

Molecular Genetics: From DNA to Proteins

Jean Brainard, Ph.D.

Say Thanks to the Authors

Click <http://www.ck12.org/saythanks>

(No sign in required)

ck-12 **flexbook**
next generation textbooks

To access a customizable version of this book, as well as other interactive content, visit www.ck12.org

CK-12 Foundation is a non-profit organization with a mission to reduce the cost of textbook materials for the K-12 market both in the U.S. and worldwide. Using an open-source, collaborative, and web-based compilation model, CK-12 pioneers and promotes the creation and distribution of high-quality, adaptive online textbooks that can be mixed, modified and printed (i.e., the FlexBook® textbooks).

Copyright © 2015 CK-12 Foundation, www.ck12.org

The names “CK-12” and “CK12” and associated logos and the terms “**FlexBook®**” and “**FlexBook Platform®**” (collectively “CK-12 Marks”) are trademarks and service marks of CK-12 Foundation and are protected by federal, state, and international laws.

Any form of reproduction of this book in any format or medium, in whole or in sections must include the referral attribution link <http://www.ck12.org/saythanks> (placed in a visible location) in addition to the following terms.

Except as otherwise noted, all CK-12 Content (including CK-12 Curriculum Material) is made available to Users in accordance with the Creative Commons Attribution-Non-Commercial 3.0 Unported (CC BY-NC 3.0) License (<http://creativecommons.org/licenses/by-nc/3.0/>), as amended and updated by Creative Commons from time to time (the “CC License”), which is incorporated herein by this reference.

Complete terms can be found at <http://www.ck12.org/about/terms-of-use>.

Printed: October 9, 2015

flexbook
next generation textbooks



AUTHOR

Jean Brainard, Ph.D.

CONTRIBUTOR

Douglas Wilkin, Ph.D.

CHAPTER

1

Molecular Genetics: From DNA to Proteins

CHAPTER OUTLINE

- 1.1 DNA and RNA
- 1.2 Protein Synthesis
- 1.3 Mutation
- 1.4 Regulation of Gene Expression
- 1.5 References



The spiral structure in the picture is a large organic molecule. Can you guess what it is? Here's a hint: molecules like this one determine who you are. They contain genetic information that controls your characteristics. They determine your eye color, facial features, and other physical attributes. What molecule is it?

You probably answered "DNA." Today, it is commonly known that DNA is the genetic material. For a long time, scientists knew such molecules existed. They were aware that genetic information was contained within organic molecules. However, they didn't know which type of molecules play this role. In fact, for many decades, scientists thought that proteins were the molecules that carry genetic information. In this chapter, you will learn how scientists discovered that DNA carries the code of life.

User:PublicDomainPictures/Pixabay. pixabay.com/en/dna-biology-medicine-gene-163466/. Public Domain.

1.1 DNA and RNA

Lesson Objectives

- State the central dogma of molecular biology.
- Outline discoveries that led to knowledge of DNA's structure and function.
- Describe the structure of RNA, and identify the three main types of RNA.

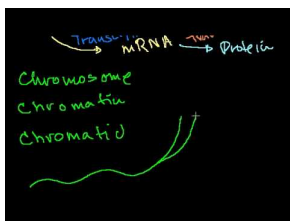
Vocabulary

- central dogma of molecular biology
- Chargaff's rules
- messenger RNA (mRNA)
- ribosomal RNA (rRNA)
- transfer RNA (tRNA)

Introduction

Your DNA, or deoxyribonucleic acid, contains the genes that determine who you are. How can this organic molecule control your characteristics? DNA contains instructions for all the proteins your body makes. Proteins, in turn, determine the structure and function of all your cells. What determines a protein's structure? It begins with the sequence of amino acids that make up the protein. Instructions for making proteins with the correct sequence of amino acids are encoded in DNA.

The vocabulary of DNA: chromosomes, chromatids, chromatin, transcription, translation, and replication is discussed at <http://www.youtube.com/watch?v=s9HPNwXd9fk> (18:23).



MEDIA

Click image to the left or use the URL below.

URL: <http://www.ck12.org/flx/render/embeddedobject/269>

Central Dogma of Molecular Biology

DNA is found in chromosomes. In eukaryotic cells, chromosomes always remain in the nucleus, but proteins are made at ribosomes in the cytoplasm. How do the instructions in DNA get to the site of protein synthesis outside

the nucleus? Another type of nucleic acid is responsible. This nucleic acid is RNA, or ribonucleic acid. RNA is a small molecule that can squeeze through pores in the nuclear membrane. It carries the information from DNA in the nucleus to a ribosome in the cytoplasm and then helps assemble the protein. In short:

DNA → RNA → Protein

Discovering this sequence of events was a major milestone in molecular biology. It is called the **central dogma of molecular biology**. You can watch a video about the central dogma and other concepts in this lesson at this link: http://www.youtube.com/watch?v=ZjRCmU0_dhY (8:07). An overview of protein synthesis can be viewed at http://www.youtube.com/watch?v=-ygpqVr7_xs (10:46).

DNA

DNA is the genetic material in your cells. It was passed on to you from your parents and determines your characteristics. The discovery that DNA is the genetic material was another important milestone in molecular biology.

Griffith Searches for the Genetic Material

Many scientists contributed to the identification of DNA as the genetic material. In the 1920s, Frederick Griffith made an important discovery. He was studying two different strains of a bacterium, called R (rough) strain and S (smooth) strain. He injected the two strains into mice. The S strain killed (virulent) the mice, but the R strain did not (nonvirulent) (see **Figure 1.1**). Griffith also injected mice with S-strain bacteria that had been killed by heat. As expected, the killed bacteria did not harm the mice. However, when the dead S-strain bacteria were mixed with live R-strain bacteria and injected, the mice died.

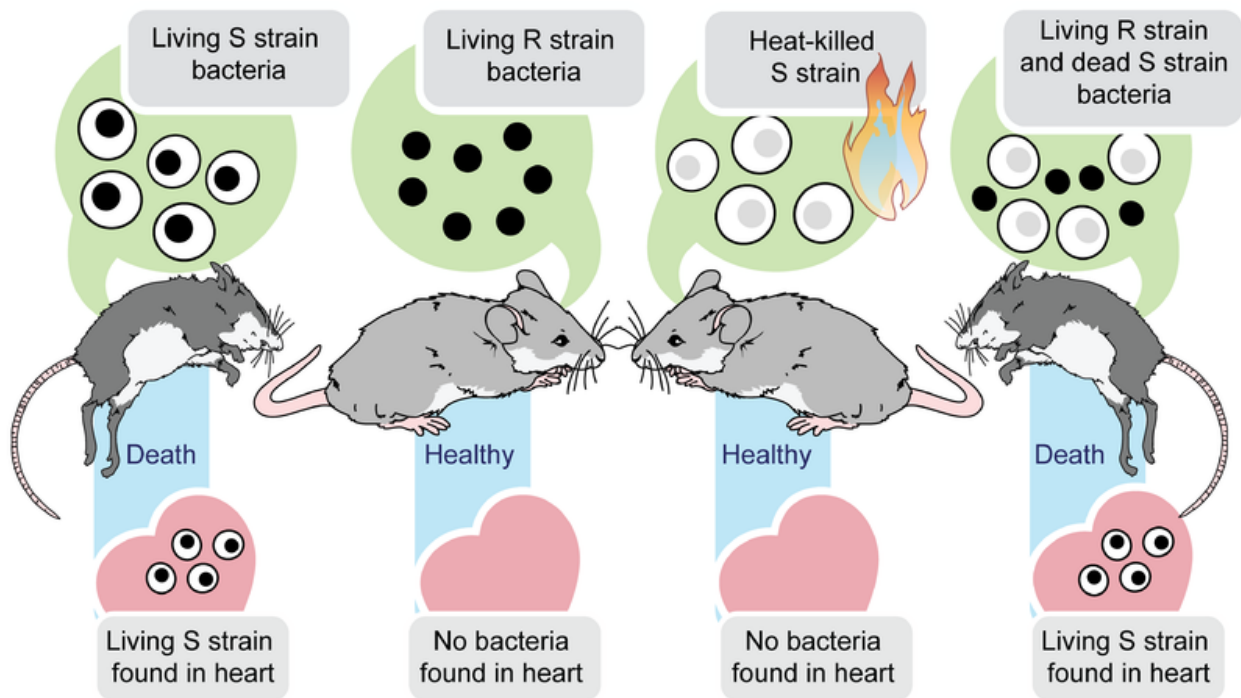
Based on his observations, Griffith deduced that something in the killed S-strain was transferred to the previously harmless R-strain, making the R-strain deadly. What was that something? What type of substance could change the characteristics of the organism that received it?

Avery's Team Makes a Major Contribution

In the early 1940s, a team of scientists led by Oswald Avery tried to answer the question raised by Griffith's results. They inactivated various substances in the S-strain bacteria. They then killed the S-strain bacteria and mixed the remains with live R-strain bacteria. (Keep in mind, the R-strain bacteria usually did not harm the mice.) When they inactivated proteins, the R-strain was deadly to the injected mice. This ruled out proteins as the genetic material. Why? Even without the S-strain proteins, the R-strain was changed, or transformed, into the deadly strain. However, when the researchers inactivated DNA in the S-strain, the R-strain remained harmless. This led to the conclusion that DNA is the substance that controls the characteristics of organisms. In other words, DNA is the genetic material. You can watch an animation about the research of both Griffith and Avery at this link: <http://www.dnalc.org/view/16375-Animation-17-A-gene-is-made-of-DNA-.html> .

Hershey and Chase Seal the Deal

The conclusion that DNA is the genetic material was not widely accepted at first. It had to be confirmed by other research. In the 1950s, Alfred Hershey and Martha Chase did experiments with viruses and bacteria. Viruses are not cells. They are basically DNA inside a protein coat. To reproduce, a virus must insert its own genetic material into a cell (such as a bacterium). Then it uses the cell's machinery to make more viruses. The researchers used different radioactive elements to label the DNA and proteins in viruses. This allowed them to identify which molecule the

**FIGURE 1.1**

Griffith's Experimental Results. Griffith showed that a substance could be transferred to harmless bacteria and make them deadly.

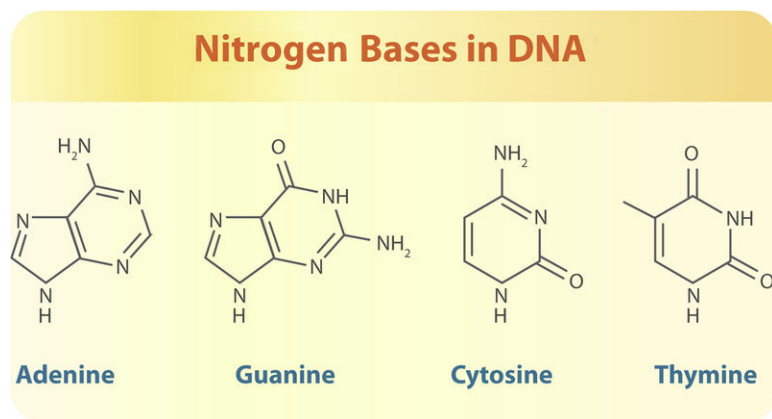
viruses inserted into bacteria. DNA was the molecule they identified. This confirmed that DNA is the genetic material.

Chargaff Writes the Rules

Other important discoveries about DNA were made in the mid-1900s by Erwin Chargaff. He studied DNA from many different species. He was especially interested in the four different nitrogen bases of DNA: adenine (A), guanine (G), cytosine (C), and thymine (T) (see **Figure 1.2**). Chargaff found that concentrations of the four bases differed from one species to another. However, within each species, the concentration of adenine was always about the same as the concentration of thymine. The same was true of the concentrations of guanine and cytosine. These observations came to be known as **Chargaff's rules**. The significance of the rules would not be revealed until the structure of DNA was discovered.

The Double Helix

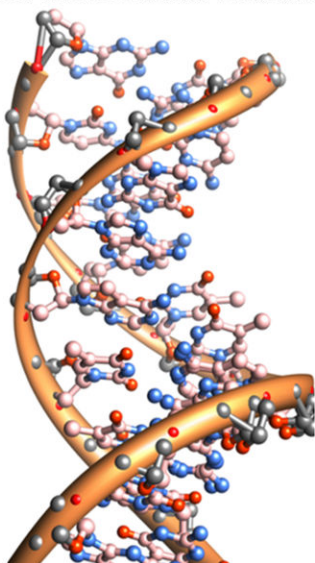
After DNA was found to be the genetic material, scientists wanted to learn more about it. James Watson and Francis Crick are usually given credit for discovering that DNA has a double helix shape, like a spiral staircase (see **Figure 1.3**). The discovery was based on the prior work of Rosalind Franklin and other scientists, who had used X rays to learn more about DNA's structure. Franklin and these other scientists have not always been given credit for their contributions. You can learn more about Franklin's work by watching the video at this link: <http://www.youtube.com>

**FIGURE 1.2**

Nitrogen Bases in DNA. The DNA of all species has the same four nitrogen bases.

[om/watch?v=s3whouvZYG8](https://www.youtube.com/watch?v=s3whouvZYG8) (7:47).

DNA Double Helix



Spiral Staircase

**FIGURE 1.3**

The DNA molecule has a double helix shape. This is the same basic shape as a spiral staircase. Do you see the resemblance? Which parts of the DNA molecule are like the steps of the spiral staircase?

The double helix shape of DNA, together with Chargaff's rules, led to a better understanding of DNA. DNA, as a nucleic acid, is made from nucleotide monomers, and the DNA double helix consists of two polynucleotide chains. Each nucleotide consists of a sugar (deoxyribose), a phosphate group, and a nitrogen-containing base (A, C, G, or T). The sugar-phosphate backbone of the double helix was discussed in the *Chemistry of Life* chapter.

Scientists concluded that bonds (hydrogen bonds) between complementary bases hold together the two polynucleotide chains of DNA. Adenine always bonds with its complementary base, thymine. Cytosine always bonds with its complementary base, guanine. If you look at the nitrogen bases in **Figure 1.2**, you will see why. Adenine and guanine have a two-ring structure. Cytosine and thymine have just one ring. If adenine were to bind with guanine and cytosine with thymine, the distance between the two DNA chains would be variable. However, when a one-ring molecule binds with a two-ring molecule, the distance between the two chains is kept constant. This maintains the uniform shape of the DNA double helix. These *base pairs* (A-T or G-C) stick into the middle of the double helix, forming, in essence, the steps of the spiral staircase.

DNA Replication

Knowledge of DNA's structure helped scientists understand how DNA replicates. DNA replication is the process in which DNA is copied. It occurs during the synthesis (S) phase of the eukaryotic cell cycle. DNA replication begins when an enzyme breaks the bonds between complementary bases in DNA (see **Figure 1.4**). This exposes the bases inside the molecule so they can be "read" by another enzyme and used to build two new DNA strands with complementary bases. The two daughter molecules that result each contain one strand from the parent molecule and one new strand that is complementary to it. As a result, the two daughter molecules are both identical to the parent molecule.

The process of DNA replication is actually much more complex than this simple summary. You can see a detailed animation of the process at this link: <http://www.youtube.com/watch?v=-mtLXpgjHL0> (2:05).

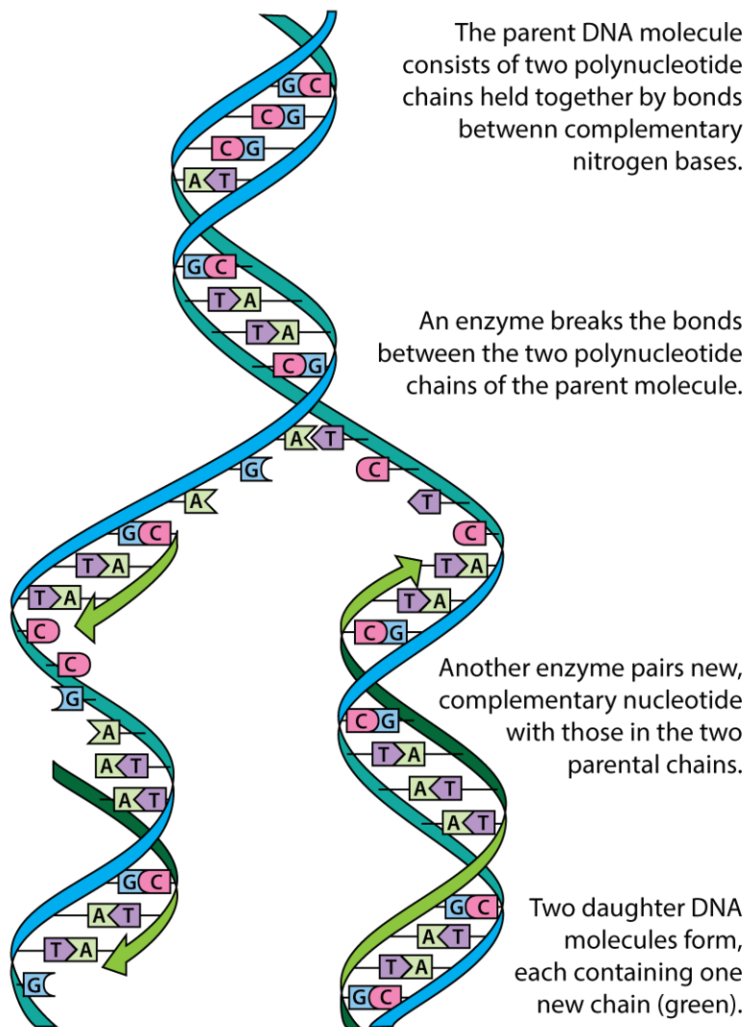


FIGURE 1.4

DNA Replication. DNA replication is a semi-conservative process. Half of the parent DNA molecule is conserved in each of the two daughter DNA molecules.

RNA

DNA alone cannot "tell" your cells how to make proteins. It needs the help of RNA, the other main player in the central dogma of molecular biology. Remember, DNA "lives" in the nucleus, but proteins are made on the ribosomes

in the cytoplasm. How does the genetic information get from the nucleus to the cytoplasm? RNA is the answer.

RNA vs. DNA

RNA, like DNA, is a nucleic acid. However, RNA differs from DNA in several ways. In addition to being smaller than DNA, RNA also

- consists of one nucleotide chain instead of two,
- contains the nitrogen base uracil (U) instead of thymine,
- contains the sugar ribose instead of deoxyribose.

Types of RNA

There are three main types of RNA, all of which are involved in making proteins.

1. **Messenger RNA (mRNA)** copies the genetic instructions from DNA in the nucleus, and carries them to the cytoplasm.
2. **Ribosomal RNA (rRNA)** helps form ribosomes, where proteins are assembled.
3. **Transfer RNA (tRNA)** brings amino acids to ribosomes, where they are joined together to form proteins.

In the next lesson, you can read in detail how these three types of RNA help cells make proteins.

Lesson Summary

- The central dogma of molecular biology states that DNA contains instructions for making a protein, which are copied by RNA. RNA then uses the instructions to make a protein. In short: $DNA \rightarrow RNA \rightarrow Protein$.
- The work of several researchers led to the discovery that DNA is the genetic material. Other researchers discovered that DNA has a double helix shape, consisting of two polynucleotide chains held together by bonds between complementary bases.
- RNA differs from DNA in several ways. There three main types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). Each type plays a different in role in making proteins.

Lesson Review Questions

Recall

1. State the central dogma of molecular biology.
2. Outline research that determined that DNA is the genetic material.
3. What are Chargaff's rules?
4. Identify the structure of the DNA molecule.
5. Why is DNA replication said to be semi-conservative?

Apply Concepts

6. Create a diagram that shows how DNA replication occurs.

Think Critically

7. Explain why complementary base pairing is necessary to maintain the double helix shape of the DNA molecule.
8. Compare and contrast DNA and RNA.

Points to Consider

All three types of RNA are needed by cells to make proteins.

- Can you develop a model in which the three types of RNA interact to make a protein?
- How do you think mRNA copies the genetic instructions in DNA? How are these instructions encoded in the DNA molecule?

1.2 Protein Synthesis

Lesson Objectives

- Give an overview of transcription.
- Describe the genetic code.
- Explain how translation occurs.

Vocabulary

- codon
- genetic code
- promoter
- protein synthesis
- transcription
- translation

Introduction

The process in which cells make proteins is called **protein synthesis**. It actually consists of two processes: transcription and translation. Transcription takes place in the nucleus. It uses DNA as a template to make an RNA molecule. RNA then leaves the nucleus and goes to a ribosome in the cytoplasm, where translation occurs. Translation reads the genetic code in mRNA and makes a protein.

Transcription

Transcription is the first part of the central dogma of molecular biology: **DNA** → **RNA**. It is the transfer of genetic instructions in DNA to mRNA. During transcription, a strand of mRNA is made that is complementary to a strand of DNA. **Figure 1.5** shows how this occurs. You can watch an animation of the process at this link: http://www.biovideo.com/d_%20Transcription.htm .

- A detailed video about transcription is available at this link: <http://vcell.ndsu.edu/animations/transcription/movie-flash.htm> .

Steps of Transcription

Transcription takes place in three steps: initiation, elongation, and termination. The steps are illustrated in **Figure 1.6**.

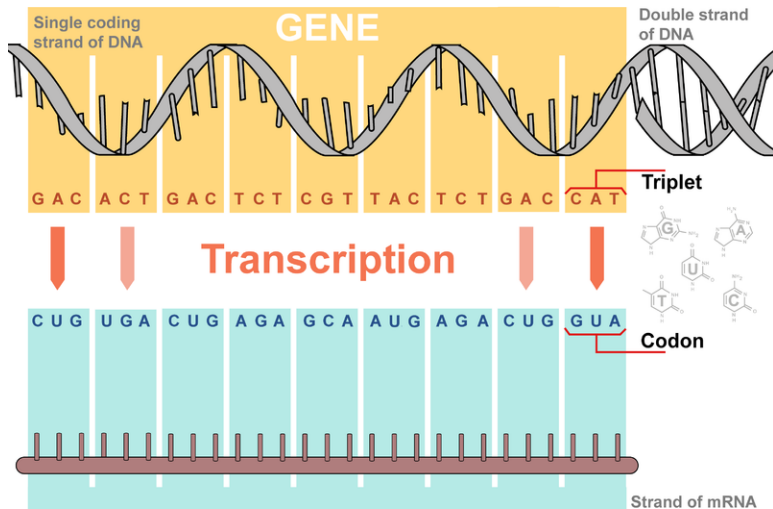


FIGURE 1.5

Overview of Transcription. Transcription uses the sequence of bases in a strand of DNA to make a complementary strand of mRNA. Triplets are groups of three successive nucleotide bases in DNA. Codons are complementary groups of bases in mRNA.

1. Initiation is the beginning of transcription. It occurs when the enzyme RNA polymerase binds to a region of a gene called the **promoter**. This signals the DNA to unwind so the enzyme can “read” the bases in one of the DNA strands. The enzyme is ready to make a strand of mRNA with a complementary sequence of bases.
2. Elongation is the addition of nucleotides to the mRNA strand.
3. Termination is the ending of transcription. The mRNA strand is complete, and it detaches from DNA.

Processing mRNA

In eukaryotes, the new mRNA is not yet ready for translation. It must go through more processing before it leaves the nucleus. This may include splicing, editing, and polyadenylation. These processes modify the mRNA in various ways. Such modifications allow a single gene to be used to make more than one protein.

- Splicing removes introns from mRNA (see **Figure 1.7**). Introns are regions that do not code for proteins. The remaining mRNA consists only of regions that do code for proteins, which are called exons. You can watch a video showing splicing in more detail at this link: <http://vcell.ndsu.edu/animations/mrnasplicing/movie-flash.htm>. Ribonucleoproteins are nucleoproteins that contains RNA. Small nuclear ribonucleoproteins are involved in pre-mRNA splicing.
- Editing changes some of the nucleotides in mRNA. For example, the human protein called APOB, which helps transport lipids in the blood, has two different forms because of editing. One form is smaller than the other because editing adds a premature stop signal in mRNA.
- Polyadenylation adds a “tail” to the mRNA. The tail consists of a string of As (adenine bases). It signals the end of mRNA. It is also involved in exporting mRNA from the nucleus. In addition, the tail protects mRNA from enzymes that might break it down.

The Genetic Code

How is the information in a gene encoded? The answer is the genetic code. The **genetic code** consists of the sequence of nitrogen bases—A, C, G, T (or U)—in a polynucleotide chain. The four bases make up the “letters”

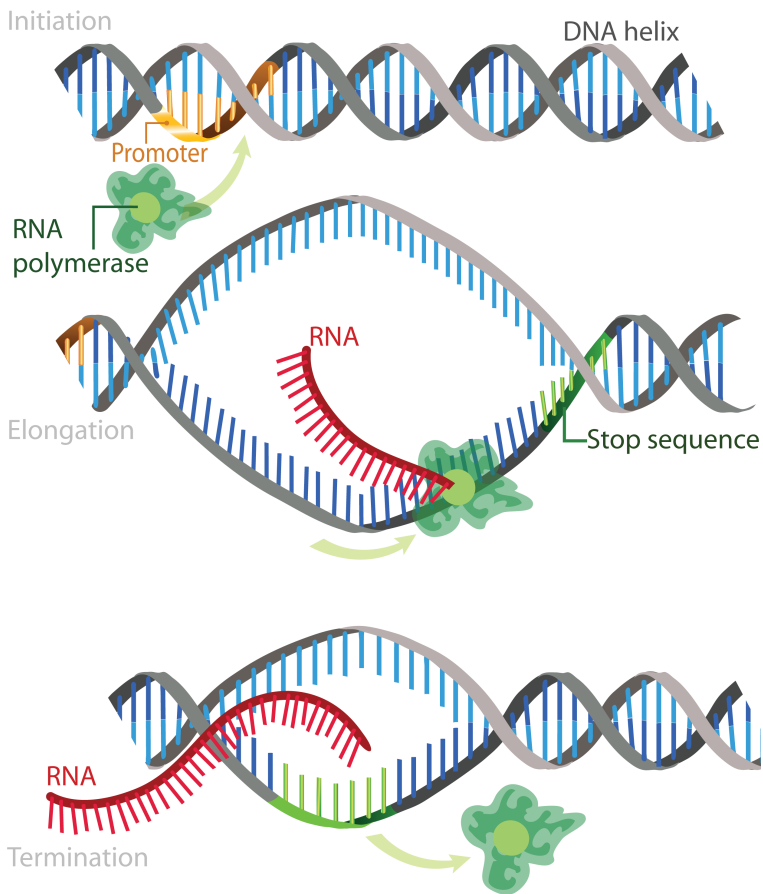


FIGURE 1.6

Steps of Transcription. Transcription occurs in the three steps - initiation, elongation, and termination - shown here.

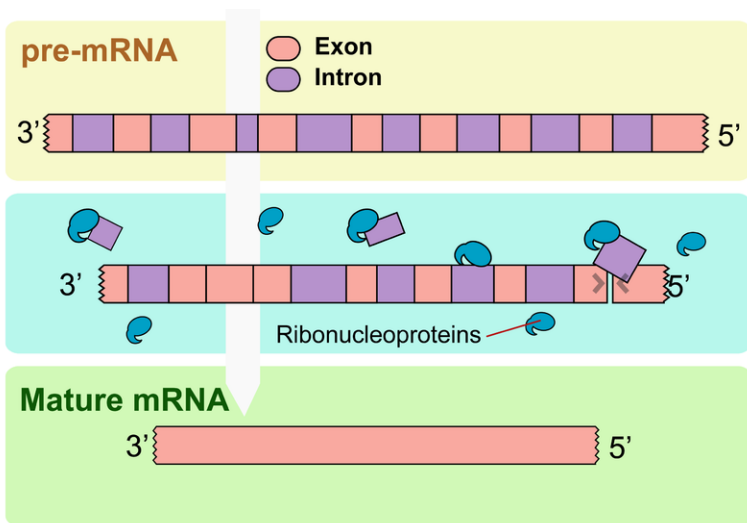


FIGURE 1.7

Splicing. Splicing removes introns from mRNA. UTR is an untranslated region of the mRNA.

of the genetic code. The letters are combined in groups of three to form code “words,” called **codons**. Each codon stands for (encodes) one amino acid, unless it codes for a start or stop signal.

There are 20 common amino acids in proteins. There are 64 possible codons, more than enough to code for the 20 amino acids. The genetic code is shown in **Figure 1.8**. To see how scientists cracked the genetic code, go to this link: <http://www.dnalc.org/view/16494-Animation-22-DNA-words-are-three-letters-long-.html> .

		2 nd base			
		U	C	A	G
1 st base	U	UUU (Phe/F) Phenylalanine	UCU (Ser/S) Serine	UAU (Tyr/Y) Tyrosine	UGU (Cys/C) Cysteine
		UUC (Phe/F) Phenylalanine	UCC (Ser/S) Serine	UAC (Tyr/Y) Tyrosine	UGC (Cys/C) Cysteine
		UUA (Leu/L) Leucine	UCA (Ser/S) Serine	UAA Ochre (Stop)	UGA Opal (Stop)
		UUG (Leu/L) Leucine	UCG (Ser/S) Serine	UAG Amber (Stop)	UGG (Trp/W) Tryptophan
	C	CUU (Leu/L) Leucine	CCU (Pro/P) Proline	CAU (His/H) Histidine	CGU (Arg/R) Arginine
		CUC (Leu/L) Leucine	CCC (Pro/P) Proline	CAC (His/H) Histidine	CGC (Arg/R) Arginine
		CUA (Leu/L) Leucine	CCA (Pro/P) Proline	CAA (Gln/Q) Glutamine	CGA (Arg/R) Arginine
		CUG (Leu/L) Leucine	CCG (Pro/P) Proline	CAG (Gln/Q) Glutamine	CGG (Arg/R) Arginine
	A	AUU (Ile/I) Isoleucine	ACU (Thr/T) Threonine	AAU (Asn/N) Asparagine	AGU (Ser/S) Serine
		AUC (Ile/I) Isoleucine	ACC (Thr/T) Threonine	AAC (Asn/N) Asparagine	AGC (Ser/S) Serine
		AUA (Ile/I) Isoleucine	ACA (Thr/T) Threonine	AAA (Lys/K) Lysine	AGA (Arg/R) Arginine
		AUG ^(A) (Met/M) Methionine	ACG (Thr/T) Threonine	AAG (Lys/K) Lysine	AGG (Arg/R) Arginine
	G	GUU (Val/V) Valine	GCU (Ala/A) Alanine	GAU (Asp/D) Aspartic acid	GGU (Gly/G) Glycine
		GUC (Val/V) Valine	GCC (Ala/A) Alanine	GAC (Asp/D) Aspartic acid	GGC (Gly/G) Glycine
		GUA (Val/V) Valine	GCA (Ala/A) Alanine	GAA (Glu/E) Glutamic acid	GGA (Gly/G) Glycine
		GUG (Val/V) Valine	GCG (Ala/A) Alanine	GAG (Glu/E) Glutamic acid	GGG (Gly/G) Glycine

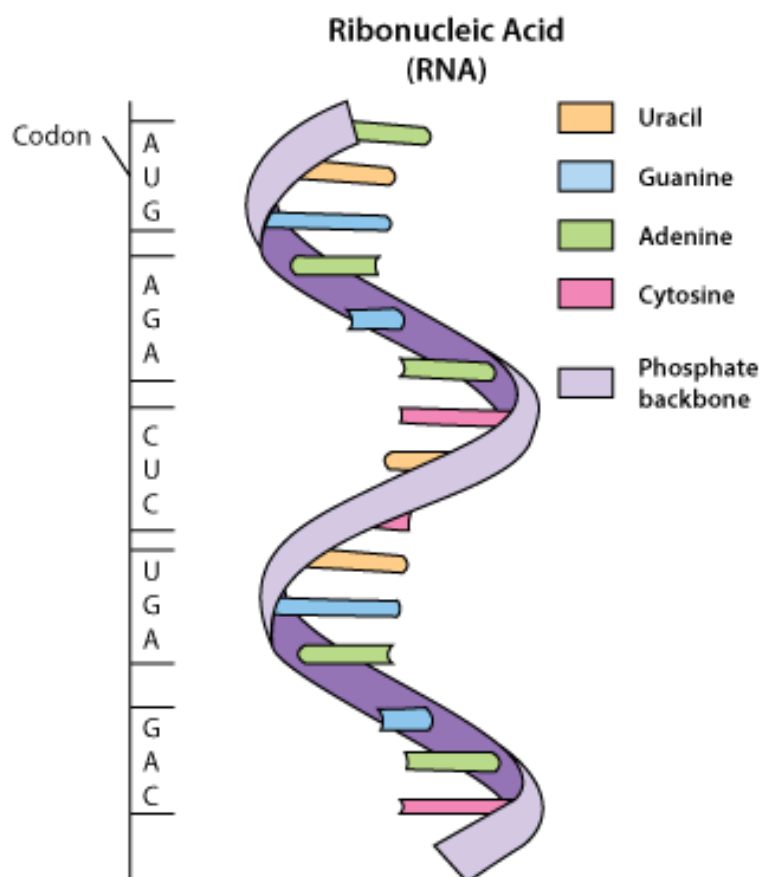
nonpolar	polar	basic	acidic	(stop codon)
----------	-------	-------	--------	--------------

FIGURE 1.8

The Genetic Code. To find the amino acid for a particular codon, find the cell in the table for the first and second bases of the codon. Then, within that cell, find the codon with the correct third base. For example CUG codes for leucine, AAG codes for lysine, and GGG codes for glycine.

Reading the Genetic Code

As shown in **Figure 1.8**, the codon AUG codes for the amino acid methionine. This codon is also the start codon that begins translation. The start codon establishes the reading frame of mRNA. The reading frame is the way the letters are divided into codons. After the AUG start codon, the next three letters are read as the second codon. The next three letters after that are read as the third codon, and so on. This is illustrated in **Figure 1.9**. The mRNA molecule is read, codon by codon, until a stop codon is reached. UAG, UGA, and UAA are all stop codons. They do not code for any amino acids.

**FIGURE 1.9**

Reading the Genetic Code. The genetic code is read three bases at a time. Codons are the code words of the genetic code. Which amino acid does codon 2 in the drawing stand for?

Characteristics of the Genetic Code

The genetic code has a number of important characteristics.

- The genetic code is universal. All known living things have the same genetic code. This shows that all organisms share a common evolutionary history.
- The genetic code is unambiguous. Each codon codes for just one amino acid (or start or stop). What might happen if codons encoded more than one amino acid?
- The genetic code is redundant. Most amino acids are encoded by more than one codon. In **Figure 1.8**, how many codons code for the amino acid threonine? What might be an advantage of having more than one codon for the same amino acid?

Translation

Translation is the second part of the central dogma of molecular biology: **RNA** → **Protein**. It is the process in which the genetic code in mRNA is read to make a protein. **Figure 1.10** shows how this happens. After mRNA leaves the nucleus, it moves to a ribosome, which consists of rRNA and proteins. The ribosome reads the sequence of codons in mRNA. Molecules of tRNA bring amino acids to the ribosome in the correct sequence.

To understand the role of tRNA, you need to know more about its structure. Each tRNA molecule has an anticodon for the amino acid it carries. An anticodon is complementary to the codon for an amino acid. For example, the amino

acid lysine has the codon AAG, so the anticodon is UUC. Therefore, lysine would be carried by a tRNA molecule with the anticodon UUC. Wherever the codon AAG appears in mRNA, a UUC anticodon of tRNA temporarily binds. While bound to mRNA, tRNA gives up its amino acid. Bonds form between the amino acids as they are brought one by one to the ribosome, forming a polypeptide chain. The chain of amino acids keeps growing until a stop codon is reached. To see how this happens, go the link below. <http://www.youtube.com/watch?v=B6O6uRb1D38> (1:29)

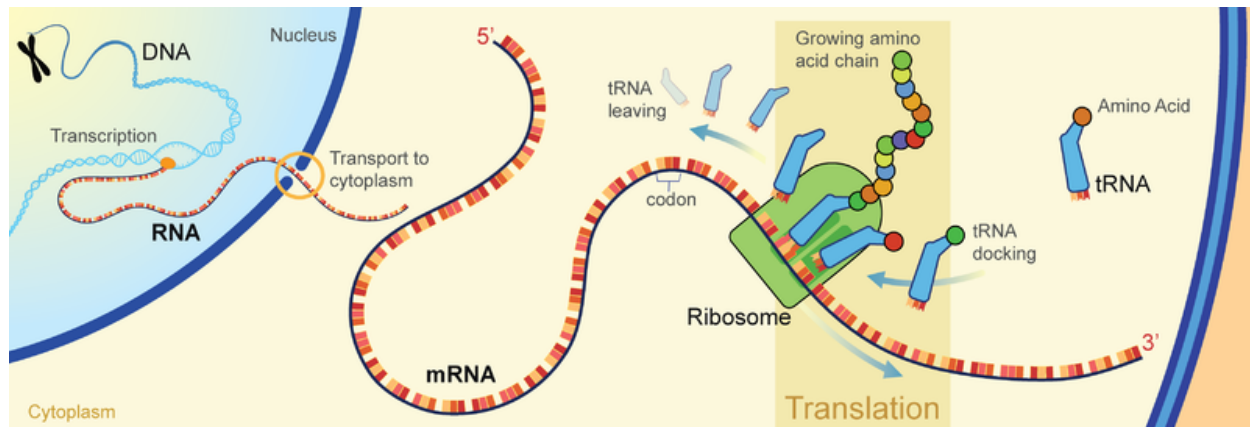


FIGURE 1.10

Translation. Translation of the codons in mRNA to a chain of amino acids occurs at a ribosome. Find the different types of RNA in the diagram. What are their roles in translation?

After a polypeptide chain is synthesized, it may undergo additional processes. For example, it may assume a folded shape due to interactions among its amino acids. It may also bind with other polypeptides or with different types of molecules, such as lipids or carbohydrates. Many proteins travel to the Golgi apparatus to be modified for the specific job they will do. You can see how this occurs by watching the animation at this link: <http://vcell.ndsu.edu/animations/proteinmodification/movie-flash.htm> .

Lesson Summary

- Transcription is the $DNA \rightarrow RNA$ part of the central dogma of molecular biology. It occurs in the nucleus. During transcription, a copy of mRNA is made that is complementary to a strand of DNA. In eukaryotes, mRNA may be modified before it leaves the nucleus.
- The genetic code consists of the sequence of bases in DNA or RNA. Groups of three bases form codons, and each codon stands for one amino acid (or start or stop). The codons are read in sequence following the start codon until a stop codon is reached. The genetic code is universal, unambiguous, and redundant.
- Translation is the $RNA \rightarrow protein$ part of the central dogma. It occurs at a ribosome. During translation, a protein is synthesized using the codons in mRNA as a guide. All three types of RNA play a role in translation.

Lesson Review Questions

Recall

1. Describe transcription.
2. How may mRNA be modified before it leaves the nucleus?
3. What is the genetic code? What are codons?
4. Outline the steps of translation.

Apply Concepts

5. Use the genetic code in **Figure 1.8** to translate the following segment of RNA into a sequence of five amino acids:
GUC-GCG-CAU-AGC-AAG

Think Critically

6. The genetic code is universal, unambiguous, and redundant. Explain what this means and why it is important.
7. How are transcription and translation related to the central dogma of molecular biology?

Points to Consider

When DNA is replicated or transcribed, accidents can happen, leading to a change in the base sequence.

- What do you think could cause such accidents to occur?
- How might the changes affect the reading frame? How might the encoded protein be affected?

1.3 Mutation

Lesson Objectives

- Identify causes of mutation.
- Compare and contrast types of mutations.
- Explain how mutations may affect the organisms in which they occur.

Vocabulary

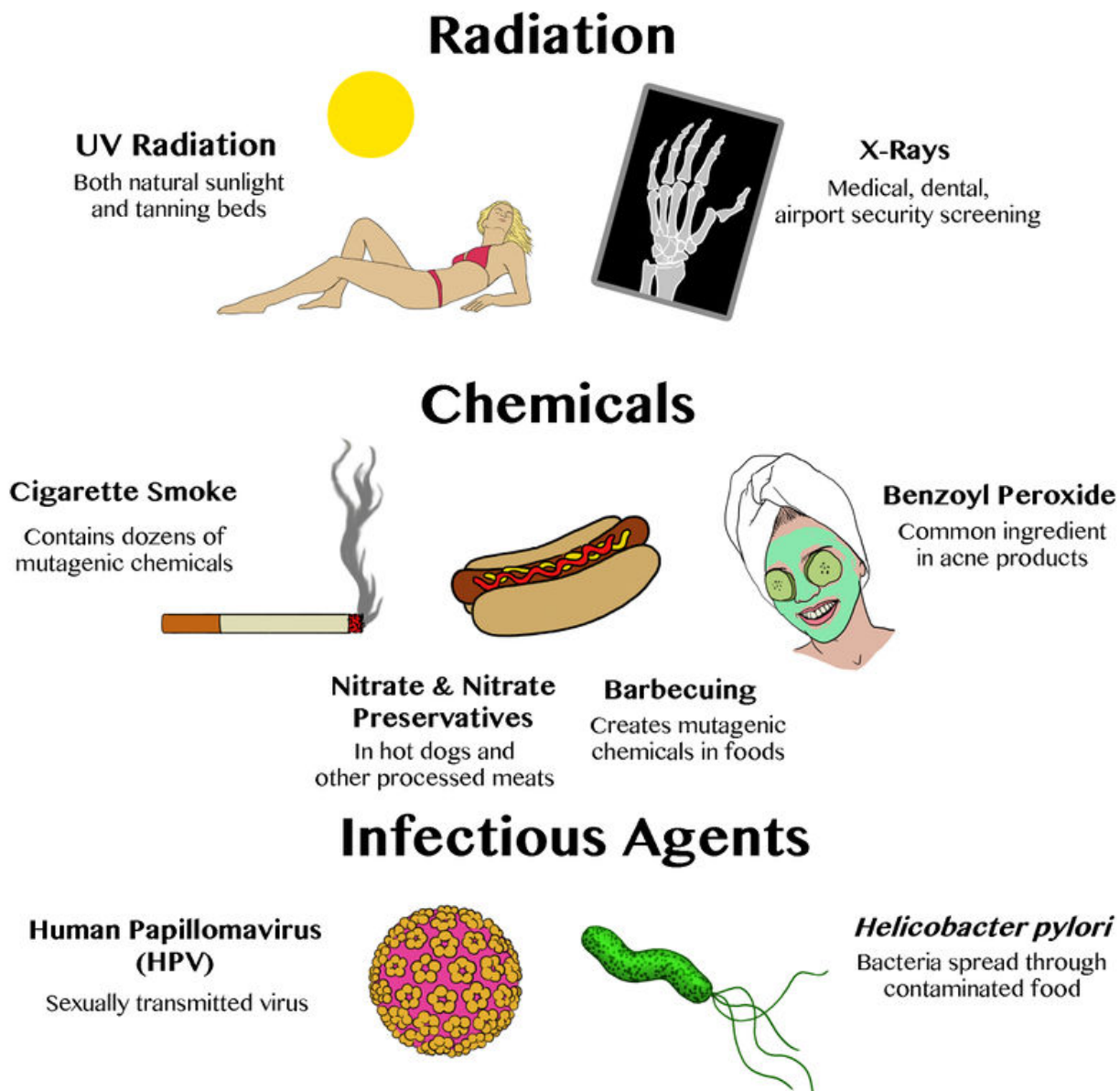
- chromosomal alteration
- frameshift mutation
- genetic disorder
- germline mutation
- mutagen
- mutation
- point mutation
- somatic mutation

Introduction

A change in the sequence of bases in DNA or RNA is called a **mutation**. Does the word mutation make you think of science fiction and bug-eyed monsters? Think again. Everyone has mutations. In fact, most people have dozens or even hundreds of mutations in their DNA. Mutations are essential for evolution to occur. They are the ultimate source of all new genetic material in a species. Although most mutations have no effect on the organisms in which they occur, some mutations are beneficial. Even harmful mutations rarely cause drastic changes in organisms.

Causes of Mutation

Mutations have many possible causes. Some mutations seem to happen spontaneously without any outside influence. They occur when mistakes are made during DNA replication or transcription. Other mutations are caused by environmental factors. Anything in the environment that can cause a mutation is known as a **mutagen**. Examples of mutagens are pictured in **Figure 1.11**. For a video about mutagens, go the link below. <http://www.youtube.com/watch?v=0wrNxCGKCws> (0:36)

**FIGURE 1.11**

Examples of Mutagens. Types of mutagens include radiation, chemicals, and infectious agents. Do you know of other examples of each type of mutagen shown here?

Types of Mutations

There are a variety of types of mutations. Two major categories of mutations are germline mutations and somatic mutations.

- **Germline mutations** occur in gametes. These mutations are especially significant because they can be transmitted to offspring and every cell in the offspring will have the mutation.
- **Somatic mutations** occur in other cells of the body. These mutations may have little effect on the organism because they are confined to just one cell and its daughter cells. Somatic mutations cannot be passed on to offspring.

Mutations also differ in the way that the genetic material is changed. Mutations may change the structure of a chromosome or just change a single nucleotide.

Chromosomal Alterations

Chromosomal alterations are mutations that change chromosome structure. They occur when a section of a chromosome breaks off and rejoins incorrectly or does not rejoin at all. Possible ways these mutations can occur are illustrated in **Figure 1.12**. Go to this link for a video about chromosomal alterations: http://www.youtube.com/watch?v=OrXRSqa_3IU (2:18).

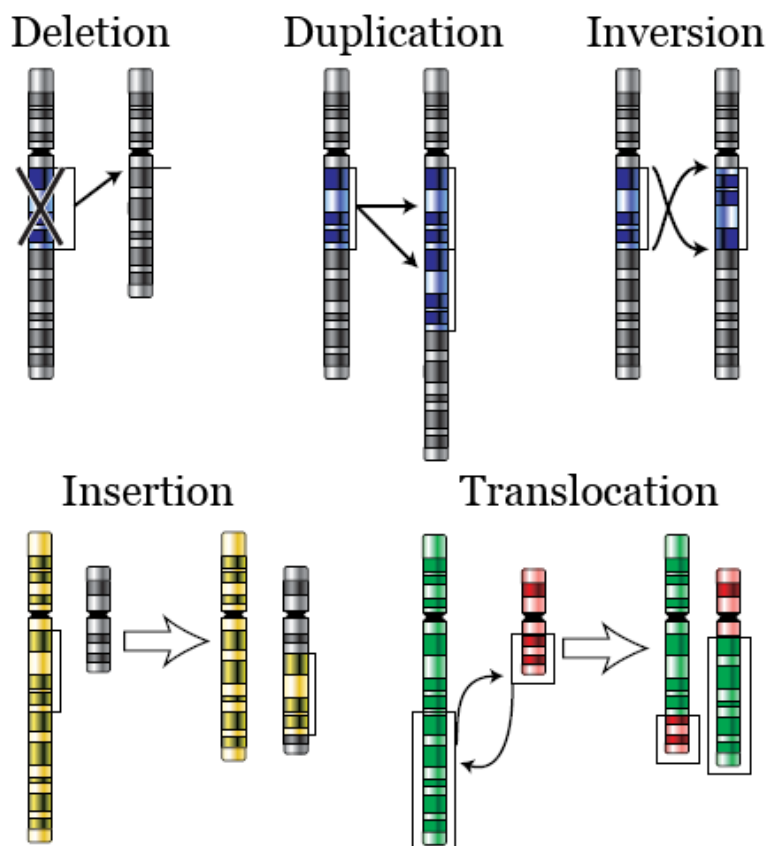


FIGURE 1.12

Chromosomal Alterations. Chromosomal alterations are major changes in the genetic material.

Chromosomal alterations are very serious. They often result in the death of the organism in which they occur. If the organism survives, it may be affected in multiple ways. An example of a human chromosomal alteration is the mutation that causes Down Syndrome. It is a duplication mutation that leads to developmental delays and other abnormalities.

Point Mutations

A **point mutation** is a change in a single nucleotide in DNA. This type of mutation is usually less serious than a chromosomal alteration. An example of a point mutation is a mutation that changes the codon UUU to the codon UCU. Point mutations can be silent, missense, or nonsense mutations, as shown in **Table 1.1**. The effects of point mutations depend on how they change the genetic code. You can watch an animation about nonsense mutations at this link: http://www.biostudio.com/d_%20Nonsense%20Suppression%20I%20Nonsense%20Mutation.htm .

TABLE 1.1: Point Mutations and Their Effects

Type	Description	Example	Effect
Silent	mutated codon codes for the same amino acid	CAA (glutamine) → CAG (glutamine)	none
Missense	mutated codon codes for a different amino acid	CAA (glutamine) → CCA (proline)	variable
Nonsense	mutated codon is a premature stop codon	CAA (glutamine) → UAA (stop) usually	serious

Frameshift Mutations

A **frameshift mutation** is a deletion or insertion of one or more nucleotides that changes the reading frame of the base sequence. Deletions remove nucleotides, and insertions add nucleotides. Consider the following sequence of bases in RNA:

AUG-AAU-ACG-GCU = start-asparagine-threonine-alanine

Now, assume an insertion occurs in this sequence. Let's say an **A** nucleotide is inserted after the start codon **AUG**:

AUG-AAA-UAC-GGC-U = start-lysine-tyrosine-glycine

Even though the rest of the sequence is unchanged, this insertion changes the reading frame and thus all of the codons that follow it. As this example shows, a frameshift mutation can dramatically change how the codons in mRNA are read. This can have a drastic effect on the protein product.

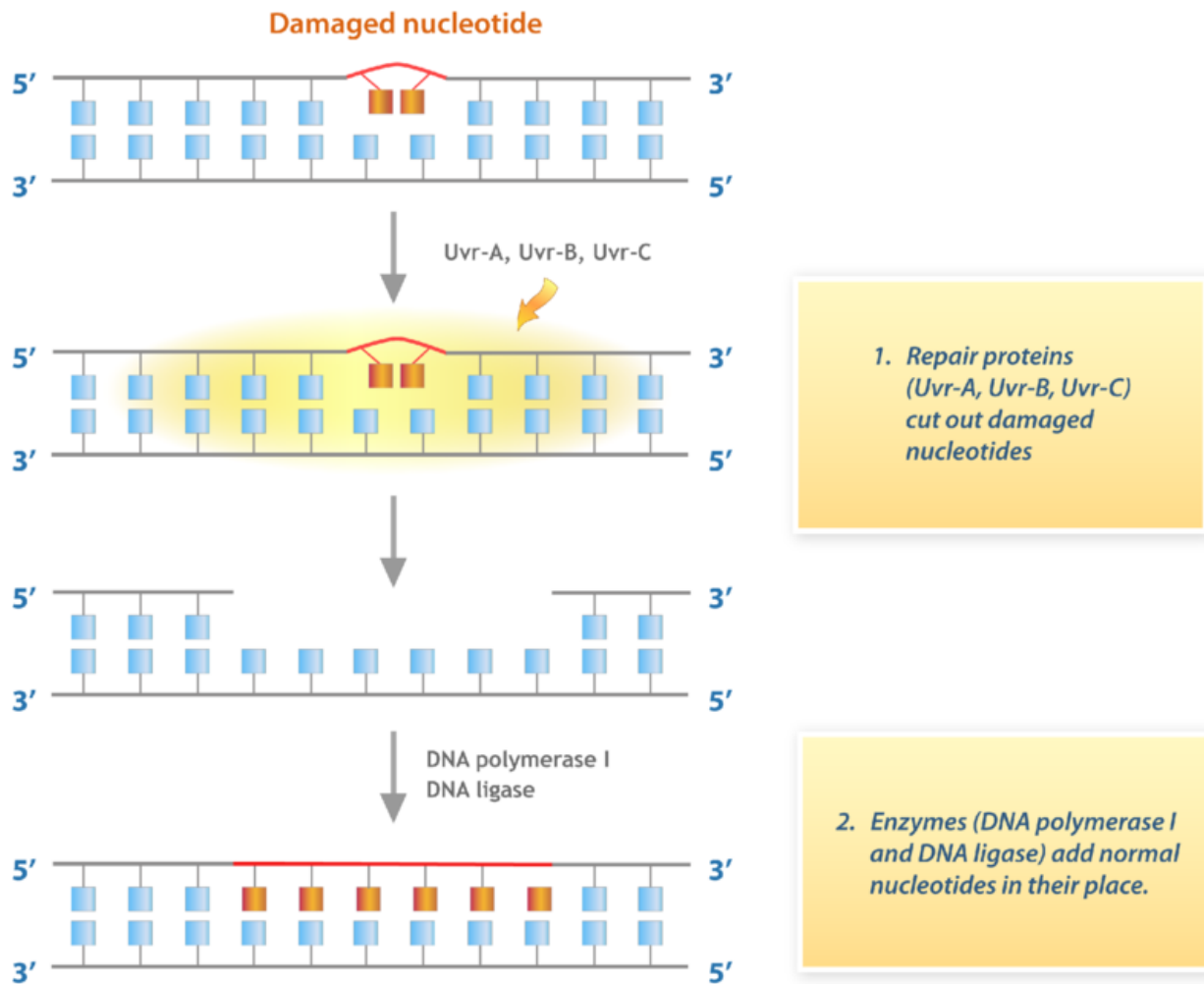
Effects of Mutations

The majority of mutations have neither negative nor positive effects on the organism in which they occur. These mutations are called neutral mutations. Examples include silent point mutations. They are neutral because they do not change the amino acids in the proteins they encode.

Many other mutations have no effect on the organism because they are repaired before protein synthesis occurs. Cells have multiple repair mechanisms to fix mutations in DNA. One way DNA can be repaired is illustrated in **Figure 1.13**. If a cell's DNA is permanently damaged and cannot be repaired, the cell is likely to be prevented from dividing.

Beneficial Mutations

Some mutations have a positive effect on the organism in which they occur. They are called beneficial mutations. They lead to new versions of proteins that help organisms adapt to changes in their environment. Beneficial mutations are essential for evolution to occur. They increase an organism's chances of surviving or reproducing, so they are

**FIGURE 1.13**

DNA Repair Pathway. This flow chart shows one way that damaged DNA is repaired in *E. coli* bacteria.

likely to become more common over time. There are several well-known examples of beneficial mutations. Here are just two:

1. Mutations in many bacteria that allow them to survive in the presence of antibiotic drugs. The mutations lead to antibiotic-resistant strains of bacteria.
2. A unique mutation is found in people in a small town in Italy. The mutation protects them from developing atherosclerosis, which is the dangerous buildup of fatty materials in blood vessels. The individual in which the mutation first appeared has even been identified.

Harmful Mutations

Imagine making a random change in a complicated machine such as a car engine. The chance that the random change would improve the functioning of the car is very small. The change is far more likely to result in a car that does not run well or perhaps does not run at all. By the same token, any random change in a gene's DNA is likely to

result in a protein that does not function normally or may not function at all. Such mutations are likely to be harmful. Harmful mutations may cause genetic disorders or cancer.

- A **genetic disorder** is a disease caused by a mutation in one or a few genes. A human example is cystic fibrosis. A mutation in a single gene causes the body to produce thick, sticky mucus that clogs the lungs and blocks ducts in digestive organs. You can watch a video about cystic fibrosis and other genetic disorders at this link: <http://www.youtube.com/watch?v=8s4he3wLgkM> (9:31).
- Cancer is a disease in which cells grow out of control and form abnormal masses of cells. It is generally caused by mutations in genes that regulate the cell cycle. Because of the mutations, cells with damaged DNA are allowed to divide without limits. Cancer genes can be inherited. You can learn more about hereditary cancer by watching the video at the following link: <http://www.youtube.com/watch?v=LWk5FplsKwM> (4:29)

Albino Redwoods, Ghosts of the Forest

What happens if a plant does not have chlorophyll? They would lack the part of the leaf that makes them green. So these plants could be referred to as albino. This would have to result from a genetic mutation. Do these plants die because they cannot photosynthesize? Not necessarily. What can these plants tell us about the biochemistry, genetics and physiology of plants?

See *Science on the SPOT: Albino Redwoods, Ghosts of the Forest* at <http://science.kqed.org/quest/video/science-on-the-spot-albino-redwoods-ghosts-of-the-forest/> , *Science on the SPOT: Revisiting Albino Redwoods, Biological Mystery* at <http://science.kqed.org/quest/video/science-on-the-spot-revisiting-albino-redwoods-biological-mystery/> , and *Science on the SPOT: Revisiting Albino Redwoods, Cracking the Code* at <http://science.kqed.org/quest/video/science-on-the-spot-revisiting-albino-redwoods-cracking-the-code/> for more information.



MEDIA

Click image to the left or use the URL below.

URL: <http://www.ck12.org/flx/render/embeddedobject/454>



MEDIA

Click image to the left or use the URL below.

URL: <http://www.ck12.org/flx/render/embeddedobject/108477>



MEDIA

Click image to the left or use the URL below.

URL: <http://www.ck12.org/flx/render/embeddedobject/455>

Lesson Summary

- Mutations are caused by environmental factors known as mutagens. Types of mutagens include radiation, chemicals, and infectious agents.
- Germline mutations occur in gametes. Somatic mutations occur in other body cells. Chromosomal alterations are mutations that change chromosome structure. Point mutations change a single nucleotide. Frameshift mutations are additions or deletions of nucleotides that cause a shift in the reading frame.
- Mutations are essential for evolution to occur because they increase genetic variation and the potential for individuals to differ. The majority of mutations are neutral in their effects on the organisms in which they occur. Beneficial mutations may become more common through natural selection. Harmful mutations may cause genetic disorders or cancer.

Lesson Review Questions

Recall

1. Define mutation and mutagen.
2. List three examples of mutagens.
3. Identify three types of chromosomal alterations.
4. Distinguish among silent, missense, and nonsense point mutations.
5. What is a frameshift mutation? What causes this type of mutation?

Apply Concepts

6. Assume that a point mutation changes the codon AUU to AUC. Why is this a neutral mutation?
7. Look at the mutation shown below. The base A was inserted following the start codon AUG. Describe how this mutation affects the encoded amino acid sequence.

AUG-GUC-CCU-AAA → AUG-AGU-CCC-UAA-A

Think Critically

8. Compare and contrast germline mutations and somatic mutations.
9. Why are mutations essential for evolution to occur?

Points to Consider

Sometimes even drastic mutations do not affect the proteins produced by a particular type of cell. The reason? The genes affected by the mutations are not normally used to make proteins in that type of cell. In all cells, some genes are turned off - they are not transcribed - while other genes are turned on.

- How do cells control which genes are turned on and used to make proteins?

- Can you think of a mechanism that might prevent transcription of a gene?

1.4 Regulation of Gene Expression

Lesson Objectives

- Identify general mechanisms that regulate gene expression.
- Describe how gene regulation occurs in prokaryotes.
- Give an overview of gene regulation in eukaryotes.

Vocabulary

- gene expression
- homeobox gene
- operator
- operon
- regulatory element
- regulatory protein
- TATA box

Introduction

Each of your cells has at least 20,000 genes. In fact, all of your cells have the same genes. Do all of your cells make the same proteins? Obviously not. If they did, then all your cells would be alike. Instead, you have cells with different structures and functions. This is because different cells make different proteins. They do this by using, or expressing, different genes. Using a gene to make a protein is called **gene expression**.

How Gene Expression is Regulated

Gene expression is regulated to ensure that the correct proteins are made when and where they are needed. Regulation may occur at any point in the expression of a gene, from the start of transcription to the processing of a protein after translation. The focus in this lesson is the regulation of transcription.

As shown in **Figure 1.14**, transcription is controlled by **regulatory proteins**. The proteins bind to regions of DNA, called **regulatory elements**, which are located near promoters. After regulatory proteins bind to regulatory elements, they can interact with RNA polymerase, the enzyme that transcribes DNA to mRNA. Regulatory proteins are typically either activators or repressors.

- Activators promote transcription by enhancing the interaction of RNA polymerase with the promoter.
- Repressors prevent transcription by impeding the progress of RNA polymerase along the DNA strand.

Other factors may also be involved in the regulation of transcription, but these are typically the key players.

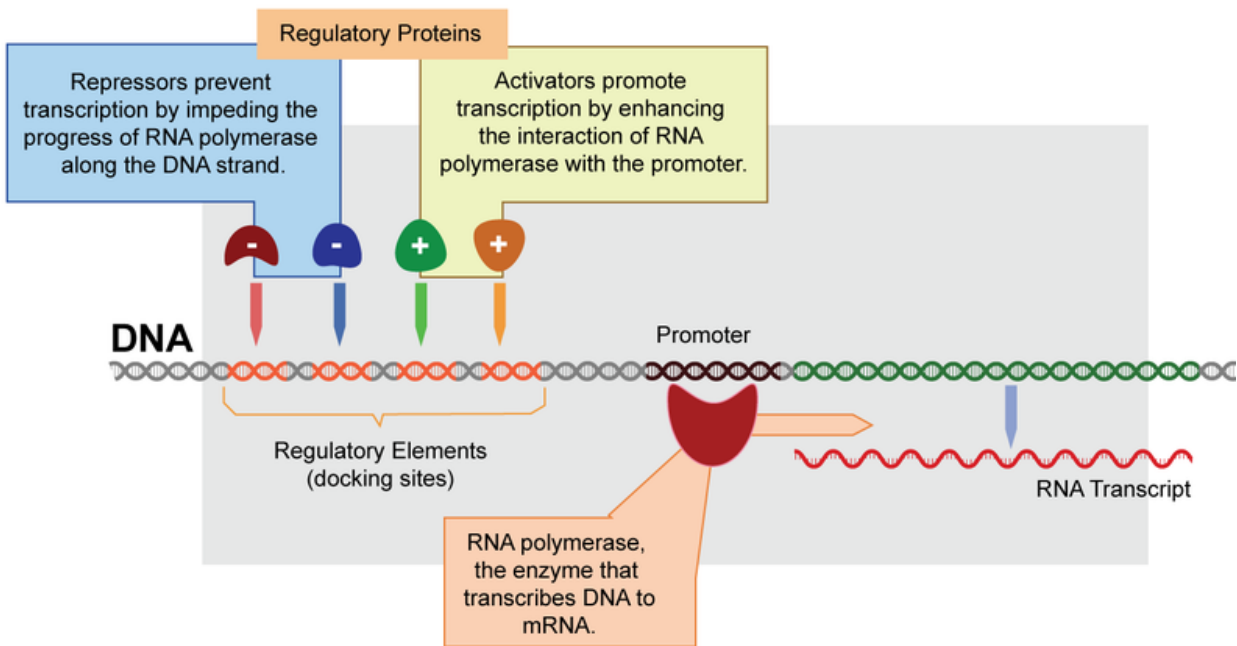


FIGURE 1.14

Regulation of Transcription. Regulatory proteins bind to regulatory elements to control transcription. The regulatory elements are embedded within the DNA.

Prokaryotic Gene Regulation

Transcription is regulated differently in prokaryotes and eukaryotes. In general, prokaryotic regulation is simpler than eukaryotic regulation.

The Role of Operons

Regulation of transcription in prokaryotes typically involves operons. An **operon** is a region of DNA that consists of one or more genes that encode the proteins needed for a specific function. The operon also includes a promoter and an operator. The **operator** is a region of the operon where regulatory proteins bind. It is located near the promoter and helps regulate transcription of the operon genes.

The Lac Operon

A well-known example of operon regulation involves the lac operon in *E. coli* bacteria (see **Figure 1.15** and the video at the link below). The lac operon consists of a promoter, an operator, and three genes that encode the enzymes needed to digest lactose, the sugar found in milk. The lac operon is regulated by lactose in the environment. <http://www.youtube.com/watch?v=oBwtxdI1zvK>

- When lactose is absent, a repressor protein binds to the operator. The protein blocks the binding of RNA polymerase to the promoter. As a result, the lac genes are not expressed.

- When lactose is present, the repressor protein does not bind to the operator. This allows RNA polymerase to bind to the promoter and begin transcription. As a result, the lac genes are expressed, and lactose is digested.

Why might it be beneficial to express genes only when they are needed? (Hint: synthesizing proteins requires energy and materials.)

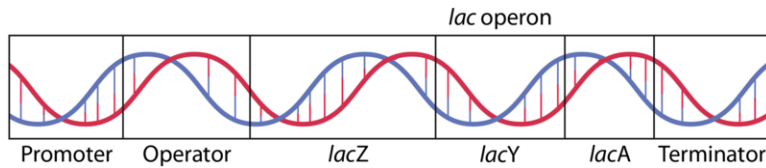


FIGURE 1.15

The three genes of the lac operon are lacZ, lacY, and lacA. They encode proteins needed to digest lactose. The genes are expressed only in the presence of lactose.

Eukaryotic Gene Regulation

In eukaryotic cells, the start of transcription is one of the most complicated parts of gene regulation. There may be many regulatory proteins and regulatory elements involved. Regulation may also involve enhancers. Enhancers are distant regions of DNA that can loop back to interact with a gene's promoter.

The TATA Box

Different types of cells have unique patterns of regulatory elements that result in only the necessary genes being transcribed. That's why a skin cell and nerve cell, for example, are so different from each other. However, some patterns of regulatory elements are common to all genes, regardless of the cells in which they occur. An example is the **TATA box**. This is a regulatory element that is part of the promoter of most eukaryotic genes. A number of regulatory proteins bind to the TATA box, forming a multi-protein complex. It is only when all of the appropriate proteins are bound to the TATA box that RNA polymerase recognizes the complex and binds to the promoter. Once RNA polymerase binds, transcription begins. To see a video showing the role of the TATA box in the initiation of transcription, go to this link: <http://www.youtube.com/watch?v=6tqPsi-9aQA> .

Regulation During Development

The regulation of gene expression is extremely important during the development of an organism. Regulatory proteins must turn on certain genes in particular cells at just the right time so the organism develops normal organs and organ systems. **Homeobox genes** are an example of genes that regulate development. They code for regulatory proteins that switch on whole series of major developmental genes. In insects, homeobox genes called hox genes ensure that body parts such as limbs develop in the correct place. **Figure 1.16** shows how a mutation in a hox gene can affect an insect's development. You can learn more about homeobox genes at this link: <http://www.youtube.com/watch?v=LFG-aLidT8s> .

Gene Expression and Cancer

The mutations that cause cancer generally occur in two types of regulatory genes: tumor-suppressor genes and proto-oncogenes (see **Figure 1.17**). These genes produce regulatory proteins that control the cell cycle. When the genes

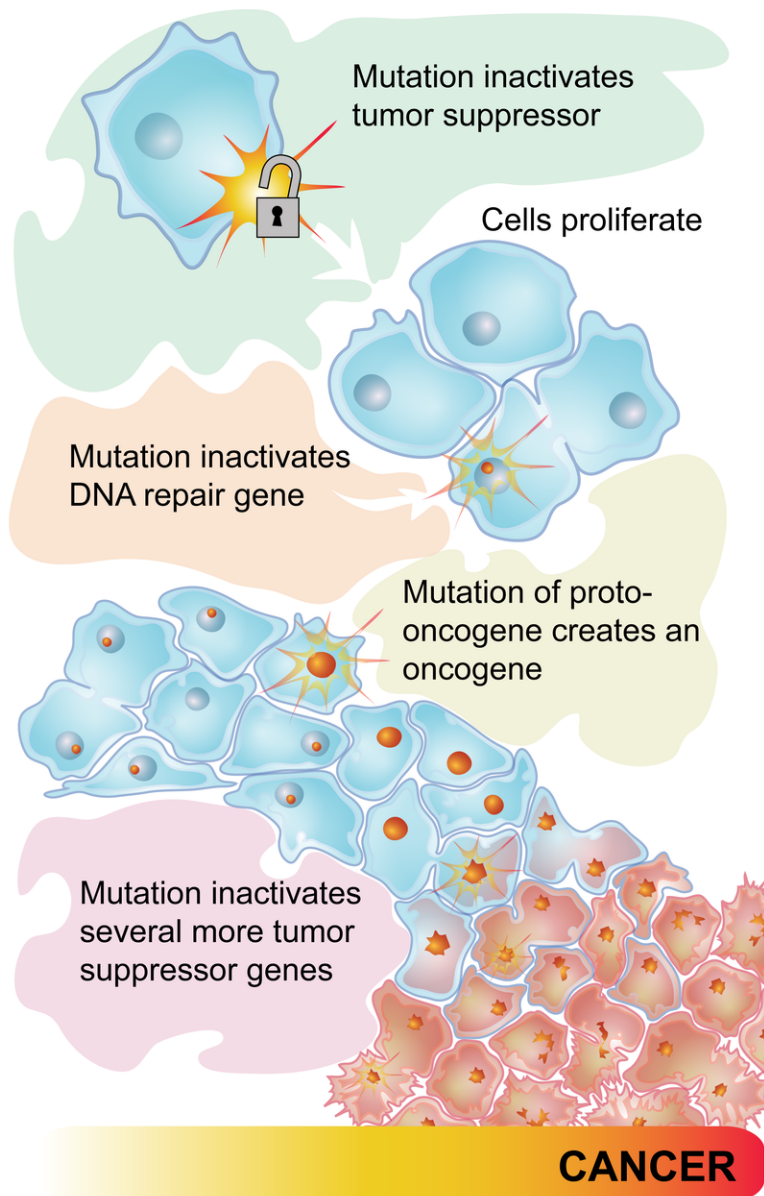
**FIGURE 1.16**

Effect of Hox Gene Mutation. Scientists caused a mutation in a hox gene of this fruit fly. As a result of the mutation, a leg grew out of its head where an antenna should have developed (fly on right). Normal fly (fly on left).

mutate, cells with mutations divide rapidly and without limits.

Lesson Summary

- Gene transcription is controlled by regulatory proteins that bind to regulatory elements on DNA. The proteins usually either activate or repress transcription.
- Regulation of transcription in prokaryotes typically involves an operon, such as the lac operon in *E. coli*. The lac operon is regulated by proteins that behave differently depending on whether lactose is present.
- Regulation of transcription in eukaryotes is generally more complex. It involves unique regulatory elements in different cells as well as common regulatory elements such as the TATA box. Regulation is especially important during development. It may involve regulatory genes such as homeobox genes that switch other regulatory genes on or off. Mutations in regulatory genes that normally control the cell cycle cause cancer.

**FIGURE 1.17**

How Cancer Develops. This flow chart shows how a series of mutations in tumor-suppressor genes and proto-oncogenes leads to cancer.

Lesson Review Questions

Recall

1. What is gene expression?
2. Describe how regulatory proteins regulate gene expression.
3. Identify the TATA box and its function in transcription.
4. What is a homeobox gene?

Apply Concepts

5. Draw a diagram to show how the lac operon is regulated.
6. Sketch how an insect with a mutated hox gene might look. Explain your sketch.

Think Critically

7. Why is gene regulation especially important during development?

Points to Consider

Scientists know more about human chromosomes and genes than they know about the genetic material of most other species. In fact, scientists have identified all of the approximately 20,000-25,000 genes in human DNA.

- What do you know about human chromosomes and genes? For example, do you know how many chromosomes humans normally have?
- Do you know how human characteristics are inherited? Can you identify characteristics that are controlled by a single gene?

1.5 References

1. Mariana Ruiz Villarreal (LadyofHats) for CK-12 Foundation. [CK-12 Foundation](#) . CC BY-NC 3.0
2. Hana Zavadska. [CK-12 Foundation](#) . CC BY-NC 3.0
3. DNA: Image copyright ermess, 2014; Staircase: Andrew Gould. DNA: <http://www.shutterstock.com>; Staircase: <http://www.flickr.com/photos/27950702@N04/3644533740> . DNA: Used under license from Shutterstock.com; Staircase: CC BY 2.0
4. Zachary Wilson. [CK-12 Foundation](#) . CC BY-NC 3.0
5. LadyofHats for CK-12. [CK-12 Foundation](#) . CC-BY-NC-SA 3.0
6. Mariana Ruiz Villarreal (LadyofHats) for CK-12 Foundation. [CK-12 Foundation](#) . CC BY-NC 3.0
7. Mariana Ruiz Villarreal (LadyofHats) for CK-12 Foundation. [CK-12 Foundation](#) . CC BY-NC 3.0
8. Hana Zavadska. [CK-12 Foundation](#) . CC BY-NC 3.0
9. Zachary Wilson. [CK-12 Foundation](#) . CC BY-NC 3.0
10. Original image by the National Human Genome Research Institute, redrawn by Mariana Ruiz Villarreal (LadyofHats) for CK-12 Foundation. [CK-12 Foundation](#); Original image from: <http://www.genome.gov/Glossary/index.cfm?id=200> . CC BY-NC 3.0
11. Laura Guerin. [CK-12 Foundation](#) . CC BY-NC 3.0
12. Courtesy of the National Human Genome Research Institute, User:Dietzel65/De.Wikipedia. <http://commons.wikimedia.org/wiki/File:Chromosomenmutationen.png> . Public Domain
13. Hana Zavadska. [CK-12 Foundation](#) . CC BY-NC 3.0
14. Mariana Ruiz Villarreal (LadyofHats) for CK-12 Foundation. [CK-12 Foundation](#) . CC BY-NC 3.0
15. Jodi So. [CK-12 Foundation](#) . CC BY-NC 3.0
16. User:Ktbn/Wikimedia Commons. <http://commons.wikimedia.org/wiki/File:Antennapedia.jpg> .
17. Mariana Ruiz Villarreal (LadyofHats) for CK-12 Foundation. [CK-12 Foundation](#) . CC BY-NC 3.0