enhancer regions and block transcription. Like the transcriptional activators, repressors respond to external stimuli to prevent the binding of activating transcription factors.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. January 3, 2017. https://cnx.org/contents/GFy_h8cu@10.120:7Ry3oRse@6/Eukaryotic-Transcription-Gene-

Eukaryotic Post-transcriptional Regulation

After RNA is transcribed, it must be processed into a mature form before translation can begin. This processing after an RNA molecule has been transcribed, but before it is translated into a protein, is called post-transcriptional modification. As with the epigenetic and transcriptional stages of processing, this post-transcriptional step can also be regulated to control gene expression in the cell. If the RNA is not processed, shuttled, or translated, then no protein will be synthesized.

ALTERNATIVE RNA SPLICING

In the 1970s, genes were first observed that exhibited **alternative RNA splicing**. Alternative RNA splicing is a mechanism that allows different protein products to be produced from one gene when different combinations of introns (and sometimes exons) are removed from the transcript (**Figure 1**). This alternative splicing can be haphazard, but more often it is controlled and acts as a mechanism of gene regulation, with the frequency of

different splicing alternatives controlled by the cell as a way to control the production of different protein products in different cells, or at different stages of development. Alternative splicing is now understood to be a common mechanism of gene regulation in eukaryotes; according to one estimate, 70% of genes in humans are expressed as multiple proteins through alternative splicing.

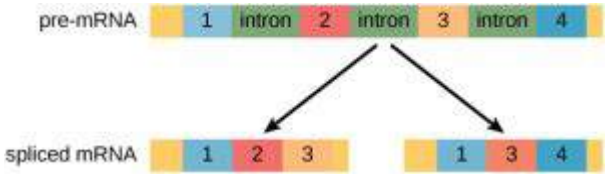


Figure 1 Pre-mRNA can be alternatively spliced to create different proteins.

How could alternative splicing evolve? Introns have a beginning and ending recognition sequence, and it is easy to imagine the failure of the splicing mechanism to identify the end of an intron and find the end of the next intron, thus removing two introns and the intervening exon. In fact, there are mechanisms in place to prevent such exon skipping, but mutations are likely to lead to their failure. Such "mistakes" would more than likely produce a nonfunctional protein. Indeed, the cause of many genetic diseases is alternative splicing rather than mutations in a sequence. However, alternative splicing would create a protein variant without the loss of the original protein, opening up possibilities for adaptation of the new variant to new functions. Gene duplication has played an important role in the evolution of new functions in a similar way—by providing genes that may evolve without eliminating the original functional protein.

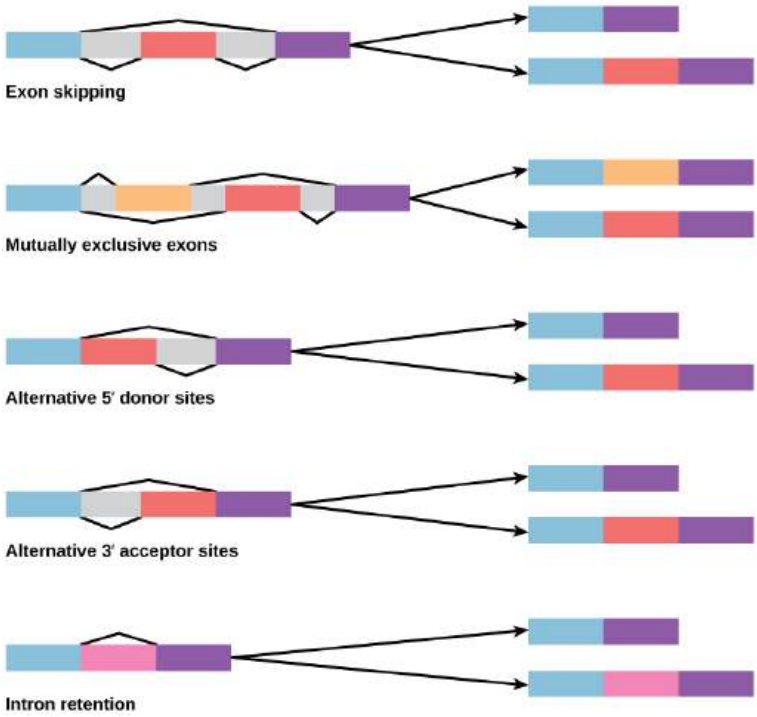


Figure 2 There are five basic modes of alternative splicing.

CONTROL OF RNA STABILITY

Before the mRNA leaves the nucleus, it is given two protective "caps" that prevent the end of the strand from degrading during its journey. The 5' cap, which is placed on the 5' end of the mRNA, is usually composed of a methylated guanosine triphosphate molecule (GTP). The poly-A tail, which is attached to the 3' end, is usually composed of a series of adenine nucleotides. Once the RNA is transported to the cytoplasm, the length of time that the RNA remains there can be controlled. Each RNA molecule has a defined lifespan and decays at a specific rate. This rate of decay can influence

how much protein is in the cell. If the RNA decays more rapidly, translation has less time to occur, so less protein will be produced. Conversely, if RNA decays less rapidly, more protein will be produced. This rate of decay is referred to as the RNA stability. If the RNA is stable, it will be detected for longer periods of time in the cytoplasm. Binding of proteins to the RNA can influence its stability (Figure 3).

Figure 3 The protein-coding region of mRNA is flanked by 5' and 3' untranslated regions (UTRs). The presence of RNA-binding proteins at the 5' or 3' UTR influences the stability of the RNA molecule.

RNA STABILITY AND MICRORNAS

In addition to proteins that bind to and control (increase or decrease) RNA stability, other elements called microRNAs can bind to the RNA molecule. These microRNAs, or miRNAs, are short RNA molecules that are only 21–24 nucleotides in length. The miRNAs are made in the nucleus as longer pre-miRNAs. These pre-miRNAs are chopped into mature miRNAs by a protein called dicer. Together, miRNAs and a large protein complex called RISC rapidly destroy the RNA molecule.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. January 3, 2017. https://cnx.org/contents/

GFy_h8cu@10.120:wf3BMwl1@9/Eukaryotic-Post-transcriptiona

Eukaryotic Translational and Post-Translational Regulation

After the RNA has been transported to the cytoplasm, it is translated into protein. Control of this process is largely dependent on the RNA molecule. As previously discussed, the stability of the RNA will have a large impact on its translation into a protein. As the stability changes, the amount of time that it is available for translation also changes.

THE INITIATION COMPLEX AND TRANSLATION RATE

Like transcription, translation is controlled by proteins that bind and initiate the process. In translation, the complex that assembles to start the process is referred to as the **initiation complex**. Regulation of the formation of this complex can increase or decrease rates of translation (Figure 1).

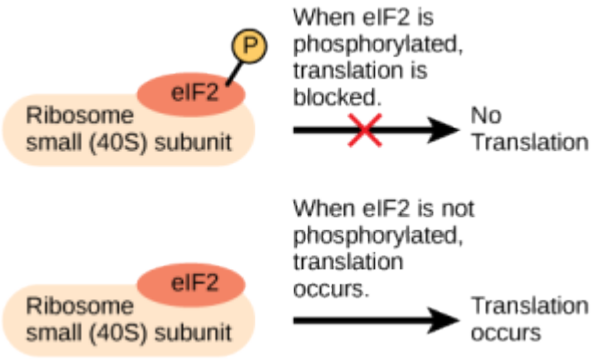


Figure 1 Gene expression can be controlled by factors that bind the translation initiation complex.

CHEMICAL MODIFICATIONS, PROTEIN ACTIVITY, AND LONGEVITY

Proteins can be chemically modified with the addition of groups including methyl, phosphate, acetyl, and ubiquitin groups. The addition or removal of these groups from proteins regulates their activity or the length of time they exist in the cell. Sometimes these modifications can regulate where a protein is found in the cell—for example, in the nucleus, the cytoplasm, or attached to the plasma membrane.

Chemical modifications occur in response to external stimuli such as stress, the lack of nutrients, heat, or ultraviolet light exposure. These changes can alter epigenetic accessibility, transcription, mRNA stability, or translation—all resulting in changes in expression of various genes. This is an efficient way for the cell to rapidly change the levels of specific proteins in response to the environment. Because proteins are involved in every stage of gene regulation, the phosphorylation of a protein (depending on the protein that is modified) can alter accessibility to the chromosome, can

alter translation (by altering transcription factor binding or function), can change nuclear shuttling (by influencing modifications to the nuclear pore complex), can alter RNA stability (by binding or not binding to the RNA to regulate its stability), can modify translation (increase or decrease), or can change post-translational modifications (add or remove phosphates or other chemical modifications).

The addition of an **ubiquitin group** to a protein marks that protein for degradation. Ubiquitin acts like a flag indicating that the protein lifespan is complete. These proteins are moved to the proteasome, an organelle that functions to remove proteins, to be degraded (Figure 2). One way to control gene expression, therefore, is to alter the longevity of the protein.

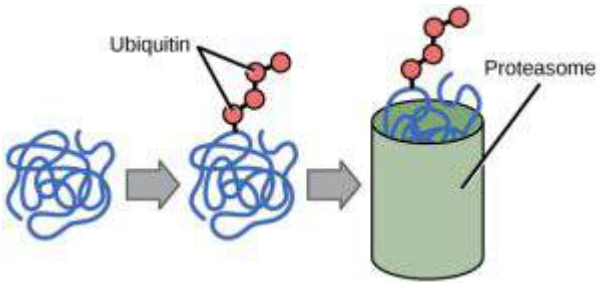


Figure 2 Proteins with ubiquitin tags are marked for degradation within the proteasome.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

Cancer and Gene Regulation

Cancer is not a single disease but includes many different diseases. In cancer cells, mutations modify cell-cycle control and cells don't stop growing as they normally would. Mutations can also alter the growth rate or the progression of the cell through the cell cycle. As a result, cells can progress through the cell cycle unimpeded, even if mutations exist in the cell and its growth should be terminated.

CANCER: DISEASE OF ALTERED GENE EXPRESSION

Cancer can be described as a disease of altered gene expression. There are many proteins that are turned on or off (gene activation or gene silencing) that dramatically alter the overall activity of the cell. A gene that is not normally expressed in that cell can be switched on and expressed at high levels. This can be the result of gene mutation or changes in any level of gene regulation (epigenetic, transcription, post-transcription, translation, or post-translation).

Changes in epigenetic regulation, transcription, RNA stability, protein translation, and post-translational control can be detected in cancer. While these changes don't occur simultaneously in one cancer, changes at each of these levels

can be detected when observing cancer at different sites in different individuals. Therefore, changes in histone acetylation (epigenetic modification that leads to gene silencing), activation of transcription factors by phosphorylation, increased RNA stability, increased translational control, and protein modification can all be detected at some point in various cancer cells. Scientists are working to understand the common changes that give rise to certain types of cancer or how a modification might be exploited to destroy a tumor cell.

TUMOR SUPPRESSOR GENES, ONCOGENES, AND CANCER

In normal cells, some genes function to prevent excess, inappropriate cell growth. These are **tumor suppressor genes**, which are active in normal cells to prevent uncontrolled cell growth. There are many tumor suppressor genes in cells. The most studied tumor suppressor gene is **p53**, which is mutated in over 50 percent of all cancer types. The p53 protein itself functions as a transcription factor. It can bind to sites in the promoters of genes to initiate transcription. Therefore, the mutation of p53 in cancer will dramatically alter the transcriptional activity of its target genes.

Proto-oncogenes are positive cell-cycle regulators (their normal function is to allow the cell cycle to progress through checkpoints). When mutated, proto-oncogenes can become oncogenes and cause cancer. Overexpression of the oncogene can lead to uncontrolled cell growth. This is because oncogenes can alter transcriptional activity, stability, or protein translation of another gene that directly or indirectly controls cell growth.

CANCER AND EPIGENETIC ALTERATIONS

Silencing genes through epigenetic mechanisms is also very common in cancer cells. There are characteristic modifications to histone proteins and DNA that are associated with silenced genes. In cancer cells, the DNA in the promoter region of silenced genes is methylated on cytosine DNA residues in CpG islands. Histone proteins that surround that region lack the acetylation modification that is present when the genes are expressed in normal cells. This combination of DNA methylation and histone deacetylation (epigenetic modifications that lead to gene silencing) is commonly found in cancer. When these modifications occur, the gene present in that chromosomal region is silenced. Increasingly, scientists understand how epigenetic changes are altered in cancer. Because these changes are temporary and can be reversed—for example, by preventing the action of the histone deacetylase protein that removes acetyl groups, or by DNA methyl transferase enzymes that add methyl groups to cytosines in DNA—it is possible to design new drugs and new therapies to take advantage of the reversible nature of these processes. Indeed, many researchers are testing how a silenced gene can be switched back on in a cancer cell to help re-establish normal growth patterns.

Genes involved in the development of many other illnesses, ranging from allergies to inflammation to autism, are thought to be regulated by epigenetic mechanisms. As our knowledge of how genes are controlled deepens, new ways to treat diseases like cancer will emerge.

CANCER AND TRANSCRIPTIONAL CONTROL

Alterations in cells that give rise to cancer can affect the transcriptional control of gene expression. Mutations that

activate transcription factors, such as increased phosphorylation, can increase the binding of a transcription factor to its binding site in a promoter. This could lead to increased transcriptional activation of that gene that results in modified cell growth. Alternatively, a mutation in the DNA of a promoter or enhancer region can increase the binding ability of a transcription factor. This could also lead to the increased transcription and aberrant gene expression that is seen in cancer cells.

Researchers have been investigating how to control the transcriptional activation of gene expression in cancer. Identifying how a transcription factor binds, or a pathway that activates where a gene can be turned off, has led to new drugs and new ways to treat cancer. In breast cancer, for example, many proteins are overexpressed. This can lead to increased phosphorylation of key transcription factors that increase transcription. One such example is the overexpression of the **epidermal growth factor receptor (EGFR)** in a subset of breast cancers. The EGFR pathway activates many protein kinases that, in turn, activate many transcription factors that control genes involved in cell growth. New drugs that prevent the activation of EGFR have been developed and are used to treat these cancers.

CANCER AND POST-TRANSCRIPTIONAL CONTROL

Changes in the post-transcriptional control of a gene can also result in cancer. Recently, several groups of researchers have shown that specific cancers have altered expression of miRNAs. Because miRNAs bind to the 3' UTR of RNA molecules to degrade them, overexpression of these miRNAs could be detrimental to normal cellular activity. Too many miRNAs could dramatically decrease the RNA population leading to a decrease in protein expression. Several studies have demonstrated a change in the miRNA population in

specific cancer types. It appears that the subset of miRNAs expressed in breast cancer cells is quite different from the subset expressed in lung cancer cells or even from normal breast cells. This suggests that alterations in miRNA activity can contribute to the growth of breast cancer cells. These types of studies also suggest that if some miRNAs are specifically expressed only in cancer cells, they could be potential drug targets. It would, therefore, be conceivable that new drugs that turn off miRNA expression in cancer could be an effective method to treat cancer.

CANCER AND TRANSLATIONAL/ POST-TRANSLATIONAL CONTROL

There are many examples of how translational or post-translational modifications of proteins arise in cancer. A variety of modifications can be found in cancer cells that cause effects that range from the increased translation of a protein to changes in protein phosphorylation to alternative splice variants of a protein. Keep in mind that translational control will result in the amount of protein that is translated, while post-translational control will result in the amount of active or modified protein found in the cell.

An example of how the expression of an alternative form of a protein can have dramatically different outcomes is seen in colon cancer cells. The c-Flip protein, a protein involved in mediating the cell death pathway, comes in two forms: long (c-FLIPL) and short (c-FLIPS). Both forms appear to be involved in initiating controlled cell death mechanisms in normal cells. However, in colon cancer cells, expression of the long form results in increased cell growth instead of cell death. Clearly, the expression of the wrong protein dramatically alters cell function and contributes to the development of cancer.

NEW DRUGS TO COMBAT CANCER: TARGETED THERAPIES

Scientists are using what is known about the regulation of gene expression in disease states, including cancer, to develop new ways to treat and prevent disease development. Many scientists are designing drugs on the basis of the gene expression patterns within individual tumors. This idea, that therapy and medicines can be tailored to an individual, has given rise to the field of personalized medicine. With an increased understanding of gene regulation and gene function, medicines can be designed to specifically target diseased cells without harming healthy cells. Some new medicines, called targeted therapies, have exploited the overexpression of a specific protein or the mutation of a gene to develop a new medication to treat disease. One such example is the use of anti-EGF receptor medications to treat the subset of breast cancer tumors that have very high levels of the EGF protein. Undoubtedly, more targeted therapies will be developed as scientists learn more about how gene expression changes can cause cancer.

REFERENCES

OpenStax, Concepts of Biology. OpenStax CNX. January 3, 2017. https://cnx.org/contents/GFy_h8cu@10.120:IbZZwBDM@6/Cancer-and-Gene-Regulation

MEIOSIS - SEXUAL REPRODUCTION

Learning Objectives

Course Outcomes for this section:

Apply biological theories and concepts to solve problems related to classical and molecular genetics

1. Describe the molecular basis of inheritance.

The ability to reproduce *in kind* is a basic characteristic of all living things. *In kind* means that the offspring of any organism closely resembles its parent or parents. Hippopotamuses give birth to hippopotamus calves; Monterey pine trees produce seeds from which Monterey pine seedlings emerge; and adult flamingos lay eggs that hatch into flamingo chicks. *In kind* does not generally mean *exactly the same*. While many single-celled organisms and a few multicellular organisms can produce genetically identical clones of themselves through mitotic cell division, many single-celled organisms

and most multicellular organisms reproduce regularly using another method.

Figure 1: Each of us, like these other large multicellular organisms, begins life as a fertilized egg. After trillions of cell divisions, each of us develops into a complex, multicellular organism. (credit a: modification of work by Frank Wouters; credit b: modification of work by Ken Cole, USGS; credit c: modification of work by Martin Pettitt)

Sexual reproduction is the production by parents of sex cells and the fusion of two sex cells to form a single, unique cell. In multicellular organisms, this new cell will then undergo mitotic cell divisions to develop into an adult organism. A type of cell division called meiosis leads to the cells that are part of the sexual reproductive cycle. Sexual reproduction, specifically meiosis and fertilization, introduces variation into offspring that may account for the evolutionary success of sexual reproduction. The vast majority of eukaryotic organisms can or must employ some form of meiosis and fertilization to reproduce.

Sexual reproduction was an early evolutionary innovation after the appearance of eukaryotic cells. The fact that most eukaryotes reproduce sexually is evidence of its evolutionary success. In many animals, it is the only mode of reproduction. And yet, scientists recognize some real disadvantages to sexual reproduction. On the surface, offspring that are

genetically identical to the parent may appear to be more advantageous. If the parent organism is successfully occupying a habitat, offspring with the same traits would be similarly successful. There is also the obvious benefit to an organism that can produce offspring by asexual budding, fragmentation, or asexual eggs. These methods of reproduction do not require another organism of the opposite sex. There is no need to expend energy finding or attracting a mate. That energy can be spent on producing more offspring. Indeed, some organisms that lead a solitary lifestyle have retained the ability to reproduce asexually. In addition, asexual populations only have female individuals, so every individual is capable of reproduction. In contrast, the males in sexual populations (half the population) are not producing offspring themselves. Because of this, an asexual population can grow twice as fast as a sexual population in theory. This means that in competition, the asexual population would have the advantage. All of these advantages to asexual reproduction, which are also disadvantages to sexual reproduction, should mean that the number of species with asexual reproduction should be more common.

However, multicellular organisms that exclusively depend on asexual reproduction are exceedingly rare. Why is sexual reproduction so common? This is one of the important questions in biology and has been the focus of much research from the latter half of the twentieth century until now. A likely explanation is that the variation that sexual reproduction creates among offspring is very important to the survival and reproduction of those offspring. The only source of variation in asexual organisms is mutation. This is the ultimate source of variation in sexual organisms. In addition, those different mutations are continually reshuffled from one generation to the next when different parents combine their unique genomes, and the genes are mixed

into different combinations by the process of **meiosis**. Meiosis is the division of the contents of the nucleus that divides the chromosomes among gametes. Variation is introduced during meiosis, as well as when the gametes combine in fertilization.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:qOUtHXNY@3/Sexual-Reproduction

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:6-3MVU-j@4/Errors-in-Meiosis

Overview of Meiosis

Sexual reproduction requires **fertilization**, a union of two cells from two individual organisms. If those two cells each contain one set of chromosomes, then the resulting cell contains two sets of chromosomes. The number of sets of chromosomes in a cell is called its ploidy level (Figure 1). **Haploid** cells contain one set of chromosomes. Cells containing two sets of chromosomes are called **diploid**. If the reproductive cycle is to continue, the diploid cell must somehow reduce its number of chromosome sets before fertilization can occur again, or there will be a continual doubling in the number of chromosome sets in every generation. So, in addition to fertilization, sexual reproduction includes a nuclear division, known as meiosis, that reduces the number of chromosome sets.

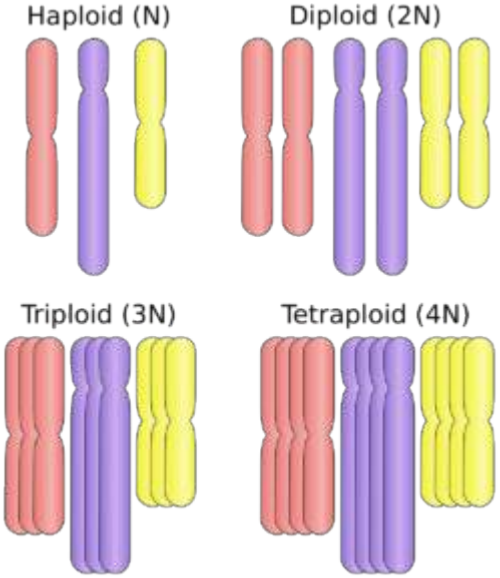


Figure 1 Number of chromosomes in a haploid and diploid cell. Note that triploid and tetraploid are not normal numbers of chromosomes in humans.

Most animals and plants are diploid, containing two sets of

chromosomes; in each **somatic cell** (the non-reproductive cells of a multicellular organism), the nucleus contains two copies of each chromosome that are referred to as homologous chromosomes. Somatic cells are sometimes referred to as "body" cells. **Homologous chromosomes** are matched pairs containing genes for the same traits in identical locations along their length (Figure 2). Diploid organisms inherit one copy of each homologous chromosome from each parent; all together, they are considered a full set of chromosomes. In animals, haploid cells containing a single copy of each homologous chromosome are found only within gametes. Gametes fuse with another haploid gamete to produce a diploid cell.

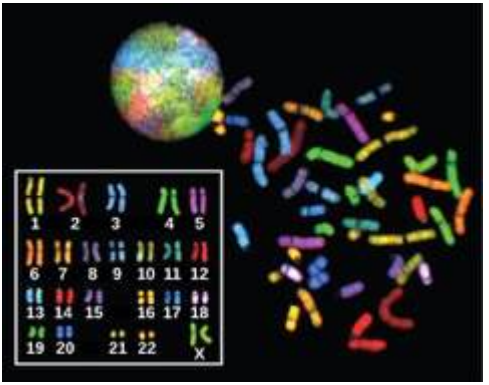


Figure 2 A karyotype displaying all of the chromosomes in the human genome. Note that there are two copies of each chromosome. These are the homologous chromosomes (one from each parent).

Nearly all animals employ a diploid-dominant life-cycle strategy in which the only haploid cells produced by the organism are the gametes. Early in the development of the embryo, specialized diploid cells, called germ cells, are produced within the gonads, such as the testes and ovaries. Germ cells are capable of mitosis to perpetuate the cell line and meiosis to produce gametes. Once the haploid gametes are formed, they lose the ability to divide again. There is

no multicellular haploid life stage. Fertilization occurs with the fusion of two gametes, usually from different individuals, restoring the diploid state (Figure 3).

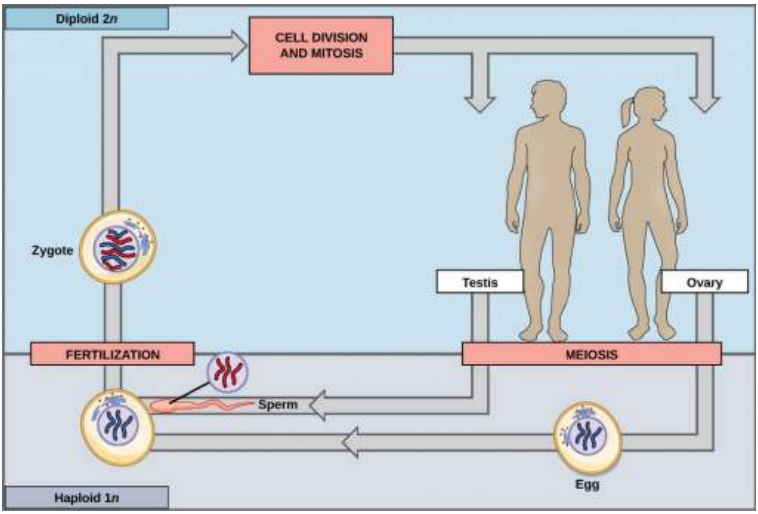


Figure 3 In animals, sexually reproducing adults form haploid gametes from diploid germ cells. Fusion of the gametes gives rise to a fertilized egg cell, or zygote. The zygote will undergo multiple rounds of mitosis to produce a multicellular offspring. The germ cells are generated early in the development of the zygote.

The nuclear division that forms haploid cells, which is called meiosis, is related to mitosis. As you have learned, mitosis is part of a cell reproduction cycle that results in identical daughter nuclei that are also genetically identical to the original parent nucleus. In mitosis, both the parent and the daughter nuclei contain the same number of chromosome sets—diploid for most plants and animals. Meiosis employs many of the same mechanisms as mitosis. However, the starting nucleus is always diploid and the nuclei that result at the end of a meiotic cell division are haploid. To achieve the reduction in chromosome number, meiosis consists of one

round of chromosome duplication and two rounds of nuclear division.

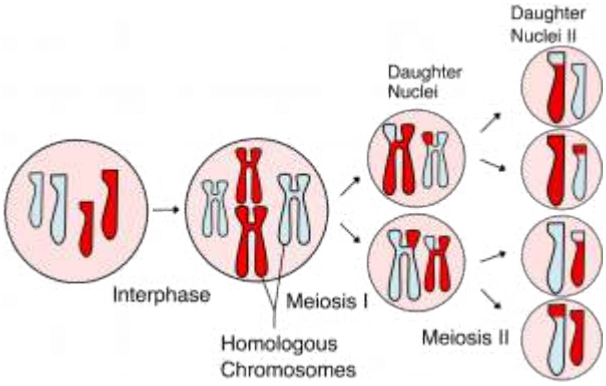


Figure 4 An overview of meiosis. Two sets of homologous chromosomes are shown. One set is comprised of a long red and a long blue chromosome. The second set is the two shorter chromosomes. During interphase, the chromosomes are duplicated so that in the second cell the look like X's. These two connected copies are called sister chromatids. Photo credit Rdbickel; Wikimedia.

Because the events that occur during each of the division stages are analogous to the events of mitosis, the same stage names are assigned. However, because there are two rounds of division, the stages are designated with a "I" or "II." Thus, **meiosis I** is the first round of meiotic division and reduces the number of chromosome sets from two to one (Figure 4). The genetic information is also mixed during this division to create unique recombinant chromosomes. **Meiosis II**, in which the second round of meiotic division takes place in a way that is similar to mitosis, separates the sister chromatids (the identical copies of each chromosome produced during DNA replication that are attached at the centromere).

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

Meiosis I

INTERPHASE

Meiosis is preceded by an interphase which is nearly identical to the interphase preceding mitosis. During interphase, the DNA of the chromosomes is replicated (during S phase). After DNA replication, each chromosome becomes composed of two identical copies (called sister chromatids) that are held together at the centromere until they are pulled apart during meiosis II (**Figure 1**).

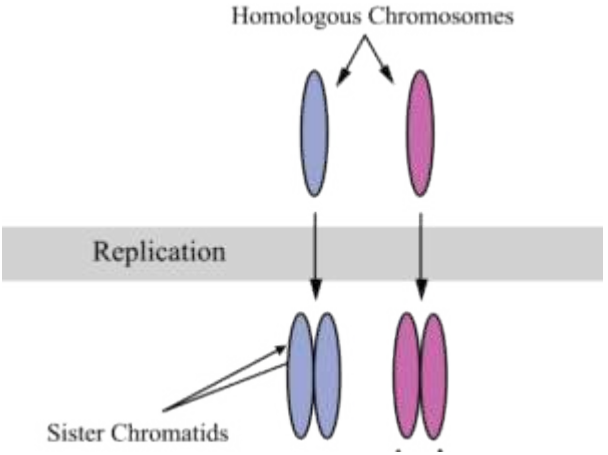


Figure 1: Sister chromatids are identical copies of a chromosome that are held together at the centromere. They are produced during DNA replication. (Credit: User:SyntaxError55, from Wikimedia)

Meiosis I

Meiosis is preceded by an interphase consisting of the G_1,

S, and G_2 phases, which are nearly identical to the phases preceding mitosis. The G_1 phase, which is also called the first gap phase, is the first phase of the interphase and is focused on cell growth. The S phase is the second phase of interphase, during which the DNA of the chromosomes is replicated. Finally, the G_2 phase, also called the second gap phase, is the third and final phase of interphase; in this phase, the cell undergoes the final preparations for meiosis.

During DNA duplication in the S phase, each chromosome is replicated to produce two identical copies, called sister chromatids, that are held together at the centromere by cohesin proteins. Cohesin holds the chromatids together until anaphase II. The centrosomes, which are the structures that organize the microtubules of the meiotic spindle, also replicate. This prepares the cell to enter prophase I, the first meiotic phase.

PROPHASE I

Early in prophase I, the chromosomes can be seen clearly microscopically. As the nuclear envelope begins to break down, the proteins associated with homologous chromosomes bring the pair close to each other. The tight pairing of the homologous chromosomes is called **synapsis** (Figure 2). In synapsis, the genes on the chromatids of the homologous chromosomes are precisely aligned with each other. Recall that synapsis does NOT occur during mitosis.

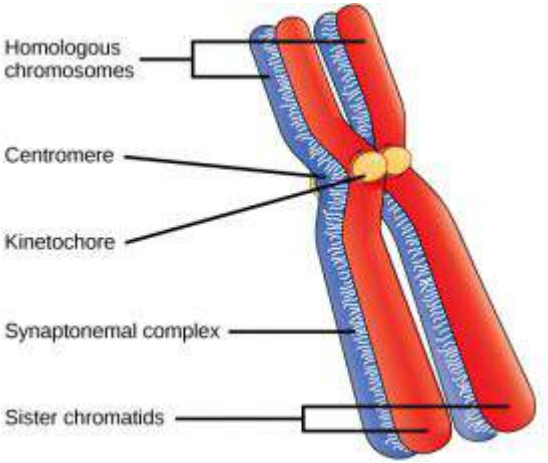


Figure 2 Early in prophase I, homologous chromosomes come together to form a synapse. The chromosomes are bound tightly together and in perfect alignment by a protein lattice called a synaptonemal complex and by cohesin proteins at the centromere.

In synapsis, the genes on the chromatids of the homologous chromosomes are aligned precisely with each other. An exchange of chromosome segments between non-sister homologous chromatids occurs and is called **crossing over (Figure 3)**. The crossover events are the first source of genetic variation produced by meiosis. A single crossover event between homologous non-sister chromatids leads to a reciprocal exchange of equivalent DNA between a maternal chromosome and a paternal chromosome. Now, when that sister chromatid is moved into a gamete, it will carry some DNA from one parent of the individual and some DNA from the other parent. The **recombinant** sister chromatid has a combination of maternal and paternal genes that did not exist before the crossover.

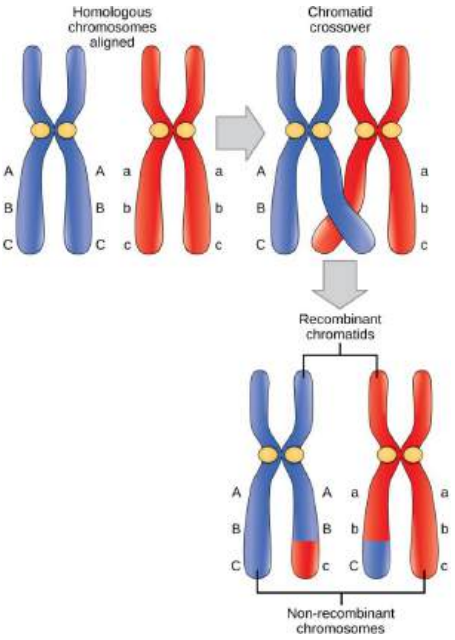


Figure 3: In this illustration of the effects of crossing over, the blue chromosome came from the individual's father and the red chromosome came from the individual's mother. Crossover occurs between non-sister chromatids of homologous chromosomes. The result is an exchange of genetic material between homologous chromosomes. The chromosomes that have a mixture of maternal and paternal sequence are called recombinant and the chromosomes that are completely paternal or maternal are called non-recombinant.

PROMETAPHASE I

The key event in prometaphase I is the attachment of the spindle fiber microtubules to the kinetochore proteins at the centromeres. Kinetochore proteins are multiprotein complexes that bind the centromeres of a chromosome to the microtubules of the mitotic spindle. Microtubules grow from centrosomes placed at opposite poles of the cell. The microtubules move toward the middle of the cell and attach to one of the two fused homologous chromosomes. The microtubules attach at each chromosomes' kinetochores. With each member of the homologous pair attached to opposite poles of the cell, in the next phase, the microtubules can pull the homologous pair apart. A spindle fiber that has attached to a kinetochore is called a kinetochore microtubule. At the end of prometaphase I, each tetrad is attached to microtubules from both poles, with one homologous chromosome facing each pole. The homologous chromosomes are still held together at chiasmata. In addition, the nuclear membrane has broken down entirely.

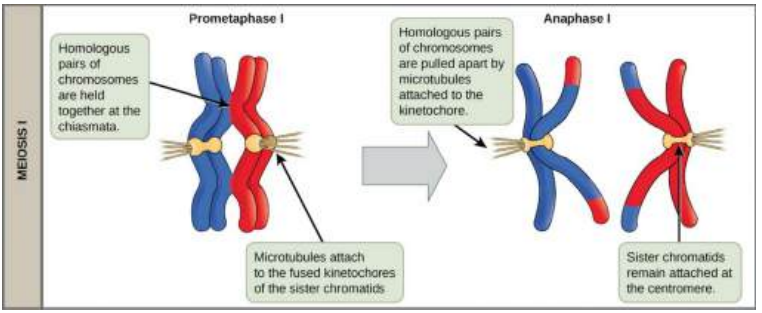


Figure 4 In prometaphase I, microtubules attach to the fused kinetochores of homologous chromosomes, and the homologous chromosomes are arranged at the midpoint of the cell in metaphase I. In anaphase I, the homologous chromosomes are separated.

METAPHASE I

During metaphase I, the homologous chromosomes are arranged in the center of the cell with the kinetochores facing opposite poles. The orientation of each pair of homologous chromosomes at the center of the cell is random. This randomness, called **independent assortment**, is the physical basis for the generation of the second form of genetic variation in offspring (Figure 5). Consider that the homologous chromosomes of a sexually reproducing organism are originally inherited as two separate sets, one from each parent in the egg and the sperm. Using humans as an example, one set of 23 chromosomes is present in the egg donated by the mother. The father provides the other set of 23 chromosomes in the sperm that fertilizes the egg. In metaphase I, these pairs line up at the midway point between the two poles of the cell. Because there is an equal chance that a microtubule fiber will encounter a maternally or paternally inherited chromosome, the arrangement of the tetrads at the metaphase plate is random. Any maternally inherited chromosome may face either pole. Any paternally inherited chromosome may also face either pole. The orientation of each tetrad is independent of the orientation of the other 22 tetrads.

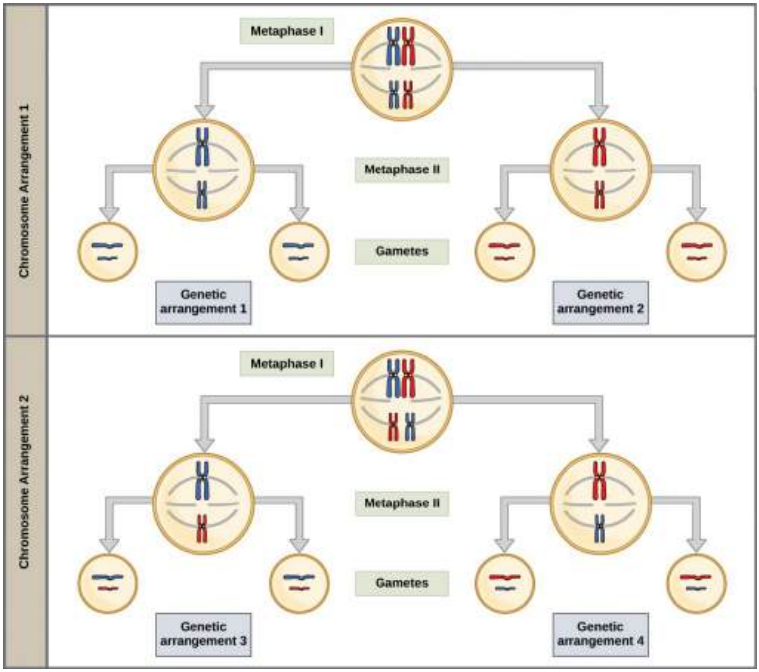


Figure 5 Random, independent assortment during metaphase I can be demonstrated by considering a cell with a set of two chromosomes (n = 2). In this case, there are two possible arrangements at the equatorial plane in metaphase I. The total possible number of different gametes is 2^n, where n equals the number of chromosomes in a set. In this example, there are four possible genetic combinations for the gametes. With n = 23 in human cells, there are over 8 million possible combinations of paternal and maternal chromosomes.

In each cell that undergoes meiosis, the arrangement of the tetrads is different. The number of variations depends on the number of chromosomes making up a set. There are two possibilities for orientation (for each tetrad); thus, the possible number of alignments equals 2^n where n is the number of chromosomes per set. Humans have 23 chromosome pairs, which results in over eight million (2^{23}) possibilities. This number does not include the variability

previously created in the sister chromatids by crossover. Given these two mechanisms, it is highly unlikely that any two haploid cells resulting from meiosis will have the same genetic composition (**Figure 5**).

To summarize the genetic consequences of meiosis I: the maternal and paternal genes are recombined by crossover events occurring on each homologous pair during prophase I; in addition, the random assortment of tetrads at metaphase produces a unique combination of maternal and paternal chromosomes that will make their way into the gametes.

Anaphase I

In anaphase I, the microtubules pull the linked chromosomes apart. The sister chromatids remain tightly bound together at the centromere. The chiasmata are broken in anaphase I as the microtubules attached to the fused kinetochores pull the homologous chromosomes apart (Figure 4).

Telophase I and Cytokinesis I

In telophase, the separated chromosomes arrive at opposite poles. The remainder of the typical telophase events may or may not occur, depending on the species. In some organisms, the chromosomes decondense and nuclear envelopes form around the chromatids in telophase I. In other organisms, cytokinesis—the physical separation of the cytoplasmic components into two daughter cells—occurs without reformation of the nuclei. In nearly all species of animals and some fungi, cytokinesis separates the cell contents via a cleavage furrow (constriction of the actin ring that leads to cytoplasmic division). In plants, a cell plate is formed during cell cytokinesis by Golgi vesicles fusing at the metaphase plate. This cell plate will ultimately lead to the formation of cell walls that separate the two daughter cells.

Two haploid cells are the end result of the first meiotic

division. The cells are haploid because at each pole, there is just one of each pair of the homologous chromosomes. Therefore, only one full set of the chromosomes is present. This is why the cells are considered haploid—there is only one chromosome set, even though each homolog still consists of two sister chromatids. Recall that sister chromatids are merely duplicates of one of the two homologous chromosomes (except for changes that occurred during crossing over). In meiosis II, these two sister chromatids will separate, creating four haploid daughter cells.

SUMMARY OF MEIOSIS I

The chromosomes are copied during interphase (prior to meiosis I). This forms two identical sister chromatids that are attached together at the centromere. During prophase I, crossing over introduces genetic variation by swapping pieces of homologous chromosomes. Additional genetic variation is introduced by independent assortment, which takes into account how the homologous chromosomes line up during metaphase I. At the end of meiosis I, two haploid cells (where each chromosome still consists of two sister chromatids) are produced.

REFERENCES

OpenStax, Biology. OpenStax CNX. January 2, 2017 http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

Meiosis II

In some species, cells enter a brief interphase, or interkinesis, before entering meiosis II. Interkinesis lacks an S phase, so chromosomes are **not** duplicated. The two cells produced in meiosis I go through the events of meiosis II at the same time. During meiosis II, the sister chromatids within the two daughter cells separate, forming four new haploid gametes. The mechanics of meiosis II is similar to mitosis, except that each dividing cell has only one set of homologous chromosomes. Therefore, each cell has half the number of sister chromatids to separate out as a diploid cell undergoing mitosis.

PROPHASE II

If the chromosomes decondensed in telophase I, they condense again. If nuclear envelopes were formed, they fragment into vesicles. The centrosomes that were duplicated during interkinesis move away from each other toward opposite poles, and new spindles are formed.

PROMETAPHASE II

The nuclear envelopes are completely broken down, and the spindle is fully formed. Each sister chromatid forms an individual kinetochore that attaches to microtubules from opposite poles.

METAPHASE II

The sister chromatids are maximally condensed and aligned at the equator of the cell.

ANAPHASE II

The sister chromatids are pulled apart by the kinetochore microtubules and move toward opposite poles (Figure 1). Non-kinetochore microtubules elongate the cell.

In meiosis II, the connected sister chromatids remaining in the haploid cells from meiosis I will be split to form four haploid cells. The two cells produced in meiosis I go through the events of meiosis II in synchrony. Overall, meiosis II resembles the mitotic division of a haploid cell. During meiosis II, the sister chromatids are pulled apart by the spindle fibers and move toward opposite poles.

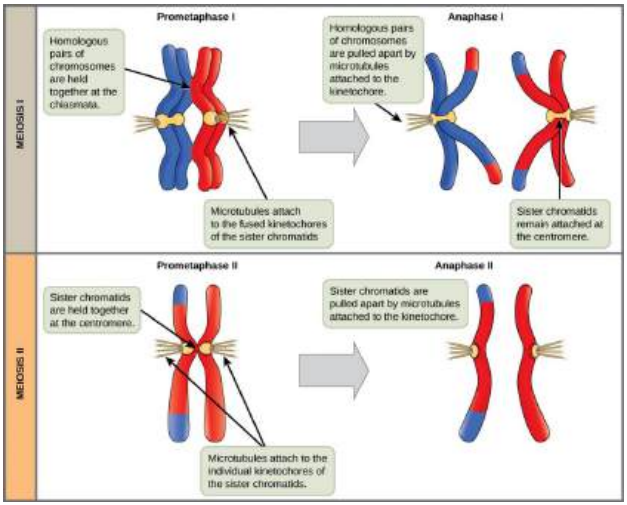


Figure 1 In prometaphase I, microtubules attach to the fused kinetochores of homologous chromosomes. In anaphase I, the homologous chromosomes are separated. In prometaphase II, microtubules attach to individual kinetochores of sister chromatids. In anaphase II, the sister chromatids are separated.

TELOPHASE II AND CYTOKINESIS

The chromosomes arrive at opposite poles and begin to decondense. Nuclear envelopes form around the chromosomes. Cytokinesis separates the two cells into four unique haploid cells. At this point, the newly formed nuclei are both haploid and have only one copy of the single set of chromosomes. The cells produced are genetically unique because of the random assortment of paternal and maternal homologs and because of the recombining of maternal and paternal segments of chromosomes (with their sets of genes) that occurs during crossover.

The entire process of meiosis is outlined in Figure 2.

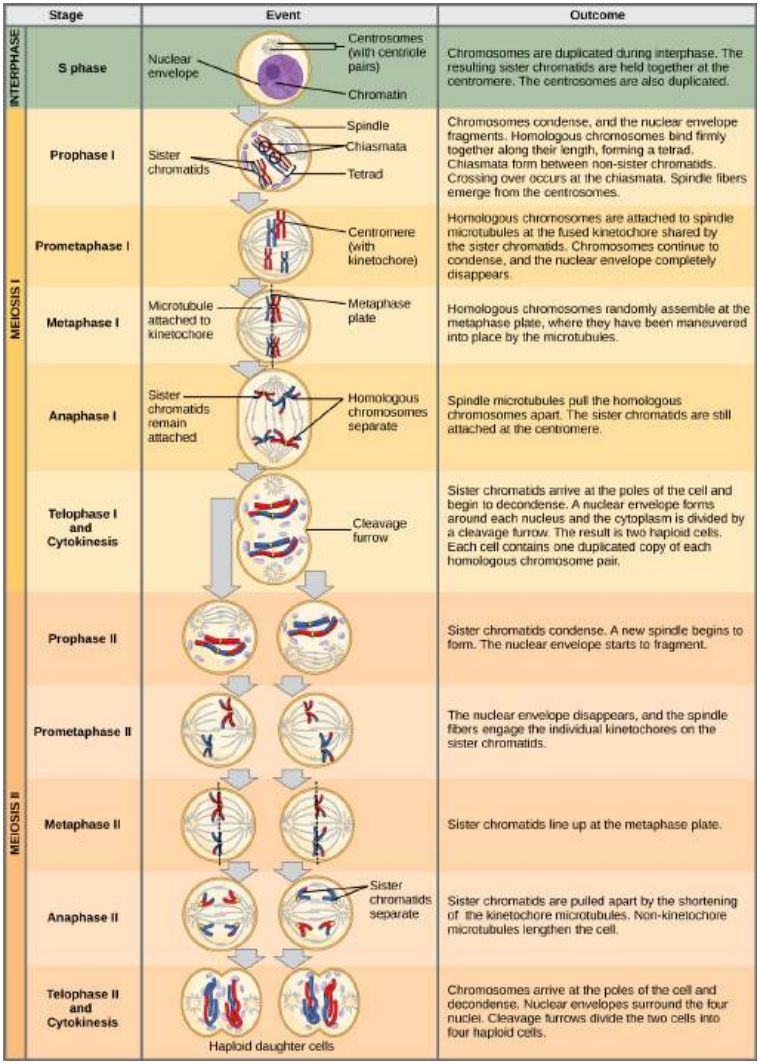


Figure 2 An animal cell with a diploid number of four (2n = 4) proceeds through the stages of meiosis to form four haploid daughter cells.

SUMMARY OF MEIOSIS II

Meiosis II begins with the 2 haploid cells where each chromosome is made up of two connected sister chromatids. DNA replication does NOT occur at the beginning of meiosis II. The sister chromatids are separated, producing 4 genetically different haploid cells.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

Comparing Meiosis and Mitosis

Mitosis and meiosis, which are both forms of division of the nucleus in eukaryotic cells, share some similarities, but also exhibit distinct differences that lead to their very different outcomes. Mitosis is a single nuclear division that results in two nuclei, usually partitioned into two new cells. The nuclei resulting from a mitotic division are genetically identical to the original. They have the same number of sets of chromosomes: one in the case of haploid cells, and two in the case of diploid cells. On the other hand, meiosis is two nuclear divisions that result in four nuclei, usually partitioned into four new cells. The nuclei resulting from meiosis are never genetically identical, and they contain one chromosome set only—this is half the number of the original cell, which was diploid.

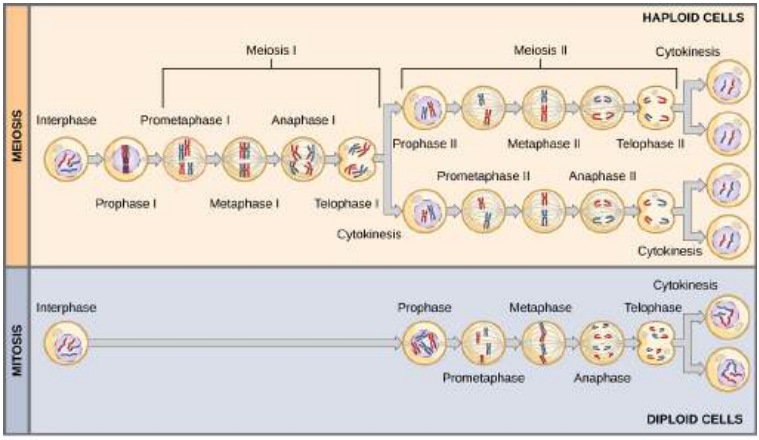
The differences in the outcomes of meiosis and mitosis occur because of differences in the behavior of the chromosomes during each process. Most of these differences in the processes occur in meiosis I, which is a very different nuclear division than mitosis. In meiosis I, the homologous chromosome pairs become associated with each other, are bound together, experience chiasmata and crossover between sister chromatids, and line up along the metaphase plate in tetrads with spindle fibers from opposite spindle poles attached to each kinetochore of a homolog in a tetrad. All of these events occur only in meiosis I, never in mitosis.

Homologous chromosomes move to opposite poles during meiosis I so the number of sets of chromosomes in each nucleus-to-be is reduced from two to one. For this reason, meiosis I is referred to as a **reduction division**. There is no such reduction in ploidy level in mitosis.

Meiosis II is much more analogous to a mitotic division. In this case, duplicated chromosomes (only one set of them) line up at the center of the cell with divided kinetochores attached to spindle fibers from opposite poles. During

anaphase II, as in mitotic anaphase, the kinetochores divide and one sister chromatid is pulled to one pole and the other sister chromatid is pulled to the other pole. If it were not for the fact that there had been crossovers, the two products of each meiosis II division would be identical as in mitosis; instead, they are different because there has always been at least one crossover per chromosome. Meiosis II is not a reduction division because, although there are fewer copies of the genome in the resulting cells, there is still one set of chromosomes, as there was at the end of meiosis I.

Cells produced by mitosis will function in different parts of the body as a part of growth or replacing dead or damaged cells. They may even be involved in asexual reproduction in some organisms. Cells produced by meiosis in a diploid-dominant organism such as an animal will only participate in sexual reproduction.



						OUTCOME
PROCESS	DNA synthesis	Synapsis of homologous chromosomes	Crossover	Homologous chromosomes line up at metaphase plate	Sister chromatids line up at metaphase plate	Number and genetic composition of daughter cells
MEIOSIS	Occurs in S phase of interphase	During prophase I	During prophase I	During metaphase I	During metaphase II	Four haploid cells at the end of meiosis II
MITOSIS	Occurs in S phase of interphase	Does not occur in mitosis	Does not occur in mitosis	Does not occur in mitosis	During metaphase	Two diploid cells at the end of mitosis

Figure 1 Meiosis and mitosis are both preceded by one round of DNA replication; however, meiosis includes two nuclear divisions. The four daughter cells resulting from meiosis are haploid and genetically distinct. The daughter cells resulting from mitosis are diploid and identical to the parent cell.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

Errors in Meiosis

Inherited disorders can arise when chromosomes behave abnormally during meiosis. Chromosome disorders can be divided into two categories: abnormalities in chromosome number and chromosome structural rearrangements. Because even small segments of chromosomes can span many genes, chromosomal disorders are characteristically dramatic and often fatal.

DISORDERS IN CHROMOSOME NUMBER

The isolation and microscopic observation of chromosomes forms the basis of cytogenetics and is the primary method by which clinicians detect chromosomal abnormalities in humans. A **karyotype** is the number and appearance of chromosomes, including their length, banding pattern, and centromere position. To obtain a view of an individual's karyotype, cytologists photograph the chromosomes and then cut and paste each chromosome into a chart, or **karyogram** (**Figure 1**).

Figure 1 This karyogram shows the chromosomes of a normal female human immune cell during mitosis. (Credit: Andreas Bolzer, et al)

By observing a karyogram, geneticists can actually visualize the chromosomal composition of an individual to confirm or predict genetic abnormalities in offspring even before birth.

Geneticists Use Karyograms to Identify Chromosomal Aberrations

Although Mendel is referred to as the "father of modern genetics," he performed his experiments with none of the tools that the geneticists of today routinely employ. One such powerful cytological technique is karyotyping, a method in which traits characterized by chromosomal abnormalities can be identified from a single cell. To observe an individual's karyotype, a person's cells (like white blood cells) are first collected from a blood sample or other tissue. In the laboratory, the isolated cells are stimulated to begin actively dividing. A chemical called colchicine is then applied to cells to arrest condensed chromosomes in metaphase. Cells are then made to swell using a hypotonic solution so the chromosomes spread apart. Finally, the sample is preserved in a fixative and applied to a slide.

The geneticist then stains chromosomes with one of several dyes to better visualize the distinct and reproducible banding patterns of each chromosome pair. Following staining, the chromosomes are viewed using bright-field microscopy. A common stain choice is the Giemsa stain. Giemsa staining results in approximately 400–800 bands (of tightly coiled DNA and condensed proteins) arranged along all of the 23 chromosome pairs; an experienced geneticist can identify each band. In addition to the banding patterns, chromosomes are further identified on the basis of size and centromere location. To obtain the classic depiction of the karyotype in which homologous pairs of chromosomes are aligned in numerical order from longest to shortest, the geneticist obtains a digital

image, identifies each chromosome, and manually arranges the chromosomes into this pattern (Figure 1).

At its most basic, the karyogram may reveal genetic abnormalities in which an individual has too many or too few chromosomes per cell. Examples of this are Down Syndrome, which is identified by a third copy of chromosome 21, and Turner Syndrome, which is characterized by the presence of only one X chromosome in women instead of the normal two. Geneticists can also identify large deletions or insertions of DNA. For instance, Jacobsen Syndrome—which involves distinctive facial features as well as heart and bleeding defects—is identified by a deletion on chromosome 11. Finally, the karyotype can pinpoint translocations, which occur when a segment of genetic material breaks from one chromosome and reattaches to another chromosome or to a different part of the same chromosome. Translocations are implicated in certain cancers, including chronic myelogenous leukemia.

During Mendel's lifetime, inheritance was an abstract concept that could only be inferred by performing crosses and observing the traits expressed by offspring. By observing a karyogram, today's geneticists can actually visualize the chromosomal composition of an individual to confirm or predict genetic abnormalities in offspring, even before birth.

Of all the chromosomal disorders, abnormalities in chromosome number are the most easily identifiable from a karyogram. Disorders of chromosome number include the duplication or loss of entire chromosomes, as well as changes in the number of complete sets of chromosomes. They are caused by **nondisjunction**, which occurs when pairs of homologous chromosomes or sister chromatids fail to separate during meiosis. The risk of nondisjunction increases with the age of the parents.

Nondisjunction can occur during either meiosis I or II, with

different results (**Figure 2**). If homologous chromosomes fail to separate during meiosis I, the result is two gametes that lack that chromosome and two gametes with two copies of the chromosome. If sister chromatids fail to separate during meiosis II, the result is one gamete that lacks that chromosome, two normal gametes with one copy of the chromosome, and one gamete with two copies of the chromosome.

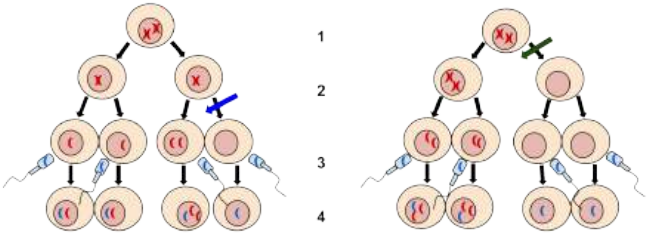


Figure 2 Nondisjunction occurs when homologous chromosomes or sister chromatids fail to separate during meiosis, resulting in an abnormal chromosome number. Nondisjunction may occur during meiosis I or meiosis II. Photo credit Tweety207; Wikimedia.

An individual with the appropriate number of chromosomes for their species is called **euploid**; in humans, euploidy corresponds to 22 pairs of **autosomes** and one pair of sex chromosomes (such as is seen in the karyotype in Figure 1). An individual with an error in chromosome number is described as **aneuploid**, a term that includes **monosomy** (loss of one chromosome) or **trisomy** (gain of an extraneous chromosome). Monosomic human zygotes missing any one copy of an autosome invariably fail to develop to birth because they have only one copy of essential genes. Most autosomal trisomies also fail to develop to birth; however, duplications of some of the smaller chromosomes (13, 15, 18, 21, or 22) can result in offspring that survive for several weeks to many years. Trisomic individuals suffer from a

different type of genetic imbalance: an excess in gene dose. Cell functions are calibrated to the amount of gene product produced by two copies (doses) of each gene; adding a third copy (dose) disrupts this balance. The most common trisomy is that of chromosome 21, which leads to Down syndrome. Individuals with this inherited disorder have characteristic physical features and developmental delays in growth and cognition.

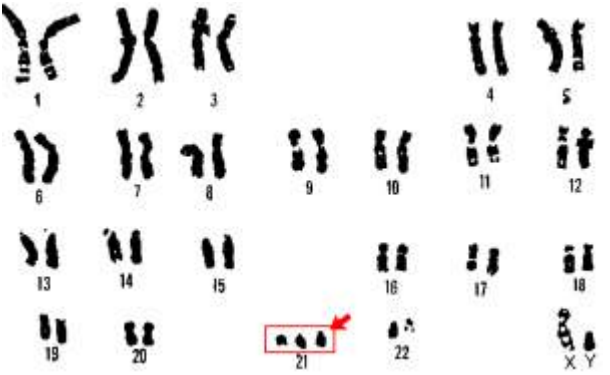


Figure 3 Karyotype of an individual with Down Syndrome. Photo credit U.S. Department of Energy Human Genome Program. Wikimedia.

The incidence of Down syndrome is correlated with maternal age, such that older women are more likely to give birth to children with Down syndrome (Figure 4).

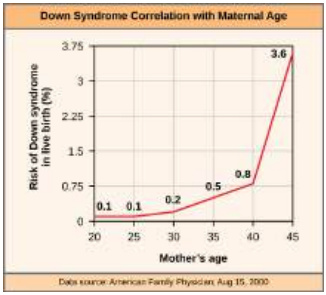


Figure 4: The incidence of having a fetus with trisomy 21 increases dramatically with maternal age.

An individual with more than the correct number of chromosome sets (two for diploid species) is called **polyploid**. For instance, fertilization of an abnormal diploid egg with a normal haploid sperm would yield a triploid zygote. Polyploid animals are extremely rare, with only a few examples among the flatworms, crustaceans, amphibians, fish, and lizards. Triploid animals are sterile because meiosis cannot proceed normally with an odd number of chromosome sets. In contrast, polyploidy is very common in the plant kingdom, and polyploid plants tend to be larger and more robust than euploids of their species (Figure 5).

Figure 5 As with many polyploid plants, this triploid orange daylily (Hemerocallis fulva) is particularly large and robust, and grows flowers with triple the number of petals of its diploid counterparts. (credit: Steve Karg)

SEX CHROMOSOME NONDISJUNCTION

Humans display dramatic deleterious effects with autosomal trisomies and monosomies. Therefore, it may seem counterintuitive that human females and males can function normally, despite carrying different numbers of the X chromosome. In part, this occurs because of a process called **X inactivation**. Early in development, when female mammalian embryos consist of just a few thousand cells, one X chromosome in each cell inactivates by condensing into a structure called a Barr body. The genes on the inactive X chromosome are not expressed. The particular X chromosome (maternally or paternally derived) that is inactivated in each cell is random, but once the inactivation occurs, all cells descended from that cell will have the same inactive X chromosome. By this process, females compensate for their double genetic dose of X chromosome.

In so-called "tortoiseshell" cats, X inactivation is observed as coat-color variegation (**Figure 6**). Females heterozygous for an X-linked coat color gene will express one of two different coat colors over different regions of their body, corresponding to whichever X chromosome is inactivated in the embryonic cell progenitor of that region. When you see a tortoiseshell cat, you will know that it has to genetically be a female.

Figure 6 Embryonic inactivation of one of two different X chromosomes encoding different coat colors gives rise to the tortoiseshell phenotype in cats. (credit: Michael Bodega)

In an individual carrying an abnormal number of X chromosomes, cellular mechanisms will inactivate all but one X in each of her cells. As a result, X-chromosomal abnormalities are typically associated with mild mental and physical defects, as well as sterility. If the X chromosome is absent altogether, the individual will not develop.

Several errors in sex chromosome number have been characterized. Individuals with three X chromosomes, called triplo-X, appear female but express developmental delays and reduced fertility. The XXY chromosome complement,

corresponding to one type of Klinefelter syndrome, corresponds to male individuals with small testes, enlarged breasts, and reduced body hair. The extra X chromosome undergoes inactivation to compensate for the excess genetic dosage. Turner syndrome, characterized as an X0 chromosome complement (i.e., only a single sex chromosome), corresponds to a female individual with short stature, webbed skin in the neck region, hearing and cardiac impairments, and sterility.

CHROMOSOME STRUCTURAL REARRANGEMENTS

Cytologists have characterized numerous structural rearrangements in chromosomes, including partial duplications, deletions, inversions, and translocations. Duplications and deletions often produce offspring that survive but exhibit physical and mental abnormalities. Cri-du-chat (from the French for "cry of the cat") is a syndrome associated with nervous system abnormalities and identifiable physical features that results from a deletion of most of the small arm of chromosome 5 (**Figure 7**). Infants with this genotype emit a characteristic high-pitched cry upon which the disorder's name is based.

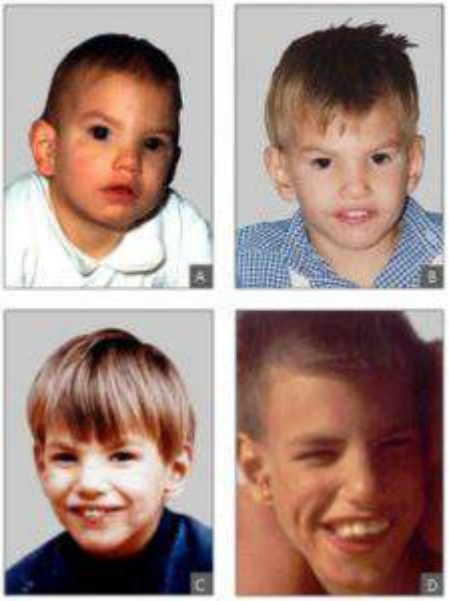


Figure 7 This individual with cri-du-chat syndrome is shown at various ages: (A) age two, (B) age four, (C) age nine, and (D) age 12. (credit: Paola Cerruti Mainardi)

Chromosome inversions and translocations can be identified by observing cells during meiosis because homologous chromosomes with a rearrangement in one of the pair must contort to maintain appropriate gene alignment and pair effectively during prophase I.

A **chromosome inversion** is the detachment, 180° rotation, and reinsertion of part of a chromosome (**Figure 8**). Unless they disrupt a gene sequence, inversions only change the orientation of genes and are likely to have more mild effects than aneuploid errors.

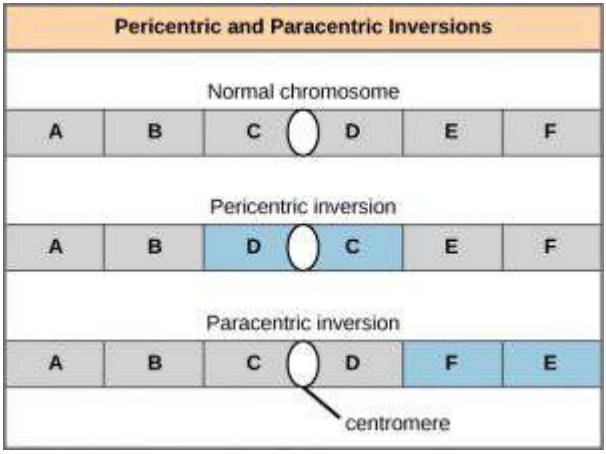


Figure 8 An inversion occurs when a chromosome segment breaks from the chromosome, reverses its orientation, and then reattaches in the original position.

A translocation occurs when a segment of a chromosome dissociates and reattaches to a different, nonhomologous chromosome. Translocations can be benign or have devastating effects, depending on how the positions of genes are altered with respect to regulatory sequences. Notably, specific translocations have been associated with several cancers and with schizophrenia. Reciprocal translocations result from the exchange of chromosome segments between two nonhomologous chromosomes such that there is no gain or loss of genetic information (**Figure 9**).

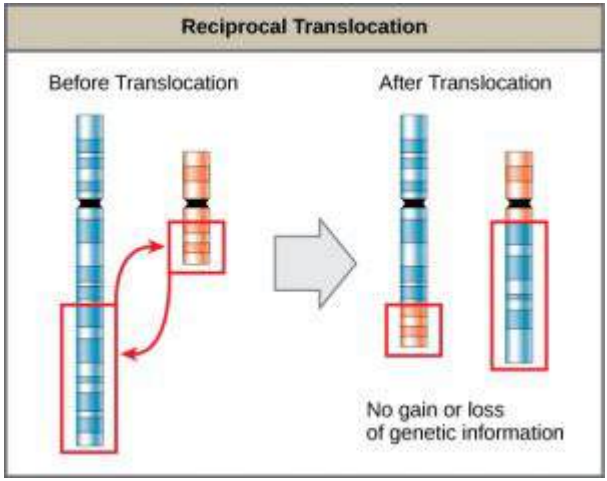


Figure 9 A reciprocal translocation occurs when a segment of DNA is transferred from one chromosome to another, nonhomologous chromosome. (credit: modification of work by National Human Genome Research/USA)

One specific example of a chromosomal translocation – the "Philadelphia chromosome" – is found in people who suffer from chronic myeloid leukemia (CML). In this translocation, a piece of chromosome 9 is swapped with a section of chromosome 22. This connects two genes on chromosome 22; one that was originally from chromosome 9 and one that was from chromosome 22. This translocation produces the BCR-ABL fusion protein, which causes white blood cells to divide out of control. BCR-ABL positive cancers can be treated with the drug Gleevac.

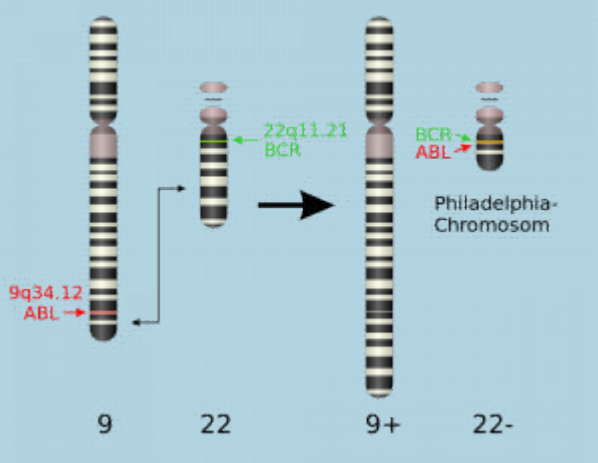


Figure 9 "Philadelphia chromosome" showing the location of the BCR-ABL fusion protein. Photo credit A Obeidat; Wikimedia.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:6-3MVU-j@4/Errors-in-Meiosis

PATTERNS OF INHERITANCE

Learning Objectives

By the end of this section, you will be able to:

- Describe the molecular basis of inheritance.
- Determine the outcome in crosses involving various types of inheritance.
- Present and decipher information about trait inheritance using a pedigree.

GENETICS: DOG COAT COLOR

Learning Objectives

By the end of this section, you will be able to:

- Describe the molecular basis of inheritance.
- Determine the outcome in crosses involving complete dominance.
- Present and decipher information about inheritance using a pedigree.

Figure 1: Experimenting with thousands of garden peas, Johann Gregor Mendel uncovered the fundamentals of genetics. (credit: modification of work by Jerry Kirkhart)

Remember that a trait is an aspect of the physical appearance of an organism that can vary. Organisms get their traits from proteins; proteins are produced using the information found in the organism's DNA. Variation in the DNA between different organisms causes the production of proteins that contain differing orders of amino acids. These proteins can have different shapes and therefore different functions. When proteins function differently, this leads to differences in traits.

Recall that diploid organisms have two copies of each chromosome: a pair of homologous chromosomes. The reason that they have two copies is because they inherited one copy of each chromosome from each parent. Each parent donates one haploid gamete (egg or sperm) to the reproductive process. A haploid gamete contains one copy of each chromosome because during meiosis the number of chromosomes is cut in half: the DNA is copied once and then divided twice. This separation of the homologous chromosomes means that only one of the copies of the gene gets moved into a gamete. The offspring are formed when

that gamete unites with one from another parent and the two copies of each gene (and chromosome) are restored.

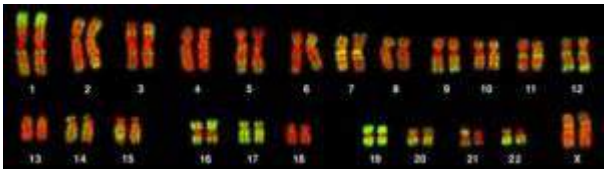


Figure 2: A karyogram is a picture of all the chromosomes in a cell, organized into homologous pairs. This is a human karyogram which shows the 46 chromosomes present in diploid human somatic cells.

A diploid organism has two copies of a given gene. The two copies may or may not encode the same version of that characteristic. For example, one individual pea plant (such as those studied by Mendel) would have two copies of the gene that controls flower color. That individual could carry one version of the gene that leads to white flower color and a second different version of that same gene that leads to violet flower color. The interaction between these two different versions of the same gene will lead to the visible flower color in the pea plant. Gene variations that arise by mutation and exist at the same relative locations on homologous chromosomes are called **alleles**. Mendel examined the inheritance of genes with just two allele forms, but it is common to encounter more than two alleles for many genes in a natural population.

Each individual (assuming it is a diploid organism) will have two alleles for a specific gene: one from each of its two parents. These two alleles are expressed and interact to produce physical characteristics. The observable traits expressed by an organism are referred to as its **phenotype**. An organism's underlying genetic makeup, consisting of both the physically visible and the non-expressed alleles, is called its **genotype**.

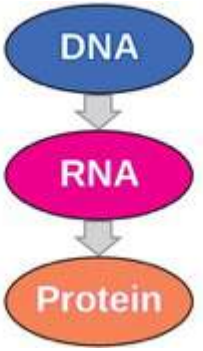
Diploid organisms that are **homozygous** for a gene have two identical alleles, one on each of their homologous chromosomes. If the organism has two different alleles, this is referred to as **heterozygous.**

This chapter will address a simple type of inheritance: complete dominance. In this type of inheritance, there are two alleles: dominant and recessive. A **dominant** allele will completely cover up a recessive allele. This means that if one dominant allele is present, the organism will have the trait conferred by that allele. In order for the recessive phenotype to be seen, the organism must have two **recessive** alleles. Just because an allele is dominant does not automatically make it better than a recessive trait. It also does not make it more common than the recessive trait. All it means for an allele to be dominant is that it is able to cover up the recessive allele.

We typically abbreviate the genotype of an organism by using single letters. The letter chosen is often the first letter of the dominant trait. A homozygous dominant genotype would be written AA, a heterozygous genotype as Aa, and a homozygous recessive genotype as aa.

Introduction to Genetics

Recall that genes are segments of DNA that are typically several hundred or thousand bases long. Each gene directs the production of a protein through the process of protein synthesis: DNA gets transcribed to produce an mRNA; mRNA provides to code for a ribosome to produce a chain of amino acids. Read this section of the book if you need to review this topic: How do genes direct the production of proteins?

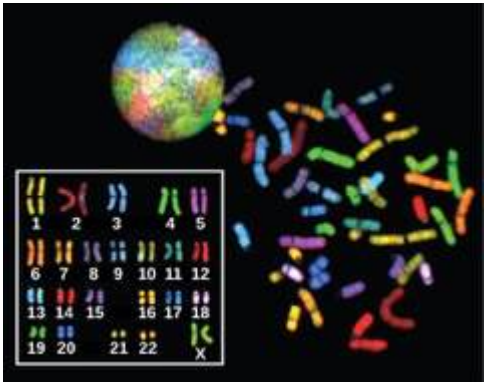


The Central Dogma – DNA is used to make RNA is used to make protein. Photo credit: ?

Recall that eukaryotic genes are found on chromosomes and that each eukaryotic chromosome typically contains hundreds or thousands of genes. In most eukaryotes, including humans and other animals, each cell contains two copies of each chromosome. The reason we have two copies of each gene is that we inherit one from each parent.

In contrast to eukaryotes, prokaryotes have one circular chromosome. This means they have one copy of each gene.

Read this section of the book if you need to review this topic: How DNA is arranged in the cell



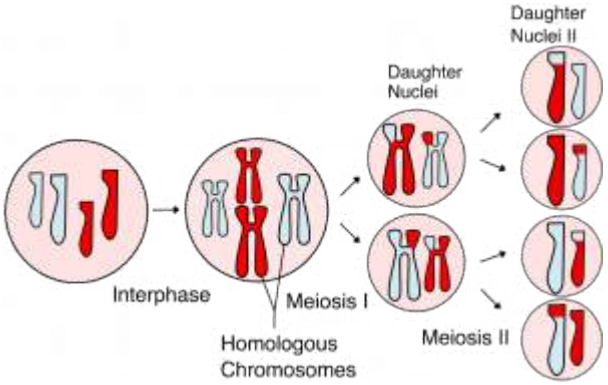
There are 23 pairs of chromosomes in a female human body cell. These chromosomes are viewed within the nucleus (top), removed from a cell during cell division (right), and arranged according to length (left) in an arrangement called a karyotype. In this image, the chromosomes were exposed to fluorescent stains to distinguish them. (credit: "718 Bot"/Wikimedia Commons, National Human Genome Research)

Chromosomes are inherited by the offspring from the parents via the egg or sperm. Inside one egg or one sperm is one copy of each chronometer (so 23 total in humans). When an egg is fertilized by a sperm, the resulting **zygote** (fertilized egg) will contain two copies of each chromosome, just like each of its parents.

Meiosis is the process that produces eggs and sperm. Eggs and sperm are also known as **gametes**. During meiosis, one copy of each paired chromosome is moved into the gamete. Cells with one copy of each chromosome are known as "**haploid**". This separation, or segregation, of the **homologous** (paired) chromosomes means also that only one of the copies of the gene gets moved into a gamete.

The offspring are formed when that gamete unites with

one from another parent and the two copies of each gene (and chromosome) are restored. Read this section of the book if you need more information on this topic: Overview of Meiosis



During meiosis, the DNA is copied once, then the cell divides twice. This produces cells with half as much genetic information as the original cell (1 copy of each chromosome). These cells become the sex cells (eggs or sperm). When two sex cells unite during fertilization, the original number of chromosomes (2 copies of each one) is restored. Photo credit: ?

The offspring will receive two copies of each gene (one from each parent), but the copies are not necessarily identical. You already knew this – you don't get identical information from your mother and your father because they have different DNA (which gives them different traits). The different versions of one specific gene are known as **alleles.** As you learn about genetics, you will learn about how the information from both alleles of a specific gene interact to give an individual their trait. The genetic information that an individual has is called their **genotype.** The genotype of an individual produces the individual's **phenotype,** or physical traits.

REFERENCES

Unless otherwise noted, text and images by Lisa Bartee, 2016.

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8v2Xzdco@5/The-Structure-of-DNA

Pedigrees and Punnett Squares

PEDIGREES

Inheritance of a trait through generations can be shown visually using a pedigree, such as is pictured in **Figure 1**. Square shapes represent males; circles represent females. Filled-in shapes are individuals that have whatever trait is being shown in the pedigree. Two individuals connected together with a horizontal line between them are the parents of the individuals that are connected by vertical lines below them. Siblings are typically shown in birth order with the oldest sibling to the left.

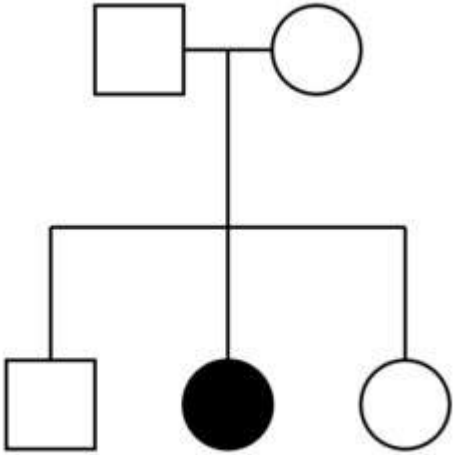


Figure 1 A simple pedigree. In this pedigree, the parents (at the top) have produced three children: a male and two females. The first female has the condition being shown in the pedigree.

PUNNETT SQUARES

As discussed above, diploid individuals have two copies of each chromosome: one from their male parent, one from their female parent. This means they have two copies of each gene. They can have two of the same alleles (homozygous) or two different alleles (heterozygous). Regardless of their genotype, they will randomly pass only one copy of each chromosome to their offspring. This is because meiosis produces haploid gametes that contain one copy of each chromosome, and those chromosomes are assorted into gametes randomly. Since genes are present on chromosomes, this means they will pass one copy of each

gene to their offspring. That means that an offspring inherits one allele of each gene from each of its two parents. This is illustrated in Figure 2. This concept is called **Mendel's Law of Segregation**.

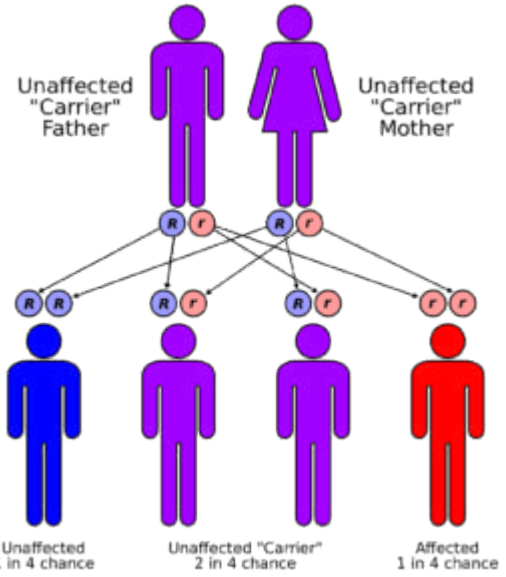


Figure 2 Two parents who are heterozygous each pass one chromosome / gene / allele to each offspring. Each resulting offspring has two of each chromosome / gene. The individual can have two of the same or two different alleles.

An easy, organized way of illustrating the offspring that can result from two specific parents is to use a Punnett square. The gametes that can be generated by each parent are represented above the rows and next to the columns of the square. Each gamete is haploid for the "A gene", meaning it only contains one copy of that gene. In the Punnett square seen in Figure 3, haploid eggs are above each column and haploid sperm are next to each row. When a haploid sperm

and a haploid egg (each with 1 copy of the "A gene") combine during the process of fertilization, a diploid offspring (with 2 copies of the A gene) is the result.

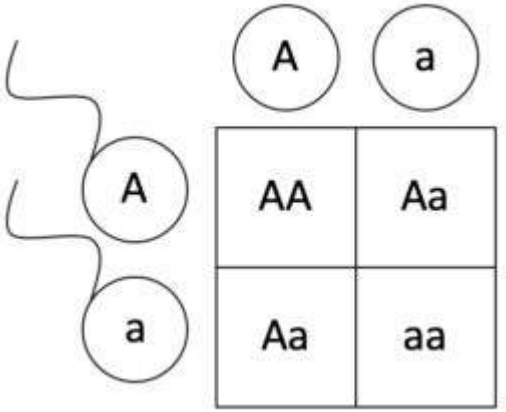


Figure 3: A Punnett square showing a cross between two individuals who are both heterozygous for A.

A Punnett square shows the probability of an offspring with a given genotype resulting from a cross. It does not show actual offspring. For example, the Punnett square in Figure 3 shows that there is a 25% chance that a homozygous recessive offspring will result from the cross Aa x Aa. It does *not* mean that these parents must have 4 offspring and that they will have the ratio 1 AA : 2 Aa : 1 aa. It's just like flipping a coin: you expect 50% heads, but you wouldn't be too surprised to see 7 heads out of 10 coin flips. Additionally, the probability does not change for successive offspring. The probability that the first offspring will have the genotype "aa" is 25% and the probability of the second offspring having the genotype "aa" is still 25%. Again, it's just like flipping a coin: if you flip heads the first time, that doesn't change the probability of getting heads on the next flip.

Organisms don't just inherit one trait at a time, though. They inherit all their traits at once. Sometimes, we want to determine the probability of an individual inheriting two different traits. The easiest way to do this is to determine the probability of the individual inheriting each trait separately, then multiply those probabilities together. An example of this can be seen in **Figure 4**. In order for this to work, we must assume that genes do not influence each other with regard to the sorting of alleles into gametes, and every possible combination of alleles for every gene is equally likely to occur. This is called **Mendel's Law of Independent Assortment**.

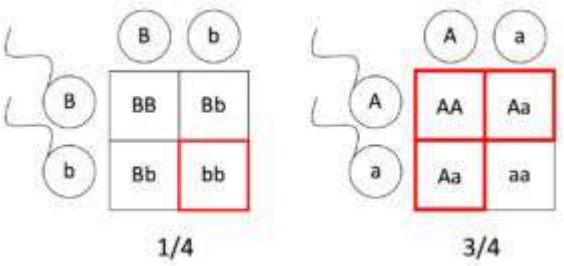


Figure 4: These two Punnett square show the cross between two individuals who are both heterozygous for two different genes: BbAa x BbAa. We can determine the probability of an offspring having the recessive trait for "B" and the dominant trait for "A". The probability of the offspring having the recessive phenotype for "B" is 1/4. The probability of the offspring having the dominant phenotype for "A" is 3/4. 1/4 x 3/4 = 3/16.

Another way of determining the probability of getting two different traits is to use a dihybrid Punnett square. **Figure 5** shows three generations of the inheritance of pea seed color and shape. Peas can be either yellow or green, and they can be either round or wrinkled. These are two of the traits that Mendel studied in his work with peas. In the first

generation (the "P" generation), two true-breeding (homozygous) individuals are crossed. Their offspring will get one allele of the Y gene and one allele of the R gene from each parent. This means that all their offspring (the "F1" generation) will be heterozygous for both genes. The results (the "F2" generation) from crossing two heterozygous individuals can be seen in the 4×4 Punnett square in **Figure 5**.

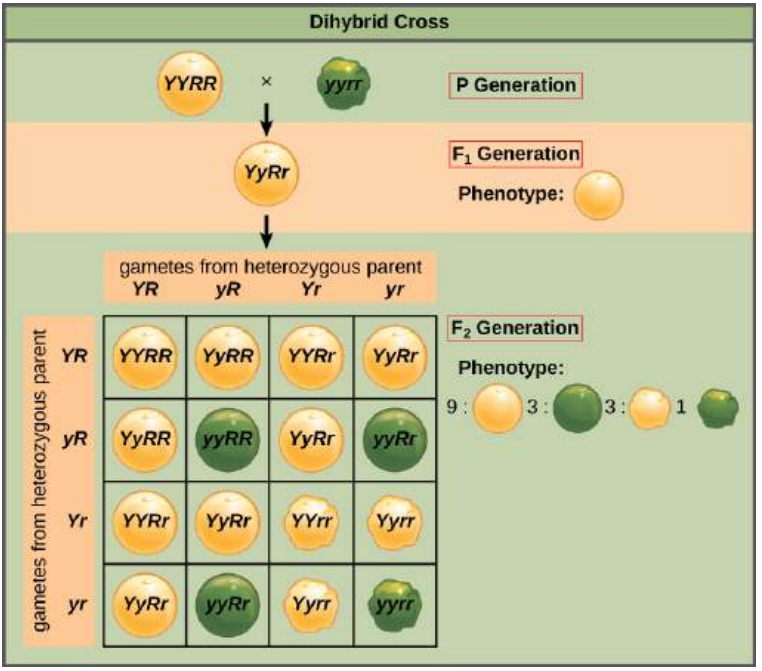


Figure 5: This dihybrid cross shows the expected offspring from the F2 generation after crossing YYRR x yyrr. Compare the results from this Punnett square to the results seen in the previous figure. They match! Photo Credit: OpenStax Biology.

The gametes produced by the F1 individuals must have one allele from each of the two genes. For example, a gamete could get an *R* allele for the seed shape gene and either a *Y* or a *y* allele for the seed color gene. It cannot get both

an *R* and an *r* allele; each gamete can have only one allele per gene. The law of independent assortment states that a gamete into which an *r* allele is sorted would be equally likely to contain either a *Y* or a *y* allele. Thus, there are four equally likely gametes that can be formed when the *RrYy* heterozygote is self-crossed, as follows: *RY*, *rY*, *Ry*, and *ry*. Arranging these gametes along the top and left of a 4 × 4 Punnett square (**Figure 5**) gives us 16 equally likely genotypic combinations. From these genotypes, we find a phenotypic ratio of 9 round–yellow:3 round–green:3 wrinkled–yellow:1 wrinkled–green (**Figure 5**). These are the offspring ratios we would expect, assuming we performed the crosses with a large enough sample size.

We can look for individuals who have the recessive phenotype for Y and the dominant phenotype for R. These individuals must have two little y's and at least one big R. The possible genotypes are yyRR or yyRr. Examining the Punnett square in **Figure 5**, we can find 3 individuals with these genotypes (they are round and green). If you compare the results from **Figure 4** and **Figure 5**, you'll see that we have arrived at the same value: 3/16!

REFERENCES

Black fur color: a dominant trait

BLACK FUR COLOR IS DOMINANT OVER BROWN

Figure 1 This chocolate lab has two recessive alleles of the TYRP1 gene. (Credit: Rob Hanson; photo from Wikimedia.)

Most of us are familiar with the labrador retriever dog breed, such as the chocolate lab seen in **Figure 8**. But have you ever thought about what makes this dog brown? The difference between brown and black coat color in dogs is caused by a mutation in the TYRP1 gene. The TYRP1 gene provides instructions for making an enzyme called tyrosinase-related protein 1. This enzyme is required to produce a pigment called eumelanin. Eumelanin is a dark colored pigment. The TYRP1 gene is located on chromosome 11 in dogs (Parker, 2001).

A group of scientists who were interested in determining what caused the difference between black and brown coats sequenced the DNA within the protein-coding region of the TYRP1 gene (Schmutz, 2002). They identified three variations in the DNA making up the TYRP1 gene between brown dogs

and black dogs. These variations in DNA sequence are examples of different **alleles** of the TYRP1 gene.

Table 1: Variations in the TYRP1 allele that lead to brown color in dogs. Data from Schmutz, 2002.

Location	Black DNA sequence	Brown DNA sequence	Effect on protein
exon 2	TGT	CGT	changes a cysteine amino acid to a serine
exon 5	CAG	TAG	introduces a premature stop codon which results in 330 of 512 amino acids in the protein
exon 5	CCT	— (deleted)	deletion of a proline amino acid

All of these variations in the DNA sequence are predicted to cause a change in the amino acid sequence of the TYRP1 protein. These changes affect the production of eumelanin pigment, which is black in color. When eumelanin is not being produced correctly, the dog appears brown instead of black.

Like other diploid organisms, dogs all have two copies of the TYRP1 gene (one from their male parent, one from their female parent). Dogs that are homozygous for the black allele (dogs that have two copies of the black allele) are obviously going to be black in color. Dogs that are homozygous for the brown allele are obviously going to be brown. Dogs that are heterozygous (dogs that have one black allele and one brown allele) appear black. The black and brown colors do not blend together: the black allele covers up the brown allele. This means that the black allele is **dominant** over the brown allele. Remember that dominant alleles cover up **recessive** alleles. If there is one dominant allele present, the dog will appear black. The brown allele is recessive to the black allele. There must be two copies of the recessive brown allele present in order for the dog to appear brown.

Figure 2: Black and brown phenotypes in labrador retrievers. (Credit: demealiffe; from Wikimedia)

Remember that genotypes can be abbreviated with a single letter and that the letter which is chosen is typically the first letter of the dominant trait. In this case, the letter "B" is used to represent the dominant black allele, while "b" represents a recessive brown allele.

The reason that the black allele is dominant over the brown allele in this specific situation is because the black allele produces functional TYRP1 protein, while the brown allele does not. The presence of one functional allele produces enough TYRP1 protein allows the cells to produce eumelanin and appear black.

Remember: dominant does not mean "better" or "more normal". Black color does not confer any special advantages on dogs compared to brown color. It's just a difference.

Figure 3: What alleles of TYRP1 does this black lab puppy have? We can't tell by looking at it. The puppy could be homozygous (BB) or heterozygous (Bb). Since black is completely dominant over brown, both options would be black. (Credit: Alice Birkin)

Let's visualize the inheritance of black and brown using a pedigree. The pedigree in **Figure 4** shows a litter of puppies. The shaded symbol shows a brown puppy, while open symbols are black individuals.

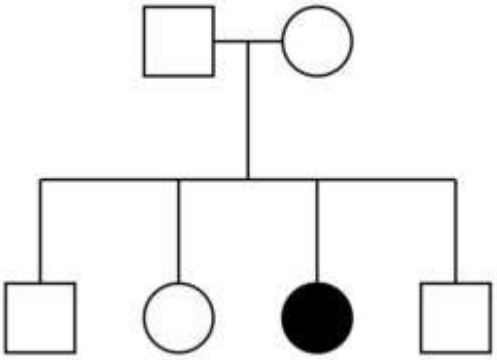


Figure 4: An example litter of puppies. The filled-in symbol shows a brown individual.

To interpret this pedigree, let's start with information that we already know:

- Brown is recessive, which means brown individuals must have the genotype bb. In this pedigree, brown individuals are filled in.
- Black is dominant, which means black individuals must have at least one B allele. Their genotype could be either BB or Bb. In this pedigree, black individuals are not filled in.

Figure 5 shows the same pedigree, but with information about the individual's genotypes filled in.

1. The shaded individual, who is a brown female puppy, must have the genotype bb. If she had any B alleles, she would be black because the black allele is dominant over the brown allele.
2. In order for the brown puppy to have the genotype bb, she must have gotten two "b" alleles: one from

each of her parents. We know that her parents are both black (because they are unshaded), which means they must have a least one "B" allele. This means that both parents must be heterozygous: Bb.

3. The three black puppies must have at least one "B" allele in order for them to be black in color. However, we can't tell whether they are homozygous dominant (BB) or heterozygous (Bb) since both of those genotypes would result in black color. One way to represent this on a pedigree is B-, meaning that the second allele could be either B or b.

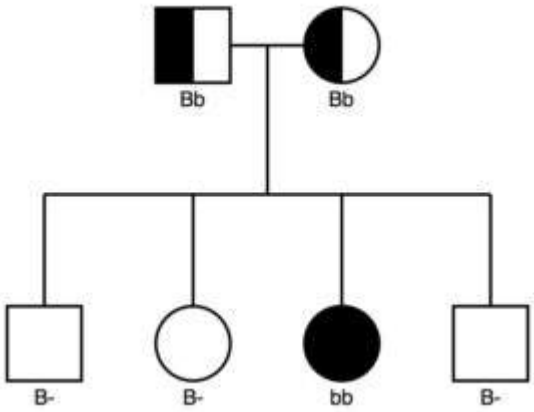


Figure 5: Genotypes of the individuals in this pedigree.

We can also show the cross between these parents as a Punnett square (**Figure 6**). We would expect 1/4 of the offspring to have the genotype bb, and that is what we see in the pedigree above.

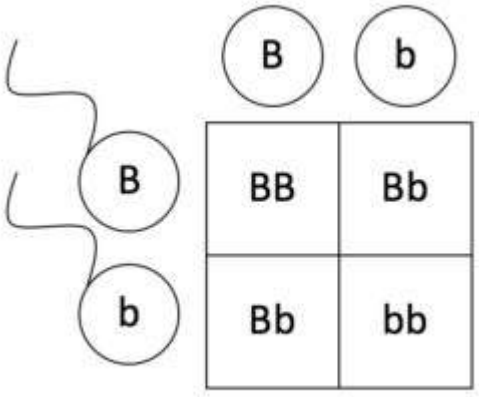


Figure 6: The information from the pedigree shown in Figure can also be shown as a Punnett square.

Human Connection

A small number of mutations in the TYRP1 gene have been found to cause oculocutaneous albinism type 3. This condition includes a form of albinism called rufous oculocutaneous albinism, which has been described primarily in dark-skinned people from southern Africa. Affected individuals have reddish-brown skin, ginger or red hair, and hazel or brown irises. Two TYRP1 mutations are known to cause this form of albinism in individuals from Africa. One mutation replaces a protein building block (amino acid) in tyrosine-related protein 1 with a signal that prematurely stops protein production. This mutation, written as Ser166Ter or S166X, affects the amino acid serine at protein position 166. The other mutation, written as 368delA, deletes a single DNA building block from the TYRP1gene. Other alterations in this gene have been reported in a few affected people of non-African heritage.

Most TYRP1 mutations lead to the production of an abnormally short, nonfunctional version of tyrosinase-related protein 1. Because this enzyme plays a role in normal pigmentation, its loss leads to the changes in skin, hair, and eye coloration that are characteristic of oculocutaneous albinism.

Photo credit: Muntuwandi; from Wikipedia.

REFERENCES

Unless otherwise noted, text and images by Lisa Bartee, 2016.

Parker HG, Yuhua X, Mellersh CS, Khan S, Shibuya H, Johnson GS, Ostrander EA. Sept 2001. Meiotic linkage mapping of 52 genes onto the canine map does not identify significant levels of microrearrangement. Mamm Genome. 12(9):713-8.

Schmutz SM, Berryere TG, Goldfinch AD. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. Mammalian Genome 13, 380-387.

OpenStax, Biology. OpenStax CNX. May 27,

2016 http://cnx.org/contents/s8Hh0oOc@9.10:FtsD6vMd@3/Mendels-Experiments

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:sbdXt0s3@4/Laws-of-Inheritance

Information about TYRP1 and oculocutaneous albinism type 3: "Tyrp1" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

Yellow fur color: a recessive trait

YELLOW COLOR IN DOGS

Labrador retrievers don't only come in brown and black, they also come in yellow. Yellow color in labs is caused by variations in a different gene: MC1R. This gene controls the production of the melanocortin 1 receptor protein. MC1R is located on chromosome 5 in dogs (Schmutz, 2001).

Figure 1: This yellow lab is producing light-colored pheomelanin instead of dark-colored eumelanin. (Credit: Djmirko; from Wikimedia)

Melanocytes make two forms of melanin, eumelanin and pheomelanin. The relative amounts of these two pigments help determine the color of an individual's hair and skin. Individuals who produce mostly eumelanin tend to have brown or black hair and dark skin that tans easily (in humans). Eumelanin also protects skin from damage caused by ultraviolet (UV) radiation in sunlight. Individuals who produce mostly pheomelanin tend to have red or blond hair,

freckles, and light-colored skin that tans poorly. Because pheomelanin does not protect skin from UV radiation, people with more pheomelanin have an increased risk of skin damage caused by sun exposure.

The melanocortin 1 receptor controls which type of melanin is produced by melanocytes. When the receptor is activated, it triggers a series of chemical reactions inside melanocytes that stimulate these cells to make eumelanin. If the receptor is not activated or is blocked, melanocytes make pheomelanin instead of eumelanin. This means that if the receptor is working correctly and is turned on, dark pigment will be produced. If the receptor is not functional or is not turned on, light pigment will be produced.

Figure 2: The three recognized colors of labs are due to black eumelanin, brown eumelanin, or pheomelanin. (Credit: Erikeltic, from Wikimedia)

Schmutz et. al. (2002) determined the DNA sequence for the MC1R gene from dogs of various colors. They determined that black and brown dogs all have one allele of MC1R, while yellow and red dogs have a different allele. The allele that leads to yellow or red color has a premature stop codon which results in a shorter-than-normal protein. This protein would be predicted to not function correctly. Remember that when the melanocortin 1 receptor is not functioning correctly, light pheomelanin pigment is produced and not dark eumelanin.

Dogs that are homozygous for the functioning allele of MC1R (which would cause eumelanin to be produced) are dark in color. Dogs that are homozygous for the non-functioning allele (which would cause pheomelanin to be produced) are light in color. Dogs that are heterozygous are dark in color. What does this tell you about which allele is dominant? If you said “the dark allele is dominant because it covers up the light allele”, you’re correct. We will use “E” to represent the genotype at MC1R because the dominant phenotype in this case is the production of eumelanin. Dogs that have the genotype EE or Ee will produce eumelanin and be dark. Dogs that have the genotype “ee” will produce pheomelanin and be light.

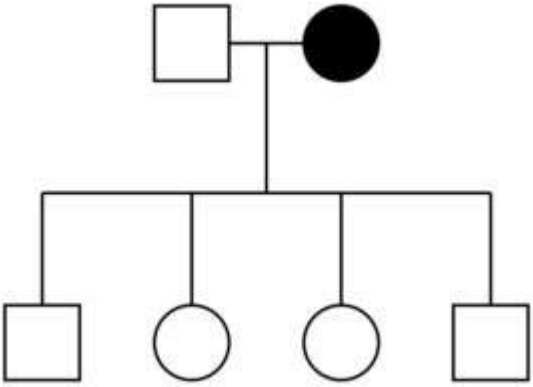


Figure 3: In this pedigree, the shaded individual is yellow. She therefore has the genotype ee and produces pheomelanin. We can't tell the genotype of her mate by looking (he could be Ee or EE), but since all of their puppies were dark in color, we would predict that his genotype was EE. In this cross: EE x ee, 100% of the puppies would have the genotype Ee, so 100% of the puppies would produce eumelanin instead of pheomelanin.

The cross shown in **Figure 3** can also be shown as a Punnett square (Figure 4). Since we are unsure whether the male dog has the genotype "EE" or "Ee", we have to make two Punnett squares. Since all of the puppies resulting from this cross were black, we would predict that the first Punnett square shows the cross. However, it is possible that the second Punnett square is correct. There are only 4 puppies, so it's not hard to imagine that they could all be black even though the Punnett square predicts only 50% black. It would be comparable to flipping a coin 4 times and getting 4 heads in a row. Getting 4 heads in a row is less likely, but definitely possible.

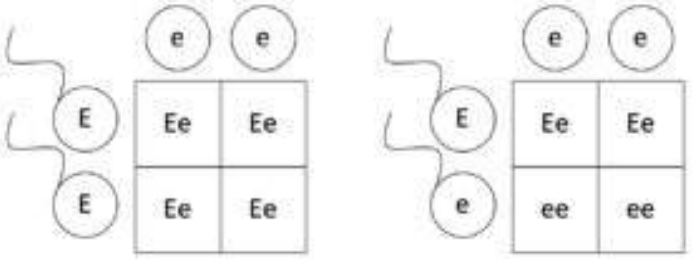


Figure 4: Cross from Figure 3 shown as Punnett squares

It is very important to note here that yellow dogs still have the TYRP1 gene, even though they are not black or brown!

Human Connection

Common variations (polymorphisms) in the MC1R gene are associated with normal differences in skin and hair color. Certain genetic variations are most common in people with red hair, fair skin, freckles, and an increased sensitivity to sun exposure. These MC1R polymorphisms reduce the ability of the melanocortin 1 receptor to stimulate eumelanin production, causing melanocytes to make mostly pheomelanin. Although MC1R is a key gene in normal human pigmentation, researchers believe that the effects of other genes also contribute to a person's hair and skin coloring.

The melanocortin 1 receptor is also active in cells other than melanocytes, including cells involved in the body's immune and inflammatory responses. The receptor's function in these cells is unknown.

Photo credit: dusdin on flickr; from Wikipedia.

RESOURCES

Unless otherwise noted, text and images by Lisa Bartee, 2016.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:FtsD6vMd@3/Mendels-Experiments

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:sbdXt0s3@4/Laws-of-Inheritance

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:zLLYW2hj@5/Extensions-of-the-Laws-of-Inhe

Human Connection – information about MC1R: "MC1R" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of

Health: U.S> National Library of Medicine is in the Public Domain

Schmutz SM, Moker JS, Berryere TG, Christison KM, Dolf G. 2001. An SNP is used to map MC1R to dog chromosome 5. Anim Genet. 32(1):43-4.

Schmutz SM, Berryere TG, Goldfinch AD. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. Mammalian Genome 13, 380-387.

Epistasis: the relationship between black, brown, and yellow fur

EPISTASIS

Dogs don't have either the TYRP1 gene *or* the MC1R gene – they have both. In fact, every dog will have two copies of the TYRP1 gene and two copies of the MC1R gene. Since both genes control aspects of coat color, it makes sense that they interact. In fact, TYRP1 and MC1R have what is called an epistatic relationship: the action of one gene controls the expression of a second gene. Another way to phrase this relationship is that the effect of one gene is dependent on another gene.

Remember that TYRP1 is required for the production of eumelanin. The dominant allele of TYRP1 (B) produces black eumelanin, while the recessive allele (b) produces brown eumelanin. However, if a dog is homozygous recessive for MC1R (ee), they lack the ability to produce eumelanin at all. If no eumelanin is being produced, it doesn't matter whether it would have been black or brown: there is none. This means that any dog that is homozygous recessive for MC1R will appear yellow regardless of its genotype at TYRP1. These two genes are epistatic: the action of MC1R controls the expression of TYRP1. The effect of TYRP1 is dependent on MC1R.

If a dog has at least one dominant functioning allele of MC1R, then its genotype at TYRP1 can be seen. If the dog has at least one dominant allele of TYRP1, it will appear black. If it has two recessive alleles, it will appear brown.

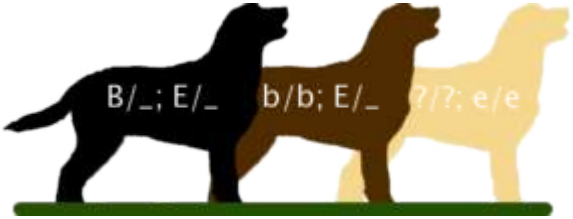


Figure 1: Genotypes for TYRP1 (B) and MC1R (E) that lead to the three recognized colors of labs. (Credit EArellano, from Wikimedia)

A pedigree can be used to show the inheritance of two different genes such as TYRP1 and MC1R.

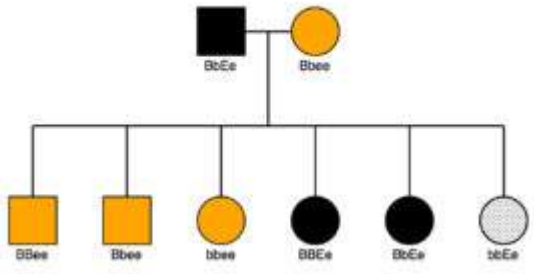


Figure 2: In this pedigree, a cross between an individual who is heterozygous for both MC1R and TYRP1 and an individual who has the genotype "Bbee"is shown. Black individuals are shaded black, yellow individuals are shaded yellow, and brown individuals are shaded grey. The 6 different possible genotypes are each shown as one offspring. This does not give you any information about the probability of getting a certain genotype of offspring – it gives you the actual number of offspring observed and their traits.

Punnett squares can also be used to show this cross. If the probability of inheriting one trait is multiplied by the

probability of inheriting the second trait, the overall probability of getting any given offspring can be determined.

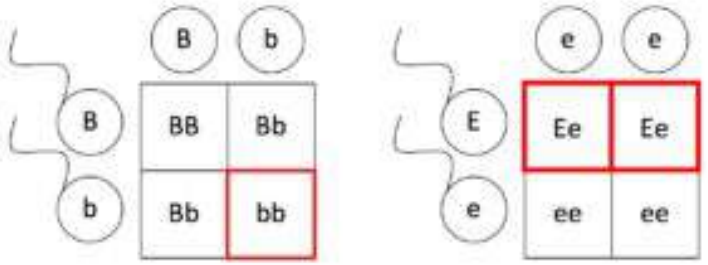


Figure 3: These two Punnett squares can be used to determine the results of a cross between these individuals: Bbee x BbEe. If you wanted to determine the probability of getting a brown dog, you would multiply the probability of getting bb by the probability of having at least one dominant E. That would equal 1/4 x 1/2 = 1/8. This gives you the probability of getting a brown dog, but doesn't tell you anything about the number of brown dogs actually observed.

Human Connection

Individuals who have albinism lack the ability to produce any pigment. If no pigment is being produced, the color that the pigment would have been is unimportant. The effect of the pigment genes is controlled by the gene that allows pigment to be produced. This is an example of epistasis.

Albinism can occur in humans (see the section on TYRP1) as well as other animals, such as the squirrel seen below.

Photo credit: Stephenkniatt from Wikipedia.

REFERENCES:

Unless otherwise noted, text and images by Lisa Bartee, 2016.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:zLLYW2hj@5/Extensions-of-the-Laws-of-Inhe

Schmutz SM, Berryere TG, Goldfinch AD. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. Mammalian Genome 13, 380-387.

Brindle color: partial dominance and epistasis

Brindle coloration is a black and brown striping pattern that is caused by different alleles at the "K locus", which is probably a gene called ASIP that controls pigment switching (Figure 1; Ciampolini, 2013). There are three alleles of the K locus: K^B, k^{br}, and k^y (Kerns, 2007). The K^B allele is dominant over the other two alleles and produces solid black color. k^{br} produces the brindle color pattern and is dominant over the k^y allele. This means that dogs with the genotype $k^{br}k^{br}$ or $k^{br}k^y$ will have the brindle color pattern. Dogs with the genotype k^yk^y are yellow in color.

Figure 1 This boxer shows the brindle color pattern, which looks sort of like tiger stripes. (Credit: Steve Henderson Location: Memphis, TN)

The K locus and MC1R (which controls the difference between dark eumelanin and light pheomelanin production) have an epistatic relationship. If a dog has two recessive

alleles for MC1R and is therefore unable to make eumelanin, the dog will appear yellow regardless of its genotype at the K locus.

REFERENCES

Unless otherwise noted, text and images by Lisa Bartee, 2016.

Ciampolini R, Cecchi F, Spaterna A, Bramante A, Bardet SM, Oulmouden A. 2013. Characterization of different 5'-untranslated exons of the ASIP gene in black-and-tan Doberman Pinscher and **brindle** Boxer **dogs**. Anim Genet. 44(1):114-7.

Kerns JA, Cargill EJ, Clark LA, Candille SI, Berryere TG, Olivier M, Lust G, Todhunter RJ, Schmutz SM, Murphy KE, Barsh GS. 2007. Linkage and segregation analysis of black and **brindle** coat color in domestic **dogs**. Genetics. 176(3):1679-89.

Incomplete dominance: when traits blend

FLOWER COLOR IN SNAPDRAGONS

Mendel's results in crossing peas, black vs brown fur color, and eumelanin production vs pheomelanin production all demonstrate traits are inherited as dominant and recessive. This contradicts the historical view that offspring always exhibited a blend of their parents' traits. However, sometimes heterozygote phenotype is intermediate between the two parents. For example, in the snapdragon, *Antirrhinum majus* (**Figure 1**), a cross between a homozygous parent with white flowers (C^WC^W) and a homozygous parent with red flowers (C^RC^R) will produce offspring with pink flowers (C^RC^W) (**Figure 2**).

Figure 1: These pink flowers of a heterozygote snapdragon result from incomplete dominance. (credit: "storebukkebruse"/Flickr)

Note that different genotypic abbreviations are used to distinguish these patterns from simple dominance and recessiveness. The abbreviation C^W can be read as "at the flower color gene (C), the white allele is present."

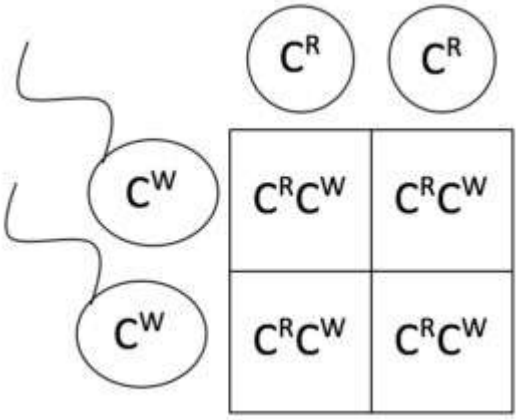


Figure 2: A cross between a red and white snapdragon will yield 100% pink offspring.

This pattern of inheritance is described as **incomplete dominance**, meaning that neither of the alleles is completely dominant over the other: both alleles can be seen at the same time. The allele for red flowers is incompletely dominant over the allele for white flowers. Red + white = pink. The results of a cross where the alleles are incompletely dominant can still be predicted, just as with complete dominant and recessive crosses. **Figure 3** shows the results from a cross between two heterozygous individuals: C^RC^W x C^RC^W. The expected offspring would have the genotypic ratio 1 C^RC^R:2 C^RC^W:1 C^WC^W, and the phenotypic ratio would be 1:2:1 for red:pink:white. The basis for the intermediate color in the heterozygote is simply that the pigment produced by the red allele (anthocyanin) is diluted in the heterozygote and therefore appears pink because of the white background of the flower petals.

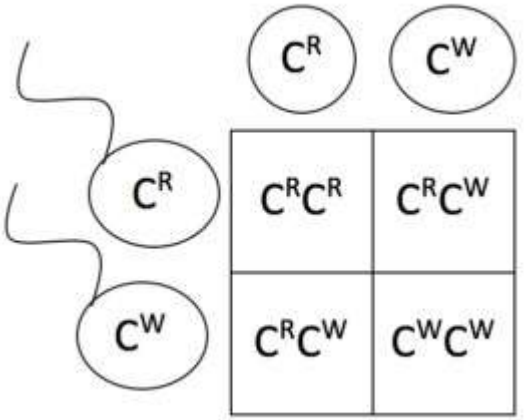


Figure 3: The results of crossing two pink snapdragons.

STRAIGHT, CURLY, AND WAVY HAIR IN DOGS

Figure 4: The wavy hair on this labradoodle is caused by incomplete dominance. (Credit: Localpups, Flickr)

Another example of incomplete dominance is the inheritance of straight, wavy, and curly hair in dogs. The KRT71 gene is used to synthesize the keratin 71 protein. Genes in the KRT family provide instructions for making proteins called keratins. Keratins are a group of tough, fibrous proteins that form the structural framework of epithelial cells, which are cells that line the surfaces and cavities of the body. Epithelial cells make up tissues such as the hair, skin, and nails. These cells also line the internal organs and are an important part of many glands.

Keratins are best known for providing strength and resilience to cells that form the hair, skin, and nails. These proteins allow tissues to resist damage from friction and minor trauma, such as rubbing and scratching. Keratins are also involved in several other critical cell functions, including cell movement (migration), regulation of cell size, cell growth and division (proliferation), wound healing, and transport of materials within cells. Different combinations of keratin proteins are found in different tissues.

The mutation which causes curly hair in dogs, such as the labradoodle seen in Figure 23, is in exon 2 of the gene and is predicted to substantially disrupt the structure of the keratin 71 protein (Cadieu, 2009). This change in protein shape prevents the keratin proteins from interacting together correctly within the hair, altering the structure of the hair and resulting in a curly coat (Runkel, 2006).

When a dog has two curly alleles ($K^{C}K^{C}$), it has a very curly coat, such as on the poodle in **Figure 5**. A dog with two straight alleles ($K^{+}K^{+}$) has a straight coat. Dogs that are heterozygous ($K^{+}K^{C}$) have an intermediate or wavy coat like the labradoodle in **Figure 4**.

Figure 24: This poodle has two copies of the curly allele of the KRT71 gene (KCKC). Compare his curly hair to the wavy hair of the labradoodle in Figure 23. The labradoodle is heterozygous (K+KC). (Credit B. Schoener; From Wikimedia)

Human Connection – Blood Type

Blood is classified into different groups according to the presence or absence of molecules called antigens on the surface of every red blood cell in a person's body. Antigens determine blood type and can either be proteins or complexes of sugar molecules (polysaccharides). The genes in the blood group antigen family provide instructions for making antigen proteins. Blood group antigen proteins serve a variety of functions within the cell membrane of red blood cells. These

protein functions include transporting other proteins and molecules into and out of the cell, maintaining cell structure, attaching to other cells and molecules, and participating in chemical reactions.

There are 29 recognized blood groups, most involving only one gene. Variations (polymorphisms) within the genes that determine blood group give rise to the different antigens for a particular blood group protein. For example, changes in a few DNA building blocks (nucleotides) in the ABO gene give rise to the A, B, and O blood types of the ABO blood group. The changes that occur in the genes that determine blood group typically affect only blood type and are not associated with adverse health conditions, although exceptions do occur.

The A and B alleles are codominant, which is similar to incomplete dominance in that heterozygotes have an intermediate phenotype. If both the A and B alleles are present, both will be seen in the phenotype. The O allele is recessive to both A and B.

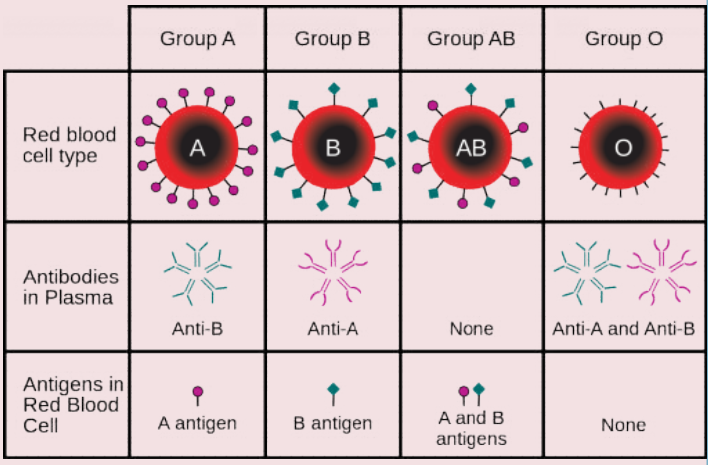


Photo credit: InvictaHOG, from Wikipedia.

REFERENCES

Unless otherwise noted, text and images by Lisa Bartee, 2016.

Cadieu E, Neff MW, Quignon P, Walsh K, Chase K, Parker HG, Vonholdt BM, Rhue A, Boyko A, Byers A, Wong A, Mosher DS, Elkahloun AG, Spady TC, André C, Lark KG, Cargill M, Bustamante CD, Wayne RK, Ostrander EA. 2009. Coat variation in the domestic **dog** is governed by variants in three genes. Science. 326(5949):150-3.

Runkel F, Klaften M, Koch K, Böhnert V, Büssow H, Fuchs H, Franz T, Hrabé de Angelis M. 2006. Morphologic and molecular characterization of two novel Krt71 (Krt2-6g) mutations: Krt71rco12 and Krt71rco13. Mamm Genome. 17(12):1172-82.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:zLLYW2hj@5/Extensions-of-the-Laws-of-Inhe

"Blood Group Antigens" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

"Keratins" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

White spotting: When there's more than two alleles

So far, we have discussed genes which have only two alleles. However, that is not always the case: there can be more than two alleles for a given gene. One example is the MITF gene, which is the major gene that controls white spotting in dogs. This protein is required for the migration and survival of melanocytes into the skin during development. If it is not functional, it impairs the ability of the skin to make pigment, thus "covering up" the effect of other color genes. There are thought to be at least four alleles that can contribute (Karlsson, 2007). Depending on which alleles are present in a dog, the amount of white can vary from none (a solid-colored dog) to mostly white (**Table 2** and **Figure 1**).

Table 2: Combinations of different alleles for MITF result in different amounts of white present in the coat.

Alleles	Amount of white
$S^{}S^{}$	None (solid colored)
$S^{}s^{i}$	Small amounts of white possible on face, chest, feet, and tail tip
$S^{}s^{p}$	More than 50% solid colored, with white on the face, chest, feet, collar, underbelly, and tail tip
$s^{i}s^{p}$	Approximately even amounts of color and white
$s^{i}s^{e}$	More than 50% white with irregular splashes of color
$s^{e}s^{e}$	Mostly white with only minimal areas of color, perhaps on one or both ears, an eye patch, or a spot near the tail

Figure 1: These dogs have different combinations of alleles of the MITF gene. The first dog probably has the genotype"Ssp"; the dog in the center is likely "SS"; the dog on the right is likely "sese". (Credits: Funny black dog by X posid from Publicdomainpictures. A black and white dog by Petr Kratochvil from Free stock photos. White dog with black ears by RetyiRetyi from Pixabay.)

There is a similar white-spotting gene in domestic cats that leads to the bicolor coat pattern commonly called "tuxedo" (Figure 2). However, the genetic causes of domestic cat coat patterns are not well understood and there is no consensus on which specific gene controls white spotting in cats. Some studies suggest that the gene responsible is the feline version of the KIT gene (Cooper, et al. 2006; Montague, et al. 2014), but further research on mammalian pigmentation is needed before a conclusion can be made (Mort, et al. 2016).

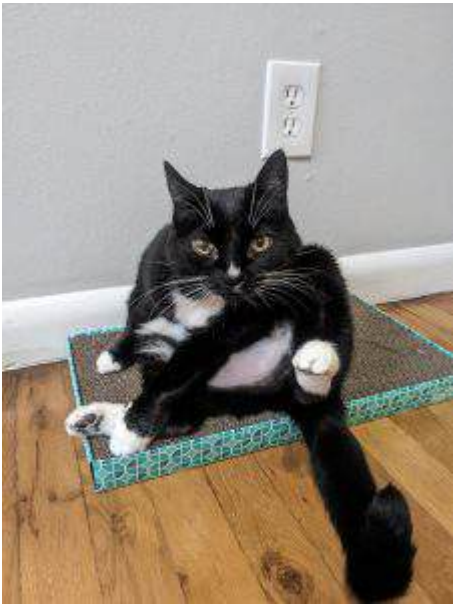


Figure 2: An image of the author's bicolored cat in an undignified pose that displays the white spotting on the underbelly, nose, and paws.

Human Connection – Blood Type

Human blood type was discussed in the previous section. You may remember that there are three alleles for the ABO gene: A, B, and O. A and B are codominant, meaning that if both alleles are present, both will be seen in the phenotype. A person with type AB blood has one A allele and one B allele.

O is recessive to A and B. A person with the genotype AO will have Type A blood. A person with the genotype BO will have type B blood. Type O blood results from two O alleles.

	A	B	O
A	AA	AB	AO
B	AB	BB	BO
O	AO	BO	OO

Photo credit: Kalaiarasy, from Wikipedia.

REFERENCES

Unless otherwise noted, text and images by Lisa Bartee (2016) and Catherine Creech (2023).

Karlsson EK, Baranowska I, Wade CM, Salmon Hillbertz NH, Zody MC, Anderson N, Biagi TM, Patterson N, Pielberg GR, Kulbokas EJ 3rd, Comstock KE, Keller ET, Mesirov JP, von Euler H, Kämpe O, Hedhammar A, Lander ES, Andersson G, Andersson L, Lindblad-Toh K. 2007. Efficient mapping of mendelian traits in **dogs** through genome-wide association. Nat Genet. 39(11):1321-8.

Montague MJ, Li G, Gandolfi B, Khan R, Aken BL, Searle SM, Minx P, Hillier LW, Koboldt DC, Davis BW, Driscoll CA, Barr CS, Blackistone K, Quilez J, Lorente-Galdos B, Marques-Bonet T, Alkan C, Thomas GW, Hahn MW, Menotti-Raymond M, O'Brien SJ, Wilson RK, Lyons LA, Murphy WJ, Warren WC. 2014. Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. Proc Natl Acad Sci USA. 111(48):17230-5.

Cooper MP, Fretwell N, Bailey SJ, Lyons LA. 2006. White spotting in the domestic cat (Felis catus) maps near KIT on feline chromosome B1. Anim Genet. 37(2):163-5.

Mort, R. L., Ross, R. J. H., Hainey, K. J., Harrison, O. J., Keighren, M. A., Landini, G., Baker, R. E., Painter, K. J., Jackson, I. J., & Yates, C. A. 2016. Reconciling diverse mammalian pigmentation patterns with a fundamental mathematical model. Nature Communications, 7(1), 10288.

Hemophilia: a sex-linked disorder

So far, all the genes we have discussed have had two copies present in all individuals. This is because the individual inherited one from the male parent's haploid gamete and one from the female parent's haploid gamete. The two gametes came together during fertilization to produce a diploid individual. There is, however, one exception to this: genes which are present on the sex chromosomes.

In humans, as well as in many other animals and some plants, the sex of the individual is determined by **sex chromosomes** – one pair of non-homologous chromosomes. Until now, we have only considered inheritance patterns among non-sex chromosomes, or **autosomes.** In addition to 22 homologous pairs of autosomes, human females have a homologous pair of X chromosomes, whereas human males have an XY chromosome pair. Although the Y chromosome contains a small region of similarity to the X chromosome so that they can pair during meiosis, the Y chromosome is much shorter and contains fewer genes. When a gene being examined is present on the X, but not the Y, chromosome, it is **X-linked**.

The X chromosome is one of two sex chromosomes. Humans and most mammals have two sex chromosomes, the X and Y. Females have two X chromosomes in their cells, while males have X and Y chromosomes in their cells. Egg cells all contain an X chromosome, while sperm cells contain an X or a Y chromosome. This arrangement means that during fertilization, it is the male that determines the sex of the offspring since the female can only give an X chromosome to the offspring.

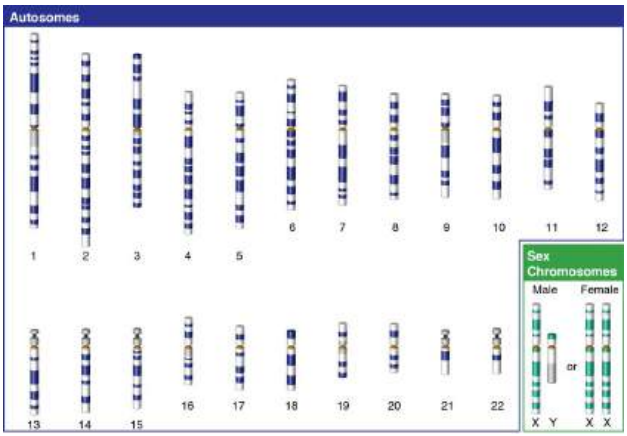


Figure 1: A diagram showing the autosomal and sex chromosomes. Remember that in a diploid cell, there would be two copies of each autosomal chromosome present. (Credit: Darryl Lega, NHGRI)

Most sex-linked genes are present on the X chromosome simply because it is much larger than the Y chromosome. The X chromosome spans about 155 million DNA base pairs and represents approximately 5 percent of the total DNA in cells. The X chromosome likely contains 800 to 900 genes. In contrast, the Y chromosome has approximately 59 million base pairs and only 50-60 genes. Sex is determined by the SRY gene, which is located on the Y chromosome and is responsible for the development of a fetus into a male. This means that the presence of a Y chromosome is what causes a fetus to develop as male. Other genes on the Y chromosome are important for male fertility.

Hemophilia is a bleeding disorder that slows the blood clotting process. People with this condition experience prolonged bleeding or oozing following an injury, surgery, or having a tooth pulled. In severe cases of hemophilia, continuous bleeding occurs after minor trauma or even in the absence of injury (spontaneous bleeding). Serious

complications can result from bleeding into the joints, muscles, brain, or other internal organs. Milder forms of hemophilia do not necessarily involve spontaneous bleeding, and the condition may not become apparent until abnormal bleeding occurs following surgery or a serious injury.

The major types of this condition are hemophilia A (also known as classic hemophilia or factor VIII deficiency) and hemophilia B (also known as Christmas disease or factor IX deficiency). Although the two types have very similar signs and symptoms, they are caused by mutations in different genes.

Hemophilia A and hemophilia B are inherited in an X-linked recessive pattern. The genes associated with these conditions are located on the X chromosome, which is one of the two sex chromosomes. In males (who have only one X chromosome), one altered copy of the gene in each cell is sufficient to cause the condition. In females (who have two X chromosomes), a mutation would have to occur in both copies of the gene to cause the disorder. Because it is unlikely that females will have two altered copies of this gene, it is very rare for females to have hemophilia. A characteristic of X-linked inheritance is that fathers cannot pass X-linked traits to their sons (Figure 2 and 3).

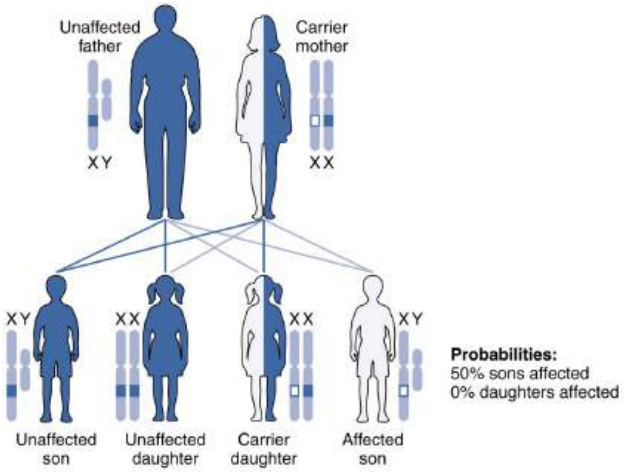


Figure 2 X-linked recessive inheritance. Photo credit OpenStax College; [OpenStax Anatomy and Physiology](#).

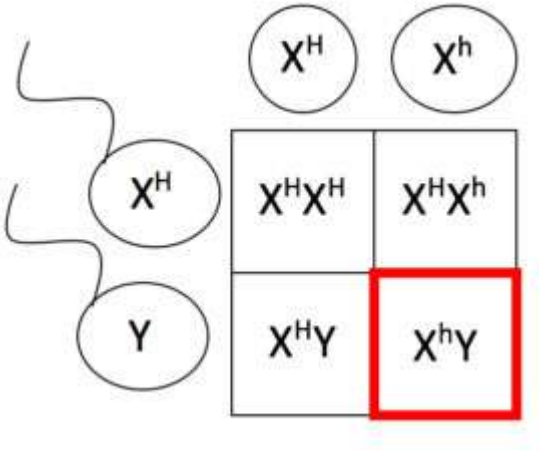


Figure 3: If a carrier female and a normal male produce offspring, there is a 25% total chance that they will have a child with hemophilia. None of their daughters will have the disease (although all will be carriers). Half their sons will be hemophiliacs.

In X-linked recessive inheritance, a female with one altered copy of the gene in each cell is called a carrier. Carrier females have about half the usual amount of coagulation factor VIII or coagulation factor IX, which is generally enough for normal blood clotting. However, about 10 percent of carrier females have less than half the normal amount of one of these coagulation factors; these individuals are at risk for abnormal bleeding, particularly after an injury, surgery, or tooth extraction.

Colorblindness is another example of a sex-linked trait in humans. The genes that produce the photopigments necessary for color vision are located on the X chromosome. If one of these genes is not functional because it contains a harmful mutation, the individual will be colorblind. Men are much more likely than women to be colorblind: up to 100 times more men than women have various types of colorblindness (http://www.colour-blindness.com/general/prevalence/).

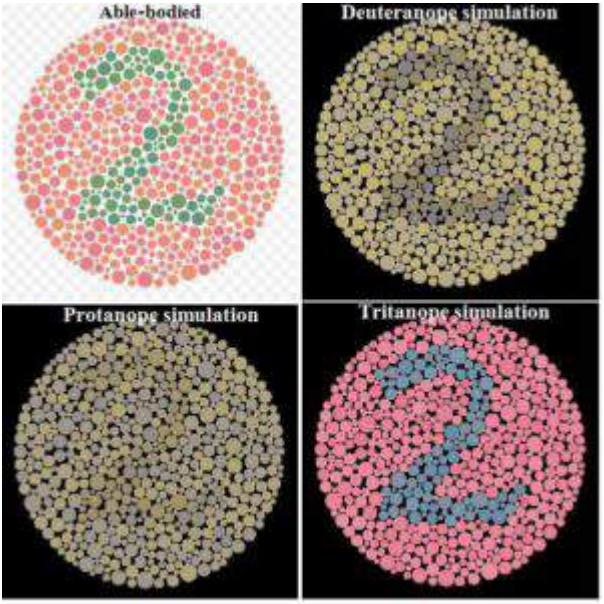


Figure 4: A test image for color-blindness as seen by someone with normal color vision and several types of colorblindness. (Credit: Sakurambo)

REFERENCES:

Unless otherwise noted, text adapted by and images by Lisa Bartee, 2016.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:zLLYW2hj@5/Extensions-of-the-Laws-of-Inhe

"X chromosome" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

"Y chromosome" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of

Health: U.S> National Library of Medicine is in the Public Domain

"Hemophilia" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

Overall phenotypes: putting it all together

None of the genes discussed in these sections occur in isolation: one individual dog would have all the genes for color, hair structure, and hemoglobin (dogs can get hemophilia too). Genes interact together to produce the overall phenotype of the individual.

EXAMPLE 1: SUGAR

For example, look at Sugar in **Figure 1**. She has short hair that is mostly white. The colored portion of her hair is the tiger-striped pattern termed “brindle.”

Figure 1 Sugar has short hair with brindle colored spots. (Credit: Lisa Bartee)

The difference between short and long hair in dogs is caused by different alleles of a gene called FGF5. This gene produces a protein that is important in regulating the hair growth cycle. When the protein doesn't function correctly, the growth phase of the hair cycle is longer, resulting in long hair. Short hair is the dominant trait. Since Sugar has short hair, we know she has at least one dominant allele of FGF5. We can use the letter "S" for short hair. Sugar's genotype for FGF5 is therefore "S-", meaning she has one dominant allele and we can't tell by looking at her what her second allele is.

Sugar's hair is also straight, which means that she has two straight alleles of KRT71. Her genotype would be $K^{+}K^{+}$.

Sugar is more than 50% white with irregular splashes of color, which means that her genotype for MITF (the gene that controls white spotting) is $s^{i}s^{e}$.

The brindle pattern is caused by the k^{br} allele at the K

locus. Sugar can't have the K^B allele or she would have solid color instead of the brindle pattern because K^B is dominant over k^{br} and k^y. She could have either the genotype $k^{br}k^{br}$ or $k^{br}k^y$, since the k^{br} allele is dominant over the yellow allele (k^y).

Sugar has black eumelanin pigment in her hair and nose. This means she has the dominant phenotype for TYPR1, so her genotype would be "B-". Because she has eumelanin and not pheomelanin in her coat, she has the dominant phenotype for MC1R, so her genotype would be "E-".

Sugar is a female dog who does not have hemophilia. This means that her genotype would be either X^HX^H or X^HX^h.

Putting all these together, we could say that Sugar's overall coat genotype is S- K^+K^+ s^is^e B- E- X^HX–

We could potentially determine some of the unknown alleles in her genotype if we knew anything about her parents, but Sugar was adopted from the Multnomah County Animal Shelter after being picked up as a stray. Therefore, her ancestry is unknown. However, it turns out that after having her ancestry determined using DNA sequencing, she is 100% American Staffordshire Terrier.

EXAMPLE 2: RAGS

Figure 2: Rags is similar in color to Sugar, but has a very different fur type. (Credit: Lisa Bartee)

Rags has "furnishings", a term used to describe his beard and mustache. Furnishings are caused by a mutation in the RSPO2 gene. This gene produces a protein that is involved in establishing hair follicles. The allele that leads to furnishings is dominant over the allele for no furnishings. Rags must therefore have the genotype "F-" at RSPO2. This allele also causes the long-ish hair on his legs and tail.

Gene	Genotype	Phenotype
RSPO2	FF or Ff	has furnishings
FGF5	SS or Ss	short fur (his longer fur is caused by the furnishings allele)
KRT71	$K^{+}K^{+}$	straight fur
MITF	$s^{i}s^{e}$	more than 50% white
K locus	$K^{B}K^{B}$, $K^{B}k^{br}$, or $K^{B}k^{y}$	Solid color, not brindle or yellow.
TYRP1	BB or Bb	Produces black eumelanin, not brown
MC1R	EE or Ee	Produces eumelanin instead of pheomelanin
F8	$X^{H}Y$	Male, no hemophilia

EXAMPLE 3: BLACK POODLE

Figure 3: Black poodle. (Credit: B. Schoener from Wikimedia)

Gene	Genotype	Phenotype
RSPO2	ff	no furnishings
FGF5	ss	long fur
KRT71	$K^{C}K^{C}$	curly fur
MITF	SS	entirely solid color
K locus	$K^{B}K^{B}$, $K^{B}k^{br}$, or $K^{B}k^{y}$	Solid color, not brindle or yellow.
TYRP1	BB or Bb	Produces black eumelanin, not brown
MC1R	EE or Ee	Produces eumelanin instead of pheomelanin

EXAMPLE 4: GOLDEN RETRIEVER

Figure 4: Golden Retriever. (Credit: Dirk Vorderstraße)

Gene	Genotype	Phenotype
RSPO2	ff	no furnishings
FGF5	ss	long fur
KRT71	$K^{+}K^{+}$	straight fur
MITF	SS	entirely solid color
K locus	$K^{B}K^{B}$, $K^{B}k^{br}$, or $K^{B}k^{y}$	Solid color, not brindle or yellow.
TYRP1	BB or Bb	Produces black eumelanin, not brown (seen in the nose)
MC1R	ee	Produces pheomelanin instead of eumelanin, so appears yellow

BUT WAIT, THERE'S MORE!

Figure 5: An English Cocker Spaniel. (Credit eNil)

We haven't exhaustively discussed all the genes that can affect dog appearance. For example, what gene (or genes) causes the English Springer Spaniel in **Figure 5** to be red? What gene(s) cause it to be speckled on it's back? Or lead to its freckles? There are estimated to be about 19,000 genes in the dog genome (Ostrander, 2005). The interactions of all these genes together lead to the overall phenotype of one individual dog.

If you're interested in learning more about the genes that are involved in the appearance of dogs, check out the Dog Coat Color Genetics website at http://www.doggenetics.co.uk.

REFERENCES

Unless otherwise noted, text and images by Lisa Bartee, 2016.

Ostrander EA, Wayne RK. 2005. The Canine Genome. Genome Res. 15: 1706-1716.

KIT - embryonic lethality

Although MITF is the major gene impacting white spotting in dogs, a second gene known as KIT has also been shown to have an impact in a subpopulation of German Shepherd dogs. This mutation in KIT arose very recently (it appeared spontaneously in a female dog born in 2000) and causes a phenotype called "panda spotting." The KIT gene produces a tyrosine kinase receptor protein that functions in the same melanogenesis pathway as MITF. The KIT receptor controls many important processes within the cells including growth and division, survival, and cell migration. Its signaling function is important in the development of many different types of cells, including melanocytes which product the pigment melanin.

Interestingly, no dogs have been identified that are homozygous for the mutation in KIT. This is probably because having two mutations in this gene is lethal because of its many important functions. This is called a homozygous lethal allele that results in **embryonic lethality**. There are other examples of embryonic lethal genes:

- The Agouti gene causes coat color changes in many animals, including mice. One mutation that has been isolated in mice causes a yellow colored coat in mice heterozygous for the mutation. Homozygous mice die in the womb (Figure 1).

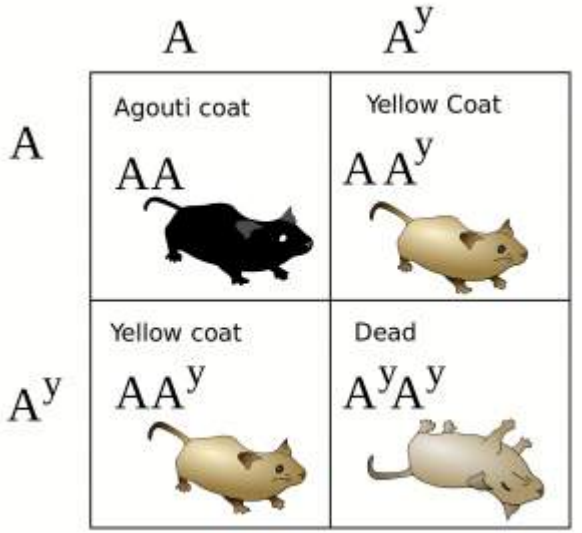


Figure 1 Lethality of homozygous agouti mutation. Photo credit Jcfidy; Wikimedia.

- Mutations in the FGFR3 gene result in Achondroplasia, a form of dwarfism in humans (Figure 2). This disorder is inherited in an autosomal dominant fashion. Two dominant alleles results in lethality.

Figure 2 Jason Acuña outside of the Waterfront Mariott in Portland, OR. on August 15, 2009. Photo credit Sakibomb222; Wikimedia.

- Short legs in dogs, called Chondrodysplasia, are associated with repeats of the FGFR4 gene (Figure 3) (Bannasch, et al. 2022). These mutations have an additive effect and breeds with the shortest legs have a higher rate of mutation frequency.

Figure 3: A corgi dog demonstrating the short leg trait. Image by Marsiyanka "Pembroke Welsh Corgi" via WikiCommons.

- Manx cats have a dominant mutation that leads to them being tailless or having short, stubby tails (Figure 4). Kittens with two dominant alleles are thought to not survive until birth (Catworld, 2009).

Figure 4: A "rumpy riser" manx kitten. Photo credit Michelle Wiegold; Wikimedia.

- Mexican hairless dogs contain a mutation in the FOXI3 gene. Heterozygous dogs are hairless (Figure

5). Homozygous mutant dogs are never observed.

Figure 5: Michelob (Michael) and Irish Mist (Misty), a hairless and a coated Xoloitzcuintli. Christopher A. and Amanda L. Dellario, Nottingham, NH, USA. Photo taken by Amanda L. Dellario, August 2006. Wikimedia.

In humans, the KIT gene is located on chromosome 4 (Figure 6), while it is located on chromosome 13 in dogs (Wong, 2012).

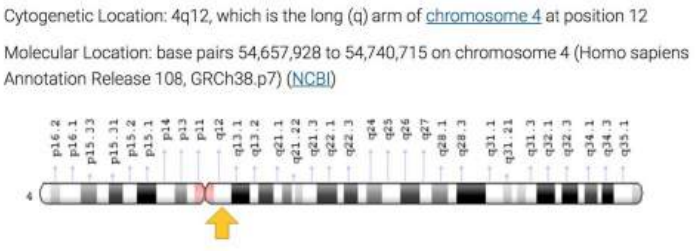


Figure 6: Chromosomal location of the KIT gene in humans. Photo credit Genetics Home Reference; Public Domain.

While KIT has only been shown to have a phenotypic effect in one family of German Shepherd dogs (at least so far), at least 69 mutations in the human KIT gene have been identified (GHR, 2018). Mutations in the human KIT gene lead to piebaldism, where melanocytes are absent from certain areas of the hair and skin (Figure 7). These mutations are inherited in an autosomal dominant fashion. I couldn't find any information about whether homozygous piebald mutations are lethal in humans. This is likely because generating a homozygous piebald human would require mating two heterozygous piebald humans, and this disorder is rare so this mating would be unlikely.

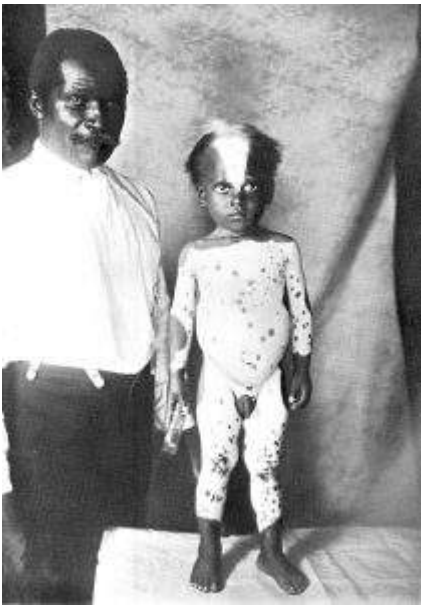


Figure 7: Piebaldism in a 5 year old boy. Neither of his parents nor any of his five siblings showed any white spotting. Photo credit Wellcome Images; Wikimedia.

REFERENCES

Bannasch, D.; Batcher, K.; Leuthard, F.; Bannasch, M.; Hug, P.; Marcellin-Little, D.J.; Dickinson, P.J.; Drögemüller, M.; Drögemüller, C.; Leeb, T. 2022. The Effects of *FGF4* Retrogenes on Canine Morphology. Genes. Available from: https://doi.org/10.3390/genes13020325

Parker, HG, et. al. 21 August 2009. An Expressed Fgf4 Retrogene Is Associated with Breed-Defining Chondrodysplasia in Domestic Dogs. Science. 325(5943): 995-998.

Wilson J. 2009. Cat World. Feline Genetic Loci Table. Available from: https://www.cat-world.com.au/feline-genetic-loci-table.html

Genetics Home Reference (GHR). 2018. National Institutes of Health. Available from https://ghr.nlm.nih.gov/gene/KIT#conditions

Wong AK, Ruhe AL, Roberson KR, Lowe ER, Williams DC, Neff MW. 2012. A *de novo* mutation in *KIT* causes white spotting in a subpopulation of German Shepherd dogs. Animal Genetics: 44:305-310.

It's not all in the genes - the effect of environment

Not all traits are directly caused by DNA alone. The environment also plays a large role in shaping an individual's traits. Some examples can be seen below.

- **Height and weight:** A number of genes interact to determine the general height and weight that a person will have. But the environment has a major influence as well. If an individual is malnourished, their growth may be slowed and they may be smaller than they would have been if they had gotten enough food. In contrast, if a person consumes more calories than they need, their weight will likely increase regardless of their genetics.

- **Fingerprints:** the general characteristics of a person's fingerprints are determined by genetics, but the specific pattern is generated randomly during development. Identical twins typically have fingerprints that are similar, but not identical.

- **Intelligence:** Like most aspects of human behavior and cognition, intelligence is a complex trait that is influenced by both genetic and environmental factors. Roughly 50% of a person's IQ appears to be determined by genetic factors. Factors related to a child's home environment and parenting, education and availability of learning resources, and nutrition, among others, also contribute to intelligence. A person's environment and genes influence each other, and it can be challenging to tease apart the effects of the environment from those of genetics. For example, if a child's IQ is similar to that of his or her parents, is that similarity due to genetic factors passed down from parent to child, to shared environmental factors, or (most likely) to a

combination of both? It is clear that both environmental and genetic factors play a part in determining intelligence.

- **Cancer Risk:** For example, a person could inherit a mutation in the BRCA1 gene, which increases the risk of developing breast or ovarian cancer. Researchers have identified more than 1,800 mutations in the BRCA1 gene. Most BRCA1 gene mutations lead to the production of an abnormally short version of the BRCA1 protein or prevent any protein from being made from one copy of the gene. As a result, less of this protein is available to help repair damaged DNA or fix mutations that occur in other genes. As these defects accumulate, they can trigger cells to grow and divide uncontrollably to form a tumor. These mutations are present in every cell in the body and can be passed from one generation to the next. As a result, they are associated with cancers that cluster in families. However, not everyone who inherits a mutation in the BRCA1 gene will develop cancer. Other genetic, environmental, and lifestyle factors also contribute to a person's cancer risk.
- In contrast, cancer can be caused by purely environmental factors. According to the CDC, cigarette smoking is the number one risk factor for lung cancer. In the United States, cigarette smoking is linked to about 90% of lung cancers and people who smoke are 15 to 30 times more likely to get lung cancer or die from lung cancer than people who do not smoke. Radon exposure also increases the likelihood that a person will develop lung cancer.

Figure: The colors on the poodle seen in this figure have no relationship to his DNA: he was dyed for a parade. (Credit: skeeze)

REFERENCES

"BRCA1" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

"Is intelligence determined by genetics?" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

Pleiotropy - one gene affects more than one trait

So far, we have discussed examples where changing the DNA sequence of one gene affects one protein, which affects one specific trait (for example, the change from brown to black fur). However, there are examples where one mutation can affect more than one trait. This is called **pleiotropy.**

MC1R

MC1R, the gene which leads to the difference between yellow and dark colored dogs, is also found in humans. Recall that MC1R helps determine whether mostly eumelanin (dark pigment) or pheomelanin (lighter reddish pigment) will be produced. Humans who produce mostly eumelanin have the active allele of MC1R and have darker skin that tans easily and darker hair. Humans who produce mostly pheomelanin have inactive MC1R proteins and have red or blonde hair and light skin with freckles that does not tan easily.

One relatively obvious secondary phenotype of having lighter skin is that rates of skin cancer are higher in those individuals compared to individuals who have darker skin that tans easily. Since MC1R affects skin pigmentation, it also has an effect on skin cancer rates. In addition, MC1R has an

effect on cancer rates that is unrelated to skin pigmentation due to its interactions with other genes that regulate inflammatory responses, DNA repair, and apoptosis (Feller, 2016).

Interestingly, red headed individuals also exhibit higher pain tolerance due to their MC1R alleles (Liem, 2004; Liem, 2006). So far, the reason for this is unknown.

Mutations in MC1R have been shown to decrease knee cartilage in mice (Lorenz, 2014). Again, the mechanism for this is not understood, but likely relates to MC1Rs signaling role.

MC1R polymorphisms have been associated with a decrease in the development of sepsis (blood poisoning) in humans (Seaton, 2017). This is probably due to the role that MC1R plays in inflammation (again, due to its role in signal transduction).

Activation of MC1R by agonists (chemicals that activate receptors) were shown to reduce several harmful symptoms in the kidneys of rats (Lindskog, 2010). The specific reason for this is not understood.

"FRIZZLED" CHICKENS

The dominant "frizzle" allele causes feathers to turn upwards rather than remain flat against the chicken's body (Figure 1). However, along with producing defective feathers, the frizzle allele also lead to abnormal body temperatures, higher metabolic and blood flow rates, and greater digestive capacity. Furthermore, chickens who had this allele also laid fewer eggs than their wild-type counterparts.

Figure 1 A "frizzled" chicken. Photo credit Joe Goldberg; Flickr.

PHENYLKETONURIA

Phenylketonuria (PKU) is a disorder that affects the levels of the amino acid phenylalanine in the body. We get phenylalanine from food, then process it within our cells. Individuals with PKU have a mutation in the enzyme required to break down phenylalanine. but levels of this amino acid build up in individuals with PKU. This build up can lead to a variety of different symptoms including intellectual disability, seizures, poor bone strength, skin rashes, behavioral and mental disorders, and an unusually small head. If you've ever seen a warning on a package that says "Phenylketonurics – contains phenylalanine", this is why.

REFERENCES

Feller L, Khammissa RAG, Kramer B, Altini M, Lemmer J. 2016. Basal cell carcinoma, squamous cell carcinoma and melanoma of the head and face. Head Face Med. 12:11.

Liem EB, et. al. 2004. Anesthetic Requirement is Increased in Redheads. Anesthesiology. 101(2): 279-283.

Liem EB, Joiner TV, Tsueda K, Sessler DI. 2006. Increased sensitivity to Thermal Pain and Reduced Subcutaneious Lidocaine Efficacy in Redheads. Anesthesiology. 102(3): 509-514.

Lindskog A, et. al. 2010. Melanocortin 1 Receptor Agonists Reduce Proteinuria. J Am Soc Néphron. 21(8): 1290-1298.

Lorenz J, et. al. 2014. Melanocortin 1 Receptor-Signaling Deficiency Results in an Articular Cartilage Phenotype and Accelerates Pathogenesis of Surgically Induced Murine Osteoarthritis. PLoS One. 9(9): e105858.

Seaton ME, Parent BA, Mood RF, Wurfel MM, Muffley LA, O'Keefe GE, Gibran NS. 2017. MELANOCORTIN-1 RECEPTOR POLYMORPHISMS AND THE RISK OF COMPLICATED SEPSIS AFTER TRAUMA: A CANDIDATE GENE ASSOCIATION STUDY. Shock. 47(1): 79-85.

BIOTECHNOLOGY

The latter half of the twentieth century began with the discovery of the structure of DNA, then progressed to the development of the basic tools used to study and manipulate DNA. These advances, as well as advances in our understanding of and ability to manipulate cells, have led some to refer to the twenty-first century as the biotechnology century. The rate of discovery and of the development of new applications in medicine, agriculture, and energy is expected to accelerate, bringing huge benefits to humankind and perhaps also significant risks. Many of these developments are expected to raise significant ethical and social questions that human societies have not yet had to consider.

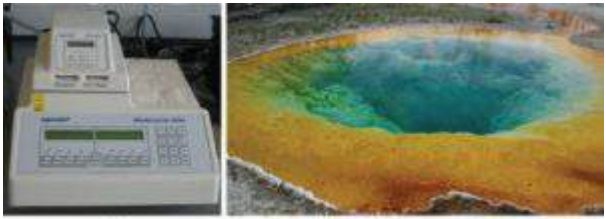


Figure 1: (a) A thermal cycler, such as the one shown here, is a basic tool used to study DNA in a process called the polymerase chain reaction (PCR). The polymerase enzyme most often used with PCR comes from a strain of bacteria that lives in (b) the hot springs of Yellowstone National Park. (credit a: modification of work by Magnus Manske; credit b: modification of work by Jon Sullivan)

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Pdu1uR8Y@2/Introduction

DNA Isolation, Gel Electrophoresis, and PCR

Biotechnology is the use of artificial methods to modify the genetic material of living organisms or cells to produce novel compounds or to perform new functions. Biotechnology has been used for improving livestock and crops since the beginning of agriculture through selective breeding. Since the discovery of the structure of DNA in 1953, and particularly since the development of tools and methods to manipulate DNA in the 1970s, biotechnology has become synonymous with the manipulation of organisms' DNA at the molecular level. The primary applications of this technology are in medicine (for the production of vaccines and antibiotics) and in agriculture (for the genetic modification of crops). Biotechnology also has many industrial applications, such as fermentation, the treatment of oil spills, and the production of biofuels, as well as many household applications such as the use of enzymes in laundry detergent.

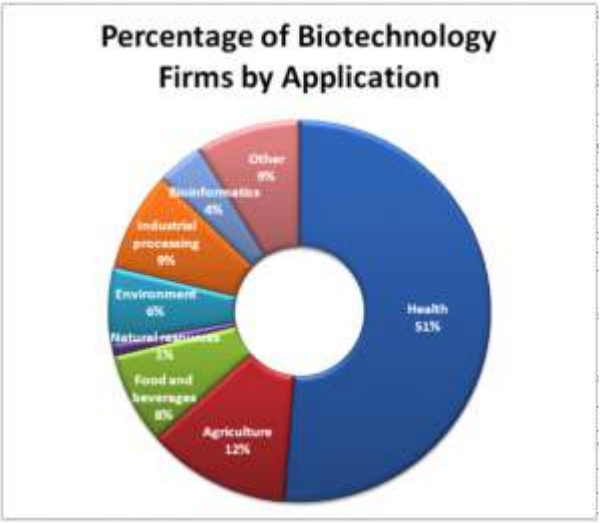


Figure 1 The majority of biotechnology companies are involved in human health. Photo credit 1Rileyw; Wikimedia.

To accomplish any of the applications described above, biotechnologists must be able to extract, manipulate, and analyze nucleic acids.

REVIEW OF NUCLEIC ACID STRUCTURE

To understand the basic techniques used to work with nucleic acids, remember that nucleic acids are macromolecules made of nucleotides (a sugar, a phosphate, and a nitrogenous base). The phosphate groups on these molecules each have a net negative charge. An entire set of DNA molecules in the nucleus of eukaryotic organisms is called the genome. DNA has two complementary strands linked by hydrogen bonds between the paired bases.

Unlike DNA in eukaryotic cells, RNA molecules leave the nucleus. Messenger RNA (mRNA) is analyzed most frequently

because it represents the protein-coding genes that are being expressed in the cell.

ISOLATION OF NUCLEIC ACIDS

To study or manipulate nucleic acids, the DNA must first be extracted from cells. Various techniques are used to extract different types of DNA (**Figure 2**). Most nucleic acid extraction techniques involve steps to break open the cell, and then the use of enzymatic reactions to destroy all undesired macromolecules. Cells are broken open using a detergent solution containing buffering compounds. To prevent degradation and contamination, macromolecules such as proteins and RNA are inactivated using enzymes. The DNA is then brought out of solution using alcohol. The resulting DNA, because it is made up of long polymers, forms a gelatinous mass. This method extracts all the nucleic acid within a cell. This includes **genomic DNA** (all the DNA in the genome), as well as RNA. If this DNA was to be used for further study, the RNA would often be digested with an enzyme to remove it.

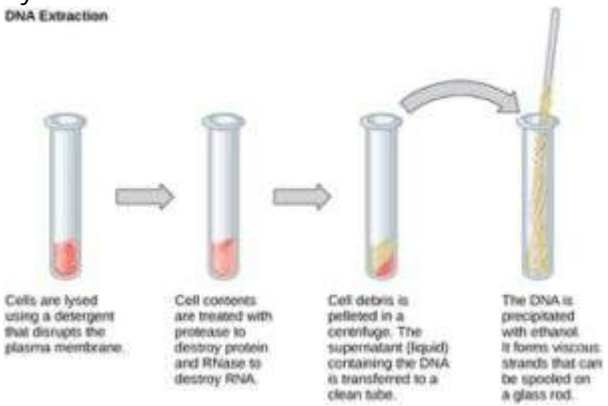


Figure 2 This diagram shows the basic method used for the extraction of DNA.

RNA is studied to understand gene expression patterns in cells. RNA is naturally very unstable because enzymes that break down RNA are commonly present in nature. Some are even secreted by our own skin and are very difficult to inactivate. Similar to DNA extraction, RNA extraction involves the use of various buffers and enzymes to inactivate other macromolecules and preserve only the RNA.

GEL ELECTROPHORESIS

Because nucleic acids are negatively charged ions at neutral or alkaline pH in an aqueous environment, they can be moved by an electric field. **Gel electrophoresis** is a technique used to separate charged molecules on the basis of size and charge. The nucleic acids can be separated as whole chromosomes or as fragments. The nucleic acids are loaded into a slot at one end of a gel matrix, an electric current is applied, and negatively charged molecules are pulled toward the opposite end of the gel (the end with the positive electrode). Smaller molecules move through the pores in the gel faster than larger molecules; this difference in the rate of migration separates the fragments on the basis of size. The nucleic acids in a gel matrix are invisible until they are stained with a compound that allows them to be seen, such as a dye. Distinct fragments of nucleic acids appear as bands at specific distances from the top of the gel (the negative electrode end) that are based on their size (**Figure 3**). A mixture of many fragments of varying sizes appear as a long smear, whereas uncut genomic DNA is usually too large to run through the gel and forms a single large band at the top of the gel.

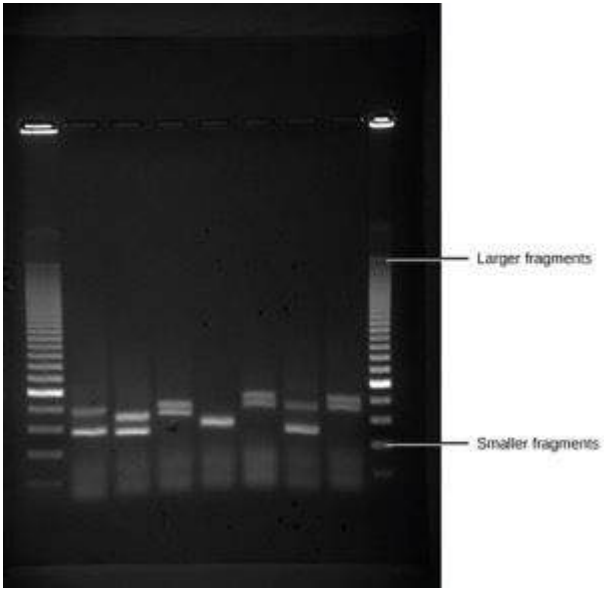


Figure 3: Shown are DNA fragments from six samples run on a gel, stained with a fluorescent dye and viewed under UV light. (credit: modification of work by James Jacob, Tompkins Cortland Community College)

POLYMERASE CHAIN REACTION

DNA analysis often requires focusing on one or more specific regions of the genome. It also frequently involves situations in which only one or a few copies of a DNA molecule are available for further analysis. These amounts are insufficient for most procedures, such as gel electrophoresis. **Polymerase chain reaction (PCR)** is a technique used to rapidly increase the number of copies of one specific region of DNA for further analyses (**Figure 4**). Typically the DNA that is used as the starting sample in a PCR reaction is genomic DNA, which would contain all the genes in the organism. PCR uses a special form of heat tolerant DNA polymerase,

the enzyme that replicates DNA, and other short nucleotide sequences called primers that base pair to a specific portion of the DNA being copied. A PCR reaction does not copy the entire genome, rather it makes millions of copies of one specific region of interest.

- *Taq* DNA polymerase – this polymerase was isolated from *Thermus aquaticus*, which is a bacteria that lives at high temperatures in hot springs and deep sea vents. Taq polymerase has an optimum temperature of 70-80ºC and can survive nearly an hour at 95ºC. Like other DNA polymerases, *Taq* polymerase assembles nucleotides only in the 5' to 3' direction.
- Primers – Just like during DNA replication, Taq polymerase needs a free 3' end to begin synthesis of the new DNA. Primers in a PCR reaction are man-made synthetic segments of DNA that match the ends of the sequence that the scientist is interested in amplifying. In order to synthesize both strands of the DNA double helix, a forward and a reverse primer must be added. These two primers have free 3' ends that point towards the sequence of interest (Figure 4, middle).

PCR is used for many purposes in laboratories. These include: 1) the identification of the owner of a DNA sample left at a crime scene; 2) paternity analysis; 3) the comparison of small amounts of ancient DNA with modern organisms; and 4) determining the sequence of nucleotides in a specific region. In all of these cases, the starting sample is genomic DNA. In some cases, the complete genome may not be present due to the DNA being old or broken down.

PCR involves the following basic steps:

1. Denaturation (94ºC): The sample is heated in a

small plastic tube inside a thermocycler (which is able to rapidly change temperatures very accurately). Heating the sample to 94ºC causes the hydrogen bonds holding the DNA double helix to denature so the strands are separated (Figure 4, top). *Taq* polymerase is heat tolerant and does not denature at this temperature.

2. Annealing (54-60ºC): The sample is cooled so that the primers can anneal (base pair) with their complementary sequence (Figure 4; middle). Because of the base pairing rules, primers can only anneal in the specific spot that contains their complementary base sequence. This allows scientists to choose which region of DNA will be amplified.
3. Extension (72ºC): The sample is warmed up to the optimum temperature for *Taq* polymerase. Taq extends off of the free 3' end of both primers, producing a double helix (Figure 4; bottom).

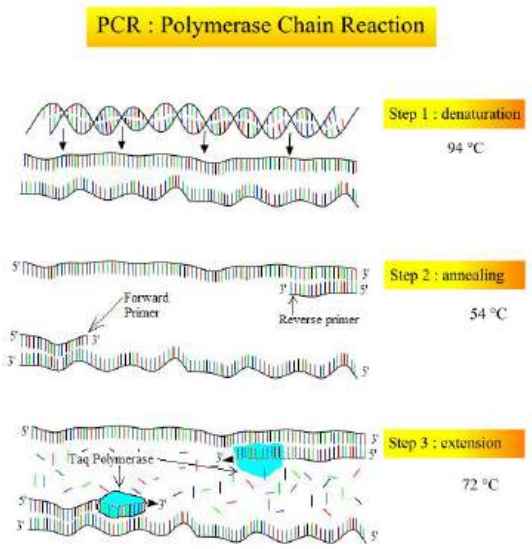


Figure 4 The steps of a PCR reaction. Photo credit Tinojasontran; Wikimedia.

The three basic steps of a PCR reaction are repeated multiple times ("cycles"). Each cycle causes a roughly 2x increase in the number of amplified copies of the specific section of interest (Figure 5). Remember that only one section of DNA is being amplified due to the specificity of the primers.

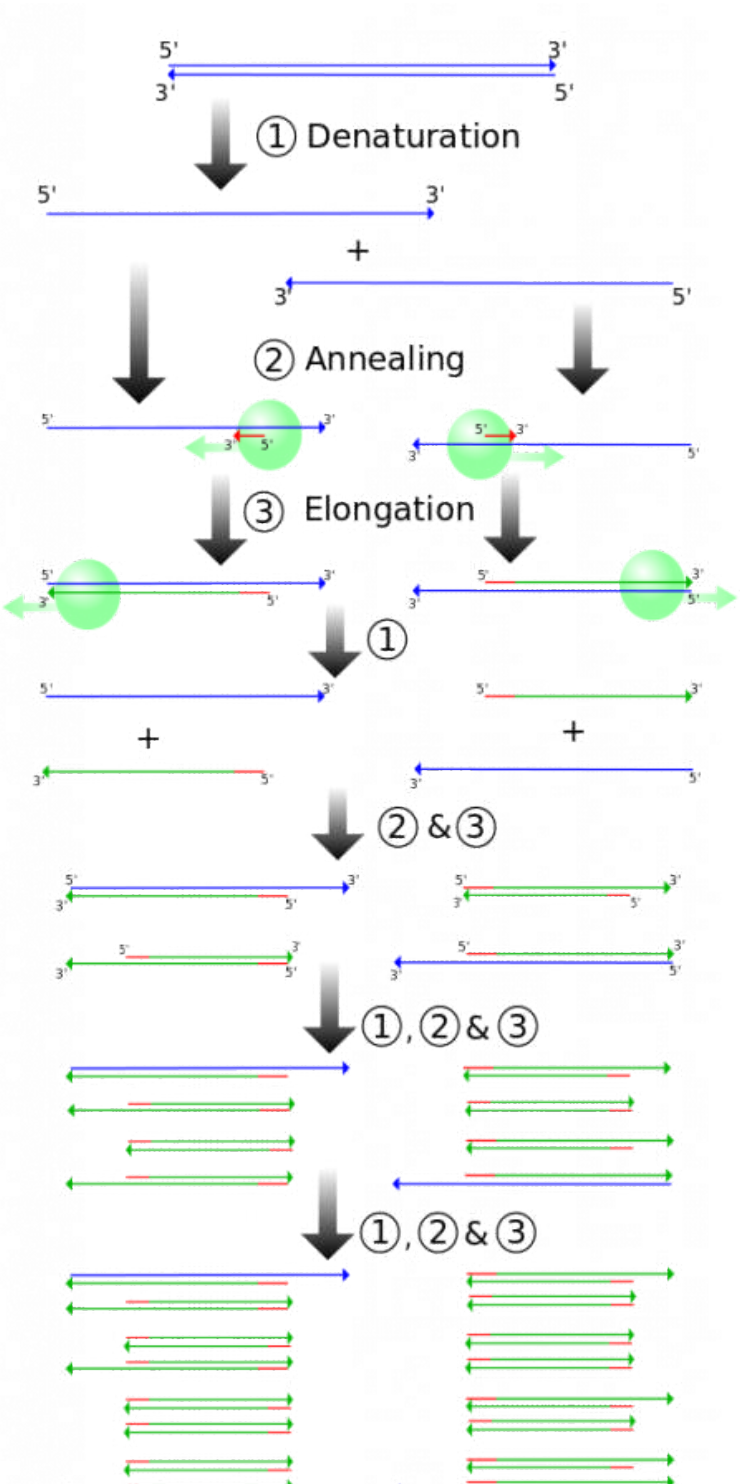

5'
3'
3'
5'
① Denaturation
5'
3'
+
3'
5'
② Annealing
③ Elongation
①
+
+
② & ③
①, ② & ③
①, ② & ③

Figure 5 Repeated cycling causes exponential growth of the desired sequence. Photo credit Madprime; Wikimedia.

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=784#h5p-2

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=784#h5p-3

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=784#h5p-4

For the question below, drag TWO primers to the appropriate

location where they would anneal. The arrowhead shows the 3' end of the primer. Keep in mind that Taq DNA polymerase can only extend from the 3' of the primer.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=784#h5p-10

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8CA_YwJq@3/Cloning-and-Genetic-Engineering

Cloning

In general, **cloning** means the creation of a perfect replica. Typically, the word is used to describe the creation of a genetically identical copy. In biology, the re-creation of a whole organism is referred to as "reproductive cloning." Long before attempts were made to clone an entire organism, researchers learned how to copy short stretches of DNA—a process that is referred to as molecular cloning.

MOLECULAR CLONING

Cloning allows for the creation of multiple copies of genes, expression of genes, and study of specific genes. To get the DNA fragment into a bacterial cell in a form that will be copied or expressed, the fragment is first inserted into a plasmid. A **plasmid** (also called a vector in this context) is a small circular DNA molecule that replicates independently of the chromosomal DNA in bacteria. In cloning, the plasmid molecules can be used to provide a "vehicle" in which to insert a desired DNA fragment. Modified plasmids are usually reintroduced into a bacterial host for replication. As the bacteria divide, they copy their own DNA (including the plasmids). The inserted DNA fragment is copied along with the rest of the bacterial DNA. In a bacterial cell, the fragment of DNA from the human genome (or another organism that is being studied) is referred to as foreign DNA to differentiate it from the DNA of the bacterium (the host DNA).

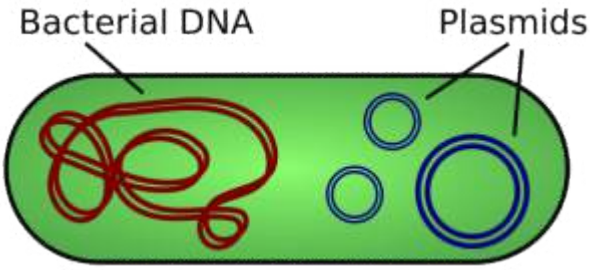


Figure 1 Plasmids occur naturally in bacteria, but can also be modified by scientists.

Plasmids occur naturally in bacterial populations (such as *Escherichia coli*) and have genes that can contribute favorable traits to the organism, such as antibiotic resistance (the ability to be unaffected by antibiotics). Plasmids have been highly engineered as vectors for molecular cloning and for the subsequent large-scale production of important molecules, such as insulin. A valuable characteristic of plasmid vectors is the ease with which a foreign DNA fragment can be introduced. These plasmid vectors contain many short DNA sequences that can be cut with different commonly available **restriction enzymes**. Restriction enzymes (also called restriction endonucleases) recognize specific DNA sequences and cut them in a predictable manner; they are naturally produced by bacteria as a defense mechanism against foreign DNA. Many restriction enzymes make staggered cuts in the two strands of DNA, such that the cut ends have a 2- to 4-nucleotide single-stranded overhang. The sequence that is recognized by the restriction enzyme is a four- to eight-nucleotide sequence that is a palindrome. Like with a word palindrome, this means the sequence reads the same forward and backward. In most cases, the sequence reads the same forward on one strand and backward on the complementary strand. When a

staggered cut is made in a sequence like this, the overhangs are complementary (**Figure 2**).

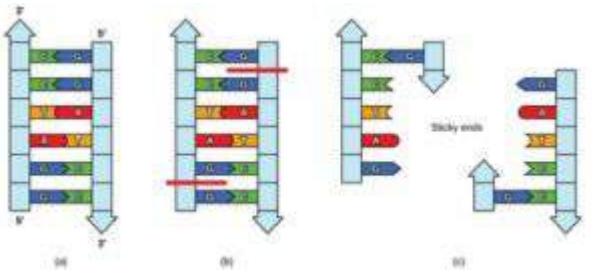


Figure 2 In this (a) six-nucleotide restriction enzyme recognition site, notice that the sequence of six nucleotides reads the same in the 5' to 3' direction on one strand as it does in the 5' to 3' direction on the complementary strand. This is known as a palindrome. (b) The restriction enzyme makes breaks in the DNA strands, and (c) the cut in the DNA results in "sticky ends". Another piece of DNA cut on either end by the same restriction enzyme could attach to these sticky ends and be inserted into the gap made by this cut.

Because these overhangs are capable of coming back together by hydrogen bonding with complementary overhangs on a piece of DNA cut with the same restriction enzyme, these are called "sticky ends." The process of forming hydrogen bonds between complementary sequences on single strands to form doublestranded DNA is called **annealing**. Addition of an enzyme called DNA ligase, which takes part in DNA replication in cells, permanently joins the DNA fragments when the sticky ends come together. In this way, any DNA fragment can be spliced between the two ends of a plasmid DNA that has been cut with the same restriction enzyme (**Figure 3**).

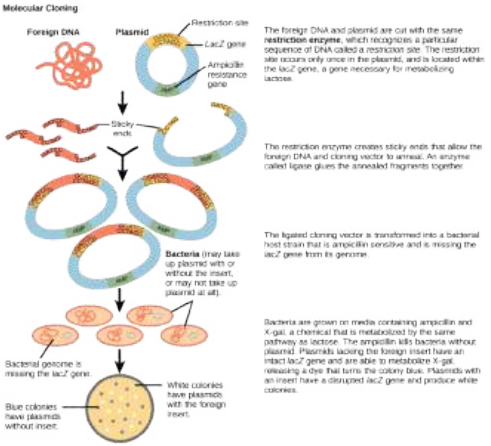


Figure 3: This diagram shows the steps involved in molecular cloning.

Plasmids with foreign DNA inserted into them are called **recombinant DNA** molecules because they contain new combinations of genetic material. Proteins that are produced from recombinant DNA molecules are called **recombinant proteins**. Not all recombinant plasmids are capable of expressing genes. Plasmids may also be engineered to express proteins only when stimulated by certain environmental factors, so that scientists can control the expression of the recombinant proteins.

CELLULAR CLONING

Unicellular organisms, such as bacteria and yeast, naturally produce clones of themselves when they replicate asexually by binary fission; this is known as cellular cloning. The nuclear DNA duplicates by the process of mitosis, which creates an exact replica of the genetic material.

REPRODUCTIVE CLONING

Reproductive cloning is a method used to make a clone or an identical copy of an entire multicellular organism. Most multicellular organisms undergo reproduction by sexual means, which involves the contribution of DNA from two individuals (parents), making it impossible to generate an identical copy or a clone of either parent. Recent advances in biotechnology have made it possible to reproductively clone mammals in the laboratory.

Natural sexual reproduction involves the union, during fertilization, of a sperm and an egg. Each of these gametes is haploid, meaning they contain one set of chromosomes in their nuclei. The resulting cell, or zygote, is then diploid and contains two sets of chromosomes. This cell divides mitotically to produce a multicellular organism. However, the union of just any two cells cannot produce a viable zygote; there are components in the cytoplasm of the egg cell that are essential for the early development of the embryo during its first few cell divisions. Without these provisions, there would be no subsequent development. Therefore, to produce a new individual, both a diploid genetic complement and an egg cytoplasm are required. The approach to producing an artificially cloned individual is to take the egg cell of one individual and to remove the haploid nucleus (**Figure 4**). Then a diploid nucleus from a body cell of a second individual, the donor, is put into the egg cell. The egg is then stimulated to divide so that development proceeds. This sounds simple, but in fact it takes many attempts before each of the steps is completed successfully.

The first cloned agricultural animal was Dolly, a sheep who was born in 1996. The success rate of reproductive cloning at the time was very low. Dolly lived for six years and died of a lung tumor. There was speculation that because the cell DNA that gave rise to Dolly came from an older individual, the

age of the DNA may have affected her life expectancy. Since Dolly, several species of animals (such as horses, bulls, and goats) have been successfully cloned.

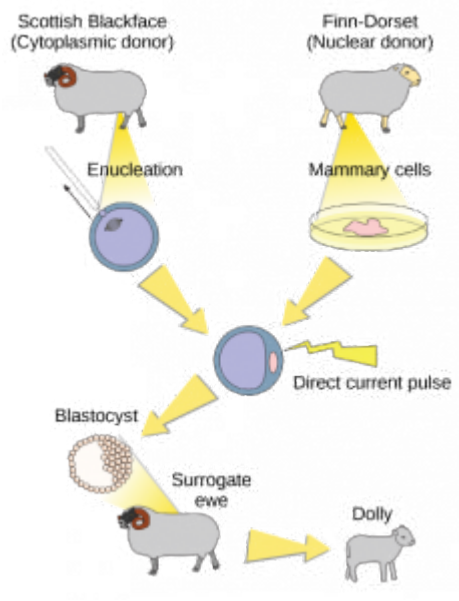


Figure 4 The creation of Dolly, the cloned sheep.

There have been attempts at producing cloned human embryos as sources of embryonic stem cells. In the procedure, the DNA from an adult human is introduced into a human egg cell, which is then stimulated to divide. The technology is similar to the technology that was used to produce Dolly, but the embryo is never implanted into a surrogate mother. The cells produced are called embryonic stem cells because they have the capacity to develop into many different kinds of cells, such as muscle or nerve cells. The stem cells could be used to research and ultimately provide therapeutic applications, such as replacing damaged tissues. The benefit of cloning in this instance is that the cells used to regenerate new tissues would be a perfect match

to the donor of the original DNA. For example, a leukemia patient would not require a sibling with a tissue match for a bone-marrow transplant.

An interactive H5P element has been excluded from this version of the text. You can view it online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=788#h5p-5

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8CA_YwJq@3/Cloning-and-Genetic-Engineerin

Genetic Engineering

Genetic engineering is the alteration of an organism's genotype using recombinant DNA technology to modify an organism's DNA to achieve desirable traits. The addition of foreign DNA in the form of recombinant DNA vectors generated by molecular cloning is the most common method of genetic engineering. The organism that receives the recombinant DNA is called a **genetically modified organism** (GMO). If the foreign DNA that is introduced comes from a different species, the host organism is called **transgenic.** Bacteria, plants, and animals have been genetically modified since the early 1970s for academic, medical, agricultural, and industrial purposes. In the US, GMOs such as Roundup-ready soybeans and borer-resistant corn are part of many common processed foods.

GENE TARGETING

Although classical methods of studying the function of genes began with a given phenotype and determined the genetic basis of that phenotype, modern techniques allow researchers to start at the DNA sequence level and ask: "What does this gene or DNA element do?" This technique, called reverse genetics, has resulted in reversing the classic genetic methodology. This method would be similar to damaging a body part to determine its function. An insect that loses a wing cannot fly, which means that the function of the wing is flight. The classical genetic method would compare insects that cannot fly with insects that can fly, and

observe that the non-flying insects have lost wings. Similarly, mutating or deleting genes provides researchers with clues about gene function. The methods used to disable gene function are collectively called gene targeting. Gene targeting is the use of recombinant DNA vectors to alter the expression of a particular gene, either by introducing mutations in a gene, or by eliminating the expression of a certain gene by deleting a part or all of the gene sequence from the genome of an organism.

GENETIC DIAGNOSIS AND GENE THERAPY

The process of testing for suspected genetic defects before administering treatment is called genetic diagnosis by genetic testing. Depending on the inheritance patterns of a disease-causing gene, family members are advised to undergo genetic testing. For example, women diagnosed with breast cancer are usually advised to have a biopsy so that the medical team can determine the genetic basis of cancer development. Treatment plans are based on the findings of genetic tests that determine the type of cancer. If the cancer is caused by inherited gene mutations, other female relatives are also advised to undergo genetic testing and periodic screening for breast cancer. Genetic testing is also offered for fetuses (or embryos with in vitro fertilization) to determine the presence or absence of disease-causing genes in families with specific debilitating diseases.

Gene therapy is a genetic engineering technique used to cure disease. In its simplest form, it involves the introduction of a good gene at a random location in the genome to aid the cure of a disease that is caused by a mutated gene. The good gene is usually introduced into diseased cells as part of a vector transmitted by a virus that can infect the host cell and deliver the foreign DNA (Figure 1). More advanced forms

of gene therapy try to correct the mutation at the original site in the genome, such as is the case with treatment of severe combined immunodeficiency (SCID).

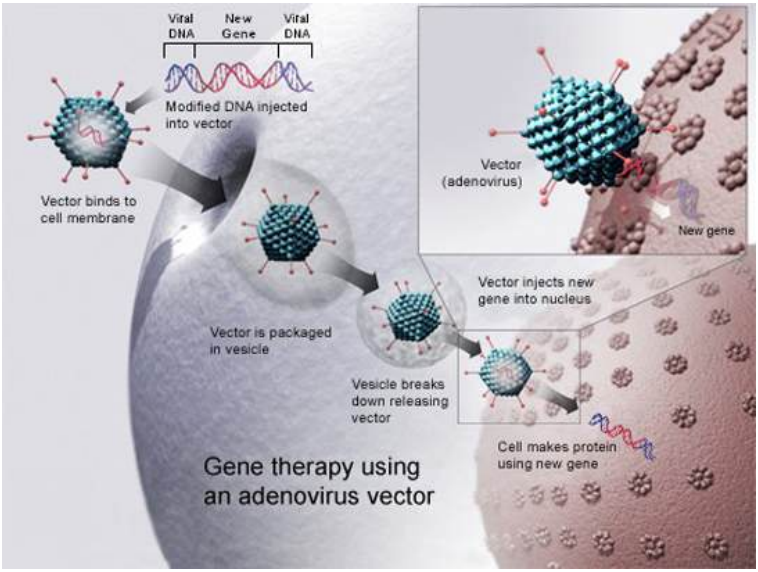


Figure 1 Gene therapy using an adenovirus vector can be used to cure certain genetic diseases in which a person has a defective gene. (credit: NIH)

PRODUCTION OF VACCINES, ANTIBIOTICS, AND HORMONES

Traditional vaccination strategies use weakened or inactive forms of microorganisms to mount the initial immune response. Modern techniques use the genes of microorganisms cloned into vectors to mass produce the desired antigen. The antigen is then introduced into the body to stimulate the primary immune response and trigger immune memory. Genes cloned from the influenza virus

have been used to combat the constantly changing strains of this virus.

Antibiotics are a biotechnological product. They are naturally produced by microorganisms, such as fungi, to attain an advantage over bacterial populations. Antibiotics are produced on a large scale by cultivating and manipulating fungal cells.

Recombinant DNA technology was used to produce large-scale quantities of human insulin in *E. coli* as early as 1978. Previously, it was only possible to treat diabetes with pig insulin, which caused allergic reactions in humans because of differences in the gene product. Currently, the vast majority of diabetes suffers who inject insulin do so with insulin produced by bacteria.

Human growth hormone (HGH) is used to treat growth disorders in children. The HGH gene was cloned from a cDNA library and inserted into *E. coli* cells by cloning it into a bacterial vector. Bacterial HGH can be used in humans to reduce symptoms of various growth disorders.

TRANSGENIC ANIMALS

Although several recombinant proteins used in medicine are successfully produced in bacteria, some proteins require a eukaryotic animal host for proper processing. For this reason, the desired genes are cloned and expressed in animals, such as sheep, goats, chickens, and mice. Animals that have been modified to express recombinant DNA are called transgenic animals. Several human proteins are expressed in the milk of transgenic sheep and goats, and some are expressed in the eggs of chickens. Mice have been used extensively for expressing and studying the effects of recombinant genes and mutations.

Figure 2 These mice have been engineered to contain the green fluorescent protein (GFP) gene, which was originally isolated from jellyfish. Photo credit Moen, 2012.

TRANSGENIC PLANTS

Manipulating the DNA of plants (i.e., creating GMOs) has helped to create desirable traits, such as disease resistance, herbicide and pesticide resistance, better nutritional value, and better shelf-life (Figure 3). Plants are the most important source of food for the human population. Farmers developed ways to select for plant varieties with desirable traits long before modern-day biotechnology practices were established.

Figure 3 Corn, a major agricultural crop used to create products for a variety of industries, is often modified through plant biotechnology. (credit: Keith Weller, USDA)

Plants that have received recombinant DNA from other species are called transgenic plants. Because they are not natural, transgenic plants and other GMOs are closely monitored by government agencies to ensure that they are fit for human consumption and do not endanger other plant and animal life. Because foreign genes can spread to other species in the environment, extensive testing is required to ensure ecological stability. Staples like corn, potatoes, and tomatoes were the first crop plants to be genetically engineered.

TRANSFORMATION OF PLANTS USING *AGROBACTERIUM TUMEFACIENS*

Gene transfer occurs naturally between species in microbial

populations. Many viruses that cause human diseases, such as cancer, act by incorporating their DNA into the human genome. In plants, tumors caused by the bacterium *Agrobacterium tumefaciens* occur by transfer of DNA from the bacterium to the plant. Although the tumors do not kill the plants, they make the plants stunted and more susceptible to harsh environmental conditions. Many plants, such as walnuts, grapes, nut trees, and beets, are affected by *A. tumefaciens*. The artificial introduction of DNA into plant cells is more challenging than in animal cells because of the thick plant cell wall.

Figure 4 Tumor caused by infection with Agrobacterium tumefaciens. Photo credit C-M; Wikimedia.

Researchers used the natural transfer of DNA from *Agrobacterium* to a plant host to introduce DNA fragments of their choice into plant hosts. In nature, the disease-causing *A. tumefaciens* have a set of plasmids, called the Ti plasmids (tumor-inducing plasmids), that contain genes for the production of tumors in plants. DNA from the Ti plasmid integrates into the infected plant cell's genome. Researchers manipulate the Ti plasmids to remove the tumor-causing genes and insert the desired DNA fragment for transfer into

the plant genome. The Ti plasmids carry antibiotic resistance genes to aid selection and can be propagated in *E. coli* cells as well.

THE ORGANIC INSECTICIDE *BACILLUS THURINGIENSIS*

Bacillus thuringiensis (Bt) is a bacterium that produces protein crystals during sporulation that are toxic to many insect species that affect plants. Bt toxin has to be ingested by insects for the toxin to be activated. Insects that have eaten Bt toxin stop feeding on the plants within a few hours. After the toxin is activated in the intestines of the insects, death occurs within a couple of days. Modern biotechnology has allowed plants to encode their own crystal Bt toxin that acts against insects. The crystal toxin genes have been cloned from Bt and introduced into plants. Bt toxin has been found to be safe for the environment, non-toxic to humans and other mammals, and is approved for use by organic farmers as a natural insecticide.

FLAVR SAVR TOMATO

The first GM crop to be introduced into the market was the Flavr Savr Tomato produced in 1994. Antisense RNA technology was used to slow down the process of softening and rotting caused by fungal infections, which led to increased shelf life of the GM tomatoes. Additional genetic modification improved the flavor of this tomato. The Flavr Savr tomato did not successfully stay in the market because of problems maintaining and shipping the crop.

An interactive H5P element has been excluded from this version of the text. You can view it online here:
https://openoregon.pressbooks.pub/mhccmajorsbio/?p=789#h5p-6

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8CA_YwJq@3/Cloning-and-Genetic-Engineerin

Moen I, Jevne C, Kalland K-H, Chekenya M, Akslen LA, Sleire L, Enger P, Reed RK, Oyan AM, Stuhr LEB. 2012. Gene expression in tumor cells and stroma in dsRed 4T1 tumors in eGFP-expressing mice with and without enhanced oxygenation. BMC Cancer. 12:21. doi:10.1186/1471-2407-12-21 PDF

Mapping Genomes

Genomics is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. Genome mapping is the process of finding the locations of genes on each chromosome. The maps created by genome mapping are comparable to the maps that we use to navigate streets. A genetic map is an illustration that lists genes and their location on a chromosome. **Genetic maps** provide the big picture (similar to a map of interstate highways) and use genetic markers (similar to landmarks). A genetic marker is a gene or sequence on a chromosome that co-segregates (shows genetic linkage) with a specific trait. Early geneticists called this linkage analysis. **Physical maps** present the intimate details of smaller regions of the chromosomes (similar to a detailed road map). A physical map is a representation of the physical distance, in nucleotides, between genes or genetic markers. Both genetic linkage maps and physical maps are required to build a complete picture of the genome. Having a complete map of the genome makes it easier for researchers to study individual genes. Human genome maps help researchers in their efforts to identify human disease-causing genes related to illnesses like cancer, heart disease, and cystic fibrosis. Genome mapping can be used in a variety of other applications, such as using live microbes to clean up pollutants or even prevent pollution. Research involving plant genome mapping may lead to producing higher crop yields or developing plants that better adapt to climate change.

GENETIC MAPS

The study of genetic maps begins with linkage analysis, a procedure that analyzes the recombination frequency

between genes to determine if they are linked or show independent assortment. The term *linkage* was used before the discovery of DNA. Early geneticists relied on the observation of phenotypic changes to understand the genotype of an organism. Shortly after Gregor Mendel (the father of modern genetics) proposed that traits were determined by what are now known as genes, other researchers observed that different traits were often inherited together, and thereby deduced that the genes were physically linked by being located on the same chromosome. The mapping of genes relative to each other based on linkage analysis led to the development of the first genetic maps.

Observations that certain traits were always linked and certain others were not linked came from studying the offspring of crosses between parents with different traits. For example, in experiments performed on the garden pea, it was discovered that the color of the flower and shape of the plant's pollen were linked traits, and therefore the genes encoding these traits were in close proximity on the same chromosome. The exchange of DNA between homologous pairs of chromosomes is called genetic recombination, which occurs by the crossing over of DNA between homologous strands of DNA, such as nonsister chromatids. Linkage analysis involves studying the recombination frequency between any two genes. The greater the distance between two genes, the higher the chance that a recombination event will occur between them, and the higher the recombination frequency between them. Two possibilities for recombination between two nonsister chromatids during meiosis are shown in Figure 1. If the recombination frequency between two genes is less than 50 percent, they are said to be linked.

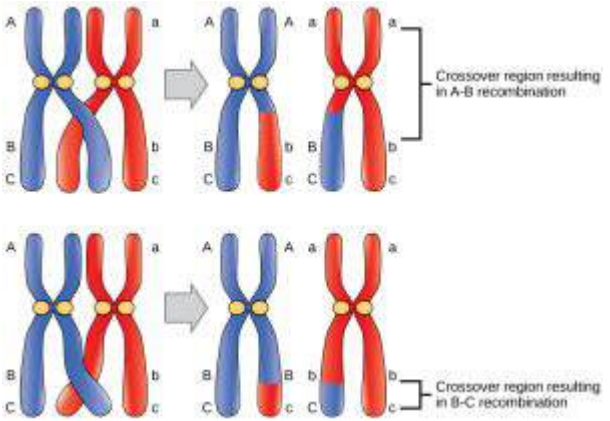


Figure 1 Crossover may occur at different locations on the chromosome. Recombination between genes A and B is more frequent than recombination between genes B and C because genes A and B are farther apart; a crossover is therefore more likely to occur between them.

The generation of genetic maps requires markers, just as a road map requires landmarks (such as rivers and mountains). Early genetic maps were based on the use of known genes as markers. More sophisticated markers, including those based on non-coding DNA, are now used to compare the genomes of individuals in a population. Although individuals of a given species are genetically similar, they are not identical; every individual has a unique set of traits. These minor differences in the genome between individuals in a population are useful for the purposes of genetic mapping. In general, a good genetic marker is a region on the chromosome that shows variability or polymorphism (multiple forms) in the population.

Some genetic markers used in generating genetic maps are restriction fragment length polymorphisms (RFLP), variable number of tandem repeats (VNTRs), microsatellite polymorphisms, and the single nucleotide polymorphisms (SNPs). RFLPs (sometimes pronounced “rif-lips”) are detected

when the DNA of an individual is cut with a restriction endonuclease that recognizes specific sequences in the DNA to generate a series of DNA fragments, which are then analyzed by gel electrophoresis. The DNA of every individual will give rise to a unique pattern of bands when cut with a particular set of restriction endonucleases; this is sometimes referred to as an individual's DNA "fingerprint." Certain regions of the chromosome that are subject to polymorphism will lead to the generation of the unique banding pattern. VNTRs are repeated sets of nucleotides present in the non-coding regions of DNA. Non-coding, or "junk," DNA has no known biological function; however, research shows that much of this DNA is actually transcribed. While its function is uncertain, it is certainly active, and it may be involved in the regulation of coding genes. The number of repeats may vary in individual organisms of a population. Microsatellite polymorphisms are similar to VNTRs, but the repeat unit is very small. SNPs are variations in a single nucleotide.

Because genetic maps rely completely on the natural process of recombination, mapping is affected by natural increases or decreases in the level of recombination in any given area of the genome. Some parts of the genome are recombination hotspots, whereas others do not show a propensity for recombination. For this reason, it is important to look at mapping information developed by multiple methods.

PHYSICAL MAPS

A physical map provides detail of the actual physical distance between genetic markers, as well as the number of nucleotides. There are three methods used to create a physical map: cytogenetic mapping, radiation hybrid

mapping, and sequence mapping. Cytogenetic mapping uses information obtained by microscopic analysis of stained sections of the chromosome (Figure 2). It is possible to determine the approximate distance between genetic markers using cytogenetic mapping, but not the exact distance (number of base pairs). Radiation hybrid mapping uses radiation, such as x-rays, to break the DNA into fragments. The amount of radiation can be adjusted to create smaller or larger fragments. This technique overcomes the limitation of genetic mapping and is not affected by increased or decreased recombination frequency. Sequence mapping resulted from DNA sequencing technology that allowed for the creation of detailed physical maps with distances measured in terms of the number of base pairs. The creation of genomic libraries and complementary DNA (cDNA) libraries (collections of cloned sequences or all DNA from a genome) has sped up the process of physical mapping. A genetic site used to generate a physical map with sequencing technology (a sequence-tagged site, or STS) is a unique sequence in the genome with a known exact chromosomal location. An expressed sequence tag (EST) and a single sequence length polymorphism (SSLP) are common STSs. An EST is a short STS that is identified with cDNA libraries, while SSLPs are obtained from known genetic markers and provide a link between genetic maps and physical maps.

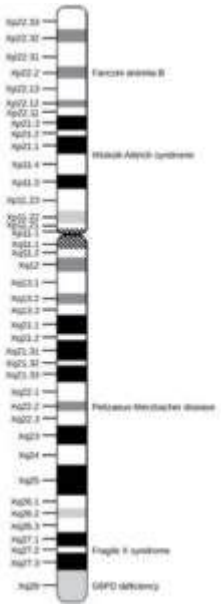


Figure 1: This is a physical map of the human X chromosome. (credit: modification of work by NCBI, NIH)

INTEGRATION OF GENETIC AND PHYSICAL MAPS

Genetic maps provide the outline and physical maps provide the details. It is easy to understand why both types of genome mapping techniques are important to show the big picture. Information obtained from each technique is used in combination to study the genome. Genomic mapping is being used with different model organisms that are used for research. Genome mapping is still an ongoing process, and as more advanced techniques are developed, more advances are expected. Genome mapping is similar to completing a complicated puzzle using every piece of available data. Mapping information generated in laboratories all over the

world is entered into central databases, such as GenBank at the National Center for Biotechnology Information (NCBI). Efforts are being made to make the information more easily accessible to researchers and the general public. Just as we use global positioning systems instead of paper maps to navigate through roadways, NCBI has created a genome viewer tool to simplify the data-mining process.

REFERENCES

OpenStax, Biology. OpenStax CNX. January 2, 2017 https://cnx.org/contents/GFy_h8cu@10.120:Qq6Y1A16@5/Mapping-Genomes

Applying Genomics

The introduction of DNA sequencing and whole genome sequencing projects, particularly the Human Genome Project, has expanded the applicability of DNA sequence information. Genomics is now being used in a wide variety of fields, such as metagenomics, pharmacogenomics, and mitochondrial genomics. The most commonly known application of genomics is to understand and find cures for diseases.

PREDICTING DISEASE RISK AT THE INDIVIDUAL LEVEL

Predicting the risk of disease involves screening and identifying currently healthy individuals by genome analysis at the individual level. Intervention with lifestyle changes and drugs can be recommended before disease onset. However, this approach is most applicable when the problem arises from a single gene mutation. Such defects only account for about 5 percent of diseases found in developed countries. Most of the common diseases, such as heart disease, are multifactorial or polygenic, which refers to a phenotypic characteristic that is determined by two or more genes, and also environmental factors such as diet. In April 2010, scientists at Stanford University published the genome analysis of a healthy individual (Stephen Quake, a scientist at Stanford University, who had his genome sequenced); the analysis predicted his propensity to acquire various diseases. A risk assessment was done to analyze Quake's percentage of risk for 55 different medical conditions. A rare genetic mutation was found that showed him to be at risk for sudden heart attack. He was also predicted to have a 23 percent risk of developing prostate cancer and a 1.4 percent risk of developing Alzheimer's disease. The scientists used databases and several publications to analyze the genomic

data. Even though genomic sequencing is becoming more affordable and analytical tools are becoming more reliable, ethical issues surrounding genomic analysis at a population level remain to be addressed. For example, could such data be legitimately used to charge more or less for insurance or to affect credit ratings?

GENOME-WIDE ASSOCIATION STUDIES

Since 2005, it has been possible to conduct a type of study called a genome-wide association study, or GWAS. A GWAS is a method that identifies differences between individuals in single nucleotide polymorphisms (SNPs) that may be involved in causing diseases. The method is particularly suited to diseases that may be affected by one or many genetic changes throughout the genome. It is very difficult to identify the genes involved in such a disease using family history information. The GWAS method relies on a genetic database that has been in development since 2002 called the International HapMap Project. The HapMap Project sequenced the genomes of several hundred individuals from around the world and identified groups of SNPs. The groups include SNPs that are located near to each other on chromosomes so they tend to stay together through recombination. The fact that the group stays together means that identifying one marker SNP is all that is needed to identify all the SNPs in the group. There are several million SNPs identified, but identifying them in other individuals who have not had their complete genome sequenced is much easier because only the marker SNPs need to be identified.

In a common design for a GWAS, two groups of individuals are chosen; one group has the disease, and the other group does not. The individuals in each group are matched in other characteristics to reduce the effect of confounding variables

causing differences between the two groups. For example, the genotypes may differ because the two groups are mostly taken from different parts of the world. Once the individuals are chosen, and typically their numbers are a thousand or more for the study to work, samples of their DNA are obtained. The DNA is analyzed using automated systems to identify large differences in the percentage of particular SNPs between the two groups. Often the study examines a million or more SNPs in the DNA. The results of GWAS can be used in two ways: the genetic differences may be used as markers for susceptibility to the disease in undiagnosed individuals, and the particular genes identified can be targets for research into the molecular pathway of the disease and potential therapies. An offshoot of the discovery of gene associations with disease has been the formation of companies that provide socalled "personal genomics" that will identify risk levels for various diseases based on an individual's SNP complement. The science behind these services is controversial.

Because GWAS looks for associations between genes and disease, these studies provide data for other research into causes, rather than answering specific questions themselves. An association between a gene difference and a disease does not necessarily mean there is a cause-and-effect relationship. However, some studies have provided useful information about the genetic causes of diseases. For example, three different studies in 2005 identified a gene for a protein involved in regulating inflammation in the body that is associated with a disease-causing blindness called age-related macular degeneration. This opened up new possibilities for research into the cause of this disease. A large number of genes have been identified to be associated with Crohn's disease using GWAS, and some of these have suggested new hypothetical mechanisms for the cause of the disease.

PHARMACOGENOMICS

Pharmacogenomics involves evaluating the effectiveness and safety of drugs on the basis of information from an individual's genomic sequence. Personal genome sequence information can be used to prescribe medications that will be most effective and least toxic on the basis of the individual patient's genotype. Studying changes in gene expression could provide information about the gene transcription profile in the presence of the drug, which can be used as an early indicator of the potential for toxic effects. For example, genes involved in cellular growth and controlled cell death, when disturbed, could lead to the growth of cancerous cells. Genome-wide studies can also help to find new genes involved in drug toxicity. The gene signatures may not be completely accurate, but can be tested further before pathologic symptoms arise.

METAGENOMICS

Traditionally, microbiology has been taught with the view that microorganisms are best studied under pure culture conditions, which involves isolating a single type of cell and culturing it in the laboratory. Because microorganisms can go through several generations in a matter of hours, their gene expression profiles adapt to the new laboratory environment very quickly. On the other hand, many species resist being cultured in isolation. Most microorganisms do not live as isolated entities, but in microbial communities known as biofilms. For all of these reasons, pure culture is not always the best way to study microorganisms. **Metagenomics** is the study of the collective genomes of multiple species that grow and interact in an environmental niche. Metagenomics can be used to identify new species more rapidly and to

analyze the effect of pollutants on the environment (**Figure 1**). Metagenomics techniques can now also be applied to communities of higher eukaryotes, such as fish.

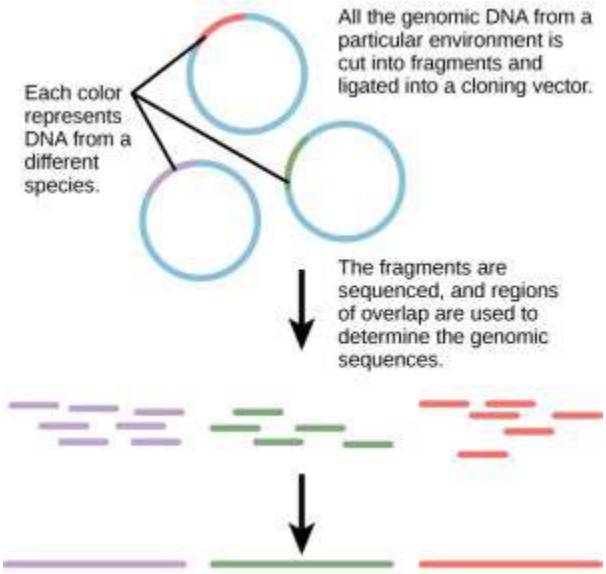


Figure 1: Metagenomics involves isolating DNA from multiple species within an environmental niche. The DNA is cut up and sequenced, allowing entire genome sequences of multiple species to be reconstructed from the sequences of overlapping pieces.

CREATION OF NEW BIOFUELS

Knowledge of the genomics of microorganisms is being used to find better ways to harness biofuels from algae and cyanobacteria. The primary sources of fuel today are coal, oil, wood, and other plant products such as ethanol. Although plants are renewable resources, there is still a need to find more alternative renewable sources of energy to meet our population's energy demands. The microbial world is one of

the largest resources for genes that encode new enzymes and produce new organic compounds, and it remains largely untapped. This vast genetic resource holds the potential to provide new sources of biofuels (**Figure 2**).

Figure 2: Renewable fuels were tested in Navy ships and aircraft at the first Naval Energy Forum. (credit: modification of work by John F Williams, US Navy)

MITOCHONDRIAL GENOMICS

Mitochondria are intracellular organelles that contain their own DNA. Mitochondrial DNA mutates at a rapid rate and is often used to study evolutionary relationships. Another feature that makes studying the mitochondrial genome interesting is that in most multicellular organisms, the mitochondrial DNA is passed on from the mother during the process of fertilization. For this reason, mitochondrial genomics is often used to trace genealogy.

GENOMICS IN FORENSIC ANALYSIS

Information and clues obtained from DNA samples found at crime scenes have been used as evidence in court cases, and genetic markers have been used in forensic analysis. Genomic analysis has also become useful in this field. In 2001, the first use of genomics in forensics was published. It was a collaborative effort between academic research institutions and the FBI to solve the mysterious cases of anthrax (**Figure 10.15**) that was transported by the US Postal Service. Anthrax bacteria were made into an infectious powder and mailed to news media and two U.S. Senators. The powder infected the administrative staff and postal workers who opened or handled the letters. Five people died, and 17 were sickened from the bacteria. Using microbial genomics, researchers determined that a specific strain of anthrax was used in all the mailings; eventually, the source was traced to a scientist at a national biodefense laboratory in Maryland.

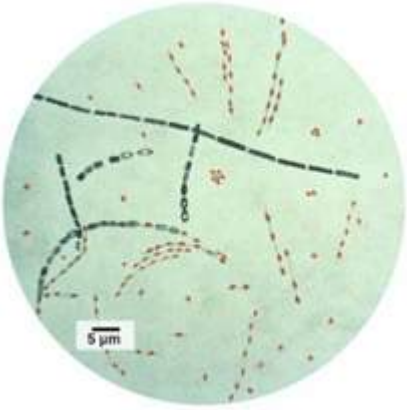


Figure 3: Bacillus anthracis is the organism that causes anthrax. (credit: modification of work by CDC; scale-bar data from Matt Russell)

GENOMICS IN AGRICULTURE

Genomics can reduce the trials and failures involved in scientific research to a certain extent, which could improve the quality and quantity of crop yields in agriculture (**Figure 10.16**). Linking traits to genes or gene signatures helps to improve crop breeding to generate hybrids with the most desirable qualities. Scientists use genomic data to identify desirable traits, and then transfer those traits to a different organism to create a new genetically modified organism, as described in the previous module. Scientists are discovering how genomics can improve the quality and quantity of agricultural production. For example, scientists could use desirable traits to create a useful product or enhance an existing product, such as making a drought-sensitive crop more tolerant of the dry season.

Figure 4: Transgenic agricultural plants can be made to resist disease. These transgenic plums are resistant to the plum pox virus. (credit: Scott Bauer, USDA ARS)

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TE1njgbY@4/Genomics-and-Proteomics

Proteomics

Proteins are the final products of genes that perform the function encoded by the gene. Proteins are composed of amino acids and play important roles in the cell. All enzymes (except ribozymes) are proteins and act as catalysts that affect the rate of reactions. Proteins are also regulatory molecules, and some are hormones. Transport proteins, such as hemoglobin, help transport oxygen to various organs. Antibodies that defend against foreign particles are also proteins. In the diseased state, protein function can be impaired because of changes at the genetic level or because of direct impact on a specific protein.

A proteome is the entire set of proteins produced by a cell type. Proteomes can be studied using the knowledge of genomes because genes code for mRNAs, and the mRNAs encode proteins. The study of the function of proteomes is called **proteomics**. Proteomics complements genomics and is useful when scientists want to test their hypotheses that were based on genes. Even though all cells in a multicellular organism have the same set of genes, the set of proteins produced in different tissues is different and dependent on gene expression. Thus, the genome is constant, but the proteome varies and is dynamic within an organism. In addition, RNAs can be alternatively spliced (cut and pasted to create novel combinations and novel proteins), and many proteins are modified after translation. Although the genome provides a blueprint, the final architecture depends on several factors that can change the progression of events that generate the proteome.

Genomes and proteomes of patients suffering from specific diseases are being studied to understand the genetic basis of the disease. The most prominent disease being studied with proteomic approaches is cancer (**Figure 1**). Proteomic approaches are being used to improve the screening and early detection of cancer; this is achieved by identifying proteins whose expression is affected by the

disease process. An individual protein is called a **biomarker**, whereas a set of proteins with altered expression levels is called a **protein signature**. For a biomarker or protein signature to be useful as a candidate for early screening and detection of a cancer, it must be secreted in body fluids such as sweat, blood, or urine, so that large-scale screenings can be performed in a noninvasive fashion. The current problem with using biomarkers for the early detection of cancer is the high rate of false-negative results. A false-negative result is a negative test result that should have been positive. In other words, many cases of cancer go undetected, which makes biomarkers unreliable. Some examples of protein biomarkers used in cancer detection are CA-125 for ovarian cancer and PSA for prostate cancer. Protein signatures may be more reliable than biomarkers to detect cancer cells. Proteomics is also being used to develop individualized treatment plans, which involves the prediction of whether or not an individual will respond to specific drugs and the side effects that the individual may have. Proteomics is also being used to predict the possibility of disease recurrence.

Figure 1: This machine is preparing to do a proteomic pattern analysis to identify specific cancers so that an accurate cancer prognosis can be made. (credit: Dorie Hightower, NCI, NIH)

The National Cancer Institute has developed programs to improve the detection and treatment of cancer. The Clinical Proteomic Technologies for Cancer and the Early Detection Research Network are efforts to identify protein signatures specific to different types of cancers. The Biomedical Proteomics Program is designed to identify protein signatures and design effective therapies for cancer patients.

REFERENCES

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TE1njgbY@4/Genomics-and-Proteomics

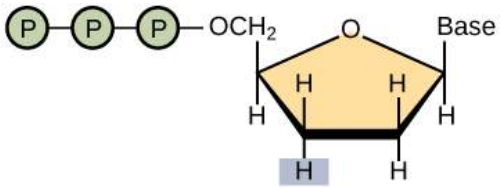
Whole Genome Sequencing

Although there have been significant advances in the medical sciences in recent years, doctors are still confounded by some diseases, and they are using whole-genome sequencing to get to the bottom of the problem. Whole-genome sequencing is a process that determines the DNA sequence of an entire genome. Whole-genome sequencing is a brute-force approach to problem solving when there is a genetic basis at the core of a disease. Several laboratories now provide services to sequence, analyze, and interpret entire genomes.

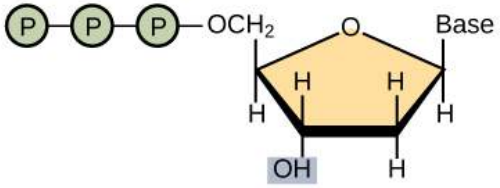
For example, whole-exome sequencing is a lower-cost alternative to whole genome sequencing. In exome sequencing, only the coding, exon-producing regions of the DNA are sequenced. In 2010, whole-exome sequencing was used to save a young boy whose intestines had multiple mysterious abscesses. The child had several colon operations with no relief. Finally, whole-exome sequencing was performed, which revealed a defect in a pathway that controls apoptosis (programmed cell death). A bone-marrow transplant was used to overcome this genetic disorder, leading to a cure for the boy. He was the first person to be successfully treated based on a diagnosis made by whole-exome sequencing. Today, human genome sequencing is more readily available and can be completed in a day or two for about $1000.

STRATEGIES USED IN SEQUENCING PROJECTS

The basic sequencing technique used in all modern day sequencing projects is the chain termination method (also known as the dideoxy method), which was developed by Fred Sanger in the 1970s. The chain termination method involves DNA replication of a single-stranded template with the use of a primer and a regular deoxynucleotide (dNTP), which is a monomer, or a single unit, of DNA. The primer and dNTP are mixed with a small proportion of fluorescently labeled dideoxynucleotides (ddNTPs). The ddNTPs are monomers that are missing a hydroxyl group (–OH) at the site at which another nucleotide usually attaches to form a chain (Figure 1).



Dideoxynucleotide (ddNTP)



Deoxynucleotide (dNTP)

Figure 1 A dideoxynucleotide is similar in structure to a deoxynucleotide, but is missing the 3' hydroxyl group (indicated by the box). When a dideoxynucleotide is incorporated into a DNA strand, DNA synthesis stops.

Each ddNTP is labeled with a different color of fluorophore. Every time a ddNTP is incorporated in the growing complementary strand, it terminates the process of DNA replication, which results in multiple short strands of replicated DNA that are each terminated at a different point during replication. When the reaction mixture is processed by gel electrophoresis after being separated into single strands, the multiple newly replicated DNA strands form a ladder because of the differing sizes. Because the ddNTPs are fluorescently labeled, each band on the gel reflects the size of the DNA strand and the ddNTP that terminated the reaction. The different colors of the fluorophore-labeled ddNTPs help identify the ddNTP incorporated at that position. Reading the gel on the basis of the color of each

band on the ladder produces the sequence of the template strand (figure 2)

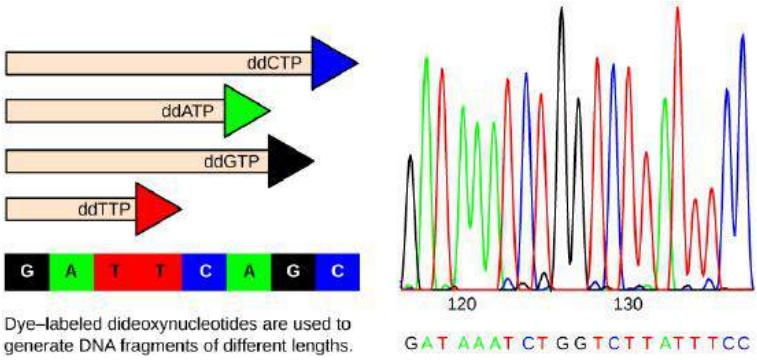


Figure 2: Frederick Sanger's dideoxy chain termination method is illustrated. Using dideoxynucleotides, the DNA fragment can be terminated at different points. The DNA is separated on the basis of size, and these bands, based on the size of the fragments, can be read.

EARLY STRATEGIES: SHOTGUN SEQUENCING AND PAIR-WISE END SEQUENCING

In shotgun sequencing method, several copies of a DNA fragment are cut randomly into many smaller pieces (somewhat like what happens to a round shot cartridge when fired from a shotgun). All of the segments are then sequenced using the chain-sequencing method. Then, with the help of a computer, the fragments are analyzed to see where their sequences overlap. By matching up overlapping sequences at the end of each fragment, the entire DNA sequence can be reformed. A larger sequence that is assembled from overlapping shorter sequences is called a

contig. As an analogy, consider that someone has four copies of a landscape photograph that you have never seen before and know nothing about how it should appear. The person then rips up each photograph with their hands, so that different size pieces are present from each copy. The person then mixes all of the pieces together and asks you to reconstruct the photograph. In one of the smaller pieces you see a mountain. In a larger piece, you see that the same mountain is behind a lake. A third fragment shows only the lake, but it reveals that there is a cabin on the shore of the lake. Therefore, from looking at the overlapping information in these three fragments, you know that the picture contains a mountain behind a lake that has a cabin on its shore. This is the principle behind reconstructing entire DNA sequences using shotgun sequencing.

Originally, shotgun sequencing only analyzed one end of each fragment for overlaps. This was sufficient for sequencing small genomes. However, the desire to sequence larger genomes, such as that of a human, led to the development of double-barrel shotgun sequencing, more formally known as pairwise-end sequencing. In pairwise-end sequencing, both ends of each fragment are analyzed for overlap. Pairwise-end sequencing is, therefore, more cumbersome than shotgun sequencing, but it is easier to reconstruct the sequence because there is more available information.

NEXT-GENERATION SEQUENCING

Since 2005, automated sequencing techniques used by laboratories are under the umbrella of next-generation sequencing, which is a group of automated techniques used for rapid DNA sequencing. These automated low-cost sequencers can generate sequences of hundreds of

thousands or millions of short fragments (25 to 500 base pairs) in the span of one day. These sequencers use sophisticated software to get through the cumbersome process of putting all the fragments in order.

USE OF WHOLE-GENOME SEQUENCES OF MODEL ORGANISMS

The first genome to be completely sequenced was of a bacterial virus, the bacteriophage fx174 (5368 base pairs); this was accomplished by Fred Sanger using shotgun sequencing. Several other organelle and viral genomes were later sequenced. The first organism whose genome was sequenced was the bacterium *Haemophilus influenzae*; this was accomplished by Craig Venter in the 1980s. Approximately 74 different laboratories collaborated on the sequencing of the genome of the yeast *Saccharomyces cerevisiae*, which began in 1989 and was completed in 1996, because it was 60 times bigger than any other genome that had been sequenced. By 1997, the genome sequences of two important model organisms were available: the bacterium *Escherichia coli* K12 and the yeast *Saccharomyces cerevisiae*. Genomes of other model organisms, such as the mouse *Mus musculus*, the fruit fly *Drosophila melanogaster*, the nematode *Caenorhabditis. elegans*, and humans *Homo sapiens* are now known. A lot of basic research is performed in model organisms because the information can be applied to genetically similar organisms. A model organism is a species that is studied as a model to understand the biological processes in other species represented by the model organism. Having entire genomes sequenced helps with the research efforts in these model organisms. The process of attaching biological information to gene sequences is called genome annotation. Annotation of gene sequences helps

with basic experiments in molecular biology, such as designing PCR primers and RNA targets.

USES OF GENOME SEQUENCES

DNA microarrays are methods used to detect gene expression by analyzing an array of DNA fragments that are fixed to a glass slide or a silicon chip to identify active genes and identify sequences. Almost one million genotypic abnormalities can be discovered using microarrays, whereas whole-genome sequencing can provide information about all six billion base pairs in the human genome. Although the study of medical applications of genome sequencing is interesting, this discipline tends to dwell on abnormal gene function. Knowledge of the entire genome will allow future onset diseases and other genetic disorders to be discovered early, which will allow for more informed decisions to be made about lifestyle, medication, and having children. Genomics is still in its infancy, although someday it may become routine to use whole-genome sequencing to screen every newborn to detect genetic abnormalities.

In addition to disease and medicine, genomics can contribute to the development of novel enzymes that convert biomass to biofuel, which results in higher crop and fuel production, and lower cost to the consumer. This knowledge should allow better methods of control over the microbes that are used in the production of biofuels. Genomics could also improve the methods used to monitor the impact of pollutants on ecosystems and help clean up environmental contaminants. Genomics has allowed for the development of agrochemicals and pharmaceuticals that could benefit medical science and agriculture.

It sounds great to have all the knowledge we can get from whole-genome sequencing; however, humans have a

responsibility to use this knowledge wisely. Otherwise, it could be easy to misuse the power of such knowledge, leading to discrimination based on a person's genetics, human genetic engineering, and other ethical concerns. This information could also lead to legal issues regarding health and privacy.

REFERENCES

OpenStax, Biology. OpenStax CNX. January 2, 2017 https://cnx.org/contents/GFy_h8cu@10.120:5l844Z38@7/Whole-Genome-Sequencing

BIOLOGY 213 - ECOLOGY AND EVOLUTION

Principles of Biology introduces biology as a scientific discipline for students planning to major in biology and other science disciplines. Laboratories and classroom activities introduce techniques used to study biological processes and provide opportunities for students to develop their ability to conduct research. The course focus of BI213 is on the interactions of living systems and the ecology and evolution of biodiversity.

Course Outcomes: Upon successful completion of this course, students should be able to

1) Apply the scientific method to biological questions by designing experiments and using the resulting data to form and communicate a conclusion.

2) Assess the strengths and weaknesses of the design of scientific studies and the conclusions drawn from such studies.

3) Select, evaluate and utilize discipline-specific information and literature to research a biological topic.

4) Use evidence to develop informed opinions on contemporary biological issues while considering cultural

and ethical implications.
5) Apply biological theories and concepts to solve problems related to ecology and evolution
6) Discuss and describe modern evolutionary theory and the observations, evidence, and conclusions used to develop the theory of evolution by natural selection
7) Describe how structure and function reflect the ecological challenges faced by organisms and reveal the processes underlying evolutionary change
8) Describe patterns of biological diversity and discuss the biotic and physical processes that have led to these patterns
9) Describe major ecological processes including the transformation of energy and nutrient cycling at the ecosystem and community levels and discuss change in ecological systems
10) Discuss the complexities of biological systems and how interactions at different ecological levels can lead to emergent properties

INTRODUCTION TO EVOLUTION

The theory of **evolution** is the unifying theory of biology, meaning it is the framework within which biologists ask questions about the living world. Its power is that it provides direction for predictions about living things that are borne out in experiment after experiment. The Ukrainian-born American geneticist Theodosius Dobzhansky famously wrote that "nothing makes sense in biology except in the light of evolution" (Dobzhansky, 1964). He meant that the tenet that all life has evolved and diversified from a common ancestor is the foundation from which we approach all questions in biology.

Evolution by natural selection describes a mechanism for how species change over time. There are two levels to this change; microevolution and macroevolution. **Microevolution** refers to changes in the frequency of a gene in a population. These changes can occur in short periods of time and may not be visible until enough data is collected over several generations. **Macroevolution** refers to changes within whole taxonomic groups over long periods of time. This can be seen as the formation of a new trait or feature, the creation of new species, or the loss of species via extinction events.

That species change had been suggested and debated well before Darwin began to explore this idea. The view that species were static and unchanging was grounded in the writings of Plato, yet there were also ancient Greeks who expressed evolutionary ideas. In the eighteenth century, ideas about the evolution of animals were reintroduced by the naturalist Georges-Louis Leclerc Comte de Buffon who observed that various geographic regions have different plant and animal populations, even when the environments are similar. It was also accepted that there were extinct species.

During this time, James Hutton, a Scottish naturalist, proposed that geological change occurred gradually by the accumulation of small changes from processes operating like they are today over long periods of time. This contrasted with the predominant view that the geology of the planet was a consequence of catastrophic events occurring during a relatively brief past. Hutton's view was popularized in the nineteenth century by the geologist Charles Lyell who became a friend to Darwin. Lyell's ideas were influential on Darwin's thinking: Lyell's notion of the greater age of Earth gave more time for gradual change in species, and the process of change provided an analogy for gradual change in species. In the early nineteenth century, Jean-Baptiste Lamarck published a book that detailed a mechanism for evolutionary change. This mechanism is now referred to as an inheritance of **acquired characteristics** by which modifications in an individual are caused by its environment, or the use or disuse of a structure during its lifetime, could be inherited by its offspring and thus bring about change in a species. While this mechanism for evolutionary change was discredited, Lamarck's ideas were an important influence on evolutionary thought.

Charles Darwin and Natural Selection

In the mid-nineteenth century, the actual mechanism for evolution was independently conceived of and described by two naturalists: **Charles Darwin** and **Alfred Russel Wallace**. Importantly, each naturalist spent time exploring the natural world on expeditions to the tropics. From 1831 to 1836, Darwin traveled around the world on *H.M.S. Beagle*, including stops in South America, Australia, and the southern tip of Africa. Wallace traveled to Brazil to collect insects in the Amazon rainforest from 1848 to 1852 and to the Malay Archipelago from 1854 to 1862. Darwin's journey, like Wallace's later journeys to the Malay Archipelago, included stops at several island chains, the last being the Galápagos Islands west of Ecuador. On these islands, Darwin observed species of organisms on different islands that were clearly similar, yet had distinct differences. For example, the ground finches inhabiting the Galápagos Islands comprised several species with a unique beak shape (Figure 1). The species on the islands had a graded series of beak sizes and shapes with very small differences between the most similar. He observed that these finches closely resembled another finch species on the mainland of South America. Darwin imagined that the island species might be species modified from one of the original mainland species. Upon further study, he realized that the varied beaks of each finch helped the birds acquire a specific type of food. For example, seed-eating finches had stronger, thicker beaks for breaking seeds, and insect-eating finches had spear-like beaks for stabbing their prey.

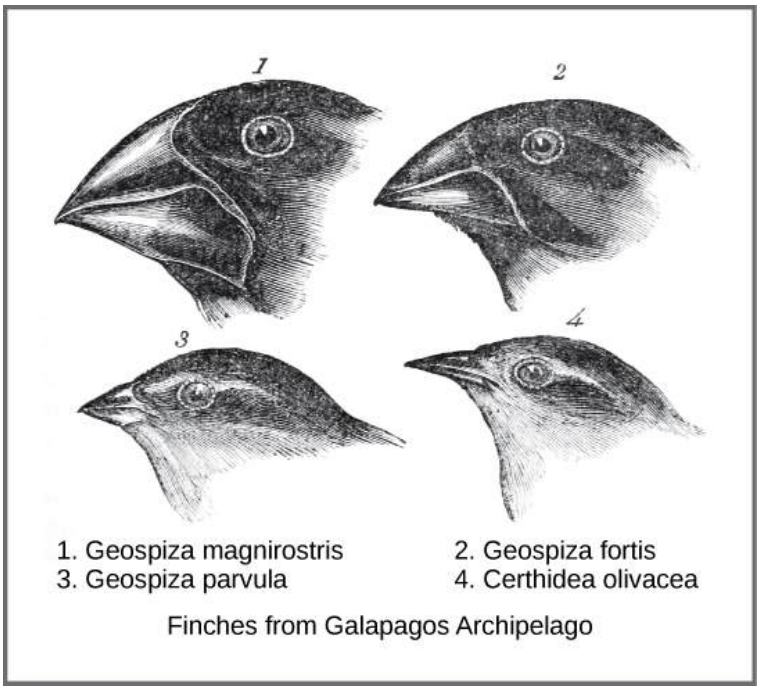


Figure 1: Darwin observed that beak shape varies among finch species. He postulated that the beak of an ancestral species had adapted over time to equip the finches to acquire different food sources. "Galápagos Island Finches" by OpenStax is licensed under CC BY 4.0

Wallace and Darwin both observed similar patterns in other organisms and they independently developed the same explanation for how and why such changes could take place. Darwin called this mechanism **natural selection**. Natural selection, also known as "survival of the fittest," is the more prolific reproduction of individuals with favorable traits that survive environmental change because of those traits; this leads to evolutionary change.

For example, a population of giant tortoises found in the Galapagos Archipelago was observed by Darwin to have longer necks than those that lived on other islands with dry lowlands. These tortoises were "selected" because they could

reach more leaves and access more food than those with short necks. In times of drought when fewer leaves would be available, those that could reach more leaves had a better chance to eat and survive than those that couldn't reach the food source. Consequently, long-necked tortoises would be more likely to be reproductively successful and pass the long-necked trait to their offspring. Over time, only long-necked tortoises would be present in the population.

Natural selection, Darwin argued, was an inevitable outcome of three principles that operated in nature. First, most characteristics of organisms are inherited, or passed from parent to offspring. Although no one, including Darwin and Wallace, knew how this happened at the time, it was a common understanding. Second, more offspring are produced than are able to survive, so resources for survival and reproduction are limited. The capacity for reproduction in all organisms outstrips the availability of resources to support their numbers. Thus, there is competition for those resources in each generation. Both Darwin and Wallace's understanding of this principle came from reading an essay by the economist Thomas Malthus who discussed this principle in relation to human populations. Third, offspring vary among each other in regard to their characteristics and those variations are inherited. Darwin and Wallace reasoned that offspring with inherited characteristics that allow them to best compete for limited resources will survive and have more offspring than those individuals with variations that are less able to compete. Because characteristics are inherited, these traits will be better represented in the next generation. This will lead to change in populations over generations in a process that Darwin called descent with modification. Ultimately, natural selection leads to greater adaptation of the population to its local environment; it is the only mechanism known for **adaptive evolution**.

Papers by Darwin and Wallace presenting the idea of

natural selection were read together in 1858 before the Linnean Society in London. The following year Darwin's book, *On the Origin of Species,* was published. His book outlined in considerable detail his arguments for evolution by natural selection. It's important to note that Darwin is a deeply problematic figure and that many of his contributions to science are remembered because of his privilege and power. Darwin had many racist and offensive beliefs, and the fact that he is often celebrated in biology classes is an example of systematic racism in the classroom. Scientists and educators need to be doing a better job of recognizing prejudice in our materials and highlighting ways to promote racial equity.

Demonstrations of evolution by natural selection are time-consuming and difficult to obtain. One of the best examples has been demonstrated in the very birds that helped to inspire Darwin's theory: the Galápagos finches. **Peter and Rosemary Grant** and their colleagues have studied Galápagos finch populations every year since 1976 and have provided important demonstrations of natural selection. The Grants found changes from one generation to the next in the distribution of beak shapes with the medium ground finch on the Galápagos island of Daphne Major. The birds have inherited variation in the bill shape with some birds having wide deep bills and others having thinner bills. During a period in which rainfall was higher than normal because of an El Niño, the large hard seeds that large-billed birds ate were reduced in number; however, there was an abundance of the small soft seeds which the small-billed birds ate. Therefore, survival and reproduction were much better in the following years for the small-billed birds. In the years following this El Niño, the Grants measured beak sizes in the population and found that the average bill size was smaller. Since bill size is an inherited trait, parents with smaller bills had more offspring and the size of bills had evolved to be smaller. As conditions improved in 1987 and larger seeds

became more available, the trend toward smaller average bill size ceased.

The book The Beak of the Finch: A Story of Evolution in Our Time by Jonathan Weiner is a wonderful exploration of Peter and Rosemary Grant's work. If you want to learn more, I highly recommend checking it out.

Processes and Patterns of Evolution

Natural selection can only take place if there is **variation**, or differences, among individuals in a **population**. Importantly, these differences must have some genetic basis; otherwise, the selection will not lead to change in the next generation. This is critical because variation among individuals can be caused by non-genetic reasons such as an individual being taller because of better nutrition rather than different genes.

Genetic diversity in a population comes from two main mechanisms: **mutation** and **sexual reproduction**. Mutation, a change in DNA, is the ultimate source of new alleles, or new genetic variation in any population. The genetic changes caused by mutation can have one of three outcomes on the phenotype. A mutation affects the phenotype of the organism in a way that gives it **reduced fitness**—lower likelihood of survival or fewer offspring. A mutation may produce a phenotype with a beneficial effect on fitness. And, many mutations will also have no effect on the fitness of the phenotype; these are called **neutral mutations**. Mutations may also have a whole range of effect sizes on the fitness of the organism that expresses them in their phenotype, from a small effect to a great effect. Sexual reproduction also leads to genetic diversity: when two parents reproduce, unique combinations of alleles assemble to produce the unique genotypes and thus phenotypes in each of the offspring.

A heritable trait that helps the survival and reproduction of an organism in its present environment is called an **adaptation**. Scientists describe groups of organisms becoming adapted to their environment when a change in the range of genetic variation occurs over time that increases or maintains the "fit" of the population to its environment. The webbed feet of platypuses are an adaptation for swimming. The snow leopards' thick fur is an adaptation for living in the cold. The cheetahs' fast speed is an adaptation for catching prey.

Whether or not a trait is favorable depends on the environmental conditions at the time. The same traits are not always selected because environmental conditions can change. For example, consider a species of plant that grew in a moist climate and did not need to conserve water. Large leaves were selected because they allowed the plant to obtain more energy from the sun. Large leaves require more water to maintain than small leaves, and the moist environment provided favorable conditions to support large leaves. After thousands of years, the climate changed, and the area no longer had excess water. The direction of natural selection shifted so that plants with small leaves were selected because those populations were able to conserve water to survive the new environmental conditions.

The evolution of species has resulted in enormous variation in form and function. Sometimes, evolution gives rise to groups of organisms that become tremendously different from each other. When two species evolve in diverse directions from a common point, it is called **divergent evolution**.

In other cases, similar phenotypes evolve independently in distantly related species. For example, flight has evolved in both bats and insects, and they both have structures we refer to as wings, which are adaptations to flight. However, the wings of bats and insects have evolved from very different

original structures. This phenomenon is called **convergent evolution**, where similar traits evolve independently in species that do not share a common ancestry. The two species came to the same function, flying, but did so separately from each other.

These physical changes occur over enormous spans of time and help explain how evolution occurs. Natural selection acts on individual organisms, which in turn can shape an entire species. Although natural selection may work in a single generation on an individual, it can take thousands or even millions of years for the genotype of an entire species to evolve. It is over these large time spans that life on earth has changed and continues to change.

REFERENCES

Dobzhansky, T. 1964. "Biology, Molecular and Organismic." *American Zoologist* 4, no. 4: 449.

OpenStax, Biology. OpenStax CNX. June 23, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Evidence of Evolution

The evidence for evolution is compelling and extensive. Looking at every level of organization in living systems, biologists see the signature of past and present evolution. Darwin dedicated a large portion of his book, *On the Origin of Species*, to identifying patterns in nature that were consistent with evolution, and since Darwin, our understanding has become clearer and broader.

Fossils

Fossils provide solid evidence that organisms from the past are not the same as those found today, and fossils show a progression of evolution. Scientists determine the age of fossils and categorize them from all over the world to determine when the organisms lived relative to each other. The resulting fossil record tells the story of the past and shows the evolution of form over millions of years. For example, scientists have recovered highly detailed records showing the evolution of humans and horses (Figure 1).

Figure 1: In this (a) display, fossil hominids are arranged from oldest (bottom) to newest (top). As hominids evolved, the shape of the skull changed. An artist's rendition of (b) extinct species of the genus Equus reveals that these ancient species resembled the modern horse (Equus ferus) but varied in size. (credit: "Fossil evidence" by OpenStax is licensed under CC BY 4.0)

Anatomy and Embryology

Another type of evidence for evolution is the presence of structures in organisms that share the same basic form. For example, the bones in the appendages of a human, dog, bird, and whale all share the same overall construction (Figure 2) resulting from their origin in the appendages of a common ancestor. Over time, evolution led to changes in the shapes and sizes of these bones in different species, but they have maintained the same overall layout. Scientists call these synonymous parts **homologous structures**.

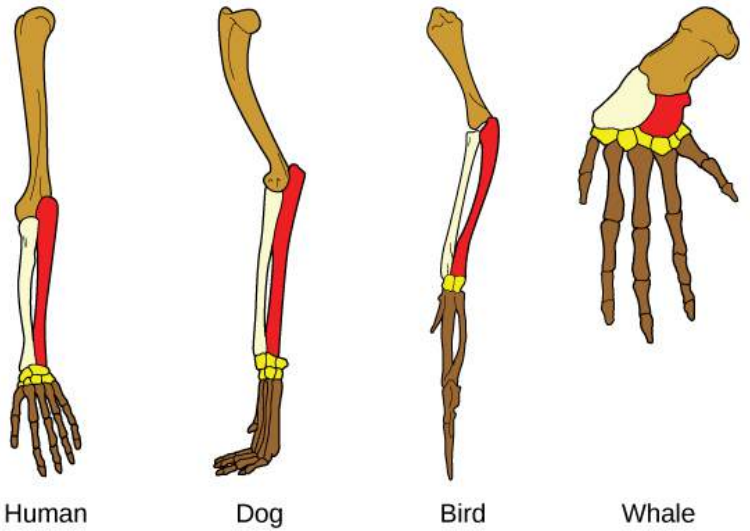


Figure 2: The similar construction of these appendages indicates that these organisms share a common ancestor. (Credit: "homologous structures" by OpenStax is licensed under CC BY 4.0)

Some structures exist in organisms that have no apparent function at all and appear to be residual parts from a past common ancestor. These unused structures without function are called **vestigial structures**. Other examples of vestigial structures are wings on flightless birds, leaves on some cacti, and hind leg bones in whales.

Another evidence of evolution is the **convergence of form** in organisms that share similar environments. For example, species of unrelated animals, such as the arctic fox and ptarmigan, living in the arctic region have been selected for seasonal white phenotypes during winter to blend with the snow and ice (Figure 3). These similarities occur not because of common ancestry, but because of similar selection pressures—the benefits of not being seen by predators.

Figure 3: The white winter coat of the (a) arctic fox and the (b) ptarmigan's plumage are adaptations to their environments. (Credit: "winter coats" by OpenStax is licensed under CC BY 4.0. Figure 4a: modification of work by Keith Morehouse)

Embryology, the study of the development of the anatomy of an organism to its adult form, also provides evidence of relatedness between now widely divergent groups of organisms. Mutational tweaking in the embryo can have such magnified consequences in the adult that embryo formation tends to be conserved. As a result, structures that are absent in some groups often appear in their embryonic forms and disappear by the time the adult or juvenile form is reached. For example, all vertebrate embryos, including humans, exhibit gill slits and tails at some point in their early development. These disappear in the adults of terrestrial groups but are maintained in adult forms of aquatic groups such as fish and some amphibians. Great ape embryos, including humans, have a tail structure during their development that is lost by the time of birth.

Biogeography

The **geographic distribution** of organisms (referred to as **biogeography**) on the planet follows patterns that are best explained by evolution in conjunction with the movement of tectonic plates over geological time. Broad groups that evolved before the breakup of the supercontinent **Pangaea** (about 200 million years ago) are distributed worldwide. Groups that evolved since the breakup appear uniquely in regions of the planet, such as the unique flora and fauna of northern continents that formed from the supercontinent Laurasia and of the southern continents that formed from the supercontinent **Gondwana**. The presence of members of the plant family Proteaceae in Australia, southern Africa, and South America is best by their presence prior to the southern supercontinent Gondwana breaking up.

The great **diversification** of marsupials in Australia and the absence of other mammals reflect Australia's long isolation. Australia has an abundance of endemic species—species found nowhere else—which is typical of islands whose isolation by expanses of water prevents species to migrate. Over time, these species diverge evolutionarily into new species that look very different from their ancestors that may exist on the mainland. The marsupials of Australia, the finches on the Galápagos, and many species on the Hawaiian Islands are all unique to their one point of origin, yet they display distant relationships to ancestral species on the mainland.

Molecular Biology

Like anatomical structures, the structures of the molecules of life reflect descent with modification. Evidence of a common ancestor for all of life is reflected in the universality of **DNA** as the genetic material and in the near universality of the

genetic code and the machinery of DNA replication and expression. Fundamental divisions in life between the three domains are reflected in major structural differences in otherwise conservative structures such as the components of ribosomes and the structures of membranes. In general, the relatedness of groups of organisms is reflected in the similarity of their DNA sequences—exactly the pattern that would be expected from descent and diversification from a common ancestor.

DNA sequences have also shed light on some of the mechanisms of evolution. For example, it is clear that the evolution of new functions for proteins commonly occurs after gene duplication events that allow the free modification of one copy by mutation, selection, or drift (changes in a population's gene pool resulting from chance), while the second copy continues to produce a functional protein.

Misconceptions of Evolution

Although the theory of evolution generated some controversy when it was first proposed, it was almost universally accepted by biologists, particularly younger biologists, within 20 years after the publication of *On the Origin of Species*. Nevertheless, the theory of evolution is a difficult concept and misconceptions about how it works abound.

Evolution Is Just a Theory

Critics of the **theory** of evolution dismiss its importance by purposefully confounding the everyday usage of the word "theory" with the way scientists use the word. In science, a "theory" is understood to be a body of thoroughly tested and verified explanations for a set of observations of the natural world. Scientists have a theory of the atom, a theory

of gravity, and the theory of relativity, each of which describes understood facts about the world. In the same way, the theory of evolution describes facts about the living world. As such, a theory in science has survived significant efforts to discredit it by scientists. In contrast, a "theory" in common vernacular is a word meaning a guess or suggested explanation; this meaning is more akin to the scientific concept of "hypothesis." When critics of evolution say evolution is "just a theory," they are implying that there is little evidence supporting it and that it is still in the process of being rigorously tested. This is a mischaracterization.

Individuals Evolve

Evolution is the change in the genetic composition of a population over time, specifically over generations, resulting from differential reproduction of individuals with certain alleles. Individuals do change over their lifetime, obviously, but this is called development and involves changes programmed by the set of genes the individual acquired at birth in coordination with the individual's environment. When thinking about the evolution of a characteristic, it is probably best to think about the change of the average value of the characteristic in the population over time. For example, when natural selection leads to a beak size change in medium-ground finches in the Galápagos, this does not mean that individual beaks on living finches are changing. Instead, it means that if one measures the average beak size among all individuals in the population at one time and then measures the average beak size in the population several years later, the average value will be different as a result of evolution. Although some individuals may survive from the first time to the second, they will still have the same beak size; however, there will be many new individuals that contribute to the shift in average beak size.

Evolution Explains the Origin of Life

It is a common misunderstanding that evolution includes an explanation of **life's origins**. The theory does not try to explain the origin of life. The theory of evolution explains how populations change over time and how life diversifies. It does not shed light on the beginnings of life including the origins of the first cells, which is how life is defined. The mechanisms of the origin of life on Earth are a particularly difficult problem because it occurred a very long time ago, and presumably, it just occurred once.

However, once a mechanism of inheritance was in place in the form of a molecule like DNA either within a cell or pre-cell, these entities would be subject to the principle of natural selection. More effective reproducers would increase in frequency at the expense of inefficient reproducers. So while evolution does not explain the origin of life, it may have something to say about some of the processes operating once pre-living entities acquired certain properties.

Organisms Evolve on Purpose

Statements such as "organisms evolve in response to a change in an environment" are quite common, but such statements can lead to two types of misunderstandings. First, the statement must not be understood to mean that individual organisms evolve. The statement is shorthand for "a population evolves in response to a changing environment." However, a second misunderstanding may arise by interpreting the statement to mean that evolution is somehow **intentional**. A changed environment results in some individuals in the population, those with particular phenotypes, benefiting and therefore producing proportionately more offspring than other phenotypes. This

results in a change in the population if the characteristics are genetically determined.

It is also important to understand that the variation that natural selection works on is already in a population and does not arise in **response** to an environmental change. For example, applying antibiotics to a population of bacteria will, over time, select a population of bacteria that are resistant to antibiotics. The resistance, which is caused by a gene, did not arise by mutation because of the application of the antibiotic. The gene for resistance was already present in the gene pool of the bacteria, likely at a low frequency. The antibiotic, which kills the bacterial cells without the resistance gene, strongly selects individuals that are resistant, since these would be the only ones that survived and divided. Experiments have demonstrated that mutations for antibiotic resistance do not arise as a result of antibiotics.

In a larger sense, **evolution is not goal-directed**. Species do not become "better" over time; they simply track their changing environment with adaptations that maximize their reproduction in a particular environment at a particular time. Evolution has no goal of making faster, bigger, more complex, or even smarter species, despite the commonness of this kind of language in popular discourse. What characteristics evolve in a species are a function of the variation present and the environment, both of which are constantly changing in a non-directional way. What trait is fit in one environment at one time may well be fatal at some point in the future. This holds equally well for a species of insect as it does the human species.

Summary

Evolution is the process of adaptation through mutation which allows more desirable characteristics to be passed to the next generation. Over time, organisms evolve more

characteristics that are beneficial to their survival. For living organisms to adapt and change to environmental pressures, genetic variation must be present. With genetic variation, individuals have differences in form and function that allow some to survive certain conditions better than others. These organisms pass their favorable traits to their offspring. Eventually, environments change, and what was once a desirable, advantageous trait may become an undesirable trait and organisms may further evolve. Evolution may be convergent with similar traits evolving in multiple species or divergent with diverse traits evolving in multiple species that came from a common ancestor. Evidence of evolution can be observed by means of DNA code and the fossil record, and also by the existence of homologous and vestigial structures.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 23, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Formation of New Species

Although all life on earth shares various genetic similarities, only certain organisms combine genetic information by sexual reproduction and have offspring that can then successfully reproduce. Scientists call such organisms members of the same biological species.

Species Concepts

It turns out that scientists don't always agree on the definition of a species. The different ideas about what does and does not constitute a species are referred to as **species concepts**. There are around twenty-six different species concepts, but four are the most accepted. The **biological species concept** states that if two organisms are able to successfully breed and produce viable, fertile offspring, then they are the same species. When populations of organisms cease to successfully breed (called **reproductive isolation**) they are then considered separate species. The biological species concept works well for scientists studying living creatures that have regular breeding patterns, such as insects or mammals. The **morphological species concept** states that if two organisms are morphologically similar enough, then they are the same species. There is no guideline about how similar counts as similar enough, so it is up to the researcher to make the judgment call. This concept

works well for organisms that don't breed regularly (such as fungi) or that are no longer living. The **genetic species concept** states that if two organisms are genetically similar enough, then they are the same species. Again, there is no rule about how genetically similar they need to be, so each discipline determines its own limits. This concept works well for organisms that are very tiny, have unusual reproduction strategies, and not morphologically distinct (such as bacteria). The **evolutionary species concept** states that if two organisms' evolutionary paths are similar enough, then they are the same species. This concept works well for long-dead organisms because fossils cannot breed, often are lacking impressions of soft tissues, and usually don't have enough DNA left to work with.

Species and the Ability to Reproduce

A species is a group of individual organisms that **interbreed** and produce **fertile**, **viable** offspring. According to this definition, one species is distinguished from another when, in nature, it is not possible for matings between individuals from each species to produce fertile offspring.

Members of the same species share both external and internal characteristics, which develop from their DNA. The closer relationship two organisms share, the more DNA they have in common, just like people and their families. People's DNA is likely to be more like their father or mother's DNA than their cousin or grandparent's DNA. Organisms of the same species have the highest level of DNA alignment and therefore share characteristics and behaviors that lead to successful reproduction.

Species' appearance can be misleading in suggesting an ability or inability to mate. For example, even though domestic dogs (*Canis lupus familiaris*) display phenotypic differences, such as size, build, and coat, most dogs can

interbreed and produce viable puppies that can mature and sexually reproduce (Figure 5).

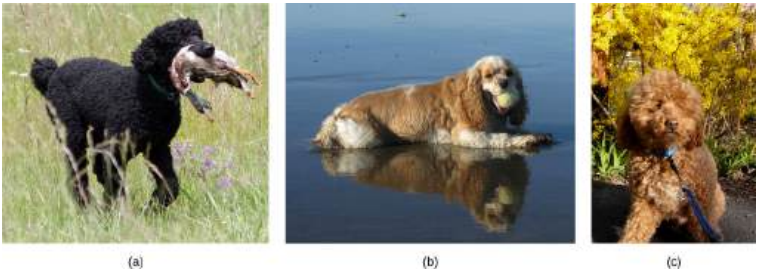


Figure 5: The (a) poodle and (b) cocker spaniel can reproduce to produce a breed known as (c) the cockapoo. (credit a: modification of work by Sally Eller, Tom Reese; credit b: modification of work by Jeremy McWilliams; credit c: modification of work by Kathleen Conklin. "this image" by OpenStax is licensed under CC BY 4.0)

In other cases, individuals may appear similar although they are not members of the same species. For example, even though bald eagles (*Haliaeetus leucocephalus*) and African fish eagles (*Haliaeetus vocifer*) are both birds and eagles, each belongs to a separate species group (Figure 6a). Mushrooms often have look-alikes that are not closely related as well. For example, the wonderful edible chanterelle mushroom (*Cantharellus cibarius*) and the toxic false chanterelle (*Hygrophoropsis aurantiaca*) look nearly identical and grow in similar habitats (Figure 6b). If humans were to artificially intervene and fertilize the egg of a bald eagle with the sperm of an African fish eagle and a chick did hatch, that offspring, called a **hybrid** (a cross between two species), would probably be **infertile**—unable to successfully reproduce after it reached maturity. Different species may have different genes that are active in development; therefore, it may not be possible to develop a viable offspring with two different sets of directions. Thus, even though hybridization may take place, the two species still remain separate.

Figure 6a: The (a) African fish eagle is similar in appearance to the (b) bald eagle, but the two birds are members of different species. (credit a: modification of work by Nigel Wedge; credit b: modification of work by U.S. Fish and Wildlife Service. "this image" by OpenStax is licensed under CC BY 4.0)

Figure 6b: the edible chanterelle mushroom Cantharellus cibarius (left) and the toxic false chanterelle Hygrophoropsis aurantiaca (right) display similar appearances. Image credit Bf5man (left) and Jerzy Opioła (right) via Wikicommons CC BY-SA 3.0

Populations of species share a **gene pool**: a collection of all the variants of genes in the species. Again, the basis for any changes in a group or population of organisms must be genetic for this is the only way to share and pass on traits. When variations occur within a species, they can only be passed to the next generation along two main pathways: **asexual reproduction** or **sexual reproduction**. The change

will be passed on asexually simply if the reproducing cell possesses the changed trait. For the changed trait to be passed on by sexual reproduction, a gamete, such as a sperm or egg cell, must possess the changed trait. In other words, sexually-reproducing organisms can experience several genetic changes in their body cells, but if these changes do not occur in a sperm or egg cell, the changed trait will never reach the next generation. Only **heritable traits** can evolve. Therefore, reproduction plays a paramount role for genetic change to take root in a population or species. In short, organisms must be able to reproduce with each other to pass new traits to offspring.

Speciation

The biological definition of species, which works for sexually reproducing organisms, is a group of actually or potentially interbreeding individuals. There are exceptions to this rule. Many species are similar enough that hybrid offspring are possible and may often occur in nature, but for the majority of species, this rule generally holds. In fact, the presence in nature of hybrids between similar species suggests that they may have descended from a single interbreeding species, and the speciation process may not yet be completed.

Given the extraordinary diversity of life on the planet, there must be mechanisms for **speciation**: the formation of two species from one original species. Darwin envisioned this process as a branching event and diagrammed the process in the only illustration found in *On the Origin of Species* (Figure 7a). Compare this illustration to the diagram of elephant evolution (Figure 7b), which shows that as one species changes over time, it branches to form more than one new species, repeatedly, as long as the population survives or until the organism becomes extinct.

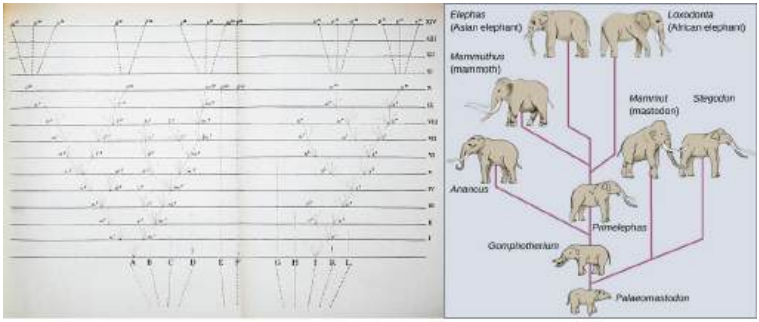


Figure 7: The only illustration in Darwin's On the Origin of Species is (a) a diagram showing speciation events leading to biological diversity. The diagram shows similarities to phylogenetic charts that are drawn today to illustrate the relationships of species. (b) Modern elephants evolved from the Palaeomastodon, a species that lived in Egypt 35–50 million years ago. (credit: "this image" by OpenStax is licensed under CC BY 4.0)

For speciation to occur, two new populations must be formed from one original population and they must evolve in such a way that it becomes impossible for individuals from the two new populations to interbreed. Biologists have proposed mechanisms by which this could occur that fall into two broad categories. **Allopatric speciation** (allo- = "other"; -patric = "homeland") involves geographic separation of populations from a parent species and subsequent evolution. **Sympatric speciation** (sym- = "same"; -patric = "homeland") involves speciation occurring within a parent species remaining in one location.

Biologists think of speciation events as the splitting of one ancestral species into two descendant species. There is no reason why there might not be more than two species formed at one time except that it is less likely and multiple events can be conceptualized as single splits occurring close in time.

Allopatric Speciation

A geographically continuous population has a gene pool that is relatively homogeneous. **Gene flow**, the movement of alleles across the range of the species, is relatively free because individuals can move and then mate with individuals in their new location. Thus, the frequency of an allele at one end of a distribution will be similar to the frequency of the allele at the other end. When populations become geographically discontinuous, that free-flow of alleles is prevented. When that separation lasts for a period of time, the two populations are able to evolve along different trajectories. Thus, their allele frequencies at numerous genetic loci gradually become more and more different as new alleles independently arise by mutation in each population. Typically, environmental conditions, such as climate, resources, predators, and competitors for the two populations will differ causing natural selection to favor divergent adaptations in each group.

Isolation of populations leading to allopatric speciation can occur in a variety of ways: a river forming a new branch, erosion forming a new valley, a group of organisms traveling to a new location without the ability to return, or seeds floating over the ocean to an island. The nature of the geographic separation necessary to isolate populations depends entirely on the biology of the organism and its potential for dispersal. If two flying insect populations took up residence in separate nearby valleys, chances are, individuals from each population would fly back and forth continuing gene flow. However, if two rodent populations became divided by the formation of a new lake, continued gene flow would be unlikely; therefore, speciation would be more likely.

Biologists group allopatric processes into two categories: **dispersal** and **vicariance**. Dispersal is when a few members

of a species move to a new geographical area, and vicariance is when a natural situation arises to physically divide organisms.

Scientists have documented numerous cases of allopatric speciation taking place. For example, along the west coast of the United States, two separate sub-species of spotted owls exist. The northern spotted owl has genetic and phenotypic differences from its close relative: the Mexican spotted owl, which lives in the south (Figure 8).

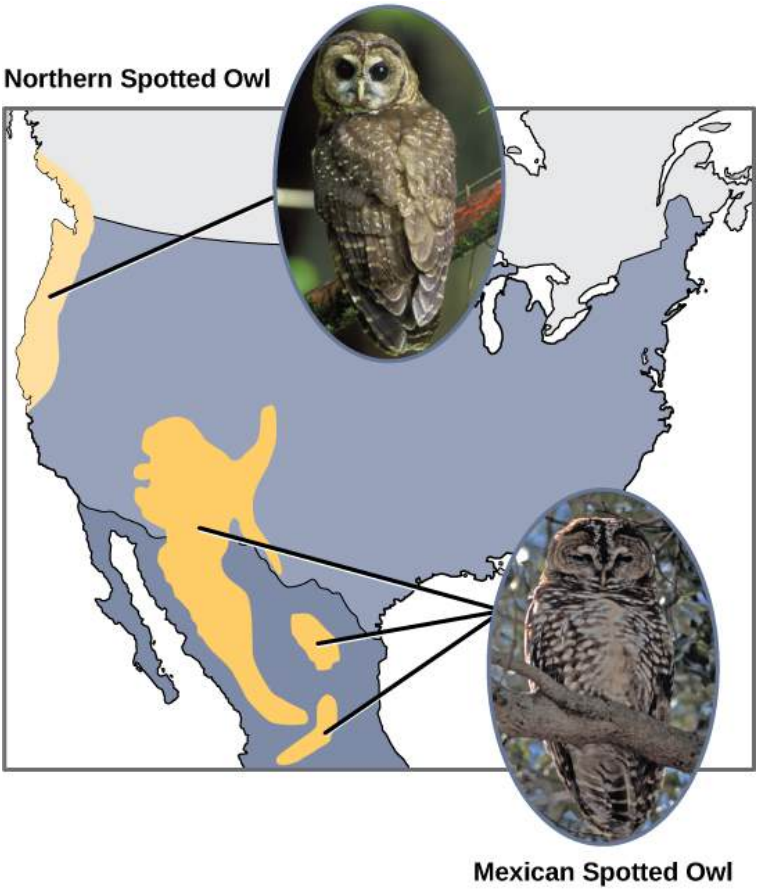


Figure 8: The northern spotted owl and the Mexican spotted owl inhabit geographically separate locations with different climates and ecosystems. The owl is an example of allopatric speciation. (credit "northern spotted owl": modification of work by John and Karen Hollingsworth; credit "Mexican spotted owl": modification of work by Bill Radke. "spotted owls" by OpenStax is licensed under CC BY 4.0)

Additionally, scientists have found that the further the distance between two groups that once were the same species, the more likely it is that speciation will occur. This

seems logical because as the distance increases, the various environmental factors would likely have less in common than locations in close proximity. Consider the two owls: in the north, the climate is cooler than in the south; the types of organisms in each ecosystem differ, as do their behaviors and habits; also, the hunting habits and prey choices of the southern owls vary from the northern owls. These variances can lead to evolved differences in the owls, and speciation likely will occur.

Adaptive Radiation

In some cases, a population of one species disperses throughout an area, and each finds a distinct niche or isolated habitat. Over time, the varied demands of their new lifestyles lead to multiple speciation events originating from a single species. This is called **adaptive radiation** because many adaptations evolve from a single point of origin; thus, causing the species to radiate into several new ones. Island archipelagos like the Hawaiian Islands provide an ideal context for adaptive radiation events because water surrounds each island which leads to geographical isolation for many organisms. The Hawaiian honeycreeper illustrates one example of adaptive radiation. From a single species, called the founder species, numerous species have evolved, including the six shown in Figure 9.

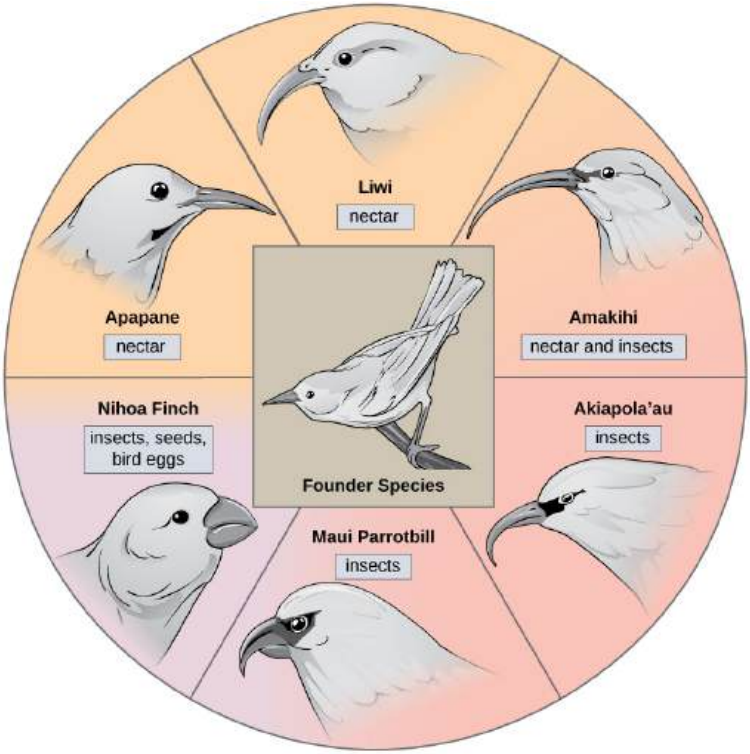


Figure 9: The honeycreeper birds illustrate adaptive radiation. From one original species of bird, multiple others evolved, each with its own distinctive characteristics. (credit: "honeycreepers" by OpenStax is licensed under CC BY 4.0)

Notice the differences in the species' beaks. Evolution in response to natural selection based on specific food sources in each new habitat led to the evolution of a different beak suited to the specific food source. The seed-eating bird has a thicker, stronger beak which is suited to break hard nuts. The nectar-eating birds have long beaks to dip into flowers to reach the nectar. The insect-eating birds have beaks like swords, appropriate for stabbing and impaling insects.

Darwin's finches are another example of adaptive radiation in an archipelago.

Sympatric Speciation

Can divergence occur if no physical barriers are in place to separate individuals who continue to live and reproduce in the same habitat? The answer is yes. The process of speciation within the same space is called sympatric speciation; the prefix "sym" means same, so "sympatric" means "same homeland" in contrast to "allopatric" meaning "other homeland." A number of mechanisms for sympatric speciation have been proposed and studied.

One form of sympatric speciation can begin with a serious chromosomal error during cell division. In a normal cell division event chromosomes replicate, pair up, and then separate so that each new cell has the same number of chromosomes. However, sometimes the pairs separate and the end cell product has too many or too few individual chromosomes in a condition called **aneuploidy** (Figure 10).

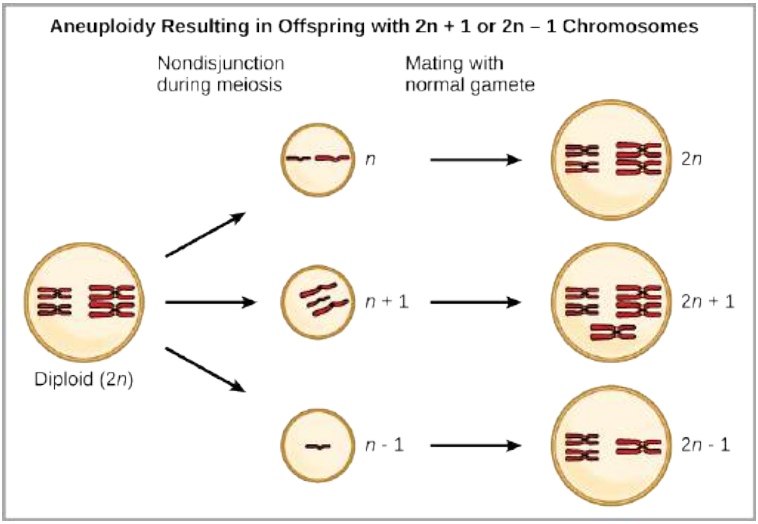


Figure 10: Aneuploidy results when the gametes have too many or too few chromosomes due to nondisjunction during meiosis. In the example shown here, the resulting offspring will have 2n+1 or 2n-1 chromosomes (credit: "Aneuploidy" by OpenStax is licensed under CC BY 4.0)

Polyploidy is a condition in which a cell or organism has an extra set, or sets, of chromosomes. Scientists have identified two main types of polyploidy that can lead to reproductive isolation of an individual in the polyploidy state. Reproductive isolation is the inability to interbreed. In some cases, a polyploid individual will have two or more complete sets of chromosomes from its own species in a condition called **autopolyploidy** (Figure 11). The prefix "auto-" means "self," so the term means multiple chromosomes from one's own species. Polyploidy results from an error in meiosis in which all of the chromosomes move into one cell instead of separating.

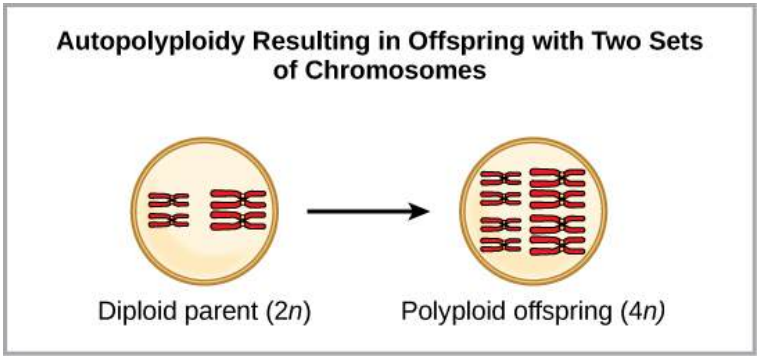


Figure 11: Autopolyploidy results when mitosis is not followed by cytokinesis. (credit: "Autopolyploidy" by OpenStax is licensed under CC BY 4.0)

For example, if a plant species with $2n = 6$ produces autopolyploid gametes that are also diploid ($2n = 6$, when they should be $n = 3$), the gametes now have twice as many chromosomes as they should have. These new gametes will be incompatible with the normal gametes produced by this plant species. However, they could either self-pollinate or reproduce with other autopolyploid plants with gametes having the same diploid number. In this way, sympatric speciation can occur quickly by forming offspring with $4n$ called a tetraploid. These individuals would immediately be able to reproduce only with those of this new kind and not those of the ancestral species.

The other form of polyploidy occurs when individuals of two different species reproduce to form a viable offspring called an **allopolyploid**. The prefix "allo-" means "other" (recall from allopatric): therefore, an allopolyploid occurs when gametes from two different species combine. Figure 12 illustrates one possible way an allopolyploid can form. Notice how it takes two generations, or two reproductive acts, before the viable fertile hybrid results.

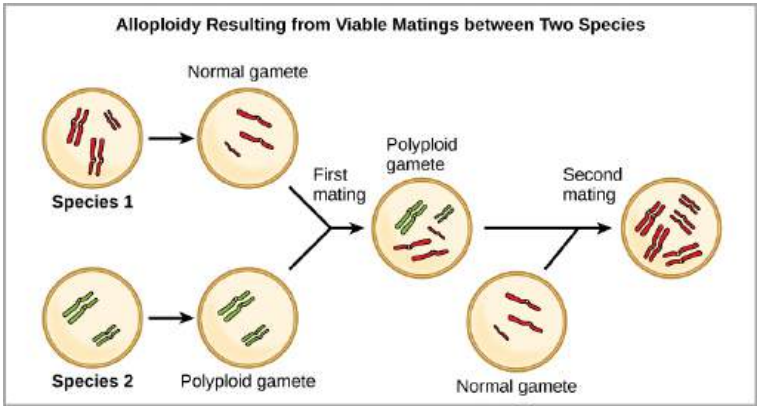


Figure 12: Alloploidy results when two species mate to produce viable offspring. In the example shown, a normal gamete from one species fuses with a polyploidy gamete from another. Two matings are necessary to produce viable offspring. (credit: "Alloploidy" by OpenStax is licensed under CC BY 4.0)

The cultivated forms of wheat, cotton, and tobacco plants are all allopolyploids. Although polyploidy occurs occasionally in animals, it takes place most commonly in plants. (Animals with any of the types of chromosomal aberrations described here are unlikely to survive and produce normal offspring.) Scientists have discovered more than half of all plant species studied relate back to a species that evolved through polyploidy. With such a high rate of polyploidy in plants, some scientists hypothesize that this mechanism takes place more as an adaptation than as an error.

Reproductive Isolation

Given enough time, the genetic and phenotypic divergence between populations will affect characters that influence reproduction: if individuals of the two populations were to be brought together, mating would be less likely, but if mating

occurred, offspring would be non-viable or infertile. Many types of diverging characters may affect the reproductive isolation, the ability to interbreed, of the two populations.

Reproductive isolation can take place in a variety of ways. Scientists organize them into two groups: **prezygotic barriers** and **postzygotic barriers**. Recall that a **zygote** is a fertilized egg: the first cell of the development of an organism that reproduces sexually. Therefore, a prezygotic barrier is a mechanism that blocks reproduction from taking place; this includes barriers that prevent fertilization when organisms attempt reproduction. A postzygotic barrier occurs after zygote formation; this includes organisms that don't survive the embryonic stage and those that are born sterile.

Some types of prezygotic barriers prevent reproduction entirely. Many organisms only reproduce at certain times of the year, often just annually. Differences in breeding schedules, called **temporal isolation**, can act as a form of reproductive isolation. For example, two species of frogs inhabit the same area, but one reproduces from January to March, whereas the other reproduces from March to May (Figure 13).

Figure 13: These two related frog species exhibit temporal reproductive isolation. (a) Rana aurora breeds earlier in the year than (b) Rana boylii. (credit a: modification of work by Mark R. Jennings, USFWS; credit b: modification of work by Alessandro Catenazzi. "frog reproductive isolation" by OpenStax is licensed under CC BY 4.0)

In some cases, populations of a species move or are moved to a new habitat and take up residence in a place that no longer overlaps with the other populations of the same species. This situation is called **habitat isolation**. Reproduction with the parent species ceases, and a new group exists that is now reproductively and genetically independent. For example, a cricket population that was divided after a flood could no longer interact with each other. Over time, the forces of natural selection, mutation, and genetic drift will likely result in the divergence of the two groups.

Behavioral isolation occurs when the presence or absence of a specific behavior prevents reproduction from taking place. For example, male fireflies use specific light patterns to attract females. Various species of fireflies display their lights differently. If a male of one species tried to attract the female of another, she would not recognize the light pattern and would not mate with the male.

Other prezygotic barriers work when differences in their gamete cells (eggs and sperm) prevent fertilization from taking place; this is called a **gametic barrier**. Similarly, in some cases closely related organisms try to mate, but their reproductive structures simply do not fit together. For example, damselfly males of different species have differently shaped reproductive organs. If one species tries to mate with the female of another, their body parts simply do not fit together. (Figure 14).

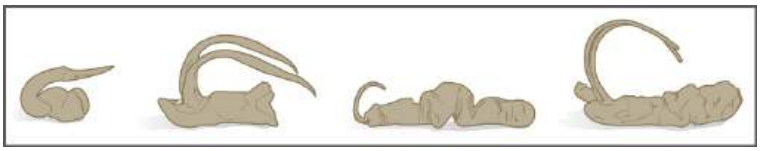


Figure 14: The shape of the male reproductive organ varies among male damselfly species, and is only compatible with the female of that species. Reproductive organ incompatibility keeps the species reproductively isolated. (credit: "damselfly reproductive organ" by OpenStax is licensed under CC BY 4.0)

In plants, certain structures aimed to attract one type of pollinator simultaneously prevent a different pollinator from accessing the pollen. The tunnel through which an animal must access nectar can vary widely in length and diameter, which prevents the plant from being cross-pollinated with a different species (Figure 15).

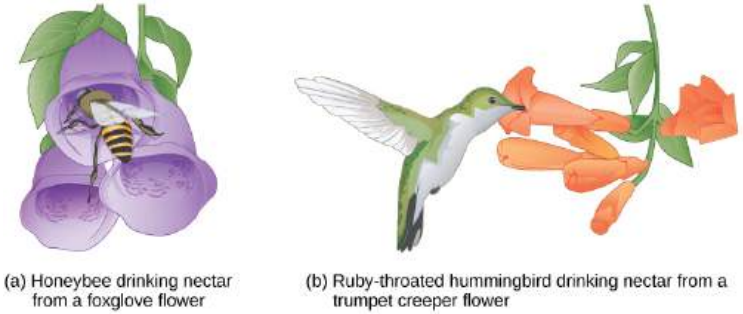


Figure 15: Some flowers have evolved to attract certain pollinators. The (a) wide foxglove flower is adapted for pollination by bees, while the (b) long, tube-shaped trumpet creeper flower is adapted for pollination by hummingbirds. (credit: "pollenation" by OpenStax is licensed under CC BY 4.0)

When fertilization takes place and a zygote forms, postzygotic barriers can prevent reproduction. Hybrid

individuals in many cases cannot form normally in the womb and simply do not survive past the embryonic stages. This is called **hybrid inviability** because the hybrid organisms simply are not viable. In another postzygotic situation, reproduction leads to the birth and growth of a hybrid that is sterile and unable to reproduce offspring of their own; this is called **hybrid sterility**.

Habitat Influence on Speciation

Sympatric speciation may also take place in ways other than polyploidy. For example, consider a species of fish that lives in a lake. As the population grows, competition for food also grows. Under pressure to find food, suppose that a group of these fish had the genetic flexibility to discover and feed off another resource that was unused by the other fish. What if this new food source was found at a different depth of the lake? Over time, those feeding on the second food source would interact more with each other than the other fish; therefore, they would breed together as well. The offspring of these fish would likely behave as their parents: feeding and living in the same area and keeping them separate from the original population. If this group of fish continued to remain separate from the first population, eventually sympatric speciation might occur as more genetic differences accumulated between them.

This scenario does play out in nature, as do others that lead to reproductive isolation. One such place is Lake Victoria in Africa, famous for its sympatric speciation of cichlid fish. Researchers have found hundreds of sympatric speciation events in these fish, which have not only happened in great number, but also over a short period of time. Figure 16 shows this type of speciation among a cichlid fish population in Nicaragua. In this locale, two types of cichlids live in the

same geographic location but have come to have different morphologies that allow them to eat various food sources.

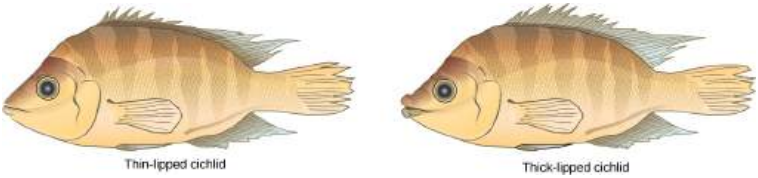


Figure 16: Cichlid fish from Lake Apoyeque, Nicaragua, show evidence of sympatric speciation. Lake Apoyeque, a crater lake, is 1800 years old, but genetic evidence indicates that the lake was populated only 100 years ago by a single population of cichlid fish. Nevertheless, two populations with distinct morphologies and diets now exist in the lake, and scientists believe these populations may be in an early stage of speciation. (credit: "Cichlid fish" by OpenStax is licensed under CC BY 4.0)

Summary

Speciation occurs along two main pathways: geographic separation (allopatric speciation) and through mechanisms that occur within a shared habitat (sympatric speciation). Both pathways isolate a population reproductively in some form. Mechanisms of reproductive isolation act as barriers between closely related species, enabling them to diverge and exist as genetically independent species. Prezygotic barriers block reproduction prior to the formation of a zygote, whereas postzygotic barriers block reproduction after fertilization occurs. For a new species to develop, something must cause a breach in the reproductive barriers. Sympatric speciation can occur through errors in meiosis that form gametes with extra chromosomes (polyploidy). Autopolyploidy occurs within a single species, whereas allopolyploidy occurs between closely related species.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 23, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Speciation

Speciation occurs over a span of evolutionary time, so when a new species arises, there is a transition period during which the closely related species continue to interact.

Reconnection

After speciation, two species may recombine or even continue interacting indefinitely. Individual organisms will mate with any nearby individual with whom they are capable of breeding. An area where two closely related species continue to interact and reproduce, forming hybrids, is called a **hybrid zone**. Over time, the hybrid zone may change depending on the fitness of the hybrids and the reproductive barriers (Figure 1). If the hybrids are less fit than the parents, reinforcement of speciation occurs, and the species continue to diverge until they can no longer mate and produce viable offspring. If reproductive barriers weaken, fusion occurs and the two species become one. Barriers remain the same if hybrids are fit and reproductive: stability may occur and hybridization continues.

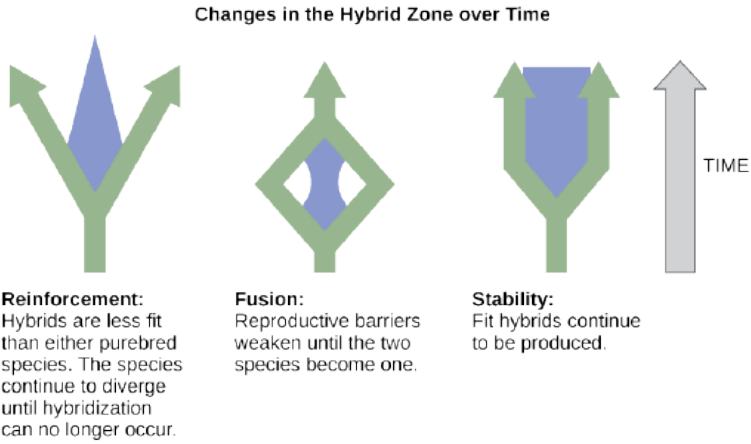


Figure 1: After speciation has occurred, the two separate but closely related species may continue to produce offspring in an area called the hybrid zone. Reinforcement, fusion, or stability may result, depending on reproductive barriers and the relative fitness of the hybrids. (credit: "speciation diagram" by OpenStax is licensed under CC BY 4.0)

If two species eat a different diet but one of the food sources is eliminated and both species are forced to eat the same foods, what change in the hybrid zone is most likely to occur? Hybrids can be either less fit than the parents, more fit, or about the same. Usually, hybrids tend to be less fit; therefore, such reproduction diminishes over time, nudging the two species to diverge further in a process called **reinforcement**. This term is used because the low success of the hybrids reinforces the original speciation. If the hybrids are as fit or more fit than the parents, the two species may fuse back into one species (Figure 1). Scientists have also observed that sometimes two species will remain separate but also continue to interact to produce some hybrid individuals; this is classified as **stability** because no real net change is taking place.

Varying Rates of Speciation

Scientists around the world study speciation, documenting observations both of living organisms and those found in the fossil record. As their ideas take shape and as research reveals new details about how life evolves, they develop models to help explain rates of speciation. In terms of how quickly speciation occurs, two patterns are currently observed: gradual speciation model and punctuated equilibrium model.

In the **gradual speciation** model, species diverge gradually over time in small steps. In the **punctuated equilibrium** model, a new species undergoes changes quickly from the parent species and then remains largely unchanged for long periods of time afterward (Figure 2). This early change model is called punctuated equilibrium because it begins with a punctuated or periodic change and then remains in balance afterward. While punctuated equilibrium suggests a faster tempo, it does not necessarily exclude gradualism.

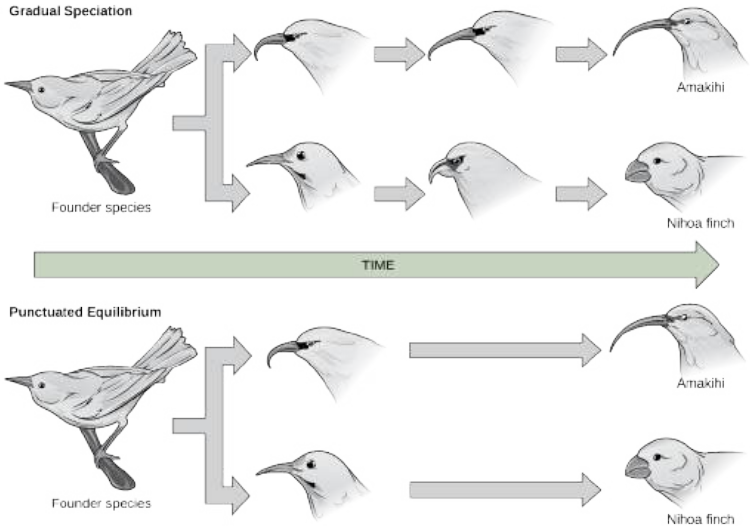


Figure 2: In (a) gradual speciation, species diverge at a slow, steady pace as traits change incrementally. In (b) punctuated equilibrium, species diverge quickly and then remain unchanged for long periods of time. (credit: "gradual speciation and punctuated equilibrium" by OpenStax is licensed under CC BY 4.0)

The primary influencing factor on changes in speciation rate is environmental conditions. Under some conditions, selection occurs quickly or radically. Consider a species of snails that had been living with the same basic form for many thousands of years. Layers of their fossils would appear similar for a long time. When a change in the environment takes place—such as a drop in the water level—a small number of organisms are separated from the rest in a brief period of time, essentially forming one large and one tiny population. The tiny population faces new environmental conditions. Because its gene pool quickly became so small, any variation that surfaces and that aids in surviving the new conditions becomes the predominant form.

Summary

Speciation is not a precise division: overlap between closely related species can occur in areas called hybrid zones. Organisms reproduce with other similar organisms. The fitness of these hybrid offspring can affect the evolutionary path of the two species. Scientists propose two models for the rate of speciation: one model illustrates how a species can change slowly over time; the other model demonstrates how change can occur quickly from a parent generation to a new species. Both models continue to follow the patterns of natural selection.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

EVOLUTION IN ACTION

All life on Earth is related. Evolutionary theory states that humans, beetles, plants, and bacteria all share a common ancestor, but that millions of years of evolution have shaped each of these organisms into the forms seen today. Scientists consider evolution a key concept to understanding life. Natural selection is one of the most dominant evolutionary forces. Natural selection acts to promote traits and behaviors that increase an organism's chances of survival and reproduction while eliminating those traits and behaviors that are to the organism's detriment. But natural selection can only, as its name implies, select—it cannot create. The introduction of novel traits and behaviors falls on the shoulders of another evolutionary force—mutation. Mutation and other sources of variation among individuals, as well as the evolutionary forces that act upon them, alter populations and species. This combination of processes has led to the world of life we see today.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020.

https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Population Evolution

The mechanisms of inheritance, or genetics, were not understood at the time Charles Darwin and Alfred Russel Wallace were developing their idea of natural selection. This lack of understanding was a stumbling block to understanding many aspects of evolution. In fact, the predominant (and incorrect) genetic theory of the time, blending inheritance, made it difficult to understand how natural selection might operate. Darwin and Wallace were unaware of the genetics work by Austrian monk Gregor Mendel, which was published in 1866, not long after the publication of Darwin's book *On the Origin of Species*. Mendel's work was rediscovered in the early twentieth century at which time geneticists were rapidly coming to an understanding of the basics of inheritance. Over the next few decades, genetics and evolution were integrated into what became known as the **modern synthesis**—the coherent understanding of the relationship between natural selection and genetics that took shape by the 1940s and is generally accepted today. In sum, the modern synthesis describes how evolutionary processes, such as natural selection, can affect a population's genetic makeup, and, in turn, how this can result in the gradual evolution of populations and species. The theory also connects this change of a population over time, called **microevolution**, with the processes that gave

rise to new species and higher taxonomic groups with widely divergent characters, called **macroevolution**.

Evolution and Flu Vaccines

Every fall, the media starts reporting on flu vaccinations and potential outbreaks. Scientists, health experts, and institutions determine recommendations for different parts of the population, predict optimal production and inoculation schedules, create vaccines, and set up clinics to provide inoculations. You may think of the annual flu shot as a lot of media hype, an important health protection, or just a briefly uncomfortable prick in your arm. But do you think of it in terms of evolution?

The media hype of annual flu shots is scientifically grounded in our understanding of evolution. Each year, scientists across the globe strive to predict the flu strains that they anticipate being most widespread and harmful in the coming year. This knowledge is based on how flu strains have evolved over time and over the past few flu seasons. Scientists then work to create the most effective vaccine to combat those selected strains. Hundreds of millions of doses are produced in a short period in order to provide vaccinations to key populations at the optimal time.

Because viruses, like the flu, evolve very quickly (especially in evolutionary time), this poses quite a challenge. Viruses mutate and replicate at a fast rate, so the vaccine developed to protect against last year's flu strain may not provide the protection needed against the coming year's strain. The evolution of these viruses means continued adaptions to ensure survival, including adaptations to survive previous vaccines.

Population Genetics

Recall that a **gene** for a particular character may have several **alleles**, or variants, that code for different traits associated with that character. For example, in the ABO blood type system in humans, three alleles determine the particular blood-type protein on the surface of red blood cells. Each individual in a population of diploid organisms can only carry two alleles for a particular gene, but more than two may be present in the individuals that make up the population. Mendel followed alleles as they were inherited from parent to offspring. In the early twentieth century, biologists in a field of study known as population genetics began to study how selective forces change a population through changes in allele and genotypic frequencies.

The **allele frequency** (or **gene frequency**) is the rate at which a specific allele appears within a population. Until now we have discussed evolution as a change in the characteristics of a population of organisms, but behind that phenotypic change is genetic change. In population genetics, the term evolution is defined as a change in the frequency of an allele in a population. Using the ABO blood type system as an example, the frequency of one of the alleles, I^A, is the number of copies of that allele divided by all the copies of the ABO gene in the population. For example, a study in Jordan found a frequency of I^A to be 26.1 percent (Hanania, Hassawi, & Irshaid, 2007). The I^B and I^0 alleles made up 13.4 percent and 60.5 percent of the alleles respectively, and all of the frequencies added up to 100 percent. A change in this frequency over time would constitute evolution in the population.

The allele frequency within a given population can change depending on environmental factors; therefore, certain alleles become more widespread than others during the process of natural selection. Natural selection can alter the

population's genetic makeup; for example, if a given allele confers a phenotype that allows an individual to better survive or have more offspring. Because many of those offspring will also carry the beneficial allele, and often the corresponding phenotype, they will have more offspring of their own that also carry the allele, thus, perpetuating the cycle. Over time, the allele will spread throughout the population. Some alleles will quickly become **fixed** in this way, meaning that every individual of the population will carry the allele, while detrimental mutations may be swiftly eliminated if derived from a dominant allele from the gene pool. The **gene pool** is the sum of all the alleles in a population.

Sometimes, allele frequencies within a population change randomly with no advantage to the population over existing allele frequencies. This phenomenon is called **genetic drift**. Natural selection and genetic drift usually occur simultaneously in populations and are not isolated events. It is hard to determine which process dominates because it is often nearly impossible to determine the cause of change in allele frequencies at each occurrence. An event that initiates an allele frequency change in an isolated part of the population, which is not typical of the original population, is called the founder effect. **Natural selection**, **random drift**, and **founder effects** can lead to significant changes in the genome of a population.

Hardy-Weinberg Principle of Equilibrium

In the early twentieth century, English mathematician Godfrey Hardy and German physician Wilhelm Weinberg stated the principle of equilibrium to describe the genetic makeup of a population. The theory, which later became known as the **Hardy-Weinberg principle of equilibrium**, states that a population's allele and genotype frequencies are

inherently stable— unless some kind of evolutionary force is acting upon the population, neither the allele nor the genotypic frequencies would change. The Hardy-Weinberg principle assumes conditions with **no mutations**, **migration**, **emigration**, or **selective pressure** for or against genotype, plus an **infinite population**. While no population can satisfy those conditions, the principle offers a useful model against which to compare real population changes.

Working under this theory, population geneticists represent different alleles as different variables in their mathematical models. The **variable p** represents the dominant allele in the population while the **variable q** represents the recessive allele. For example, when looking at Mendel's peas, the variable p represents the frequency of y alleles that confer the color yellow and the variable q represents the frequency of y alleles that confer the color green. If these are the only two possible alleles for a given locus in the population, **p + q = 1**. In other words, all the p alleles and all the q alleles make up all of the alleles for that locus that are found in the population.

But what ultimately interests most biologists is not the frequencies of different alleles, but the frequencies of the resulting genotypes, known as the population's genetic structure, from which scientists can surmise the distribution of phenotypes. If the phenotype is observed, only the genotype of the homozygous recessive alleles can be known; the calculations provide an estimate of the remaining genotypes. Since each individual carries two alleles per gene, if the allele frequencies (p and q) are known, predicting the frequencies of these genotypes is a simple mathematical calculation to determine the probability of getting these genotypes if two alleles are drawn at random from the gene pool. So in the above scenario, an individual pea plant could be pp (YY), and thus produce yellow peas; pq (Yy), also yellow; or qq (yy), and thus producing green peas (Figure 1). In other

words, the frequency of pp individuals is simply p^2; the frequency of pq individuals is 2pq; and the frequency of qq individuals is q^2. And, again, if p and q are the only two possible alleles for a given trait in the population, these genotype frequencies will sum to one: $\mathbf{p^2 + 2pq + q^2 = 1}$.

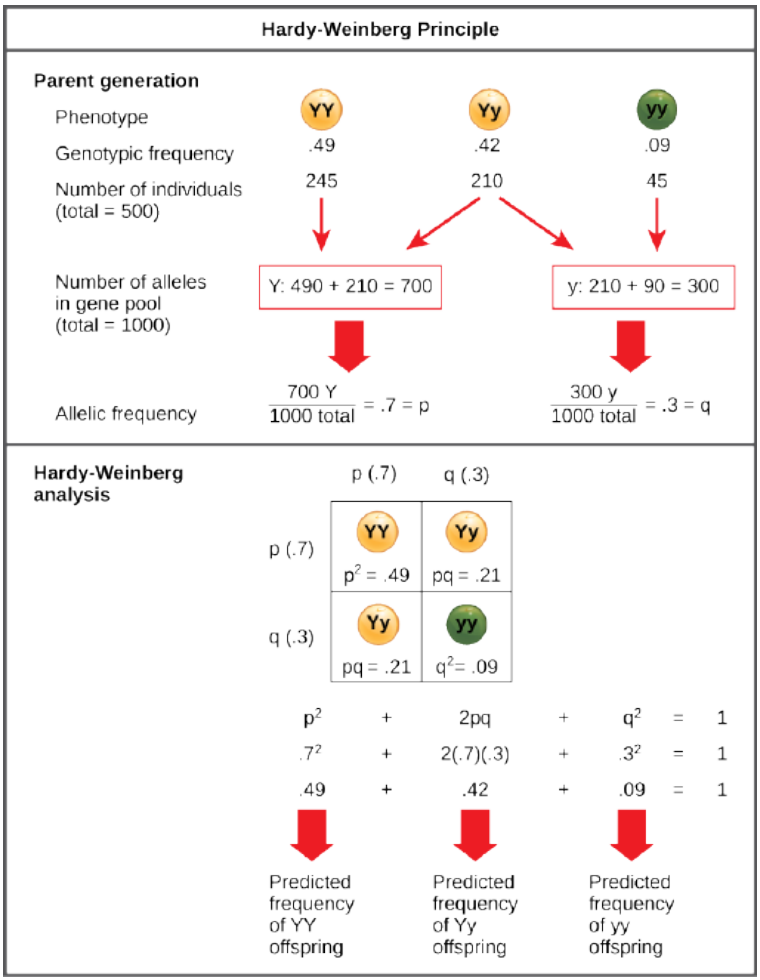


Figure 1: When populations are in the Hardy-Weinberg equilibrium, the allelic frequency is stable from generation to generation and the distribution of alleles can be determined from the Hardy-Weinberg equation. If the allelic frequency measured in the field differs from the predicted value, scientists can make inferences about what evolutionary forces are at play. (credit: "Hardy-Weinberg equilibrium" by OpenStax is licensed under CC BY 4.0)

In plants, violet flower color (V) is dominant over white (v). If p = 0.8 and q = 0.2 in a population of 500 plants, how many individuals would you expect to be homozygous dominant (VV), heterozygous (Vv), and homozygous recessive (vv)? How many plants would you expect to have violet flowers, and how many would have white flowers?

In theory, if a population is at equilibrium—that is, there are no evolutionary forces acting upon it—generation after generation would have the same gene pool and genetic structure, and these equations would all hold true all of the time. Of course, even Hardy and Weinberg recognized that no natural population is immune to evolution. Populations in nature are constantly changing in genetic makeup due to drift, mutation, possibly migration, and selection. As a result, the only way to determine the exact distribution of phenotypes in a population is to go out and count them. But the Hardy-Weinberg principle gives scientists a mathematical baseline of a non-evolving population to which they can compare evolving populations and thereby infer what evolutionary forces might be at play. If the frequencies of alleles or genotypes deviate from the value expected from the Hardy-Weinberg equation, then the population is evolving.

Summary

The modern synthesis of evolutionary theory grew out of the cohesion of Darwin's, Wallace's, and Mendel's thoughts on evolution and heredity, along with the more modern study of population genetics. It describes the evolution of populations and species, from small-scale changes among individuals to large-scale changes over paleontological time periods. To understand how organisms evolve, scientists can track populations' allele frequencies over time. If they differ from

generation to generation, scientists can conclude that the population is not in Hardy-Weinberg equilibrium, and is thus evolving.

REFERENCES

Hanania, S., Hassawi, D., and Irshaid, N. 2007. "Allele Frequency and Molecular Genotypes of ABO Blood Group System in a Jordanian Population," *Journal of Medical Sciences* 7: 51-58, doi:10.3923/jms.2007.51.58.

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Population Genetics

Individuals of a population often display different phenotypes, or express different alleles of a particular gene, referred to as **polymorphisms**. Populations with two or more variations of particular characteristics are called **polymorphic**. The distribution of phenotypes among individuals, known as the population variation, is influenced by a number of factors, including the population's genetic structure and the environment (Figure 1). Understanding the sources of phenotypic variation in a population is important for determining how a population will evolve in response to different evolutionary pressures.

Figure 1: The distribution of phenotypes in this litter of kittens illustrates population variation. (credit: Pieter Lanser. "cats in a basket" by OpenStax is licensed under CC BY 4.0)

Genetic Variance

Natural selection and some of the other evolutionary forces can only act on heritable traits, namely an organism's genetic code. Because alleles are passed from parent to offspring, those that confer beneficial traits or behaviors may be selected for, while deleterious alleles may be selected against. **Acquired traits**, for the most part, are not heritable. For example, if an athlete works out in the gym every day, building up muscle strength, the athlete's offspring will not necessarily grow up to be a bodybuilder. If there is a genetic basis for the ability to run fast, on the other hand, this may be passed to a child.

Heritability is the fraction of phenotype variation that can be attributed to genetic differences, or genetic variance,

among individuals in a population. The greater the hereditability of a population's phenotypic variation, the more susceptible it is to the evolutionary forces that act on heritable variation.

The diversity of alleles and genotypes within a population is called **genetic variance**. When scientists are involved in the breeding of a species, such as with animals in zoos and nature preserves, they try to increase a population's genetic variance to preserve as much of the phenotypic diversity as they can. This also helps reduce the risks associated with **inbreeding**, the mating of closely related individuals, which can have the undesirable effect of bringing together deleterious recessive mutations that can cause abnormalities and susceptibility to disease. For example, a disease that is caused by a rare, recessive allele might exist in a population, but it will only manifest itself when an individual carries two copies of the allele. Because the allele is rare in a normal, healthy population with unrestricted habitat, the chance that two carriers will mate is low, and even then, only 25 percent of their offspring will inherit the disease allele from both parents. While it is likely to happen at some point, it will not happen frequently enough for natural selection to be able to swiftly eliminate the allele from the population, and as a result, the allele will be maintained at low levels in the gene pool. However, if a family of carriers begins to interbreed with each other, this will dramatically increase the likelihood of two carriers mating and eventually producing diseased offspring, a phenomenon known as inbreeding depression.

Changes in allele frequencies that are identified in a population can shed light on how it is evolving. In addition to natural selection, there are other evolutionary forces that could be in play: genetic drift, gene flow, mutation, nonrandom mating, and environmental variances.

Genetic Drift

The theory of natural selection stems from the observation that some individuals in a population are more likely to survive longer and have more offspring than others; thus, they will pass on more of their genes to the next generation. A big, powerful male gorilla, for example, is much more likely than a smaller, weaker one to become the population's silverback, the pack's leader who mates far more than the other males of the group. The pack leader will father more offspring, who share half of his genes, and are likely to also grow bigger and stronger like their father. Over time, the genes for bigger size will increase in frequency in the population, and the population will, as a result, grow larger on average. That is, this would occur if this particular selection pressure, or driving selective force, were the only one acting on the population. In other examples, better camouflage or a stronger resistance to drought might pose a selection pressure.

Another way a population's allele and genotype frequencies can change is **genetic drift** (Figure 2), which is simply the effect of chance. By chance, some individuals will have more offspring than others—not due to an advantage conferred by some genetically-encoded trait, but just because one male happened to be in the right place at the right time (when the receptive female walked by) or because the other one happened to be in the wrong place at the wrong time (when a fox was hunting).

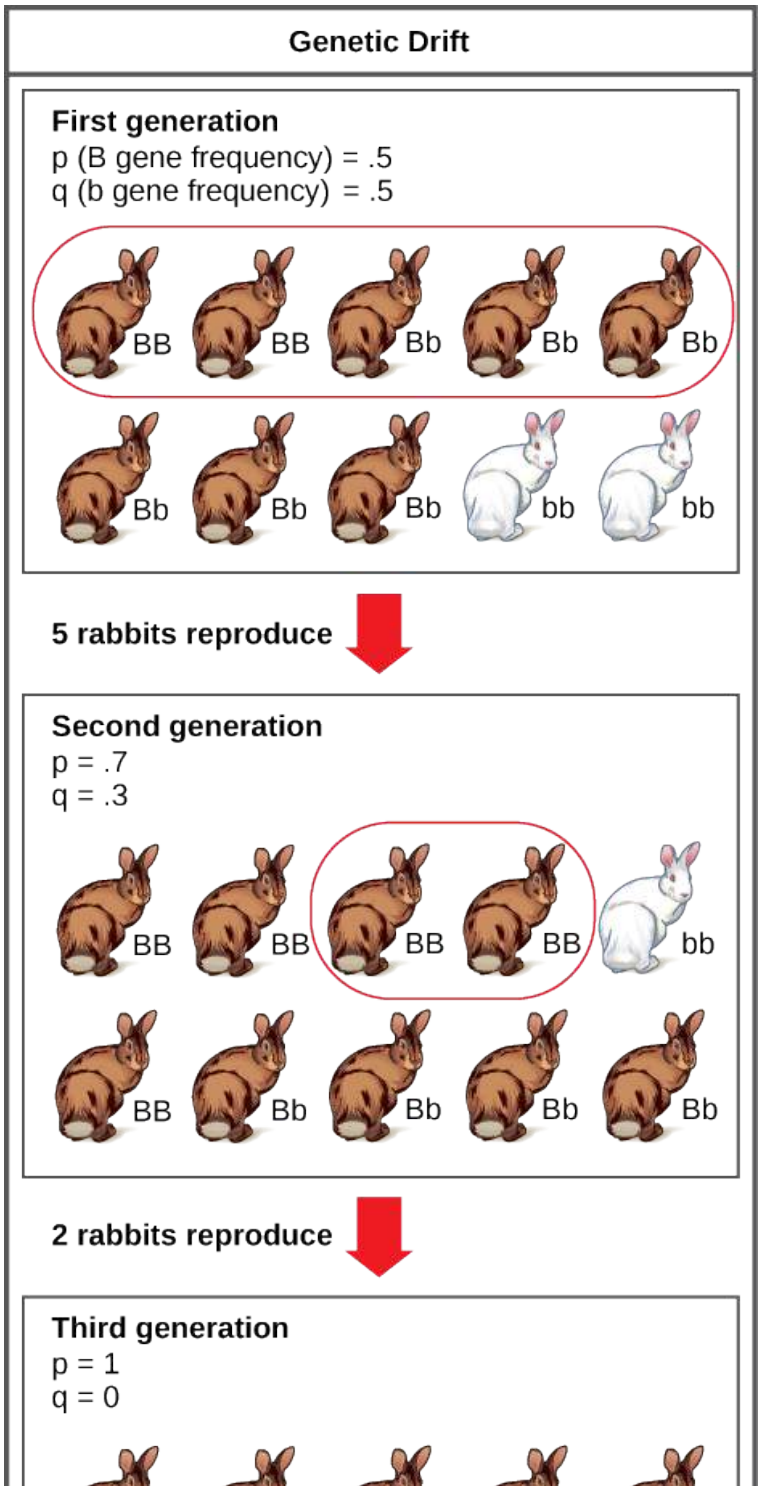

Genetic Drift
First generation
p (B gene frequency) = .5
q (b gene frequency) = .5
BB
BB
Bb
Bb
Bb
Bb
Bb
Bb
bb
bb
5 rabbits reproduce
Second generation
p = .7
q = .3
BB
BB
BB
BB
bb
BB
Bb
Bb
Bb
Bb
2 rabbits reproduce
Third generation
p = 1
q = 0

Figure 2: Genetic drift in a population can lead to the elimination of an allele from a population by chance. In this example, rabbits with the brown coat color allele (B) are dominant over rabbits with the white coat color allele (b). In the first generation, the two alleles occur with equal frequency in the population, resulting in p and q values of .5. Only half of the individuals reproduce, resulting in a second generation with p and q values of .7 and .3, respectively. Only two individuals in the second generation reproduce, and by chance these individuals are homozygous dominant for brown coat color. As a result, in the third generation the recessive b allele is lost. (credit: "genetic drift in fur color" by OpenStax is licensed under CC BY 4.0)

Do you think genetic drift would happen more quickly on an island or on the mainland? Small populations are more susceptible to the forces of genetic drift. Large populations, on the other hand, are buffered against the effects of chance. If one individual of a population of 10 individuals happens to die at a young age before it leaves any offspring to the next generation, all of its genes—1/10 of the population's gene pool—will be suddenly lost. In a population of 100, that's only 1 percent of the overall gene pool; therefore, it is much less impactful on the population's genetic structure.

Genetic drift can also be magnified by natural events, such as a natural disaster that kills—at random—a large portion of the population. Known as the **bottleneck effect**, it results in a large portion of the genome suddenly being wiped out (Figure 3). In one fell swoop, the genetic structure of the survivors becomes the genetic structure of the entire

population, which may be very different from the pre-disaster population.

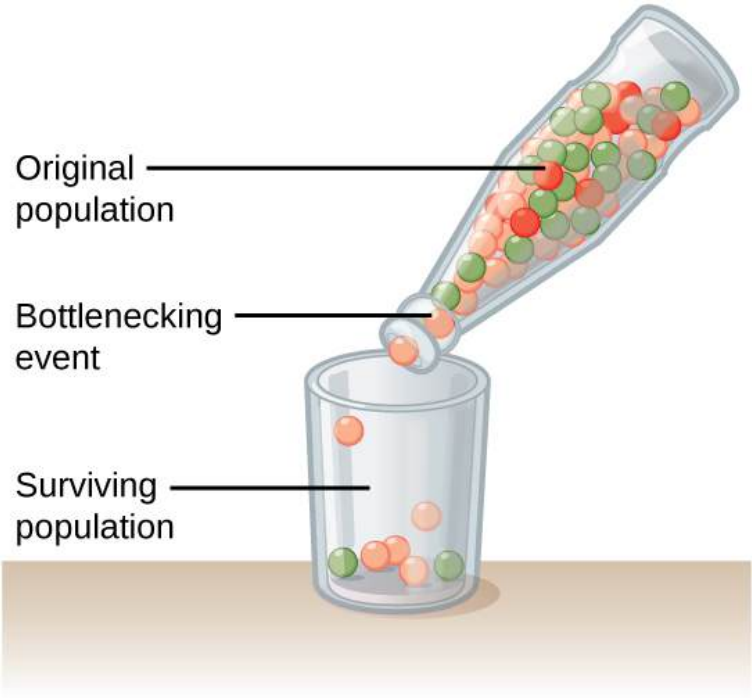


Figure 3: A chance event or catastrophe can reduce the genetic variability within a population. (credit: "the bottleneck effect" by OpenStax is licensed under CC BY 4.0)

Another scenario in which populations might experience a strong influence of genetic drift is if some portion of the population leaves to start a new population in a new location or if a population gets divided by a physical barrier of some kind. In this situation, those individuals are unlikely to be representative of the entire population, which results in the **founder effect**. The founder effect occurs when the genetic structure changes to match that of the new population's

founding fathers and mothers. The founder effect is believed to have been a key factor in the genetic history of the Afrikaner population of Dutch settlers in South Africa, as evidenced by mutations that are common in Afrikaners but rare in most other populations. This is likely due to the fact that a higher-than-normal proportion of the founding colonists carried these mutations. As a result, the population expresses unusually high incidences of Huntington's disease (HD) and Fanconi anemia (FA), a genetic disorder known to cause blood marrow and congenital abnormalities—even cancer (Tipping et al., 2001).

Watch this short video to learn more about the founder and bottleneck effects.

One or more interactive elements has been excluded from this version of the text. You can view them online here:

https://openoregon.pressbooks.pub/mhccmajorsbio/?p=1314#oembed-1

Source URL: https://www.youtube.com/watch?v=hEYV9WEvwaI&feature=youtu.be"

Scientific Method Practice: Testing the Bottleneck Effect

Question: How do natural disasters affect the genetic structure of a population?

Background: When much of a population is suddenly wiped out by an earthquake or hurricane, the individuals that survive the event are usually a random sampling of the original group. As a result, the genetic makeup of the

population can change dramatically. This phenomenon is known as the bottleneck effect.

Hypothesis: Repeated natural disasters will yield different population genetic structures; therefore, each time this experiment is run, the results will vary.

Test the hypothesis: Count out the original population using different colored beads. For example, red, blue, and yellow beads might represent red, blue, and yellow individuals. After recording the number of each individual in the original population, place them all in a bottle with a narrow neck that will only allow a few beads out at a time. Then, pour 1/3 of the bottle's contents into a bowl. This represents the surviving individuals after a natural disaster kills a majority of the population. Count the number of the different colored beads in the bowl, and record it. Then, place all of the beads back in the bottle and repeat the experiment four more times.

Analyze the data: Compare the five populations that resulted from the experiment. Do the populations all contain the same number of different colored beads, or do they vary? Remember, these populations all came from the same exact parent population.

Form a conclusion: Most likely, the five resulting populations will differ quite dramatically. This is because natural disasters are not selective—they kill and spare individuals at random. Now think about how this might affect a real population. What happens when a hurricane hits the Mississippi Gulf Coast? How do the seabirds that live on the beach fare?

Gene Flow

Another important evolutionary force is **gene flow**: the flow of alleles in and out of a population due to the migration of individuals or gametes (Figure 4). While some populations

are fairly stable, others experience more flux. Many plants, for example, send their pollen far and wide, by wind or by bird, to pollinate other populations of the same species some distance away. Even a population that may initially appear to be stable, such as a pride of lions, can experience its fair share of **immigration** and **emigration** as developing males leave their mothers to seek out a new pride with genetically unrelated females. This variable flow of individuals in and out of the group not only changes the gene structure of the population, but it can also introduce new genetic variation to populations in different geological locations and habitats.

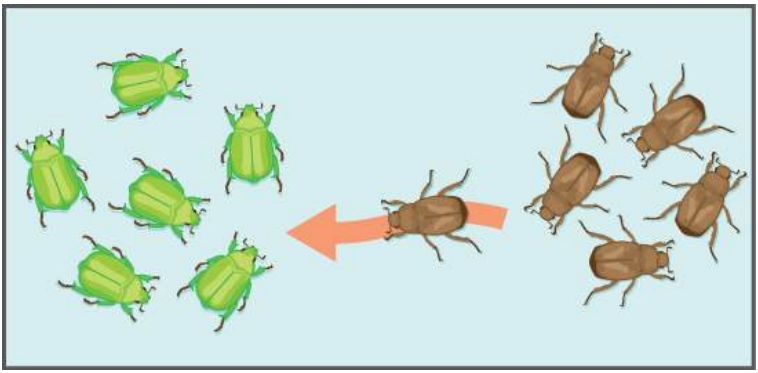


Figure 4: Gene flow can occur when an individual travels from one geographic location to another. (credit: "Gene flow" by OpenStax is licensed under CC BY 4.0)

Mutation

Mutations are changes to an organism's DNA and are an important driver of diversity in populations. Species evolve because of the accumulation of mutations that occur over time. The appearance of new mutations is the most common way to introduce novel genotypic and phenotypic variance. Some mutations are unfavorable or harmful and are quickly

eliminated from the population by natural selection. Others are beneficial and will spread through the population. Whether or not a mutation is beneficial or harmful is determined by whether it helps an organism survive to sexual maturity and reproduce. Some mutations do not do anything and can linger, unaffected by natural selection, in the genome. Some can have a dramatic effect on a gene and the resulting phenotype.

Nonrandom Mating

There are many reasons **nonrandom mating** occurs. One reason is simple mate choice; for example, female peahens may prefer peacocks with bigger, brighter tails. Traits that lead to more matings for an individual become selected for by natural selection. One common form of mate choice, called **assortative mating**, is an individual's preference to mate with partners who are phenotypically similar to themselves.

Another cause of nonrandom mating is physical location. This is especially true in large populations spread over large geographic distances where not all individuals will have equal access to one another. Some might be miles apart through woods or over rough terrain, while others might live immediately nearby.

Environmental Variance

Genes are not the only players involved in determining population variation. Phenotypes are also influenced by other factors, such as the **environment** (Figure 5). A beachgoer is likely to have darker skin than a city dweller, for example, due to regular exposure to the sun, an environmental factor. Some major characteristics, such as sex, are determined by the environment for some species.

For example, some turtles and other reptiles have temperature-dependent sex determination (TSD). TSD means that individuals develop into males if their eggs are incubated within a certain temperature range, or females at a different temperature range.

Figure 5: The sex of the American alligator (Alligator mississippiensis) is determined by the temperature at which the eggs are incubated. Eggs incubated at 30°C produce females, and eggs incubated at 33°C produce males. (credit: Steve Hillebrand, USFWS. "Alligator mississippiens" by OpenStax is licensed under CC BY 4.0)

Geographic separation between populations can lead to differences in the phenotypic variation between those populations. Such geographical variation is seen between most populations and can be significant. One type of geographic variation, called a **cline**, can be seen as populations of a given species vary gradually across an ecological gradient. Species of warm-blooded animals, for example, tend to have larger bodies in the cooler climates closer to the earth's poles, allowing them to better conserve

heat. This is considered a **latitudinal cline**. Alternatively, flowering plants tend to bloom at different times depending on where they are along the slope of a mountain, known as an **altitudinal cline**.

If there is gene flow between the populations, the individuals will likely show gradual differences in phenotype along the cline. Restricted gene flow, on the other hand, can lead to abrupt differences, even speciation.

Summary

Both genetic and environmental factors can cause phenotypic variation in a population. Different alleles can confer different phenotypes, and different environments can also cause individuals to look or act differently. Only those differences encoded in an individual's genes, however, can be passed to its offspring and, thus, be a target of natural selection. Natural selection works by selecting for alleles that confer beneficial traits or behaviors while selecting against those for deleterious qualities. Genetic drift stems from the chance occurrence that some individuals in the germline have more offspring than others. When individuals leave or join the population, allele frequencies can change as a result of gene flow. Mutations to an individual's DNA may introduce new variation into a population. Allele frequencies can also be altered when individuals do not randomly mate with others in the group.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Tipping, A.J., et al. 2001. "Molecular and Genealogical

Evidence for a Founder Effect in Fanconi Anemia Families of the Afrikaner Population of South Africa," *PNAS* 98, no. 10: 5734-5739, doi: 10.1073/pnas.091402398.

Adaptive Evolution

Natural selection only acts on the population's heritable traits: selecting for beneficial alleles and thus increasing their frequency in the population, while selecting against deleterious alleles and thereby decreasing their frequency—a process known as **adaptive evolution**. Natural selection does not act on individual alleles, however, but on entire organisms. An individual may carry a very beneficial genotype with a resulting phenotype that, for example, increases the ability to reproduce (**fecundity**), but if that same individual also carries an allele that results in a fatal childhood disease, that fecundity phenotype will not be passed on to the next generation because the individual will not live to reach reproductive age. Natural selection acts at the level of the individual; it selects for individuals with greater contributions to the gene pool of the next generation, known as an organism's **evolutionary (Darwinian) fitness**.

Fitness is often quantifiable and is measured by scientists in the field. However, it is not the absolute fitness of an individual that counts, but rather how it compares to the other organisms in the population. This concept, called **relative fitness**, allows researchers to determine which individuals are contributing additional offspring to the next generation, and thus, how the population might evolve.

There are several ways selection can affect population variation: stabilizing selection, directional selection,

diversifying selection, frequency-dependent selection, and sexual selection. As natural selection influences the allele frequencies in a population, individuals can either become more or less genetically similar and the phenotypes displayed can become more similar or more disparate.

Stabilizing Selection

If natural selection favors an average phenotype, selecting against extreme variation, the population will undergo **stabilizing selection** (Figure 1). In a population of mice that live in the woods, for example, natural selection is likely to favor individuals that best blend in with the forest floor and are less likely to be spotted by predators. Assuming the ground is a fairly consistent shade of brown, those mice whose fur is most closely matched to that color will be most likely to survive and reproduce, passing on their genes for their brown coat. Mice that carry alleles that make them a bit lighter or a bit darker will stand out against the ground and be more likely to fall victim to predation. As a result of this selection, the population's genetic variance will decrease.

Directional Selection

When the environment changes, populations will often undergo **directional selection** (Figure 1), which selects for phenotypes at one end of the spectrum of existing variation. A classic example of this type of selection is the evolution of the peppered moth in eighteenth- and nineteenth-century England. Prior to the Industrial Revolution, the moths were predominately light in color, which allowed them to blend in with the light-colored trees and lichens in their environment. But as soot began spewing from factories, the trees became darkened, and the light-colored moths became easier for predatory birds to spot. Over time, the frequency of the

melanic form of the moth increased because they had a higher survival rate in habitats affected by air pollution because their darker coloration blended with the sooty trees. Similarly, the hypothetical mouse population may evolve to take on a different coloration if something were to cause the forest floor where they live to change color. The result of this type of selection is a shift in the population's genetic variance toward the new, fit phenotype.

Diversifying Selection

Sometimes two or more distinct phenotypes can each have their advantages and be selected for by natural selection, while the intermediate phenotypes are, on average, less fit. Known as **diversifying selection** (Figure 1), this is seen in many populations of animals that have multiple male forms. Large, dominant alpha males obtain mates by brute force, while small males can sneak in for furtive copulations with the females in an alpha male's territory. In this case, both the alpha males and the "sneaking" males will be selected for, but medium-sized males, which can't overtake the alpha males and are too big to sneak copulations, are selected against. Diversifying selection can also occur when environmental changes favor individuals on either end of the phenotypic spectrum. Imagine a population of mice living at the beach where there is light-colored sand interspersed with patches of tall grass. In this scenario, light-colored mice that blend in with the sand would be favored, as well as dark-colored mice that can hide in the grass. Medium-colored mice, on the other hand, would not blend in with either the grass or the sand, and would thus be more likely to be eaten by predators. The result of this type of selection is increased genetic variance as the population becomes more diverse.

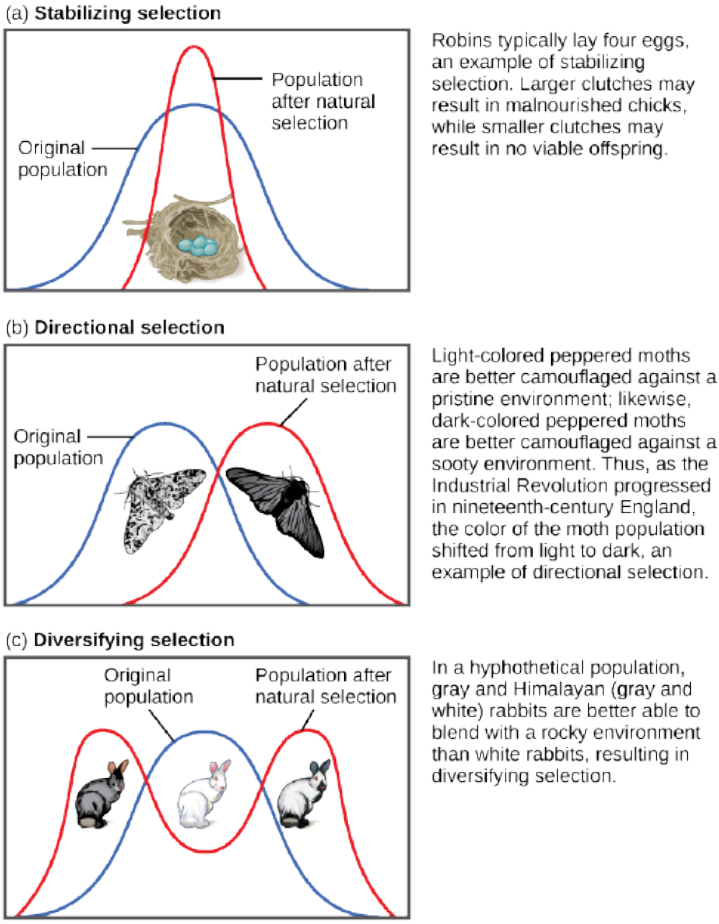


Figure 1: Different types of natural selection can impact the distribution of phenotypes within a population. In (a) stabilizing selection, an average phenotype is favored. In (b) directional selection, a change in the environment shifts the spectrum of phenotypes observed. In (c) diversifying selection, two or more extreme phenotypes are selected for, while the average phenotype is selected against. (credit: "types of selection" by OpenStax is licensed under CC BY 4.0)

In recent years, factories have become cleaner, and less soot is released into the environment. What impact do you think this has had on the distribution of moth color in the population?

Frequency-dependent Selection

Another type of selection, called **frequency-dependent selection**, favors phenotypes that are either **common (positive frequency-dependent selection)** or **rare (negative frequency-dependent selection)**. An interesting example of this type of selection is seen in a unique group of lizards of the Pacific Northwest. Male common side-blotched lizards come in three throat-color patterns: orange, blue, and yellow. Each of these forms has a different reproductive strategy: orange males are the strongest and can fight other males for access to their females; blue males are medium-sized and form strong pair bonds with their mates; and yellow males (Figure 2) are the smallest, and look a bit like females, which allows them to sneak copulations. Like a game of rock-paper-scissors, orange beats blue, blue beats yellow, and yellow beats orange in the competition for females. That is, the big, strong orange males can fight off the blue males to mate with the blue's pair-bonded females, the blue males are successful at guarding their mates against yellow sneaker males, and the yellow males can sneak copulations from the potential mates of the large, polygynous orange males.

Figure 2: A yellow-throated side-blotched lizard is smaller than either the blue-throated or orange-throated males and appears a bit like the females of the species, allowing it to sneak copulations. (credit: "tinyfroglet"/Flickr. "tinyfroglet" by OpenStax is licensed under CC BY 4.0)

In this scenario, orange males will be favored by natural selection when the population is dominated by blue males, blue males will thrive when the population is mostly yellow males, and yellow males will be selected for when orange males are the most populous. As a result, populations of side-blotched lizards cycle in the distribution of these phenotypes—in one generation, orange might be

predominant, and then yellow males will begin to rise in frequency. Once yellow males make up a majority of the population, blue males will be selected for. Finally, when blue males become common, orange males will once again be favored.

Negative frequency-dependent selection serves to increase the population's genetic variance by selecting for rare phenotypes, whereas positive frequency-dependent selection usually decreases genetic variance by selecting for common phenotypes.

Sexual Selection

Males and females of certain species are often quite different from one another in ways beyond the reproductive organs. Males are often larger, for example, and display many elaborate colors and adornments, like the peacock's tail, while females tend to be smaller and duller in decoration. Such differences are known as **sexual dimorphisms** (Figure 3), which arise from the fact that in many populations, particularly animal populations, there is more variance in the reproductive success of the males than there is of the females. That is, some males—often the bigger, stronger, or more decorated males—get the vast majority of the total matings, while others receive none. This can occur because the males are better at fighting off other males, or because females will choose to mate with the bigger or more decorated males. In either case, this variation in reproductive success generates a strong selection pressure among males to get those matings, resulting in the evolution of bigger body size and elaborate ornaments to get the females' attention. Females, on the other hand, tend to get a handful of selected matings; therefore, they are more likely to select more desirable males.

Sexual dimorphism varies widely among species, of course,

and some species are even sex-role reversed. In such cases, females tend to have a greater variance in their reproductive success than males and are correspondingly selected for the bigger body size and elaborate traits usually characteristic of males.

Figure 3: Sexual dimorphism is observed in (a) peacocks and peahens, (b) Argiope appensa spiders (the female spider is the large one), and in (c) wood ducks. (credit "spiders": modification of work by "Sanba38"/Wikimedia Commons; credit "duck": modification of work by Kevin Cole. "this image" by OpenStax is licensed under CC BY 4.0)

The selection pressures on males and females to obtain matings is known as **sexual selection**; it can result in the development of secondary sexual characteristics that do not benefit the individual's likelihood of survival but help to maximize its reproductive success. Sexual selection can be so strong that it selects for traits that are actually detrimental to the individual's survival. Think, once again, about the peacock's tail. While it is beautiful and the male with the largest, most colorful tail is more likely to win the female, it is not the most practical appendage. In addition to being more visible to predators, it makes the males slower in their attempted escapes. There is some evidence that this risk, in fact, is why females like the big tails in the first place. The speculation is that large tails carry risk, and only the best

males survive that risk: the bigger the tail, the more fit the male. This idea is known as the **handicap principle**.

The **good genes hypothesis** states that males develop these impressive ornaments to show off their efficient metabolism or their ability to fight disease. Females then choose males with the most impressive traits because it signals their genetic superiority, which they will then pass on to their offspring. Though it might be argued that females should not be picky because it will likely reduce their number of offspring, if better males father more fit offspring, it may be beneficial. Fewer, healthier offspring may increase the chances of survival more than many, weaker offspring. In 1915, biologist Ronald Fisher proposed another model of sexual selection: the **Fisherian runaway model**, which suggests that selection of certain traits is a result of sexual preference.

In both the handicap principle and the good genes hypothesis, the trait is said to be an honest signal of the males' quality, thus giving females a way to find the fittest mates— males that will pass the best genes to their offspring.

No Perfect Organism

Natural selection is a driving force in evolution and can generate populations that are better adapted to survive and successfully reproduce in their environments. But natural selection cannot produce the perfect organism. Natural selection can only select on existing variation in the population; it does not create anything from scratch. Thus, it is limited by a population's existing genetic variance and whatever new alleles arise through mutation and gene flow.

Natural selection is also limited because it works at the level of individuals, not alleles, and some alleles are linked

due to their physical proximity in the genome, making them more likely to be passed on together (**linkage disequilibrium**). Any given individual may carry some beneficial alleles and some unfavorable alleles. It is the net effect of these alleles, or the organism's fitness, upon which natural selection can act. As a result, good alleles can be lost if they are carried by individuals that also have several overwhelmingly bad alleles; likewise, bad alleles can be kept if they are carried by individuals that have enough good alleles to result in an overall fitness benefit.

Furthermore, natural selection can be constrained by the relationships between different polymorphisms. One morph may confer a higher fitness than another, but may not increase in frequency due to the fact that going from the less beneficial to the more beneficial trait would require going through a less beneficial phenotype. Think back to the mice that live at the beach. Some are light-colored and blend in with the sand, while others are dark and blend in with the patches of grass. The dark-colored mice may be, overall, more fit than the light-colored mice, and at first glance, one might expect the light-colored mice be selected for a darker coloration. But remember that the intermediate phenotype, a medium-colored coat, is very bad for the mice—they cannot blend in with either the sand or the grass and are more likely to be eaten by predators. As a result, the light-colored mice would not be selected for a dark coloration because those individuals that began moving in that direction (began being selected for a darker coat) would be less fit than those that stayed light.

Finally, it is important to understand that not all evolution is **adaptive**. While natural selection selects the fittest individuals and often results in a more fit population overall, other forces of evolution, including genetic drift and gene flow, often do the opposite: introducing deleterious alleles to the population's gene pool. Evolution has no purpose—it

is not changing a population into a preconceived ideal. It is simply the sum of the various forces described in this chapter and how they influence the genetic and phenotypic variance of a population.

Summary

Because natural selection acts to increase the frequency of beneficial alleles and traits while decreasing the frequency of deleterious qualities, it is adaptive evolution. Natural selection acts at the level of the individual, selecting for those that have a higher overall fitness compared to the rest of the population. If the fit phenotypes are those that are similar, natural selection will result in stabilizing selection, and an overall decrease in the population's variation. Directional selection works to shift a population's variance toward a new, fit phenotype, as environmental conditions change. In contrast, diversifying selection results in increased genetic variance by selecting for two or more distinct phenotypes.

Other types of selection include frequency-dependent selection, in which individuals with either common (positive frequency-dependent selection) or rare (negative frequency-dependent selection) are selected for. Finally, sexual selection results from the fact that one sex has more variance in the reproductive success than the other. As a result, males and females experience different selective pressures, which can often lead to the evolution of phenotypic differences, or sexual dimorphisms, between the two.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020.

https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

EVOLUTIONARY RELATIONSHIPS

Figure 1: The life of a bee is very different from the life of a flower, but the two organisms are related. Both are members the domain Eukarya and have cells containing many similar organelles, genes, and proteins. (credit: modification of work by John Beetham. "bee and Echinacea" by OpenStax is licensed under CC BY 4.0)

This bee and *Echinacea* flower (Figure 1) could not look more different, yet they are related, as are all living organisms on Earth. By following pathways of similarities and changes—both visible and genetic—scientists seek to map

the evolutionary past of how life developed from single-celled organisms to the tremendous collection of creatures that have germinated, crawled, floated, swam, flown, and walked on this planet.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Organizing Life on Earth

In scientific terms, the evolutionary history and relationship of an organism or group of organisms is called its **phylogeny**. A phylogeny describes the relationships of an organism, such as from which organisms it is thought to have evolved, to which species it is most closely related, and so forth. Phylogenetic relationships provide information on shared ancestry but not necessarily on how organisms are similar or different.

Phylogenetic Trees

Scientists use a tool called a **phylogenetic tree** to show the evolutionary pathways and connections among organisms. A phylogenetic tree is a diagram used to reflect evolutionary relationships among organisms or groups of organisms. Scientists consider phylogenetic trees to be a hypothesis of the evolutionary past since one cannot go back to confirm the proposed relationships. In other words, a "tree of life" can be constructed to illustrate when different organisms evolved and to show the relationships among different organisms (Figure 1).

Unlike a **taxonomic classification diagram**, a phylogenetic tree can be read like a map of evolutionary history. Many phylogenetic trees have a single lineage at the base representing a common ancestor. Scientists call

such trees **rooted**, which means there is a single ancestral lineage (typically drawn from the bottom or left) to which all organisms represented in the diagram relate. Notice in the rooted phylogenetic tree that the three domains— Bacteria, Archaea, and Eukarya—diverge from a single point and branch off. The small branch that plants and animals (including humans) occupy in this diagram shows how recent and minuscule these groups are compared with other organisms. **Unrooted trees** don't show a common ancestor but do show relationships among species.

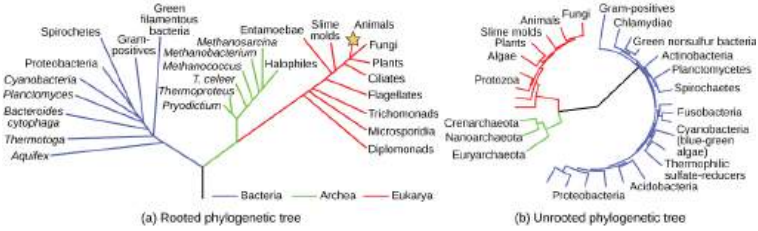


Figure 1: Both of these phylogenetic trees show the relationship of the three domains of life—Bacteria, Archaea, and Eukarya—but the (a) rooted tree attempts to identify when various species diverged from a common ancestor while the (b) unrooted tree does not. (credit a: modification of work by Eric Gaba. "three domains of life" by OpenStax is licensed under CC BY 4.0)

In a rooted tree, the branching indicates evolutionary relationships (Figure 2). The point where a split occurs, called a **branch point**, represents where a single lineage evolved into a distinct new one. A lineage that evolved early from the root and remains unbranched is called **basal taxon**. When two lineages stem from the same branch point, they are called **sister taxa**. A branch with more than two lineages is called a **polytomy** and serves to illustrate where scientists have not definitively determined all of the relationships. It is important to note that although sister taxa and polytomy do share an ancestor, it does not mean that the groups of

organisms split or evolved from each other. Organisms in two taxa may have split apart at a specific branch point, but neither taxa gave rise to the other.

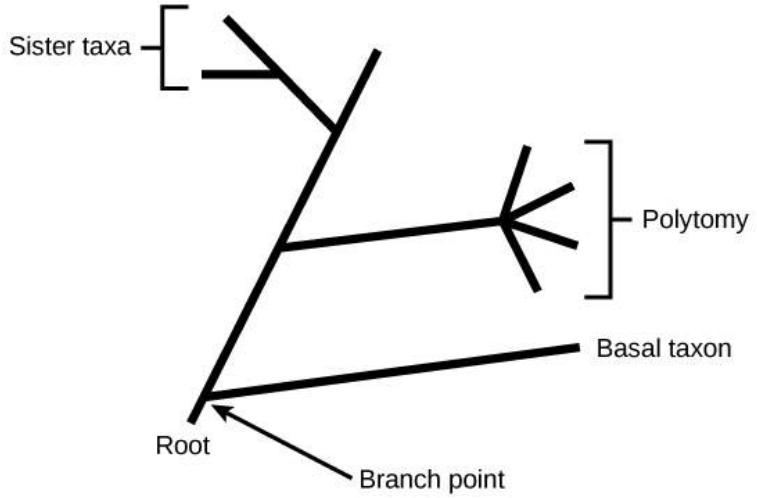


Figure 2: The root of a phylogenetic tree indicates that an ancestral lineage gave rise to all organisms on the tree. A branch point indicates where two lineages diverged. A lineage that evolved early and remains unbranched is a basal taxon. When two lineages stem from the same branch point, they are sister taxa. A branch with more than two lineages is a polytomy. (credit: "rooted phylogenetic tree" by OpenStax is licensed under CC BY 4.0)

The diagrams above can serve as a pathway to understanding evolutionary history. The pathway can be traced from the origin of life to any individual species by navigating through the evolutionary branches between the two points. Also, by starting with a single species and tracing back towards the "trunk" of the tree, one can discover that species' ancestors, as well as where lineages share a common ancestry. In addition, the tree can be used to study entire groups of organisms.

Another point to mention on phylogenetic tree structure is that rotation at branch points does not change the information. For example, if a branch point was rotated and the taxon order changed, this would not alter the information because the evolution of each taxon from the branch point was independent of the other.

Many disciplines within the study of biology contribute to understanding how past and present life evolved over time; these disciplines together contribute to building, updating, and maintaining the "tree of life." Information is used to organize and classify organisms based on evolutionary relationships in a scientific field called systematics. Data may be collected from fossils, from studying the structure of body parts or molecules used by an organism, and by DNA analysis. By combining data from many sources, scientists can put together the phylogeny of an organism; since phylogenetic trees are hypotheses, they will continue to change as new types of life are discovered and new information is learned.

Limitations of Phylogenetic Trees

It may be easy to assume that more closely related organisms look more alike, and while this is often the case, it is not always true. If two closely related lineages evolved under significantly varied surroundings or after the evolution of a major new adaptation, it is possible for the two groups to appear more different than other groups that are not as closely related. For example, the phylogenetic tree in Figure 3 shows that lizards and rabbits both have amniotic eggs, whereas frogs do not; yet lizards and frogs appear more similar than lizards and rabbits.

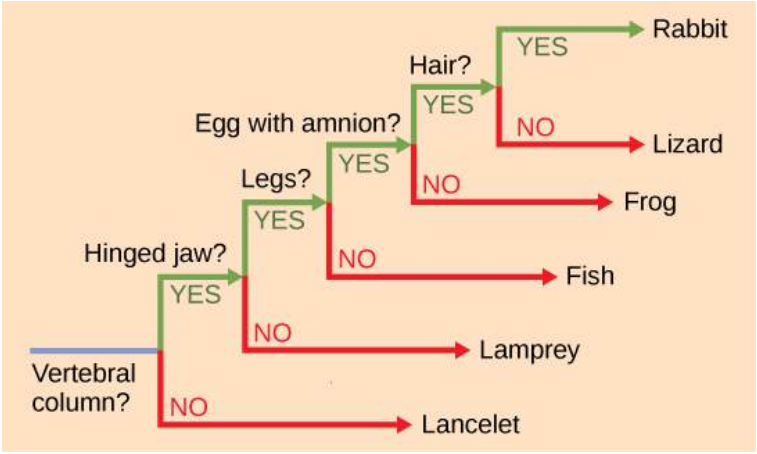


Figure 3: This ladder-like phylogenetic tree of vertebrates is rooted by an organism that lacked a vertebral column. At each branch point, organisms with different characters are placed in different groups based on the characteristics they share. (credit: "ladder-like phylogenetic tree" by OpenStax is licensed under CC BY 4.0

Another aspect of phylogenetic trees is that, unless otherwise indicated, the branches do not account for length of time, only the evolutionary order. In other words, the length of a branch does not typically mean more time passed, nor does a short branch mean less time passed—unless specified on the diagram. For example, in Figure 3, the tree does not indicate how much time passed between the evolution of amniotic eggs and hair. What the tree does show is the order in which things took place. Again using Figure 3, the tree shows that the oldest trait is the vertebral column, followed by hinged jaws, and so forth. Remember that any phylogenetic tree is a part of the greater whole, and like a real tree, it does not grow in only one direction after a new branch develops. So, for the organisms in Figure 3, just because a vertebral column evolved does not mean that invertebrate evolution ceased, it only means that a new

branch formed. Also, groups that are not closely related, but evolve under similar conditions, may appear more phenotypically similar to each other than to a close relative.

The Levels of Classification

Taxonomy (which literally means "arrangement law") is the science of classifying organisms to construct internationally shared classification systems with each organism placed into more and more inclusive groupings. Think about how a grocery store is organized. One large space is divided into departments, such as produce, dairy, and meats. Then each department further divides into aisles, then each aisle into categories and brands, and then finally a single product. This organization from larger to smaller, more specific categories is called a hierarchical system.

The **taxonomic classification system** (also called the **Linnaean system** after its inventor, Carl Linnaeus, a Swedish botanist, zoologist, and physician) uses a hierarchical model. Moving from the point of origin, the groups become more specific, until one branch ends as a single species. For example, after the common beginning of all life, scientists divide organisms into three large categories called a **domain**: Bacteria, Archaea, and Eukarya. Within each domain is a second category called a **kingdom**. After kingdoms, the subsequent categories of increasing specificity are: **phylum**, **class**, **order**, **family**, **genus**, and **species** (Figure 4).

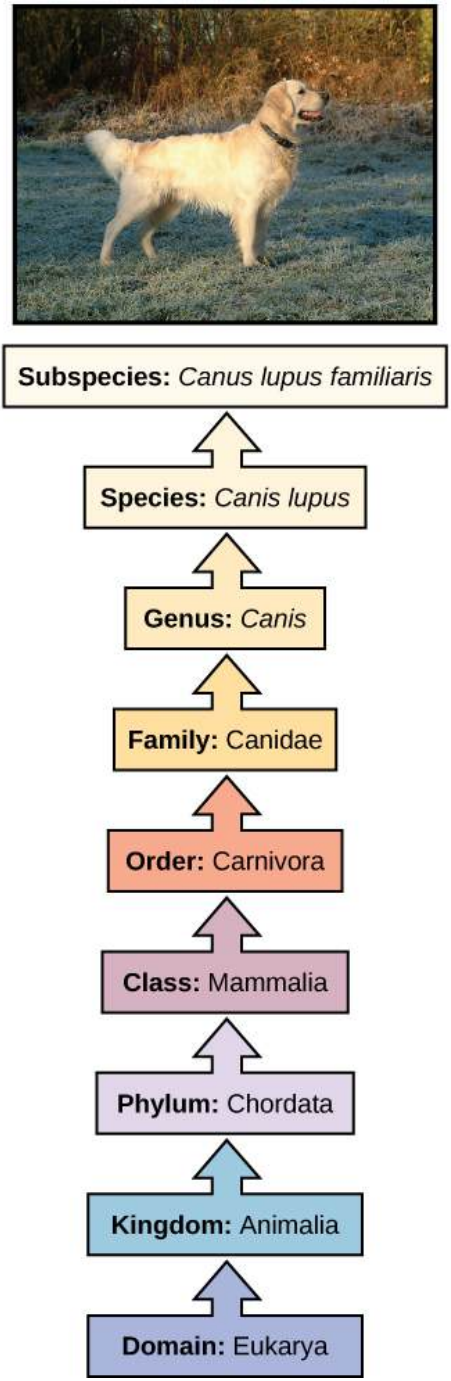

Subspecies: Canus lupus familiaris
Species: Canis lupus
Genus: Canis
Family: Canidae
Order: Carnivora
Class: Mammalia
Phylum: Chordata
Kingdom: Animalia
Domain: Eukarya

Figure 4: The taxonomic classification system uses a hierarchical model to organize living organisms into increasingly specific categories. The common dog, Canis lupus familiaris, is a subspecies of Canis lupus, which also includes the wolf and dingo. (credit "dog": modification of work by Janneke Vreugdenhil. "dog" by OpenStax is licensed under CC BY 4.0)

The **kingdom Animalia** stems from the Eukarya domain. For the common dog, the classification levels would be as shown in Figure 4. Therefore, the full name of an organism technically has eight terms. For the dog, it is: Eukarya, Animalia, Chordata, Mammalia, Carnivora, Canidae, *Canis,* and *lupus*. Notice that each name is capitalized except for species, and the genus and species names are italicized. Scientists generally refer to an organism only by its genus and species, which is its two-word scientific name, in what is called **binomial nomenclature**. Therefore, the scientific name of the dog is *Canis lupus*. The name at each level is also called a **taxon**. In other words, dogs are in order Carnivora. Carnivora is the name of the taxon at the order level; Canidae is the taxon at the family level, and so forth. Organisms also have a common name that people typically use, in this case, dog. Note that the dog is additionally a subspecies: the "*familiaris*" in *Canis lupus familiaris.* Subspecies are members of the same species that are capable of mating and reproducing viable offspring, but they are considered separate subspecies due to geographic or behavioral isolation or other factors.

Figure 5 shows how the levels move toward specificity with other organisms. Notice how the dog shares a domain with the widest diversity of organisms, including plants and butterflies. At each sublevel, the organisms become more similar because they are more closely related. Historically, scientists classified organisms using characteristics, but as

DNA technology developed, more precise phylogenies have been determined.

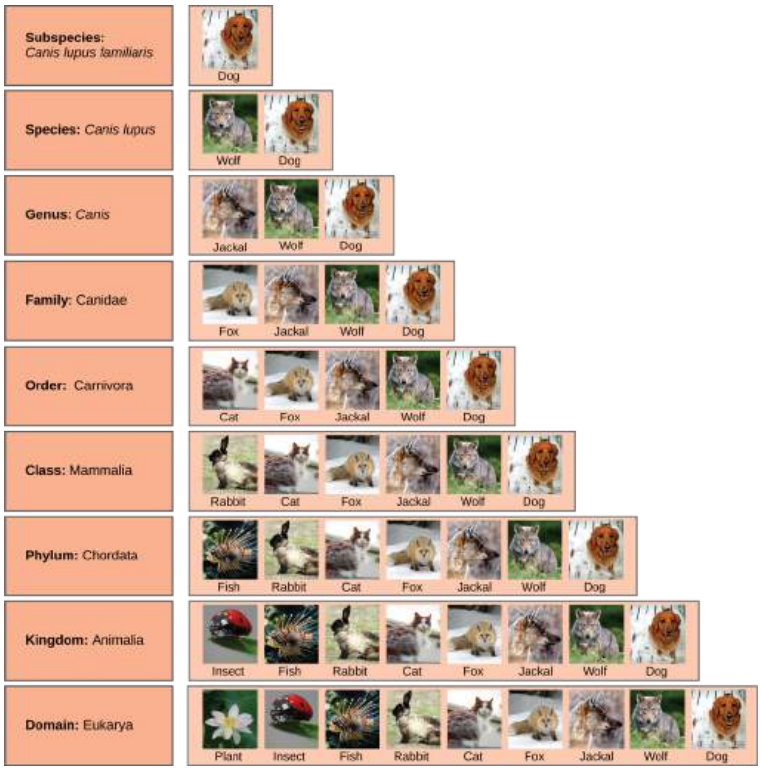


Figure 5: At each sublevel in the taxonomic classification system, organisms become more similar. Dogs and wolves are the same species because they can breed and produce viable offspring, but they are different enough to be classified as different subspecies. (credit "plant": modification of work by "berduchwal"/Flickr; credit "insect": modification of work by Jon Sullivan; credit "fish": modification of work by Christian Mehlführer; credit "rabbit": modification of work by Aidan Wojtas; credit "cat": modification of work by Jonathan Lidbeck; credit "fox": modification of work by Kevin Bacher, NPS; credit "jackal": modification of work by Thomas A. Hermann, NBII, USGS; credit "wolf": modification of work by Robert Dewar; credit "dog": modification of work by "digital_image_fan"/Flickr. "this image" by OpenStax is licensed under CC BY 4.0)

At what levels are cats and dogs considered to be part of the same group?

Recent genetic analysis and other advancements have found that some earlier phylogenetic classifications do not align with the evolutionary past; therefore, changes and updates must be made as new discoveries occur. Recall that phylogenetic trees are hypotheses and are modified as data becomes available. In addition, classification historically has focused on grouping organisms mainly by shared characteristics and does not necessarily illustrate how the various groups relate to each other from an evolutionary perspective. For example, despite the fact that a hippopotamus resembles a pig more than a whale, the hippopotamus may be the closest living relative of the whale.

Summary

Scientists continually gain new information that helps understand the evolutionary history of life on Earth. Each group of organisms went through its own evolutionary journey, called its phylogeny. Each organism shares relatedness with others, and based on morphologic and genetic evidence, scientists attempt to map the evolutionary pathways of all life on Earth. Historically, organisms were organized into a taxonomic classification system. However, today many scientists build phylogenetic trees to illustrate evolutionary relationships.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Determining Evolutionary Relationships

Scientists must collect accurate information that allows them to make evolutionary connections among organisms. Similar to detective work, scientists must use evidence to uncover the facts. In the case of phylogeny, evolutionary investigations focus on two types of evidence: morphologic (form and function) and genetic.

Two Options for Similarities

In general, organisms that share similar physical features and genomes tend to be more closely related than those that do not. Such features that overlap both **morphologically** (in form) and **genetically** are referred to as **homologous structures**; they stem from developmental similarities that are based on evolution. For example, the bones in the wings of bats and birds have homologous structures (Figure 1).

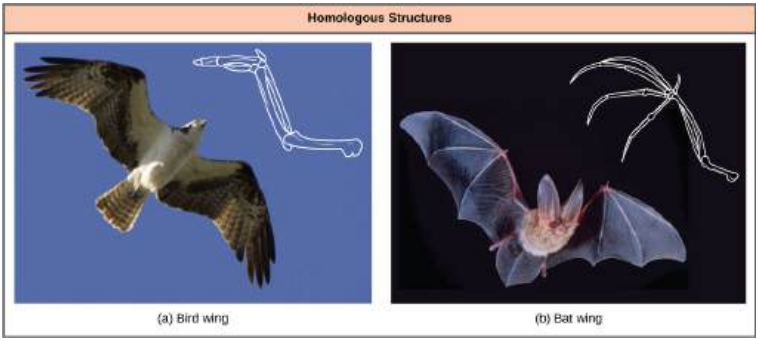


Figure 1: Bat and bird wings are homologous structures, indicating that bats and birds share a common evolutionary past. (credit a: modification of work by Steve Hillebrand, USFWS; credit b: modification of work by U.S. DOI BLM. "homologous structures" by OpenStax is licensed under CC BY 4.0)

Notice it is not simply a single bone, but rather a grouping of several bones arranged in a similar way. The more complex the feature, the more likely any kind of overlap is due to a common evolutionary past. Imagine two people from different countries both inventing a car with all the same parts and in exactly the same arrangement without any previous or shared knowledge. That outcome would be highly improbable. However, if two people both invented a hammer, it would be reasonable to conclude that both could have the original idea without the help of the other. The same relationship between complexity and shared evolutionary history is true for homologous structures in organisms.

Misleading Appearances

Some organisms may be very closely related, even though a minor genetic change caused a major morphological difference to make them look quite different. Similarly,

unrelated organisms may be distantly related, but appear very much alike. This usually happens because both organisms were in common adaptations that evolved within similar environmental conditions. When similar characteristics occur because of environmental constraints and not due to a close evolutionary relationship, it is called an **analogy** or **homoplasy**. For example, insects use wings to fly like bats and birds, but the wing structure and embryonic origin is completely different. These are called **analogous structures** (Figure 2).

Similar traits can be either homologous or analogous. Homologous structures share a similar embryonic origin; analogous organs have a similar function. For example, the bones in the front flipper of a whale are homologous to the bones in the human arm. These structures are not analogous. The wings of a butterfly and the wings of a bird are analogous but not homologous. Some structures are both analogous and homologous: the wings of a bird and the wings of a bat are both homologous and analogous. Scientists must determine which type of similarity a feature exhibits to decipher the phylogeny of the organisms being studied.

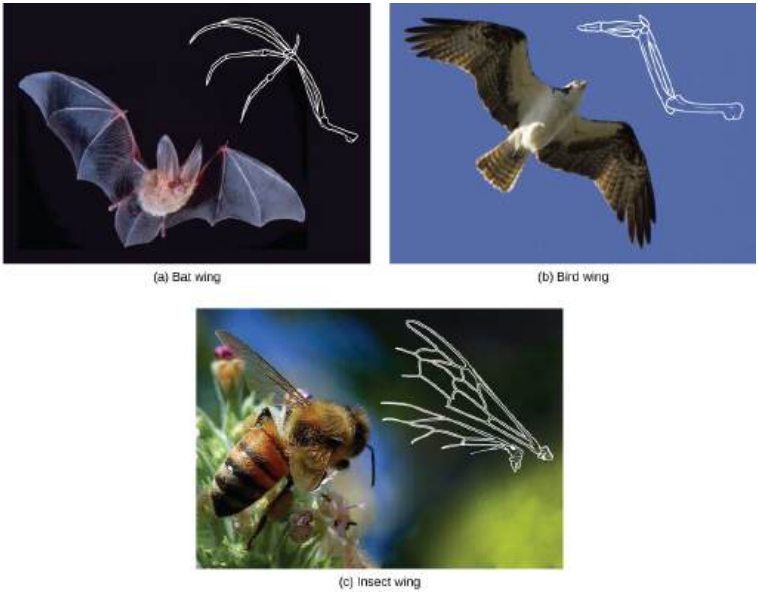


Figure 2: The (c) wing of a honeybee is similar in shape to a (b) bird wing and (a) bat wing, and it serves the same function. However, the honeybee wing is not composed of bones and has a distinctly different structure and embryonic origin. These wing types (insect versus bat and bird) illustrate an analogy—similar structures that do not share an evolutionary history. (credit a: modification of work by Steve Hillebrand, USFWS; credit b: modification of work by U.S. DOI BLM; credit c: modification of work by Jon Sullivan. "analogy" by OpenStax is licensed under CC BY 4.0)

Molecular Comparisons

With the advancement of DNA technology, the area of molecular systematics, which describes the use of information on the molecular level including DNA analysis, has blossomed. New computer programs not only confirm many earlier classified organisms, but also uncover

previously made errors. As with physical characteristics, even the DNA sequence can be tricky to read in some cases. For some situations, two very closely related organisms can appear unrelated if a mutation occurred that caused a shift in the genetic code. An insertion or deletion mutation would move each nucleotide base over one place, causing two similar codes to appear unrelated.

Sometimes two segments of DNA code in distantly related organisms randomly share a high percentage of bases in the same locations, causing these organisms to appear closely related when they are not. For both of these situations, computer technologies have been developed to help identify the actual relationships, and, ultimately, the coupled use of both morphologic and molecular information is more effective in determining phylogeny.

Why Does Phylogeny Matter?

Evolutionary biologists could list many reasons why understanding phylogeny is important to everyday life in human society. For botanists, phylogeny acts as a guide to discovering new plants that can be used to benefit people. Think of all the ways humans use plants—food, medicine, and clothing are a few examples. If a plant contains a compound that is effective in treating cancer, scientists might want to examine all of the relatives of that plant for other useful drugs.

A research team in China identified a segment of DNA thought to be common to some medicinal plants in the family Fabaceae (the legume family) and worked to identify which species had this segment (Figure 3). After testing plant species in this family, the team found a DNA marker (a known location on a chromosome that enabled them to identify the species) present. Then, using the DNA to uncover phylogenetic relationships, the team could identify whether

a newly discovered plant was in this family and assess its potential medicinal properties.

Figure 3: Dalbergia sissoo (D. sissoo) is in the Fabaceae, or legume family. Scientists found that D. sissoo shares a DNA marker with species within the Fabaceae family that have antifungal properties. Subsequently, D. sissoo was shown to have fungicidal activity, supporting the idea that DNA markers can be used to screen for plants

with potential medicinal properties. (credit: "Dalbergia sissoo" by OpenStax is licensed under CC BY 4.0)

Building Phylogenetic Trees

How do scientists construct phylogenetic trees? After the homologous and analogous traits are sorted, scientists often organize the homologous traits using a system called **cladistics**. This system sorts organisms into **clades**: groups of organisms that descended from a single ancestor. For example, in Figure 4, all of the organisms in the orange region evolved from a single ancestor that had amniotic eggs. Consequently, all of these organisms also have amniotic eggs and make a single clade, also called a **monophyletic group**. Clades must include all of the descendants from a branch point.

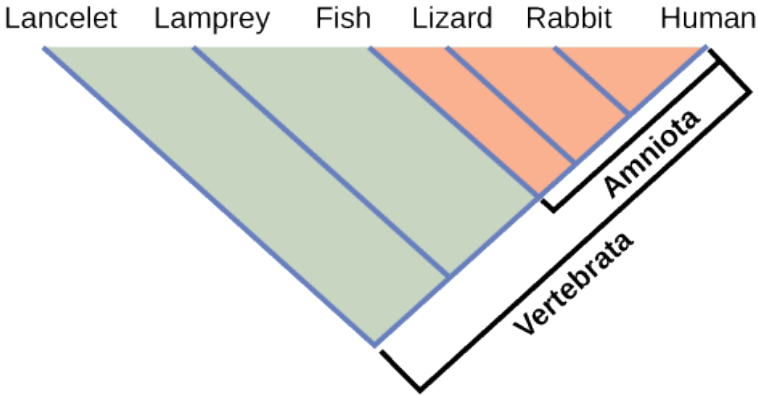


Figure 4: Lizards, rabbits, and humans all descend from a common ancestor that had an amniotic egg. Thus, lizards, rabbits, and humans all belong to the clade Amniota. Vertebrata is a larger clade that also includes fish and lamprey. (credit:"monophyletic groups" by OpenStax is licensed under CC BY 4.0)

Which animals in this figure belong to a clade that includes animals with hair? Which evolved first, hair or the amniotic egg?

Clades can vary in size depending on which branch point is being referenced. The important factor is that all of the organisms in the clade or monophyletic group stem from a single point on the tree. This can be remembered because monophyletic breaks down into “mono,” meaning one, and “phyletic,” meaning evolutionary relationship. Figure 5 shows various examples of clades. Notice how each clade comes from a single point, whereas the non-clade groups show branches that do not share a single point.

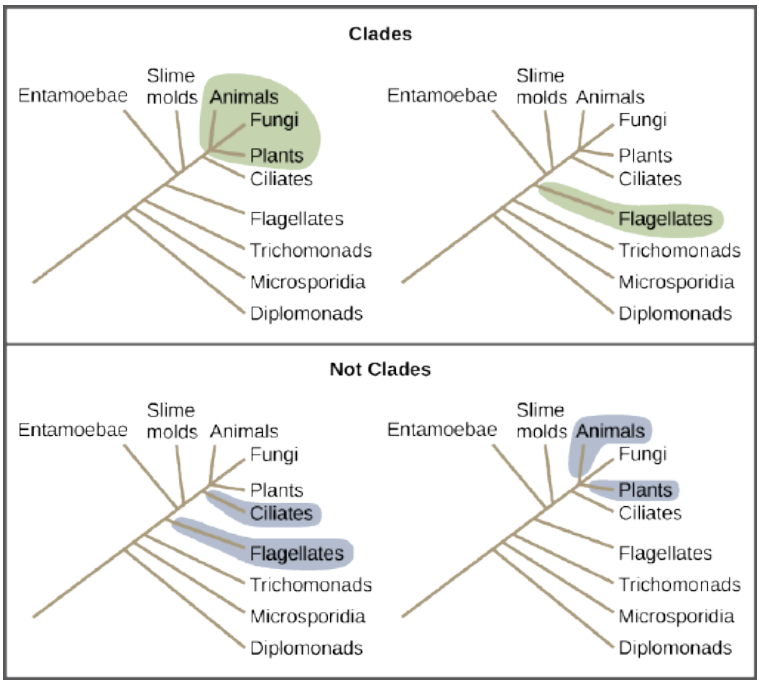


Figure 5: All the organisms within a clade stem from a single point on the tree. A clade may contain multiple groups, as in the case of animals, fungi and plants, or a single group, as in the case of flagellates. Groups that diverge at a different branch point, or that do not include all groups in a single branch point, are not considered clades. (credit: "clades" by OpenStax is licensed under CC BY 4.0)

Groups that do not include all organisms that descended from a single ancestor have different names. A **paraphyletic** group includes the most recent common ancestor, but not all of its descendants (Figure 6). A **polyphyletic** group includes unrelated organisms descended from more than one ancestor.

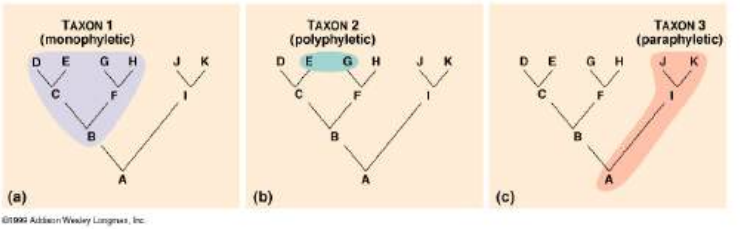


Figure 6: A visual representation of monophyletic, polyphyletic, and paraphyletic groups. (Credit: 1999 by Addison Wesley Longman)

Shared Characteristics

Organisms evolve from common ancestors and then diversify. Scientists use the phrase "descent with modification" because even though related organisms have many of the same characteristics and genetic codes, changes occur. This pattern repeats over and over as one goes through the phylogenetic tree of life:

1. A change in the genetic makeup of an organism leads to a new trait which becomes prevalent in the group.
2. Many organisms descend from this point and have this trait.
3. New variations continue to arise: some are adaptive and persist, leading to new traits.
4. With new traits, a new branch point is determined (go back to step 1 and repeat).

If a characteristic is found in the ancestor of a group, it is considered a **shared ancestral character** because all of the organisms in the taxon or clade have that trait. The vertebrate in Figure 4 is a shared ancestral character. Now consider the amniotic egg characteristic in the same figure.

Only some of the organisms in Figure 4 have this trait, and to those that do, it is called a **shared derived character** because this trait derived at some point but does not include all of the ancestors in the tree.

The tricky aspect to shared ancestral and shared derived characters is the fact that these terms are relative. The same trait can be considered one or the other depending on the particular diagram being used. Returning to Figure 4, note that the amniotic egg is a shared ancestral character for the Amniota clade, while having hair is a shared derived character for some organisms in this group. These terms help scientists distinguish between clades in the building of phylogenetic trees.

Choosing the Right Relationships

Imagine being the person responsible for organizing all of the items in a department store properly—an overwhelming task. Organizing the evolutionary relationships of all life on Earth proves much more difficult: scientists must span enormous blocks of time and work with information from long-extinct organisms. Trying to decipher the proper connections, especially given the presence of homologies and analogies, makes the task of building an accurate tree of life extraordinarily difficult. Add to that the advancement of DNA technology, which now provides large quantities of genetic sequences to be used and analyzed. Taxonomy is a subjective discipline: many organisms have more than one connection to each other, so each taxonomist will decide the order of connections.

To aid in the tremendous task of describing phylogenies accurately, scientists often use a concept called **maximum parsimony**, which means that events occurred in the simplest, most obvious way. For example, if a group of people entered a forest preserve to go hiking, based on the

principle of maximum parsimony, one could predict that most of the people would hike on established trails rather than forge new ones.

For scientists deciphering evolutionary pathways, the same idea is used: the pathway of evolution probably includes the fewest major events that coincide with the evidence at hand. Starting with all of the homologous traits in a group of organisms, scientists look for the most obvious and simple order of evolutionary events that led to the occurrence of those traits.

These tools and concepts are only a few of the strategies scientists use to tackle the task of revealing the evolutionary history of life on Earth. Recently, newer technologies have uncovered surprising discoveries with unexpected relationships, such as the fact that people seem to be more closely related to fungi than fungi are to plants. Sound unbelievable? As the information about DNA sequences grows, scientists will become closer to mapping the evolutionary history of all life on Earth.

Summary

To build phylogenetic trees, scientists must collect accurate information that allows them to make evolutionary connections between organisms. Using morphologic and molecular data, scientists work to identify homologous characteristics and genes. Similarities between organisms can stem either from shared evolutionary history (homologies) or from separate evolutionary paths (analogies). Newer technologies can be used to help distinguish homologies from analogies. After homologous information is identified, scientists use cladistics to organize these events as a means to determine an evolutionary timeline. Scientists apply the concept of maximum parsimony, which states that the order of events probably

occurred in the most obvious and simple way with the least amount of steps. For evolutionary events, this would be the path with the least number of major divergences that correlate with the evidence.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Perspectives on the Phylogenetic Tree

The concepts of phylogenetic modeling are constantly changing. It is one of the most dynamic fields of study in all of biology. Over the last several decades, new research has challenged scientists' ideas about how organisms are related. New models of these relationships have been proposed for consideration by the scientific community.

Many phylogenetic trees have been shown as models of the evolutionary relationship among species. Phylogenetic trees originated with Charles Darwin, who sketched the first phylogenetic tree in 1837 (Figure 1a), which served as a pattern for subsequent studies for more than a century. The concept of a phylogenetic tree with a single trunk representing a common ancestor, with the branches representing the divergence of species from this ancestor, fits well with the structure of many common trees, such as the oak (Figure 1b). However, evidence from modern DNA sequence analysis and newly developed computer algorithms has caused skepticism about the validity of the standard tree model in the scientific community.

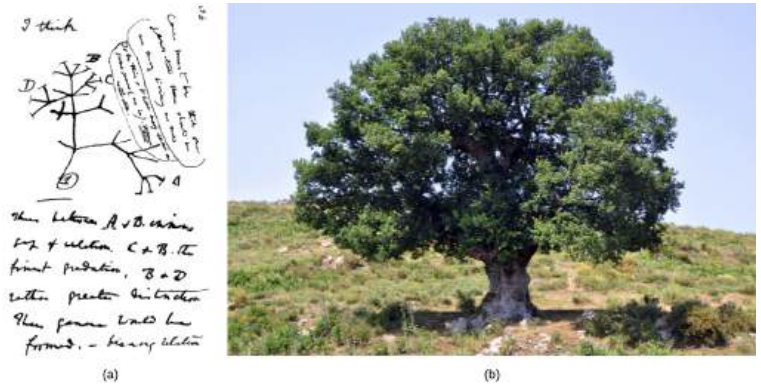


Figure 1: The (a) concept of the "tree of life" goes back to an 1837 sketch by Charles Darwin. Like an (b) oak tree, the "tree of life" has a single trunk and many branches. (credit b: modification of work by "Amada44"/Wikimedia Commons. "tree of life" by OpenStax is licensed under CC BY 4.0)

Limitations to the Classic Model

Classical thinking about prokaryotic evolution, included in the classic tree model, is that species evolve clonally. That is, they produce offspring themselves with only random mutations causing the descent into the variety of modern-day and extinct species known to science. This view is somewhat complicated in eukaryotes that reproduce sexually, but the laws of Mendelian genetics explain the variation in offspring, again, to be a result of a mutation within the species. The concept of genes being transferred between unrelated species was not considered as a possibility until relatively recently. **Horizontal gene transfer** (HGT), also known as **lateral gene transfer**, is the transfer of genes between unrelated species. HGT has been shown to be an ever-present phenomenon, with many evolutionists postulating a major role for this process in evolution, thus complicating

the simple tree model. Genes have been shown to be passed between species which are only distantly related using standard phylogeny, thus adding a layer of complexity to the understanding of phylogenetic relationships.

The various ways that HGT occurs in prokaryotes is important to understanding phylogenies. Although at present HGT is not viewed as important to eukaryotic evolution, HGT does occur in this domain as well. Finally, as an example of the ultimate gene transfer, theories of genome fusion between symbiotic or endosymbiotic organisms have been proposed to explain an event of great importance—the evolution of the first eukaryotic cell, without which humans could not have come into existence.

Horizontal Gene Transfer

Horizontal gene transfer (HGT) is the introduction of genetic material from one species to another species by mechanisms other than the vertical transmission from parent(s) to offspring. These transfers allow even distantly related species to share genes, influencing their phenotypes. It is thought that HGT is more prevalent in prokaryotes, but that only about 2% of the prokaryotic genome may be transferred by this process. Some researchers believe such estimates are premature: the actual importance of HGT to evolutionary processes must be viewed as a work in progress. As the phenomenon is investigated more thoroughly, it may be revealed to be more common. Many scientists believe that HGT and mutation appear to be (especially in prokaryotes) a significant source of genetic variation, which is the raw material for the process of natural selection. These transfers may occur between any two species that share an intimate relationship (Table 1).

Summary of Mechanisms of Prokaryotic and Eukaryotic HGT

	Mechanism	Mode of Transmission	Example
Prokaryotes	transformation	DNA uptake	many prokaryotes
	transduction	bacteriophage (virus)	bacteria
	conjugation	pilus	many prokaryotes
	gene transfer agents	phage-like particles	purple non-sulfur bacteria
Eukaryotes	from food organisms	unknown	aphid
	jumping genes	transposons	rice and millet plants
	epiphytes/ parasites	unknown	yew tree fungi
	from viral infections		

HGT in Prokaryotes

The mechanism of HGT has been shown to be quite common in the prokaryotic domains of Bacteria and Archaea, significantly changing the way their evolution is viewed. The majority of evolutionary models, such as in the **Endosymbiont Theory**, propose that eukaryotes descended from multiple prokaryotes, which makes HGT all the more important to understanding the phylogenetic relationships of all extant and extinct species.

The fact that genes are transferred among common bacteria is well known to microbiology students. These gene transfers between species are the major mechanism whereby bacteria acquire resistance to antibiotics. Classically, this type of transfer has been thought to occur by three different mechanisms:

1. **Transformation**: naked DNA is taken up by a bacteria

2. **Transduction**: genes are transferred using a virus
3. **Conjugation**: the use a hollow tube called a pilus to transfer genes between organisms

More recently, a fourth mechanism of gene transfer between prokaryotes has been discovered. Small, virus-like particles called **gene transfer agents** (GTAs) transfer random genomic segments from one species of prokaryote to another. GTAs have been shown to be responsible for genetic changes, sometimes at a very high frequency compared to other evolutionary processes. The first GTA was characterized in 1974 using purple, non-sulfur bacteria. These GTAs, which are thought to be bacteriophages that lost the ability to reproduce on their own, carry random pieces of DNA from one organism to another. The ability of GTAs to act with high frequency has been demonstrated in controlled studies using marine bacteria. Gene transfer events in marine prokaryotes, either by GTAs or by viruses, have been estimated to be as high as 10^{13} per year in the Mediterranean Sea alone. GTAs and viruses are thought to be efficient HGT vehicles with a major impact on prokaryotic evolution.

As a consequence of this modern DNA analysis, the idea that eukaryotes evolved directly from Archaea has fallen out of favor. While eukaryotes share many features that are absent in bacteria, such as the TATA box (found in the promoter region of many genes), the discovery that some eukaryotic genes were more homologous with bacterial DNA than Archaea DNA made this idea less tenable. Furthermore, the fusion of genomes from Archaea and Bacteria by endosymbiosis has been proposed as the ultimate event in eukaryotic evolution.

HGT in Eukaryotes

Although it is easy to see how prokaryotes exchange genetic material by HGT, it was initially thought that this process was absent in eukaryotes. After all, prokaryotes are but single cells exposed directly to their environment, whereas the sex cells of multicellular organisms are usually sequestered in protected parts of the body. It follows from this idea that the gene transfers between multicellular eukaryotes should be more difficult. Indeed, it is thought that this process is rarer in eukaryotes and has a much smaller evolutionary impact than in prokaryotes. In spite of this fact, HGT between distantly related organisms has been demonstrated in several eukaryotic species, and it is possible that more examples will be discovered in the future.

In plants, gene transfer has been observed in species that cannot cross-pollinate by normal means. Transposons or "jumping genes" have been shown to transfer between rice and millet plant species. Furthermore, fungal species feeding on yew trees, from which the anti-cancer drug TAXOL® is derived from the bark, have acquired the ability to make taxol themselves, a clear example of gene transfer.

In animals, a particularly interesting example of HGT occurs within the aphid species (Figure 2). Aphids are insects that vary in color based on carotenoid content. Carotenoids are pigments made by a variety of plants, fungi, and microbes, and they serve a variety of functions in animals, who obtain these chemicals from their food. Humans require carotenoids to synthesize vitamin A, and we obtain them by eating orange fruits and vegetables such as carrots, apricots, mangoes, and sweet potatoes. On the other hand, aphids have acquired the ability to make carotenoids on their own. According to DNA analysis, this ability is due to the transfer of fungal genes into the insect by HGT, presumably as the insect consumed fungi for food. A carotenoid enzyme called

a desaturase is responsible for the red coloration seen in certain aphids, and it has been further shown that when this gene is inactivated by mutation, the aphids revert back to their more common green color (Figure 2).

Figure 2: (a) Red aphids get their color from red carotenoid pigment. Genes necessary to make this pigment are present in certain fungi, and scientists speculate that aphids acquired these genes through HGT after consuming fungi for food. If genes for making carotenoids are inactivated by mutation, the aphids revert back to (b) their green color. Red coloration makes the aphids a lot more conspicuous to predators, but evidence suggests that red aphids are more resistant to insecticides than green ones. Thus, red aphids may be more fit to survive in some environments than green ones. (credit a: modification of work by Benny Mazur; credit b: modification of work by Mick Talbot. "aphids" by OpenStax is licensed under CC BY 4.0)

Genome Fusion and the Evolution of Eukaryotes

Scientists believe the ultimate in HGT occurs through genome fusion between different species of prokaryotes when two symbiotic organisms become endosymbiotic. This occurs when one species is taken inside the cytoplasm of another species, which ultimately results in a genome consisting of genes from both the endosymbiont and the host. This mechanism is an aspect of the Endosymbiont

Theory, which is accepted by a majority of biologists as the mechanism whereby eukaryotic cells obtained their mitochondria and chloroplasts. However, the role of endosymbiosis in the development of the nucleus is more controversial. Nuclear and mitochondrial DNA are thought to be of different (separate) evolutionary origin, with the mitochondrial DNA being derived from the circular genomes of bacteria that were engulfed by ancient prokaryotic cells. Mitochondrial DNA can be regarded as the smallest chromosome. Interestingly enough, mitochondrial DNA is inherited only from the mother. The mitochondrial DNA degrades in sperm when the sperm degrades in the fertilized egg or in other instances when the mitochondria located in the flagellum of the sperm fails to enter the egg.

Within the past decade, the process of **genome fusion** by endosymbiosis has been proposed by James Lake of the UCLA/NASA Astrobiology Institute to be responsible for the evolution of the first eukaryotic cells (Figure 3a). Using DNA analysis and a new mathematical algorithm called conditioned reconstruction (CR), his laboratory proposed that eukaryotic cells developed from an endosymbiotic gene fusion between two species, one an Archaea and the other a Bacteria. As mentioned, some eukaryotic genes resemble those of Archaea, whereas others resemble those from Bacteria. An endosymbiotic fusion event, such as Lake has proposed, would clearly explain this observation. On the other hand, this work is new and the CR algorithm is relatively unsubstantiated, which causes many scientists to resist this hypothesis.

More recent work by Lake (Figure 3b) proposes that gram-negative bacteria, which are unique within their domain in that they contain two lipid bilayer membranes, indeed resulted from an endosymbiotic fusion of archaeal and bacterial species. The double membrane would be a direct result of the endosymbiosis, with the endosymbiont picking

up the second membrane from the host as it was internalized. This mechanism has also been used to explain the double membranes found in mitochondria and chloroplasts. Lake's work is not without skepticism, and the ideas are still debated within the biological science community. In addition to Lake's hypothesis, there are several other competing theories as to the origin of eukaryotes. How did the eukaryotic nucleus evolve? One theory is that the prokaryotic cells produced an additional membrane that surrounded the bacterial chromosome. Some bacteria have the DNA enclosed by two membranes; however, there is no evidence of a nucleolus or nuclear pores. Other proteobacteria also have membrane-bound chromosomes. If the eukaryotic nucleus evolved this way, we would expect one of the two types of prokaryotes to be more closely related to eukaryotes.

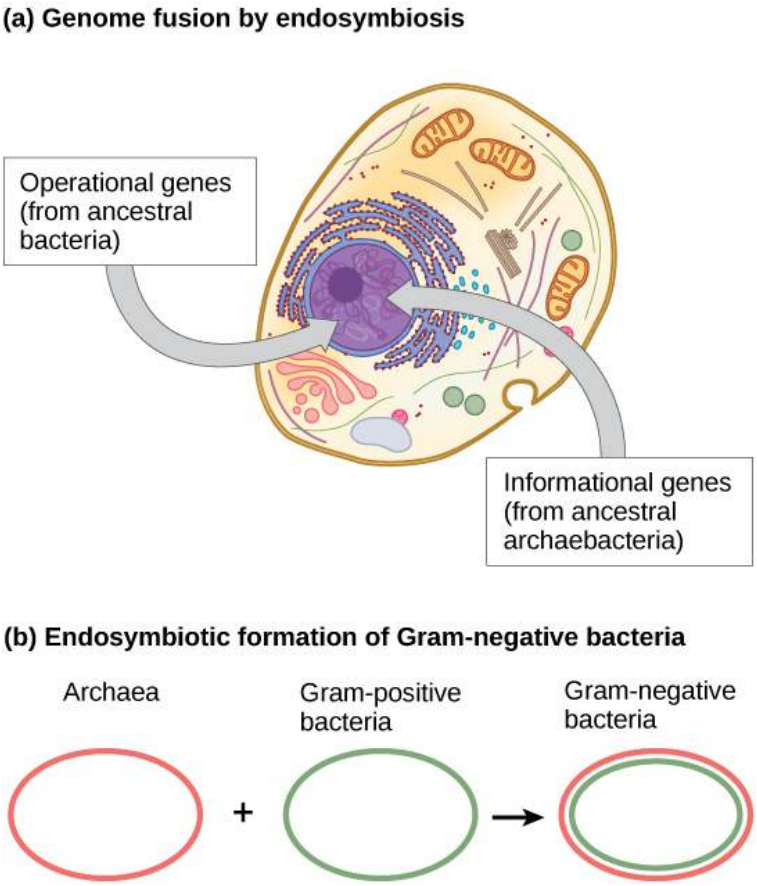


Figure 3: The theory that mitochondria and chloroplasts are endosymbiotic in origin is now widely accepted. More controversial is the proposal that (a) the eukaryotic nucleus resulted from the fusion of archaeal and bacterial genomes, and that (b) Gram-negative bacteria, which have two membranes, resulted from the fusion of Archaea and Gram-positive bacteria, each of which has a single membrane. (credit: "endosymbiosis" by OpenStax is licensed under CC BY 4.0)

The **nucleus-first hypothesis** proposes that the nucleus evolved in prokaryotes first (Figure 4a), followed by a later fusion of the new eukaryote with bacteria that became

mitochondria. The **mitochondria-first hypothesis** proposes that mitochondria were first established in a prokaryotic host (Figure 4b), which subsequently acquired a nucleus, by fusion or other mechanisms, to become the first eukaryotic cell. The **eukaryote-first hypothesis** proposes that prokaryotes actually evolved from eukaryotes by losing genes and complexity (Figure 4c). All of these hypotheses are testable. Only time and more experimentation will determine which hypothesis is best supported by data.

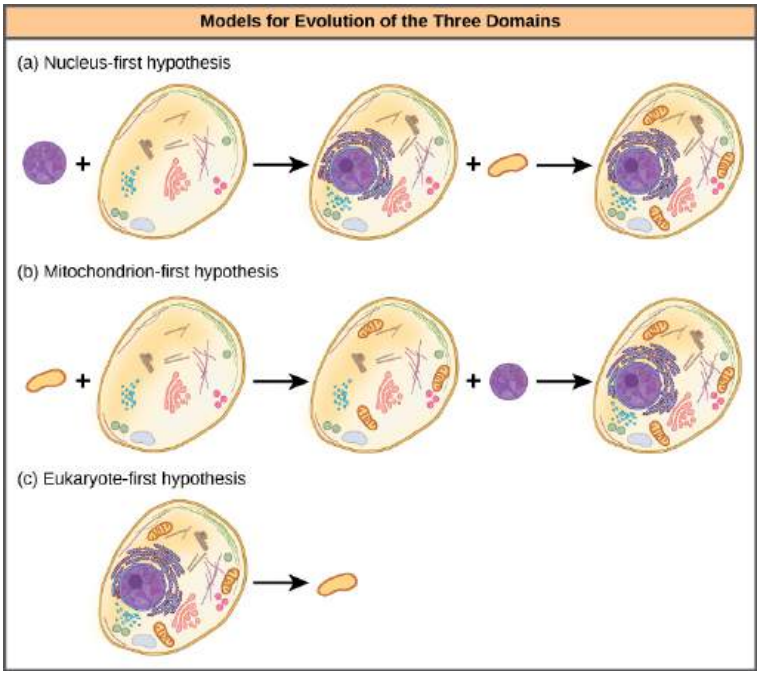


Figure 4: Three alternate hypotheses of eukaryotic and prokaryotic evolution are (a) the nucleus-first hypothesis, (b) the mitochondrion-first hypothesis, and (c) the eukaryote-first hypothesis. (credit: "Three alternate hypotheses" by OpenStax is licensed under CC BY 4.0)

Web and Network Models

The recognition of the importance of HGT, especially in the evolution of prokaryotes, has caused some to propose abandoning the classic "tree of life" model. In 1999, W. Ford Doolittle proposed a phylogenetic model that resembles a web or a network more than a tree. The hypothesis is that eukaryotes evolved not from a single prokaryotic ancestor, but from a pool of many species that were sharing genes by HGT mechanisms. As shown in Figure 5a, some individual prokaryotes were responsible for transferring the bacteria that caused mitochondrial development to the new eukaryotes, whereas other species transferred the bacteria that gave rise to chloroplasts. This model is often called the "**web of life.**" In an effort to save the tree analogy, some have proposed using the *Ficus* tree (Figure 5b) with its multiple trunks as a phylogenetic to represent a diminished evolutionary role for HGT.

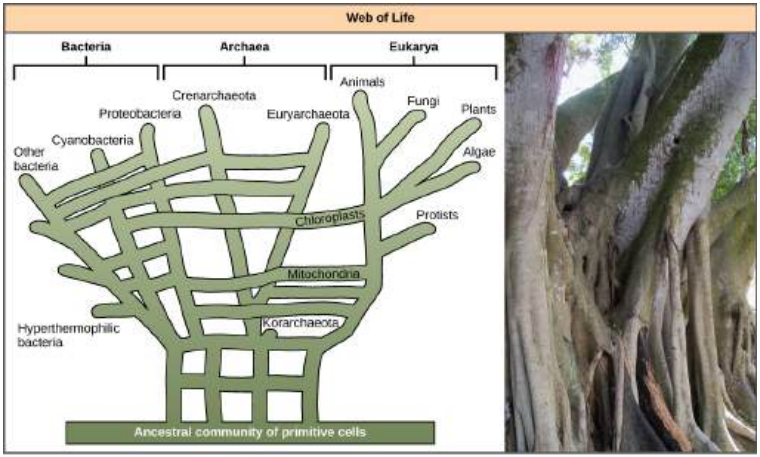


Figure 5: In the (a) phylogenetic model proposed by W. Ford Doolittle, the "tree of life" arose from a community of ancestral cells, has multiple trunks, and has connections between branches where horizontal gene transfer has occurred. Visually, this concept is better represented by (b) the multi-trunked Ficus than by the single trunk of the oak similar to the tree drawn by Darwin Figure. (credit b: modification of work by "psyberartist"/Flickr. "Doolittle, the "tree of life"" by OpenStax is licensed under CC BY 4.0)

Ring of Life Models

Others have proposed abandoning any tree-like model of phylogeny in favor of a ring structure, the so-called "**ring of life**" (Figure 6); a phylogenetic model where all three domains of life evolved from a pool of primitive prokaryotes. Lake, again using the conditioned reconstruction algorithm, proposes a ring-like model in which species of all three domains—Archaea, Bacteria, and Eukarya—evolved from a single pool of gene-swapping prokaryotes. His laboratory proposes that this structure is the best fit for data from

extensive DNA analyses performed in his laboratory and that the ring model is the only one that adequately takes HGT and genomic fusion into account. However, other phylogeneticists remain highly skeptical of this model.

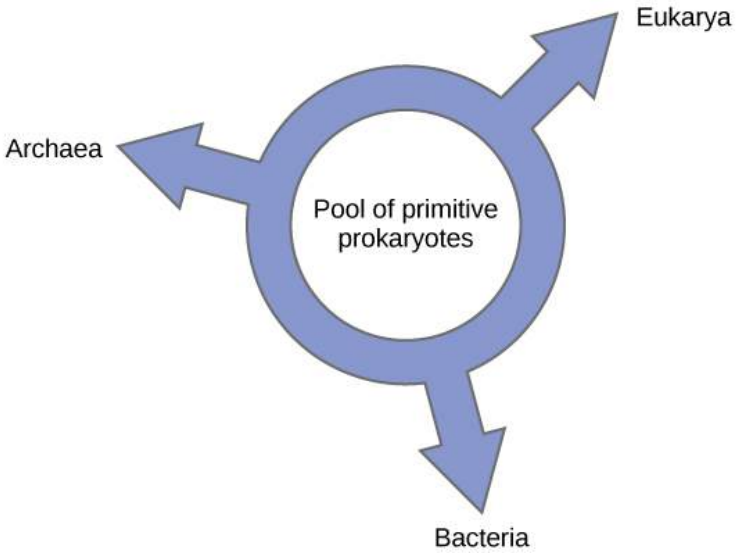


Figure 6: According to the "ring of life" phylogenetic model, the three domains of life evolved from a pool of primitive prokaryotes. (credit: "ring of life" by OpenStax is licensed under CC BY 4.0)

In summary, the "**tree of life**" model proposed by Darwin must be modified to include HGT. Does this mean abandoning the tree model completely? Even Lake argues that all attempts should be made to discover some modification of the tree model to allow it to accurately fit his data, and only the inability to do so will sway people toward his ring proposal.

This doesn't mean a tree, web, or a ring will correlate completely to an accurate description of phylogenetic relationships of life. A consequence of the new thinking about phylogenetic models is the idea that Darwin's original

conception of the phylogenetic tree is too simple, but made sense based on what was known at the time. However, the search for a more useful model moves on: each model serving as hypotheses to be tested with the possibility of developing new models. This is how science advances. These models are used as visualizations to help construct hypothetical evolutionary relationships and understand the massive amount of data being analyzed.

Summary

The phylogenetic tree, first used by Darwin, is the classic "tree of life" model describing phylogenetic relationships among species, and the most common model used today. New ideas about HGT and genome fusion have caused some to suggest revising the model to resemble webs or rings.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

ECOLOGY AND THE BIOSPHERE

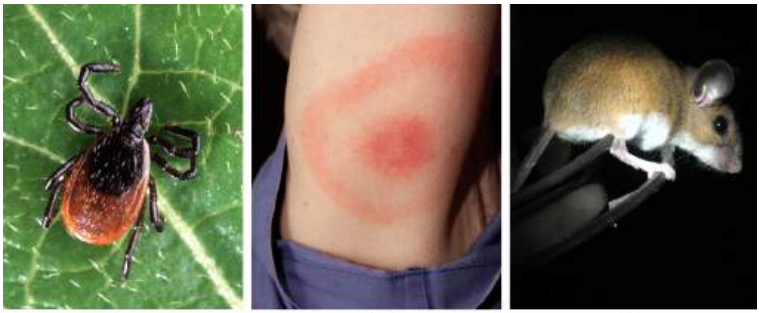


Figure 1: The (a) deer tick carries the bacterium that produces Lyme disease in humans, often evident in (b) a symptomatic bull's eye rash. The (c) white-footed mouse is one well-known host to deer ticks carrying the Lyme disease bacterium. (credit a: modification of work by Scott Bauer, USDA ARS; credit b: modification of work by James Gathany, CDC; credit c: modification of work by Rob Ireton. "this image" by OpenStax is licensed under CC BY 4.0)

Why study ecology? Perhaps you are interested in learning about the natural world and how living things have adapted to the physical conditions of their environment. Or, perhaps

you're a future physician seeking to understand the connection between human health and ecology.

Humans are a part of the ecological landscape, and human health is one important part of human interaction with our physical and living environment. Lyme disease, for instance, serves as one modern-day example of the connection between our health and the natural world (Figure 1). More formally known as Lyme borreliosis, Lyme disease is a bacterial infection that can be transmitted to humans when they are bitten by the deer tick (*Ixodes scapularis*), which is the primary vector for this disease. However, not all deer ticks carry the bacteria that will cause Lyme disease in humans, and *I. scapularis* can have other hosts besides deer. In fact, it turns out that the probability of infection depends on the type of host upon which the tick develops: a higher proportion of ticks that live on white-footed mice carry the bacterium than do ticks that live on deer. Knowledge about the environments and population densities in which the host species is abundant would help a physician or an epidemiologist better understand how Lyme disease is transmitted and how its incidence could be reduced.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

The Scope of Ecology

Ecology is the study of the interactions of living organisms with their environment. One core goal of ecology is to understand the distribution and abundance of living things in the physical environment. Attainment of this goal requires the integration of scientific disciplines inside and outside of biology, such as biochemistry, physiology, evolution, biodiversity, molecular biology, geology, and climatology. Some ecological research also applies aspects of chemistry and physics, and it frequently uses mathematical models.

Levels of Ecological Study

When a discipline such as biology is studied, it is often helpful to subdivide it into smaller, related areas. For instance, cell biologists interested in cell signaling need to understand the chemistry of the signal molecules (which are usually proteins) as well as the result of cell signaling. Ecologists interested in the factors that influence the survival of an endangered species might use mathematical models to predict how current conservation efforts affect endangered organisms. To produce a sound set of management options, a conservation biologist needs to collect accurate data, including current population size, factors affecting reproduction (like physiology and behavior), habitat requirements (such as plants and soils), and potential human

influences on the endangered population and its habitat (which might be derived through studies in sociology and urban ecology). Within the discipline of ecology, researchers work at four specific levels, sometimes discretely and sometimes with overlap: organism, population, community, and ecosystem (Figure 1).

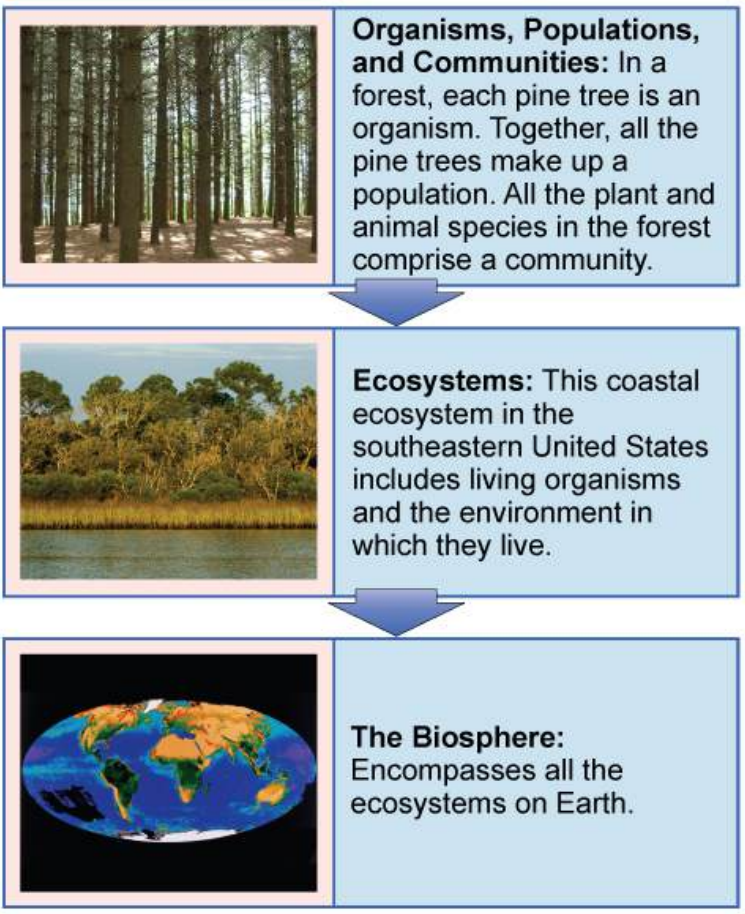


Figure 1: Ecologists study within several biological levels of organization. (credit "organisms": modification of work by "Crystl"/Flickr; credit "ecosystems": modification of work by Tom Carlisle, US Fish and Wildlife Service Headquarters; credit "biosphere": NASA. "this image" by OpenStax is licensed under CC BY 4.0)

Organismal Ecology

Researchers studying ecology at the **organismal level** are interested in the adaptations that enable individuals to live in specific habitats. These adaptations can be morphological, physiological, and behavioral. For instance, the Karner blue butterfly (*Lycaeides melissa samuelis*) (Figure 2) is considered a specialist because the females preferentially oviposit (that is, lay eggs) on wild lupine. This preferential adaptation means that the Karner blue butterfly is highly dependent on the presence of wild lupine plants for its continued survival.

Figure 2: The Karner blue butterfly (Lycaeides melissa samuelis) is a rare butterfly that lives only in open areas with few trees or shrubs, such as pine barrens and oak savannas. It can only lay its eggs on lupine plants. (credit: modification of work by J & K Hollingsworth, USFWS. "Lycaeides melissa samuelis" by OpenStax is licensed under CC BY 4.0)

After hatching, the larval caterpillars emerge and spend four to six weeks feeding solely on wild lupine (Figure 3). The caterpillars pupate (undergo metamorphosis) and emerge as butterflies after about four weeks. The adult butterflies feed on the nectar of flowers of wild lupine and other plant species. A researcher interested in studying Karner blue butterflies at the organismal level might, in addition to asking questions about egg laying, ask questions about the butterflies' preferred temperature (a **physiological** question) or the behavior of the caterpillars when they are at different larval stages (a **behavioral** question).

Figure 3: The wild lupine (Lupinus perennis) is the host plant for the Karner blue butterfly. (credit: "Lupinus perennis" by OpenStax is licensed under CC BY 4.0)

Population Ecology

A **population** is a group of interbreeding organisms that are members of the same species living in the same area at the same time. (Organisms that are all members of the same species are called conspecifics.) A population is identified, in part, by where it lives. Its area may have natural or artificial boundaries: natural boundaries might be rivers, mountains, or deserts, while examples of artificial boundaries include mowed grass, manmade structures, or roads. The study of population ecology focuses on the number of individuals in an area and how and why population size changes over time. Population ecologists are particularly interested in counting the Karner blue butterfly, for example, because it is classified as federally endangered. However, the distribution and density of this species is highly influenced by the distribution and abundance of wild lupine. Researchers might ask questions about the factors leading to the decline of wild lupine and how these affect Karner blue butterflies. For example, ecologists know that wild lupine thrives in open areas where trees and shrubs are largely absent. In natural settings, intermittent wildfires regularly remove trees and shrubs, helping to maintain the open areas that wild lupine requires. Mathematical models can be used to understand how wildfire suppression by humans has led to the decline of this important plant for the Karner blue butterfly.

Community Ecology

A biological **community** consists of the different species within an area, typically a three-dimensional space, and the interactions within and among these species. Community ecologists are interested in the processes driving these interactions and their consequences. Questions about conspecific interactions often focus on competition among members of the same species for a limited resource. Ecologists also study interactions among various species; members of different species are called **heterospecifics**. Examples of heterospecific interactions include predation, parasitism, herbivory, competition, and pollination. These interactions can have regulating effects on population sizes and can impact ecological and evolutionary processes affecting diversity.

For example, Karner blue butterfly larvae form mutualistic relationships with ants. A **mutualism** is a form of a long-term relationship that has coevolved between two species and from which each species benefits. For mutualism to exist between individual organisms, each species must receive some benefit from the other as a consequence of the relationship. Researchers have shown that there is an increase in the probability of survival when Karner blue butterfly larvae (caterpillars) are tended by ants. This might be because the larvae spend less time in each life stage when tended by ants, which provides an advantage for the larvae. Meanwhile, the Karner blue butterfly larvae secrete a carbohydrate-rich substance that is an important energy source for the ants. Both the Karner blue larvae and the ants benefit from their interaction.

Ecosystem Ecology

Ecosystem ecology is an extension of organismal, population,

and community ecology. The **ecosystem** is composed of all the **biotic** components (living things) in an area along with the **abiotic** components (non-living things) of that area. Some of the abiotic components include air, water, and soil. Ecosystem biologists ask questions about how nutrients and energy are stored and how they move among organisms and the surrounding atmosphere, soil, and water.

The Karner blue butterflies and the wild lupine live in an oak-pine barren habitat. This habitat is characterized by natural disturbance and nutrient-poor soils that are low in nitrogen. The availability of nutrients is an important factor in the distribution of the plants that live in this habitat. Researchers interested in ecosystem ecology could ask questions about the importance of limited resources and the movement of resources, such as nutrients, though the biotic and abiotic portions of the ecosystem.

Career Connection: Ecologist

A career in ecology contributes to many facets of human society. Understanding ecological issues can help society meet the basic human needs of food, shelter, and health care. Ecologists can conduct their research in the laboratory and outside in natural environments (Figure 4). These natural environments can be as close to home as the stream running through your campus or as far away as the hydrothermal vents at the bottom of the Pacific Ocean. Ecologists manage natural resources such as white-tailed deer populations (*Odocoileus virginianus*) for hunting or aspen (*Populus* spp.) timber stands for paper production. Ecologists also work as educators who teach children and adults at various institutions including universities, high schools, museums, and nature centers. Ecologists may also work in advisory positions assisting local, state, and federal policymakers to develop laws that are ecologically sound, or they may

develop those policies and legislation themselves. Becoming an ecologist requires an undergraduate degree, usually in a natural science. The undergraduate degree is often followed by specialized training or an advanced degree, depending on the area of ecology selected. Ecologists should also have a broad background in the physical sciences, as well as a sound foundation in mathematics and statistics.

Figure 4: This landscape ecologist is releasing a black-footed ferret into its native habitat as part of a study. (credit: USFWS Mountain Prairie Region, NPS. "this image" by OpenStax is licensed under CC BY 4.0)

Summary

Ecology is the study of the interactions of living things with their environment. Ecologists ask questions across four levels of biological organization—organismal, population, community, and ecosystem. At the organismal level, ecologists study individual organisms and how they interact

with their environments. At the population and community levels, ecologists explore, respectively, how a population of organisms changes over time and the ways in which that population interacts with other species in the community. Ecologists studying an ecosystem examine the living species (the biotic components) of the ecosystem as well as the nonliving portions (the abiotic components), such as air, water, and soil, of the environment.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Biogeography

Many forces influence the communities of living organisms present in different parts of the biosphere (all of the parts of Earth inhabited by life). The biosphere extends into the atmosphere (several kilometers above Earth) and into the depths of the oceans. Despite its apparent vastness to an individual human, the biosphere occupies only a minute space when compared to the known universe. Many abiotic forces influence where life can exist and the types of organisms found in different parts of the biosphere. The abiotic factors influence the distribution of biomes: large areas of land with similar climate, flora, and fauna.

Biogeography

Biogeography is the study of the geographic distribution of living things and the abiotic factors that affect their distribution. Abiotic factors such as temperature and rainfall vary based mainly on latitude and elevation. As these abiotic factors change, the composition of plant and animal communities also changes. For example, if you were to begin a journey at the equator and walk north, you would notice gradual changes in plant communities. At the beginning of your journey, you would see tropical wet forests with broad-leaved evergreen trees, which are characteristic of plant communities found near the equator. As you continued to

travel north, you would see these broad-leaved evergreen plants eventually give rise to seasonally dry forests with scattered trees. You would also begin to notice changes in temperature and moisture. At about 30 degrees north, these forests would give way to deserts, which are characterized by low precipitation.

Moving farther north, you would see that deserts are replaced by grasslands or prairies. Eventually, grasslands are replaced by deciduous temperate forests. These deciduous forests give way to the boreal forests found in the subarctic, the area south of the Arctic Circle. Finally, you would reach the Arctic tundra, which is found at the most northern latitudes. This trek north reveals gradual changes in both climate and the types of organisms that have adapted to environmental factors associated with ecosystems found at different latitudes. However, different ecosystems exist at the same latitude due in part to abiotic factors such as jet streams, the Gulf Stream, and ocean currents. If you were to hike up a mountain, the changes you would see in the vegetation would parallel those as you move to higher latitudes.

Ecologists who study biogeography examine patterns of species distribution. No species exists everywhere; for example, the Venus flytrap is endemic to a small area in North and South Carolina. An **endemic** species is one that is naturally found only in a specific geographic area that is usually restricted in size. Other species are **generalists**: species that live in a wide variety of geographic areas; the raccoon, for example, is native to most of North and Central America.

Species distribution patterns are based on biotic and abiotic factors and their influences during the very long periods of time required for species evolution; therefore, early studies of biogeography were closely linked to the emergence of evolutionary thinking in the eighteenth

century. Some of the most distinctive assemblages of plants and animals occur in regions that have been physically separated for millions of years by geographic barriers. Biologists estimate that Australia, for example, has between 600,000 and 700,000 species of plants and animals. Approximately 3/4 of living plant and mammal species are endemic species found solely in Australia (Figure 1).

Figure 1: Australia is home to many endemic species. The (a) wallaby (Wallabia bicolor), a medium-sized member of the kangaroo family, is a pouched mammal, or marsupial. The (b) echidna (Tachyglossus aculeatus) is an egg-laying mammal. (credit a: modification of work by Derrick Coetzee; credit b: modification of work by Allan Whittome. "this image" by OpenStax is licensed under CC BY 4.0)

Sometimes ecologists discover unique patterns of species distribution by determining where species are *not* found. Hawaii, for example, has no native land species of reptiles or amphibians and has only one native terrestrial mammal, the hoary bat. Most of New Guinea, as another example, lacks placental mammals.

Plants can be endemic or generalists: endemic plants are found only on specific regions of the Earth, while generalists are found in many regions. Isolated landmasses—such as

Australia, Hawaii, and Madagascar—often have large numbers of endemic plant species. Some of these plants are endangered due to human activity. The forest gardenia (*Gardenia brighamii*), for instance, is endemic to Hawaii; only an estimated 15–20 trees are thought to exist (Figure 2).

Figure 2: Listed as federally endangered, the forest gardenia is a small tree with distinctive flowers. It is found only in five of the Hawaiian Islands in small populations consisting of a few individual specimens. (credit: Forest & Kim Starr. "forest gardenia" by OpenStax is licensed under CC BY 4.0)

Energy Sources

Energy from the sun is captured by green plants, algae, cyanobacteria, and photosynthetic protists. These organisms convert **solar energy** into the **chemical energy** needed by all living things. Light availability can be an important force directly affecting the evolution of adaptations in photosynthesizers. For instance, plants in the understory of a temperate forest are shaded when the trees above them in the canopy completely leaf out in the late spring. Not surprisingly, understory plants have adaptations to successfully capture available light. One such adaptation is the rapid growth of spring ephemeral plants such as the spring beauty (Figure 3). These spring flowers achieve much of their growth and finish their life cycle (reproduce) early in the season before the trees in the canopy develop leaves.

Figure 3: The spring beauty is an ephemeral spring plant that flowers early in the spring to avoid competing with larger forest trees for sunlight. (credit: John Beetham. "spring beauty" by OpenStax is licensed under CC BY 4.0)

In aquatic ecosystems, the availability of light may be limited because sunlight is absorbed by water, plants, suspended particles, and resident microorganisms. Toward the bottom of a lake, pond, or ocean, there is a zone that light cannot reach. **Photosynthesis** cannot take place there and, as a result, a number of adaptations have evolved that enable living things to survive without light. For instance, aquatic plants have photosynthetic tissue near the surface of the

water; for example, think of the broad, floating leaves of a water lily—water lilies cannot survive without light. In environments such as **hydrothermal** vents, some bacteria extract energy from inorganic chemicals because there is no light for photosynthesis.

The availability of **nutrients** in aquatic systems is also an important aspect of energy or photosynthesis. Many organisms sink to the bottom of the ocean when they die in the open water; when this occurs, the energy found in that living organism is sequestered for some time unless **ocean upwelling** occurs. Ocean upwelling is the rising of deep ocean waters that occurs when prevailing winds blow along surface waters near a coastline (Figure 4). As the wind pushes ocean waters offshore, water from the bottom of the ocean moves up to replace this water. As a result, the nutrients once contained in dead organisms become available for reuse by other living organisms.

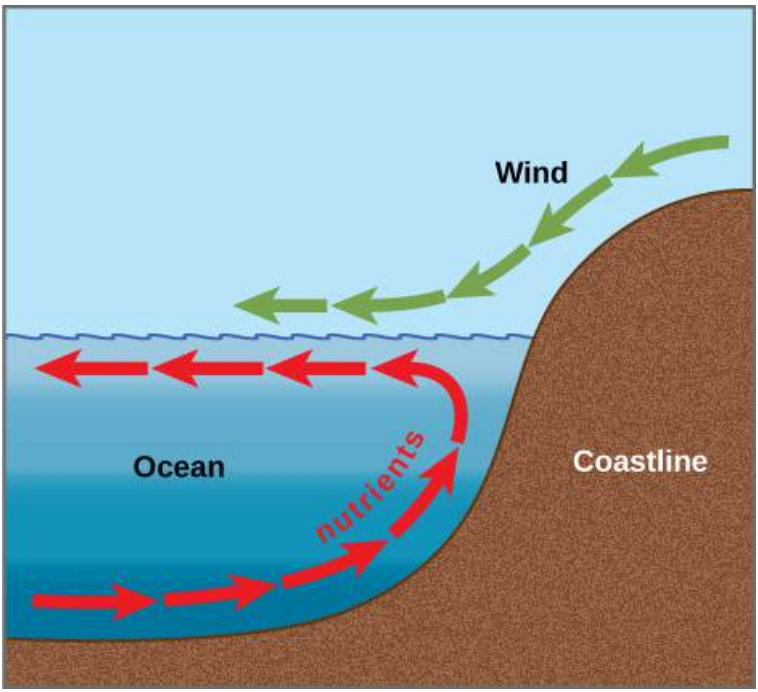


Figure 4: Ocean upwelling is an important process that recycles nutrients and energy in the ocean. As wind (green arrows) pushes offshore, it causes water from the ocean bottom (red arrows) to move to the surface, bringing up nutrients from the ocean depths. (credit: "Ocean upwelling" by OpenStax is licensed under CC BY 4.0)

In freshwater systems, the recycling of nutrients occurs in response to air temperature changes. The nutrients at the bottom of lakes are recycled twice each year: in the spring and fall turnover. The spring and fall turnover is a seasonal process that recycles nutrients and oxygen from the bottom of a freshwater ecosystem to the top of a body of water (Figure 5). These turnovers are caused by the formation of a **thermocline**: a layer of water with a temperature that is significantly different from that of the surrounding layers. In wintertime, the surface of lakes found in many northern

regions is frozen. However, the water under the ice is slightly warmer, and the water at the bottom of the lake is warmer yet at 4 °C to 5 °C (39.2 °F to 41 °F). Water is densest at 4 °C; therefore, the deepest water is also the densest. The deepest water is oxygen-poor because the decomposition of organic material at the bottom of the lake uses up available oxygen that cannot be replaced by means of oxygen diffusion into the water due to the surface ice layer.

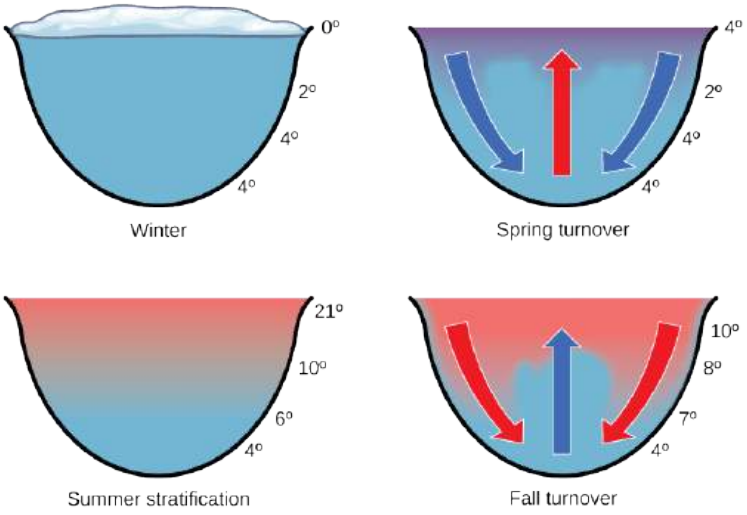


Figure 5: The spring and fall turnovers are important processes in freshwater lakes that act to move the nutrients and oxygen at the bottom of deep lakes to the top. Turnover occurs because water has a maximum density at 4 °C. Surface water temperature changes as the seasons progress, and denser water sinks. (credit: "Turnover" by OpenStax is licensed under CC BY 4.0)

In springtime, air temperatures increase, and surface ice melts. When the temperature of the surface water begins to reach 4 °C, the water becomes heavier and sinks to the bottom. The water at the bottom of the lake is then displaced

by the heavier surface water and, thus, rises to the top. As that water rises to the top, the sediments and nutrients from the lake bottom are brought along with it. During the summer months, the lake water stratifies, or forms layers, with the warmest water at the lake surface.

As air temperatures drop in the fall, the temperature of the lake water cools to 4 °C; therefore, this causes fall turnover as the heavy cold water sinks and displaces the water at the bottom. The oxygen-rich water at the surface of the lake then moves to the bottom of the lake, while the nutrients at the bottom of the lake rise to the surface (Figure 5). During the winter, the oxygen at the bottom of the lake is used by decomposers and other organisms requiring oxygen, such as fish.

Temperature

Temperature affects the physiology of living things as well as the density and state of water. Temperature exerts an important influence on living things because few living things can survive at temperatures below 0 °C (32 °F) due to metabolic constraints. It is also rare for living things to survive at temperatures exceeding 45 °C (113 °F); this is a reflection of evolutionary response to typical temperatures. Enzymes are most efficient within a narrow and specific range of temperatures; enzyme degradation can occur at higher temperatures. Therefore, organisms either must maintain an internal temperature or they must inhabit an environment that will keep the body within a temperature range that supports metabolism. Some animals have adapted to enable their bodies to survive significant temperature fluctuations, such as seen in hibernation or reptilian torpor. Similarly, some bacteria are adapted to surviving in extremely hot temperatures such as geysers.

Such bacteria are examples of extremophiles: organisms that thrive in extreme environments.

Temperature can limit the distribution of living things. Animals faced with temperature fluctuations may respond with adaptations, such as migration, in order to survive. **Migration**, the movement from one place to another, is an adaptation found in many animals, including many that inhabit seasonally cold climates. Migration solves problems related to temperature, locating food, and finding a mate. In migration, for instance, the Arctic Tern (*Sterna paradisaea*) makes a 40,000 km (24,000 mi) round trip flight each year between its feeding grounds in the southern hemisphere and its breeding grounds in the Arctic Ocean. Monarch butterflies (*Danaus plexippus*) live in the eastern United States in the warmer months and migrate to Mexico and the southern United States in the wintertime. Some species of mammals also make migratory forays. Reindeer (*Rangifer tarandus*) travel about 5,000 km (3,100 mi) each year to find food. Amphibians and reptiles are more limited in their distribution because they lack migratory ability. Not all animals that can migrate do so: migration carries risk and comes at a high energy cost.

Some animals hibernate or estivate to survive hostile temperatures. **Hibernation** enables animals to survive cold conditions, and estivation allows animals to survive the hostile conditions of a hot, dry climate. Animals that hibernate or estivate enter a state known as **torpor**: a condition in which their metabolic rate is significantly lowered. This enables the animal to wait until its environment better supports its survival. Some amphibians, such as the wood frog (*Rana sylvatica*), have an antifreeze-like chemical in their cells, which retains the cells' integrity and prevents them from bursting.

Water

Water is required by all living things because it is critical for cellular processes. Since terrestrial organisms lose water to the environment by simple diffusion, they have evolved many adaptations to retain water.

- Plants have a number of interesting features on their leaves, such as leaf hairs and a waxy cuticle, that serve to decrease the rate of water loss via **transpiration**.
- Freshwater organisms are surrounded by water and are constantly in danger of having water rush into their cells because of osmosis. Many adaptations of organisms living in freshwater environments have evolved to ensure that solute concentrations in their bodies remain within appropriate levels. One such adaptation is the excretion of dilute urine.
- Marine organisms are surrounded by water with a higher solute concentration than the organism and, thus, are in danger of losing water to the environment because of osmosis. These organisms have morphological and physiological adaptations to retain water and release solutes into the environment. For example, Marine iguanas (*Amblyrhynchus cristatus*), sneeze out water vapor that is high in salt in order to maintain solute concentrations within an acceptable range while swimming in the ocean and eating marine plants.

Inorganic Nutrients and Soil

Inorganic nutrients, such as **nitrogen** and **phosphorus**, are

important in the distribution and abundance of living things. Plants obtain these inorganic nutrients from the soil when water moves into the plant through the roots. Therefore, soil structure (particle size of soil components), soil pH, and soil nutrient content play an important role in the distribution of plants. Animals obtain inorganic nutrients from the food they consume. Therefore, animal distributions are related to the distribution of what they eat. In some cases, animals will follow their food resource as it moves through the environment.

Other Aquatic Factors

Some abiotic factors, such as **oxygen**, are important in aquatic ecosystems as well as terrestrial environments. Terrestrial animals obtain oxygen from the air they breathe. Oxygen availability can be an issue for organisms living at very high elevations, however, where there are fewer molecules of oxygen in the air. In aquatic systems, the concentration of dissolved oxygen is related to water temperature and the speed at which the water moves. Cold water has more dissolved oxygen than warmer water. In addition, salinity, current, and tide can be important abiotic factors in aquatic ecosystems.

Other Terrestrial Factors

Wind can be an important abiotic factor because it influences the rate of evaporation and transpiration. The physical force of wind is also important because it can move soil, water, or other abiotic factors, as well as an ecosystem's organisms.

Fire is another terrestrial factor that can be an important agent of disturbance in terrestrial ecosystems. Some organisms are adapted to fire and, thus, require the high heat associated with fire to complete a part of their life cycle.

For example, the jack pine—a coniferous tree—requires heat from fire for its seed cones to open (Figure 6). Through the burning of pine needles, fire adds nitrogen to the soil and limits competition by destroying undergrowth.

Figure 6: The mature cones of the jack pine (Pinus banksiana) open only when exposed to high temperatures, such as during a forest fire. A fire is likely to kill most vegetation, so a seedling that germinates after a fire is more likely to receive ample sunlight than one that germinates under normal conditions. (credit: USDA. "Pinus banksiana" by OpenStax is licensed under CC BY 4.0)

Abiotic Factors Influencing Plant Growth

Temperature and moisture are important influences on plant production (**primary productivity**) and the amount of organic matter available as food (**net primary productivity**). Net primary productivity is an estimation of all of the organic matter available as food; it is calculated as the total amount of carbon fixed per year minus the amount that is oxidized

during cellular respiration. In terrestrial environments, net primary productivity is estimated by measuring the aboveground biomass per unit area, which is the total mass of living plants, excluding roots. This means that a large percentage of plant biomass that exists underground is not included in this measurement. Net primary productivity is an important variable when considering differences in biomes. Very productive biomes have a high level of aboveground biomass.

Annual biomass production is directly related to the abiotic components of the environment. Environments with the greatest amount of biomass have conditions in which photosynthesis, plant growth, and the resulting net primary productivity are optimized. The climate of these areas is warm and wet. Photosynthesis can proceed at a high rate, enzymes can work most efficiently, and stomata can remain open without the risk of excessive transpiration; together, these factors lead to the maximal amount of carbon dioxide (CO_2) moving into the plant, resulting in high biomass production. The aboveground biomass produces several important resources for other living things, including habitat and food. Conversely, dry and cold environments have lower photosynthetic rates and therefore less biomass. The animal communities living there will also be affected by the decrease in available food.

Summary

Biogeography is the study of the geographic distribution of living things and the abiotic factors that affect their distribution. Endemic species are species that are naturally found only in a specific geographic area. The distribution of living things is influenced by several environmental factors that are, in part, controlled by the latitude or elevation at which an organism is found. Ocean upwelling and spring

and fall turnovers are important processes regulating the distribution of nutrients and other abiotic factors important in aquatic ecosystems. Energy sources, temperature, water, inorganic nutrients, and soil are factors limiting the distribution of living things in terrestrial systems. Net primary productivity is a measure of the amount of biomass produced by a biome.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Terrestrial Biomes

The Earth's **biomes** are categorized into two major groups: **terrestrial** and **aquatic**. Terrestrial biomes are based on land, while aquatic biomes include both ocean and freshwater biomes. The eight major terrestrial biomes on Earth are each distinguished by characteristic temperatures and amount of precipitation. Comparing the annual totals of precipitation and fluctuations in precipitation from one biome to another provides clues as to the importance of abiotic factors in the distribution of biomes. Temperature variation on a daily and seasonal basis is also important for predicting the geographic distribution of the biome and the vegetation type in the biome. The distribution of these biomes shows that the same biome can occur in geographically distinct areas with similar climates (Figure 1).

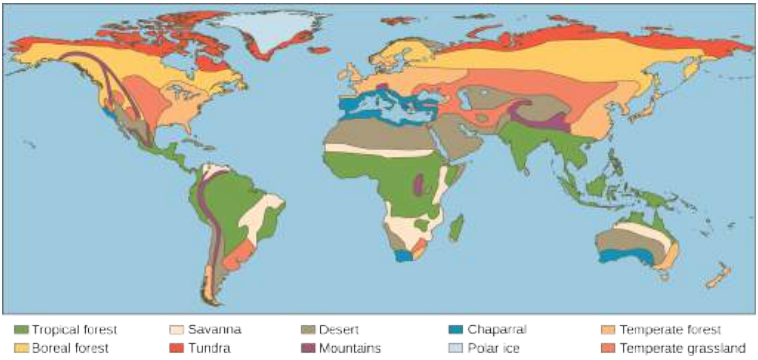


Figure 1: Each of the world's major biomes is distinguished by characteristic temperatures and amounts of precipitation. Polar ice and mountains are also shown. (credit: "biome distribution" by OpenStax is licensed under CC BY 4.0)

Tropical Wet Forest

Tropical wet forests are also referred to as **tropical rainforests**. This biome is found in equatorial regions (Figure 1). The vegetation is characterized by plants with broad leaves that fall off throughout the year. Unlike the trees of deciduous forests, the trees in this biome do not have a seasonal loss of leaves associated with variations in temperature and sunlight; these forests are "evergreen" year-round.

The temperature and sunlight profiles of tropical wet forests are very stable in comparison to that of other terrestrial biomes, with the temperatures ranging from 20 °C to 34 °C (68 °F to 93 °F). When one compares the annual temperature variation of tropical wet forests with that of other forest biomes, the lack of seasonal temperature variation in the tropical wet forest becomes apparent. This lack of seasonality leads to year-round plant growth, rather

than the seasonal (spring, summer, and fall) growth seen in other biomes. In contrast to other ecosystems, tropical ecosystems do not have long days and short days during the yearly cycle. Instead, a constant daily amount of sunlight (11–12 hrs per day) provides more solar radiation, thereby, a longer period of time for plant growth.

The annual rainfall in tropical wet forests ranges from 125 to 660 cm (50–200 in) with some monthly variation. While sunlight and temperature remain fairly consistent, annual rainfall is highly variable. Tropical wet forests have wet months in which there can be more than 30 cm (11–12 in) of precipitation, as well as dry months in which there is fewer than 10 cm (3.5 in) of rainfall. However, the driest month of a tropical wet forest still exceeds the *annual* rainfall of some other biomes, such as deserts.

Tropical wet forests have high net primary productivity because the annual temperatures and precipitation values in these areas are ideal for plant growth. Therefore, the extensive biomass present in the tropical wet forest leads to plant communities with very high species diversities (Figure 2). Tropical wet forests have more species of trees than any other biome; on average between 100 and 300 species of trees are present in a single hectare (2.5 acres) of South America. One way to visualize this is to compare the distinctive horizontal layers within the tropical wet forest biome. On the forest floor is a sparse layer of plants and decaying plant matter. Above that is an understory of short shrubby foliage. A layer of trees rises above this understory and is topped by a closed upper canopy—the uppermost overhead layer of branches and leaves. Some additional trees emerge through this closed upper canopy. These layers provide diverse and complex habitats for the variety of plants, fungi, animals, and other organisms within the tropical wet forests. For instance, epiphytes are plants that grow on other plants, which typically are not harmed.

Epiphytes are found throughout tropical wet forest biomes. Many species of animals use the variety of plants and the complex structure of the tropical wet forests for food and shelter. Some organisms live several meters above ground and have adapted to this arboreal lifestyle.

Figure 2: Tropical wet forests, such as these forests of Madre de Dios, Peru, near the Amazon River, have high species diversity. (credit: Roosevelt Garcia. "Tropical wet forests" by OpenStax is licensed under CC BY 4.0)

Savannas

Savannas are grasslands with scattered trees, and they are located in Africa, South America, and northern Australia (Figure 1). Savannas are hot, tropical areas with temperatures averaging from 24 °C to 29 °C (75 °F to 84 °F) and an annual rainfall of 10–40 cm (3.9–15.7 in). Savannas

have an extensive dry season; for this reason, forest trees do not grow as well as they do in the tropical wet forest (or other forest biomes). As a result, within the grasses and forbs (herbaceous flowering plants) that dominate the savanna, there are relatively few trees (Figure 3). Since fire is an important source of disturbance in this biome, plants have evolved well-developed root systems that allow them to quickly re-sprout after a fire.

Figure 3: Savannas, like this one in Taita Hills Wildlife Sanctuary in Kenya, are dominated by grasses. (credit: Christopher T. Cooper. "Savannas" by OpenStax is licensed under CC BY 4.0)

Subtropical Deserts

Subtropical deserts exist between 15 ° and 30 ° north and south latitude and are centered on the Tropics of Cancer and Capricorn (Figure 1). This biome is very dry; in some years, evaporation exceeds precipitation. Subtropical hot deserts can have daytime soil surface temperatures above 60 °C (140

°F) and nighttime temperatures approaching 0 °C (32 °F). In cold deserts, temperatures can be as high as 25 °C and can drop below -30 °C (-22 °F). Subtropical deserts are characterized by low annual precipitation of fewer than 30 cm (12 in) with little monthly variation and lack of predictability in rainfall. In some cases, the annual rainfall can be as low as 2 cm (0.8 in) in subtropical deserts located in central Australia ("the Outback") and northern Africa.

The vegetation and low animal diversity of this biome are closely related to this low and unpredictable precipitation. Very dry deserts lack perennial vegetation that lives from one year to the next; instead, many plants are annuals that grow quickly and reproduce when rainfall does occur, then they die. Many other plants in these areas are characterized by having a number of adaptations that conserve water, such as deep roots, reduced foliage, and water-storing stems (Figure 4). Seed plants in the desert produce seeds that can be in dormancy for extended periods between rains. Adaptations in desert animals include nocturnal behavior and burrowing.

Figure 4: To reduce water loss, many desert plants have tiny leaves or no leaves at all. The leaves of ocotillo (Fouquieria splendens), shown here in the Sonora Desert near Gila Bend, Arizona, appear only after rainfall, and then are shed. (credit: "Subtropical deserts" by OpenStax is

Chaparral

The **chaparral** is also called the **scrub forest** and is found in California, along the Mediterranean Sea, and along the southern coast of Australia (Figure 1). The annual rainfall in this biome ranges from 65 cm to 75 cm (25.6–29.5 in), and the majority of the rain falls in the winter. Summers are very dry and many chaparral plants are dormant during the summertime. The chaparral vegetation, shown in Figure 5, is dominated by shrubs and is adapted to periodic fires, with some plants producing seeds that only germinate after a hot fire. The ashes left behind after a fire are rich in nutrients like nitrogen that fertilize the soil and promote plant regrowth.

Figure 5: The chaparral is dominated by shrubs. (credit: Miguel Vieira. "chaparral" by OpenStax is licensed under CC BY 4.0)

Temperate Grasslands

Temperate grasslands are found throughout central North America, where they are also known as **prairies**; they are also in Eurasia, where they are known as steppes (Figure 1). Temperate grasslands have pronounced annual fluctuations in temperature with hot summers and cold winters. The annual temperature variation produces specific growing seasons for plants. Plant growth is possible when temperatures are warm enough to sustain plant growth and when ample water is available, which occurs in the spring, summer, and fall. During much of the winter, temperatures are low, and water, which is stored in the form of ice, is not available for plant growth.

Annual precipitation ranges from 25 cm to 75 cm (9.8–29.5 in). Because of relatively lower annual precipitation in temperate grasslands, there are few trees except for those found growing along rivers or streams. The dominant vegetation tends to consist of grasses and some prairies sustain populations of grazing animals Figure 6. The vegetation is very dense and the soils are fertile because the subsurface of the soil is packed with the roots and **rhizomes** (underground stems) of these grasses. The roots and rhizomes act to anchor plants into the ground and replenish the organic material (**humus**) in the soil when they die and decay.

Figure 6: The American bison (Bison bison), more commonly called the buffalo, is a grazing mammal that once populated American prairies in huge numbers. (credit: Jack Dykinga, USDA Agricultural Research Service. "Temperate grasslands" by OpenStax is licensed under CC BY 4.0)

Fires, mainly caused by lightning, are a natural disturbance in temperate grasslands. When fire is suppressed in temperate grasslands, the vegetation eventually converts to scrub and dense forests. Often, the restoration or management of temperate grasslands requires the use of controlled burns to suppress the growth of trees and maintain the grasses.

Temperate Forests

Temperate forests are the most common biome in eastern North America, Western Europe, Eastern Asia, Chile, and New Zealand (Figure 1). This biome is found throughout mid-latitude regions. Temperatures range between -30 °C and 30 °C (-22 °F to 86 °F) and drop to below freezing on an annual basis. These temperatures mean that temperate forests have

defined growing seasons during the spring, summer, and early fall. Precipitation is relatively constant throughout the year and ranges between 75 cm and 150 cm (29.5–59 in).

Because of the moderate annual rainfall and temperatures, deciduous trees are the dominant plant in this biome (Figure 7). **Deciduous trees** lose their leaves each fall and remain leafless in the winter. Thus, no photosynthesis occurs in the deciduous trees during the dormant winter period. Each spring, new leaves appear as the temperature increases. Because of the dormant period, the net primary productivity of temperate forests is less than that of tropical wet forests. In addition, temperate forests show less diversity of tree species than tropical wet forest biomes.

Figure 7: Deciduous trees are the dominant plant in the temperate forest. (credit: Oliver Herold. "Temperate forests" by OpenStax is licensed under CC BY 4.0)

The trees of the temperate forests leaf out and shade much of the ground; however, this biome is more open than

tropical wet forests because trees in the temperate forests do not grow as tall as the trees in tropical wet forests. The soils of the temperate forests are rich in inorganic and organic nutrients. This is due to the thick layer of leaf litter on forest floors. As this leaf litter decays, nutrients are returned to the soil. The leaf litter also protects soil from erosion, insulates the ground, and provides habitats for **invertebrates** (such as the pill bug or roly-poly, *Armadillidium vulgare*) and their predators, such as the red-backed salamander (*Plethodon cinereus*).

Boreal Forests

The **boreal forest**, also known as **taiga** or **coniferous forest**, is found south of the Arctic Circle and across most of Canada, Alaska, Russia, and northern Europe (Figure 1). This biome has cold, dry winters and short, cool, wet summers. The annual precipitation is from 40 cm to 100 cm (15.7–39 in) and usually takes the form of snow. Little evaporation occurs because of the cold temperatures.

The long and cold winters in the boreal forest have led to the predominance of cold-tolerant cone-bearing plants. These are **evergreen** coniferous trees like pines, spruce, and fir, which retain their needle-shaped leaves year-round. Evergreen trees can photosynthesize earlier in the spring than deciduous trees because less energy from the sun is required to warm a needle-like leaf than a broad leaf. This benefits evergreen trees, which grow faster than deciduous trees in the boreal forest. In addition, soils in boreal forest regions tend to be acidic with little available nitrogen. Leaves are a nitrogen-rich structure and deciduous trees must produce a new set of these nitrogen-rich structures each year. Therefore, coniferous trees that retain nitrogen-rich needles may have a competitive advantage over the broad-leafed deciduous trees.

The net primary productivity of boreal forests is lower than that of temperate forests and tropical wet forests. The aboveground biomass of boreal forests is high because these slow-growing tree species are long-lived and accumulate standing biomass over time. Plant species diversity is less than that seen in temperate forests and tropical wet forests. Boreal forests lack the pronounced elements of the layered forest structure seen in tropical wet forests. The structure of a boreal forest is often only a tree layer and a ground layer (Figure 8). When conifer needles are dropped, they decompose more slowly than broad leaves; therefore, fewer nutrients are returned to the soil to fuel plant growth.

Figure 8: The boreal forest (taiga) has low lying plants and conifer trees. (credit: L.B. Brubaker. "boreal forest" by OpenStax is licensed under CC BY 4.0)

Arctic Tundra

The **Arctic tundra** lies north of the subarctic boreal forest

and is located throughout the Arctic regions of the northern hemisphere (Figure 1). The average winter temperature is -34 °C (-34 °F) and the average summer temperature is from 3 °C to 12 °C (37 °F–52 °F). Plants in the arctic tundra have a very short growing season of approximately 10–12 weeks. However, during this time, there are almost 24 hours of daylight and plant growth is rapid. The annual precipitation of the Arctic tundra is very low with little annual variation in precipitation. And, as in the boreal forests, there is little evaporation due to the cold temperatures.

Plants in the Arctic tundra are generally low to the ground (Figure 9). There is little species diversity, low net primary productivity, and low aboveground biomass. The soils of the Arctic tundra may remain in a perennially frozen state referred to as permafrost. The permafrost makes it impossible for roots to penetrate deep into the soil and slows the decay of organic matter, which inhibits the release of nutrients from organic matter. During the growing season, the ground of the Arctic tundra can be completely covered with plants or lichens.

Figure 9: Low-growing plants such as shrub willow dominate the tundra landscape, shown here in the Arctic National Wildlife Refuge. (credit: USFWS Arctic National Wildlife Refuge. "Arctic tundra" by OpenStax is licensed under CC BY 4.0)

Summary

The Earth has terrestrial biomes and aquatic biomes. Aquatic biomes include both freshwater and marine environments. There are eight major terrestrial biomes: tropical wet forests, savannas, subtropical deserts, chaparral, temperate grasslands, temperate forests, boreal forests, and Arctic tundra. The same biome can occur in different geographic locations with similar climates. Temperature and precipitation, and variations in both, are key abiotic factors that shape the composition of animal and plant communities in terrestrial biomes. Some biomes, such as temperate grasslands and temperate forests, have distinct seasons, with cold weather and hot weather alternating throughout

the year. In warm, moist biomes, such as the tropical wet forest, net primary productivity is high, as warm temperatures, abundant water, and a year-round growing season fuel plant growth. Other biomes, such as deserts and tundra, have low primary productivity due to extreme temperatures and a shortage of available water.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 25, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Aquatic Biomes

Abiotic Factors Influencing Aquatic Biomes

Like terrestrial biomes, aquatic biomes are influenced by a series of abiotic factors. The aquatic medium—water— has different physical and chemical properties than air, however. Even if the water in a pond or other body of water is perfectly clear (there are no suspended particles), water, on its own, absorbs light. As one descends into a deep body of water, there will eventually be a depth which the sunlight cannot reach. While there are some abiotic and biotic factors in a terrestrial ecosystem that might obscure light (like fog, dust, or insect swarms), usually these are not permanent features of the environment. The importance of light in aquatic biomes is central to the communities of organisms found in both freshwater and marine ecosystems. In freshwater systems, **stratification** due to differences in density is perhaps the most critical abiotic factor and is related to the energy aspects of light. The **thermal** properties of water (rates of heating and cooling) are significant to the function of marine systems and have major impacts on global climate and weather patterns. Marine systems are also influenced by large-scale physical water movements, such as **currents**; these are less important in most freshwater lakes.

The ocean is categorized into several areas or zones (Figure 1). All of the ocean's open water is referred to as the **pelagic**

realm (or zone). The **benthic realm** (or zone) extends along the ocean bottom from the shoreline to the deepest parts of the ocean floor. Within the pelagic realm is the **photic zone**, which is the portion of the ocean that light can penetrate (approximately 200 m or 650 ft). At depths greater than 200 m, light cannot penetrate; thus, this is referred to as the **aphotic zone**. The majority of the ocean is aphotic and lacks sufficient light for photosynthesis. The deepest part of the ocean, the Challenger Deep (in the Mariana Trench, located in the western Pacific Ocean), is about 11,000 m (about 6.8 mi) deep. To give some perspective on the depth of this trench, the ocean is, on average, 4267 m or 14,000 ft deep. These realms and zones are relevant to freshwater lakes as well.

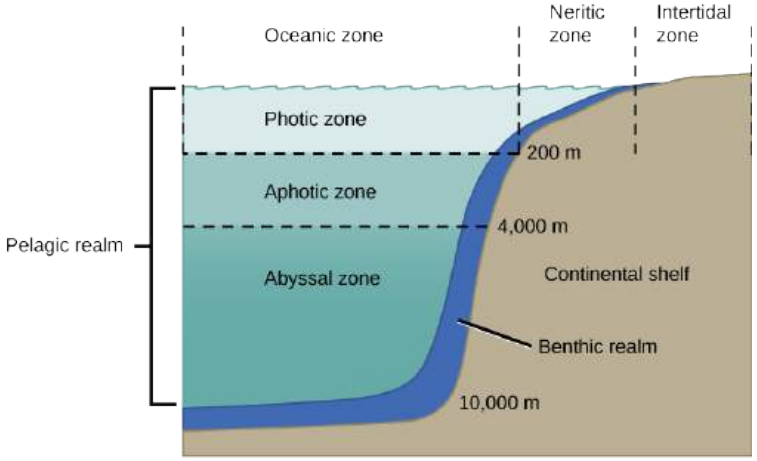


Figure 1: The ocean is divided into different zones based on water depth and distance from the shoreline. (credit: "Ocean zones" by OpenStax is licensed under CC BY 4.0)

Marine Biomes

The **ocean** is the largest marine biome. It is a continuous body of saltwater that is relatively uniform in chemical composition; it is a weak solution of mineral salts and decayed biological matter. Within the ocean, **coral reefs** are the second kind of marine biome. **Estuaries**, coastal areas where salt water and fresh water mix, form a third unique marine biome.

Ocean

The physical diversity of the ocean is a significant influence on plants, animals, and other organisms. The ocean is categorized into different zones based on how far light reaches into the water. Each zone has a distinct group of species adapted to the biotic and abiotic conditions particular to that zone.

The **intertidal zone**, which is the zone between high and low tide, is the oceanic region that is closest to land (Figure 1). Generally, most people think of this portion of the ocean as a sandy beach. In some cases, the intertidal zone is indeed a sandy beach, but it can also be rocky or muddy. The intertidal zone is an extremely variable environment because of tides. Organisms are exposed to air and sunlight at low tide and are underwater most of the time, especially during high tide. Therefore, living things that thrive in the intertidal zone are adapted to being dry for long periods of time. The shore of the intertidal zone is also repeatedly struck by waves, and the organisms found there are adapted to withstand damage from the pounding action of the waves (Figure 2). The exoskeletons of shoreline crustaceans (such as the shore crab, *Carcinus maenas*) are tough and protect them from desiccation (drying out) and wave damage. Another consequence of the pounding waves is that few algae and

plants establish themselves in the constantly moving rocks, sand, or mud.

Figure 2: Sea urchins, mussel shells, and starfish are often found in the intertidal zone, shown here in Kachemak Bay, Alaska. (credit: NOAA. "seastars" by OpenStax is licensed under CC BY 4.0)

The **neritic zone** (Figure 1) extends from the intertidal zone to depths of about 200 m (or 650 ft) at the edge of the continental shelf. Since light can penetrate this depth, photosynthesis can occur in the neritic zone. The water here contains silt and is well-oxygenated, low in pressure, and stable in temperature. **Phytoplankton** and floating *Sargassum* (a type of free-floating marine seaweed) provide a habitat for some sea life found in the neritic zone. **Zooplankton**, protists, small fishes, and shrimp are found in the neritic zone and are the base of the food chain for most of the world's fisheries.

Beyond the neritic zone is the open ocean area known as the **oceanic zone** (Figure 1). Within the oceanic zone, there

is thermal stratification where warm and cold waters mix because of ocean currents. Abundant plankton serves as the base of the food chain for larger animals such as whales and dolphins. Nutrients are scarce and this is a relatively less productive part of the marine biome. When photosynthetic organisms and the protists and animals that feed on them die, their bodies fall to the bottom of the ocean where they remain; unlike freshwater lakes, the open ocean lacks a process for bringing the organic nutrients back up to the surface. The majority of organisms in the aphotic zone include sea cucumbers (phylum Echinodermata) and other organisms that survive on the nutrients contained in the dead bodies of organisms in the photic zone.

Beneath the pelagic zone is the **benthic realm**, the deepwater region beyond the continental shelf (Figure 1). The bottom of the benthic realm is comprised of sand, silt, and dead organisms. Temperature decreases, remaining above freezing, as water depth increases. This is a nutrient-rich portion of the ocean because of the dead organisms that fall from the upper layers of the ocean. Because of this high level of nutrients, a diversity of fungi, sponges, sea anemones, marine worms, sea stars, fishes, and bacteria exist.

The deepest part of the ocean is the **abyssal zone**, which is at depths of 4000 m or greater. The abyssal zone (Figure 1) is very cold and has very high pressure, high oxygen content, and low nutrient content. There are a variety of invertebrates and fishes found in this zone, but the abyssal zone does not have plants because of the lack of light. **Hydrothermal** vents are found primarily in the abyssal zone; **chemosynthetic** bacteria utilize the hydrogen sulfide and other minerals emitted from the vents. These chemosynthetic bacteria use hydrogen sulfide as an energy source and serve as the base of the food chain found in the abyssal zone.

Coral Reefs

Coral reefs are ocean ridges formed by marine invertebrates living in warm shallow waters within the photic zone of the ocean. They are found within 30˚ north and south of the equator. The Great Barrier Reef is a well-known reef system located several miles off the northeastern coast of Australia. Other coral reef systems are fringing islands, which are directly adjacent to land, or atolls, which are circular reef systems surrounding a former landmass that is now underwater. The coral organisms (members of phylum Cnidaria) are colonies of saltwater polyps that secrete a calcium carbonate skeleton. These calcium-rich skeletons slowly accumulate, forming the underwater reef (Figure 3). Corals found in shallower waters (at a depth of approximately 60 m or about 200 ft) have a mutualistic relationship with photosynthetic unicellular algae. The relationship provides corals with the majority of the nutrition and the energy they require. The waters in which these corals live are nutritionally poor and, without this mutualism, it would not be possible for large corals to grow. Some corals living in deeper and colder water do not have a mutualistic relationship with algae; these corals attain energy and nutrients using stinging cells on their tentacles to capture prey.

It is estimated that more than 4,000 fish species inhabit coral reefs. These fishes can feed on coral, the cryptofauna (invertebrates found within the calcium carbonate substrate of the coral reefs), or the seaweed and algae that are associated with the coral. In addition, some fish species inhabit the boundaries of a coral reef; these species include predators, herbivores, or planktivores. Predators are animal species that hunt and are **carnivores** or "flesh-eaters." **Herbivores** eat plant material, and **planktivores** eat plankton.

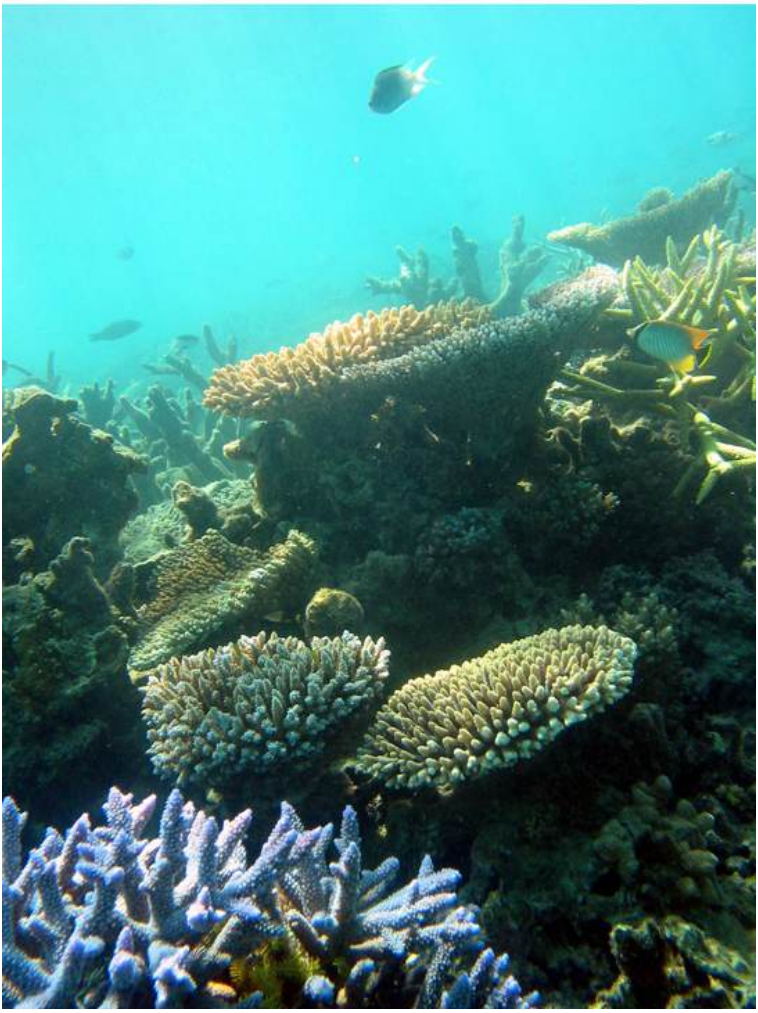


Figure 3: Coral reefs are formed by the calcium carbonate skeletons of coral organisms, which are marine invertebrates in the phylum Cnidaria. (credit: Terry Hughes. "Coral reefs" by OpenStax is licensed under CC BY 4.0)

Evolution Connection – Global Decline of Coral Reefs

It takes a long time to build a coral reef. The animals that create coral reefs have evolved over millions of years, continuing to slowly deposit the calcium carbonate that forms their characteristic ocean homes. Bathed in warm tropical waters, the coral animals and their symbiotic algal partners evolved to survive at the upper limit of ocean water temperature.

Together, climate change and human activity pose dual threats to the long-term survival of the world's coral reefs. As global warming due to fossil fuel emissions raises ocean temperatures, coral reefs are suffering. The excessive warmth causes the reefs to expel their symbiotic, food-producing algae, resulting in a phenomenon known as **bleaching**. When bleaching occurs, the reefs lose much of their characteristic color as the algae and the coral animals die if the loss of the symbiotic zooxanthellae is prolonged.

Rising levels of atmospheric carbon dioxide further threaten the corals in other ways; as CO_2 dissolves in ocean waters, it lowers the pH and increases **ocean acidity**. As acidity increases, it interferes with the calcification that normally occurs as coral animals build their calcium carbonate homes.

When a coral reef begins to die, species diversity plummets as animals lose food and shelter. Coral reefs are also economically important tourist destinations, so the decline of coral reefs poses a serious threat to coastal economies.

Human population growth has damaged corals in other ways, too. As human coastal populations increase, the runoff of sediment and agricultural chemicals has increased, too, causing some of the once-clear tropical waters to become cloudy. At the same time, overfishing of popular fish species has allowed the predator species that eat corals to go unchecked.

Although a rise in global temperatures of 1–2˚C (a conservative scientific projection) in the coming decades may not seem large, it is very significant to this biome. When change occurs rapidly, species can become extinct before evolution leads to new adaptations. Many scientists believe that global warming, with its rapid (in terms of evolutionary time) and inexorable increases in temperature, is tipping the balance beyond the point at which many of the world's coral reefs can recover.

Estuaries: Where the Ocean Meets Fresh Water

Estuaries are biomes that occur where a source of fresh water, such as a river, meets the ocean. Therefore, both fresh water and salt water are found in the same vicinity; mixing results in a diluted (**brackish**) saltwater. Estuaries form protected areas where many of the young offspring of **crustaceans**, **mollusks**, and fish begin their lives. Salinity is a very important factor that influences the organisms and the adaptations of the organisms found in estuaries. The salinity of estuaries varies and is based on the rate of flow of its freshwater sources. Once or twice a day, high tides bring salt water into the estuary. Low tides occurring at the same frequency reverse the current of salt water.

The short-term and rapid variation in salinity due to the mixing of fresh water and salt water is a difficult physiological challenge for the plants and animals that inhabit estuaries. Many estuarine plant species are **halophytes**: plants that can tolerate salty conditions. Halophytic plants are adapted to deal with the salinity resulting from saltwater on their roots or from sea spray. In some halophytes, filters in the roots remove the salt from the water that the plant absorbs. Other plants are able to pump oxygen into their roots. Animals, such as mussels and clams (phylum Mollusca), have developed behavioral adaptations that expend a lot of

energy to function in this rapidly changing environment. When these animals are exposed to low salinity, they stop feeding, close their shells, and switch from aerobic respiration (in which they use gills) to anaerobic respiration (a process that does not require oxygen). When high tide returns to the estuary, the salinity and oxygen content of the water increases and these animals open their shells, begin feeding, and return to aerobic respiration.

Freshwater Biomes

Freshwater biomes include **lakes** and **ponds** (standing water) as well as **rivers** and **streams** (flowing water). They also include **wetlands**, which will be discussed later. Humans rely on freshwater biomes to provide aquatic resources for drinking water, crop irrigation, sanitation, and industry. These various roles and human benefits are referred to as ecosystem services. Lakes and ponds are found in terrestrial landscapes and are, therefore, connected with abiotic and biotic factors influencing these terrestrial biomes.

Lakes and Ponds

Lakes and **ponds** can range in area from a few square meters to thousands of square kilometers. Temperature is an important abiotic factor affecting living things found in lakes and ponds. In the summer, thermal stratification of lakes and ponds occurs when the upper layer of water is warmed by the sun and does not mix with deeper, cooler water. Light can penetrate within the photic zone of the lake or pond. **Phytoplankton** (algae and cyanobacteria) are found here and carry out photosynthesis, providing the base of the food web of lakes and ponds. **Zooplankton**, such as rotifers and small crustaceans, consume these phytoplankton. At the bottom of lakes and ponds, bacteria in

the aphotic zone break down dead organisms that sink to the bottom.

Nitrogen and phosphorus are important limiting nutrients in lakes and ponds. Because of this, they are **determining factors** in the amount of phytoplankton growth in lakes and ponds. When there is a large input of nitrogen and phosphorus (from sewage and runoff from fertilized lawns and farms, for example), the growth of algae skyrockets, resulting in a large accumulation of algae called an **algal bloom**. Algal blooms (Figure 4) can become so extensive that they reduce light penetration in water. As a result, the lake or pond becomes aphotic and photosynthetic plants cannot survive. When the algae die and decompose, severe oxygen depletion of the water occurs. Fishes and other organisms that require oxygen are then more likely to die, and resulting dead zones are found across the globe. Lake Erie and the Gulf of Mexico represent freshwater and marine habitats where phosphorus control and storm water runoff pose significant environmental challenges.

Figure 4: The uncontrolled growth of algae in this lake has resulted in an algal bloom. (credit: Jeremy Nettleton. "uncontrolled growth of algae" by OpenStax is licensed under CC BY 4.0)

Rivers and Streams

Rivers and **streams** are continuously moving bodies of water that carry large amounts of water from the source, or headwater, to a lake or ocean. The largest rivers include the Nile River in Africa, the Amazon River in South America, and the Mississippi River in North America.

Abiotic features of rivers and streams vary along the length of the river or stream. Streams begin at a point of origin referred to as source water. The source water is usually cold, low in nutrients, and clear. The **channel** (the width of the river or stream) is narrower than at any other place along the length of the river or stream. Because of this, the current

is often faster here than at any other point of the river or stream.

The fast-moving water results in minimal silt accumulation at the bottom of the river or stream; therefore, the water is clear. Photosynthesis here is mostly attributed to algae that are growing on rocks; the swift current inhibits the growth of phytoplankton. An additional input of energy can come from leaves or other organic material that falls into the river or stream from trees and other plants that border the water. When the leaves decompose, the organic material and nutrients in the leaves are returned to the water. Plants and animals have adapted to this fast-moving water. For instance, leeches (phylum Annelida) have elongated bodies and suckers on both ends. These suckers attach to the substrate, keeping the leech anchored in place. Freshwater trout species (phylum Chordata) are important predators in these fast-moving rivers and streams.

As the river or stream flows away from the source, the width of the channel gradually widens and the current slows. This slow-moving water, caused by the gradient decrease and the volume increase as tributaries unite, has more **sedimentation**. Phytoplankton can also be suspended in slow-moving water. Therefore, the water will not be as clear as it is near the source. The water is also warmer. Worms (phylum Annelida) and insects (phylum Arthropoda) can be found burrowing into the mud. The higher-order predator vertebrates (phylum Chordata) include waterfowl, frogs, and fishes. These predators must find food in these slow-moving, sometimes murky, waters, and, unlike the trout in the waters at the source, these vertebrates may not be able to use vision as their primary sense to find food. Instead, they are more likely to use taste or chemical cues to find prey.

Wetlands

Wetlands are environments in which the soil is either permanently or periodically saturated with water. Wetlands are different from lakes because wetlands are shallow bodies of water whereas lakes vary in depth. Emergent vegetation consists of wetland plants that are rooted in the soil but have portions of leaves, stems, and flowers extending above the water's surface. There are several types of wetlands including marshes, swamps, bogs, mudflats, and salt marshes (Figure 5). The three shared characteristics among these types—what makes them wetlands—are their hydrology, hydrophytic vegetation, and hydric soils.

Figure 5: Located in southern Florida, Everglades National Park is vast array of wetland environments, including sawgrass marshes, cypress swamps, and estuarine mangrove forests. Here, a great egret walks among cypress trees. (credit: NPS. "Everglades National Park" by OpenStax is licensed under CC BY 4.0)

Freshwater marshes and **swamps** are characterized by slow and steady water flow. **Bogs** develop in depressions where water flow is low or nonexistent. Bogs usually occur in areas where there is a clay bottom with poor percolation. **Percolation** is the movement of water through the pores in the soil or rocks. The water found in a bog is stagnant and oxygen-depleted because the oxygen that is used during the decomposition of organic matter is not replaced. As the oxygen in the water is depleted, decomposition slows. This leads to organic acids and other acids building up and lowering the pH of the water. At a lower pH, nitrogen becomes unavailable to plants. This creates a challenge for plants because nitrogen is an important limiting resource. Some types of bog plants (such as sundews, pitcher plants, and Venus flytraps) capture insects and extract nitrogen from their bodies. Bogs have low net primary productivity because the water found in bogs has low levels of nitrogen and oxygen.

Summary

Aquatic ecosystems include both saltwater and freshwater biomes. The abiotic factors important for the structuring of aquatic ecosystems can be different than those seen in terrestrial systems. Sunlight is a driving force behind the structure of forests and also is an important factor in bodies of water, especially those that are very deep, because of the role of photosynthesis in sustaining certain organisms. Density and temperature shape the structure of aquatic systems. Oceans may be thought of as consisting of different zones based on water depth and distance from the shoreline and light penetrance. Different kinds of organisms are adapted to the conditions found in each zone. Coral reefs are unique marine ecosystems that are home to a wide variety of species. Estuaries are found where rivers meet the ocean;

their shallow waters provide nourishment and shelter for young crustaceans, mollusks, fishes, and many other species. Freshwater biomes include lakes, ponds, rivers, streams, and wetlands. Bogs are an interesting type of wetland characterized by standing water, lower pH, and a lack of nitrogen.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Climate and the Effects of Global Climate Change

All biomes are universally affected by global conditions, such as climate, that ultimately shape each biome's environment. Scientists who study climate have noted a series of marked changes that have gradually become increasingly evident during the last sixty years. **Global climate change** is the term used to describe altered global weather patterns, including a worldwide increase in temperature, due largely to rising levels of atmospheric carbon dioxide.

Climate and Weather

A common misconception about global climate change is that a specific weather event occurring in a particular region (for example, a very cool week in June in central Indiana) is evidence of global climate change. However, a cold week in June is a weather-related event and not a climate-related one. These misconceptions often arise because of confusion over the terms climate and weather.

Climate refers to the long-term, predictable atmospheric conditions of a specific area. The climate of a biome is characterized by having consistent temperature and annual rainfall ranges. Climate does not address the amount of rain that fell on one particular day in a biome or the colder-than-

average temperatures that occurred on one day. In contrast, **weather** refers to the conditions of the atmosphere during a short period of time. Weather forecasts are usually made for 48-hour cycles. Long-range weather forecasts are available but can be unreliable.

To better understand the difference between climate and weather, imagine that you are planning an outdoor event in northern Wisconsin. You would be thinking about *climate* when you plan the event in the summer rather than the winter because you have long-term knowledge that any given Saturday in the months of May to August would be a better choice for an outdoor event in Wisconsin than any given Saturday in January. However, you cannot determine the specific day that the event should be held on because it is difficult to accurately predict the weather on a specific day. Climate can be considered "average" weather.

Global Climate Change

Climate change can be understood by approaching three areas of study:

- current and past global climate change
- causes of past and present-day global climate change
- ancient and current results of climate change

It is helpful to keep these three different aspects of climate change clearly separated when consuming media reports about global climate change. It is common for reports and discussions about global climate change to confuse the data showing that Earth's climate is changing with the factors that drive this climate change.

Evidence for Global Climate Change

Since scientists cannot go back in time to directly measure climatic variables, such as average temperature and precipitation, they must instead indirectly measure temperature. To do this, scientists rely on historical evidence of Earth's past climate.

Antarctic ice cores are a key example of such evidence. These ice cores are samples of polar ice obtained by means of drills that reach thousands of meters into ice sheets or high mountain glaciers. Viewing the ice cores is like traveling backward through time; the deeper the sample, the earlier the time period. Trapped within the ice are bubbles of air and other biological evidence that can reveal temperature and carbon dioxide data. Antarctic ice cores have been collected and analyzed to indirectly estimate the temperature of the Earth over the past 400,000 years (Figure 1). The 0 °C on this graph refers to the long-term average. Temperatures that are greater than 0 °C exceed Earth's long-term average temperature. Conversely, temperatures that are less than 0 °C are less than Earth's average temperature. This figure shows that there have been periodic cycles of increasing and decreasing temperature.

Before the late 1800s, the Earth has been as much as 9 °C cooler and about 3 °C warmer. Note that the graph in Figure 1 shows that the atmospheric concentration of carbon dioxide has also risen and fallen in periodic cycles; note the relationship between carbon dioxide concentration and temperature. Figure 1 shows that carbon dioxide levels in the atmosphere have historically cycled between 180 and 300 parts per million (ppm) by volume.

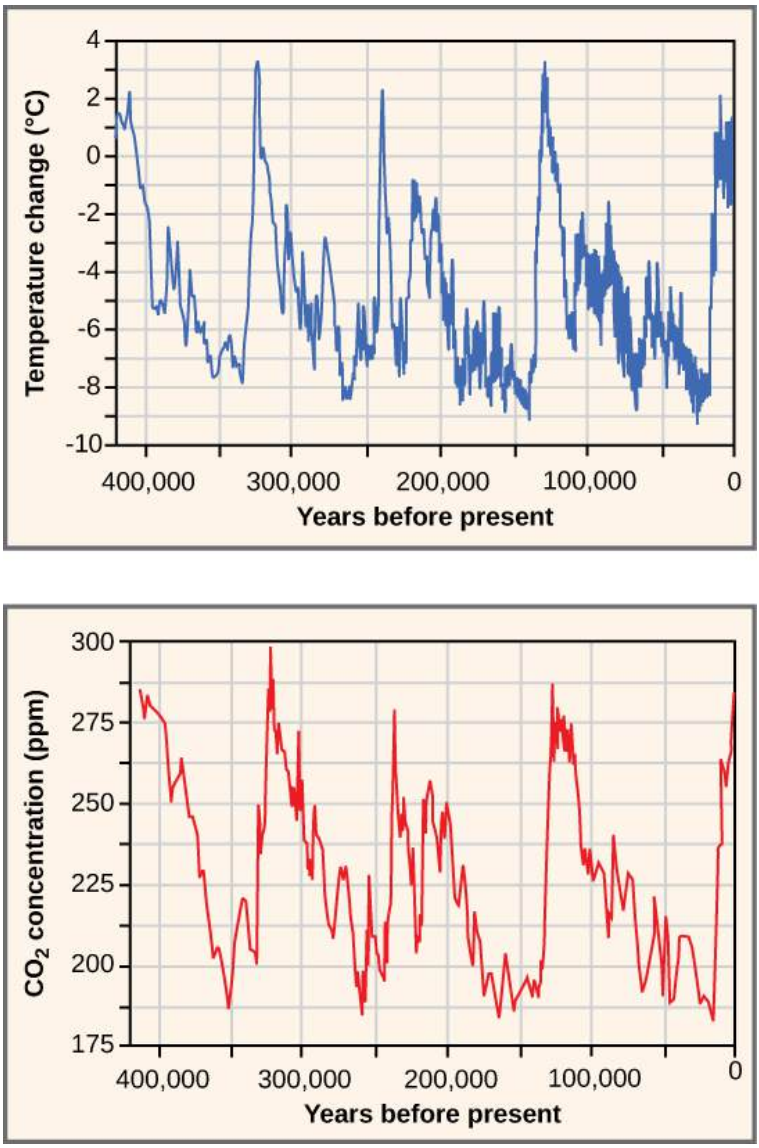


Figure 1: Ice at the Russian Vostok station in East Antarctica was laid down over the course 420,000 years and reached a depth of over 3,000 m. By measuring the amount of CO2 trapped in the ice, scientists have determined past atmospheric CO2 concentrations. Temperatures

relative to modern day were determined from the amount of deuterium (an isotope of hydrogen) present. (credit: "CO2 and temp" by OpenStax is licensed under CC BY 4.0)

Figure 1 does not show the last 2,000 years with enough detail to compare the changes of Earth's temperature during the last 400,000 years with the temperature change that has occurred in the more recent past. Two significant temperature **anomalies**, or irregularities, have occurred in the last 2000 years. These are the Medieval Climate Anomaly (or the Medieval Warm Period) and the Little Ice Age. A third temperature anomaly aligns with the Industrial Era. **The Medieval Climate Anomaly** occurred between 900 and 1300 AD. During this time period, many climate scientists think that slightly warmer weather conditions prevailed in many parts of the world; the higher-than-average temperature changes varied between 0.10 °C and 0.20 °C above the norm. Although 0.10 °C does not seem large enough to produce any noticeable change, it did free seas of ice. Because of this warming, the Vikings were able to colonize Greenland.

The Little Ice Age was a cold period that occurred between 1550 AD and 1850 AD. During this time, a slight cooling of a little less than 1 °C was observed in North America, Europe, and possibly other areas of the Earth. This 1 °C change in global temperature is a seemingly small deviation in temperature (as was observed during the Medieval Climate Anomaly); however, it also resulted in noticeable changes. Historical accounts reveal a time of exceptionally harsh winters with much snow and frost.

The Industrial Revolution, which began around 1750, was characterized by changes in much of human society. Advances in agriculture increased the food supply, which improved the standard of living for people in Europe and the

United States. New technologies were invented and provided jobs and cheaper goods. These new technologies were powered using **fossil fuels**, especially **coal**. The Industrial Revolution starting in the early nineteenth century ushered in the beginning of the Industrial Era. When fossil fuel is burned, **carbon dioxide** is released. With the beginning of the Industrial Era, atmospheric carbon dioxide began to rise (Figure 2).

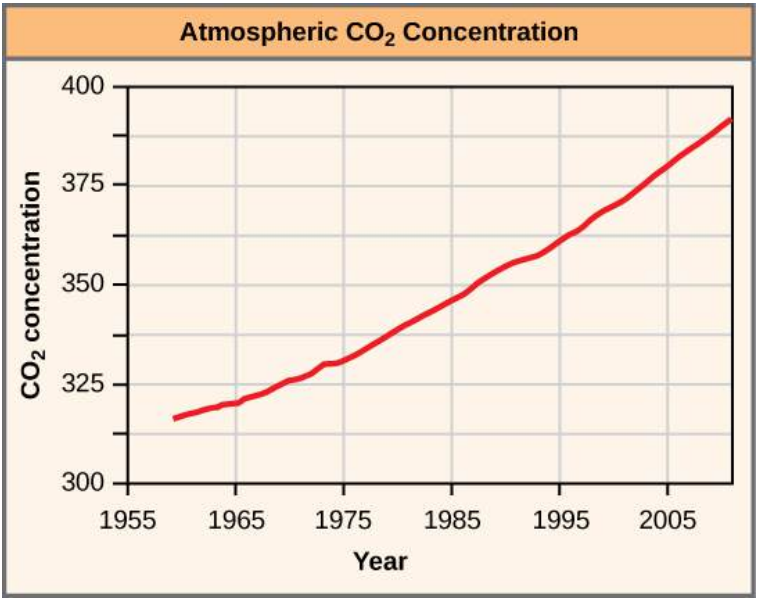


Figure 2: The atmospheric concentration of CO2 has risen steadily since the beginning of industrialization. (credit: "CO2 1955-2005" by OpenStax is licensed under CC BY 4.0)

Current and Past Drivers of Global Climate Change

Since it is not possible to go back in time to directly observe and measure climate, scientists use indirect evidence to

determine the **drivers**, or factors, that may be responsible for climate change. The indirect evidence includes data collected using ice cores, boreholes (a narrow shaft bored into the ground), tree rings, glacier lengths, pollen remains, and ocean sediments. The data shows a correlation between the timing of temperature changes and drivers of climate change: before the Industrial Era (pre-1780), there were three drivers of climate change that were not related to human activity or atmospheric gases. The first of these is the **Milankovitch cycles**. The Milankovitch cycles describe the effects of slight changes in the Earth's orbit on Earth's climate. The length of the Milankovitch cycles ranges between 19,000 and 100,000 years. In other words, one could expect to see some predictable changes in the Earth's climate associated with changes in the Earth's orbit at a minimum of every 19,000 years.

The variation in the **sun's intensity** is the second natural factor responsible for climate change. Solar intensity is the amount of solar power or energy the sun emits in a given amount of time. There is a direct relationship between solar intensity and temperature. As solar intensity increases (or decreases), the Earth's temperature correspondingly increases (or decreases). Changes in solar intensity have been proposed as one of several possible explanations for the Little Ice Age.

Finally, **volcanic eruptions** are a third natural driver of climate change. Volcanic eruptions can last a few days, but the solids and gases released during an eruption can influence the climate over a period of a few years, causing short-term climate changes. The gases and solids released by volcanic eruptions can include carbon dioxide, water vapor, sulfur dioxide, hydrogen sulfide, hydrogen, and carbon monoxide. Generally, volcanic eruptions cool the climate. This occurred in 1783 when volcanos in Iceland erupted and caused the release of large volumes of sulfuric oxide. This

led to haze-effect cooling, a global phenomenon that occurs when dust, ash, or other suspended particles block out sunlight and trigger lower global temperatures as a result; haze-effect cooling usually extends for one or more years. In Europe and North America, haze-effect cooling produced some of the lowest average winter temperatures on record in 1783 and 1784.

Greenhouse gases are probably the most significant drivers of the climate. When heat energy from the sun strikes the Earth, gases known as greenhouse gases trap the heat in the atmosphere, as do the glass panes of a greenhouse keep heat from escaping. The greenhouse gases that affect Earth include **carbon dioxide**, **methane**, **water vapor**, **nitrous oxide**, and **ozone**. Approximately half of the radiation from the sun passes through these gases in the atmosphere and strikes the Earth. This radiation is converted into thermal radiation on the Earth's surface, and then a portion of that energy is re-radiated back into the atmosphere. Greenhouse gases, however, reflect much of the thermal energy back to the Earth's surface. The more greenhouse gases there are in the atmosphere, the more thermal energy is reflected back to the Earth's surface. Greenhouse gases absorb and emit radiation and are an important factor in **the greenhouse effect**: the warming of Earth due to carbon dioxide and other greenhouse gases in the atmosphere.

Evidence supports the relationship between atmospheric concentrations of carbon dioxide and temperature: as carbon dioxide rises, global temperature rises. Since 1950, the concentration of atmospheric carbon dioxide has increased from about 280 ppm to 382 ppm in 2006. In 2011, the atmospheric carbon dioxide concentration was 392 ppm. However, the planet would not be inhabitable by current life forms if water vapor did not produce its drastic greenhouse warming effect.

Scientists look at patterns in data and try to explain

differences or deviations from these patterns. The atmospheric carbon dioxide data reveal a historical pattern of carbon dioxide increasing and decreasing, cycling between a low of 180 ppm and a high of 300 ppm. Scientists have concluded that it took around 50,000 years for the atmospheric carbon dioxide level to increase from its low minimum concentration to its higher maximum concentration. However, starting recently, atmospheric carbon dioxide concentrations have increased beyond the historical maximum of 300 ppm. The current increases in atmospheric carbon dioxide have happened very quickly—in a matter of hundreds of years rather than thousands of years. What is the reason for this difference in the rate of change and the amount of increase in carbon dioxide? A key factor that must be recognized when comparing the historical data and the current data is the presence of modern human society; no other driver of climate change has yielded changes in atmospheric carbon dioxide levels at this rate or to this magnitude.

Human activity releases carbon dioxide and methane, two of the most important greenhouse gases, into the atmosphere in several ways. The primary mechanism that releases carbon dioxide is the burning of fossil fuels, such as **gasoline**, **coal**, and **natural gas** (Figure 3). Deforestation, cement manufacture, animal agriculture, the clearing of land, and the burning of forests are other human activities that release carbon dioxide. **Methane** (CH_4) is produced when bacteria break down organic matter under anaerobic conditions. Anaerobic conditions can happen when organic matter is trapped underwater (such as in rice paddies) or in the intestines of herbivores. Methane can also be released from natural gas fields and the decomposition that occurs in landfills. Another source of methane is the melting of clathrates. **Clathrates** are frozen chunks of ice and methane found at the bottom of the ocean. When the water warms,

these chunks of ice melt and methane is released. As the ocean's water temperature increases, the rate at which clathrates melt is increasing, releasing even more methane. This leads to increased levels of methane in the atmosphere, which further accelerates the rate of global warming. This is an example of the positive feedback loop that is leading to the rapid rate of increase in global temperatures.

Figure 3: The burning of fossil fuels in industry and by vehicles releases carbon dioxide and other greenhouse gases into the atmosphere. (credit: "Pöllö"/Wikimedia Commons. "Pöllö" by OpenStax is licensed under CC BY 4.0)

Documented Results of Climate Change: Past and Present

Scientists have geological evidence of the consequences of long-ago climate change. Modern-day phenomena such as retreating glaciers and melting polar ice cause a continual rise in sea level. Meanwhile, changes in climate can negatively affect organisms.

Geological Climate Change

Global warming has been associated with at least one planet-wide extinction event during the geological past. The **Permian extinction** event occurred about 251 million years ago toward the end of the roughly 50-million-year-long geological time span known as the Permian period. This geologic time period was one of the three warmest periods in Earth's geologic history. Scientists estimate that approximately 70 percent of the terrestrial plant and animal species and 84 percent of marine species became extinct, vanishing forever near the end of the Permian period. Organisms that had adapted to wet and warm climatic conditions, such as annual rainfall of 300–400 cm (118–157 in) and 20 °C–30 °C (68 °F–86 °F) in the tropical wet forest, may not have been able to survive the Permian climate change.

Present Climate Change

A number of global events have occurred that may be attributed to climate change during our lifetimes. Glacier National Park in Montana is undergoing the **retreat** of many of its glaciers, a phenomenon known as glacier recession. In 1850, the area contained approximately 150 glaciers. By 2010, however, the park contained only about 24 glaciers

greater than 25 acres in size. One of these glaciers is the Grinnell Glacier (Figure 4) at Mount Gould. Between 1966 and 2005, the size of Grinnell Glacier shrank by 40 percent. Similarly, the mass of the ice sheets in Greenland and the Antarctic is decreasing: Greenland lost 150–250 km^3 of ice per year between 2002 and 2006. In addition, the size and thickness of the Arctic **sea ice** is decreasing.

Figure 4: The effect of global warming can be seen in the continuing retreat of Grinnel Glacier. The mean annual temperature in the park has increased 1.33 °C since 1900. The loss of a glacier results in the loss of summer meltwaters, sharply reducing seasonal water supplies and severely affecting local ecosystems. (credit: modification of work by USGS. "Grinnel Glacier retreat" by OpenStax is licensed under CC BY 4.0)

This loss of ice is leading to increases in the global sea level. On average, the **sea is rising** at a rate of 1.8 mm per year. However, between 1993 and 2010 the rate of sea level increase ranged between 2.9 and 3.4 mm per year. A variety of factors affect the volume of water in the ocean, including the temperature of the water (the density of water is related to its temperature) and the amount of water found in rivers, lakes, glaciers, polar ice caps, and sea ice. As **glaciers and polar ice caps melt**, there is a significant contribution of liquid water that was previously frozen.

In addition to some abiotic conditions changing in response to climate change, many organisms are also being affected by the changes in temperature. Temperature and precipitation play key roles in determining the geographic distribution and phenology of plants and animals. (**Phenology** is the study of the effects of climatic conditions on the timing of periodic lifecycle events, such as flowering in plants or migration in birds.) Researchers have shown that 385 plant species in Great Britain are flowering 4.5 days sooner than was recorded earlier during the previous 40 years. In addition, insect-pollinated species were more likely to flower earlier than wind-pollinated species. The impact of changes in flowering dates would be mitigated if the insect pollinators emerged earlier. This mismatched timing of plants and pollinators could result in injurious ecosystem effects because, for continued survival, insect-pollinated plants must flower when their pollinators are present.

Summary

The Earth has gone through periodic cycles of increases and decreases in temperature. During the past 2000 years, the Medieval Climate Anomaly was a warmer period, while the Little Ice Age was unusually cool. Both of these irregularities can be explained by natural causes of changes in climate, and, although the temperature changes were small, they had significant effects. Natural drivers of climate change include Milankovitch cycles, changes in solar activity, and volcanic eruptions. None of these factors, however, leads to rapid increases in global temperature or sustained increases in carbon dioxide. The burning of fossil fuels is an important source of greenhouse gases, which plays a major role in the greenhouse effect. Long ago, global warming resulted in the Permian extinction: a large-scale extinction event that is documented in the fossil record. Currently, modern-day

climate change is associated with the increased melting of glaciers and polar ice sheets, resulting in a gradual increase in sea level. Plants and animals can also be affected by global climate change when the timing of seasonal events, such as flowering or pollination, is affected by global warming.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

POPULATION AND COMMUNITY ECOLOGY

Figure 1: Asian carp jump out of the water in response to electrofishing. The Asian carp in the inset photograph were harvested from the Little Calumet River in Illinois in May, 2010, using rotenone, a toxin often used as an insecticide, in an effort to learn more about the population of the species. (credit main image: modification of work by USGS; credit inset: modification of work by Lt. David French, USCG. "Asian carp" by OpenStax is licensed under CC BY 4.0)

Imagine sailing down a river in a small motorboat on a weekend afternoon; the water is smooth and you are enjoying the warm sunshine and cool breeze when suddenly you are hit in the head by a 20-pound silver carp. This is a risk now on many rivers and canal systems in Illinois and Missouri because of the presence of Asian carp (Figure 1).

This fish—actually a group of species including the silver, black, grass, and big head carp—has been farmed and eaten in China for over 1000 years. It is one of the most important aquaculture food resources worldwide. In the United States, however, Asian carp is considered a dangerous invasive species that disrupts community structure and composition to the point of threatening native species.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Population Demography

Populations are dynamic entities. **Populations** consist of all the species living within a specific area, and populations fluctuate based on a number of factors: seasonal and yearly changes in the environment, natural disasters such as forest fires and volcanic eruptions, and competition for resources between and within species. The statistical study of population dynamics, **demography**, uses a series of mathematical tools to investigate how populations respond to changes in their biotic and abiotic environments. Many of these tools were originally designed to study human populations. For example, life tables, which detail the life expectancy of individuals within a population, were initially developed by life insurance companies to set insurance rates. In fact, while the term "demographics" is commonly used when discussing humans, all living populations can be studied using this approach.

Population Size and Density

The study of any population usually begins by determining how many individuals of a particular species exist, and how closely associated they are with each other. Within a particular habitat, a population can be characterized by its **population size (*N*)**, the total number of individuals, and its **population density**, the number of individuals within a

specific area or volume. Population size and density are the two main characteristics used to describe and understand populations. For example, populations with more individuals may be more stable than smaller populations based on their genetic variability, and thus their potential to adapt to the environment. Alternatively, a member of a population with low population density (more spread out in the habitat), might have more difficulty finding a mate to reproduce compared to a population of higher density. As is shown in Figure 1, smaller organisms tend to be more densely distributed than larger organisms.

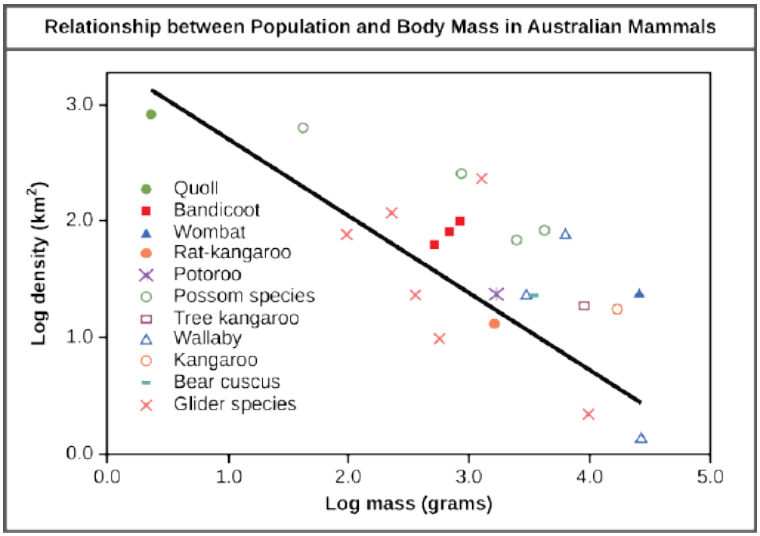


Figure 1: Australian mammals show a typical inverse relationship between population density and body size. (credit: "population density and body size" by OpenStax is licensed under CC BY 4.0)

As this graph shows, population density typically decreases with increasing body size. Why do you think this is the case?

Population Research Methods

The most accurate way to determine population size is to simply count all of the individuals within the habitat. However, this method is often not logistically or economically feasible, especially when studying large habitats. Thus, scientists usually study populations by sampling a representative portion of each habitat and using this data to make inferences about the habitat as a whole. A variety of methods can be used to sample populations to determine their size and density. For immobile organisms such as plants, or for very small and slow-moving organisms, a **quadrat** may be used (Figure 2). A quadrat is a way of marking off square areas within a habitat, either by staking out an area with sticks and string, or by the use of a wood, plastic, or metal square placed on the ground. After setting the quadrats, researchers then count the number of individuals that lie within their boundaries. Multiple quadrat samples are performed throughout the habitat at several random locations. All of this data can then be used to estimate the population size and population density within the entire habitat. The number and size of quadrat samples depend on the type of organisms under study and other factors, including the density of the organism. For example, if sampling daffodils, a 1 m^2 quadrat might be used whereas with giant redwoods, which are larger and live much further apart from each other, a larger quadrat of 100 m^2 might be employed. This ensures that enough individuals of the species are counted to get an accurate sample that correlates with the habitat, including areas not sampled.

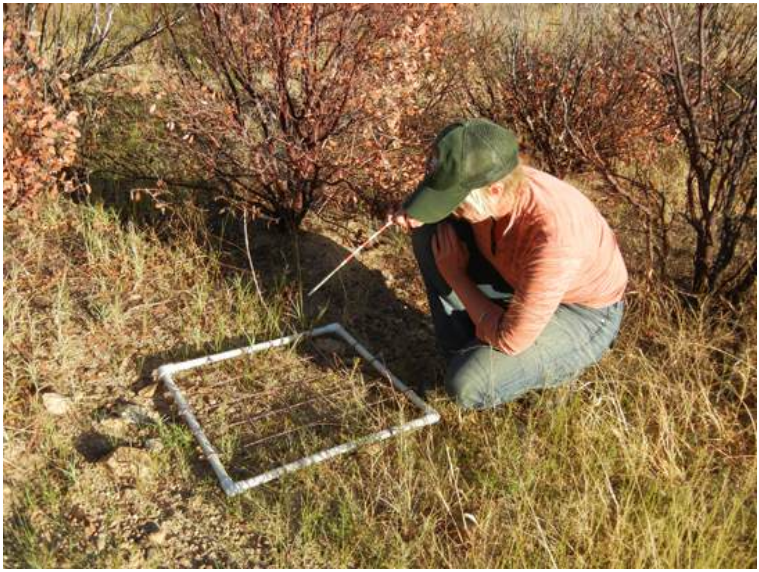


Figure 2: A scientist uses a quadrat to measure population size and density. (credit: NPS Sonoran Desert Network. "quadrat use" by OpenStax is licensed under CC BY 4.0)

For mobile organisms, such as mammals, birds, or fish, a technique called **mark and recapture** is often used. This method involves marking a sample of captured animals in some way (such as tags, bands, paint, or other body markings), and then releasing them back into the environment to allow them to mix with the rest of the population; later, a new sample is collected, including some individuals that are marked (recaptures) and some individuals that are unmarked (Figure 3).

Figure 3: Mark and recapture is used to measure the population size of mobile animals such as (a) bighorn sheep, (b) the California condor, and (c) salmon. (credit a: modification of work by Neal Herbert, NPS; credit b: modification of work by Pacific Southwest Region USFWS; credit c: modification of work by Ingrid Taylar. "this image" by OpenStax is licensed under CC BY 4.0)

Using the ratio of marked and unmarked individuals, scientists determine how many individuals are in the sample. From this, calculations are used to estimate the total population size. This method assumes that the larger the population, the lower the percentage of tagged organisms that will be recaptured since they will have mixed with more untagged individuals. For example, if 80 deer are captured, tagged, and released into the forest, and later 100 deer are captured and 20 of them are already marked, we can determine the population size (*N*) using the following equation:

(number marked first catch x total number of second catch)
number marked second catch = N(number marked first catch x total number of second catch)number marked second catch = N">(number marked first catch x total number of second catch) /
number marked second catch = N

Using our example, the population size would be estimated at 400.

(80 x 100)20 = 400(80 x 100)20 = 400">(80 x 100) / 20 = 400
Therefore, there are an estimated 400 total individuals in the original population.

There are some limitations to the mark and recapture method. Some animals from the first catch may learn to avoid capture in the second round, thus inflating population estimates. Alternatively, animals may preferentially be retrapped (especially if a food reward is offered), resulting in an underestimate of population size. Also, some species may be harmed by the marking technique, reducing their survival. A variety of other techniques have been developed, including the electronic tracking of animals tagged with radio transmitters and the use of data from commercial fishing and trapping operations to estimate the size and health of populations and communities.

Species Distribution

In addition to measuring simple density, further information about a population can be obtained by looking at the distribution of the individuals. **Species dispersion** patterns (or **distribution patterns**) show the spatial relationship between members of a population within a habitat at a particular point in time. In other words, they show whether members of the species live close together or far apart, and what patterns are evident when they are spaced apart.

Individuals in a population can be more or less equally spaced apart, dispersed randomly with no predictable pattern, or clustered in groups. These are known as **uniform**, **random**, and **clumped** dispersion patterns, respectively (Figure 4). Uniform dispersion is observed in plants that secrete substances inhibiting the growth of nearby individuals (such as the release of toxic chemicals by the sage plant *Salvia leucophylla,* a phenomenon called allelopathy)

and in animals like the penguin that maintain a defined territory. An example of random dispersion occurs with dandelion and other plants that have wind-dispersed seeds that germinate wherever they happen to fall in a favorable environment. A clumped dispersion may be seen in plants that drop their seeds straight to the ground, such as oak trees, or animals that live in groups (schools of fish or herds of elephants). Clumped dispersions may also be a function of habitat heterogeneity. Thus, the dispersion of the individuals within a population provides more information about how they interact with each other than does a simple density measurement. Just as lower density species might have more difficulty finding a mate, solitary species with a random distribution might have a similar difficulty when compared to social species clumped together in groups.

Figure 4: Species may have uniform, random, or clumped distribution. Territorial birds such as penguins tend to have uniform distribution. Plants such as dandelions with wind-dispersed seeds tend to be randomly distributed. Animals such as elephants that travel in groups exhibit clumped distribution. (credit a: modification of work by Ben Tubby; credit b: modification of work by Rosendahl; credit c: modification of work by Rebecca Wood. "this image" by OpenStax is licensed under CC BY 4.0)

Demography

While population size and density describe a population at one particular point in time, scientists must use demography to study the dynamics of a population. **Demography** is the statistical study of population changes over time: birth rates, death rates, and life expectancies. Each of these measures, especially birth rates, may be affected by the population characteristics described above. For example, a large population size results in a higher birth rate because more potentially reproductive individuals are present. In contrast, a large population size can also result in a higher death rate because of competition, disease, and the accumulation of waste. Similarly, a higher population density or a clumped dispersion pattern results in more potential reproductive encounters between individuals, which can increase birth rate. Lastly, a female-biased sex ratio (the ratio of males to females) or **age structure** (the proportion of population members at specific age ranges) composed of many individuals of reproductive age can increase birth rates.

In addition, the demographic characteristics of a population can influence how the population grows or declines over time. If birth and death rates are equal, the population remains stable. However, the population size will increase if birth rates exceed death rates; the population will decrease if birth rates are less than death rates. Life expectancy is another important factor; the length of time individuals remain in the population impacts local resources, reproduction, and the overall health of the population. These demographic characteristics are often displayed in the form of a life table.

Life Tables

Life tables provide important information about the life

history of an organism. Life tables divide the population into age groups and often sexes, and show how long a member of that group is likely to live. They are modeled after actuarial tables used by the insurance industry for estimating human life expectancy. Life tables may include the probability of individuals dying before their next birthday (i.e., their **mortality rate**), the percentage of surviving individuals dying at a particular age interval, and their **life expectancy** at each interval. An example of a life table is shown in Table 1 from a study of Dall mountain sheep, a species native to northwestern North America. Notice that the population is divided into age intervals (column A). The mortality rate (per 1000), shown in column D, is based on the number of individuals dying during the age interval (column B) divided by the number of individuals surviving at the beginning of the interval (Column C), multiplied by 1000.

mortality rate = number of individuals dyingnumber of individuals surviving x 1000mortality rate = number of individuals dyingnumber of individuals surviving x 1000">mortality rate = (number of individuals dying / number of individuals surviving) x 1000

For example, between ages three and four, 12 individuals die out of the 776 that were remaining from the original 1000 sheep. This number is then multiplied by 1000 to get the **mortality rate per thousand**.

mortality rate = 12776 x 1000 ≈ 15.5mortality rate = 12776 x 1000 ≈ 15.5">mortality rate = (12 / 776) x 1000 ≈ 15.5

As can be seen from the mortality rate data (column D), a high death rate occurred when the sheep were between 6 and 12 months old, and then increased even more from 8 to 12 years old, after which there were few survivors. The data indicate that if a sheep in this population were to survive to age one, it could be expected to live another 7.7 years on

average, as shown by the life expectancy numbers in column E.

This life table of *Ovis dalli* shows the number of deaths, number of survivors, mortality rate, and life expectancy at each age interval for the Dall mountain sheep.

Life Table of Dall Mountain Sheep (Data Adapted from Deevey, D. 1947)

Age interval (years)	Number dying in age interval out of 1000 born	Number surviving at beginning of age interval out of 1000 born	Mortality rate per 1000 alive at beginning of age interval	Life expectancy or mean lifetime remaining to those attaining age interval
0-0.5	54	1000	54.0	7.06
0.5-1	145	946	153.3	—
1-2	12	801	15.0	7.7
2-3	13	789	16.5	6.8
3-4	12	776	15.5	5.9
4-5	30	764	39.3	5.0
5-6	46	734	62.7	4.2
6-7	48	688	69.8	3.4
7-8	69	640	107.8	2.6
8-9	132	571	231.2	1.9
9-10	187	439	426.0	1.3
10-11	156	252	619.0	0.9
11-12	90	96	937.5	0.6
12-13	3	6	500.0	1.2
13-14	3	3	1000	0.7

Survivorship Curves

Another tool used by population ecologists is a **survivorship curve**, which is a graph of the number of individuals surviving at each age interval plotted versus time (usually with data compiled from a life table). These curves allow us to compare the life histories of different populations (Figure

5). Humans and most primates exhibit a **Type I survivorship curve** because a high percentage of offspring survive their early and middle years—death occurs predominantly in older individuals. These types of species usually have small numbers of offspring at one time, and they give a high amount of parental care to them to ensure their survival. Birds are an example of an intermediate or **Type II survivorship curve** because birds die more or less equally at each age interval. These organisms also may have relatively few offspring and provide significant parental care. Trees, marine invertebrates, and most fishes exhibit a **Type III survivorship curve** because very few of these organisms survive their younger years; however, those that make it to an old age are more likely to survive for a relatively long period of time. Organisms in this category usually have a very large number of offspring, but once they are born, little parental care is provided. Thus these offspring are "on their own" and vulnerable to predation, but their sheer numbers assure the survival of enough individuals to perpetuate the species.

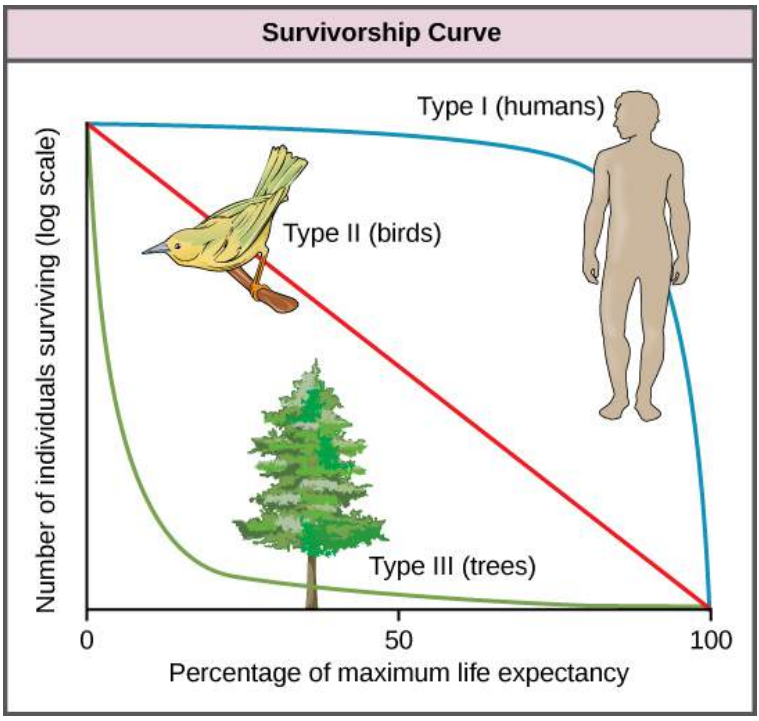


Figure 5: Survivorship curves show the distribution of individuals in a population according to age. Humans and most mammals have a Type I survivorship curve because death primarily occurs in the older years. Birds have a Type II survivorship curve, as death at any age is equally probable. Trees have a Type III survivorship curve because very few survive the younger years, but after a certain age, individuals are much more likely to survive. (credit: "survivorship curves" by OpenStax is licensed under CC BY 4.0)

Summary

Populations are individuals of a species that live in a particular habitat. Ecologists measure characteristics of populations: size, density, dispersion pattern, age structure,

and sex ratio. Life tables are useful to calculate life expectancies of individual population members. Survivorship curves show the number of individuals surviving at each age interval plotted versus time.

REFERENCES

Deevey, D. 1947. "Life Tables for Natural Populations of Animals," *The Quarterly Review of Biology* 22, no. 4: 283-314.

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Population Dynamics and Regulation

The logistic model of population growth, while valid in many natural populations and a useful model, is a simplification of real-world population dynamics. Implicit in the model is that the carrying capacity of the environment does not change, which is not the case. The carrying capacity varies annually: for example, some summers are hot and dry whereas others are cold and wet. In many areas, the carrying capacity during the winter is much lower than it is during the summer. Also, natural events such as earthquakes, volcanoes, and fires can alter an environment and hence its carrying capacity. Additionally, populations do not usually exist in isolation. They engage in interspecific competition: that is, they share the environment with other species, competing with them for the same resources. These factors are also important to understanding how a specific population will grow.

Nature regulates population growth in a variety of ways. These are grouped into density-dependent factors, in which the density of the population at a given time affects growth rate and mortality, and density-independent factors, which influence mortality in a population regardless of population density. Note that in the former, the effect of the factor on the population depends on the density of the population at onset. Conservation biologists want to understand both

types because this helps them manage populations and prevent extinction or overpopulation.

Density-dependent Regulation

Most **density-dependent factors** are biological in nature (biotic) and include predation, inter-and intraspecific competition, accumulation of waste, and diseases such as those caused by parasites. Usually, the denser a population is, the greater its mortality rate. For example, during intra- and interspecific competition, the reproductive rates of the individuals will usually be lower, reducing their population's rate of growth. In addition, low prey density increases the mortality of its predator because it has more difficulty locating its food source.

An example of density-dependent regulation is shown in Figure 1 with results from a study focusing on the giant intestinal roundworm (*Ascaris lumbricoides*), a parasite of humans and other mammals (Croll et al. 1982). Denser populations of the parasite exhibited lower fecundity: they contained fewer eggs. One possible explanation for this is that females would be smaller in more dense populations (due to limited resources) and that smaller females would have fewer eggs. This hypothesis was tested and disproved in a 2009 study which showed that female weight had no influence (Walker et al, 2009). The actual cause of the density-dependence of fecundity in this organism is still unclear and awaiting further investigation.

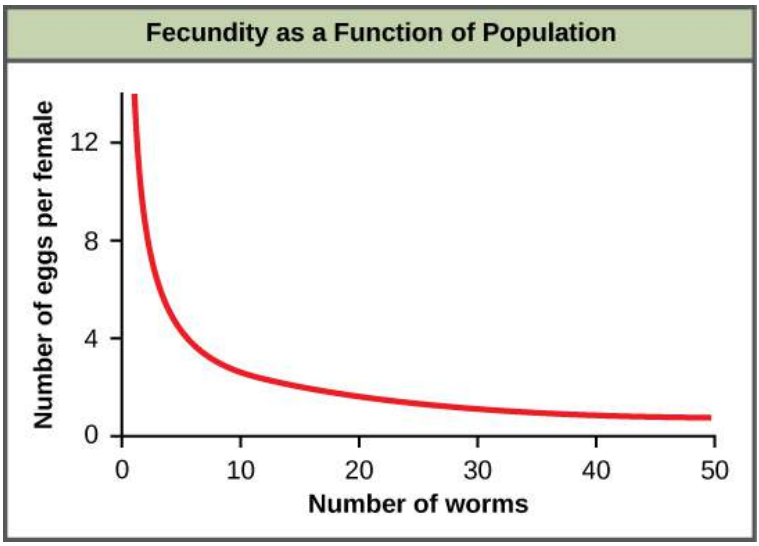


Figure 1: In this population of roundworms, fecundity (number of eggs) decreases with population density. (credit: Croll et al. 1982)

Density-independent Regulation and Interaction with Density-dependent Factors

Most **density-independent factors** are physical or chemical in nature (abiotic) These factors influence the mortality of a population regardless of its density, including weather, natural disasters, and pollution. An individual deer may be killed in a forest fire regardless of how many deer happen to be in that area. Its chances of survival are the same whether the population density is high or low. The same holds true for cold winter weather.

In real-life situations, population regulation is very complicated and density-dependent and independent factors can interact. A dense population that is reduced in a density-independent manner by some environmental

factor(s) will be able to recover differently than a sparse population. For example, a population of deer affected by a harsh winter will recover faster if there are more deer remaining to reproduce.

Evolution Connection – Why Did the Woolly Mammoth Go Extinct?

It's easy to get lost in the discussion of dinosaurs and theories about why they went extinct 65 million years ago. Was it due to a meteor slamming into Earth near the coast of modern-day Mexico, or was it from some long-term weather cycle that is not yet understood? One hypothesis that will never be proposed is that humans had something to do with it. Mammals were small, insignificant creatures of the forest 65 million years ago, and no humans existed.

Figure 2: 1916 mural of a mammoth herd from the American Museum of Natural History. Modification of work by Charles R. Knight.

Woolly mammoths, however, began to go extinct about 10,000 years ago, when they shared the Earth with humans who were no different anatomically than humans today (Figure 2). Mammoths survived in isolated island populations as recently as 1700 BC. We know a lot about these animals from carcasses found frozen in the ice of Siberia and other regions of the north. Scientists have sequenced at least 50 percent of its genome and believe mammoths are between 98 and 99 percent identical to modern elephants.

It is commonly thought that climate change and human hunting led to their extinction. A 2008 study estimated that climate change reduced the mammoth's range from 3,000,000 square miles 42,000 years ago to 310,000 square miles 6,000 years ago (Nogués-Bravo et al. 2008). It is also well documented that humans hunted these animals. A 2012 study showed that no single factor was exclusively responsible for the extinction of these magnificent creatures (MacDonald et al. 2012). In addition to human hunting, climate change, and reduction of habitat, these scientists demonstrated another important factor in the mammoth's extinction was the migration of humans across the Bering Strait to North America during the last ice age 20,000 years ago.

The maintenance of stable populations was and is very complex, with many interacting factors determining the outcome. It is important to remember that humans are also part of nature. Once we contributed to a species' decline using primitive hunting technology only.

Life Histories of *K*-selected and *r*-selected Species

While reproductive strategies play a key role in life histories, they do not account for important factors like limited resources and competition. The regulation of population growth by these factors can be used to introduce a classical concept in population biology, that of *K*-selected versus *r*-selected species.

Early Theories about Life History: K-selected and r-selected Species

By the second half of the twentieth century, the concept of K- and r-selected species was used extensively and successfully to study populations. The concept relates not only to

reproductive strategies but also to a species' habitat and behavior, especially in the way that they obtain resources and care for their young. It includes length of life and survivorship factors as well. For this analysis, population biologists have grouped species into the two large categories—*K*-selected and *r*-selected—although they are really two ends of a continuum.

***K*-selected species** are species selected by stable, predictable environments. Populations of *K*-selected species tend to exist close to their **carrying capacity** (hence the term *K*-selected) where intraspecific competition is high. These species have few, large offspring, a long gestation period, and often give long-term care to their offspring (Table 1). While larger in size when born, the offspring are relatively helpless and immature at birth. By the time they reach adulthood, they must develop skills to compete for natural resources. In plants, scientists think of parental care more broadly: how long fruit takes to develop or how long it remains on the plant are determining factors in the time to the next reproductive event. Examples of *K*-selected species are primates (including humans), elephants, and plants such as oak trees (Figure 3a).

Oak trees grow very slowly and take, on average, 20 years to produce their first seeds, known as acorns. As many as 50,000 acorns can be produced by an individual tree, but the germination rate is low as many of these rot or are eaten by animals such as squirrels. In some years, oaks may produce an exceptionally large number of acorns, and these years may be on a two- or three-year cycle depending on the species of oak (*r*-selection).

As oak trees grow to a large size and for many years before they begin to produce acorns, they devote a large percentage of their energy budget to growth and maintenance. The tree's height and size allow it to dominate other plants in the competition for sunlight, the oak's primary energy resource.

Furthermore, when it does reproduce, the oak produces large, energy-rich seeds that use their energy reserve to become quickly established (*K*-selection).

In contrast, ***r*-selected species** have a large number of small offspring (hence their *r* designation (Table 1). This strategy is often employed in unpredictable or changing environments. Animals that are *r*-selected do not give long-term parental care and the offspring are relatively mature and self-sufficient at birth. Examples of *r*-selected species are marine invertebrates, such as jellyfish, and plants, such as the dandelion (Figure 3b). Dandelions have small seeds that are wind-dispersed long distances. Many seeds are produced simultaneously to ensure that at least some of them reach a hospitable environment. Seeds that land in inhospitable environments have little chance for survival since their seeds are low in energy content. Note that survival is not necessarily a function of energy stored in the seed itself.

Characteristics of *K*-selected and *r*-selected species

Characteristics of *K*-selected species	**Characteristics of *r*-selected species**
Mature late	Mature early
Greater longevity	Lower longevity
Increased parental care	Decreased parental care
Increased competition	Decreased competition
Fewer offspring	More offspring
Larger offspring	Smaller offspring

Table 1: Characteristics of *K*-selected and *r*-selected species

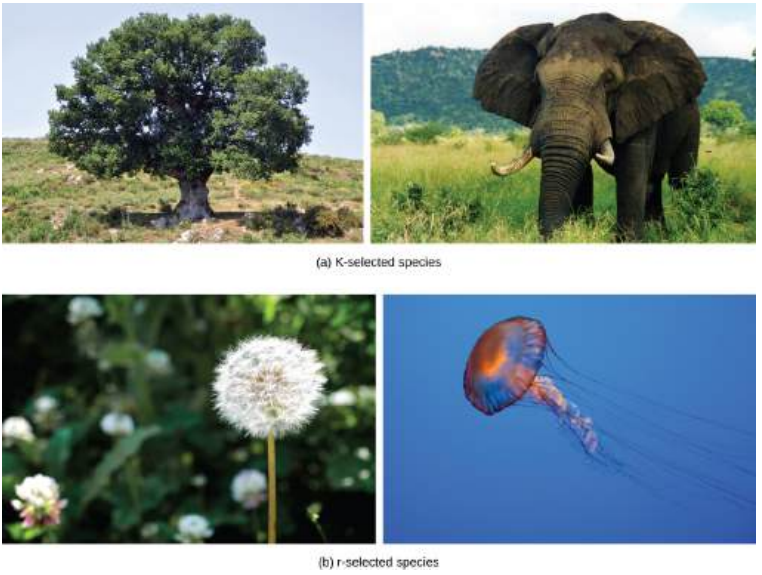


Figure 3: (a) Elephants are considered K-selected species as they live long, mature late, and provide long-term parental care to few offspring. Oak trees produce many offspring that do not receive parental care, but are considered K-selected species based on longevity and late maturation. (b) Dandelions and jellyfish are both considered r-selected species as they mature early, have short lifespans, and produce many offspring that receive no parental care. (credit: "this image" by OpenStax is licensed under CC BY 4.0)

Modern Theories of Life History

The *r*– and *K*-selection theory, although accepted for decades and used for much groundbreaking research, has now been reconsidered, and many population biologists have abandoned or modified it. Over the years, several studies attempted to confirm the theory, but these attempts have largely failed. Many species were identified that did not follow the theory's predictions. Furthermore, the theory

ignored the age-specific mortality of the populations which scientists now know is very important. New demographic-based models of life history evolution have been developed which incorporate many ecological concepts included in *r*– and *K*-selection theory as well as population age structure and mortality factors.

Body Size correlates with Generation Time

Generation time is the average span of time between the birth of an individual and the birth of its offspring. The general trend is that the larger the species, the longer the generation time (figure 4).

Relationship between generation time and body length

Figure 4: Relationship between generation time and body length in 47 taxonomic groups, ranging from bacteria to long-lived trees. Data transformed by log10. Data derived from Bonner (1965; c.f., McNaughton and Wolf, 1979; also Sammarco, pers. obs.; Strychar, pers. obs.). Image from Sammarco, P., & Strychar, K. 2009.

Summary

Populations are regulated by a variety of density-dependent and density-independent factors. Species are divided into two categories based on a variety of features of their life history patterns: *r*-selected species, which have large numbers of offspring, and *K*-selected species, which have few offspring. The *r*– and *K*-selection theory has fallen out of use; however, many of its key features are still used in newer, demographically-based models of population dynamics.

REFERENCES

Croll, N.A. et al. 1982. "The Population Biology and Control of *Ascaris lumbricoides* in a Rural Community in Iran." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 76, no. 2: 187-197, doi:10.1016/0035-9203(82)90272-3.

MacDonald, G.M., et al. 2012. "Pattern of Extinction of the Woolly Mammoth in Beringia." *Nature Communications* 3, no. 893, doi:10.1038/ncomms1881.

Nogués-Bravo, D. et al. 2008. "Climate Change, Humans, and the Extinction of the Woolly Mammoth." *PLoS Biol* 6: e79, doi:10.1371/journal.pbio.0060079.

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Walker, M. et al. 2009. "Density-Dependent Effects on the Weight of Female *Ascaris lumbricoides* Infections of Humans and its Impact on Patterns of Egg Production." *Parasites & Vectors* 2, no. 11, doi:10.1186/1756-3305-2-11.

Sammarco, P., & Strychar, K. 2009. Effects of Climate Change/Global Warming on Coral Reefs: Adaptation/Exaptation in Corals, Evolution in Zooxanthellae, and Biogeographic Shifts. Environmental Bioindicators, 4, 9–45. https://doi.org/10.1080/15555270902905377

Life Histories and Natural Selection

A species' life history describes the series of events over its lifetime, such as how resources are allocated for growth, maintenance, and reproduction. Life history traits affect the life table of an organism. A species' life history is genetically determined and shaped by the environment and natural selection.

Life History Patterns and Energy Budgets

Energy is required by all living organisms for their growth, maintenance, and reproduction; at the same time, energy is often a major limiting factor in determining an organism's survival. Plants, for example, acquire energy from the sun via photosynthesis but must expend this energy to grow, maintain health, and produce energy-rich seeds to produce the next generation. Animals have the additional burden of using some of their energy reserves to acquire food. Furthermore, some animals must expend energy caring for their offspring. Thus, all species have an **energy budget**: they must balance energy intake with their use of energy for metabolism, reproduction, parental care, and energy storage (such as bears building up body fat for winter hibernation).

Parental Care and Fecundity

Fecundity is the potential reproductive capacity of an individual within a population. Fecundity describes how many offspring could ideally be produced if an individual has as many offspring as possible, repeating the reproductive cycle as soon as possible after the birth of the offspring. In animals, fecundity is inversely related to the amount of parental care given to an individual offspring. Species, such as many marine invertebrates, that produce many offspring usually provide little if any care for the offspring (they would not have the energy or the ability to do so anyway). Most of their energy budget is used to produce many tiny offspring. Animals with this strategy are often self-sufficient at a very early age. This is because of the energy tradeoff these organisms have made to maximize their evolutionary fitness. Because their energy is used for producing offspring instead of parental care, it makes sense that these offspring have some ability to be able to move within their environment and find food and perhaps shelter. Even with these abilities, their small size makes them extremely vulnerable to predation, so the production of many offspring allows enough of them to survive to maintain the species.

Animal species that have few offspring during a reproductive event usually give extensive parental care, devoting much of their energy budget to these activities, sometimes at the expense of their own health. This is the case with many mammals, such as humans, kangaroos, and pandas. The offspring of these species are relatively helpless at birth and need to develop before they achieve self-sufficiency.

Plants with low fecundity produce few energy-rich seeds (such as coconuts and chestnuts) with each having a good chance to germinate into a new organism; plants with high fecundity usually have many small, energy-poor seeds (like

orchids) that have a relatively poor chance of surviving. Although it may seem that coconuts and chestnuts have a better chance of surviving, the energy tradeoff of the orchid is also very effective. It is a matter of where the energy is used, for large numbers of seeds or for fewer seeds with more energy.

Early versus Late Reproduction

The **timing of reproduction** in a life history also affects species survival. Organisms that reproduce at an early age have a greater chance of producing offspring, but this is usually at the expense of their growth and the maintenance of their health. Conversely, organisms that start reproducing later in life often have greater fecundity or are better able to provide parental care, but they risk that they will not survive to reproductive age. Examples of this can be seen in fishes. Small fish like guppies use their energy to reproduce rapidly, but never attain the size that would give them defense against some predators. Larger fish, like the bluegill or shark, use their energy to attain a large size but do so with the risk that they will die before they can reproduce or at least reproduce to their maximum. These different energy strategies and tradeoffs are key to understanding the evolution of each species as it maximizes its fitness and fills its niche. In terms of energy budgeting, some species "blow it all" and use up most of their energy reserves to reproduce early before they die. Other species delay having reproduction to become stronger, more experienced individuals and to make sure that they are strong enough to provide parental care if necessary.

Single versus Multiple Reproductive Events

Some life history traits, such as fecundity, the timing of

reproduction, and parental care, can be grouped together into general strategies that are used by multiple species. **Semelparity** occurs when a species reproduces only once during its lifetime and then dies. Such species use most of their resource budget during a single reproductive event, sacrificing their health to the point that they do not survive. Examples of semelparity are bamboo, which flowers once and then dies, and the Chinook salmon (Figure 1), which uses most of its energy reserves to migrate from the ocean to its freshwater nesting area, where it reproduces and then dies. Scientists have posited alternate explanations for the evolutionary advantage of the Chinook's post-reproduction death: a programmed suicide caused by a massive release of corticosteroid hormones, presumably so the parents can become food for the offspring or simple exhaustion caused by the energy demands of reproduction; these are still being debated.

Iteroparity describes species that reproduce repeatedly during their lives. Some animals are able to mate only once per year but survive multiple mating seasons. The pronghorn antelope is an example of an animal that goes into a seasonal **estrus cycle** ("heat"): a hormonally induced physiological condition preparing the body for successful mating (Figure 1). Females of these species mate only during the estrus phase of the cycle. A different pattern is observed in primates, including humans and chimpanzees, which may attempt reproduction at any time during their reproductive years, even though their menstrual cycles make pregnancy likely only a few days per month during ovulation (Figure 1).

Figure 1a: The Chinook salmon mates once and dies. (credit a: modification of work by Roger Tabor, USFWS.)

Figure 1b: The pronghorn antelope mates during specific times of the year during its reproductive life. (credit b: modification of work by Mark Gocke, USDA.)

Figure 1c: Primates, such as humans and chimpanzees, may mate on any day, independent of ovulation. (credit c: modification of work by "Shiny Things"/Flickr)

Evolution Connection – Energy Budgets, Reproductive Costs, and Sexual Selection in Drosophila

Research into how animals allocate their energy resources for growth, maintenance, and reproduction has used a variety of experimental animal models. Some of this work has been done using the common fruit fly, *Drosophila melanogaster*. Studies have shown that not only does reproduction have a cost as far as how long male fruit flies live, but also fruit flies that have already mated several times have limited sperm remaining for reproduction. Fruit flies maximize their last chances at reproduction by selecting optimal mates.

In a 1981 study, male fruit flies were placed in enclosures with either virgin or inseminated females. The males that

mated with virgin females had shorter life spans than those in contact with the same number of inseminated females with which they were unable to mate. This effect occurred regardless of how large (indicative of their age) the males were. Thus, males that did not mate lived longer, allowing them more opportunities to find mates in the future.

More recent studies, performed in 2006, show how males select the female with which they will mate and how this is affected by previous matings (Figure 2) (Byrne & Rice, 2006). Males were allowed to select between smaller and larger females. Findings showed that larger females had greater fecundity, producing twice as many offspring per mating as the smaller females did. Males that had previously mated, and thus had lower supplies of sperm, were termed "resource-depleted," while males that had not mated were termed "non-resource-depleted." The study showed that although non-resource-depleted males preferentially mated with larger females, this selection of partners was more pronounced in the resource-depleted males. Thus, males with depleted sperm supplies, which were limited in the number of times that they could mate before they replenished their sperm supply, selected larger, more fecund females, thus maximizing their chances for offspring. This study was one of the first to show that the physiological state of the male affected its mating behavior in a way that clearly maximizes its use of limited reproductive resources.

	Ratio large/small females mated
Non sperm-depleted	8 ± 5
Sperm-depleted	15 ± 5

Figure 2: Male fruit flies that had previously mated (sperm-depleted) picked larger, more fecund females more often than those that had not mated (non-sperm-depleted). This change in behavior causes an increase in the efficiency of a limited reproductive resource: sperm. (credit: Byrne, P. and Rice, R. 2006. "this image" by OpenStax is licensed under CC BY 4.0)

These studies demonstrate two ways in which the energy budget is a factor in reproduction. First, energy expended on mating may reduce an animal's lifespan, but by this time they have already reproduced, so in the context of natural selection this early death is not of much evolutionary importance. Second, when resources such as sperm (and the energy needed to replenish it) are low, an organism's behavior can change to give them the best chance of passing their genes on to the next generation. These changes in behavior, so important to evolution, are studied in a discipline known as behavioral biology, or ethology, at the interface between population biology and psychology.

Summary

All species have evolved a pattern of living, called a life history strategy, in which they partition energy for growth, maintenance, and reproduction. These patterns evolve through natural selection; they allow species to adapt to their environment to obtain the resources they need to successfully reproduce. There is an inverse relationship between fecundity and parental care. A species may

reproduce early in life to ensure surviving to a reproductive age or reproduce later in life to become larger and healthier and better able to give parental care. A species may reproduce once (semelparity) or many times (iteroparity) in its life.

REFERENCES

Byrne, P. and Rice, R. 2006. "Evidence for adaptive male mate choice in the fruit fly *Drosophila melanogaster,*" Proc Biol Sci. 273, no. 1589: 917-922, doi: 10.1098/rspb.2005.3372.

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Environmental Limits to Population Growth

Although life histories describe the way many characteristics of a population (such as their age structure) change over time in a general way, population ecologists make use of a variety of methods to model population dynamics mathematically. These more precise models can then be used to accurately describe changes occurring in a population and better predict future changes. Certain models that have been accepted for decades are now being modified or even abandoned due to their lack of predictive ability, and scholars strive to create effective new models.

Exponential Growth

Charles Darwin, in his theory of natural selection, was greatly influenced by the English clergyman Thomas Malthus. Malthus published a book in 1798 stating that populations with unlimited natural resources grow very rapidly, and then population growth decreases as resources become depleted. This accelerating pattern of increasing population size is called **exponential growth**.

The best example of exponential growth is seen in bacteria. Bacteria are prokaryotes that reproduce by prokaryotic fission. This division takes about an hour for

many bacterial species. If 1000 bacteria are placed in a large flask with an unlimited supply of nutrients (so the nutrients will not become depleted), after an hour, there is one round of division and each organism divides, resulting in 2000 organisms—an increase of 1000. In another hour, each of the 2000 organisms will double, producing 4000, an increase of 2000 organisms. After the third hour, there should be 8000 bacteria in the flask, an increase of 4000 organisms. The important concept of exponential growth is that the population growth rate—the number of organisms added in each reproductive generation—is accelerating; that is, it is increasing at a greater and greater rate. After 1 day and 24 of these cycles, the population would have increased from 1000 to more than 16 billion. When the **population size**, *N*, is plotted over time, a **J-shaped growth curve** is produced (Figure 1).

The bacteria example is not representative of the real world where resources are limited. Furthermore, some bacteria will die during the experiment and thus not reproduce, lowering the growth rate. Therefore, when calculating the growth rate of a population, the **death rate (*D*)** (number of organisms that die during a particular time interval) is subtracted from the **birth rate (*B*)** (number of organisms that are born during that interval). This is shown in the following formula:

ΔN (change in number)ΔT (change in time) = B (birth rate) - D (death rate)ΔN (change in number)ΔT (change in time) = B (birth rate) - D (death rate)">**ΔN** (change in number) / **ΔT** (change in time) = B (birth rate) – D (death rate)

The birth rate is usually expressed on a per capita (for each individual) basis. Thus, *B* (birth rate) = ***bN*** (the per capita birth rate "*b*" multiplied by the number of individuals "*N*") and *D* (death rate) =***dN*** (the per capita death rate "d" multiplied by

the number of individuals "*N*"). Additionally, ecologists are interested in the population at a particular point in time, an infinitely small time interval. For this reason, the terminology of differential calculus is used to obtain the "**instantaneous**" growth rate, replacing the *change* in number and time with an instant-specific measurement of number and time.

dNdT = bN − dN = (b - d)NdNdT = bN − dN = (b - d)N">dN / dT = bN − dN = (b − d)N

Notice that the "*d*" associated with the first term refers to the derivative (as the term is used in calculus) and is different from the death rate, also called "*d*." The difference between birth and death rates is further simplified by substituting the term "*r*" (**intrinsic rate of increase**) for the relationship between birth and death rates:

dNdT = rNdNdT = rN">dN / dT = rN

The value "*r*" can be positive, meaning the population is increasing in size; or negative, meaning the population is decreasing in size; or zero, where the population's size is unchanging, a condition known as zero population growth. A further refinement of the formula recognizes that different species have inherent differences in their intrinsic rate of increase (often thought of as the potential for reproduction), even under ideal conditions. Obviously, a bacterium can reproduce more rapidly and have a higher intrinsic rate of growth than a human. The **maximal growth rate** for a species is its biotic potential, or r_{max}, thus changing the equation to:

dNdT=rmaxNdNdT=rmaxN">dN / dT=rmaxN

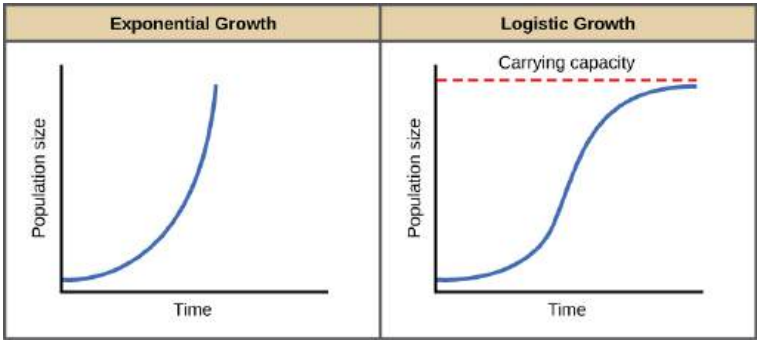


Figure 1: When resources are unlimited, populations exhibit exponential growth, resulting in a J-shaped curve. When resources are limited, populations exhibit logistic growth. In logistic growth, population expansion decreases as resources become scarce, and it levels off when the carrying capacity of the environment is reached, resulting in an S-shaped curve. (credit: "population growth" by OpenStax is licensed under CC BY 4.0)

Logistic Growth

Exponential growth is possible only when infinite natural resources are available; this is not the case in the real world. Charles Darwin recognized this fact in his description of the "struggle for existence," which states that individuals will compete (with members of their own or other species) for limited resources. The successful ones will survive to pass on their own characteristics and traits (which we know now are transferred by genes) to the next generation at a greater rate (natural selection). To model the reality of limited resources, population ecologists developed the logistic growth model.

Carrying Capacity and the Logistic Model

In the real world, with its limited resources, exponential

growth cannot continue indefinitely. Exponential growth may occur in environments where there are few individuals and plentiful resources, but when the number of individuals gets large enough, resources will be depleted, slowing the growth rate. Eventually, the growth rate will plateau or level off (Figure 1). This population size, which represents the maximum population size that a particular environment can support, is called the **carrying capacity**, or *K*.

The formula we use to **calculate logistic growth** adds the carrying capacity as a moderating force in the growth rate. The expression "*K* – *N*" is indicative of how many individuals may be added to a population at a given stage, and "*K* – *N*" divided by "*K*" is the fraction of the carrying capacity available for further growth. Thus, the exponential growth model is restricted by this factor to generate the logistic growth equation:

dNdT=rmaxdNdT=rmaxN(K - N)KdNdT=rmaxdNdT=rmaxN(K - N)K">dN / dT = rmax (dN / dT) =
rmaxN[(K − N) / K]

Notice that when *N* is very small, (*K*-*N*)/*K* becomes close to *K*/*K* or 1, and the right side of the equation reduces to $r_{max}N$, which means the population is growing exponentially and is not influenced by carrying capacity. On the other hand, when *N* is large, (*K*-*N*)/*K* come close to zero, which means that population growth will be slowed greatly or even stopped. Thus, population growth is greatly slowed in large populations by the carrying capacity *K*. This model also allows for the population of a negative population growth, or a population decline. This occurs when the number of individuals in the population exceeds the carrying capacity (because the value of (K-N)/K is negative).

A graph of this equation yields an S-shaped curve (Figure 1), and it is a more realistic model of population growth than exponential growth. There are three different sections to an S-shaped curve. Initially, growth is exponential because there

are few individuals and ample resources available. Then, as resources begin to become limited, the growth rate decreases. Finally, growth levels off at the carrying capacity of the environment, with little change in population size over time.

Role of Intraspecific Competition

The logistic model assumes that every individual within a population will have equal access to resources and, thus, an equal chance for survival. For plants, the amount of water, sunlight, nutrients, and the space to grow are the important resources, whereas in animals, important resources include food, water, shelter, nesting space, and mates.

In the real world, phenotypic variation among individuals within a population means that some individuals will be better adapted to their environment than others. The resulting competition between population members of the same species for resources is termed **intraspecific competition** (intra- = "within"; -specific = "species"). Intraspecific competition for resources may not affect populations that are well below their carrying capacity—resources are plentiful and all individuals can obtain what they need. However, as population size increases, this competition intensifies. In addition, the accumulation of waste products can reduce an environment's carrying capacity.

Examples of Logistic Growth

Yeast, a microscopic fungus used to make bread and alcoholic beverages, exhibits the classical S-shaped curve when grown in a test tube (Figure 2a). Its growth levels off as the population depletes the nutrients that are necessary for its growth. In the real world, however, there are variations

to this idealized curve. Examples in wild populations include sheep and harbor seals (Figure 2b). In both examples, the population size exceeds the carrying capacity for short periods of time and then falls below the carrying capacity afterward. This fluctuation in population size continues to occur as the population oscillates around its carrying capacity. Still, even with this oscillation, the logistic model is confirmed.

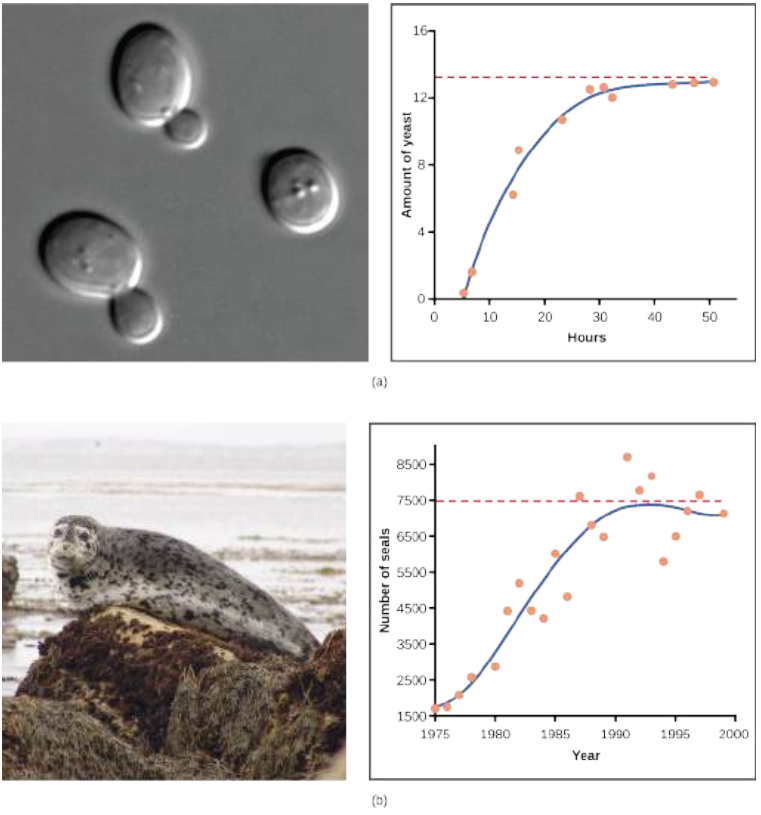


Figure 2: (a) Yeast grown in ideal conditions in a test tube show a classical S-shaped logistic growth curve, whereas (b) a natural population of seals shows real-world fluctuation. (credit: "examples of population growth" by OpenStax is licensed under CC BY 4.0)

Summary

Populations with unlimited resources grow exponentially, with an accelerating growth rate. When resources become limiting, populations follow a logistic growth curve. The population of a species will level off at the carrying capacity of its environment.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Human Population Growth

Concepts of animal population dynamics can be applied to human population growth. Humans are not unique in their ability to alter their environment. For example, beaver dams alter the stream environment where they are built. Humans, however, have the ability to alter their environment to increase its carrying capacity sometimes to the detriment of other species (e.g., via artificial selection for crops that have a higher yield). Earth's human population is growing rapidly, to the extent that some worry about the ability of the earth's environment to sustain this population, as long-term exponential growth carries the potential risks of famine, disease, and large-scale death.

Although humans have increased the carrying capacity of their environment, the technologies used to achieve this transformation have caused unprecedented changes to Earth's environment, altering ecosystems to the point where some may be in danger of collapse. The depletion of the ozone layer, erosion due to acid rain, and damage from global climate change are caused by human activities. The ultimate effect of these changes on our carrying capacity is unknown. As some point out, it is likely that the negative effects of increasing carrying capacity will outweigh the positive ones—the carrying capacity of the world for human beings might actually decrease.

The world's human population is currently experiencing

exponential growth even though human reproduction is far below its **biotic potential** (Figure 1). To reach its biotic potential, all females would have to become pregnant every nine months or so during their reproductive years. Also, resources would have to be such that the environment would support such growth. Neither of these two conditions exists. In spite of this fact, the human population is still growing exponentially.

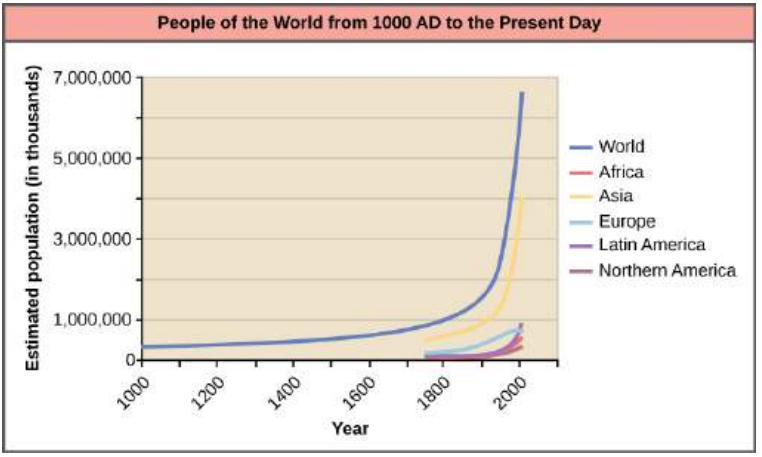


Figure 1: Human population growth since 1000 AD is exponential (dark blue line). Notice that while the population in Asia (yellow line), which has many economically underdeveloped countries, is increasing exponentially, the population in Europe (light blue line), where most of the countries are economically developed, is growing much more slowly. (credit: "Human population growth" by OpenStax is licensed under CC BY 4.0)

A consequence of exponential human population growth is the time that it takes to add a particular number of humans to the Earth is becoming shorter. Figure 2 shows that 123 years were necessary to add 1 billion humans in 1930, but it only took 24 years to add two billion people between 1975 and 1999. As already discussed, at some point it would

appear that our ability to increase our carrying capacity indefinitely on a finite world is uncertain. Without new technological advances, the human growth rate has been predicted to slow in the coming decades. However, the population will still be increasing and the threat of overpopulation remains.

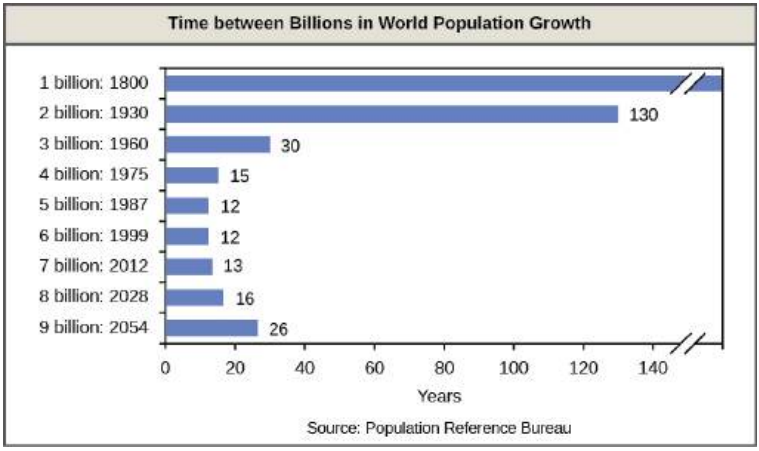


Figure 2: The time between the addition of each billion human beings to Earth decreases over time. (credit: modification of work by Ryan T. Cragun. "time between each billion" by OpenStax is licensed under CC BY 4.0)

Overcoming Density-Dependent Regulation

Humans are unique in their ability to alter their environment with the conscious purpose of increasing its carrying capacity. This ability is a major factor responsible for human population growth and a way of overcoming density-dependent growth regulation. Much of this ability is related to human intelligence, society, and communication. Humans can construct shelter to protect themselves from the elements and have developed agriculture and domesticated animals to increase their food supplies. In addition, humans

use language to communicate this technology to new generations, allowing them to improve upon previous accomplishments.

Other factors in human population growth are migration and public health. Humans originated in Africa, but have since migrated to nearly all inhabitable land on the Earth. Public health, sanitation, and the use of antibiotics and vaccines have decreased the ability of infectious disease to limit human population growth. In the past, diseases such as the bubonic plague of the fourteenth century killed between 30 and 60 percent of Europe's population and reduced the overall world population by as many as 100 million people. Today, the threat of infectious disease, while not gone, is certainly less severe. According to the World Health Organization, global death from infectious disease declined from 16.4 million in 1993 to 14.7 million in 1992. To compare to some of the epidemics of the past, the percentage of the world's population killed between 1993 and 2002 decreased from 0.30 percent of the world's population to 0.24 percent. Thus, it appears that the influence of infectious disease on human population growth is becoming less significant.

Age Structure, Population Growth, and Economic Development

The **age structure** of a population is an important factor in population dynamics. Age structure is the proportion of a population at different age ranges. Age structure allows better prediction of population growth, plus the ability to associate this growth with the level of economic development in the region. Countries with rapid growth have a pyramidal shape in their age structure diagrams, showing a preponderance of younger individuals, many of whom are of reproductive age or will be soon (Figure 3). This pattern is most often observed in underdeveloped countries where

individuals do not live to old age because of less-than-optimal living conditions. Age structures of areas with slow growth, including developed countries such as the United States, still have a pyramidal structure, but with many fewer young and reproductive-aged individuals and a greater proportion of older individuals. Other developed countries, such as Italy, have zero population growth. The age structure of these populations is more conical, with an even greater percentage of middle-aged and older individuals. The actual growth rates in different countries are shown in Figure 4, with the highest rates tending to be in the less economically developed countries of Africa and Asia.

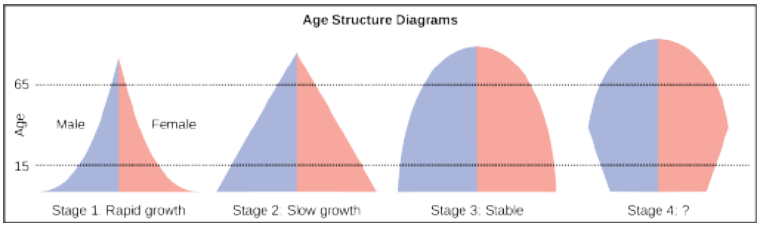


Figure 3: Typical age structure diagrams are shown. The rapid growth diagram narrows to a point, indicating that the number of individuals decreases rapidly with age. In the slow growth model, the number of individuals decreases steadily with age. Stable population diagrams are rounded on the top, showing that the number of individuals per age group decreases gradually, and then increases for the older part of the population. (credit: "age structure diagram" by OpenStax is licensed under CC BY 4.0)

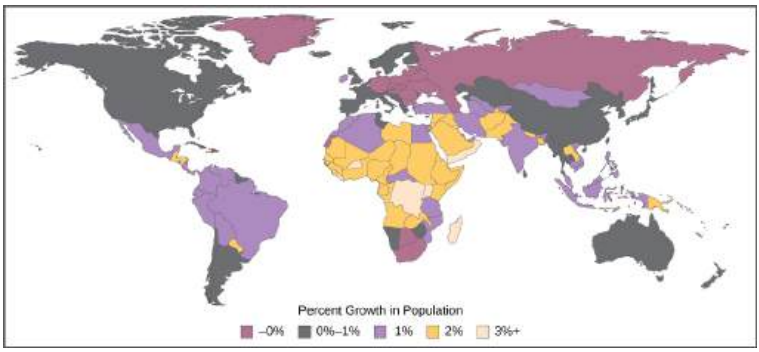


Figure 4: The percent growth rate of population in different countries is shown. Notice that the highest growth is occurring in less economically developed countries in Africa and Asia. (credit:"growth rate map" by OpenStax is licensed under CC BY 4.0)

Long-Term Consequences of Exponential Human Population Growth

Many dire predictions have been made about the world's population leading to a major crisis called the "**population explosion**." In the 1968 book *The Population Bomb*, biologist Dr. Paul R. Ehrlich wrote, "The battle to feed all of humanity is over. In the 1970s hundreds of millions of people will starve to death in spite of any crash programs embarked upon now. At this late date, nothing can prevent a substantial increase in the world death rate" (Erlich, 1968). While many critics view this statement as an exaggeration, the laws of exponential population growth are still in effect, and unchecked human population growth cannot continue indefinitely.

In spite of population control policies such as the one, the human population continues to grow. At some point, the food supply may run out because of the subsequent need to produce more and more food to feed our population.

The United Nations estimates that future world population growth may vary from 6 billion (a decrease) to 16 billion people by the year 2100. There is no way to know whether human population growth will moderate to the point where the crisis described by Dr. Ehrlich will be averted.

Another result of population growth is the endangerment of the natural environment. Many countries have attempted to reduce the human impact on climate change by reducing their emission of the greenhouse gas carbon dioxide. However, these treaties have not been ratified by every country, and many underdeveloped countries trying to improve their economic condition may be less likely to agree with such provisions if it means slower economic development. Furthermore, the role of human activity in causing climate change has become a hotly debated socio-political issue in some developed countries, including the United States. Thus, we enter the future with considerable uncertainty about our ability to curb human population growth and protect our environment.

Summary

The world's human population is growing at an exponential rate. Humans have increased the world's carrying capacity through migration, agriculture, medical advances, and communication. The age structure of a population allows us to predict population growth. Unchecked human population growth could have dire long-term effects on our environment.

REFERENCES

Erlich, P. 1968. prologue to *The Population Bomb*, (1968; repr., New York: Ballantine, 1970).

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Community Ecology

Populations rarely, if ever, live in isolation from populations of other species. In most cases, numerous species share a habitat. The interactions between these populations play a major role in regulating population growth and abundance. All populations occupying the same habitat form a community: populations inhabiting a specific area at the same time. The number of species occupying the same habitat and their relative abundance is known as species diversity. Areas with low diversity, such as the glaciers of Antarctica, still contain a wide variety of living things, whereas the diversity of tropical rainforests is so great that it cannot be counted. Ecology is studied at the community level to understand how species interact with each other and compete for the same resources.

Predation and Herbivory

Perhaps the classical example of species interaction is **predation**: the hunting of prey by its predator. Nature shows on television highlight the drama of one living organism killing another. Populations of predators and prey in a community are not constant over time: in most cases, they vary in cycles that appear to be related. The most often cited example of predator-prey dynamics is seen in the cycling of the lynx (predator) and the snowshoe hare (prey), using

nearly 200 year-old trapping data from North American forests (Figure 1). This cycle of predator and prey last approximately 10 years, with the predator population lagging 1–2 years behind that of the prey population. As the hare numbers increase, there is more food available for the lynx, allowing the lynx population to increase as well. When the lynx population grows to a threshold level, however, they kill so many hares that hare population begins to decline, followed by a decline in the lynx population because of scarcity of food. When the lynx population is low, the hare population size begins to increase due, at least in part, to low predation pressure, starting the cycle anew.

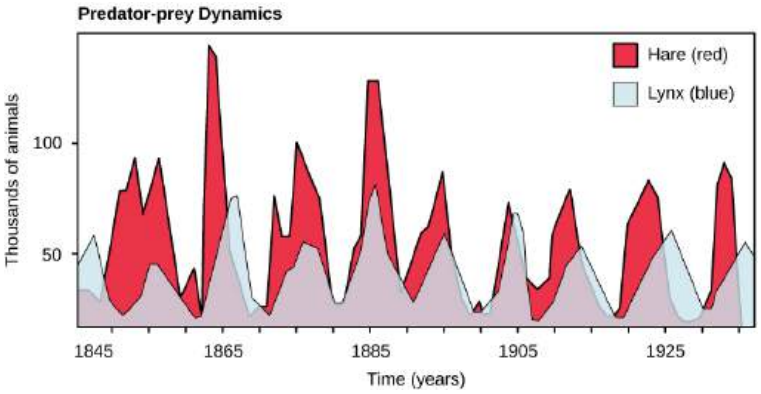


Figure 1: The cycling of lynx and snowshoe hare populations in Northern Ontario is an example of predator-prey dynamics. (credit: "lynx and snowshoe hare populations" by OpenStax is licensed under CC BY 4.0)

The idea that the population cycling of the two species is entirely controlled by predation models has come under question. More recent studies have pointed to undefined density-dependent factors as being important in the cycling, in addition to predation. One possibility is that the cycling is inherent in the hare population due to density-dependent

effects such as lower fecundity (maternal stress) caused by crowding when the hare population gets too dense. The hare cycling would then induce the cycling of the lynx because it is the lynxes' major food source. The more we study communities, the more complexities we find, allowing ecologists to derive more accurate and sophisticated models of population dynamics.

Herbivory describes the consumption of plants by insects and other animals, and it is another interspecific relationship that affects populations. Unlike animals, most plants cannot outrun predators or use mimicry to hide from hungry animals. Some plants have developed mechanisms to defend against herbivory. Other species have developed mutualistic relationships; for example, herbivory provides a mechanism of seed distribution that aids in plant reproduction.

Defense Mechanisms against Predation and Herbivory

The study of communities must consider evolutionary forces that act on the members of the various populations contained within it. Species are not static, but slowly changing and adapting to their environment by natural selection and other evolutionary forces. Species have evolved numerous mechanisms to escape predation and herbivory. These defenses may be mechanical, chemical, physical, or behavioral.

Mechanical defenses, such as the presence of thorns on plants or the hard shell on turtles, discourage animal predation and herbivory by causing physical pain to the predator or by physically preventing the predator from being able to eat the prey. **Chemical defenses** are produced by many animals as well as plants, such as the foxglove which is extremely toxic when eaten. Figure 2 shows some organisms' defenses against predation and herbivory.

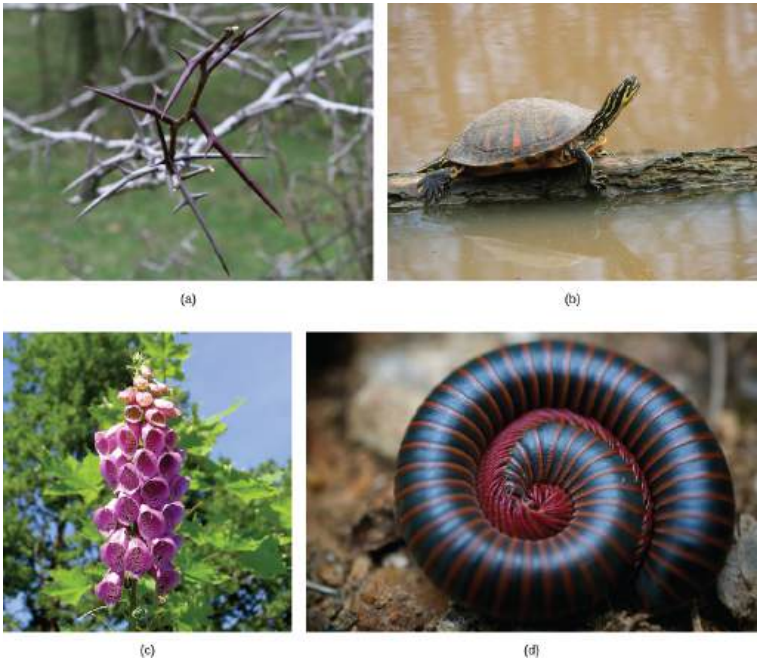


Figure 2: The (a) honey locust tree (Gleditsia triacanthos) uses thorns, a mechanical defense, against herbivores, while the (b) Florida red-bellied turtle (Pseudemys nelsoni) uses its shell as a mechanical defense against predators. (c) Foxglove (Digitalis sp.) uses a chemical defense: toxins produced by the plant can cause nausea, vomiting, hallucinations, convulsions, or death when consumed. (d) The North American millipede (Narceus americanus) uses both mechanical and chemical defenses: when threatened, the millipede curls into a defensive ball and produces a noxious substance that irritates eyes and skin. (credit a: modification of work by Huw Williams; credit b: modification of work by "JamieS93"/Flickr; credit c: modification of work by Philip Jägenstedt; credit d: modification of work by Cory Zanker. "this image" by OpenStax is licensed under CC BY 4.0)

Many species use their body shape and coloration to avoid being detected by predators. The tropical walking stick is an

insect with the coloration and body shape of a twig which makes it very hard to see when stationary against a background of real twigs (Figure 3). In another example, the chameleon can change its color to match its surroundings. Both of these are examples of **camouflage** or avoiding detection by blending in with the background.

Figure 3a: The tropical walking stick and (bottom) the chameleon use body shape and/or coloration to prevent detection by predators. (credit a: modification of work by Linda Tanner; credit b: modification of work by Frank Vassen. "this image" by OpenStax is licensed under CC BY 4.0)

Figure 3b: The chameleon uses body coloration to prevent detection by predators.

Some species use coloration as a way of warning predators that they are not good to eat. For example, the cinnabar moth caterpillar, the fire-bellied toad, and many species of beetle have bright colors that warn of a foul taste, the presence of toxic chemicals, and/or the ability to sting or bite, respectively. Predators that ignore this coloration and eat the organisms will experience their unpleasant taste or presence of toxic chemicals and learn not to eat them in the future. This type of defensive mechanism is called **aposematic coloration**, or **warning coloration** (Figure 4).

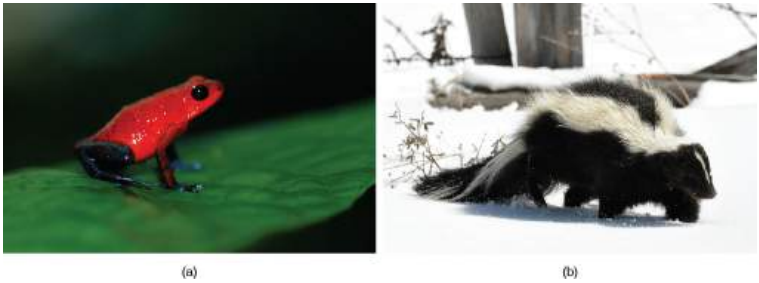


Figure 4: (a) The strawberry poison dart frog (Oophaga pumilio) uses aposematic coloration to warn predators that it is toxic, while the (b) striped skunk (Mephitis mephitis) uses aposematic coloration to warn predators of the unpleasant odor it produces. (credit a: modification of work by Jay Iwasaki; credit b: modification of work by Dan Dzurisin. "warning coloration" by OpenStax is licensed under CC BY 4.0)

While some predators learn to avoid eating certain potential prey because of their coloration, other species have evolved mechanisms to mimic this coloration to avoid being eaten, even though they themselves may not be unpleasant to eat or contain toxic chemicals. In **Batesian mimicry**, a harmless species imitates the warning coloration of a harmful one. Assuming they share the same predators, this coloration then protects the harmless ones, even though they do not have the same level of physical or chemical defenses against predation as the organism they mimic. Many insect species mimic the coloration of wasps or bees, which are stinging, venomous insects, thereby discouraging predation (Figure 5).

Figure 5a: Batesian mimicry occurs when a harmless species mimics the coloration of a harmful species, as is seen with the (top) bumblebee and (bottom) bee-like robber fly. (credit a, b: modification of work by Cory Zanker. "Batesian mimicry" by OpenStax is licensed under CC BY 4.0)

Figure 5b: Batesian mimicry occurs when a harmless species mimics the coloration of a harmful species, as is seen with the (top) bumblebee and (bottom) bee-like robber fly. (credit a, b: modification of work by Cory Zanker. "Batesian mimicry" by OpenStax is licensed under CC BY 4.0)

In **Müllerian mimicry**, multiple species share the same warning coloration, but all of them actually have defenses. Figure 6 shows a variety of foul-tasting butterflies with similar coloration. In **Emsleyan/Mertensian mimicry**, a deadly prey mimics a less dangerous one, such as the venomous coral snake mimicking the non-venomous milk snake. This type of mimicry is extremely rare and more difficult to understand than the previous two types. For this type of mimicry to work, it is essential that eating the milk snake has unpleasant but not fatal consequences. Then, these predators learn not to eat snakes with this coloration, protecting the coral snake as well. If the snake were fatal to the predator, there would be no opportunity for the predator to learn not to eat it, and the benefit for the less toxic species would disappear.

Figure 6: Several unpleasant-tasting Heliconius butterfly species share a similar color pattern with better-tasting varieties, an example of Müllerian mimicry. (credit: Joron M, Papa R, Beltrán M, Chamberlain N, Mavárez J, et al. "Heliconius" by OpenStax is licensed under CC BY 4.0)

Competitive Exclusion Principle

Resources are often limited within a habitat and multiple species may compete to obtain them. All species have an ecological niche in the ecosystem, which describes how they acquire the resources they need and how they interact with other species in the community. The **competitive exclusion principle** states that two species cannot occupy the same

niche in a habitat. In other words, different species cannot coexist in a community if they are competing for all the same resources. An example of this principle is shown in Figure 7, with two protozoan species, *Paramecium aurelia* and *Paramecium caudatum*. When grown individually in the laboratory, they both thrive. But when they are placed together in the same test tube (habitat), *P. aurelia* outcompetes *P. caudatum* for food, leading to the latter's eventual extinction.

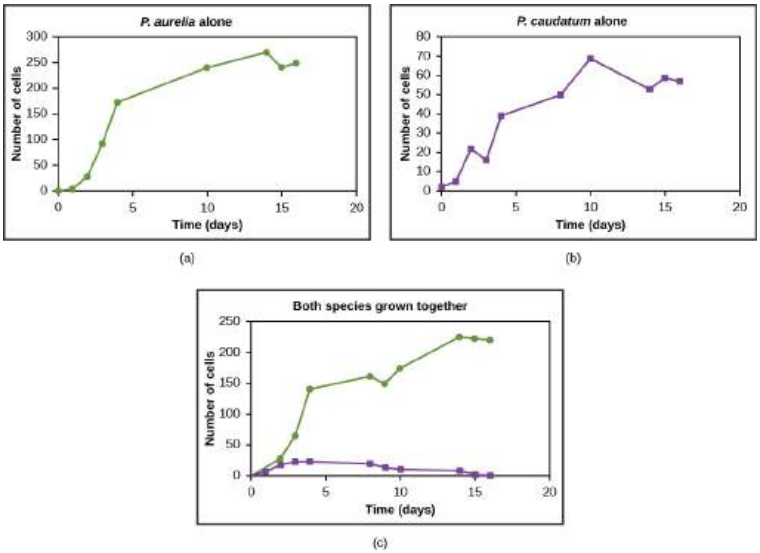


Figure 7: Paramecium aurelia and Paramecium caudatum grow well individually, but when they compete for the same resources, the P. aurelia outcompetes the P. caudatum. (credit: "Paramecium competition" by OpenStax is licensed under CC BY 4.0)

This exclusion may be avoided if a population evolves to make use of a different resource, a different area of the habitat, or feeds during a different time of day, called **resource partitioning**. The two organisms are then said to

occupy different microniches. These organisms coexist by minimizing direct competition.

Symbiosis

Symbiotic relationships, or **symbioses** (plural), are close interactions between individuals of different species over an extended period of time which impact the abundance and distribution of the associating populations. Most scientists accept this definition, but some restrict the term to only those species that are mutualistic, where both individuals benefit from the interaction. In this discussion, the broader definition will be used.

Commensalism

A **commensal relationship** occurs when one species benefits from the close, prolonged interaction, while the other neither benefits nor is harmed. Birds nesting in trees provide an example of a commensal relationship (Figure 8). The tree is not harmed by the presence of the nest among its branches. The nests are light and produce little strain on the structural integrity of the branch, and most of the leaves, which the tree uses to get energy by photosynthesis, are above the nest so they are unaffected. The bird, on the other hand, benefits greatly. If the bird had to nest in the open, its eggs and young would be vulnerable to predators. Another example of a commensal relationship is the clownfish and the sea anemone. The sea anemone is not harmed by the fish and the fish benefits with protection from predators who would be stung upon nearing the sea anemone.

Figure 8: The southern masked-weaver bird is starting to make a nest in a tree in Zambezi Valley, Zambia. This is an example of a commensal relationship, in which one species (the bird) benefits, while the other (the tree) neither benefits nor is harmed. (credit: "Hanay"/Wikimedia Commons. "Hanay" by OpenStax is licensed under CC BY 4.0)

Mutualism

A second type of symbiotic relationship is called **mutualism**, where two species benefit from their interaction. Some scientists believe that these are the only true examples of symbiosis. For example, termites have a mutualistic relationship with protozoa that live in the insect's gut (Figure 9). The termite benefits from the ability of bacterial

symbionts within the protozoa to digest cellulose. The termite itself cannot do this, and without the protozoa, it would not be able to obtain energy from its food (cellulose from the wood it chews and eats). The protozoa and the bacterial symbionts benefit by having a protective environment and a constant supply of food from the wood chewing actions of the termite. Lichens have a mutualistic relationship between fungus and photosynthetic algae or bacteria (Figure 9). As these symbionts grow together, the glucose produced by the algae provides nourishment for both organisms, whereas the physical structure of the lichen protects the algae from the elements and makes certain nutrients in the atmosphere more available to the algae.

Figure 9a: Termites form a mutualistic relationship with symbiotic protozoa in their guts, which allow both organisms to obtain energy from the cellulose the termite consumes. (credit a: modification of work by Scott Bauer, USDA)

Figure 9b: Lichen is a fungus that has symbiotic photosynthetic algae living inside its cells. (credit b: modification of work by Cory Zanker. "mutualism" by OpenStax is licensed under CC BY 4.0)

Parasitism

A **parasite** is an organism that lives in or on another living organism and derives nutrients from it. In a **parasitic relationship**, the parasite benefits, but the organism being fed upon, the host is harmed. The host is usually weakened by the parasite as it siphons resources the host would normally use to maintain itself. The parasite, however, is unlikely to kill the host, especially not quickly, because this would allow no time for the organism to complete its reproductive cycle by spreading to another host.

The reproductive cycles of parasites are often very complex, sometimes requiring more than one host species. A tapeworm is a parasite that causes disease in humans when contaminated, undercooked meat such as pork, fish, or beef

is consumed (Figure 10). The tapeworm can live inside the intestine of the host for several years, benefiting from the food the host is bringing into its gut by eating, and may grow to be over 50 ft long by adding segments. The parasite moves from species to species in a cycle, making two hosts necessary to complete its life cycle. Another common parasite is *Plasmodium falciparum*, the protozoan cause of malaria, a significant disease in many parts of the world. Living in human liver and red blood cells, the organism reproduces asexually in the gut of blood-feeding mosquitoes to complete its life cycle. Thus malaria is spread from human to human by mosquitoes, one of many arthropod-borne infectious diseases.

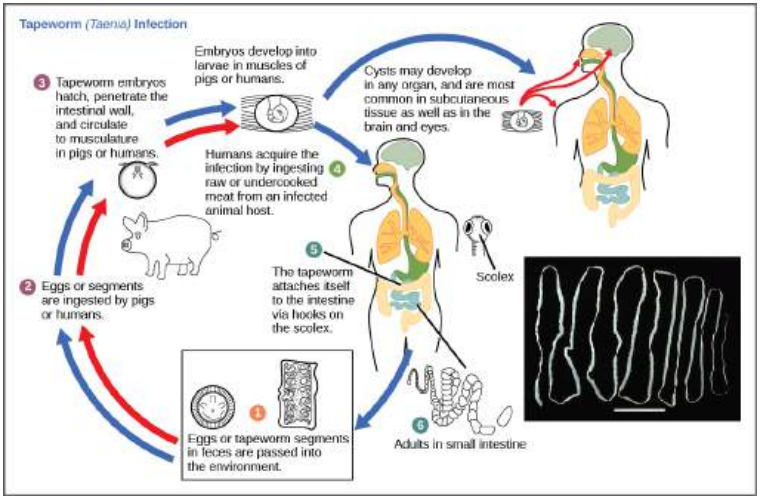


Figure 10: This diagram shows the life cycle of a pork tapeworm (Taenia solium), a human worm parasite. (credit: modification of work by CDC. "Taenia solium" by OpenStax is licensed under CC BY 4.0)

Characteristics of Communities

Communities are complex entities that can be characterized by their **structure** (the types and numbers of species present) and **dynamics** (how communities change over time). Understanding community structure and dynamics enable community ecologists to manage ecosystems more effectively.

Foundation Species

Foundation species are considered the "base" or "bedrock" of a community, having the greatest influence on its overall structure. They are usually the **primary producers**: organisms that bring most of the energy into the community. Kelp, brown algae, is a foundation species, forming the basis of the kelp forests off the coast of California.

Foundation species may physically modify the environment to produce and maintain habitats that benefit the other organisms that use them. An example is the photosynthetic corals of the coral reef (Figure 11). Corals themselves are not photosynthetic, but harbor symbionts within their body tissues (dinoflagellates called zooxanthellae) that perform photosynthesis; this is another example of a mutualism. The exoskeletons of living and dead coral make up most of the reef structure, which protects many other species from waves and ocean currents.

Figure 11: Coral is the foundation species of coral reef ecosystems. (credit: Jim E. Maragos, USFWS. "coral reflection" by OpenStax is licensed under CC BY 4.0)

Biodiversity, Species Richness, and Relative Species Abundance

Biodiversity describes a community's **biological complexity**: it is measured by the number of different species (**species richness**) in a particular area and their relative abundance (**species evenness**). The area in question could be a habitat, a biome, or the entire biosphere. Species richness is the term that is used to describe the number of species living in a habitat or biome. Species richness varies across the globe (Figure 12). One factor in determining species richness is latitude, with the greatest species richness occurring in ecosystems near the equator, which often have warmer temperatures, large amounts of rainfall, and low seasonality. The lowest species richness occurs near the poles, which

are much colder, drier, and thus less conducive to life in Geologic time (time since glaciations). The predictability of climate or productivity is also an important factor. Other factors influence species richness as well. For example, the study of island biogeography attempts to explain the relatively high species richness found in certain isolated island chains, including the Galápagos Islands that inspired the young Darwin. Relative species abundance is the number of individuals in a species relative to the total number of individuals in all species within a habitat, ecosystem, or biome. Foundation species often have the highest relative abundance of species.

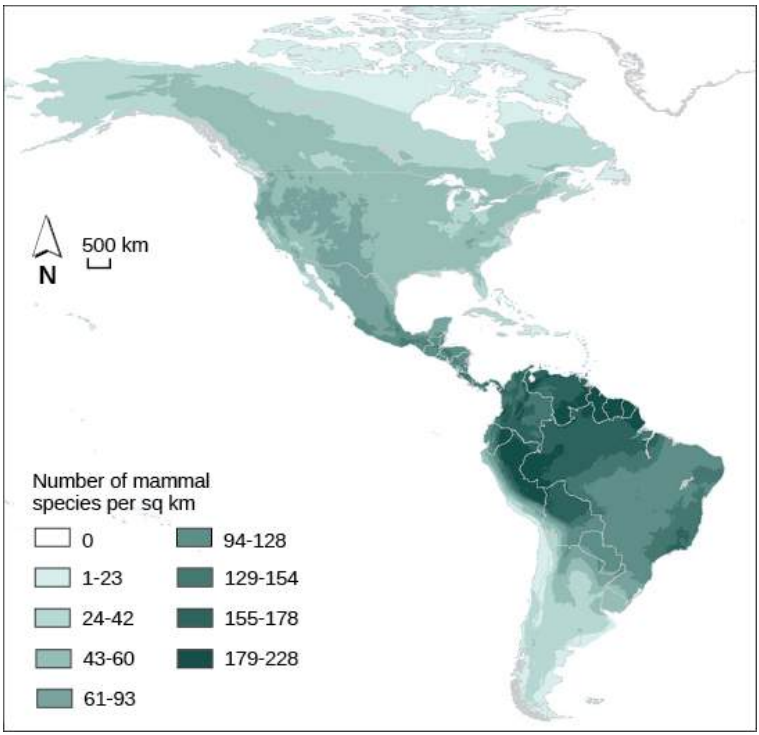


Figure 12: The greatest species richness for mammals in North and South America is associated with the equatorial latitudes. (credit: modification of work by NASA, CIESIN, Columbia University. "species richness map" by OpenStax is licensed under CC BY 4.0)

Keystone Species

A keystone species is one whose presence is key to maintaining biodiversity within an ecosystem and to upholding an ecological community's structure. The intertidal sea star, *Pisaster ochraceus*, of the northwestern United States is a keystone species (Figure 13). Studies have shown that when this organism is removed from communities, populations of their natural prey (mussels) increase, completely altering the species composition and reducing

biodiversity. Another keystone species is the banded tetra, a fish in tropical streams, which supplies nearly all of the phosphorus, a necessary inorganic nutrient, to the rest of the community. If these fish were to become extinct, the community would be greatly affected.

Figure 13: The Pisaster ochraceus sea star is a keystone species. (credit: Jerry Kirkhart. "Pisaster ochraceus" by OpenStax is licensed under CC BY 4.0)

Everyday Connection – Invasive Species

Invasive species are non-native organisms that, when introduced to an area out of their native range, threaten the ecosystem balance of that habitat. Many such species exist in the United States, as shown in Figure 14. Whether enjoying a forest hike, taking a summer boat trip, or simply walking down an urban street, you have likely encountered an invasive species.

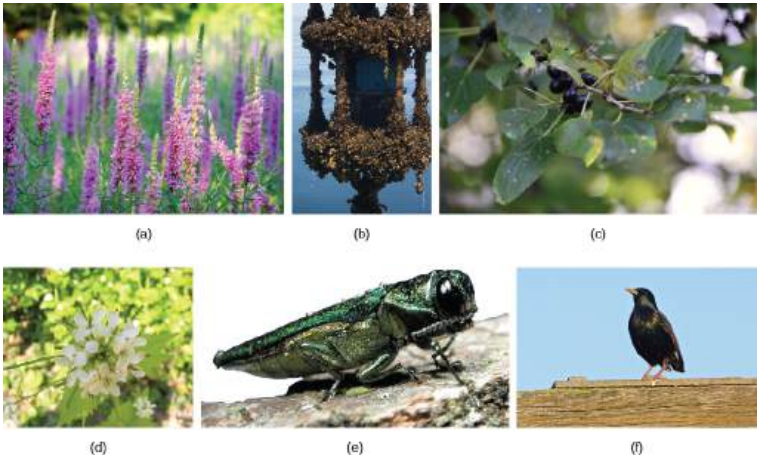


Figure 14: In the United States, invasive species like (a) purple loosestrife (Lythrum salicaria) and the (b) zebra mussel (Dreissena polymorpha) threaten certain aquatic ecosystems. Some forests are threatened by the spread of (c) common buckthorn (Rhamnus cathartica), (d) garlic mustard (Alliaria petiolata), and (e) the emerald ash borer (Agrilus planipennis). The (f) European starling (Sturnus vulgaris) may compete with native bird species for nest holes. (credit a: modification of work by Liz West; credit b: modification of work by M. McCormick, NOAA; credit c: modification of work by E. Dronkert; credit d: modification of work by Dan Davison; credit e: modification of work by USDA; credit f: modification of work by Don DeBold. "this image" by OpenStax is licensed under CC BY 4.0)

One of the many recent proliferations of an invasive species concerns the growth of **Asian carp** populations. Asian carp were introduced to the United States in the 1970s by fisheries and sewage treatment facilities that used the fish's excellent filter feeding capabilities to clean their ponds of excess plankton. Some of the fish escaped, however, and by the 1980s they had colonized many waterways of the Mississippi River basin, including the Illinois and Missouri Rivers.

Voracious eaters and rapid reproducers, Asian carp may

outcompete native species for food, potentially leading to their extinction. For example, black carp are voracious eaters of native mussels and snails, limiting this food source for native fish species. Silver carp eat plankton that native mussels and snails feed on, reducing this food source by a different alteration of the food web. In some areas of the Mississippi River, Asian carp species have become the most predominant, effectively outcompeting native fishes for habitat. In some parts of the Illinois River, Asian carp constitute 95 percent of the community's biomass. Although edible, the fish is bony and not a desired food in the United States. Moreover, their presence threatens the native fish and fisheries of the Great Lakes, which are important to local economies and recreational anglers. Asian carp have even injured humans. The fish, frightened by the sound of approaching motorboats, thrust themselves into the air, often landing in the boat or directly hitting the boaters.

The Great Lakes and their prized salmon and lake trout fisheries are also being threatened by these invasive fish. Asian carp have already colonized rivers and canals that lead into Lake Michigan. One infested waterway of particular importance is the Chicago Sanitary and Ship Channel, the major supply waterway linking the Great Lakes to the Mississippi River. To prevent the Asian carp from leaving the canal, a series of electric barriers have been successfully used to discourage their migration; however, the threat is significant enough that several states and Canada have sued to have the Chicago channel permanently cut off from Lake Michigan. Local and national politicians have weighed in on how to solve the problem, but no one knows whether the Asian carp will ultimately be considered a nuisance, like other invasive species such as the water hyacinth and zebra mussel, or whether it will be the destroyer of the largest freshwater fishery of the world.

The issues associated with Asian carp show how

population and community ecology, fisheries management, and politics intersect on issues of vital importance to the human food supply and economy. Socio-political issues like this make extensive use of the sciences of population ecology (the study of members of a particular species occupying a particular area known as a habitat) and community ecology (the study of the interaction of all species within a habitat).

Community Dynamics

Community dynamics are the changes in community structure and composition over time. Sometimes these changes are induced by environmental disturbances such as volcanoes, earthquakes, storms, fires, and climate change. Communities with a stable structure are said to be at equilibrium. Following a disturbance, the community may or may not return to the equilibrium state.

Succession describes the sequential appearance and disappearance of species in a community over time. In primary succession, newly exposed or newly formed land is colonized by living things; in secondary succession, part of an ecosystem is disturbed and remnants of the previous community remain.

Primary Succession and Pioneer Species

Primary succession occurs when new land is formed or rock is exposed: for example, following the eruption of volcanoes, such as those on the Big Island of Hawaii. As lava flows into the ocean, new land is continually being formed. On the Big Island, approximately 32 acres of land is added each year. First, weathering and other natural forces break down the substrate enough for the establishment of certain hearty plants and lichens with few soil requirements, known as **pioneer species**. These species help to further break down

the mineral rich lava into soil where other, less hardy species will grow and eventually replace the pioneer species. In addition, as these early species grow and die, they add to an ever-growing layer of decomposing organic material and contribute to soil formation. Over time the area will reach an equilibrium state, with a set of organisms quite different from the pioneer species.

Secondary succession

A classic example of **secondary succession** occurs in oak and hickory forests cleared by wildfire (Figure 15). Wildfires will burn most vegetation and kill those animals unable to flee the area. Their nutrients, however, are returned to the ground in the form of ash. Thus, even when areas are devoid of life due to severe fires, the area will soon be ready for new life to take hold.

Before the fire, the vegetation was dominated by tall trees with access to the major plant energy resource: sunlight. Their height gave them access to sunlight while also shading the ground and other low-lying species. After the fire, though, these trees are no longer dominant. Thus, the first plants to grow back are usually annual plants followed within a few years by quickly growing and spreading grasses and other pioneer species. Due to, at least in part, changes in the environment brought on by the growth of the grasses and other species, over many years, shrubs will emerge along with small pine, oak, and hickory trees. These organisms are called intermediate species. Eventually, over 150 years, the forest will reach its equilibrium point where species composition is no longer changing and resembles the community before the fire. This equilibrium state is referred to as the climax community, which will remain stable until the next disturbance.

Figure 15: Secondary succession is shown in an oak and hickory forest after a forest fire. (credit: "Secondary succession" by OpenStax is licensed under CC BY 4.0)

Summary

Communities include all the different species living in a given area. The variety of these species is called species richness. Many organisms have developed defenses against predation and herbivory, including mechanical defenses, warning coloration, and mimicry, as a result of evolution and the interaction with other members of the community. Two species cannot exist in the same habitat competing directly for the same resources. Species may form symbiotic relationships such as commensalism or mutualism. Community structure is described by its foundation and keystone species. Communities respond to environmental disturbances by succession (the predictable appearance of

different types of plant species) until a stable community structure is established.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Behavioral Biology: Proximate and Ultimate Causes of Behavior

Behavior is the change in activity of an organism in response to a stimulus. Behavioral biology is the study of the biological and evolutionary bases for such changes. The idea that behaviors evolved as a result of the pressures of natural selection is not new. Animal behavior has been studied for decades, by biologists in the science of ethology, by psychologists in the science of comparative psychology, and by scientists of many disciplines in the study of neurobiology. Although there is overlap between these disciplines, scientists in these behavioral fields take different approaches. Comparative psychology is an extension of work done in human and behavioral psychology. Ethology is an extension of genetics, evolution, anatomy, physiology, and other biological disciplines. Still, one cannot study behavioral biology without touching on both comparative psychology and ethology.

One goal of behavioral biology is to dissect out the innate behaviors, which have a strong genetic component and are largely independent of environmental influences, from the learned behaviors, which result from environmental conditioning. Innate behavior, or instinct, is important

because there is no risk of an incorrect behavior being learned. They are "hard wired" into the system. On the other hand, learned behaviors, although riskier, are flexible, dynamic, and can be altered according to changes in the environment.

Innate Behaviors: Movement and Migration

Innate or **instinctual behaviors** rely on response to stimuli. The simplest example of this is a **reflex action**, an involuntary and rapid response to stimulus. To test the "knee-jerk" reflex, a doctor taps the patellar tendon below the kneecap with a rubber hammer. The stimulation of the nerves there leads to the reflex of extending the leg at the knee. This is similar to the reaction of someone who touches a hot stove and instinctually pulls his or her hand away. Even humans, with our great capacity to learn, still exhibit a variety of innate behaviors.

Kinesis and Taxis

Another activity or movement of innate behavior is **kinesis**, or the undirected movement in response to a stimulus. **Orthokinesis** is the increased or decreased speed of movement of an organism in response to a stimulus. Woodlice, for example, increase their speed of movement when exposed to high or low temperatures. This movement, although random, increases the probability that the insect spends less time in the unfavorable environment. Another example is **klinokinesis**, an increase in turning behaviors. It is exhibited by bacteria such as *E. coli* which, in association with orthokinesis, helps the organisms randomly find a more hospitable environment.

A similar, but more directed version of kinesis is **taxis**: the directed movement towards or away from a stimulus.

This movement can be in response to light (**phototaxis**), chemical signals (**chemotaxis**), or gravity (**geotaxis**) and can be directed toward (**positive**) or away (**negative**) from the source of the stimulus. An example of a positive chemotaxis is exhibited by the unicellular protozoan *Tetrahymena thermophila*. This organism swims using its cilia, at times moving in a straight line, and at other times making turns. The attracting chemotactic agent alters the frequency of turning as the organism moves directly toward the source, following the increasing concentration gradient.

Fixed Action Patterns

A fixed action pattern is a series of movements elicited by a stimulus such that even when the stimulus is removed, the pattern goes on to completion. An example of such a behavior occurs in the three-spined stickleback, a small freshwater fish (Figure 1). Males of this species develop a red belly during breeding season and show instinctual aggressiveness to other males during this time. In laboratory experiments, researchers exposed such fish to objects that in no way resemble a fish in their shape, but which were painted red on their lower halves. The male sticklebacks responded aggressively to the objects just as if they were real male sticklebacks.

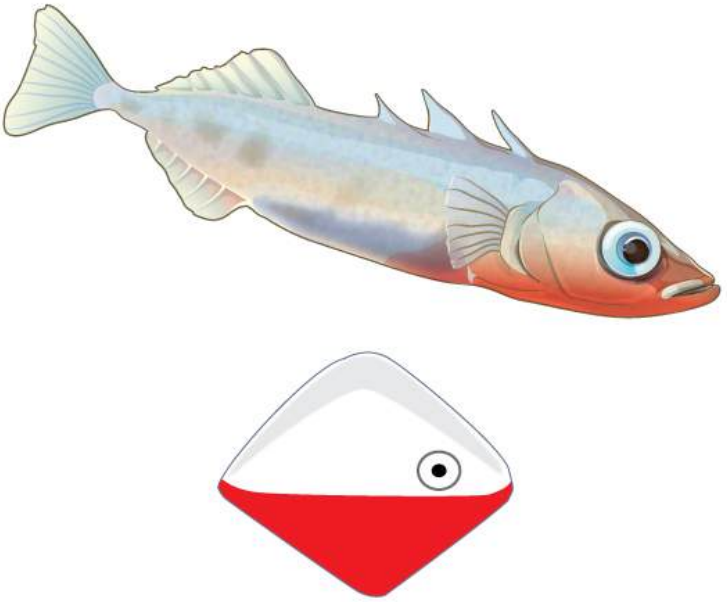


Figure 1: Male three-spined stickleback fish exhibit a fixed action pattern. During mating season, the males, which develop a bright red belly, react strongly to red-bottomed objects that in no way resemble fish. (credit: "fixed action pattern" by OpenStax is licensed under CC BY 4.0)

Migration

Migration is the long-range seasonal movement of animals. It is an evolved, adapted response to variation in resource availability, and it is a common phenomenon found in all major groups of animals. Birds fly south for the winter to get to warmer climates with sufficient food, and salmon migrate to their spawning grounds. The popular 2005 documentary *March of the Penguins* followed the 62-mile migration of emperor penguins through Antarctica to bring food back to their breeding site and to their young.

Although migration is thought of as innate behavior, only

some migrating species always migrate (**obligate migration**). Animals that exhibit facultative migration can choose to migrate or not. Additionally, in some animals, only a portion of the population migrates, whereas the rest does not migrate (**incomplete migration)**. For example, owls that live in the tundra may migrate in years when their food source, small rodents, is relatively scarce, but not migrate during the years when rodents are plentiful.

Foraging

Foraging is the act of searching for and exploiting food resources. Feeding behaviors that maximize energy gain and minimize energy expenditure are called optimal foraging behaviors, and these are favored by natural section.

Innate Behaviors: Living in Groups

Not all animals live in groups, but even those that live relatively solitary lives, with the exception of those that can reproduce asexually, must mate. Mating usually involves one animal signaling another so as to communicate the desire to mate. There are several types of energy-intensive behaviors or displays associated with mating, called **mating rituals**. Other behaviors found in populations that live in groups are described in terms of which animal benefits from the behavior. In **selfish behavior**, only the animal in question benefits; in **altruistic behavior**, one animal's actions benefit another animal; cooperative behavior describes when both animals benefit. All of these behaviors involve some sort of communication between population members.

Communication within a Species

Animals communicate with each other using stimuli known

as **signals**. An example of this is seen in the three-spined stickleback, where the visual signal of a red region in the lower half of a fish signals males to become aggressive and signals females to mate. Other signals are chemical (**pheromones**), **aural** (sound), **visual** (courtship and aggressive displays), or **tactile** (touch). These types of communication may be instinctual or learned or a combination of both. These are not the same as the communication we associate with language, which has been observed only in humans and perhaps in some species of primates and cetaceans.

A pheromone is a secreted chemical signal used to obtain a response from another individual of the same species. The purpose of pheromones is to elicit a specific behavior from the receiving individual. Pheromones are especially common among social insects, but they are used by many species to attract the opposite sex, to sound alarms, to mark food trails, and to elicit other, more complex behaviors. Even humans are thought to respond to certain pheromones called **axillary steroids**. These chemicals influence human perception of other people, and in one study were responsible for a group of women synchronizing their menstrual cycles. The role of pheromones in human-to-human communication is still somewhat controversial and continues to be researched.

Songs are an example of an aural signal, one that needs to be heard by the recipient. Perhaps the best known of these are songs of birds, which identify the species and are used to attract mates. Other well-known songs are those of whales, which are of such low frequency that they can travel long distances underwater. Dolphins communicate with each other using a wide variety of vocalizations. Male crickets make chirping sounds using a specialized organ to attract a mate, repel other males, and to announce a successful mating.

Courtship displays are a series of ritualized visual behaviors (signals) designed to attract and convince a member of the opposite sex to mate. These displays are ubiquitous in the animal kingdom. Often these displays involve a series of steps, including an initial display by one member followed by a response from the other. If at any point, the display is performed incorrectly or a proper response is not given, the mating ritual is abandoned and the mating attempt will be unsuccessful.

Aggressive displays are also common in the animal kingdom. An example is when a dog bares its teeth when it wants another dog to back down. Presumably, these displays communicate not only the willingness of the animal to fight but also it's fighting ability. Although these displays do signal aggression on the part of the sender, it is thought that these displays are actually a mechanism to reduce the amount of actual fighting that occurs between members of the same species: they allow individuals to assess the fighting ability of their opponent and thus decide whether it is "worth the fight." The testing of certain hypotheses using game theory has led to the conclusion that some of these displays may overstate an animal's actual fighting ability and are used to "bluff" the opponent. This type of interaction, even if "dishonest," would be favored by natural selection if it is successful more times than not.

Distraction displays are seen in birds and some fish. They are designed to attract a predator away from the nest that contains their young. This is an example of an altruistic behavior: it benefits the young more than the individual performing the display, which is putting itself at risk by doing so.

Many animals, especially primates, communicate with other members in the group through **touch**. Activities such as grooming, touching the shoulder or root of the tail, embracing, lip contact, and greeting ceremonies have all

been observed in the Indian langur, an Old World monkey. Similar behaviors are found in other primates, especially in the great apes.

Altruistic Behaviors

Behaviors that lower the fitness of the individual but increase the fitness of another individual are termed altruistic. Examples of such behaviors are seen widely across the animal kingdom. Social insects such as worker bees have no ability to reproduce, yet they maintain the queen so she can populate the hive with her offspring. Meerkats keep a **sentry** standing guard to warn the rest of the colony about intruders, even though the sentry is putting itself at risk. Wolves and wild dogs bring meat to pack members not present during a hunt. Lemurs take care of infants unrelated to them. Although on the surface, these behaviors appear to be altruistic, it may not be so simple.

There has been much discussion over why altruistic behaviors exist. Do these behaviors lead to overall evolutionary advantages for their species? Do they help the altruistic individual pass on its own genes? And what about such activities between unrelated individuals? One explanation for altruistic-type behaviors is found in the genetics of natural selection. In the 1976 book, *The Selfish Gene,* scientist Richard Dawkins attempted to explain many seemingly altruistic behaviors from the viewpoint of the gene itself. Although a gene obviously cannot be selfish in the human sense, it may appear that way if the sacrifice of an individual benefits related individuals that share genes that are identical by descent (present in relatives because of common lineage). Mammal parents make this sacrifice to take care of their offspring. Emperor penguins migrate miles in harsh conditions to bring food back for their young. The

selfish gene concept has been controversial over the years and is still discussed among scientists in related fields.

Even less-related individuals, those with less genetic identity than that shared by parent and offspring, benefit from seemingly altruistic behavior. The activities of social insects such as bees, wasps, ants, and termites are good examples. Sterile workers in these societies take care of the queen because they are closely related to it, and as the queen has offspring, she is passing on genes from the workers indirectly. Thus, it is of fitness benefit for the worker to maintain the queen without having any direct chance of passing on its genes due to its sterility. The lowering of individual fitness to enhance the reproductive fitness of a relative and thus one's **inclusive fitness** evolves through kin selection. This phenomenon can explain many superficially altruistic behaviors seen in animals. However, these behaviors may not be truly defined as altruism in these cases because the actor is actually increasing its own fitness either directly (through its own offspring) or indirectly (through the inclusive fitness it gains through relatives that share genes with it).

Unrelated individuals may also act altruistically to each other, and this seems to defy the selfish gene explanation. An example of this observed in many monkey species where a monkey will present its back to an unrelated monkey to have that individual pick the parasites from its fur. After a certain amount of time, the roles are reversed and the first monkey now grooms the second monkey. Thus, there is **reciprocity** in the behavior. Both benefit from the interaction and their fitness is raised more than if neither cooperated nor if one cooperated and the other did not cooperate. This behavior is still not necessarily altruism, as the giving behavior of the actor is based on the expectation that it will be the receiver of the behavior in the future, termed reciprocal altruism. Reciprocal altruism requires that individuals repeatedly

encounter each other, often the result of living in the same social group and that cheaters (those that never give back) are punished.

Evolutionary game theory, a modification of classical game theory in mathematics, has shown that many of these so-called altruistic behaviors are not altruistic at all. The definition of pure altruism, based on human behavior, is an action that benefits another without any direct benefit to oneself. Most of the behaviors previously described do not seem to satisfy this definition, and game theorists are good at finding selfish components in them. Others have argued that the terms "selfish" and "altruistic" should be dropped completely when discussing animal behavior, as they describe human behavior and may not be directly applicable to instinctual animal activity. What is clear, though, is that heritable behaviors that improve the chances of passing on one's genes or a portion of one's genes are favored by natural selection and will be retained in future generations as long as those behaviors convey a fitness advantage. These instinctual behaviors may then be applied, in special circumstances, to other species, as long as it doesn't lower the animal's fitness.

Finding Sex Partners

Not all animals reproduce sexually, but many that do have the same challenge: they need to find a suitable mate and often have to compete with other individuals to obtain one. Significant energy is spent in the process of locating, attracting, and mating with the sex partner. Two types of selection occur during this process and can lead to traits that are important to reproduction called **secondary sexual characteristics**: **intersexual selection**, the choosing of a mate where individuals of one sex choose mates of the other sex, and **intrasexual selection**, the competition for mates

between species members of the same sex. Intersexual selection is often complex because choosing a mate may be based on a variety of visual, aural, tactile, and chemical cues. An example of intersexual selection is when female peacocks choose to mate with the male with the brightest plumage. This type of selection often leads to traits in the chosen sex that do not enhance survival but are those traits most attractive to the opposite sex (often at the expense of survival). Intrasexual selection involves mating displays and aggressive mating rituals such as rams butting heads—the winner of these battles is the one that is able to mate. Many of these rituals use up considerable energy but result in the selection of the healthiest, strongest, and/or most dominant individuals for mating. Three general mating systems, all involving innate as opposed to learned behaviors, are seen in animal populations: monogamous, polygynous, and polyandrous.

In **monogamous systems**, one male and one female are paired for at least one breeding season. In some animals, such as the gray wolf, these associations can last much longer, even a lifetime. Several explanations have been proposed for this type of mating system. The "**mate-guarding hypothesis**" states that males stay with the female to prevent other males from mating with her. This behavior is advantageous in such situations where mates are scarce and difficult to find. Another explanation is the "**male-assistance hypothesis**," where males that remain with a female to help guard and rear their young will have more and healthier offspring. Monogamy is observed in many bird populations where, in addition to the parental care from the female, the male is also a major provider of parental care for the chicks. A third explanation for the evolutionary advantages of monogamy is the "**female-enforcement hypothesis**." In this scenario, the female ensures that the male does not have other offspring that might compete with her own, so she

actively interferes with the male's signaling to attract other mates.

Polygynous mating refers to one male mating with multiple females. In these situations, the female must be responsible for most of the parental care as the single male is not capable of providing care to that many offspring. In resourced-based polygyny, males compete for territories with the best resources and then mate with females that enter the territory, drawn to its resource richness. The female benefits by mating with a dominant, genetically fit male; however, it is at the cost of having no male help in caring for the offspring. An example is seen in the yellow-rumped honeyguide, a bird whose males defend beehives because the females feed on their wax. As the females approach, the male defending the nest will mate with them. **Harem mating structures** are a type of polygynous system where certain males dominate mating while controlling a territory with resources. Elephant seals, where the alpha male dominates the mating within the group are an example. A third type of polygyny is a **lek system**. Here there is a communal courting area where several males perform elaborate displays for females, and the females choose their mate from this group. This behavior is observed in several bird species including the sage grouse and the prairie chicken.

In **polyandrous mating systems**, one female mates with many males. These types of systems are much rarer than monogamous and polygynous mating systems. In pipefishes and seahorses, males receive the eggs from the female, fertilize them, protect them within a pouch, and give birth to the offspring (Figure 2). Therefore, the female is able to provide eggs to several males without the burden of carrying the fertilized eggs.

Figure 2a: Polyandrous mating, in which one female mates with many males, occurs in the (a) seahorse and the (b) pipefish. (credit a: modification of work by Brian Gratwicke; credit b: modification of work by Stephen Childs. "this image" by OpenStax is licensed under CC BY 4.0)

Figure 2b: Polyandrous mating, in which one female mates with many males, occurs in the (a) seahorse and the (b) pipefish. (credit a: modification of work by Brian Gratwicke; credit b: modification of work by Stephen Childs. "this image" by OpenStax is licensed under CC BY 4.0)

Simple Learned Behaviors

The majority of the behaviors previously discussed were **innate** or at least have an innate component (variations on the innate behaviors may be learned). They are inherited and the behaviors do not change in response to signals from the environment. Conversely, **learned behaviors**, even though they may have instinctive components, allow an organism to adapt to changes in the environment and are modified by previous experiences. Simple learned behaviors include habituation and imprinting—both are important to the maturation process of young animals.

Habituation

Habituation is a simple form of learning in which an animal stops responding to a stimulus after a period of repeated exposure. This is a form of non-associative learning, as the stimulus is not associated with any punishment or reward. Prairie dogs typically sound an alarm call when threatened by a predator, but they become habituated to the sound of human footsteps when no harm is associated with this sound, therefore, they no longer respond to them with an alarm call. In this example, habituation is specific to the sound of human footsteps, as the animals still respond to the sounds of potential predators.

Imprinting

Imprinting is a type of learning that occurs at a particular age or a life stage that is rapid and independent of the species involved. Hatchling ducks recognize the first adult they see, their mother, and make a bond with her. A familiar sight is ducklings walking or swimming after their mothers (Figure 3). This is another type of non-associative learning but is very important in the maturation process of these animals as it encourages them to stay near their mother so they will be protected, greatly increasing their chances of survival. However, if newborn ducks see a human before they see their mother, they will imprint on the human and follow it in just the same manner as they would follow their real mother. The International Crane Foundation has helped raise the world's population of whooping cranes from 21 individuals to about 600. Imprinting hatchlings has been a key to success: biologists wear full crane costumes so the birds never "see" humans.

Figure 3: The attachment of ducklings to their mother is an example of imprinting. (credit: modification of work by Mark Harkin, "imprinting" by OpenStax is licensed under CC BY 4.0)

Conditioned Behavior

Conditioned behaviors are types of associative learning, where a stimulus becomes associated with a consequence. During **operant conditioning**, the behavioral response is modified by its consequences, with regards to its form, strength, or frequency.

Classical Conditioning

In **classical conditioning**, a response called the conditioned response is associated with a stimulus that it had previously not been associated with, the conditioned stimulus. The response to the original, unconditioned stimulus is called the unconditioned response. The most cited example of classical conditioning is Ivan Pavlov's experiments with dogs (Figure 4). In Pavlov's experiments, the unconditioned response was

the salivation of dogs in response to the unconditioned stimulus of seeing or smelling their food. The conditioning stimulus that researchers associated with the unconditioned response was the ringing of a bell. During conditioning, every time the animal was given food, the bell was rung. This was repeated during several trials. After some time, the dog learned to associate the ringing of the bell with food and to respond by salivating. After the conditioning period was finished, the dog would respond by salivating when the bell was rung, even when the unconditioned stimulus, the food, was absent. Thus, the ringing of the bell became the conditioned stimulus and the salivation became the conditioned response. Although it is thought by some scientists that the unconditioned and conditioned responses are identical, even Pavlov discovered that the saliva in the conditioned dogs had characteristic differences when compared to the unconditioned dog.

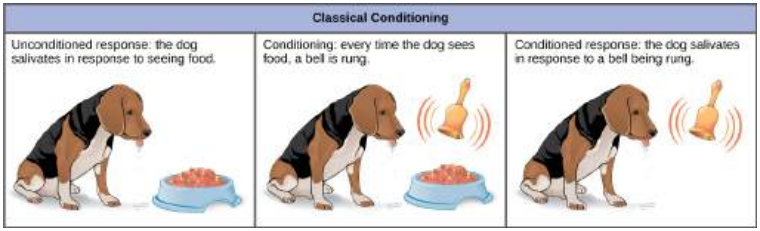


Figure 4: In the classic Pavlovian response, the dog becomes conditioned to associate the ringing of the bell with food. (credit: "Pavlovian response" by OpenStax is licensed under CC BY 4.0)

It had been thought by some scientists that this type of conditioning required multiple exposures to the paired stimulus and response, but it is now known that this is not necessary in all cases and that some conditioning can be learned in a single pairing experiment. Classical conditioning is a major tenet of behaviorism, a branch of psychological philosophy that proposes that all actions, thoughts, and

emotions of living things are behaviors that can be treated by behavior modification and changes in the environment.

Operant Conditioning

In **operant conditioning**, the conditioned behavior is gradually modified by its consequences as the animal responds to the stimulus. A major proponent of such conditioning was psychologist B.F. Skinner, the inventor of the **Skinner box**. Skinner put rats in his boxes that contained a lever that would dispense food to the rat when depressed. While initially the rat would push the lever a few times by accident, it eventually associated pushing the lever with getting the food. This type of learning is an example of operant conditioning. Operant learning is the basis of most animal training. The conditioned behavior is continually modified by **positive** or **negative reinforcement**, often a reward such as food or some type of punishment, respectively. In this way, the animal is conditioned to associate a type of behavior with the punishment or reward, and, over time, can be induced to perform behaviors that they would not have done in the wild, such as the tricks dolphins perform at marine amusement park shows (Figure 5).

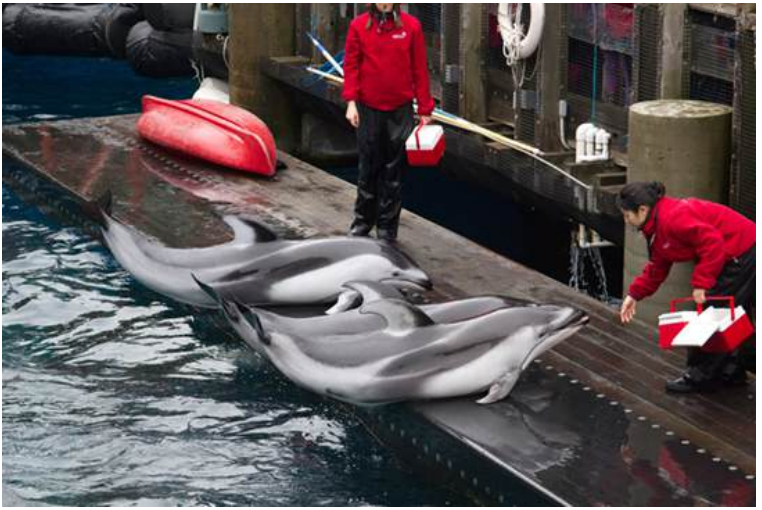


Figure 5: The training of dolphins by rewarding them with food is an example of positive reinforcement operant conditioning. (credit: Roland Tanglao. "operant conditioning" by OpenStax is licensed under CC BY 4.0)

Cognitive Learning

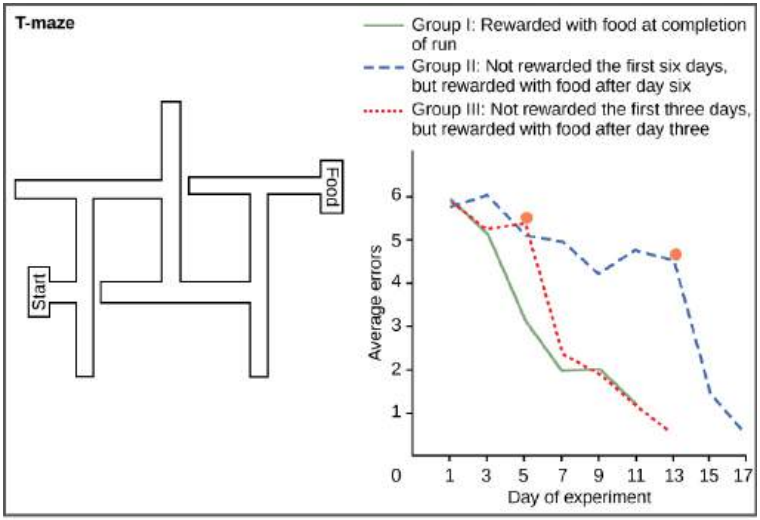
Classical and operant conditioning are inefficient ways for humans and other intelligent animals to learn. Some primates, including humans, are able to learn by imitating the behavior of others and by taking instructions. The development of complex language by humans has made **cognitive learning**, the manipulation of information using the mind, the most prominent method of human learning. In fact, that is how students are learning right now by reading this book. As students read, they can make mental images of objects or organisms and imagine changes to them, or behaviors by them, and anticipate the consequences. In addition to visual processing, cognitive learning is also

enhanced by remembering past experiences, touching physical objects, hearing sounds, tasting food, and a variety of other sensory-based inputs. Cognitive learning is so powerful that it can be used to understand conditioning in detail. In the reverse scenario, conditioning cannot help someone learn about cognition.

Classic work on cognitive learning was done by Wolfgang Köhler with chimpanzees. He demonstrated that these animals were capable of abstract thought by showing that they could learn how to solve a puzzle. When a banana was hung in their cage too high for them to reach, and several boxes were placed randomly on the floor, some of the chimps were able to stack the boxes one on top of the other, climb on top of them, and get the banana. This implies that they could visualize the result of stacking the boxes even before they had performed the action. This type of learning is much more powerful and versatile than conditioning.

Cognitive learning is not limited to primates, although they are the most efficient in using it. Maze running experiments done with rats by H.C. Blodgett in the 1920s were the first to show cognitive skills in a simple mammal. The motivation for the animals to work their way through the maze was a piece of food at its end. In these studies, the animals in Group I were run in one trial per day and had food available to them each day on completion of the run (Figure 6). Group II rats were not fed in the maze for the first six days and then subsequent runs were done with food for several days after. Group III rats had food available on the third day and every day thereafter. The results were that the control rats, Group I, learned quickly, and figured out how to run the maze in seven days. Group III did not learn much during the three days without food but rapidly caught up to the control group when given the food reward. Group II learned very slowly for the six days with no reward to motivate them, and they did not begin to catch up to the control group until the day

food was given, and then it took two days longer to learn the maze.



Redrawn after H. C. Blodgett, The effect of the introduction of reward upon the maze performance of rats. Univ. Calif. Publ. Psychol., 1929, 4, No. 8, pages 117 and 120.

Figure 6: Group I (the green solid line) found food at the end of each trial, group II (the blue dashed line) did not find food for the first 6 days, and group III (the red dotted line) did not find food during runs on the first three days. Notice that rats given food earlier learned faster and eventually caught up to the control group. The orange dots on the group II and III lines show the days when food rewards were added to the mazes.

It may not be immediately obvious that this type of learning is different than conditioning. Although one might be tempted to believe that the rats simply learned how to find their way through a conditioned series of right and left turns, E.C. Tolman proved a decade later that the rats were making a representation of the maze in their minds, which he called a "cognitive map." This was an early demonstration of the power of cognitive learning and how these abilities were not just limited to humans.

Sociobiology

Sociobiology is an interdisciplinary science originally popularized by social insect researcher E.O. Wilson in the 1970s. Wilson defined the science as "the extension of population biology and evolutionary theory to social organization" (Wilson, 1978). The main thrust of sociobiology is that animal and human behavior, including aggressiveness and other social interactions, can be explained almost solely in terms of genetics and natural selection. This science is controversial; noted scientists such as the late Stephen Jay Gould criticized the approach for ignoring the environmental effects on behavior. This is another example of the **nature versus nurture** debate of the role of genetics versus the role of the environment in determining an organism's characteristics.

Sociobiology also links genes with behaviors and has been associated with "**biological determinism**," the belief that all behaviors are hardwired into our genes. No one disputes that certain behaviors can be inherited and that natural selection plays a role in retaining them. It is the application of such principles to human behavior that sparks this controversy, which remains active today.

Summary

Behaviors are responses to stimuli. They can either be instinctual/innate behaviors, which are not influenced by the environment, or learned behaviors, which are influenced by environmental changes. Instinctual behaviors include mating systems and methods of communication. Learned behaviors include imprinting and habituation, conditioning, and, most powerfully, cognitive learning. Although the connection between behavior, genetics, and evolution is well

established, the explanation of human behavior as entirely genetic is controversial.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Wilson, E. 1978. *On Human Nature* (repr., Cambridge: Harvard University Press, 2004), xx.

ECOSYSTEMS

Figure 1: In the southwestern United States, rainy weather causes an increase in production of pinyon nuts, causing the deer mouse population to explode. Deer mice may carry a virus called Sin Nombre (a hantavirus) that causes respiratory disease in humans and has a high fatality rate. In 1992–1993, wet El Niño weather caused a Sin Nombre epidemic. Navajo healers, who were aware of the link between this disease and weather, predicted the outbreak. (credit "highway": modification of work by Phillip Capper; credit "mouse": modification of work by USFWS. "this image" by OpenStax is licensed under CC BY 4.0)

In 1993, an interesting example of ecosystem dynamics occurred when a rare lung disease struck inhabitants of the southwestern United States (Figure 1). This disease had an alarming rate of fatalities, killing more than half of early patients, many of whom were Native Americans. These

formerly healthy young adults died from complete respiratory failure. The disease was unknown, and the Centers for Disease Control (CDC), the United States government agency responsible for managing potential epidemics, was brought in to investigate. The scientists could have learned about the disease had they known to talk with the Navajo healers who lived in the area and who had observed the connection between rainfall and mice populations, thereby predicting the 1993 outbreak.

The cause of the disease, determined within a few weeks by the CDC investigators, was the hantavirus known as *Sin Nombre*, the virus with "no name." With insights from traditional Navajo medicine, scientists were able to characterize the disease rapidly and institute effective health measures to prevent its spread. This example illustrates the importance of understanding the complexities of ecosystems and how they respond to changes in the environment.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Ecology of Ecosystems

Life in an ecosystem is often about competition for limited resources, a characteristic of the theory of natural selection. Competition in **communities** (all living things within specific habitats) is observed both within species and among different species. The resources for which organisms compete include organic material from living or previously living organisms, sunlight, and mineral nutrients, which provide the energy for living processes and the matter to make up organisms' physical structures. Other critical factors influencing community dynamics are the components of its physical and geographic environment: a habitat's latitude, amount of rainfall, topography (elevation), and available species. These are all important environmental variables that determine which organisms can exist within a particular area.

An **ecosystem** is a community of living organisms and their interactions with their **abiotic** (non-living) environment. Ecosystems can be small, such as the tide pools found near the rocky shores of many oceans, or large, such as the Amazon Rainforest in Brazil (Figure 1).

Figure 1: A (a) tidal pool ecosystem in Matinicus Island in Maine is a small ecosystem, while the (b) Amazon Rainforest in Brazil is a large ecosystem. (credit a: modification of work by "takomabibelot"/Flickr; credit b: modification of work by Ivan Mlinaric. "this image" by OpenStax is licensed under CC BY 4.0)

There are three broad categories of ecosystems based on their general environment: **freshwater**, **ocean water**, and **terrestrial**. Within these broad categories are individual ecosystem types based on the organisms present and the type of environmental habitat.

Ocean ecosystems are the most common, comprising 75 percent of the Earth's surface and consisting of three basic types: **shallow ocean**, **deep ocean water**, and **deep ocean surfaces** (the low depth areas of the deep oceans). The shallow ocean ecosystems include extremely biodiverse **coral reef ecosystems**, and the deep ocean surface is known for its large numbers of **plankton** and **krill** (small crustaceans) that support it. These two environments are especially important to aerobic respirators worldwide as the **phytoplankton** perform 40 percent of all photosynthesis on Earth. Although not as diverse as the other two, deep ocean ecosystems contain a wide variety of marine organisms. Such ecosystems exist even at the bottom of the ocean where light is unable to penetrate through the water.

Freshwater ecosystems are the rarest, occurring on only

1.8 percent of the Earth's surface. Lakes, rivers, streams, and springs comprise these systems; they are quite diverse, and they support a variety of fish, amphibians, reptiles, insects, phytoplankton, fungi, and bacteria.

Terrestrial ecosystems, also known for their diversity, are grouped into large categories called **biomes**, such as tropical rain forests, savannas, deserts, coniferous forests, deciduous forests, and tundra. Grouping these ecosystems into just a few biome categories obscures the great diversity of the individual ecosystems within them. For example, there is great variation in desert vegetation: the saguaro cacti and other plant life in the Sonoran Desert, in the United States, are relatively abundant compared to the desolate rocky desert of Boa Vista, an island off the coast of Western Africa (Figure 2).

Figure 2: Desert ecosystems, like all ecosystems, can vary greatly. The desert in (a) Saguaro National Park, Arizona, has abundant plant life, while the rocky desert of (b) Boa Vista island, Cape Verde, Africa, is devoid of plant life. (credit a: modification of work by Jay Galvin; credit b: modification of work by Ingo Wölbern. "Desert ecosystems" by OpenStax is licensed under CC BY 4.0)

Ecosystems are complex with many interacting parts. They are routinely exposed to various **disturbances** or changes in the environment that affect their compositions: yearly variations in rainfall and temperature and the slower processes of plant growth, which may take several years.

Many of these disturbances are a result of natural processes. For example, when lightning causes a forest fire and destroys part of a forest ecosystem, the ground is eventually populated by grasses, then by bushes and shrubs, and later by mature trees, restoring the forest to its former state. The impact of environmental disturbances caused by human activities is as important as the changes wrought by natural processes. Human agricultural practices, air pollution, acid rain, global deforestation, overfishing, eutrophication, oil spills, and illegal dumping on land and into the ocean are all issues of concern to conservationists.

Equilibrium is the steady-state of an ecosystem where all organisms are in balance with their environment and with each other. In ecology, two parameters are used to measure changes in ecosystems: **resistance** and **resilience**. The ability of an ecosystem to remain at equilibrium in spite of disturbances is called resistance. The speed at which an ecosystem recovers equilibrium after being disturbed is called its resilience. Ecosystem resistance and resilience are especially important when considering human impact. The nature of an ecosystem may change to such a degree that it can lose its resilience entirely. This process can lead to the complete destruction or irreversible altering of the ecosystem.

Food Chains and Food Webs

The term **food chain** is sometimes used metaphorically to describe human social situations. It is not surprising that in our competitive society, individuals who are considered successful are seen as being at the top of the food chain, consuming all others for their benefit, whereas the less successful are seen as being at the bottom.

The scientific understanding of a food chain is more precise than in its everyday usage. In ecology, a food chain is

a linear sequence of organisms through which nutrients and energy pass: primary producers, primary consumers, and higher-level consumers are used to describe ecosystem structure and dynamics. There is a single path through the chain. Each organism in a food chain occupies what is called a **trophic level**. Depending on their role as producers or consumers, species or groups of species can be assigned to various trophic levels.

In many ecosystems, the bottom of the food chain consists of photosynthetic organisms (plants and/or phytoplankton), which are called **primary producers**. The organisms that consume the primary producers are **herbivores**: the **primary consumers**. **Secondary consumers** are usually **carnivores** that eat the primary consumers. **Tertiary consumers** are carnivores that eat other carnivores. Higher-level consumers feed on the next lower trophic levels, and so on, up to the organisms at the top of the food chain: the **apex consumers**. In the Lake Ontario food chain shown in Figure 3, the Chinook salmon is the apex consumer at the top of this food chain.

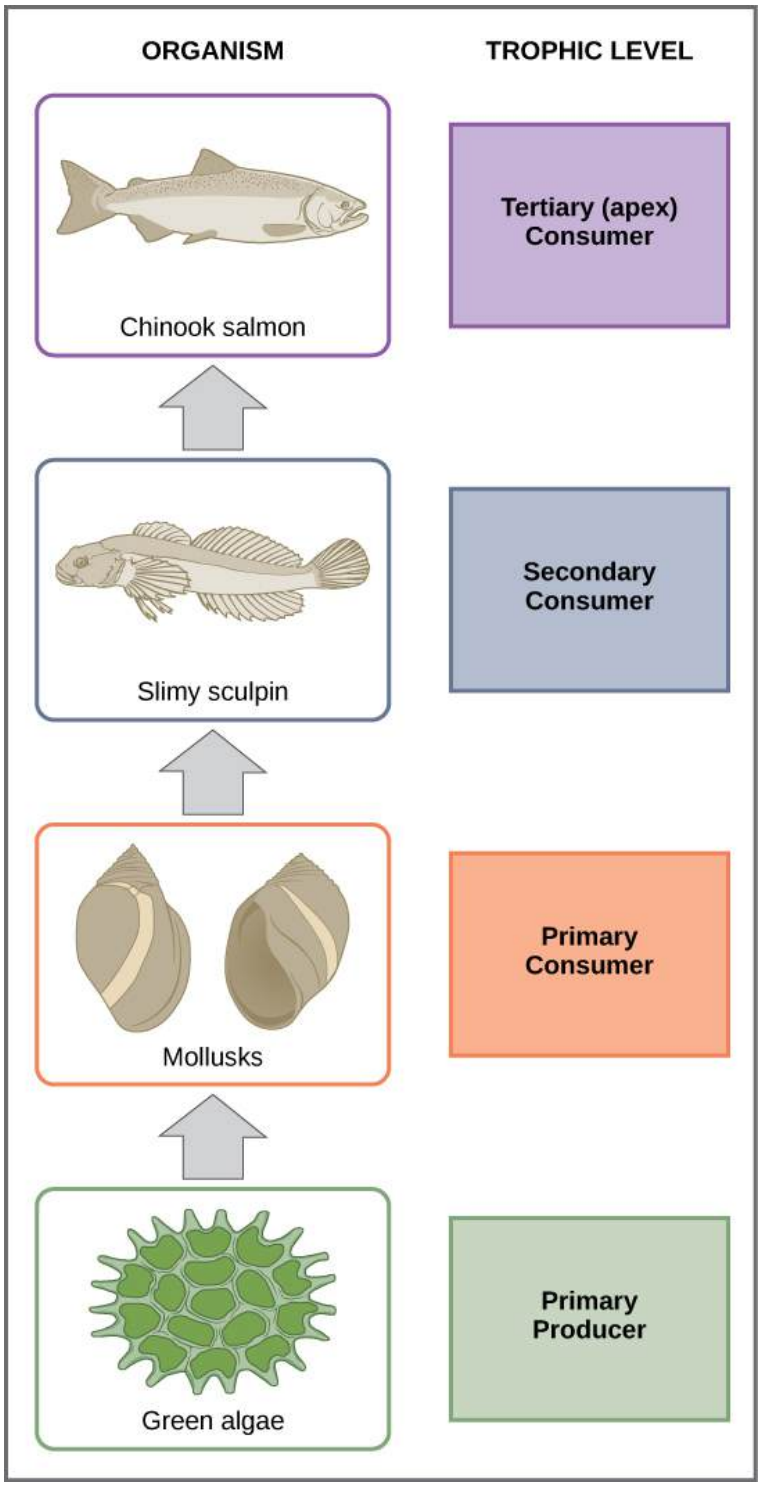

ORGANISM
TROPHIC LEVEL
Chinook salmon
Tertiary (apex) Consumer
Slimy sculpin
Secondary Consumer
Mollusks
Primary Consumer
Green algae
Primary Producer

Figure 3: These are the trophic levels of a food chain in Lake Ontario at the United States-Canada border. Energy and nutrients flow from photosynthetic green algae at the bottom to the top of the food chain: the Chinook salmon. (credit: "food chain in Lake Ontario" by OpenStax is licensed under CC BY 4.0)

One major factor that limits the length of food chains is energy. Energy is lost as **heat** between each trophic level due to the second law of thermodynamics. Thus, after a limited number of trophic energy transfers, the amount of energy remaining in the food chain may not be great enough to support viable populations at yet a higher trophic level.

The loss of energy between trophic levels is illustrated in the 1940s by the pioneering studies of Howard T. Odum of the Silver Springs ecosystem located in Florida (Figure 4). The primary producers generated 20,819 kcal/m^2/yr (kilocalories per square meter per year), the primary consumers generated 3368 kcal/m^2/yr, the secondary consumers generated 383 kcal/m^2/yr, and the tertiary consumers only generated 21 kcal/m^2/yr. Thus, there is little energy remaining for another level of consumers in this ecosystem.

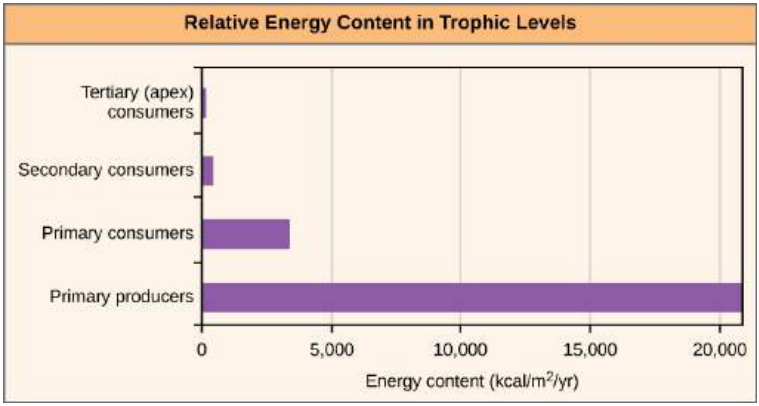


Figure 4: The relative energy in trophic levels in a Silver Springs, Florida, ecosystem is shown. Each trophic level has less energy available and supports fewer organisms at the next level. (credit: "trophic levels Silver Springs" by OpenStax is licensed under CC BY 4.0)

There is one problem when using food chains to accurately describe most ecosystems. Even when all organisms are grouped into appropriate trophic levels, some of these organisms can feed on species from more than one trophic level; likewise, some of these organisms can be eaten by species from multiple trophic levels. In other words, the linear model of ecosystems, the food chain, is not completely descriptive of ecosystem structure. A holistic model—which accounts for all the interactions between different species and their complex interconnected relationships with each other and with the environment—is a more accurate and descriptive model for ecosystems. A **food web** is a graphic representation of a holistic, non-linear web of primary producers, primary consumers, and higher-level consumers used to describe ecosystem structure and dynamics (Figure 5).

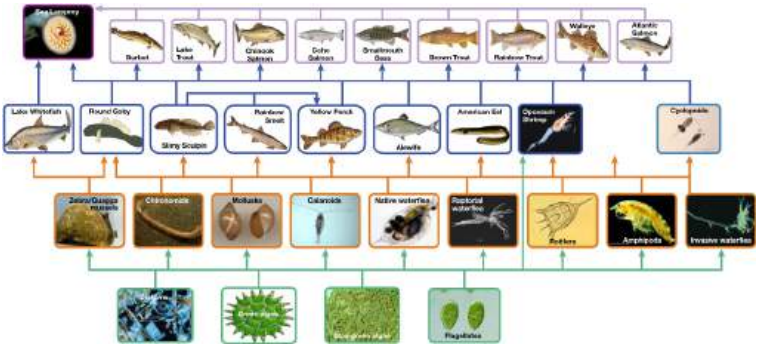


Figure 5: This food web shows the interactions between organisms across trophic levels in the Lake Ontario ecosystem. Primary producers are outlined in green, primary consumers in orange, secondary consumers in blue, and tertiary (apex) consumers in purple. Arrows point from an organism that is consumed to the organism that consumes it. Notice how some lines point to more than one trophic level. For example, the opossum shrimp eats both primary producers and primary consumers. (credit: NOAA, GLERL. "food web" by OpenStax is licensed under CC BY 4.0)

A comparison of the two types of structural ecosystem models shows strength in both. Food chains are more flexible for analytical modeling, are easier to follow, and are easier to experiment with, whereas food web models more accurately represent ecosystem structure and dynamics, and data can be directly used as input for simulation modeling.

Two general types of food webs are often shown interacting within a single ecosystem. A **grazing food web** has plants or other photosynthetic organisms at its base, followed by herbivores and various carnivores. A **detrital food web** consists of a base of organisms that feed on decaying organic matter (dead organisms), called **decomposers** or **detritivores**. These organisms are usually bacteria or fungi that recycle organic material back into the biotic part of the ecosystem as they themselves are

consumed by other organisms. As all ecosystems require a method to recycle material from dead organisms, most grazing food webs have an associated detrital food web. For example, in a meadow ecosystem, plants may support a grazing food web of different organisms, primary and other levels of consumers, while at the same time supporting a detrital food web of bacteria, fungi, and detrivorous invertebrates feeding off dead plants and animals.

Evolution Connection – Three-spined Stickleback

It is well established by the theory of natural selection that changes in the environment play a major role in the evolution of species within an ecosystem. However, little is known about how the evolution of species within an ecosystem can alter the ecosystem environment. In 2009, Dr. Luke Harmon, from the University of Idaho in Moscow, published a paper that for the first time showed that the evolution of organisms into subspecies can have direct effects on their ecosystem environment (Harmon, 2009).

The three-spines stickleback (*Gasterosteus aculeatus*) is a freshwater fish that evolved from a saltwater fish to live in freshwater lakes about 10,000 years ago, which is considered a recent development in evolutionary time (Figure 6). Over the last 10,000 years, these freshwater fish then became isolated from each other in different lakes. Depending on which lake population was studied, findings showed that these sticklebacks then either remained as one species or evolved into two species. The divergence of species was made possible by their use of different areas of the pond for feeding called micro niches.

Dr. Harmon and his team created artificial pond microcosms in 250-gallon tanks and added muck from freshwater ponds as a source of zooplankton and other invertebrates to sustain the fish. In different experimental

tanks, they introduced one species of stickleback from either a single-species or double-species lake.

Over time, the team observed that some of the tanks bloomed with algae while others did not. This puzzled the scientists, and they decided to measure the water's dissolved organic carbon (DOC), which consists of mostly large molecules of decaying organic matter that give pond-water its slightly brownish color. It turned out that the water from the tanks with two-species fish contained larger particles of DOC (and hence darker water) than water with single-species fish. This increase in DOC blocked the sunlight and prevented algal blooming. Conversely, the water from the single-species tank contained smaller DOC particles, allowing more sunlight penetration to fuel the algal blooms.

This change in the environment, which is due to the different feeding habits of the stickleback species in each lake type, probably has a great impact on the survival of other species in these ecosystems, especially other photosynthetic organisms. Thus, the study shows that, at least in these ecosystems, the environment and the evolution of populations have reciprocal effects that may now be factored into simulation models.

Figure 6: The three-spined stickleback evolved from a saltwater fish to freshwater fish. (credit: Barrett Paul, USFWS. "three-spined stickleback" by OpenStax is licensed under CC BY 4.0)

Research into Ecosystem Dynamics: Ecosystem Experimentation and Modeling

The study of the changes in ecosystem structure caused by changes in the environment (disturbances) or by internal forces is called ecosystem dynamics. Ecosystems are characterized using a variety of research methodologies. Some ecologists study ecosystems using controlled experimental systems, while some study entire ecosystems in their natural state, and others use both approaches.

A **holistic ecosystem model** attempts to quantify the composition, interaction, and dynamics of entire ecosystems; it is the most representative of the ecosystem in its natural state. A food web is an example of a holistic ecosystem model. However, this type of study is limited by time and expense, as well as the fact that it is neither feasible nor ethical to do experiments on large natural ecosystems. To

quantify all different species in an ecosystem and the dynamics in their habitat is difficult, especially when studying large habitats such as the Amazon Rainforest, which covers 1.4 billion acres (5.5 million km^2) of the Earth's surface.

For these reasons, scientists study ecosystems under more controlled conditions. **Experimental systems** usually involve either partitioning a part of a natural ecosystem that can be used for experiments, termed a **mesocosm**, or by re-creating an ecosystem entirely in an indoor or outdoor laboratory environment, which is referred to as a **microcosm**. A major limitation to these approaches is that removing individual organisms from their natural ecosystem or altering a natural ecosystem through partitioning may change the dynamics of the ecosystem. These changes are often due to differences in species numbers and diversity and also to environmental alterations caused by partitioning (mesocosm) or re-creating (microcosm) the natural habitat. Thus, these types of experiments are not totally predictive of changes that would occur in the ecosystem from which they were gathered.

As both of these approaches have their limitations, some ecologists suggest that results from these experimental systems should be used only in conjunction with holistic ecosystem studies to obtain the most representative data about ecosystem structure, function, and dynamics.

Scientists use the data generated by these experimental studies to develop **ecosystem models** that demonstrate the structure and dynamics of ecosystems. Three basic types of ecosystem modeling are routinely used in research and ecosystem management: a **conceptual model**, an **analytical model**, and a **simulation model**. A conceptual model is an ecosystem model that consists of flow charts to show interactions of different compartments of the living and nonliving components of the ecosystem. A conceptual model describes ecosystem structure and dynamics and shows how

environmental disturbances affect the ecosystem; however, its ability to predict the effects of these disturbances is limited. Analytical and simulation models, in contrast, are mathematical methods of describing ecosystems that are indeed capable of predicting the effects of potential environmental changes without direct experimentation, although with some limitations as to accuracy. An analytical model is an ecosystem model that is created using simple mathematical formulas to predict the effects of environmental disturbances on ecosystem structure and dynamics. A simulation model is an ecosystem model that is created using complex computer algorithms to holistically model ecosystems and to predict the effects of environmental disturbances on ecosystem structure and dynamics. Ideally, these models are accurate enough to determine which components of the ecosystem are particularly sensitive to disturbances, and they can serve as a guide to ecosystem managers (such as conservation ecologists or fisheries biologists) in the practical maintenance of ecosystem health.

Conceptual Models

Conceptual models are useful for describing ecosystem structure and dynamics and for demonstrating the relationships between different organisms in a community and their environment. Conceptual models are usually depicted graphically as **flow charts**. The organisms and their resources are grouped into specific compartments with arrows showing the relationship and transfer of energy or nutrients between them. Thus, these diagrams are sometimes called compartment models.

To model the cycling of mineral nutrients, organic and inorganic nutrients are subdivided into those that are bioavailable (ready to be incorporated into biological

macromolecules) and those that are not. For example, in a terrestrial ecosystem near a deposit of coal, carbon will be available to the plants of this ecosystem as carbon dioxide gas in a short-term period, not from the carbon-rich coal itself. However, over a longer period, microorganisms capable of digesting coal will incorporate its carbon or release it as natural gas (methane, CH_4), changing this unavailable organic source into an available one. This conversion is greatly accelerated by the combustion of fossil fuels by humans, which releases large amounts of carbon dioxide into the atmosphere. This is thought to be a major factor in the rise of atmospheric carbon dioxide levels in the industrial age. The carbon dioxide released from burning fossil fuels is produced faster than photosynthetic organisms can use it. This process is intensified by the reduction of photosynthetic trees because of worldwide deforestation. Most scientists agree that high atmospheric carbon dioxide is a major cause of global climate change.

Conceptual models are also used to show the flow of energy through particular ecosystems. Figure 7 is based on Howard T. Odum's classical study of Silver Springs, Florida, in the mid-twentieth century (Odum. 1957). This study shows the energy content and transfer between various ecosystem compartments.

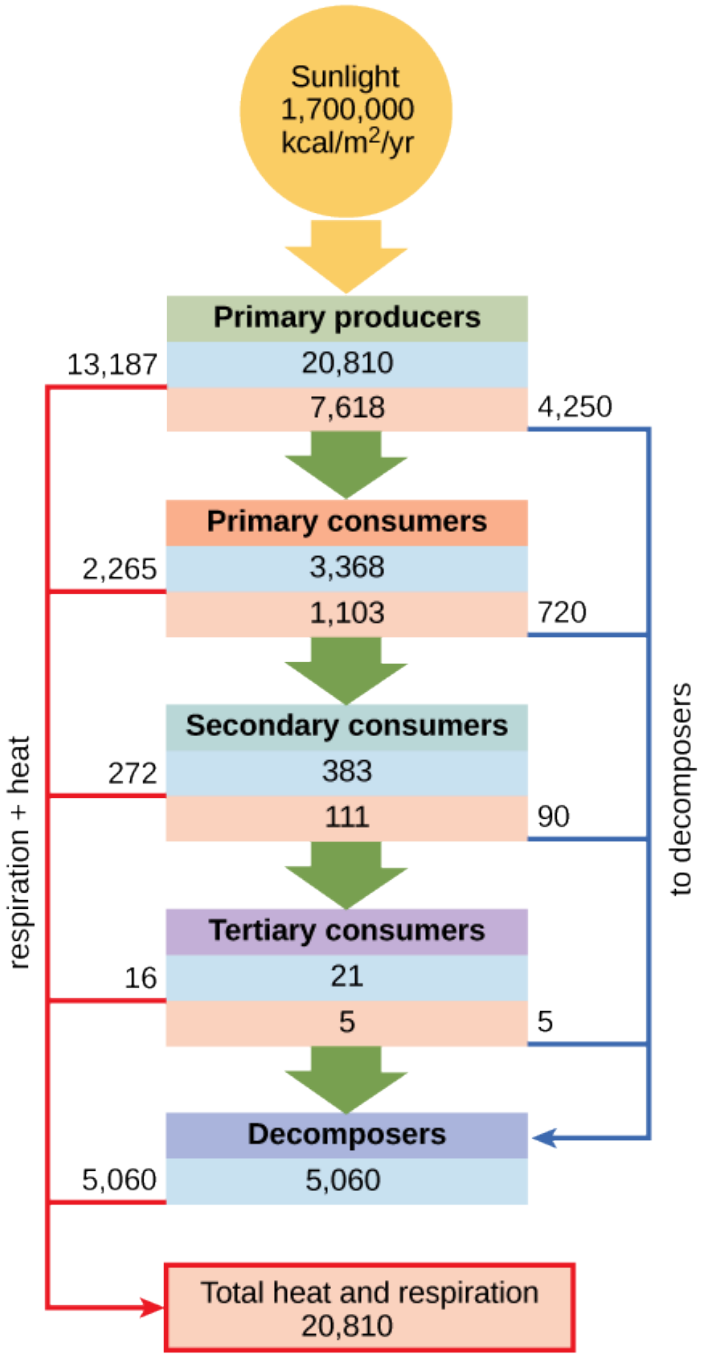

Sunlight
1,700,000
kcal/m2/yr
Primary producers
20,810
7,618
13,187
4,250
Primary consumers
3,368
1,103
2,265
720
Secondary consumers
383
111
272
90
Tertiary consumers
21
5
16
5
Decomposers
5,060
5,060
Total heat and respiration
20,810
respiration + heat
to decomposers
Gross productivity
Net productivity

Figure 7: This conceptual model shows the flow of energy through a spring ecosystem in Silver Springs, Florida. Notice that the energy decreases with each increase in trophic level. (credit: "energy through a spring ecosystem" by OpenStax is licensed under CC BY 4.0)

Analytical and Simulation Models

The major limitation of conceptual models is their inability to predict the consequences of changes in ecosystem species and/or environment. Ecosystems are dynamic entities and subject to a variety of abiotic and biotic disturbances caused by natural forces and/or human activity. Ecosystems altered from their initial equilibrium state can often recover from such disturbances and return to a state of equilibrium. As most ecosystems are subject to periodic disturbances and are often in a state of change, they are usually either moving toward or away from their **equilibrium state**. There are many of these equilibrium states among the various components of an ecosystem, which affects the ecosystem overall. Furthermore, as humans have the ability to greatly and rapidly alter the species content and habitat of an ecosystem, the need for predictive models that enable understanding of how ecosystems respond to these changes becomes more crucial.

Analytical models often use simple, linear components of ecosystems, such as food chains, and are known to be complex mathematically; therefore, they require a significant amount of mathematical knowledge and expertise. Although analytical models have great potential, their simplification of complex ecosystems is thought to limit their accuracy. **Simulation models** that use computer programs are better able to deal with the complexities of ecosystem structure.

A recent development in simulation modeling uses

supercomputers to create and run individual-based simulations, which account for the behavior of individual organisms and their effects on the ecosystem as a whole. These simulations are considered to be the most accurate and predictive of the complex responses of ecosystems to disturbances.

Summary

Ecosystems exist on land, at sea, in the air, and underground. Different ways of modeling ecosystems are necessary to understand how environmental disturbances will affect ecosystem structure and dynamics. Conceptual models are useful to show the general relationships between organisms and the flow of materials or energy between them. Analytical models are used to describe linear food chains, and simulation models work best with holistic food webs.

REFERENCES

Harmon, L. 2009. *Nature.* (Vol. 458, April)

Odum, H. 1957. "Trophic Structure and Productivity of Silver Springs, Florida," *Ecological Monographs* 27, no. 1: 47–112.

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Energy Flow through Ecosystems

All living things require energy in one form or another. Energy is required by most complex metabolic pathways (often in the form of adenosine triphosphate, ATP), especially those responsible for building large molecules from smaller compounds, and life itself is an energy-driven process. Living organisms would not be able to assemble macromolecules (proteins, lipids, nucleic acids, and complex carbohydrates) from their monomeric subunits without a constant energy input.

It is important to understand how organisms acquire energy and how that energy is passed from one organism to another through food webs and their constituent food chains. Food webs illustrate how energy flows directionally through ecosystems, including how efficiently organisms acquire it, use it, and how much remains for use by other organisms of the food web.

How Organisms Acquire Energy in a Food Web

Energy is acquired by living things in three ways: photosynthesis, chemosynthesis, and the consumption and digestion of other living or previously living organisms by heterotrophs.

Photosynthetic and **chemosynthetic** organisms are both grouped into a category known as **autotrophs**: organisms capable of synthesizing their own food (more specifically, capable of using inorganic carbon as a carbon source). Photosynthetic autotrophs (**photoautotrophs**) use sunlight as an energy source, whereas chemosynthetic autotrophs (**chemoautotrophs**) use inorganic molecules as an energy source. Autotrophs are critical for all ecosystems. Without these organisms, energy would not be available to other living organisms and life itself would not be possible.

Photoautotrophs, such as plants, algae, and photosynthetic bacteria, serve as the energy source for a majority of the world's ecosystems. These ecosystems are often described by grazing food webs. Photoautotrophs harness the solar energy of the sun by converting it to chemical energy in the form of ATP (and NADP). The energy stored in ATP is used to synthesize complex organic molecules, such as glucose.

Chemoautotrophs are primarily bacteria that are found in rare ecosystems where sunlight is not available, such as in those associated with dark caves or hydrothermal vents at the bottom of the ocean (Figure 1). Many chemoautotrophs in hydrothermal vents use hydrogen sulfide (H_2S), which is released from the vents as a source of chemical energy. This allows chemoautotrophs to synthesize complex organic molecules, such as glucose, for their own energy and in turn supplies energy to the rest of the ecosystem.

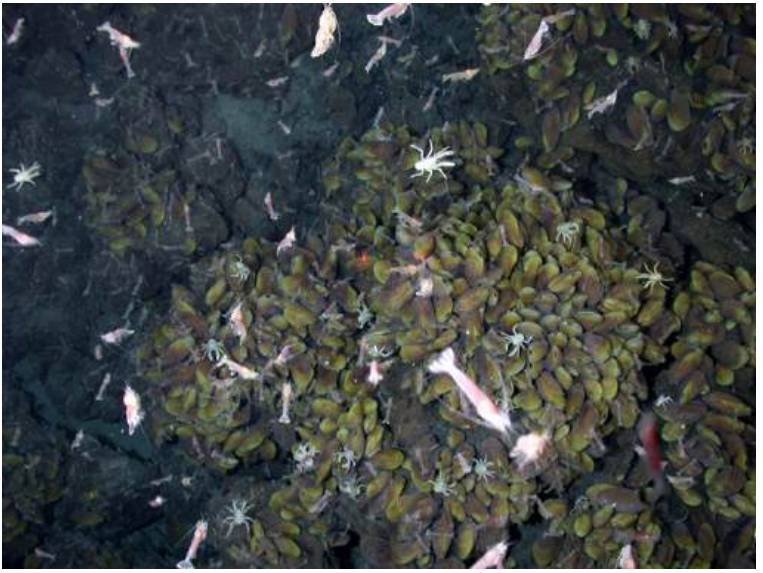


Figure 1: Swimming shrimp, a few squat lobsters, and hundreds of vent mussels are seen at a hydrothermal vent at the bottom of the ocean. As no sunlight penetrates to this depth, the ecosystem is supported by chemoautotrophic bacteria and organic material that sinks from the ocean's surface. This picture was taken in 2006 at the submerged NW Eifuku volcano off the coast of Japan by the National Oceanic and Atmospheric Administration (NOAA). The summit of this highly active volcano lies 1535 m below the surface. (credit: NOAA. "Chemoautotrophs" by OpenStax is licensed under CC BY 4.0)

Productivity within Trophic Levels

Productivity within an ecosystem can be defined as the percentage of energy entering the ecosystem incorporated into biomass in a particular trophic level. **Biomass** is the total mass, in a unit area at the time of measurement, of living or previously living organisms within a trophic level. Ecosystems have characteristic amounts of biomass at each trophic level.

For example, in the English Channel ecosystem, the primary producers account for a biomass of 4 g/m^2 (grams per meter squared), while the primary consumers exhibit a biomass of 21 g/m^2.

The productivity of the primary producers is especially important in any ecosystem because these organisms bring energy to other living organisms by photoautotrophy or chemoautotrophy. The rate at which photosynthetic primary producers incorporate energy from the sun is called **gross primary productivity**. An example of gross primary productivity is shown in the compartment diagram of energy flow within the Silver Springs aquatic ecosystem as shown (Figure 2). In this ecosystem, the total energy accumulated by the primary producers (gross primary productivity) was shown to be 20,810 $kcal/m^2/yr$.

Because all organisms need to use some of this energy for their own functions (like respiration and resulting metabolic heat loss) scientists often refer to the net primary productivity of an ecosystem. **Net primary productivity** is the energy that remains in the primary producers after accounting for the organisms' respiration and heat loss. The net productivity is then available to the primary consumers at the next trophic level. In our Silver Spring example, 13,187 of the 20,810 $kcal/m^2/yr$ were used for respiration or were lost as heat, leaving 7,632 $kcal/m^2/yr$ of energy for use by the primary consumers.

Ecological Efficiency: The Transfer of Energy between Trophic Levels

As illustrated in Figure 2, large amounts of energy are lost from the ecosystem from one trophic level to the next level as energy flows from the primary producers through the various trophic levels of consumers and decomposers.

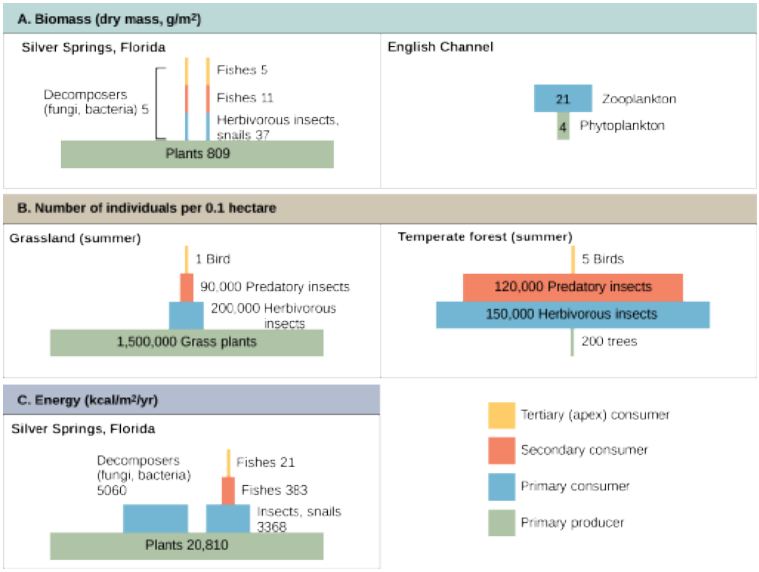


Figure 2: This conceptual model shows the flow of energy through ecosystems. Notice that the energy decreases with each increase in trophic level. Part A: on the left is a pyramid diagram of the number of individuals per 0.1 hectare in a summer grassland. There are 1,500,000 grass plants, 200,000 herbivorous insects, 90,000 predatory insects, and 1 bird. Part A: on the right is a pyramid diagram of organisms per 0.1 hectare in a temperate forest. There are 200 trees, 150,000 herbivorous insects, 120,000 predatory insects, and 5 birds. Part B: on the left is a pyramid diagram of dry biomass in grams per meter squared in the English Channel. The biomass is 4 phytoplankton and 21 zooplankton. Part B: on the right is a pyramid diagram of dry biomass in grams per meter squared in Silver Springs, Florida. The biomass of plants is 809. The biomass of primary consumers, including herbivorous insects and snails is 37. The biomass of secondary consumer fishes is 11, and the biomass of tertiary consumer fishes is 5. Primary, secondary and tertiary decomposers have a combined biomass of 5. Part C is a pyramid diagram of energy in kilocalories per meter squared per year. The energy of plants is 20,810. The energy of primary consumers, including insects and snails, is 3,368. The energy of primary consumer fishes is 383, and the energy of secondary consumer

fishes is 21. The energy of decomposers, including fungi and bacteria, is 5,060. (credit: "energy through a spring ecosystem" by OpenStax is licensed under CC BY 4.0)

The main reason for this loss is **the second law of thermodynamics**, which states that whenever energy is converted from one form to another, there is a tendency toward disorder (**entropy**) in the system. In biologic systems, this means a great deal of energy is lost as metabolic heat when the organisms from one trophic level consume the next level. In the Silver Springs ecosystem example, we see that the primary consumers produced 1103 kcal/m^2/yr from the 7618 kcal/m^2/yr of energy available to them from the primary producers. The measurement of energy transfer efficiency between two successive trophic levels is termed the **trophic level transfer efficiency** (**TLTE**) and is defined by the formula:

TLTE = production at present trophic levelproduction at previous trophic level × 100TLTE = production at present trophic levelproduction at previous trophic level × 100">TLTE = production at present trophic level / production at previous trophic level × 100

In Silver Springs, the TLTE between the first two trophic levels was approximately 14.8 percent. The low efficiency of energy transfer between trophic levels is usually the major factor that limits the length of food chains observed in a food web. The fact is, after four to six energy transfers, there is not enough energy left to support another trophic level. In the Lake Ontario example, only three energy transfers occurred between the primary producer, (green algae), and the apex consumer (Chinook salmon).

Ecologists have many different methods of measuring energy transfers within ecosystems. Some transfers are easier or more difficult to measure depending on the

complexity of the ecosystem and how much access scientists have to observe the ecosystem. In other words, some ecosystems are more difficult to study than others, and sometimes the quantification of energy transfers has to be estimated.

Another main parameter that is important in characterizing energy flow within an ecosystem is the net production efficiency. **Net production efficiency** (**NPE**) allows ecologists to quantify how efficiently organisms of a particular trophic level incorporate the energy they receive into biomass; it is calculated using the following formula:
NPE = net consumer productivityassimilation × 100NPE&nbs p;= net consumer productivityassimilatio n × 100">NPE = net consumer productivity / assimilation × 100
Net consumer productivity is the energy content available to the organisms of the next trophic level. **Assimilation** is the biomass (energy content generated per unit area) of the present trophic level after accounting for the energy lost due to incomplete ingestion of food, energy used for respiration, and energy lost as waste. I**ncomplete ingestion** refers to the fact that some consumers eat only a part of their food. For example, when a lion kills an antelope, it will eat everything except the hide and bones. The lion is missing the energy-rich bone marrow inside the bone, so the lion does not make use of all the calories its prey could provide.

Thus, NPE measures how efficiently each trophic level uses and incorporates the energy from its food into biomass to fuel the next trophic level. In general, cold-blooded animals (**ectotherms**), such as invertebrates, fish, amphibians, and reptiles, use less of the energy they obtain for respiration and heat than warm-blooded animals (**endotherms**), such as birds and mammals. The extra heat generated in endotherms, although an advantage in terms of the activity of these organisms in colder environments, is a major

disadvantage in terms of NPE. Therefore, many endotherms have to eat more often than ectotherms to get the energy they need for survival. In general, NPE for ectotherms is an order of magnitude (10x) higher than for endotherms. For example, the NPE for a caterpillar eating leaves has been measured at 18 percent, whereas the NPE for a squirrel eating acorns may be as low as 1.6 percent.

The **inefficiency of energy** use by warm-blooded animals has broad implications for the world's food supply. It is widely accepted that the meat industry uses large amounts of crops to feed livestock, and because the NPE is low, much of the energy from animal feed is lost. For example, it costs about 1¢ to produce 1000 dietary calories (kcal) of corn or soybeans, but approximately $0.19 to produce a similar number of calories growing cattle for beef consumption. The same energy content of milk from cattle is also costly, at approximately $0.16 per 1000 kcal. Much of this difference is due to the low NPE of cattle. Thus, there has been a growing movement worldwide to promote the consumption of non-meat and non-dairy foods so that less energy is wasted feeding animals for the meat industry.

Modeling Ecosystems Energy Flow: Ecological Pyramids

The structure of ecosystems can be visualized with ecological pyramids, which were first described by the pioneering studies of Charles Elton in the 1920s. **Ecological pyramids** show the relative amounts of various parameters (such as the number of organisms, energy, and biomass) across trophic levels.

Pyramids of numbers can be either upright or inverted, depending on the ecosystem. As shown in Figure 2, typical grassland during the summer has a base of many plants and the numbers of organisms decrease at each trophic level. However, during the summer in a temperate forest, the base

of the pyramid consists of few trees compared with the number of primary consumers, mostly insects. Because trees are large, they have great photosynthetic capability and dominate other plants in this ecosystem to obtain sunlight. Even in smaller numbers, primary producers in forests are still capable of supporting other trophic levels.

Another way to visualize ecosystem structure is with pyramids of biomass. This pyramid measures the amount of energy converted into living tissue at the different trophic levels. Using the Silver Springs ecosystem example, this data exhibits an upright biomass pyramid (Figure 2), whereas the pyramid from the English Channel example is inverted. The plants (primary producers) of the Silver Springs ecosystem make up a large percentage of the biomass found there. However, the phytoplankton in the English Channel example make up less biomass than the primary consumers, the zooplankton. As with inverted pyramids of numbers, this inverted pyramid is not due to a lack of productivity from the primary producers but results from the high turnover rate of the phytoplankton. The phytoplankton are consumed rapidly by the primary consumers, thus, minimizing their biomass at any particular point in time. However, phytoplankton reproduce quickly, thus they are able to support the rest of the ecosystem.

Pyramid ecosystem modeling can also be used to show energy flow through the trophic levels. **Pyramids of energy** are always upright, and an ecosystem without sufficient primary productivity cannot be supported. All types of ecological pyramids are useful for characterizing ecosystem structure. However, in the study of energy flow through the ecosystem, pyramids of energy are the most consistent and representative models of ecosystem structure.

Consequences of Food Webs: Biological Magnification

One of the most important environmental consequences of ecosystem dynamics is **biomagnification**. Biomagnification is the increasing concentration of persistent, toxic substances in organisms at each trophic level, from the primary producers to the apex consumers. Many substances have been shown to **bioaccumulate**, including classical studies with the pesticide **dichlorodiphenyltrichloroethane** (**DDT**), which was published in the 1960s bestseller, *Silent Spring*, by Rachel Carson. DDT was a commonly used pesticide before its dangers became known. In some aquatic ecosystems, organisms from each trophic level consumed many organisms of the lower level, which caused DDT to increase in birds (apex consumers) that ate fish. Thus, birds accumulated sufficient amounts of DDT to cause fragility in their eggshells. This effect increased egg breakage during nesting and was shown to have adverse effects on these bird populations. The use of DDT was banned in the United States in the 1970s.

Other substances that biomagnify are **polychlorinated biphenyls** (**PCBs**), which were used in coolant liquids in the United States until their use was banned in 1979, and heavy metals, such as mercury, lead, and cadmium. These substances were best studied in aquatic ecosystems, where fish species at different trophic levels accumulate toxic substances brought through the ecosystem by the primary producers. As illustrated in a study performed by the National Oceanic and Atmospheric Administration (NOAA) in the Saginaw Bay of Lake Huron (Figure 3), PCB concentrations increased from the ecosystem's primary producers (phytoplankton) through the different trophic levels of fish species. The apex consumer (walleye) has more than four times the amount of PCBs compared to phytoplankton. Also, based on results from other studies,

birds that eat these fish may have PCB levels at least one order of magnitude higher than those found in the lake fish.

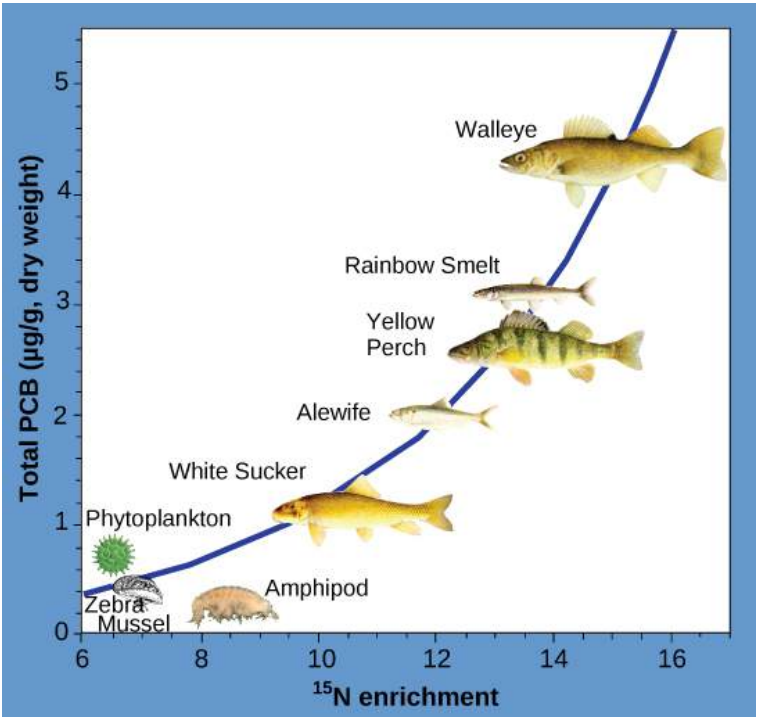


Figure 3: This chart shows the PCB concentrations found at the various trophic levels in the Saginaw Bay ecosystem of Lake Huron. Numbers on the x-axis reflect enrichment with heavy isotopes of nitrogen (15N), which is a marker for increasing trophic level. Notice that the fish in the higher trophic levels accumulate more PCBs than those in lower trophic levels. (credit: Patricia Van Hoof, NOAA, GLERL. "PCB concentrations" by OpenStax is licensed under CC BY 4.0)

Other concerns have been raised by the accumulation of **heavy metals**, such as mercury and cadmium, in certain types of seafood. The United States Environmental Protection Agency (EPA) recommends that pregnant women and young children should not consume any swordfish, shark, king mackerel, or tilefish because of their high

mercury content. These individuals are advised to eat fish low in mercury: salmon, tilapia, shrimp, pollock, and catfish. Biomagnification is a good example of how ecosystem dynamics can affect our everyday lives, even influencing the food we eat.

Summary

Organisms in an ecosystem acquire energy in a variety of ways, which is transferred between trophic levels as the energy flows from the bottom to the top of the food web, with energy being lost at each transfer. The efficiency of these transfers is important for understanding the different behaviors and eating habits of warm-blooded versus cold-blooded animals. Modeling of ecosystem energy is best done with ecological pyramids of energy, although other ecological pyramids provide other vital information about ecosystem structure.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Biogeochemical Cycles

Energy flows directionally through ecosystems, entering as sunlight (or inorganic molecules for chemoautotrophs) and leaving as heat during the many transfers between trophic levels. However, the matter that makes up living organisms is conserved and recycled. The six most common elements associated with organic molecules—carbon, nitrogen, hydrogen, oxygen, phosphorus, and sulfur—take a variety of chemical forms and may exist for long periods in the atmosphere, on land, in water, or beneath the Earth's surface. Geologic processes, such as weathering, erosion, water drainage, and the subduction of the continental plates, all play a role in this recycling of materials. Because geology and chemistry have major roles in the study of this process, the recycling of inorganic matter between living organisms and their environment is called a biogeochemical cycle.

Water contains hydrogen and oxygen, which is essential to all living processes. The hydrosphere is the area of the Earth where water movement and storage occurs: as liquid water on the surface and beneath the surface or frozen (rivers, lakes, oceans, groundwater, polar ice caps, and glaciers), and as water vapor in the atmosphere. Carbon is found in all organic macromolecules and is an important constituent of fossil fuels. Nitrogen is a major component of our nucleic acids and proteins and is critical to human agriculture. Phosphorus, a major component of nucleic acid (along with

nitrogen), is one of the main ingredients in artificial fertilizers used in agriculture and their associated environmental impacts on our surface water. Sulfur, critical to the 3–D folding of proteins (as in disulfide binding), is released into the atmosphere by the burning of fossil fuels, such as coal.

The cycling of these elements is interconnected. For example, the movement of water is critical for the leaching of nitrogen and phosphate into rivers, lakes, and oceans. Furthermore, the ocean itself is a major reservoir for carbon. Thus, mineral nutrients are cycled, either rapidly or slowly, through the entire biosphere, from one living organism to another, and between the biotic and abiotic world.

The Water (Hydrologic) Cycle

Water is the basis of all living processes. The human body is more than 1/2 water and human cells are more than 70 percent water. Thus, most land animals need a supply of fresh water to survive. However, when examining the stores of water on Earth, 97.5 percent of it is non-potable salt water (Figure 1). Of the remaining water, 99 percent is locked underground as water or as ice. Thus, less than 1 percent of fresh water is easily accessible from lakes and rivers. Many living things, such as plants, animals, and fungi, are dependent on the small amount of fresh surface water supply, a lack of which can have massive effects on ecosystem dynamics. Humans, of course, have developed technologies to increase water availability, such as digging wells to harvest groundwater, storing rainwater, and using desalination to obtain drinkable water from the ocean. Although this pursuit of drinkable water has been ongoing throughout human history, the supply of fresh water is still a major issue in modern times.

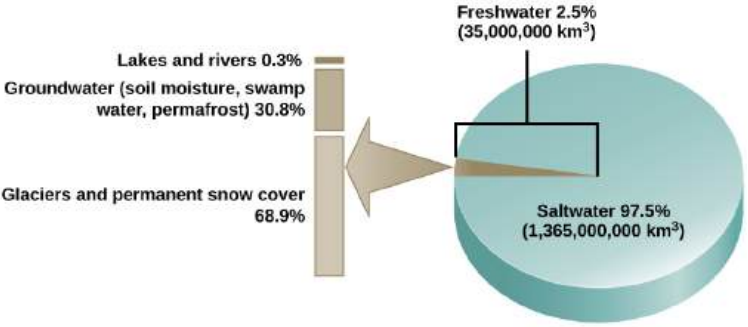


Figure 1: Only 2.5 percent of water on Earth is fresh water, and less than 1 percent of fresh water is easily accessible to living things. (credit: "fresh water" by OpenStax is licensed under CC BY 4.0)

Water cycling is extremely important to ecosystem dynamics. Water has a major influence on climate and, thus, on the environments of ecosystems, some located on distant parts of the Earth. Most of the water on Earth is stored for long periods in the **oceans**, **underground**, and as **ice.** Figure 2 illustrates the average time that an individual water molecule may spend in the Earth's major water reservoirs. Residence time is a measure of the average time an individual water molecule stays in a particular reservoir. A large amount of the Earth's water is locked in place in these reservoirs as ice, beneath the ground, and in the ocean, and, thus, is unavailable for short-term cycling (only surface water can evaporate**).**

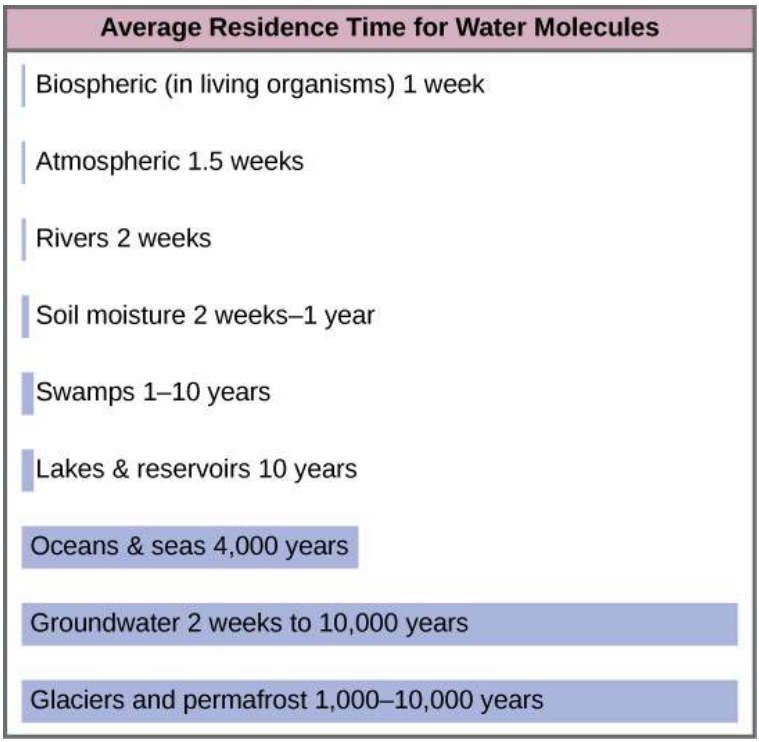


Figure 2: This graph shows the average residence time for water molecules in the Earth's water reservoirs. (credit: "time water molecule" by OpenStax is licensed under CC BY 4.0)

There are various processes that occur during the cycling of water, shown in Figure 3. These processes include the following:

- **evaporation/sublimation**
- **condensation/precipitation**
- **subsurface water flow**
- **surface runoff/snowmelt**
- **streamflow**

The water cycle is driven by the sun's energy as it warms the oceans and other surface waters. This leads to the **evaporation** (water to water vapor) of liquid surface water and the **sublimation** (ice to water vapor) of frozen water, which deposits large amounts of water vapor into the atmosphere. Over time, this water vapor condenses into clouds as liquid or frozen droplets and is eventually followed by **precipitation** (rain or snow), which returns water to the Earth's surface. Rain eventually permeates into the ground, where it may evaporate again if it is near the surface, flow beneath the surface, or be stored for long periods. More easily observed is **surface runoff**: the flow of fresh water either from rain or melting ice. Runoff can then make its way through streams and lakes to the oceans or flow directly to the oceans themselves.

Rain and surface runoff are major ways in which minerals, including carbon, nitrogen, phosphorus, and sulfur, are cycled from land to water. The environmental effects of runoff will be discussed later as these cycles are described.

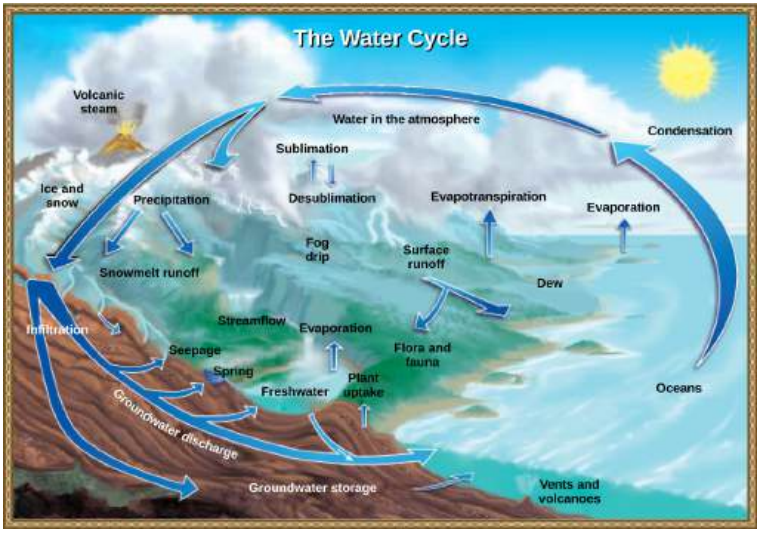


Figure 3: Water from the land and oceans enters the atmosphere by evaporation or sublimation, where it condenses into clouds and falls as rain or snow. Precipitated water may enter freshwater bodies or infiltrate the soil. The cycle is complete when surface or groundwater reenters the ocean. (credit: modification of work by John M. Evans and Howard Perlman, USGS. "the water cycle" by OpenStax is licensed under CC BY 4.0)

The Carbon Cycle

Carbon is the second most abundant element in living organisms. Carbon is present in all organic molecules, and its role in the structure of macromolecules is of primary importance to living organisms. Carbon compounds contain especially high energy, particularly those derived from fossilized organisms, mainly plants, which humans use as fuel. Since the 1800s, the number of countries using massive amounts of fossil fuels has increased. Since the beginning of the Industrial Revolution, global demand for the Earth's

limited fossil fuel supplies has risen; therefore, the amount of carbon dioxide in our atmosphere has increased. This increase in carbon dioxide has been associated with climate change and other disturbances of the Earth's ecosystems and is a major environmental concern worldwide. Thus, the "carbon footprint" is based on how much carbon dioxide is produced and how much fossil fuel countries consume.

The **carbon cycle** is most easily studied as two interconnected sub-cycles: one dealing with rapid carbon exchange among living organisms and the other dealing with the long-term cycling of carbon through geologic processes. The entire carbon cycle is shown in Figure 4.

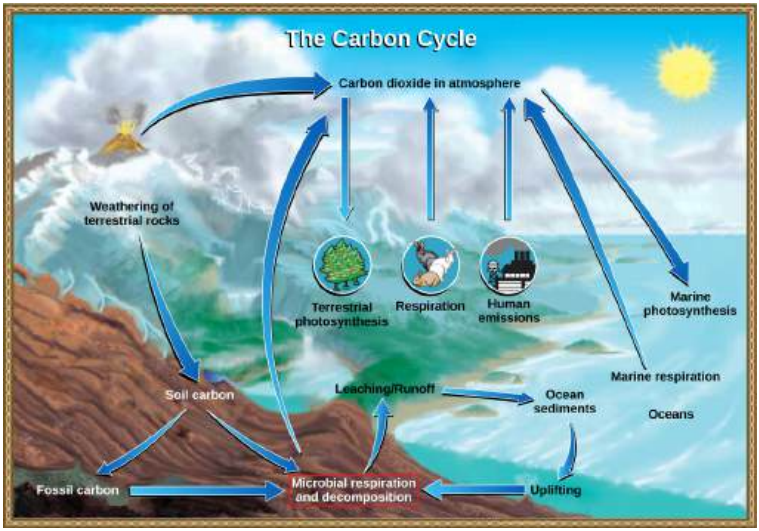


Figure 4: Carbon dioxide gas exists in the atmosphere and is dissolved in water. Photosynthesis converts carbon dioxide gas to organic carbon, and respiration cycles the organic carbon back into carbon dioxide gas. Long-term storage of organic carbon occurs when matter from living organisms is buried deep underground and becomes fossilized. Volcanic activity and, more recently, human emissions, bring this stored carbon back into the carbon cycle. (credit: modification of work by John M. Evans and Howard Perlman, USGS. "the carbon cycle" by OpenStax is licensed under CC BY 4.0)

The Biological Carbon Cycle

Living organisms are connected in many ways, even between ecosystems. A good example of this connection is the exchange of carbon between autotrophs and heterotrophs within and between ecosystems by way of atmospheric **carbon dioxide**. Carbon dioxide is the basic building block that most autotrophs use to build multi-carbon, high-energy compounds, such as glucose. The energy harnessed from the sun is used by these organisms to form the covalent bonds

that link carbon atoms together. These chemical bonds thereby store this energy for later use in the process of respiration. Most terrestrial autotrophs obtain their carbon dioxide directly from the atmosphere, while marine autotrophs acquire it in the dissolved form (carbonic acid, $H_2CO_3^-$). However carbon dioxide is acquired, a by-product of the process is oxygen. The photosynthetic organisms are responsible for depositing approximately 21 percent oxygen content of the atmosphere that we observe today.

Heterotrophs and autotrophs are partners in biological carbon exchange (especially the primary consumers, largely herbivores). Heterotrophs acquire the high-energy carbon compounds from the autotrophs by consuming them and breaking them down by respiration to obtain cellular energy, such as ATP. The most efficient type of respiration, aerobic respiration, requires oxygen obtained from the atmosphere or dissolved in water. Thus, there is a constant exchange of oxygen and carbon dioxide between the autotrophs (which need the carbon) and the heterotrophs (which need the oxygen). Gas exchange through the atmosphere and water is one way that the carbon cycle connects all living organisms on Earth.

The Biogeochemical Carbon Cycle

The movement of carbon through the land, water, and air is complex, and in many cases, it occurs much more slowly geologically than as seen between living organisms. Carbon is stored for long periods in what are known as **carbon reservoirs**, which include the atmosphere, bodies of liquid water (mostly oceans), ocean sediment, soil, land sediments (including fossil fuels), and the Earth's interior.

As stated, the atmosphere is a major reservoir of carbon in the form of carbon dioxide and is essential to the process of photosynthesis. The level of carbon dioxide in the

atmosphere is greatly influenced by the reservoir of carbon in the oceans. The exchange of carbon between the atmosphere and water reservoirs influences how much carbon is found in each location, and each one affects the other reciprocally. Carbon dioxide (CO_2) from the atmosphere dissolves in water and combines with water molecules to form **carbonic acid**, and then it ionizes to **carbonate** and **bicarbonate ions** (Figure 5)

Step 1: CO_2 (atmospheric) $\rightleftharpoons$ CO_2 (dissolved)
Step 2: CO_2 (dissolved) + H_2O $\rightleftharpoons$ H_2CO_3 (carbonic acid)
Step 3: H_2CO_3 $\rightleftharpoons$ + H^+ + HCO_3^- (biocarbonate ion)
Step 4: HCO_3^- $\rightleftharpoons$ H^+ + CO_3^{2-} (carbonate ion)

Figure 5: Carbon dioxide reacts with water to form bicarbonate and carbonate ions. (credit: "Biogeochemical Carbon Cycle" by OpenStax is licensed under CC BY 4.0)

The equilibrium coefficients are such that more than 90 percent of the carbon in the ocean is found as bicarbonate ions. Some of these ions combine with seawater calcium to form **calcium carbonate** (**$CaCO_3$**), a major component of marine organism shells. These organisms eventually form sediments on the ocean floor. Over geologic time, the calcium carbonate forms **limestone**, which comprises the largest carbon reservoir on Earth.

On land, carbon is stored in soil as a result of the decomposition of living organisms (by decomposers) or from weathering of terrestrial rock and minerals. This carbon can be leached into the water reservoirs by surface runoff. Deeper underground, on land and at sea, are fossil fuels: the anaerobically decomposed remains of plants that take millions of years to form. **Fossil fuels** are considered a non-

renewable resource because their use far exceeds their rate of formation. A **non-renewable resource**, such as fossil fuel, is either regenerated very slowly or not at all. Another way for carbon to enter the atmosphere is from land (including land beneath the surface of the ocean) by the eruption of volcanoes and other geothermal systems. Carbon sediments from the ocean floor are taken deep within the Earth by the process of subduction: the movement of one tectonic plate beneath another. Carbon is released as carbon dioxide when a volcano erupts or from volcanic hydrothermal vents.

Carbon dioxide is also added to the atmosphere by the animal husbandry practices of humans. The large numbers of land animals raised to feed the Earth's growing population results in increased carbon dioxide levels in the atmosphere due to farming practices, respiration, and methane production. This is another example of how human activity indirectly affects biogeochemical cycles in a significant way. Although much of the debate about the future effects of increasing atmospheric carbon on climate change focus on fossil fuels, scientists take natural processes, such as volcanoes and respiration, into account as they model and predict the future impact of this increase.

The Nitrogen Cycle

Getting nitrogen into the living world is difficult. Plants and phytoplankton are not equipped to incorporate nitrogen from the atmosphere (which exists as tightly bonded, triple covalent N_2) even though this molecule comprises approximately 78 percent of the atmosphere. Nitrogen enters the living world via free-living and symbiotic bacteria, which incorporate nitrogen into their macromolecules through **nitrogen fixation** (conversion of N_2). **Cyanobacteria** live in most aquatic ecosystems where sunlight is present; they play a key role in nitrogen fixation.

Cyanobacteria are able to use inorganic sources of nitrogen to fix nitrogen. ***Rhizobium*** bacteria live symbiotically in the **root nodules** of **legumes** (such as peas, beans, and peanuts) and provide them with the organic nitrogen they need. **Free-living bacteria**, such as *Azotobacter*, are also important nitrogen fixers.

Organic nitrogen is especially important to the study of ecosystem dynamics since many ecosystem processes, such as primary production and decomposition, are limited by the available supply of nitrogen. As shown in Figure 6, the nitrogen that enters living systems by nitrogen fixation is successively converted from organic nitrogen back into nitrogen gas by bacteria. This process occurs in three steps in terrestrial systems: **ammonification**, **nitrification**, and **denitrification**. First, the ammonification process converts nitrogenous waste from living animals or from the remains of dead animals into ammonium (NH_4^+) by certain bacteria and fungi. Second, the ammonium is converted to nitrites (NO_2^-) by nitrifying bacteria, such as *Nitrosomonas*, through nitrification. Subsequently, nitrites are converted to nitrates (NO_3^-) by similar organisms. Third, the process of denitrification occurs, whereby bacteria, such as *Pseudomonas* and *Clostridium*, convert the nitrates into nitrogen gas, allowing it to re-enter the atmosphere.

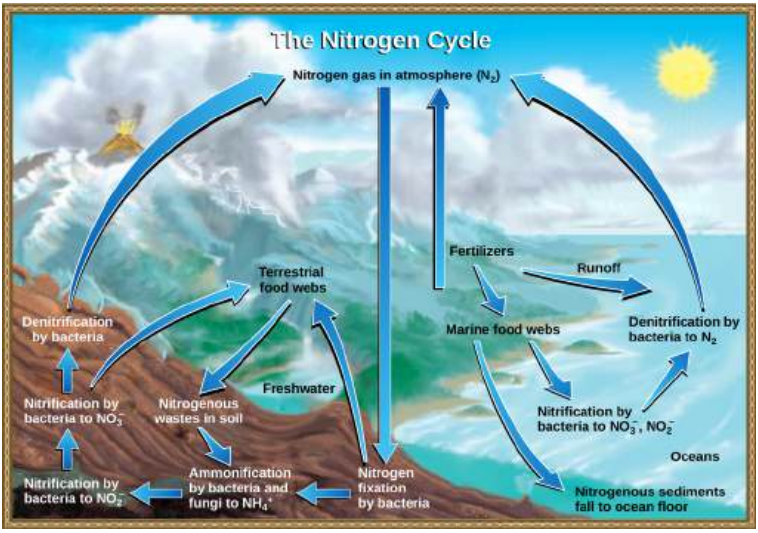


Figure 6: Nitrogen enters the living world from the atmosphere via nitrogen-fixing bacteria. This nitrogen and nitrogenous waste from animals is then processed back into gaseous nitrogen by soil bacteria, which also supply terrestrial food webs with the organic nitrogen they need. (credit: modification of work by John M. Evans and Howard Perlman, USGS. "Nitrogen Cycle" by OpenStax is licensed under CC BY 4.0)

Human activity can release nitrogen into the environment by two primary means: the combustion of fossil fuels, which releases different nitrogen oxides, and by the use of **artificial fertilizers** in agriculture, which are then washed into lakes, streams, and rivers by surface runoff. Atmospheric nitrogen is associated with several effects on Earth's ecosystems including the production of **acid rain** (as nitric acid, HNO_3) and greenhouse gas (as nitrous oxide, N_2O) potentially causing climate change. A major effect of fertilizer runoff is saltwater and freshwater **eutrophication**, a process whereby nutrient runoff causes the excess growth of

microorganisms, depleting dissolved oxygen levels and killing ecosystem fauna.

A similar process occurs in the marine nitrogen cycle, where the ammonification, nitrification, and denitrification processes are performed by marine bacteria. Some of this nitrogen falls to the ocean floor as sediment, which can then be moved to land in geologic time by the uplift of the Earth's surface and thereby incorporated into terrestrial rock. Although the movement of nitrogen from rock directly into living systems has been traditionally seen as insignificant compared with nitrogen fixed from the atmosphere, a recent study showed that this process may indeed be significant and should be included in any study of the global nitrogen cycle (Morford, Houlton, & Dahlgren, 2011).

The Phosphorus Cycle

Phosphorus is an essential nutrient for living processes; it is a major component of nucleic acid and phospholipids, and, as calcium phosphate, makes up the supportive components of our bones. Phosphorus is often the limiting nutrient (necessary for growth) in aquatic ecosystems (Figure 7).

Phosphorus occurs in nature as the **phosphate ion** (PO_4^{3-}). In addition to phosphate runoff as a result of human activity, natural surface runoff occurs when it is leached from phosphate-containing rock by weathering, thus sending phosphates into rivers, lakes, and the ocean. This rock has its origins in the ocean. Phosphate-containing ocean sediments form primarily from the bodies of ocean organisms and from their excretions. However, in remote regions, volcanic ash, aerosols, and mineral dust may also be significant phosphate sources. This sediment then is moved to land over geologic time by the uplifting of areas of the Earth's surface.

Phosphorus is also reciprocally exchanged between phosphate dissolved in the ocean and marine ecosystems.

The movement of phosphate from the ocean to the land and through the soil is extremely slow, with the average phosphate ion having an oceanic residence time between 20,000 and 100,000 years.

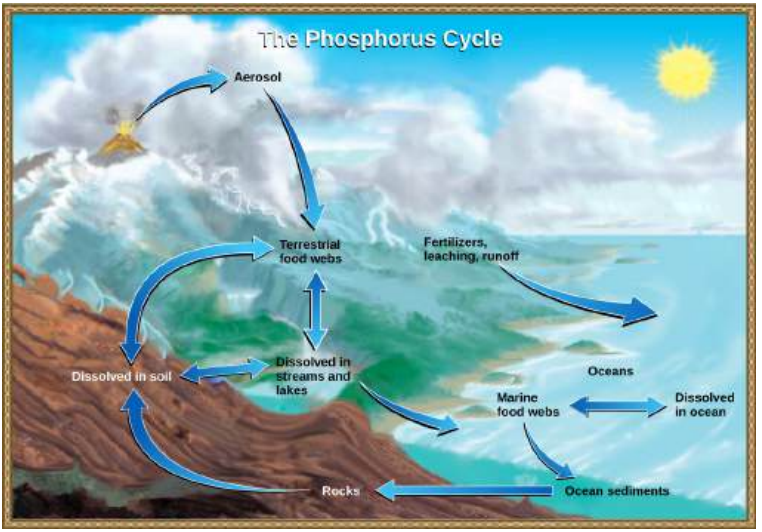


Figure 7: In nature, phosphorus exists as the phosphate ion (PO_4^{3-}). Weathering of rocks and volcanic activity releases phosphate into the soil, water, and air, where it becomes available to terrestrial food webs. Phosphate enters the oceans via surface runoff, groundwater flow, and river flow. Phosphate dissolved in ocean water cycles into marine food webs. Some phosphate from the marine food webs falls to the ocean floor, where it forms sediment. (credit: modification of work by John M. Evans and Howard Perlman, USGS. "phosphorus cycle" by OpenStax is licensed under CC BY 4.0)

Excess phosphorus and nitrogen that enters these ecosystems from fertilizer runoff and from sewage causes excessive growth of microorganisms and depletes the dissolved oxygen, which leads to the death of many ecosystem fauna, such as shellfish and finfish. This process

is responsible for **dead zones** in lakes and at the mouths of many major rivers (Figure 8).

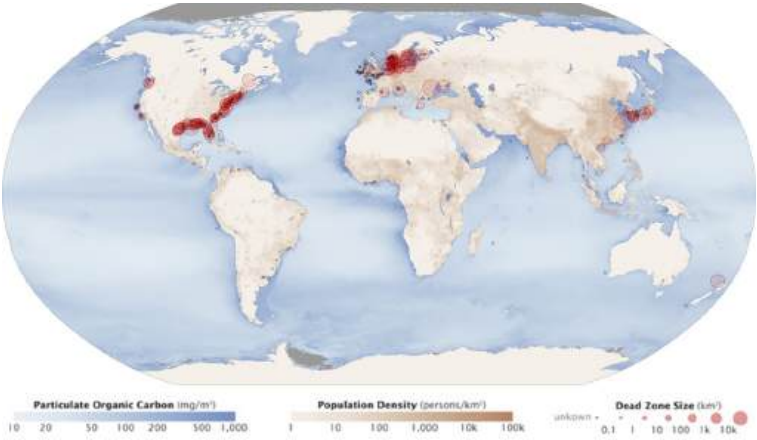


Figure 8: Dead zones occur when phosphorus and nitrogen from fertilizers cause excessive growth of microorganisms, which depletes oxygen and kills fauna. Worldwide, large dead zones are found in coastal areas of high population density. (credit: NASA Earth Observatory. "dead zones" by OpenStax is licensed under CC BY 4.0)

A dead zone is an area within a freshwater or marine ecosystem where large areas are depleted of their normal flora and fauna; these zones can be caused by , oil spills, dumping of toxic chemicals, and other human activities. The number of dead zones has been increasing for several years, and more than 400 of these zones were present as of 2008. One of the worst dead zones is off the coast of the United States in the Gulf of Mexico, where fertilizer runoff from the Mississippi River basin has created a dead zone of over 8463 square miles. Phosphate and nitrate runoff from fertilizers also negatively affect several lake and bay ecosystems including the Chesapeake Bay in the eastern United States.

Everyday Connection – Chesapeake Bay

The Chesapeake Bay has long been valued as one of the most scenic areas on Earth. It is now in distress and is recognized as a declining ecosystem. In the 1970s, the Chesapeake Bay was one of the first ecosystems to have identified dead zones, which continue to kill many fish and bottom-dwelling species, such as clams, oysters, and worms. Several species have declined in the Chesapeake Bay due to surface water runoff containing excess nutrients from artificial fertilizer used on land. The source of the fertilizers (with high nitrogen and phosphate content) is not limited to agricultural practices. There are many nearby urban areas and more than 150 rivers and streams empty into the bay that are carrying **fertilizer runoff** from lawns and gardens. Thus, the decline of the Chesapeake Bay is a complex issue and requires the cooperation of industry, agriculture, and everyday homeowners.

Of particular interest to conservationists is the oyster population; it is estimated that more than 200,000 acres of oyster reefs existed in the bay in the 1700s, but that number has now declined to only 36,000 acres. Oyster harvesting was once a major industry for Chesapeake Bay, but it declined 88 percent between 1982 and 2007. This decline was due not only to fertilizer runoff and dead zones but also to **overharvesting**. Oysters require a certain minimum population density because they must be in close proximity to reproduce. Human activity has altered the oyster population and locations, greatly disrupting the ecosystem.

The restoration of the oyster population in the Chesapeake Bay has been ongoing for several years with mixed success. Not only do many people find oysters good to eat, but they also clean up the bay. Oysters are filter feeders, and as they eat, they clean the water around them. In the 1700s, it was estimated that it took only a few days for the oyster

population to filter the entire volume of the bay. Today, with changed water conditions, it is estimated that the present population would take nearly a year to do the same job.

Restoration efforts have been ongoing for several years by non-profit organizations, such as the Chesapeake Bay Foundation. The restoration goal is to find a way to increase population density so the oysters can reproduce more efficiently. Many disease-resistant varieties (developed at the Virginia Institute of Marine Science for the College of William and Mary) are now available and have been used in the construction of experimental oyster reefs. Efforts to clean and restore the bay by Virginia and Delaware have been hampered because much of the pollution entering the bay comes from other states, which stresses the need for inter-state cooperation to gain successful restoration.

The new, hearty oyster strains have also spawned a new and economically viable industry—oyster aquaculture—which not only supplies oysters for food and profit, but also has the added benefit of cleaning the bay.

The Sulfur Cycle

Sulfur is an essential element for the macromolecules of living things. As a part of the amino acid cysteine, it is involved in the formation of **disulfide bonds** within proteins, which help to determine their 3-D folding patterns, and hence their functions. As shown in Figure 9, sulfur cycles between the oceans, land, and atmosphere. Atmospheric sulfur is found in the form of sulfur dioxide (SO_2) and enters the atmosphere in three ways: from the decomposition of organic molecules, from volcanic activity and geothermal vents, and from the burning of fossil fuels by humans.

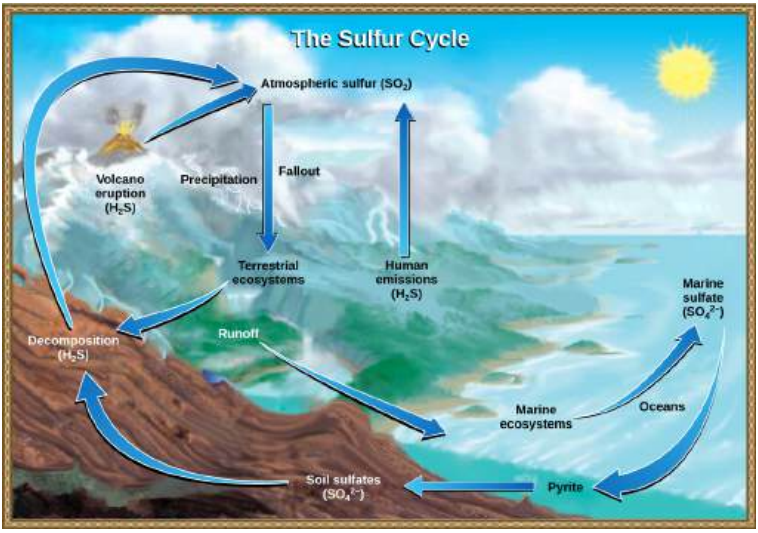


Figure 9: Sulfur dioxide from the atmosphere becomes available to terrestrial and marine ecosystems when it is dissolved in precipitation as weak sulfuric acid or when it falls directly to the Earth as fallout. Weathering of rocks also makes sulfates available to terrestrial ecosystems. Decomposition of living organisms returns sulfates to the ocean, soil and atmosphere. (credit: modification of work by John M. Evans and Howard Perlman, USGS. "The Sulfur Cycle" by OpenStax is licensed under CC BY 4.0)

On land, sulfur is deposited in four major ways: **precipitation**, **direct fallout** from the atmosphere, **rock weathering**, and **geothermal vents** (Figure 10). Atmospheric sulfur is found in the form of **sulfur dioxide** (SO_2), and as rain falls through the atmosphere, sulfur is dissolved in the form of weak **sulfuric acid** (H_2SO_4). Sulfur can also fall directly from the atmosphere in a process called fallout. Also, the weathering of sulfur-containing rocks releases sulfur into the soil. These rocks originate from ocean sediments that are moved to land by the geologic uplifting of ocean sediments. Terrestrial ecosystems can then make

use of these soil **sulfates** (SO4−SO4−">SO4−), and upon the death and decomposition of these organisms, release the sulfur back into the atmosphere as **hydrogen sulfide** (H_2S) gas.

Figure 10: At this sulfur vent in Lassen Volcanic National Park in northeastern California, the yellowish sulfur deposits are visible near the mouth of the vent. (credit: "sulfur vent" by OpenStax is licensed under CC BY 4.0)

Sulfur enters the ocean via runoff from land, atmospheric fallout, and underwater geothermal vents. Some ecosystems rely on chemoautotrophs using sulfur as a biological energy source. This sulfur then supports marine ecosystems in the form of sulfates.

Human activities have played a major role in altering the balance of the global sulfur cycle. The burning of large quantities of fossil fuels, especially coal, releases larger amounts of hydrogen sulfide gas into the atmosphere. As rain falls through this gas, it creates the phenomenon known

as **acid rain**. Acid rain is corrosive rain caused by rainwater falling to the ground through sulfur dioxide gas, turning it into weak sulfuric acid, which causes damage to aquatic ecosystems. Acid rain damages the natural environment by lowering the pH of lakes, which kills many of the resident fauna; it also affects the man-made environment through the chemical degradation of buildings. For example, many marble monuments, such as the Lincoln Memorial in Washington, DC, have suffered significant damage from acid rain over the years. These examples show the wide-ranging effects of human activities on our environment and the challenges that remain for our future.

Summary

Mineral nutrients are cycled through ecosystems and their environment. Of particular importance are water, carbon, nitrogen, phosphorus, and sulfur. All of these cycles have major impacts on ecosystem structure and function. As human activities have caused major disturbances to these cycles, their study and modeling is especially important. A variety of human activities, such as pollution, oil spills, and events) have damaged ecosystems, potentially causing global climate change. The health of Earth depends on understanding these cycles and how to protect the environment from irreversible damage.

REFERENCES

Morford, S., Houlton, B. and Dahlgren, R. 2011. "Increased Forest Ecosystem Carbon and Nitrogen Storage from Nitrogen Rich Bedrock," *Nature* 477, no. 7362: 78–81.

OpenStax, Biology. OpenStax CNX. June 26, 2020.

https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

CONSERVATION BIOLOGY AND BIODIVERSITY

Figure 1: Lake Victoria in Africa, shown in this satellite image, was the site of one of the most extraordinary evolutionary findings on the planet, as well as a casualty of devastating biodiversity loss. (credit: modification of work by Rishabh Tatiraju, using NASA World Wind software. "Lake Victoria" by OpenStax is licensed under CC BY 4.0)

In the 1980s, biologists working in Lake Victoria (Figure 1) in Africa discovered one of the most extraordinary products of evolution on the planet. Located in the Great Rift Valley,

Lake Victoria is a large lake about 68,900 km^2 in area (larger than Lake Huron, the second largest of North America's Great Lakes). Biologists were studying species of a family of fish called cichlids. They found that as they sampled for fish in different locations of the lake, they never stopped finding new species, and they identified nearly 500 evolved types of cichlids. But while studying these variations, they quickly discovered that the invasive Nile Perch was destroying the lake's cichlid population, bringing hundreds of cichlid species to extinction with devastating rapidity.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

The Biodiversity Crisis

Traditionally, ecologists have measured **biodiversity**, a general term for the variety present in the biosphere, by taking into account both the number of species and their commonness. Biodiversity can be estimated at a number of levels of organization of living things. These estimation indexes, which came from information theory, are most useful as a first step in quantifying biodiversity between and within ecosystems; they are less useful when the main concern among conservation biologists is simply the loss of biodiversity. However, biologists recognize that measures of biodiversity, in terms of species diversity, may help focus efforts to preserve the biologically or technologically important elements of biodiversity.

The Lake Victoria cichlids provide an example through which we can begin to understand biodiversity. The biologists studying cichlids in the 1980s discovered hundreds of cichlid species representing a variety of specializations to particular habitat types and specific feeding strategies: eating plankton floating in the water, scraping and then eating algae from rocks, eating insect larvae from the bottom, and eating the eggs of other species of cichlid. The cichlids of Lake Victoria are the product of an adaptive radiation. An **adaptive radiation** is a rapid (less than three million years in the case of the Lake Victoria cichlids) branching through speciation of a phylogenetic tree into many closely related species;

typically, the species "radiate" into different habitats and niches. The Galápagos finches are an example of a modest adaptive radiation with 15 species. The cichlids of Lake Victoria are an example of a spectacular adaptive radiation that includes about 500 species.

At the time biologists were making this discovery, some species began to quickly disappear. A culprit in these declines was a species of large fish that was introduced to Lake Victoria by fisheries to feed the people living around the lake. The Nile perch was introduced in 1963, but lay low until the 1980s when its populations began to surge. The Nile perch population grew by consuming cichlids, driving species after species to the point of extinction (the disappearance of a species). In fact, there were several factors that played a role in the extinction of perhaps 200 cichlid species in Lake Victoria: the Nile perch, declining lake water quality due to agriculture and land clearing on the shores of Lake Victoria, and increased fishing pressure. Scientists had not even cataloged all of the species present—so many were lost that were never named. The diversity is now a shadow of what it once was.

The cichlids of Lake Victoria are a thumbnail sketch of contemporary rapid species loss that occurs all over Earth and is caused by human activity. **Extinction** is a natural process of **macroevolution** that occurs at the rate of about one out of 1 million species becoming extinct per year. The fossil record reveals that there have been five periods of mass extinction in history with much higher rates of species loss, and the rate of species loss today is comparable to those periods of mass extinction. However, there is a major difference between the previous mass extinctions and the current extinction we are experiencing: human activity. Specifically, three human activities have a major impact: the **destruction of habitat**, the **introduction of exotic species**, and **over-harvesting**. Predictions of species loss within the

next century, a tiny amount of time on geological timescales, ranging from 10 percent to 50 percent. Extinctions on this scale have only happened five other times in the history of the planet, and they have been caused by cataclysmic events that changed the course of the history of life in each instance. Many scientists believe Earth is now entering a sixth mass extinction.

Types of Biodiversity

Scientists generally accept that the term **biodiversity** describes the number and kinds of species in a location or on the planet. Species can be difficult to define, but most biologists still feel comfortable with the concept and are able to identify and count eukaryotic species in most contexts. Biologists have also identified alternate measures of biodiversity, some of which are important for planning how to preserve biodiversity.

Genetic diversity is one of those alternate concepts. Genetic diversity or variation is the raw material for adaptation in a species. A species' future potential for adaptation depends on the genetic diversity held in the genomes of the individuals in populations that make up the species. The same is true for higher taxonomic categories. A genus with very different types of species will have more genetic diversity than a genus with species that look alike and have similar ecologies. If there were a choice between one of these genera of species being preserved, the one with the greatest potential for subsequent evolution is the most genetically diverse one. It would be ideal not to have to make such choices, but increasingly this may be the norm.

Many genes code for proteins, which in turn carry out the metabolic processes that keep organisms alive and reproducing. Genetic diversity can be measured as chemical diversity in that different species produce a variety of

chemicals in their cells, both the proteins as well as the products and byproducts of metabolism. This **chemical diversity** has potential benefits for humans as a source of pharmaceuticals, so it provides one way to measure diversity that is important to human health and welfare.

Humans have generated diversity in domestic animals, plants, and fungi. This diversity is also suffering losses because of migration, market forces, and increasing globalism in agriculture, especially in heavily populated regions such as China, India, and Japan. The human population directly depends on this diversity as a stable food source, and its decline is troubling biologists and agricultural scientists.

It is also useful to define **ecosystem diversity**, meaning the number of different ecosystems on the planet or in a given geographic area. Whole ecosystems can disappear even if some of the species might survive by adapting to other ecosystems. The loss of an ecosystem means the loss of interactions between species, the loss of unique features of **coadaptation**, and the loss of biological productivity that an ecosystem is able to create. An example of a largely extinct ecosystem in North America is the prairie ecosystem. Prairies once spanned central North America from the boreal forest in northern Canada down into Mexico. They are now all but gone, replaced by crop fields, pasture lands, and suburban sprawl. Many of the species survive, but the hugely productive ecosystem that was responsible for creating the most productive agricultural soils is now gone. As a consequence, soils are disappearing or must be maintained at greater expense.

Current Species Diversity

Despite considerable effort, knowledge of the species that inhabit the planet is limited. A recent estimate suggests that

the eukaryote species for which science has names, about 1.5 million species, account for less than 20 percent of the total number of eukaryote species present on the planet (8.7 million species, by one estimate). Estimates of numbers of prokaryotic species are largely guesses, but biologists agree that science has only begun to catalog their diversity. Even with what is known, there is no central repository of names or samples of the described species; therefore, there is no way to be sure that the 1.5 million descriptions is an accurate number. It is a best guess based on the opinions of experts in different taxonomic groups. Given that Earth is losing species at an accelerating pace, science is very much in the place it was with the Lake Victoria cichlids: knowing little about what is being lost. Table 1 presents recent estimates of biodiversity in different groups.

Table 1: Estimates of the Numbers of Described and Predicted Species by Taxonomic Group

	Mora et al. 2011		Chapman 2009		Groombridge & Jenkins 2002	
	Described	Predicted	Described	Predicted	Described	Predicted
Animalia	1,124,516	9,920,000	1,424,153	6,836,330	1,225,500	10,820,000
Chromista	17,892	34,900	25,044	200,500	—	—
Fungi	44,368	616,320	98,998	1,500,000	72,000	1,500,000
Plantae	224,244	314,600	310,129	390,800	270,000	320,000
Protozoa	16,236	72,800	28,871	1,000,000	80,000	600,000
Prokaryotes	—	—	10,307	1,000,000	10,175	—
Total	1,438,769	10,960,000	1,897,502	10,897,630	1,657,675	13,240,000

There are various initiatives to catalog described species in accessible ways, and the internet is facilitating that effort. Nevertheless, it has been pointed out that at the current rate of species description, which according to the State of Observed Species Report is 17,000 to 20,000 new species per year, it will take close to 500 years to finish describing life

on this planet (IISE, 2011). Over time, the task becomes both increasingly impossible and increasingly easier as extinction removes species from the planet.

Naming and counting species may seem an unimportant pursuit given the other needs of humanity, but it is not simply an accounting. Describing species is a complex process by which biologists determine an organism's unique characteristics and whether or not that organism belongs to any other described species. It allows biologists to find and recognize the species after the initial discovery and allows them to follow up on questions about its biology. In addition, the unique characteristics of each species make it potentially valuable to humans or other species on which humans depend. Understanding these characteristics is the value of finding and naming species.

Patterns of Biodiversity

Biodiversity is not evenly distributed on Earth. Lake Victoria contained almost 500 species of cichlids alone, ignoring the other fish families present in the lake. All of these species were found only in Lake Victoria; therefore, the 500 species of cichlids were endemic. **Endemic** species are found in only one location. Endemics with highly restricted distributions are particularly vulnerable to extinction. Higher taxonomic levels, such as genera and families, can also be endemic. Lake Huron contains about 79 species of fish, all of which are found in many other lakes in North America. What accounts for the difference in fish diversity in these two lakes? Lake Victoria is a tropical lake, while Lake Huron is a temperate lake. Lake Huron in its present form is only about 7,000 years old, while Lake Victoria in its present form is about 15,000 years old. Biogeographers have suggested these two factors, latitude and age, are two of several hypotheses to explain biodiversity patterns on the planet.

Biogeography is the study of the distribution of the world's species—both in the past and in the present. The work of biogeographers is critical to understanding our physical environment, how the environment affects species, and how environmental changes impact the distribution of a species; it has also been critical to developing evolutionary theory. Biogeographers need to understand both biology and ecology. They also need to be well-versed in evolutionary studies, soil science, and climatology.

There are three main fields of study under the heading of biogeography: **ecological biogeography**, **historical biogeography** (called **paleobiogeography**), and **conservation biogeography**. Ecological biogeography studies the current factors affecting the distribution of plants and animals. Historical biogeography, as the name implies, studies the past distribution of species. Conservation biogeography, on the other hand, is focused on the protection and restoration of species based upon known historical and current ecological information. Each of these fields considers both **zoogeography** and **phytogeography**—the past and present distribution of animals and plants.

One of the oldest observed patterns in ecology is that species biodiversity in almost every taxonomic group increases as latitude declines. In other words, biodiversity increases closer to the equator (Figure 1).

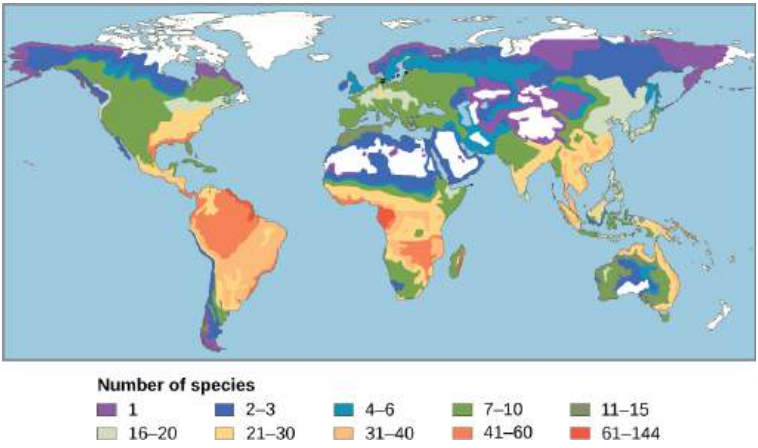


Figure 1: This map illustrates the number of amphibian species across the globe and shows the trend toward higher biodiversity at lower latitudes. A similar pattern is observed for most taxonomic groups. (credit: "amphibian map" by OpenStax is licensed under CC BY 4.0)

It is not yet clear why biodiversity increases closer to the equator, but hypotheses include the greater age of the ecosystems in the tropics versus temperate regions that were largely devoid of life or drastically impoverished during the last glaciation. The idea is that greater age provides more time for speciation. Another possible explanation is the increased energy the tropics receive from the sun versus the decreased energy that temperate and polar regions receive. It is not entirely clear how greater energy input could translate into more species. The complexity of tropical ecosystems may promote speciation by increasing the **heterogeneity**, or number of ecological niches, in the tropics relative to higher latitudes. The greater heterogeneity provides more opportunities for coevolution, specialization, and perhaps greater selection pressures leading to population differentiation. However, this hypothesis suffers from some circularity—ecosystems with more species

encourage speciation, but how did they get more species to begin with? The tropics have been perceived as being more stable than temperate regions, which have a pronounced climate and day-length seasonality. The tropics have their own forms of seasonality, such as rainfall, but they are generally assumed to be more stable environments and this stability might promote speciation.

Regardless of the mechanisms, it is certainly true that all levels of biodiversity are greatest in the tropics. Additionally, the rate of endemism is the highest, and there are more biodiversity hotspots. However, this richness of diversity also means that our knowledge of species is lowest, and there is a high potential for biodiversity loss.

Conservation of Biodiversity

In 1988, British environmentalist Norman Myers developed a conservation concept to identify areas rich in species and at significant risk for species loss: **biodiversity hotspots**. Biodiversity hotspots are geographical areas that contain high numbers of endemic species. The purpose of the concept was to identify important locations on the planet for conservation efforts, a kind of **conservation triage**. By protecting hotspots, governments are able to protect a larger number of species. The original criteria for a hotspot included the presence of 1500 or more endemic plant species and 70 percent of the area disturbed by human activity. There are now 34 biodiversity hotspots (Figure 2) containing large numbers of endemic species, which include half of Earth's endemic plants.

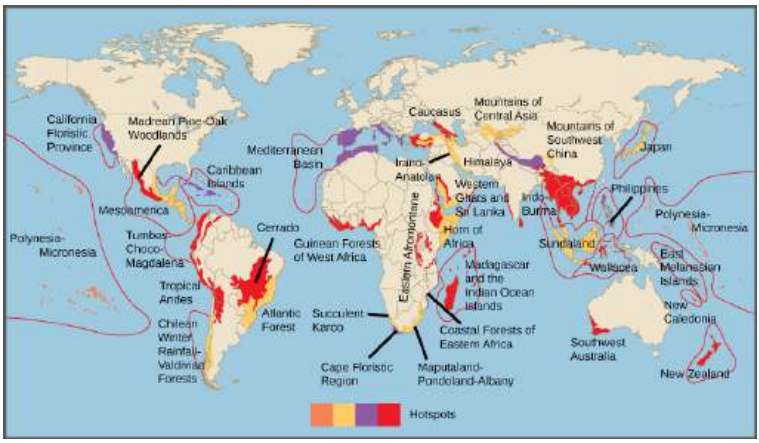


Figure 2: Conservation International has identified 34 biodiversity hotspots, which cover only 2.3 percent of the Earth's surface but have endemic to them 42 percent of the terrestrial vertebrate species and 50 percent of the world's plants. (credit: "biodiversity hotspots" by OpenStax is licensed under CC BY 4.0)

Biodiversity Change through Geological Time

The number of species on the planet, or in any geographical area, is the result of an equilibrium of two evolutionary processes that are ongoing: **speciation** and **extinction**. Both are natural "birth" and "death" processes of macroevolution. When speciation rates begin to outstrip extinction rates, the number of species will increase; likewise, the number of species will decrease when extinction rates begin to overtake speciation rates. Throughout Earth's history, these two processes have fluctuated—sometimes leading to dramatic changes in the number of species on Earth as reflected in the fossil record (Figure 3).

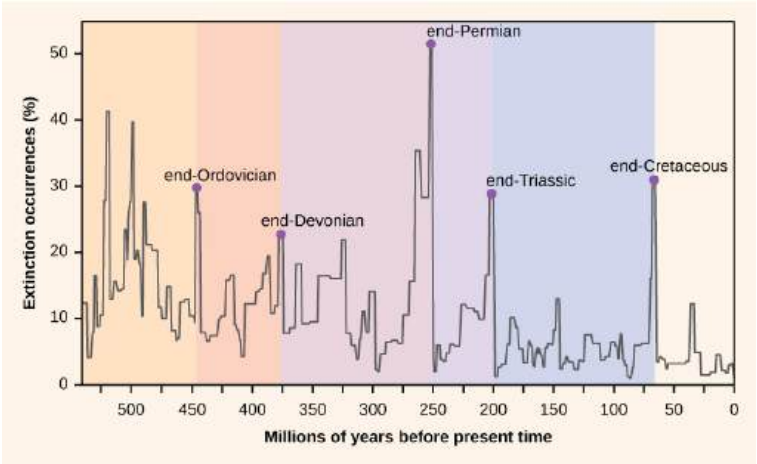


Figure 3: Percent extinction occurrences as reflected in the fossil record have fluctuated throughout Earth's history. Sudden and dramatic losses of biodiversity, called mass extinctions, have occurred five times. (credit: "extinction occurrences" by OpenStax is licensed under CC BY 4.0)

Paleontologists have identified five strata in the fossil record that appear to show sudden and dramatic (greater than half of all extant species disappearing from the fossil record) losses in biodiversity. These are called **mass extinctions**. There are many lesser, yet still dramatic, extinction events, but the five mass extinctions have attracted the most research. An argument can be made that the five mass extinctions are only the five most extreme events in a continuous series of large extinction events throughout the Phanerozoic (since 542 million years ago). In most cases, the hypothesized causes are still controversial.

The Five Mass Extinctions

The fossil record of the mass extinctions was the basis for

defining periods of geological history, so they typically occur at the transition point between geological periods. The transition in fossils from one period to another reflects the dramatic loss of species and the gradual origin of new species. These transitions can be seen in the rock strata. Table 2 provides data on the five mass extinctions.

Table 2: the names and dates for the five mass extinctions in Earth's history.

Mass Extinctions

Geological Period	Mass Extinction Name	Time (millions of years ago)
Ordovician–Silurian	end-Ordovician O–S	450–440
Late Devonian	end-Devonian	375–360
Permian–Triassic	end-Permian	251
Triassic–Jurassic	end-Triassic	205
Cretaceous–Paleogene	end-Cretaceous K–Pg (K–T)	65.5

The **Ordovician-Silurian extinction event** is the first recorded mass extinction and the second largest. During this period, about 85 percent of marine species (few species lived outside the oceans) became extinct. The main hypothesis for its cause is a period of glaciation and then warming. The extinction event actually consists of two extinction events separated by about 1 million years. The first event was caused by cooling, and the second event was due to the subsequent warming. The climate changes affected temperatures and sea levels. Some researchers have suggested that a gamma-ray burst, caused by a nearby supernova, is a possible cause of the Ordovician-Silurian extinction. The gamma-ray burst would have stripped away the Earth's ozone layer causing intense ultraviolet radiation from the sun and may account for climate changes observed at the time. The hypothesis is speculative, but extraterrestrial influences on Earth's history are an active line of research. Recovery of biodiversity after the mass extinction took from 5 to 20 million years, depending on the location.

The **late Devonian extinction** may have occurred over a relatively long period of time. It appears to have affected marine species and not the plants or animals inhabiting terrestrial habitats. The causes of this extinction are poorly understood.

The **end-Permian extinction** was the largest in the history of life. Indeed, an argument could be made that Earth nearly became devoid of life during this extinction event. The planet looked very different before and after this event. Estimates are that 96 percent of all marine species and 70 percent of all terrestrial species were lost. It was at this time, for example, that the trilobites, a group that survived the Ordovician–Silurian extinction, became extinct. The causes for this mass extinction are not clear, but the leading suspect is extended and widespread volcanic activity that led to a runaway global-warming event. The oceans became largely anoxic, suffocating marine life. Terrestrial tetrapod diversity took 30 million years to recover after the end-Permian extinction. The Permian extinction dramatically altered Earth's biodiversity makeup and the course of evolution.

The causes of the **Triassic–Jurassic extinction** event are not clear and hypotheses of climate change, asteroid impact, and volcanic eruptions have been argued. The extinction event occurred just before the breakup of the supercontinent Pangaea, although recent scholarship suggests that the extinction may have occurred more gradually throughout the Triassic.

The causes of the **end-Cretaceous extinction** event are the ones that are best understood. It was during this extinction event about 65 million years ago that the **dinosaurs**, the dominant vertebrate group for millions of years, disappeared from the planet (with the exception of a theropod clade that gave rise to birds). Every land animal that weighed more than 25 kg became extinct. The cause of this extinction is now understood to be the result of a

cataclysmic impact of a large meteorite, or asteroid, off the coast of what is now the Yucatán Peninsula. This hypothesis, proposed first in 1980, was a radical explanation based on a sharp spike in the levels of iridium (which rains down from space in meteors at a fairly constant rate but is otherwise absent on Earth's surface) at the rock stratum that marks the boundary between the Cretaceous and Paleogene periods (Figure 4). This boundary marked the disappearance of the dinosaurs in fossils as well as many other taxa. The researchers who discovered the iridium spike interpreted it as a rapid influx of iridium from space to the atmosphere (in the form of a large asteroid) rather than a slowing in the deposition of sediments during that period. It was a radical explanation, but the report of an appropriately aged and sized impact crater in 1991 made the hypothesis more believable. Now an abundance of geological evidence supports the theory. Recovery times for biodiversity after the end-Cretaceous extinction are shorter, in geological time, than for the end-Permian extinction, on the order of 10 million years.

Figure 4: In 1980, Luis and Walter Alvarez, Frank Asaro, and Helen Michels discovered, across the world, a spike in the concentration of iridium within the sedimentary layer at the K–Pg boundary. These researchers hypothesized that this iridium spike was caused by an asteroid impact that resulted in the K–Pg mass extinction. In the photo, the iridium layer is the light band. (credit: USGS. "K–Pg boundary" by OpenStax is licensed under CC BY 4.0)

The Pleistocene Extinction

The **Pleistocene Extinction** is one of the lesser extinctions and a recent one. It is well known that the North American, and to some degree Eurasian, megafauna, or large animals, disappeared toward the end of the last glaciation period. The extinction appears to have happened in a relatively restricted time period of 10,000–12,000 years ago. In North America, the losses were quite dramatic and included the **woolly mammoths** (last dated about 4,000 years ago in an isolated population), **mastodon**, **giant beavers**, **giant ground**

sloths, **saber-toothed cats**, and the **North American camel**, just to name a few. The possibility that the rapid extinction of these large animals was caused by over-hunting was first suggested in the 1900s. Research into this hypothesis continues today. It seems likely that over-hunting caused many pre-written history extinctions in many regions of the world.

In general, the timing of the Pleistocene extinctions correlated with the arrival of humans and not with climate-change events, which is the main competing hypothesis for these extinctions. The extinctions began in Australia about 40,000 to 50,000 years ago, just after the arrival of humans in the area: a marsupial lion, a giant one-ton wombat, and several giant kangaroo species disappeared. In North America, the extinction of almost all of the large mammals occurred 10,000–12,000 years ago. All that are left are the smaller mammals such as bears, elk, moose, and cougars. Finally, on many remote oceanic islands, the extinction of many species occurred coincident with human arrivals. Not all of the islands had large animals, but when there were large animals, they were lost. Madagascar was colonized about 2,000 years ago and the large mammals that lived there became extinct. Eurasia and Africa do not show this pattern, but they also did not experience a recent arrival of humans. Humans arrived in Eurasia hundreds of thousands of years ago after the origin of the species in Africa. This topic remains an area of active research and hypothesizing. It seems clear that even if climate played a role, in most cases human hunting precipitated the extinctions.

Present-Time Extinctions

The **sixth mass extinction** (also called the **Holocene extinction**) appears to have begun earlier than previously believed and has mostly to do with the activities of *Homo*

sapiens. Since the beginning of the Holocene period, there are numerous recent extinctions of individual species that are recorded in human writings. Most of these are coincident with the expansion of the European colonies since the 1500s.

One of the earlier and popularly known examples is the dodo bird. The dodo bird lived in the forests of Mauritius, an island in the Indian Ocean. The dodo bird became extinct around 1662. It was hunted for its meat by sailors and was easy prey because the dodo, which did not evolve with humans, would approach people without fear. Introduced pigs, rats, and dogs brought to the island by European ships also killed dodo young and eggs.

Steller's sea cow became extinct in 1768; it was related to the manatee and probably once lived along the northwest coast of North America. Steller's sea cow was first discovered by Europeans in 1741 and was hunted for meat and oil. The last sea cow was killed in 1768. That amounts to 27 years between the sea cow's first contact with Europeans and the extinction of the species.

In 1914, the last living passenger pigeon died in a zoo in Cincinnati, Ohio. This species had once darkened the skies of North America during its migrations, but it was hunted and suffered from habitat loss through the clearing of forests for farmland. In 1918, the last living Carolina parakeet died in captivity. This species was once common in the eastern United States, but it suffered from habitat loss. The species was also hunted because it ate orchard fruit when its native foods were destroyed to make way for farmland. The Japanese sea lion, which inhabited a broad area around Japan and the coast of Korea, became extinct in the 1950s due to fishermen. The Caribbean monk seal was distributed throughout the Caribbean Sea but was driven to extinction via hunting by 1952.

These are only a few of the recorded extinctions in the past 500 years. The International Union for Conservation of

Nature (IUCN) keeps a list of extinct and endangered species called the **Red List**. The list is not complete, but it describes 380 extinct species of vertebrates after 1500 AD, 86 of which were driven extinct by overhunting or overfishing.

Estimates of Present-Time Extinction Rates

Estimates of **extinction rates** are hampered by the fact that most extinctions are probably happening without observation. The extinction of a bird or mammal is likely to be noticed by humans, especially if it has been hunted or used in some other way. But there are many organisms that are of less interest to humans (not necessarily of less value) and many that are undescribed.

The **background extinction rate** is estimated to be about one per million species per year (E/MSY). For example, assuming there are about ten million species in existence, the expectation is that ten species would become extinct each year (each year represents ten million species per year).

One contemporary extinction rate estimate uses the extinctions in the written record since the year 1500. For birds alone, this method yields an estimate of 26 E/MSY. However, this value may be underestimated for three reasons. First, many species would not have been described until much later in the time period, so their loss would have gone unnoticed. Second, the number of recently extinct species is increasing because extinct species now are being described from skeletal remains. And third, some species are probably already extinct even though conservationists are reluctant to name them as such. Taking these factors into account raises the estimated extinction rate closer to 100 E/MSY. The predicted rate by the end of the century is 1500 E/MSY.

A second approach to estimating present-time extinction rates is to correlate species loss with habitat loss by

measuring forest-area loss and understanding species-area relationships. The **species-area relationship** is the rate at which new species are seen when the area surveyed is increased. Studies have shown that the number of species present increases as the size of the island increases. This phenomenon has also been shown to hold true in other habitats as well. Turning this relationship around, if the habitat area is reduced, the number of species living there will also decline. Estimates of extinction rates based on habitat loss and species-area relationships have suggested that with about 90 percent habitat loss an expected 50 percent of species would become extinct. Species-area estimates have led to species extinction rate calculations of about 1000 E/MSY and higher. In general, actual observations do not show this amount of loss and suggestions have been made that there is a delay in extinction. Recent work has also called into question the applicability of the species-area relationship when estimating the loss of species. This work argues that the species-area relationship leads to an overestimate of extinction rates. A better relationship to use may be the endemics-area relationship. Using this method would bring estimates down to around 500 E/MSY in the coming century. Note that this value is still 500 times the background rate.

Summary

Biodiversity exists at multiple levels of organization and is measured in different ways depending on the goals of those taking the measurements. These measurements include numbers of species, genetic diversity, chemical diversity, and ecosystem diversity. The number of described species is estimated to be 1.5 million with about 17,000 new species being described each year. Estimates for the total number of species on Earth vary but are on the order of 10 million.

Biodiversity is negatively correlated with latitude for most taxa, meaning that biodiversity is higher in the tropics. The mechanism for this pattern is not known with certainty, but several plausible hypotheses have been advanced.

Five mass extinctions with losses of more than 50 percent of extant species are observable in the fossil record. Biodiversity recovery times after mass extinctions vary, but have been up to 30 million years. Recent extinctions are recorded in written history and are the basis for one method of estimating contemporary extinction rates. The other method uses measures of habitat loss and species-area relationships. Estimates of contemporary extinction rates vary, but some rates are as high as 500 times the background rate, as determined from the fossil record, and are predicted to rise.

REFERENCES

Camilo, M. et al. 2011. "How Many Species Are There on Earth and in the Ocean?" *PLoS Biology*, doi:10.1371/journal.pbio.1001127.

Chapman, A. 2009. *Numbers of Living Species in Australia and the World*, 2nd ed. (Canberra, AU: Australian Biological Resources Study). http://www.environment.gov.au/biodiversity/abrs/publications/other/species-numbers/2009/pubs/nlsaw-2nd-complete.pdf.

Groombridge, B. and Jenkins, M. 2002. *World Atlas of Biodiversity: Earth's Living Resources in the 21st Century*. Berkeley: University of California Press.

International Institute for Species Exploration (IISE), 2011. *State of Observed Species (SOS)*. Tempe, AZ: IISE, 2011. Accessed May, 20, 2012. http://species.asu.edu/SOS.

OpenStax, Biology. OpenStax CNX. June 26, 2020.

https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

The Importance of Biodiversity to Human Life

It may not be clear why biologists are concerned about biodiversity loss. When biodiversity loss is thought of as the extinction of the passenger pigeon, the dodo bird, and even the woolly mammoth, the loss may appear to be an emotional one. But is the loss practically important for the welfare of the human species? From the perspective of evolution and ecology, the loss of a particular individual species is unimportant (however, the loss of a keystone species can lead to ecological disaster). Extinction is a normal part of macroevolution. But the accelerated extinction rate means the loss of tens of thousands of species within our lifetimes, and it is likely to have dramatic effects on human welfare through the collapse of ecosystems and in added costs to maintain food production, clean air and water, and human health.

Agriculture began after early hunter-gatherer societies first settled in one place and heavily modified their immediate environment. This cultural transition has made it difficult for humans to recognize their dependence on undomesticated living things on the planet. Biologists recognize the human species is embedded in ecosystems and is dependent on them, just as every other species on the planet is dependent. Technology smoothes out the extremes of existence, but

ultimately the human species cannot exist without its ecosystem.

Human Health

Contemporary societies that live close to the land often have a broad knowledge of the medicinal uses of plants growing in their area. Most plants produce secondary plant compounds, which are toxins used to protect the plant from insects and other animals that eat them, but some of which also work as medication. For centuries cultures around the globe have organized knowledge about the medical uses of plants into ceremonies and traditions. Humans are not the only species to use plants for medicinal reasons: the great apes, orangutans, chimpanzees, bonobos, and gorillas have all been observed self-medicating with plants.

Modern pharmaceutical science also recognizes the importance of these plant compounds. Examples of significant **medicines** derived from plant compounds include aspirin, codeine, digoxin, atropine, and vincristine (Figure 1). Many medicines were once derived from plant extracts but are now synthesized. It is estimated that, at one time, 25 percent of modern drugs contained at least one plant extract. That number has probably decreased to about 10 percent as natural plant ingredients are replaced by synthetic versions. **Antibiotics**, which are responsible for extraordinary improvements in health and lifespans in developed countries, are compounds largely derived from fungi and bacteria.

Figure 1: Catharanthus roseus, the Madagascar periwinkle, has various medicinal properties. Among other uses, it is a source of vincristine, a drug used in the treatment of lymphomas. (credit: Forest and Kim Starr. "Catharanthus roseus" by OpenStax is licensed under CC BY 4.0)

In recent years, animal **venoms** and **poisons** have excited intense research for their medicinal potential. By 2007, the FDA had approved five drugs based on animal toxins to treat diseases such as hypertension, chronic pain, and diabetes. Another five drugs are undergoing clinical trials, and at least six drugs are being used in other countries. Other toxins under investigation come from mammals, snakes, lizards, various amphibians, fish, snails, octopuses, and scorpions.

Aside from representing billions of dollars in profits, these medicines improve people's lives. Pharmaceutical companies are actively looking for new compounds synthesized by living organisms that can function as medicine. It is estimated that 1/3 of pharmaceutical research and development is spent

on natural compounds and that about 35 percent of new drugs brought to market between 1981 and 2002 were from natural compounds. The opportunities for new medications will be reduced in direct proportion to the disappearance of species.

Agricultural Diversity

Since the beginning of human agriculture more than 10,000 years ago, human groups have been breeding and selecting crop varieties. This crop diversity matched the cultural diversity of highly subdivided populations of humans. For example, potatoes were domesticated beginning around 7,000 years ago in the central Andes of Peru and Bolivia. The potatoes grown in that region belong to seven species and the number of varieties likely is in the thousands. Each variety has been bred to thrive at particular elevations and soil and climate conditions. The diversity is driven by the diverse demands of the topography, the limited movement of people, and the demands created by crop rotation for different varieties that will do well in different fields.

Potatoes are only one example of human-generated diversity. Every plant, animal, and fungus that has been cultivated by humans has been bred from original wild ancestor species into diverse varieties arising from the demands for food value, adaptation to growing conditions, and resistance to pests. The potato demonstrates a well-known example of the risks of low crop diversity: the tragic Irish potato famine when the single variety grown in Ireland became susceptible to a potato blight, wiping out the crop. The loss of the crop led to famine, death, and mass emigration. Resistance to disease is a chief benefit to maintaining crop biodiversity, and lack of diversity in contemporary crop species carries similar risks. Seed companies, which are the source of most crop varieties in

developed countries, must continually breed new varieties to keep up with evolving pest organisms. These same seed companies, however, have participated in the decline of the number of varieties available as they focus on selling fewer varieties in more areas of the world.

The ability to create new crop varieties relies on the diversity of varieties available and the accessibility of wild forms related to the crop plant. These wild forms are often the source of new gene variants that can be bred with existing varieties to create varieties with new attributes. Loss of wild species related to a crop will mean the loss of potential in crop improvement. Maintaining the genetic diversity of wild species related to domesticated species ensures our continued food supply.

Since the 1920s, government agriculture departments have maintained seed banks of crop varieties as a way to maintain crop diversity. This system has flaws because, over time, seed banks are lost through accidents, and there is no way to replace them. In 2008, the **Svalbard Global Seed Vault** (Figure 2) began storing seeds from around the world as a backup system to the regional seed banks. If a regional seed bank stores varieties in Svalbard, losses can be replaced from Svalbard. The seed vault is located deep into the rock of an arctic island. Conditions within the vault are maintained at ideal temperature and humidity for seed survival, but the deep underground location of the vault in the arctic means that failure of the vault's systems will not compromise the climatic conditions inside the vault.

Figure 2: The Svalbard Global Seed Vault is a storage facility for seeds of Earth's diverse crops. (credit: Mari Tefre, Svalbard Global Seed Vault. "Svalbard Global Seed Vault" by OpenStax is licensed under CC BY 4.0)

Crop success s is largely dependent on the quality of the **soil**. Although some agricultural soils are rendered sterile using controversial cultivation and chemical treatments, most contain a huge diversity of organisms that maintain nutrient cycles—breaking down organic matter into nutrient compounds that crops need for growth. These organisms also maintain soil texture that affects water and oxygen dynamics in the soil that are necessary for plant growth. If farmers had to maintain arable soil using alternate means, the cost of food would be much higher than it is now. These kinds of processes are called **ecosystem services**. They occur within ecosystems, such as **soil ecosystems**, as a result of the diverse metabolic activities of the organisms

living there, but they provide benefits to human food production, drinking water availability, and breathable air.

Other key ecosystem services related to food production are **plant pollination** and **crop pest control**. Over 150 crops in the United States require pollination to produce food. One estimate of the benefit of honeybee pollinations within the United States is $1.6 billion per year. Other pollinators contribute up to $6.7 billion more.

Many honeybee populations are managed by **apiarists** who rent out their hives' services to farmers. Honeybee populations in North America have been suffering large losses caused by a syndrome known as **colony collapse disorder**, whose cause is unclear. Other pollinators include a diverse array of other bee species and various insects and birds. Loss of these species would make growing crops requiring pollination impossible, increasing dependence on other crops.

Finally, humans compete for their food with crop pests, most of which are insects. **Pesticides** control these competitors; however, pesticides are costly and lose their effectiveness over time as pest populations adapt. They also lead to collateral damage by killing non-pest species and risking the health of consumers and agricultural workers. Ecologists believe that the bulk of the work in removing pests is actually done by predators and parasites of those pests, but the impact has not been well studied. A review found that in 74 percent of studies that looked for an effect of landscape complexity on natural enemies of pests, the greater the complexity, the greater the effect of pest-suppressing organisms. An experimental study found that introducing multiple enemies of pea aphids (an important alfalfa pest) increased the yield of alfalfa significantly. This study shows the importance of landscape diversity via the question of whether a diversity of pests is more effective at control than one single pest; the results showed this to be the case. Loss

of diversity in pest enemies will inevitably make it more difficult and costly to grow food.

Wild Food Sources

In addition to growing crops and raising animals for food, humans obtain food resources from wild populations, primarily **fish populations**. For approximately 1 billion people, aquatic resources provide the main source of animal protein. But since 1990, global fish production has declined. Despite considerable effort, few fisheries on the planet are managed for sustainability.

Fishery extinctions rarely lead to the complete extinction of the harvested species, but rather to a radical restructuring of the marine ecosystem in which a dominant species is so over-harvested that it becomes a minor player, ecologically. In addition to humans losing the food source, these alterations affect many other species in ways that are difficult or impossible to predict. The collapse of fisheries has dramatic and long-lasting effects on local populations that work in the fishery. In addition, the loss of an inexpensive protein source to populations that cannot afford to replace it will increase the cost of living and limit societies in other ways. In general, the fish taken from fisheries have shifted to smaller species as larger species are fished to extinction. The ultimate outcome could clearly be the loss of aquatic systems as food sources.

Psychological and Moral Value

Finally, it has been argued that humans benefit psychologically from living in a biodiverse world. A chief proponent of this idea is entomologist E. O. Wilson. He argues that human evolutionary history has adapted us to live in a natural environment and that built environments

generate stressors that affect human health and well-being. There is considerable research into the **psychological regenerative benefits** of natural landscapes that suggests the hypothesis may hold some truth. In addition, there is a moral argument that humans have a responsibility to inflict as little harm as possible on other species.

Summary

Humans use many compounds that were first discovered or derived from living organisms as medicines: secondary plant compounds, animal toxins, and antibiotics produced by bacteria and fungi. More medicines are expected to be discovered in nature. Loss of biodiversity will impact the number of pharmaceuticals available to humans.

Crop diversity is a requirement for food security, and it is being lost. The loss of wild relatives to crops also threatens breeders' abilities to create new varieties. Ecosystems provide ecosystem services that support human agriculture: pollination, nutrient cycling, pest control, and soil development and maintenance. Loss of biodiversity threatens these ecosystem services and risks making food production more expensive or impossible. Wild food sources are mainly aquatic, but few are being managed for sustainability. Fisheries' ability to provide protein to human populations is threatened when extinction occurs.

Biodiversity may provide important psychological benefits to humans. Additionally, there are moral arguments for the maintenance of biodiversity.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020.

https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Threats to Biodiversity

The core threat to biodiversity on the planet, and therefore a threat to human welfare, is the combination of human population growth and resource exploitation. The human population requires resources to survive and grow, and those resources are being removed unsustainably from the environment. The three greatest proximate threats to biodiversity are habitat loss, overharvesting, and the introduction of exotic species. The first two of these are a direct result of human population growth and resource use. The third results from increased mobility and trade. A fourth major cause of extinction, anthropogenic climate change, is predicted to become significant during this century. Global climate change is also a consequence of the human population's need for energy and the use of fossil fuels to meet those needs (Figure 1). Environmental issues, such as toxic pollution, have specific targeted effects on species, but they are not generally seen as threats at the magnitude of the others.

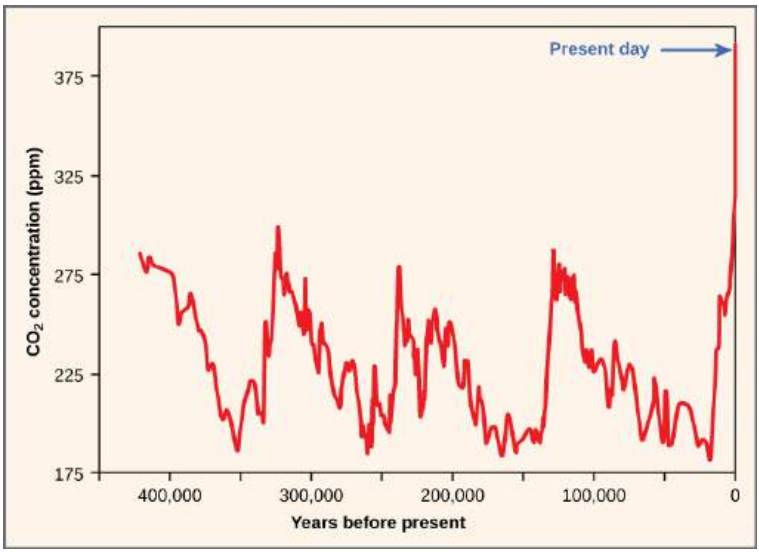


Figure 1: Atmospheric carbon dioxide levels fluctuate in a cyclical manner. However, the burning of fossil fuels in recent history has caused a dramatic increase in the levels of carbon dioxide in the Earth's atmosphere, which have now reached levels never before seen in human history. Scientists predict that the addition of this "greenhouse gas" to the atmosphere is resulting in climate change that will significantly impact biodiversity in the coming century. (credit: "Atmospheric carbon dioxide levels" by OpenStax is licensed under CC BY 4.0)

Habitat Loss

Habitat loss is a major threat to biodiversity. Remove the entire habitat within the range of a species and, unless they are one of the few species that do well in human-built environments, the species will become extinct. Human destruction of habitats accelerated in the latter half of the twentieth century. Consider the exceptional biodiversity of Sumatra: it is home to one species of orangutan, a species of critically endangered elephant, and the Sumatran tiger,

but half of Sumatra's forest is now gone. The neighboring island of Borneo, home to the other species of orangutan, has lost a similar area of forest. Forest loss continues even in protected areas of Borneo. The orangutan in Borneo is listed as endangered by the International Union for Conservation of Nature (IUCN), but it is simply the most visible of thousands of species that will not survive the disappearance of the forests of Borneo. The forests are removed for timber and to plant palm oil plantations (Figure 2). Palm oil is used in many products including food products, cosmetics, and biodiesel in Europe. A five-year estimate of global forest cover loss for the years 2000–2005 was 3.1 percent. In the humid tropics where forest loss is primarily from timber extraction, 272,000 km^2 was lost out of a global total of 11,564,000 km^2 (or 2.4 percent). In the tropics, these losses certainly also represent the extinction of species because of high levels of endemism.

(a)

(b)

(c)

(d)

(e)

Figure 2: (a) One species of orangutan, Pongo pygmaeus, is found only in the rainforests of Borneo, and the other species of orangutan (Pongo abelii) is found only in the rainforests of Sumatra. These animals are examples of the exceptional biodiversity of (c) the islands of Sumatra and Borneo. Other species include the (b) Sumatran tiger (Panthera

tigris sumatrae) and the (d) Sumatran elephant (Elephas maximus sumatranus), both critically endangered species. Rainforest habitat is being removed to make way for (e) oil palm plantations such as this one in Borneo's Sabah Province. (credit a: modification of work by Thorsten Bachner; credit b: modification of work by Dick Mudde; credit c: modification of work by U.S. CIA World Factbook; credit d: modification of work by "Nonprofit Organizations"/Flickr; credit e: modification of work by Dr. Lian Pin Koh. "this image" by OpenStax is licensed under CC BY 4.0)

Preventing Habitat Destruction with Wise Wood Choices

Most consumers do not imagine that the home improvement products they buy might be contributing to habitat loss and species extinctions. Yet the market for **illegally harvested timber** is huge, and the wood products often find themselves in building supply stores in the United States. One estimate is that 10 percent of the imported timber stream in the United States, which is the world's largest consumer of wood products, is potentially illegally logged. In 2006, this amounted to $3.6 billion in wood products. Most of the illegal products are imported from countries that act as intermediaries and are not the originators of the wood.

How is it possible to determine if a wood product, such as flooring, was harvested sustainably or even legally? The Forest Stewardship Council (FSC) certifies sustainably harvested forest products, therefore, looking for their certification on flooring and other hardwood products is one way to ensure that the wood has not been taken illegally from a tropical forest. Certification applies to specific products, not to a producer; some producers' products may not have certification while other products are certified. While there are other industry-backed certifications other than the FSC, these are unreliable due to lack of independence from the industry. Another approach is to buy

domestic wood species. While it would be great if there was a list of legal versus illegal wood products, it is not that simple. Logging and forest management laws vary from country to country; what is illegal in one country may be legal in another. Where and how a product is harvested and whether the forest from which it comes is being maintained sustainably all factor into whether a wood product will be certified by the FSC. It is always a good idea to ask questions about where a wood product came from and how the supplier knows that it was harvested legally.

Habitat destruction can affect ecosystems other than forests. Rivers and streams are important ecosystems and are frequently modified through land development and from damming or water removal. Damming of rivers affects the water flow and access to all parts of a river. Differing flow regimes can reduce or eliminate populations that are adapted to these changes in flow patterns. For example, an estimated 91 percent of river lengths in the United States have been developed: they have modifications like dams, to create energy or store water; levees, to prevent flooding; or dredging or rerouting, to create land that is more suitable for human development. Many fish species in the United States, especially rare species or species with restricted distributions, have seen declines caused by river damming and habitat loss. Research has confirmed that species of amphibians that must carry out parts of their life cycles in both aquatic and terrestrial habitats have a greater chance of suffering population declines and extinction because of the increased likelihood that one of their habitats or access between them will be lost.

Overharvesting

Overharvesting is a serious threat to many species, but particularly to aquatic species. There are many examples of

regulated commercial fisheries monitored by fisheries scientists that have nevertheless collapsed. The western Atlantic cod fishery is the most spectacular recent collapse. While it was a hugely productive fishery for 400 years, the introduction of modern factory trawlers in the 1980s and the pressure on the fishery led to it becoming unsustainable. The causes of fishery collapse are both economic and political in nature. Most fisheries are managed as a common (shared) resource even when the fishing territory lies within a country's territorial waters. Common resources are subject to an economic pressure known as the tragedy of the commons in which essentially no fisher has a motivation to exercise restraint in harvesting a fishery when it is not owned by that fisher. The natural outcome of harvests of resources held in common is their **overexploitation**. While large fisheries are regulated to attempt to avoid this pressure, it still exists in the background. This overexploitation is exacerbated when access to the fishery is open and unregulated and when technology gives fishers the ability to overfish. In a few fisheries, the biological growth of the resource is less than the potential growth of the profits made from fishing if that time and money were invested elsewhere. In these cases—whales are an example—economic forces will always drive toward fishing the population to extinction.

For the most part, fishery extinction is not equivalent to biological extinction—the last fish of a species is rarely fished out of the ocean. At the same time, fishery extinction is still harmful to fish species and their ecosystems. There are some instances in which true extinction is a possibility. Whales have slow-growing populations and are at risk of complete extinction through hunting. There are some species of sharks with restricted distributions that are at risk of extinction. The groupers are another population of generally slow-growing

fishes that, in the Caribbean, includes a number of species that are at risk of extinction from overfishing.

Coral reefs are extremely diverse marine ecosystems that face peril from several processes. Reefs are home to 1/3 of the world's marine fish species—about 4,000 species—despite making up only 1 percent of marine habitat. Most home marine aquaria are stocked with wild-caught organisms, not cultured organisms. Although no species is known to have been driven extinct by the **pet trade** in marine species, there are studies showing that populations of some species have declined in response to harvesting, indicating that the harvest is not sustainable at those levels. There are concerns about the effect of the pet trade on some terrestrial species such as turtles, amphibians, birds, plants, and even the orangutan.

Bush meat is the generic term used for wild animals killed for food. Hunting is practiced throughout the world, but hunting practices, particularly in equatorial Africa and parts of Asia, are believed to threaten several species with extinction. Traditionally, bush meat in Africa was hunted to feed families directly; however, recent commercialization of the practice now has bush meat available in grocery stores, which has increased harvest rates to the level of unsustainability. Additionally, human population growth has increased the need for protein foods that are not being met from agriculture. Species threatened by the bush meat trade are mostly mammals including many primates living in the Congo basin.

Exotic Species

Exotic species are species that have been intentionally or unintentionally introduced by humans into an ecosystem in which they did not evolve. Such introductions likely occur frequently as natural phenomena. For example, Kudzu

(*Pueraria lobata*), which is native to Japan, was introduced in the United States in 1876. It was later planted for soil conservation. Problematically, it grows too well in the southeastern United States—up to a foot a day. It is now a pest species and covers over 7 million acres in the southeastern United States. If an introduced species is able to survive in its new habitat, that introduction is now reflected in the observed range of the species. Human transportation of people and goods, including the intentional transport of organisms for trade, has dramatically increased the introduction of species into new ecosystems, sometimes at distances that are well beyond the capacity of the species to ever travel itself and outside the range of the species' natural predators.

Most exotic species introductions probably fail because of the low number of individuals introduced or poor adaptation to the ecosystem they enter. Some species, however, possess preadaptations that can make them especially successful in a new ecosystem. These exotic species often undergo dramatic population increases in their new habitat and reset the ecological conditions in the new environment, threatening the species that exist there. For this reason, exotic species are also called invasive species. Exotic species can threaten other species through competition for resources, predation, or disease.

Lakes and islands are particularly vulnerable to extinction threats from introduced species. In Lake Victoria, the intentional introduction of the Nile perch was largely responsible for the extinction of about 200 species of cichlids. The accidental introduction of the brown tree snake via aircraft (Figure 3) from the Solomon Islands to Guam in 1950 has led to the extinction of three species of birds and three to five species of reptiles endemic to the island. Several other species are still threatened. The brown tree snake is adept at exploiting human transportation as a means to

migrate; one was even found on an aircraft arriving in Corpus Christi, Texas. Constant vigilance on the part of airport, military, and commercial aircraft personnel is required to prevent the snake from moving from Guam to other islands in the Pacific, especially Hawaii. Islands do not make up a large area of land on the globe, but they do contain a disproportionate number of endemic species because of their isolation from mainland ancestors.

Figure 3: The brown tree snake, Boiga irregularis, is an exotic species that has caused numerous extinctions on the island of Guam since its accidental introduction in 1950. (credit: NPS. "Boiga irregularis" by OpenStax is licensed under CC BY 4.0)

It now appears that the global decline in amphibian species recognized in the 1990s is, in some part, caused by the fungus *Batrachochytrium dendrobatidis*, which causes the disease **chytridiomycosis** (Figure 4). There is evidence that the fungus is native to Africa and may have been spread throughout the world by transport of a commonly used

laboratory and pet species: the African clawed toad (*Xenopus laevis*). It may well be that biologists themselves are responsible for spreading this disease worldwide. The North American bullfrog, *Rana catesbeiana*, which has also been widely introduced as a food animal but which easily escapes captivity, survives most infections of *Batrachochytrium dendrobatidis* and can act as a reservoir for the disease.

Figure 4: This Limosa Harlequin Frog (Atelopus limosus), an endangered species from Panama, died from a fungal disease called chytridiomycosis. The red lesions are symptomatic of the disease. (credit: Brian Gratwicke. "chytridiomycosis" by OpenStax is licensed under CC BY 4.0)

Early evidence suggests that another fungal pathogen, *Geomyces destructans*, introduced from Europe is responsible for **white-nose syndrome**, which infects cave-hibernating bats in eastern North America and has spread from a point of origin in western New York State (Figure 5). The disease has decimated bat populations and threatens the extinction of

species already listed as endangered: the Indiana bat, *Myotis sodalis*, and potentially the Virginia big-eared bat, *Corynorhinus townsendii virginianus*. How the fungus was introduced is unclear, but one logical presumption would be that recreational cavers unintentionally brought the fungus on clothes or equipment from Europe.

Figure 5: This little brown bat in Greeley Mine, Vermont, March 26, 2009, was found to have white-nose syndrome. (credit: Marvin Moriarty, USFWS. "white-nose syndrome" by OpenStax is licensed under CC BY 4.0)

Climate Change

Climate change, and specifically the **anthropogenic** (meaning, caused by humans) warming trend presently underway, is recognized as a major extinction threat, particularly when combined with other threats such as habitat loss. Scientists disagree about the likely magnitude of the effects, with extinction rate estimates ranging from 15 percent to 40 percent of species committed to extinction by 2050. Scientists do agree, however, that climate change will alter regional climates, including rainfall and snowfall patterns, making habitats less hospitable to the species living in them. The warming trend will shift colder climates toward the north and south poles, forcing species to move with their adapted climate norms while facing habitat gaps along the way. The shifting ranges will impose new competitive regimes on species as they find themselves in contact with other species not present in their historic range. One such unexpected species contact is between polar bears and grizzly bears (Figure 6). Previously, these two species had separate ranges. Now, their ranges are overlapping and there are documented cases of these two species mating and producing viable offspring. Changing climates also throw off species' delicate timing adaptations to seasonal food resources and breeding times. Many contemporary mismatches to shifts in resource availability and timing have already been documented.

Figure 6: Since 2008, grizzly bears (Ursus arctos horribilis) have been spotted farther north than their historic range, a possible consequence of climate change. As a result, grizzly bear habitat now overlaps polar bear (Ursus maritimus) habitat. The two kinds of bears, which are capable of mating and producing viable offspring, are considered separate species as historically they lived in different habitats and never met. However, in 2006 a hunter shot a wild grizzly-polar bear

hybrid known as a grolar bear, the first wild hybrid ever found. (credit: "this image" by OpenStax is licensed under CC BY 4.0)

Range shifts are already being observed: for example, some European bird species ranges have moved 91 km northward. The same study suggested that the optimal shift based on warming trends was double that distance, suggesting that the populations are not moving quickly enough. Range shifts have also been observed in plants, butterflies, other insects, freshwater fishes, reptiles, and mammals.

Climate gradients will also move up mountains, eventually crowding species higher in altitude and eliminating the habitat for those species adapted to the highest elevations. Some climates will completely disappear. The rate of warming appears to be accelerated in the arctic, which is recognized as a serious threat to polar bear populations that require sea ice to hunt seals during the winter months: seals are the only source of protein available to polar bears. A trend to decreasing sea ice coverage has occurred since observations began in the mid-twentieth century. The rate of decline observed in recent years is far greater than previously predicted by climate models.

Finally, global warming will **raise ocean levels** due to melt water from glaciers and the greater volume of warmer water. Shorelines will be inundated, reducing island size, which will have an effect on some species, and a number of islands will disappear entirely. Additionally, the gradual melting and subsequent refreezing of the poles, glaciers, and higher elevation mountains—a cycle that has provided freshwater to environments for centuries—will also be jeopardized. This could result in an overabundance of salt water and a shortage of fresh water.

Summary

The core threats to biodiversity are human population growth and unsustainable resource use. To date, the most significant causes of extinctions are habitat loss, introduction of exotic species, and overharvesting. Climate change is predicted to be a significant cause of extinctions in the coming century. Habitat loss occurs through deforestation, damming of rivers, and other activities. Overharvesting is a threat particularly to aquatic species, while the taking of bush meat in the humid tropics threatens many species in Asia, Africa, and the Americas. Exotic species have been the cause of a number of extinctions and are especially damaging to islands and lakes. Exotic species' introductions are increasing because of the increased mobility of human populations and growing global trade and transportation. Climate change is forcing range changes that may lead to extinction. It is also affecting adaptations to the timing of resource availability that negatively affects species in seasonal environments. The impacts of climate change are greatest in the arctic. Global warming will also raise sea levels, eliminating some islands and reducing the area of all others.

REFERENCES

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

Preserving Biodiversity

Preserving biodiversity is an extraordinary challenge that must be met by greater understanding of biodiversity itself, changes in human behavior and beliefs, and various preservation strategies.

Measuring Biodiversity

The technology of molecular genetics and data processing and storage are maturing to the point where cataloguing the planet's species in an accessible way is close to feasible. **DNA barcoding** is one molecular genetic method, which takes advantage of rapid evolution in a mitochondrial gene present in eukaryotes, excepting the plants, to identify species using the sequence of portions of the gene. Plants may be barcoded using a combination of chloroplast genes. Rapid mass sequencing machines make the molecular genetics portion of the work relatively inexpensive and quick. Computer resources store and make available the large volumes of data. Projects are currently underway to use DNA barcoding to catalog museum specimens, which have already been named and studied, as well as testing the method on less studied groups. As of mid 2012, close to 150,000 named species had been barcoded. Early studies suggest there are significant numbers of undescribed species that looked too much like sibling species to previously be recognized as

different, which can now can be identified with DNA barcoding.

Numerous computer databases now provide information about named species and a framework for adding new species. However, as already noted, at the present rate of description of new species, it will take close to 500 years before the complete catalog of life is known. Many, perhaps most, species on the planet do not have that much time.

There is also the problem of understanding which species known to science are threatened and to what degree they are threatened. This task is carried out by the non-profit IUCN which, maintains the **Red List**—an online listing of endangered species categorized by taxonomy, type of threat, and other criteria (Figure 1). The Red List is supported by scientific research. In 2011, the list contained 61,000 species, all with supporting documentation.

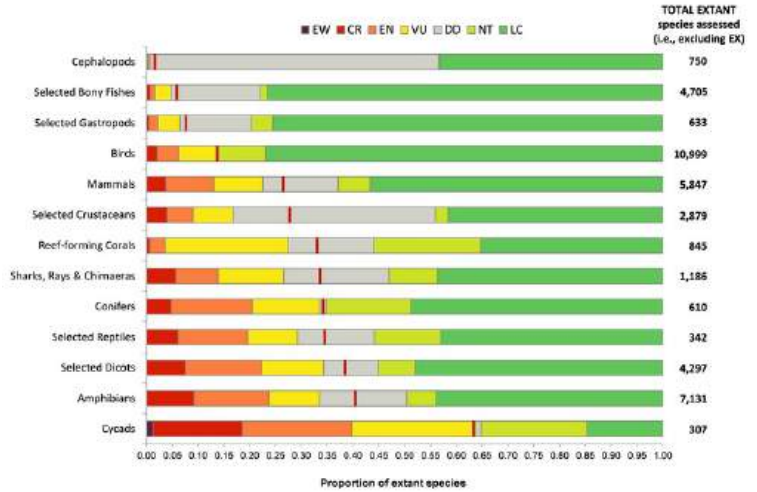


Figure 2. The proportion of extant (i.e., excluding Extinct) species in The IUCN Red List of Threatened Species. Version 2020-3 assessed in each category for the more comprehensively assessed (i.e., at least 80% of the group has been assessed) groups containing ≥150 species. Species are grouped into classes (with the exception of reef-forming corals, which includes species from classes Hydrozoa and Anthozoa), and are ordered according to the vertical red lines, which indicate the best estimate for proportion of extant species considered threatened (CR, EN, or VU). Best estimates of percentage threatened species (with lower and upper estimates) for each group are: cycads 63% (63-64%); amphibians 40% (34-50%); selected dicots (birches; cacti; magnolias; maples; oaks; protea family; southern beeches; teas) 38% (34-45%); selected reptiles (marine turtles; seasnakes; chameleons; crocodiles & alligators) 34% (29-44%); conifers 34% (34-35%); sharks, rays & chimeras 33% (27-47%); reef-forming corals 33% (27-44%); selected crustaceans (lobsters; freshwater crabs; freshwater crayfishes; freshwater shrimps) 27.5% (17-56%); mammals 26% (23-37%); birds 14% (13.5-14%); selected bony fishes (anchovies; angelfishes; billfishes; blennies; bonefishes; butterflyfishes; cornetfishes; croakers and drums; denticle herring; dragonfishes, lightfishes and relatives; filefishes; ghost pipefishes; groupers; gulpers, snipe eels and relatives; jacks, pompanos and relatives; ladyfishes; lanternfishes; lizardfishes and allies;

pristigasterids; pufferfishes; round herrings; sardines and relatives; seabreams, porgies and picarels; seahorses, pipefishes and relatives; shrimpfishes; sturgeons; Sundaland noodlefishes; surgeonfishes, tangs and unicornfishes; swordfish; tarpons; trumpetfishes; tunas; wolf herrings; wrasses) 6% (5-22%); selected gastropods (cone snails) 7.5% (6-20%); cephalopods (nautiluses; octopuses; squids) 1.5% (1-57%). The numbers to the right of each bar represent the total number of extant species assessed for each group. EW – Extinct in the Wild, CR – Critically Endangered, EN – Endangered, VU – Vulnerable, NT – Near Threatened, DD – Data Deficient, LC – Least Concern. Available from: https://www.iucnredlist.org/resources/summary-statistics#Summary%20Tables

Changing Human Behavior

Legislation throughout the world has been enacted to protect species. The legislation includes international treaties as well as national and state laws. The **Convention on International Trade in Endangered Species of Wild Fauna and Flora** (CITES) treaty came into force in 1975. The treaty, and the national legislation that supports it, provides a legal framework for preventing approximately 33,000 listed species from being transported across nations' borders, thus protecting them from being caught or killed when international trade is involved. The treaty is limited in its reach because it only deals with international movement of organisms or their parts. It is also limited by various countries' ability or willingness to enforce the treaty and supporting legislation. The illegal trade in organisms and their parts is probably a market in the hundreds of millions of dollars. Illegal wildlife trade is monitored by another non-profit: **Trade Records Analysis of Flora and Fauna in Commerce** (TRAFFIC).

Within many countries, there are laws that protect endangered species and regulate hunting and fishing. In the

United States, the **Endangered Species Act** (ESA) was enacted in 1973. Species at risk are listed by the Endangered Species Act; the U.S. Fish & Wildlife Service is required by law to develop management plans that protect the listed species and bring them back to sustainable numbers. The Endangered Species Act, and others like it in other countries, is a useful tool, but it suffers because it is often difficult to get a species listed, or to get an effective management plan in place once it is listed. Additionally, species may be controversially taken off the list without necessarily having had a change in their situation. More fundamentally, the approach to protecting individual species rather than entire ecosystems is both inefficient and focuses efforts on a few highly visible and often charismatic species, perhaps at the expense of other species that go unprotected. At the same time, the Endangered Species Act has a critical habitat provision outlined in the recovery mechanism that may benefit species other than the one targeted for management.

The **Migratory Bird Treaty Act** (MBTA) is an agreement between the United States and Canada that was signed into law in 1918 in response to declines in North American bird species caused by hunting. The Migratory Bird Treaty Act now lists over 800 protected species. It makes it illegal to disturb or kill the protected species or distribute their parts (much of the hunting of birds in the past was for their feathers).

The international response to global warming has been mixed. The **Kyoto Protocol**, an international agreement that came out of the United Nations Framework Convention on Climate Change that committed countries to reducing greenhouse gas emissions by 2012, was ratified by some countries, but spurned by others. Two important countries in terms of their potential impact that did not ratify the Kyoto Protocol were the United States and China. The United States rejected it as a result of a powerful fossil fuel industry and China because of a concern it would stifle the nation's

growth. Some goals for reduction in greenhouse gasses were met and exceeded by individual countries, but worldwide, the effort to limit greenhouse gas production is not succeeding. The intended replacement for the Kyoto Protocol has not materialized because governments cannot agree on timelines and benchmarks. Meanwhile, climate scientists predict the resulting costs to human societies and biodiversity will be high.

The private non-profit sector plays a large role in the conservation effort both in North America and around the world. The approaches range from species-specific organizations to the broadly focused IUCN and TRAFFIC. The **Nature Conservancy** takes a novel approach. It purchases land and protects it in an attempt to set up preserves for ecosystems. Ultimately, human behavior will change when human values change. At present, the growing urbanization of the human population is a force that poses challenges to the valuing of biodiversity.

Conservation in Preserves

The establishment of **wildlife and ecosystem preserves** is one of the key tools in conservation efforts. A **preserve** is an area of land set aside with varying degrees of protection for the organisms that exist within the boundaries of the preserve. Preserves can be effective in the short term for protecting both species and ecosystems, but they face challenges that scientists are still exploring to strengthen their viability as long-term solutions.

How Much Area to Preserve?

Due to the way protected lands are allocated (they tend to contain less economically valuable resources rather than being set aside specifically for the species or ecosystems at

risk) and the way biodiversity is distributed, determining a target percentage of land or marine habitat that should be protected to maintain biodiversity levels is challenging. The IUCN World Parks Congress estimated that 11.5 percent of Earth's land surface was covered by preserves of various kinds in 2003. This area is greater than previous goals; however, it only represents 9 out of 14 recognized major biomes. Research has shown that 12 percent of all species live only outside preserves; these percentages are much higher when only threatened species and high quality preserves are considered. For example, high quality preserves include only about 50 percent of threatened amphibian species. The conclusion must be that either the percentage of area protected must increase, or the percentage of high quality preserves must increase, or preserves must be targeted with greater attention to biodiversity protection. Researchers argue that more attention to the latter solution is required.

Preserve Design

There has been extensive research into optimal preserve designs for maintaining biodiversity. The fundamental principle behind much of the research has been the seminal theoretical work of Robert H. MacArthur and Edward O. Wilson published in 1967 on island biogeography (MacArthur & Wilson, 1967). This work sought to understand the factors affecting biodiversity on islands. The fundamental conclusion was that biodiversity on an island was a function of the origin of species through migration, speciation, and extinction on that island. Islands farther from a mainland are harder to get to, so migration is lower and the equilibrium number of species is lower. Within island populations, evidence suggests that the number of species gradually increases to a level similar to the numbers on the mainland from which the

species is suspected to have migrated. In addition, smaller islands are harder to find, so their immigration rates for new species are lower. Smaller islands are also less geographically diverse so there are fewer niches to promote speciation. And finally, smaller islands support smaller populations, so the probability of extinction is higher.

As islands get larger, the number of species accelerates, although the effect of island area on species numbers is not a direct correlation. Conservation preserves can be seen as "islands" of habitat within "an ocean" of non-habitat. For a species to persist in a preserve, the preserve must be large enough. The critical size depends, in part, on the home range that is characteristic of the species. A preserve for wolves, which range hundreds of kilometers, must be much larger than a preserve for butterflies, which might range within ten kilometers during its lifetime. But larger preserves have more core area of optimal habitat for individual species, they have more niches to support more species, and they attract more species because they can be found and reached more easily.

Preserves perform better when there are buffer zones around them of suboptimal habitat. The buffer allows organisms to exit the boundaries of the preserve without immediate negative consequences from predation or lack of resources. One large preserve is better than the same area of several smaller preserves because there is more core habitat unaffected by edges. For this same reason, preserves in the shape of a square or circle will be better than a preserve with many thin "arms." If preserves must be smaller, then providing wildlife corridors between them so that individuals and their genes can move between the preserves, for example along rivers and streams, will make the smaller preserves behave more like a large one. All of these factors are taken into consideration when planning the nature of a preserve before the land is set aside.

In addition to the physical, biological, and ecological

specifications of a preserve, there are a variety of policy, legislative, and enforcement specifications related to uses of the preserve for functions other than the protection of species. These can include anything from timber extraction, mineral extraction, regulated hunting, human habitation, and nondestructive human recreation. Many of these policy decisions are made based on political pressures rather than conservation considerations. In some cases, wildlife protection policies have been so strict that subsistence-living indigenous populations have been forced from ancestral lands that fell within a preserve. In other cases, even if a preserve is designed to protect wildlife, if the protections are not or cannot be enforced, the preserve status will have little meaning in the face of illegal poaching and timber extraction. This is a widespread problem with preserves in areas of the tropics.

Limitations on Preserves

Some of the limitations on preserves as conservation tools are evident from the discussion of preserve design. Political and economic pressures typically make preserves smaller, never larger, so setting aside areas that are large enough is difficult. If the area set aside is sufficiently large, there may not be sufficient area to create a buffer around the preserve. In this case, an area on the outer edges of the preserve inevitably becomes a riskier suboptimal habitat for the species in the preserve. Enforcement of protections is also a significant issue in countries without the resources or political will to prevent poaching and illegal resource extraction.

Climate change will create inevitable problems with the location of preserves. The species within them will migrate to higher latitudes as the habitat of the preserve becomes less favorable. Scientists are planning for the effects of global

warming on future preserves and striving to predict the need for new preserves to accommodate anticipated changes to habitats; however, the end effectiveness is tenuous since these efforts are prediction based.

Finally, an argument can be made that conservation preserves reinforce the cultural perception that humans are separate from nature, that humans can exist outside of nature, and that humans can only operate in ways that damage biodiversity. There is concern that creating wildlife preserves reduces the pressure on human activities outside the preserves to be sustainable and non-damaging to biodiversity. Ultimately, the political, economic, and human demographic pressures will degrade and reduce the size of conservation preserves if the activities outside them are not altered to be less damaging to biodiversity.

Habitat Restoration

Habitat restoration holds considerable promise as a mechanism for restoring and maintaining biodiversity. Of course, once a species has become extinct, its restoration is impossible. However, restoration can improve the biodiversity of **degraded ecosystems**. Reintroducing wolves, a top predator, to Yellowstone National Park in 1995 led to dramatic changes in the ecosystem that increased biodiversity. The wolves (Figure 2) function to suppress elk and coyote populations and provide more abundant resources to the guild of carrion eaters. Reducing elk populations has allowed revegetation of riparian areas, which has increased the diversity of species in that habitat. Decreasing the coyote population has increased the populations of species that were previously suppressed by this predator. The number of species of carrion eaters has increased because of the predatory activities of the wolves. In this habitat, the wolf is a keystone species, meaning a

species that is instrumental in maintaining diversity in an ecosystem. Removing a **keystone species** from an ecological community may cause a collapse in diversity. The results from the Yellowstone experiment suggest that restoring a keystone species can have the effect of restoring biodiversity in the community. Ecologists have argued for the identification of keystone species where possible and for focusing protection efforts on those species; likewise, it also makes sense to attempt to return them to their ecosystem if they have been removed.

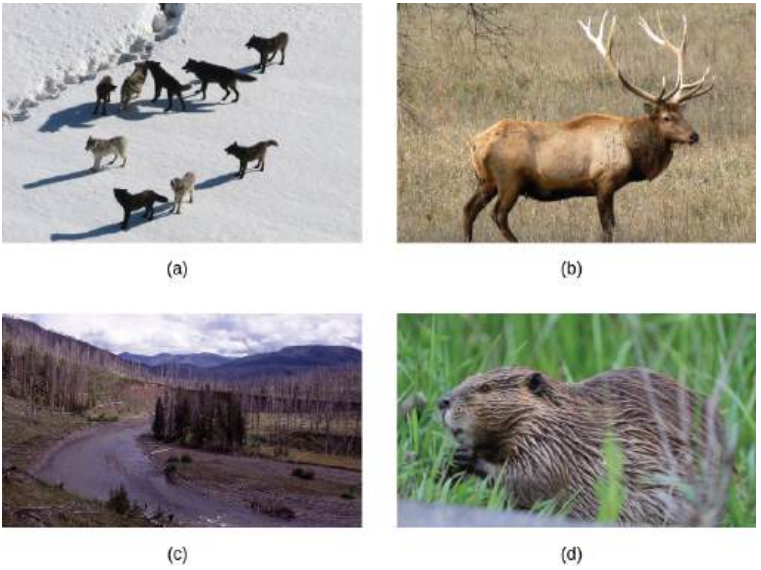


Figure 2: (a) The Gibbon wolf pack in Yellowstone National Park, March 1, 2007, represents a keystone species. The reintroduction of wolves into Yellowstone National Park in 1995 led to a change in the grazing behavior of (b) elk. To avoid predation, the elk no longer grazed exposed stream and riverbeds, such as (c) the Lamar Riverbed in Yellowstone. This allowed willow and cottonwood seedlings to grow. The seedlings decreased erosion and provided shading to the creek, which improved fish habitat. A new colony of (d) beaver may also have benefited from the habitat change. (credit a: modification of work by Doug Smith, NPS; credit c: modification of work by Jim Peaco, NPS; credit d: modification of work by "Shiny Things"/Flickr. "this image" by OpenStax is licensed under CC BY 4.0)

Other large-scale restoration experiments underway involve **dam removal**. In the United States, since the mid-1980s, many aging dams are being considered for removal rather than replacement because of shifting beliefs about the ecological value of free-flowing rivers and because many dams no longer provide the benefit and functions that they did when they were first built. The measured benefits of dam

removal include restoration of naturally fluctuating water levels (the purpose of dams is frequently to reduce variation in river flows), which leads to increased fish diversity and improved water quality. In the Pacific Northwest, dam removal projects are expected to increase populations of salmon, which is considered a keystone species because it transports key nutrients to inland ecosystems during its annual spawning migrations. In other regions such as the Atlantic coast, dam removal has allowed the return of spawning anadromous fish species (species that are born in fresh water, live most of their lives in salt water, and return to fresh water to spawn). Some of the largest dam removal projects have yet to occur or have happened too recently for the consequences to be measured. The large-scale ecological experiments that these removal projects constitute will provide valuable data for other dam projects slated either for removal or construction.

The Role of Captive Breeding

Zoos have sought to play a role in conservation efforts both through captive breeding programs and education. The transformation of the missions of zoos from collection and exhibition facilities to organizations that are dedicated to conservation is ongoing. In general, it has been recognized that, except in some specific targeted cases, **captive breeding programs** for endangered species are inefficient and often prone to failure when the species are reintroduced to the wild. Zoo facilities are far too limited to contemplate captive breeding programs for the numbers of species that are now at risk. Education is another potential positive impact of zoos on conservation efforts, particularly given the global trend to urbanization and the consequent reduction in contacts between people and wildlife. A number of studies have been performed to look at the effectiveness of zoos

on people's attitudes and actions regarding conservation; at present, the results tend to be mixed.

Summary

New technological methods such as DNA barcoding and information processing and accessibility are facilitating the cataloging of the planet's biodiversity. There is also a legislative framework for biodiversity protection. International treaties such as CITES regulate the transportation of endangered species across international borders. Legislation within individual countries protecting species and agreements on global warming have had limited success; there is at present no international agreement on targets for greenhouse gas emissions. In the United States, the Endangered Species Act protects listed species but is hampered by procedural difficulties and a focus on individual species. The Migratory Bird Act is an agreement between Canada and the United States to protect migratory birds. The non-profit sector is also very active in conservation efforts in a variety of ways.

Conservation preserves are a major tool in biodiversity protection. Presently, 11percent of Earth's land surface is protected in some way. The science of island biogeography has informed the optimal design of preserves; however, preserves have limitations imposed by political and economic forces. In addition, climate change will limit the effectiveness of preserves in the future. A downside of preserves is that they may lessen the pressure on human societies to function more sustainably outside the preserves.

Habitat restoration has the potential to restore ecosystems to previous biodiversity levels before species become extinct. Examples of restoration include reintroduction of keystone species and removal of dams on rivers. Zoos have attempted to take a more active role in conservation and can have a

limited role in captive breeding programs. Zoos also may have a useful role in education.

REFERENCES

MacArthur, R. and Wilson, E. O. 1967. *The Theory of Island Biogeography* (Princeton, N.J.: Princeton University Press)

OpenStax, Biology. OpenStax CNX. June 26, 2020. https://cnx.org/contents/GFy_h8cu@10.137:noBcfThl@7/Understanding-Evolution.

IUCN Red List of Threatened Species. Version 2020-3. [Cited Feb 22, 2021]. Available from https://www.iucnredlist.org/resources/summary-statistics#Summary%20Tables