Human Biology 2nd edition

Human Biology 2nd edition

Sarah Malmquist and Kristina Prescott

Contents

I

Theme 1: What Makes Us Unique?

As a species, Homo sapiens has accomplished amazing things. Human beings have built complex societies, created beautiful art and music, and traveled to the moon and beyond. We are capable of committing cruel and destructive acts, but also of living with great love and compassion. Do any of these characteristics make us fundamentally different from other life on Earth?

In this theme, we will explore our place within life on Earth. What makes us human and what makes each of us unique? We will examine the DNA "blueprint" for building a human being, the evolutionary origins of our species, and the sources of complex thought within the human brain and nervous system. We will also consider the uniqueness of human DNA and the ethics of collecting and using several types of DNA evidence in criminal investigations.

1.

1.1 Structural Organization of the Human Body

Before you begin to study the different structures and functions of the human body, it is helpful to consider its basic architecture; that is, how its smallest parts are assembled into larger structures. It is convenient to consider the structures of the body in terms of fundamental levels of organization that increase in complexity: subatomic particles, atoms, molecules, organelles, cells, tissues, organs, organ systems, organisms and biosphere (***Figure 1***).

Levels of Structural Organization of the Human Body

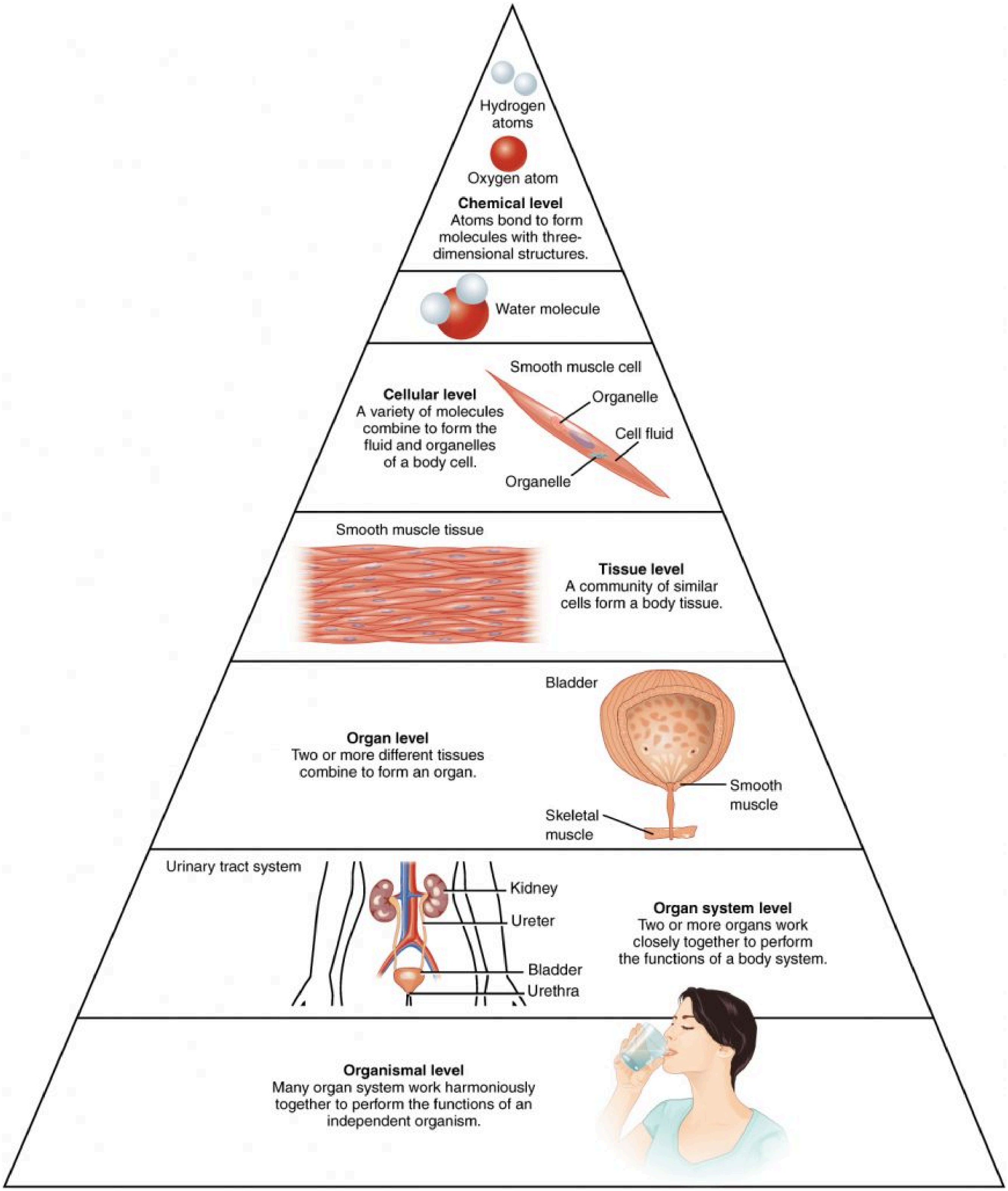


Figure 1. *The organization of the body often is discussed in terms of six distinct levels of increasing complexity, from the smallest chemical building blocks to a unique human organism.*

The Levels of Organization

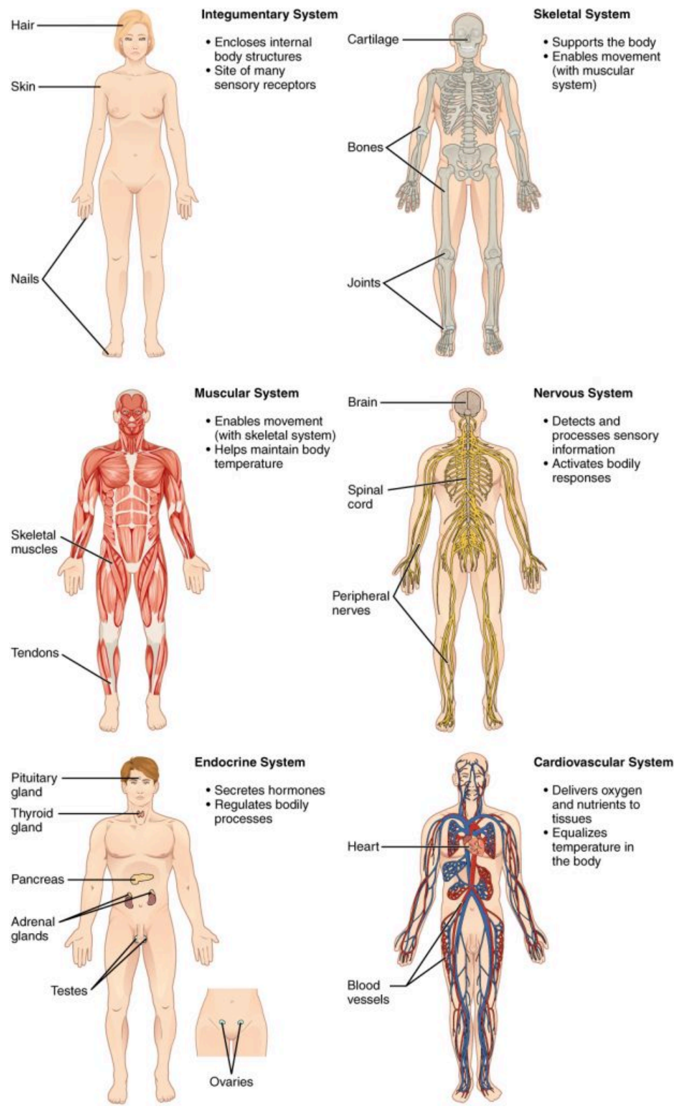
To study the chemical level of organization, scientists consider the simplest building blocks of matter: subatomic particles, atoms and molecules. All matter in the universe is composed of one or more unique pure substances called elements, familiar examples of which are hydrogen, oxygen, carbon, nitrogen, calcium, and iron. The smallest unit of any of these pure substances (elements) is an atom. Atoms are made up of subatomic particles such as the proton, electron and neutron. Two or more atoms combine to form a molecule, such as the water molecules, proteins, and sugars found in living things. Molecules are the chemical building blocks of all body structures.

A **cell** is the smallest independently functioning unit of a living organism. Even bacteria, which are extremely small, independently-living organisms, have a cellular structure. Each bacterium is a single cell. All living structures of human anatomy contain cells, and almost all functions of human physiology are performed in cells or are initiated by cells.

A human cell typically consists of flexible membranes that enclose cytoplasm, a water-based cellular fluid together with a variety of tiny functioning units called **organelles**. In humans, as in all organisms, cells perform all functions of life. A **tissue** is a group of many similar cells (though sometimes composed of a few related types) that work together to perform a specific function. An **organ** is an anatomically distinct structure of the body composed of two or more tissue types. Each organ performs one or more specific physiological functions. An **organ system** is a group of organs that work together to perform major functions or meet physiological needs of the body.

Figure 2 shows some of the organ systems of the body that we will consider over the course of this semester. Many organs have functions integral to more than one organ system.

Organ Systems of the Human Body


Integumentary System
• Encloses internal body structures
• Site of many sensory receptors
Hair
Skin
Nails
Skeletal System
• Supports the body
• Enables movement (with muscular system)
Cartilage
Bones
Joints
Muscular System
• Enables movement (with skeletal system)
• Helps maintain body temperature
Skeletal muscles
Tendons
Nervous System
• Detects and processes sensory information
• Activates bodily responses
Brain
Spinal cord
Peripheral nerves
Endocrine System
• Secretes hormones
• Regulates bodily processes
Pituitary gland
Thyroid gland
Pancreas
Adrenal glands
Testes
Ovaries
Cardiovascular System
• Delivers oxygen and nutrients to tissues
• Equalizes temperature in the body
Heart
Blood vessels

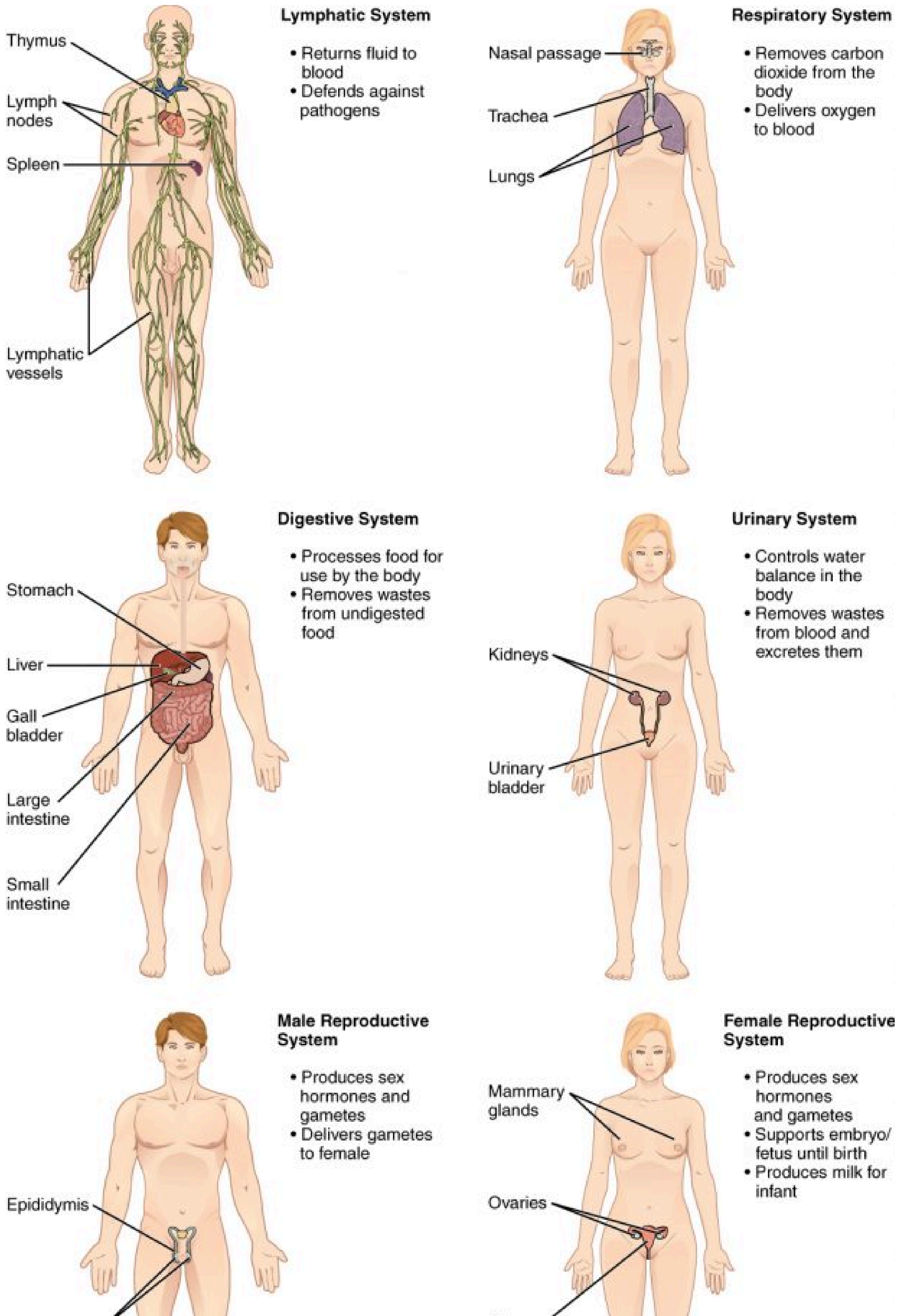

Lymphatic System
• Returns fluid to blood
• Defends against pathogens
Thymus
Lymph nodes
Spleen
Lymphatic vessels
Respiratory System
• Removes carbon dioxide from the body
• Delivers oxygen to blood
Nasal passage
Trachea
Lungs
Digestive System
• Processes food for use by the body
• Removes wastes from undigested food
Stomach
Liver
Gall bladder
Large intestine
Small intestine
Urinary System
• Controls water balance in the body
• Removes wastes from blood and excretes them
Kidneys
Urinary bladder
Male Reproductive System
• Produces sex hormones and gametes
• Delivers gametes to female
Epididymis
Female Reproductive System
• Produces sex hormones and gametes
• Supports embryo/fetus until birth
• Produces milk for infant
Mammary glands
Ovaries

Figure 2. *Organs that work together are grouped into organ systems.*

The organism level is the highest level of organization considered in anatomy/physiology. An **organism** is a living being that has a cellular structure and that can independently perform all physiologic functions necessary for life. In multicellular organisms, including humans, all cells, tissues, organs, and organ systems of the body work together to maintain the life and health of the organism.

Section Summary

Life processes of the human body are maintained at several levels of structural organization. These include the chemical, cellular, tissue, organ, organ system, and the organism level. Higher levels of organization are built from lower levels. Therefore, molecules combine to form cells, cells combine to form tissues, tissues combine to form organs, organs combine to form organ systems, and organ systems combine to form organisms.

Glossary

cell
: smallest independently functioning unit of all organisms; in animals, a cell contains cytoplasm, composed of fluid and organelles

organ
: functionally distinct structure composed of two or more types of tissues

organ system
: group of organs that work together to carry out a particular function

organism
: living being that has a cellular structure and that can independently perform all physiologic functions necessary for life

tissue
: group of similar or closely related cells that act together to perform a specific function

2.

1.2 DNA Overview

All living things have genetic material that they pass on to their offspring. This genetic material, **deoxyribonucleic acid**, or DNA, contains a code that serves as the blueprint for building each unique organism. In humans, and many other organisms such as plants and fungi, DNA is housed in a membrane-bound organelle called a nucleus. Every cell in your body contains a complete copy of the DNA instructions for building all of the parts of your body. This complete copy is referred to as your genome. Unless you have an identical twin, your genome is unique.

Despite their identical DNA, your body cells can be very different from each other. For example, the cells lining the inner surface of your cheek have a different structure and function than do the cells that make up the cheek muscles that allow you to smile. This is because each cell uses only a subset of the DNA instructions to make the specific proteins that it needs.

DNA structure

DNA is a macromolecule made up of a string of smaller units called nucleotides. Each nucleotide is made up of a sugar, a phosphate group and a base (***Figure 1***).

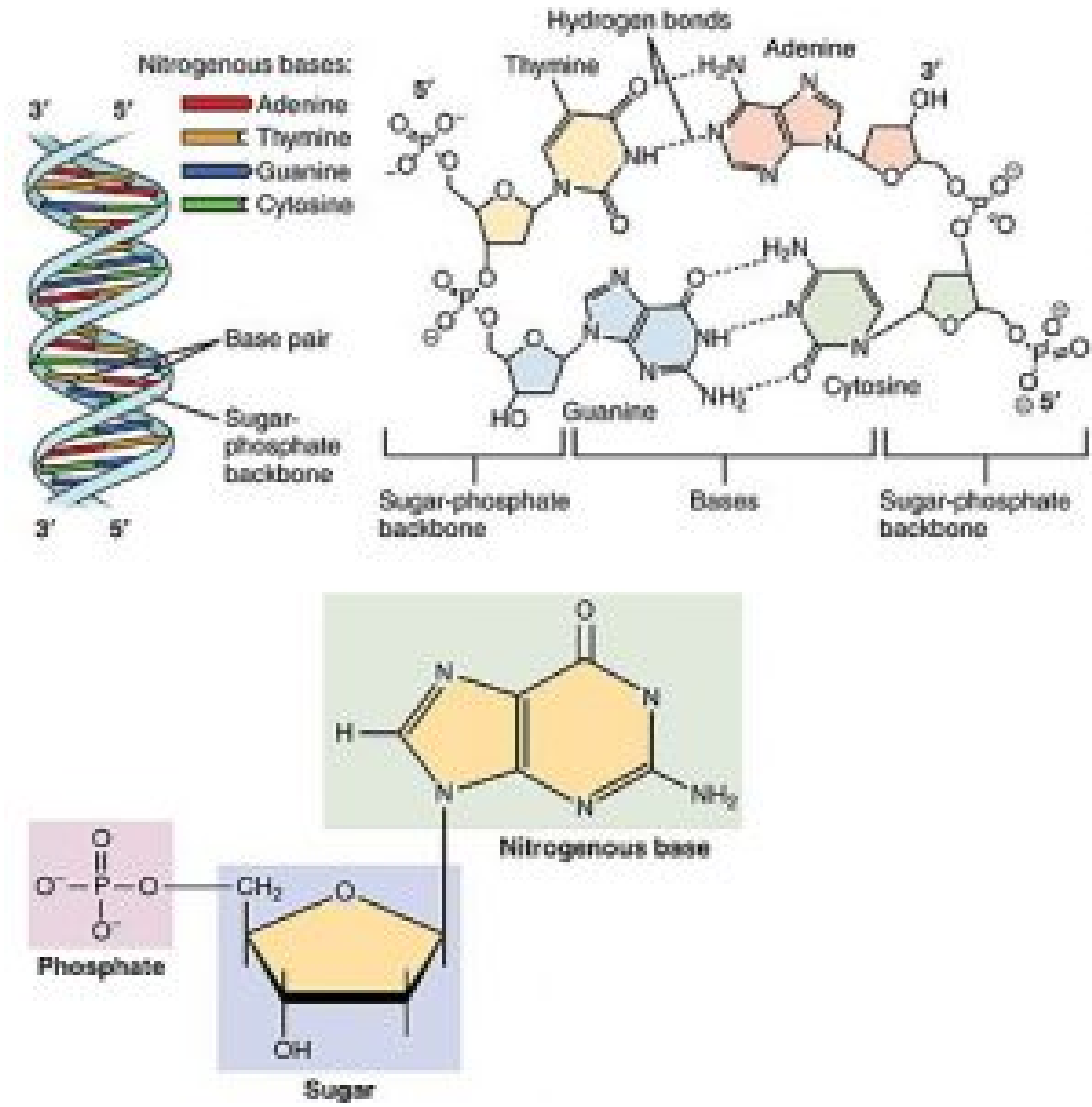


Figure 1. *Basic DNA structure.*

There are four different possible bases, adenine, thymine, cytosine and guanine. It is the order of these bases (often referred to simply as As, Ts, Cs and Gs) which make up the DNA “code.” DNA is found in the nuclei of all human cells, and serves as the blueprint for the production of the proteins necessary to maintain life.

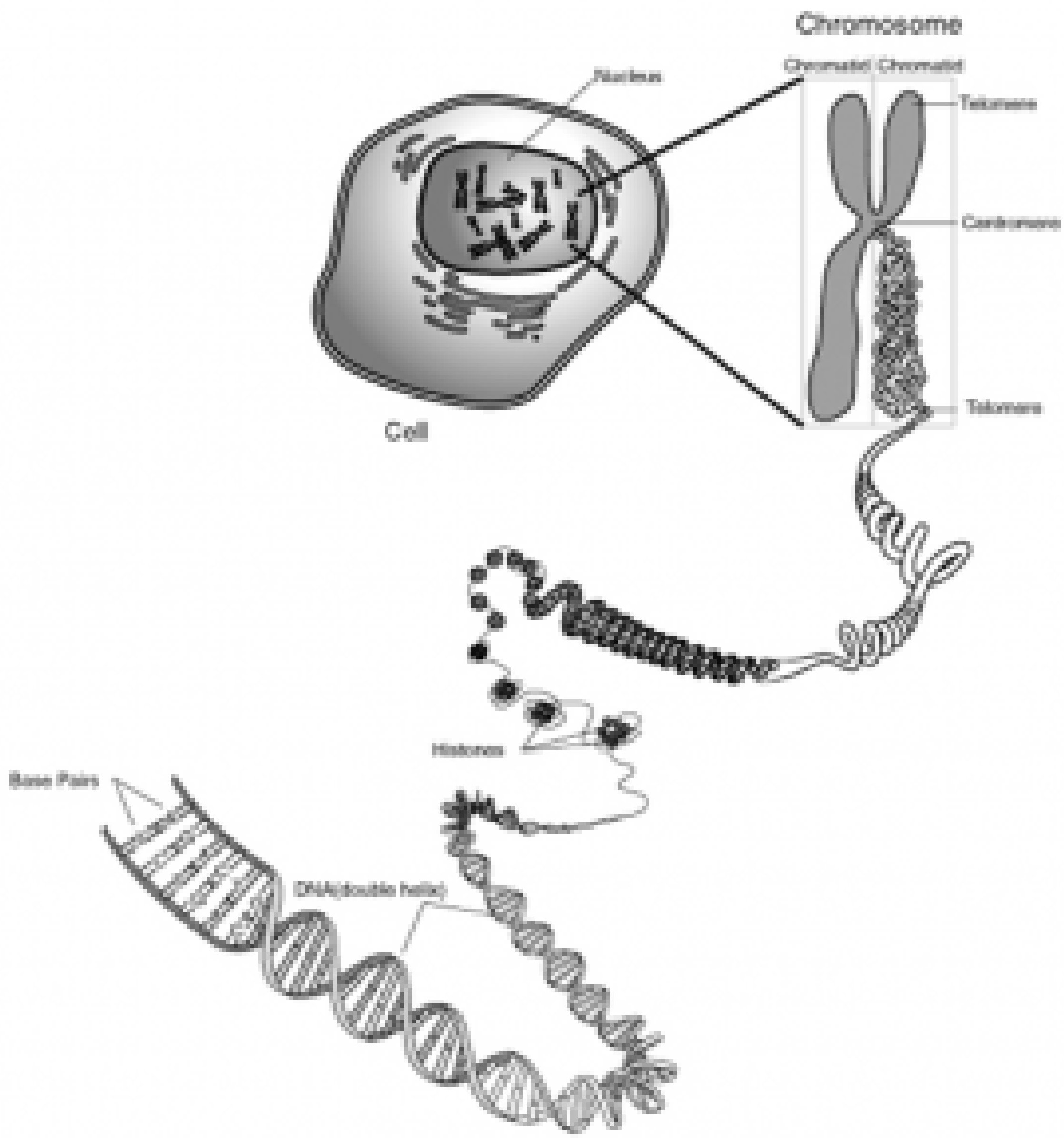


Figure 2. *DNA organized into chromosomes*

Human DNA is organized into 23 pairs of chromosomes (***Figure 2***). Before a cell divides, it replicates its DNA so that both daughter cells have a complete copy of the DNA blueprint. Occasionally, errors are made in DNA replication. Because the order of base pairs determine the proteins a cell makes, these errors, if left uncorrected, can lead to the production of different proteins. Errors in DNA replication that result in changes in the DNA molecule are called **mutations**. Mutations can happen in the DNA of any cell in the body. If a mutation happens in a sex cell, which is responsible for makings gametes (sperm or eggs), that mutation could be passed on to the organism's offspring. This videoclip offers a brief overview of the organization of human DNA and how small variations in human DNA caused by mutations can result in the variation we see in human populations.

3.

1.3 Human DNA and Genetic Diversity

Human DNA and Genetic Diversity

Small differences in DNA among humans can give rise to differences in traits. A **gene** is a segment of DNA that codes for a specific trait. For example, a person with blood type A has a gene that codes for the production of the blood surface protein A. A person lacking this gene would not produce A blood protein. On each of our 23 chromosomes, we have many different genes, each coding for different proteins and resulting in a variety of traits. An individual's complete set of genes is referred to as their **genome**.

DNA is used to study the origin of human characteristics, variation among human beings and the relatedness between humans and other species. From 1990-2003, scientists worked to sequence the human genome, as part of the Human Genome Project (https:// www.genome.gov/12011238/an-overview-of-the-human-genome-project/). The project, led by the National Institute of Health, was based on collaborative research, with contributions by scientists from China, France, Germany, Japan, Spain, the U.K. and the U.S., among others. After collecting DNA samples from many individuals, researchers sequenced the DNA of a subset of individuals and patched it together to create a mosaic of the human genome.

This complete human genome is about 3 billion base pairs long and is freely available for public and research use. The identities of the individuals whose genes were sequenced was not revealed to the researchers working on the project, and participants do not know if their DNA is the DNA that was sequenced. Within this human genome, there are about 22,300 protein-coding genes. Researchers are actively working to figure out what many of these genes do, as well as how the human genome might differ in populations not included in the sample. A "map" of chromosome 22, one of the shortest human chromosomes, is shown at above, with some genes of interest marked.

Human beings share almost all of their DNA. Thus, although we often focus on differences, you and the person sitting beside you have almost identical DNA sequences. Most variation between individuals' DNA comes in the form of single nucleotide polymorphisms (often referred to as "SNPS"). This refers to a single nucleotide difference between two people's DNA (for example, base pair A, instead of C). About 1 base pair out of every 1,000 will be different between two individuals. Human genetic variation seems to be continuous- it is very difficult to identify genetic distinctions among groups of people. Research results suggest that although a few traits can be mapped to specific to geographic regions (like arsenic resistance in isolated populations in South America, for example), variation in most traits is distributed throughout the global human population. In other words, humans are pretty much one big, interbreeding group, with few genetic characteristics that could distinguish any "groups." (See https://www.ncbi.nlm.nih.gov/books/NBK20363/)

Where does Human Genetic Variation Come From?

Genetic variation refers to differences in genomes across members of a population or species. Genetic variation arises in a population in several ways. Mutations (uncorrected replication errors) that occur during the division of sex cells can result in alterations of gene activity or protein function that are passed down to the offspring. Genetic recombination (the re-arranging of genes during sexual reproduction) and gene flow (the introduction of genes from other populations) also increase genetic variation.

Given our population size- over 7 billion humans living on the planet today (see popclock for the most up to date numbers)- we have relatively low genetic variation. This may be because the human species is relatively new, that human populations were historically very small, or that small groups of humans moved out of Africa to populate the entire world. The Human Genome Project and subsequent research into human genetic variation has allowed scientists to learn much more about human diversity, but many questions remain unanswered.

4.

1.4 The Genetic Basis of Evolution

Evolution is defined as a change in traits in a population over time. Small changes in the frequencies of specific traits from one generation to the next are typically referred to as **micro-evolution**. Bigger changes- such as one species diverging into two over many, many generations, are typically referred to as **macro-evolution**. In this course we will examine several mechanisms by small changes can happen within populations over generations (micro-evolution) as well as look at how these changes accumulate to create the diverse species we see today (macro-evolution). We will focus on human evolution- understanding how humans relate to other species and what this means about our characteristics, as well as how humans continue to evolve. In addition to this text, your lab manual provides a basic overview of mechanisms of evolution with examples.

Evolution is the Source of New Species

All species of living organisms evolved at some point from a common ancestor. Although it may seem that living things today stay much the same from generation to generation, that is not the case: evolution is ongoing. Evolution is the process through which the characteristics of species change and through which new species arise.

The theory of evolution is a unifying theory of biology, meaning it is a 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.” [1] He meant that the principle that all life has evolved and diversified from a common ancestor is the foundation from which we understand all other questions in biology. This chapter will explain some of the mechanisms for evolutionary change and the kinds of questions that biologists can and have answered using evolutionary theory.

Natural Selection is a Mechanism of Evolution

The theory of evolution by natural selection describes a mechanism for species change over time. That species change had been suggested and debated well before Darwin. The view that species were static and unchanging was grounded in the writings of Plato, yet there were also ancient Greeks that expressed evolutionary ideas.

1. Theodosius Dobzhansky. “Biology, Molecular and Organismic.” American Zoologist 4, no. 4 (1964): 449.

Charles Darwin and Natural Selection

Natural selection as a mechanism for evolution was independently conceived of and described by two naturalists, Charles Darwin and Alfred Russell Wallace, in the mid-nineteenth century. Importantly, each 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*, visiting 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 in 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 that each had a unique beak shape (***Figure 1***). He observed both that these finches closely resembled another finch species on the mainland of South America and that the group of species in the Galápagos formed a graded series of beak sizes and shapes, with very small differences between the most similar. Darwin imagined that the island species might be all species modified from one original mainland species. In 1860, he wrote, "Seeing this gradation and diversity of structure in one small, intimately related group of birds, one might really fancy that from an original paucity of birds in this archipelago, one species had been taken and modified for different ends.[2]

2. Charles Darwin, Journal of Researches into the Natural History and Geology of the Countries Visited during the Voyage of H.M.S. Beagle Round the World, under the Command of Capt. Fitz Roy, R.N, 2nd. ed. (London: John Murray, 1860), http://www.archive.org/details/journalofresea00darw.

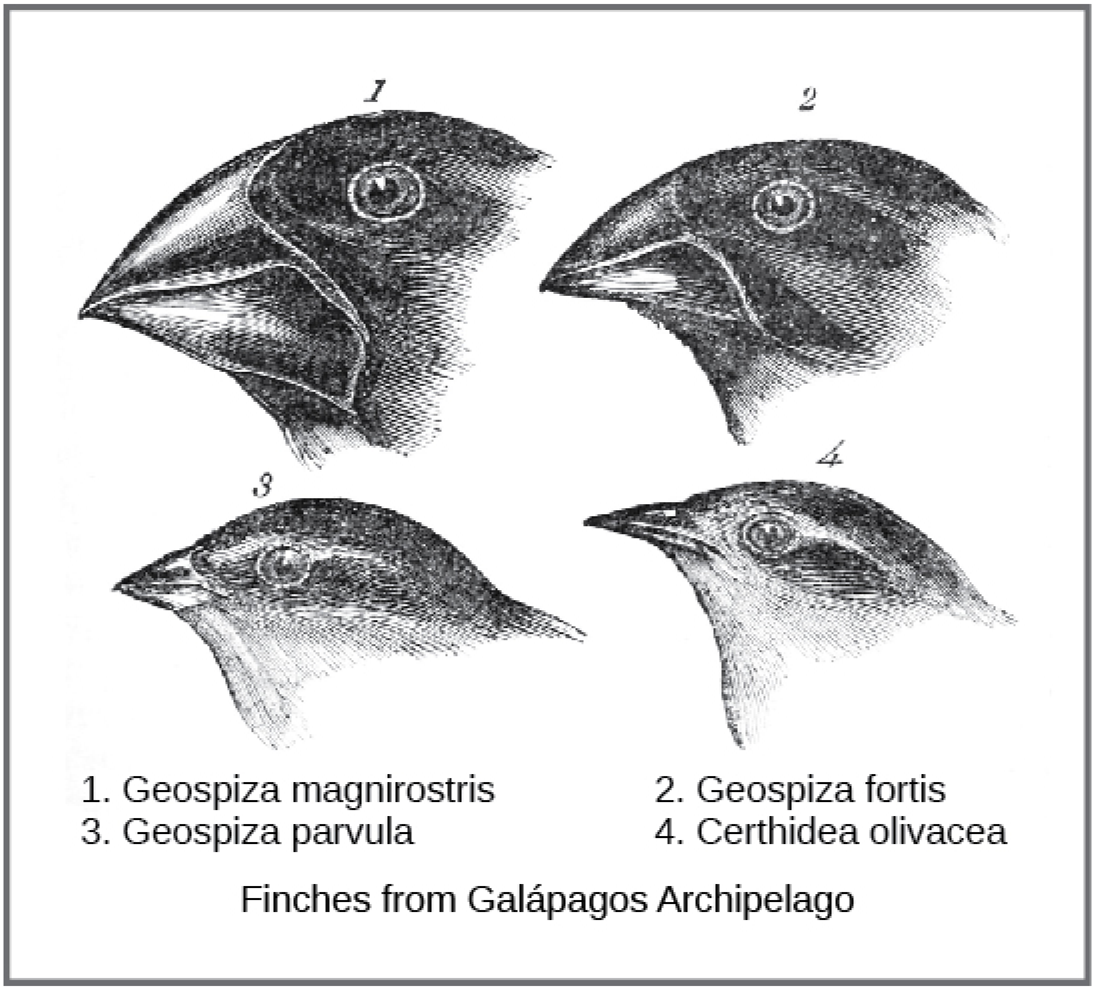


Figure 1. *Darwin observed that beak shape varies among finch species. He postulated that the beak of an ancestral species had adapted over time to equip the finches to acquire different food sources. This illustration shows the beak shapes for four species of ground finch: 1. Geospiza magnirostris (the large ground finch), 2. G. fortis (the medium ground finch), 3. G. parvula (the small tree finch), and 4. Certhidea olivacea (the green-warbler finch).*

Wallace and Darwin both observed similar patterns in other organisms and independently conceived a mechanism to explain how and why such changes could take place. Darwin called this mechanism natural selection. **Natural selection**, Darwin argued, was an inevitable outcome of three principles that operated in nature. First, the characteristics of organisms are inherited, or passed from parent to offspring. Second, more offspring are produced than are able to survive; in other words, 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 a 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. Out of these three principles, 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.”

Papers by Darwin and Wallace (***Figure 2***) presenting the idea of natural selection were read together in 1858 before the Linnaean Society in London. The following year Darwin’s book, *On the Origin of Species,* was published, which outlined in considerable detail his arguments for evolution by natural selection.

(a) (b)

Figure 2. (a) *Charles Darwin and* ***(b)*** *Alfred Wallace wrote scientific papers on natural selection that were presented together before the Linnean Society in 1858.*

Demonstrations of evolution by natural selection can be time consuming. One of the best demonstrations has been in the very birds that helped to inspire the 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 the operation of natural selection. The Grants found changes from one generation to the next in the beak shapes of the medium ground finches on the Galápagos island of Daphne Major. The medium ground finch feeds on seeds. The birds have inherited variation in the bill shape with some individuals having wide, deep bills and others having thinner bills. Large-billed birds feed more efficiently on large, hard seeds, whereas smaller billed birds feed more efficiently on small, soft seeds. During 1977, a drought period altered vegetation on the island. After this period, the number of seeds declined dramatically: the decline in small, soft seeds was greater than the decline in large, hard seeds. The large-billed birds were able to survive better than the small-billed birds the following year. The year following the drought when the Grants measured beak sizes in the much-reduced population, they found that the average bill size was larger (***Figure 3***). This was clear evidence for natural selection (differences in survival) of bill size caused by the availability of seeds. The

Grants had studied the inheritance of bill sizes and knew that the surviving large-billed birds would tend to produce offspring with larger bills, so the selection would lead to evolution of bill size. Subsequent studies by the Grants have demonstrated selection on and evolution of bill size in this species in response to changing conditions on the island. The evolution has occurred both to larger bills, as in this case, and to smaller bills when large seeds became rare.

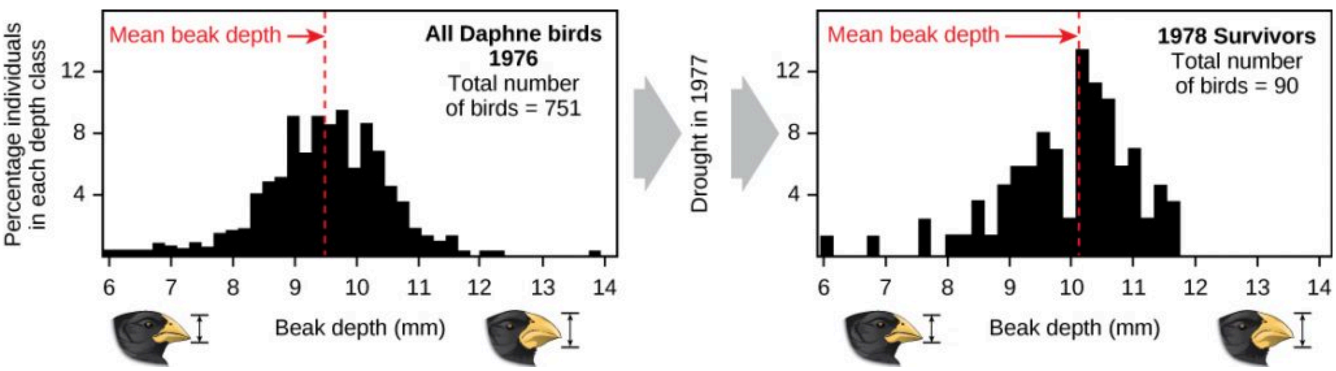


Figure 3. *A drought on the Galápagos island of Daphne Major in 1977 reduced the number of small seeds available to finches, causing many of the small-beaked finches to die. This caused an increase in the finches' average beak size between 1976 and 1978.*

Variation and Adaptation

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, 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 sources: mutation and sexual reproduction. Mutation, a change in DNA, is the ultimate source of new genetic variation in any population. An individual that has a mutated gene might have a different trait than other individuals in the population. However, this is not always the case. A mutation can have one of three outcomes on the organisms' appearance (or phenotype):

- A mutation may affect an organism's traits in a way that gives it reduced fitness—lower likelihood of survival, resulting in fewer offspring.
- A mutation may produce a trait with a beneficial effect on fitness.
- Many mutations, called neutral mutations, will have no effect on fitness.

Sexual reproduction can also generate novel combinations of traits that may have positive or negative effects on the survival of offspring. For example, your DNA is organized into 23 pairs of chromosomes–one member of each pair is from your mother, and one from your father. Since you inherit only half of your mother's chromosomes and only half of your father's chromosomes, and the exact chromosomes you get from each is determined by chance, you are a unique combination of your parents, with traits slightly different from either of parent. This re-combination of DNA at each generation gives sexually reproducing organisms like us some guaranteed variation in our populations.

A heritable trait that aids the survival and reproduction of an organism in its present environment is called an **adaptation**. An adaptation is a "match" of the organism to the environment. Adaptation to an environment comes about when a change in the range of genetic variation occurs over time that increases or maintains the match of the population with its environment. The variations in finch beaks shifted from generation to generation providing adaptation to food availability.

Whether or not a trait is favorable depends on the environment at the time. The same traits do not always have the same relative benefit or disadvantage because environmental conditions can change. For example, finches with large bills were benefited in one climate, while small bills were a disadvantage; in a different climate, the relationship reversed.

Patterns of Evolution

The evolution of species has resulted in enormous variation in form and function. When two species evolve in different directions from a common point, it is called **divergent evolution**. Such divergent evolution can be seen in the forms of the reproductive organs of flowering plants, which share the same basic anatomies; however, they can look very different as a result of selection in different physical environments, and adaptation to different kinds of pollinators (***Figure 4***).

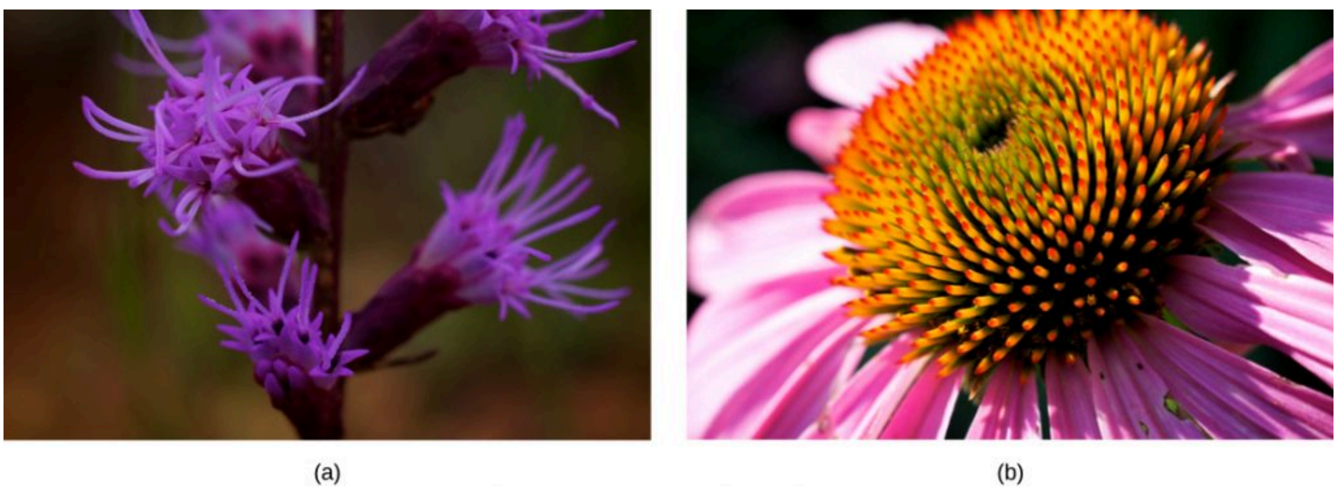


Figure 4. *Flowering plants evolved from a common ancestor. Notice that the (a) dense blazing star and (b) purple coneflower vary in appearance, yet both share a similar basic morphology. (credit a, b: modification of work by Cory Zanker)*

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. The wings of bats and insects, however, evolved from very different original structures. When similar structures arise through evolution independently in different species it is called **convergent evolution**. The wings of bats and insects are called **analogous structures**; they are similar in function and appearance, but do not share an origin in a common ancestor. Instead they evolved independently in the two lineages. The wings of a hummingbird and an ostrich are **homologous structures,** meaning they share similarities (despite their differences resulting from evolutionary divergence). The wings of hummingbirds and ostriches did not evolve independently in the hummingbird lineage and the ostrich lineage—they descended from a common ancestor with wings.

The Modern Synthesis

The mechanisms of inheritance, genetics, were not understood at the time Darwin and Wallace were developing their idea of natural selection. This lack of understanding was a stumbling block to comprehending many aspects of evolution. Darwin and Wallace were unaware of the genetics work by Austrian monk Gregor Mendel, which was published in 1866, not long after publication of *On the Origin of Species*. Mendel's work, which described the genetic basis of inheritance, was rediscovered in the early twentieth century. and integrated in what became known as the **modern synthesis**—the coherent understanding of the relationship between natural selection and genetics. The modern synthesis describes how evolutionary pressures, 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 gradual 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**.

Population Genetics

Until now, we have defined evolution as a change in the characteristics of a population of organisms, but behind that change in characteristics is genetic change. In population genetic terms, evolution is defined as a change in the frequency of specific gene in a population. Using the ABO blood system as an example, the frequency of the gene that codes for A blood protein, I^A, is the number of copies of that gene divided by the total number of all A, B or O blood protein coding genes in the population. For example, a study in Jordan found a frequency of I^A to be 26.1 percent.[3] The I^B, I^0 blood coding genes made up 13.4 percent and 60.5 percent of the blood protein coding genes respectively, and all of the frequencies add up to 100 percent. A change in this frequency over time would constitute evolution in the population.

One way the frequency of a particular gene in a population can change is natural selection. If the gene confers a trait that allows an individual to have more offspring that survive and reproduce, that gene, by virtue of being inherited by those offspring, will be in greater frequency in the next generation. Since gene frequencies always add up to 100 percent, an increase in the frequency of one gene always means a corresponding decrease in one or more of the other genes. Highly beneficial genes may, over a very few generations, become "fixed" in this way, meaning that every individual of the population will carry the gene. Similarly, detrimental genes may be swiftly eliminated from the **gene pool**, the sum of all the genes in a population. Part of the study of population genetics is tracking how selective forces change the frequencies of certain genes in a population over time, which can give scientists clues regarding the selective forces that may be operating on a given population. The studies of changes in wing coloration in the peppered moth from mottled white to dark in response to soot-covered tree trunks and then back to

3. Sahar S. Hanania, Dhia S. Hassawi, and Nidal M. Irshaid, "Allele Frequency and Molecular Genotypes of ABO Blood Group System in a Jordanian Population," Journal of Medical Sciences 7 (2007): 51-58, doi:10.3923/jms.2007.51.58

mottled white when factories stopped producing so much soot is a classic example of studying evolution in natural populations (***Figure 5***).

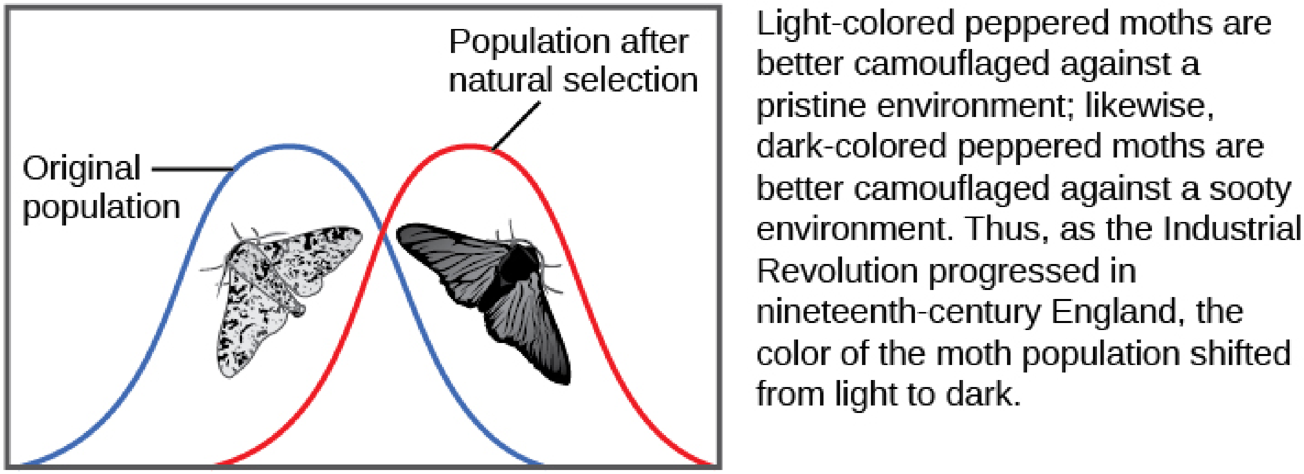


Figure 5. *As the Industrial Revolution caused trees to darken from soot, darker colored peppered moths were better camouflaged than the lighter colored ones, which caused there to be more of the darker colored moths in the population.*

Section Summary

Evolution by natural selection arises from three conditions: individuals within a species vary, some of those variations are heritable, and organisms have more offspring than resources can support. The consequence is that individuals with relatively advantageous variations will be more likely to survive and have higher reproductive rates than those individuals with different traits. The advantageous traits will be passed on to offspring in greater proportion. Thus, the trait will have higher representation in the next and subsequent generations leading to genetic change in the population.

Glossary

adaptation
: a heritable trait or behavior in an organism that aids in its survival in its present environment

analogous structure
: a structure that is similar because of evolution in response to similar selection pressures resulting in convergent evolution, not similar because of descent from a common ancestor

convergent evolution
: an evolution that results in similar forms on different species

divergent evolution
: an evolution that results in different forms in two species with a common ancestor

gene pool
: all of the alleles carried by all of the individuals in the population

homologous structure
a structure that is similar because of descent from a common ancestor

macroevolution
a broader scale of evolutionary changes seen over paleontological time

microevolution
the changes in a population's genetic structure (i.e., allele frequency)

modern synthesis
the overarching evolutionary paradigm that took shape by the 1940s and is generally accepted today

natural selection
the greater relative survival and reproduction of individuals in a population that have favorable heritable traits, leading to evolutionary change

population genetics
the study of how selective forces change the allele frequencies in a population over time

variation
the variety of traits in a population

5.

1.5 Mechanisms of Evolution

When certain genes become more or less common in the population over generations, we refer to this change as evolution. Although natural selection is the mechanism of evolution most commonly discussed, other evolutionary mechanisms also change the frequencies of traits (and the genes that control them) in populations. These include mutation, genetic drift and migration.

Natural Selection

Natural selection, discussed in the previous chapter, is the mechanism of evolution that explains how species can become better adapted to their environment. Depending on the environmental conditions, certain traits may confer an advantage or disadvantage to the individuals that possess them, relative to others in the population. If a certain trait confers an advantage, then the individual possessing the trait may have more offspring than those with other traits. If the trait is heritable, then the genes that give rise to the trait will be more common in the next generation. If conditions remain the same, those offspring, which are carrying the same trait, will also benefit, and pass the genes that give rise to this trait on to their own offspring. Over time, the advantageous trait (aka adaptation) will become more common in the population.

Mutation

Mutation is a source of variation in a population. Mutation is a change in the DNA sequence of the gene. In some cases a change in the DNA will change the protein produced. The change in frequency resulting from a mutation in one individual is small, so its effect on evolution is small unless it interacts with one of the other factors, such as selection. A mutation may produce an allele that is selected against, selected for, or selectively neutral. Harmful mutations are removed from the population by selection and will generally only be found in very low frequencies equal to the mutation rate. Beneficial mutations will spread through the population through selection, although that initial spread is slow. Whether or not a mutation is beneficial or harmful is determined by whether it helps an organism survive to sexual maturity and reproduce. It should be noted that mutation is the ultimate source of genetic variation in all populations—new alleles, and, therefore, new genetic variations arise through mutation.

Genetic Drift

Another way the frequencies of certain genes can change is genetic drift (***Figure 1***), which is simply the effect of chance. Genetic drift is most important in small populations. Because the genes in an offspring generation are a random sample of the genes in the parent generation, some versions of a gene may not make it into the next generation due to chance events. If one individual in a population of ten individuals happens to die before it leaves any offspring to the next generation, all of its genes—a tenth of the population's gene pool—will be suddenly lost. In a population of 100, that 1 individual represents only 1 percent of the overall gene pool; therefore, it has much less impact on the population's genetic structure and is unlikely to remove all copies of even a relatively rare gene.

Imagine a population of ten individuals, half with a version of a gene we will call *A* and half with a version of a gene we will call *a*. In a stable population, the next generation will also have ten individuals. Choose that generation randomly by flipping a coin ten times and let heads be *A* and tails be *a*. It is unlikely that the next generation will have exactly half of each gene. There might be six of one and four of the other, or some different set of frequencies. Thus, the frequencies have changed and evolution has occurred. A coin will no longer work to choose the next generation (because the odds are no longer one half for each gene). The frequency in each generation will drift up and down on what is known as a random walk until at one point either all *A* or all *a* are chosen and that version of the gene is fixed from that point on. This could take a very long time for a large population. The effect of drift on frequencies is greater the smaller a population is.

ART CONNECTION

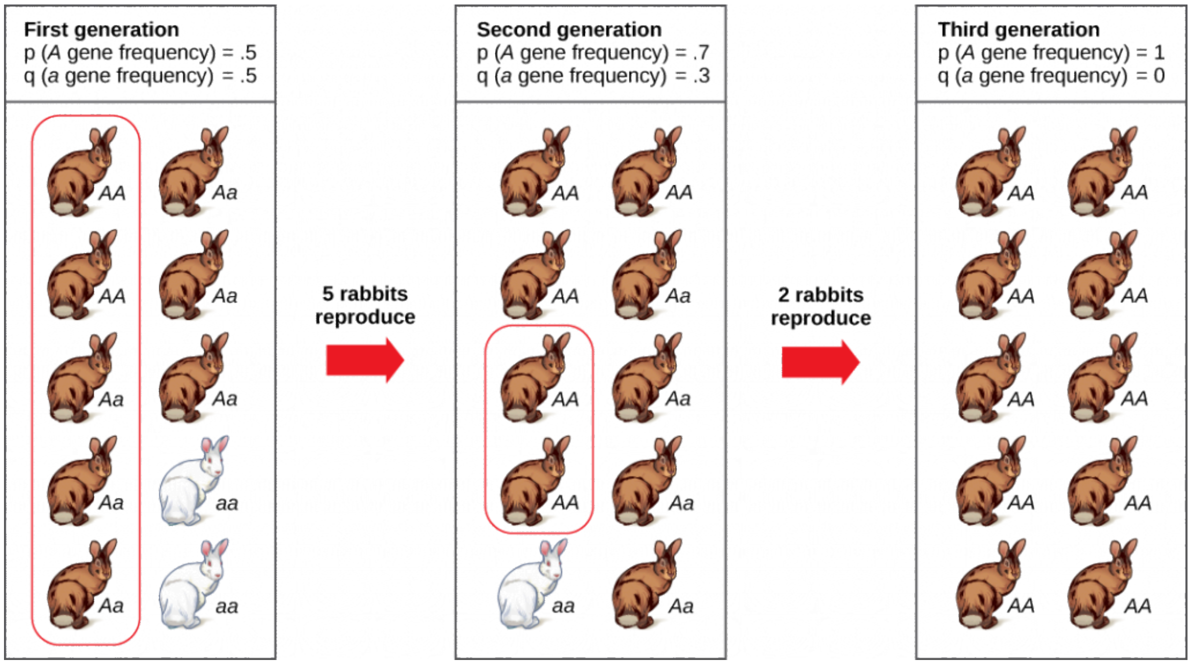


Figure 1. *Genetic drift in a population can lead to the elimination of an allele from a population by chance. In each generation, a random set of individuals reproduces to produce the next generation. The frequency of alleles in the next generation is equal to the frequency of alleles among the individuals reproducing.*

Genetic drift can also be magnified by natural or human-caused events, such as a disaster that randomly kills a large portion of the population, which is known as the **bottleneck effect** that results in a large portion of the gene pool suddenly being wiped out (***Figure 2***). 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. The disaster must be one that kills for reasons unrelated to the organism's traits, such as a hurricane or lava flow.

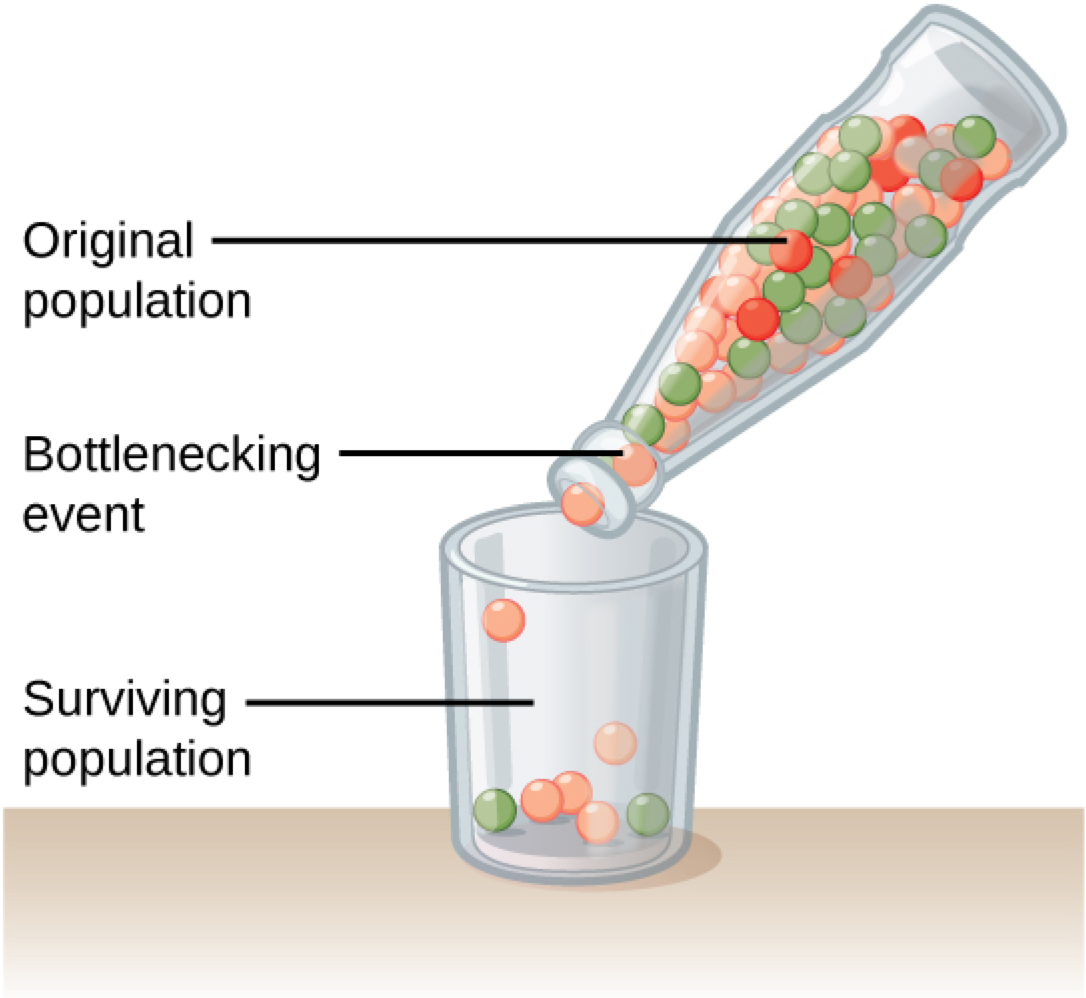


Figure 2. *A chance event or catastrophe can reduce the genetic variability within a population.*

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 matches 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 a higher-than-normal proportion of the founding colonists, which were a small sample of the original population, 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 bone marrow and congenital abnormalities, and even cancer[1].

1. A. J. Tipping et al., "Molecular and Genealogical Evidence for a Founder Effect in Fanconi Anemia Families of the Afrikaner Population of South Africa," *PNAS* 98, no. 10 (2001): 5734-5739, doi: 10.1073/pnas.091402398.

Gene Flow

Another important evolutionary force is **gene flow**, or the flow of genes in and out of a population resulting from the migration of individuals or gametes (***Figure 3***). While some populations are fairly stable, others experience more flux. Many plants, for example, send their seeds far and wide, by wind or in the guts of animals; these seeds may introduce genes common in the source population to a new population in which they are rare.

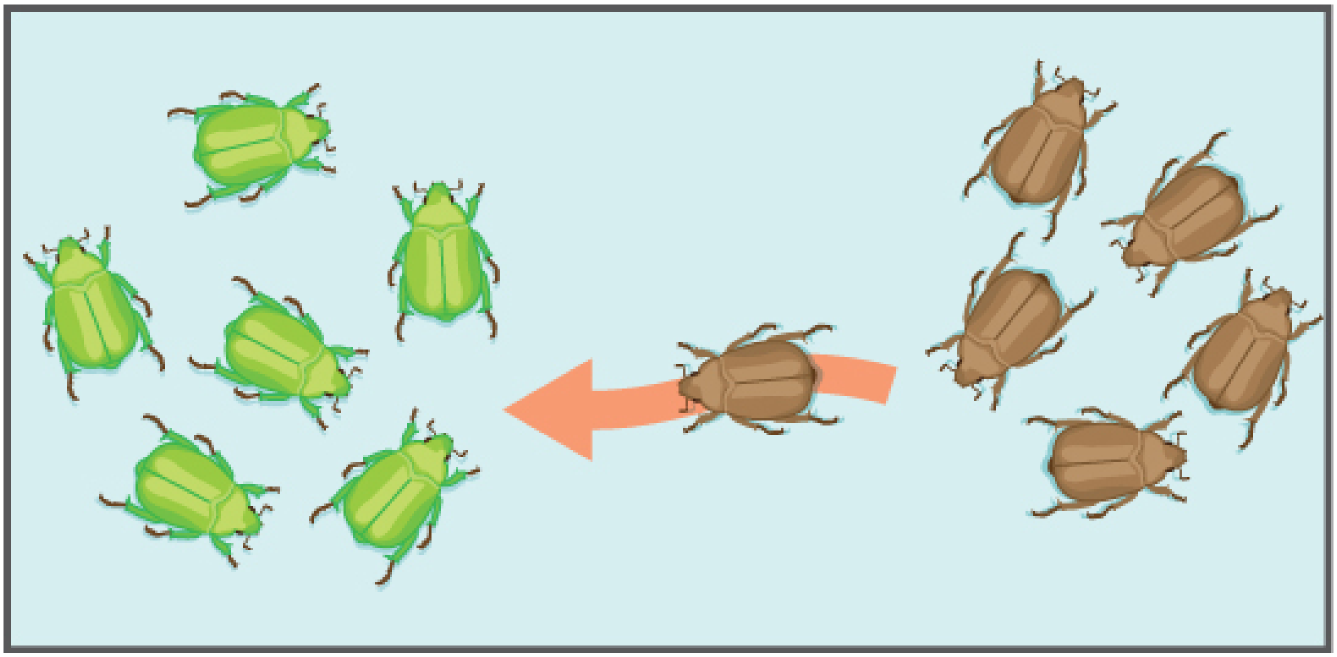


Figure 3. *Gene flow can occur when an individual travels from one geographic location to another and joins a different population of the species. In the example shown here, the brown allele is introduced into the green population.*

Section Summary

There are four factors that can change the frequencies of genes in a population. Natural selection works by selecting for genes that confer beneficial traits or behaviors, while selecting against those for deleterious qualities. Mutations introduce new versions of genes into populations. Genetic drift stems from the chance occurrence that some individuals have more offspring than others and results in changes in gene frequencies that are random in direction. When individuals leave or join the population, gene frequencies can change as a result of gene flow.

Glossary

bottleneck effect
: the magnification of genetic drift as a result of natural events or catastrophes

founder effect
: a magnification of genetic drift in a small population that migrates away from a large parent population carrying with it an unrepresentative set of alleles

gene flow
: the movement of genetic information across populations due to the migration of individuals or gametes (sex cells- such as a plant’s pollen)

genetic drift
: random changes in trait frequencies over time

6.

1.6 Introduction to Phylogenies

Evolution is defined as the gradual change in characteristics of a population of organisms over generations. As changes accumulate, new species can form. A **phylogeny** describes the relationships among groups of organisms (such as which groups are most closely related, which diverged most recently from a common ancestor, etc. The groups of organisms being compared in a phylogeny might be populations, species or groups of species such as genera, kingdoms, etc. When comparing groups, we often use the generic term **taxa** (singular, taxon) which could refer to any of these levels of groups. Phylogenetic relationships provide information on shared ancestry among taxa.

Phylogenetic Trees

Scientists use a diagram called a phylogenetic tree to show the evolutionary pathways and connections among taxa. A phylogenetic tree is 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***).

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 miniscule 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 shows 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)*

The diagrams can help us understand evolutionary history. The pathway can be traced from the origin of life to any individual taxon by navigating through the evolutionary branches between the two points. Also, by starting with a single taxon and tracing back towards the "trunk" of the tree, one can discover that taxon's ancestors, as well as where lineages share a common ancestry.

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. 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. Many phylogenietic trees are being re-evaluated in light of our increased ability to examine DNA evidence.

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 (***Figure 2***) shows that lizards and rabbits both have amniotic eggs, whereas frogs do not; yet lizards and frogs appear more similar than lizards and rabbits.

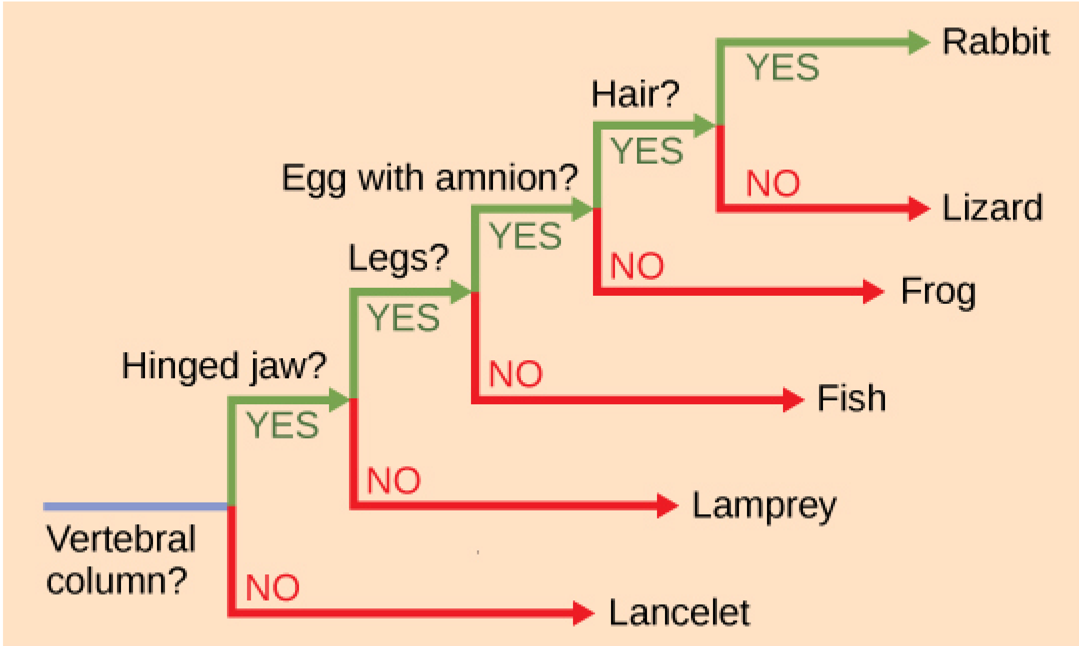


Figure 2. *This ladder-like phylogenetic tree of vertebrates is rooted by an organism that lacked a vertebral column. At each branch point, organisms with different characters are placed in different groups based on the characteristics they share.*

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 2***, 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 2***, 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 2***, 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 3***).

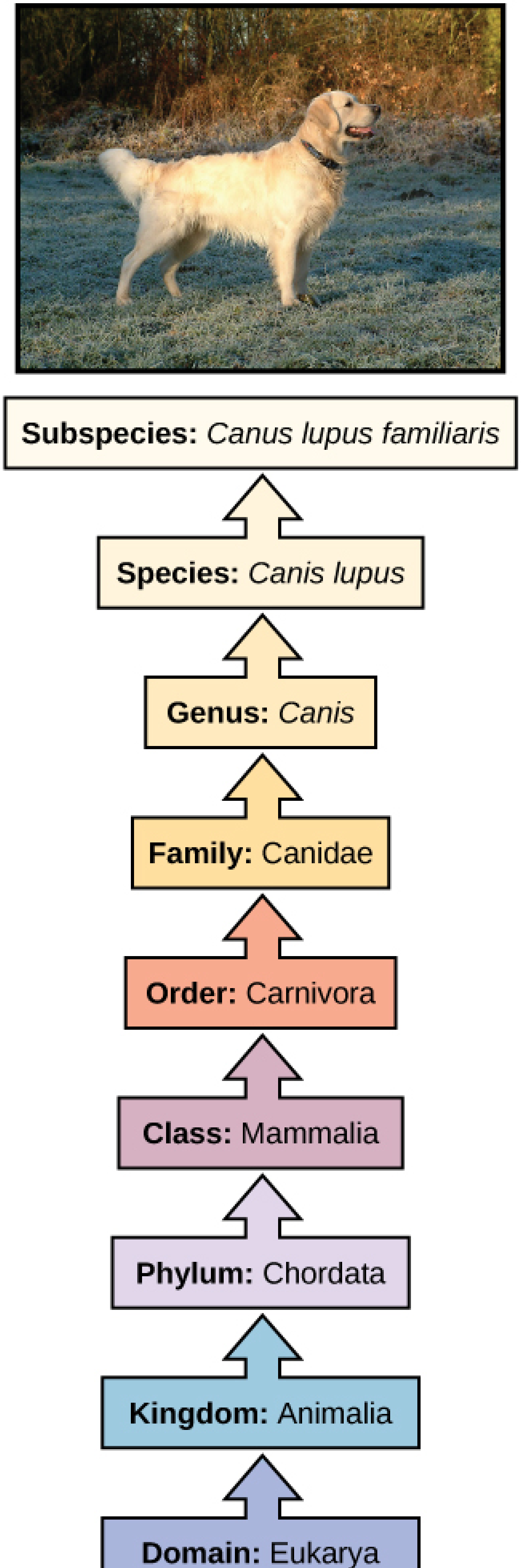

Subspecies: Canus lupus familiaris
Species: Canis lupus
Genus: Canis
Family: Canidae
Order: Carnivora
Class: Mammalia
Phylum: Chordata
Kingdom: Animalia
Domain: Eukarya

Figure 3. *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)*

Figure 4 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.

ART CONNECTION

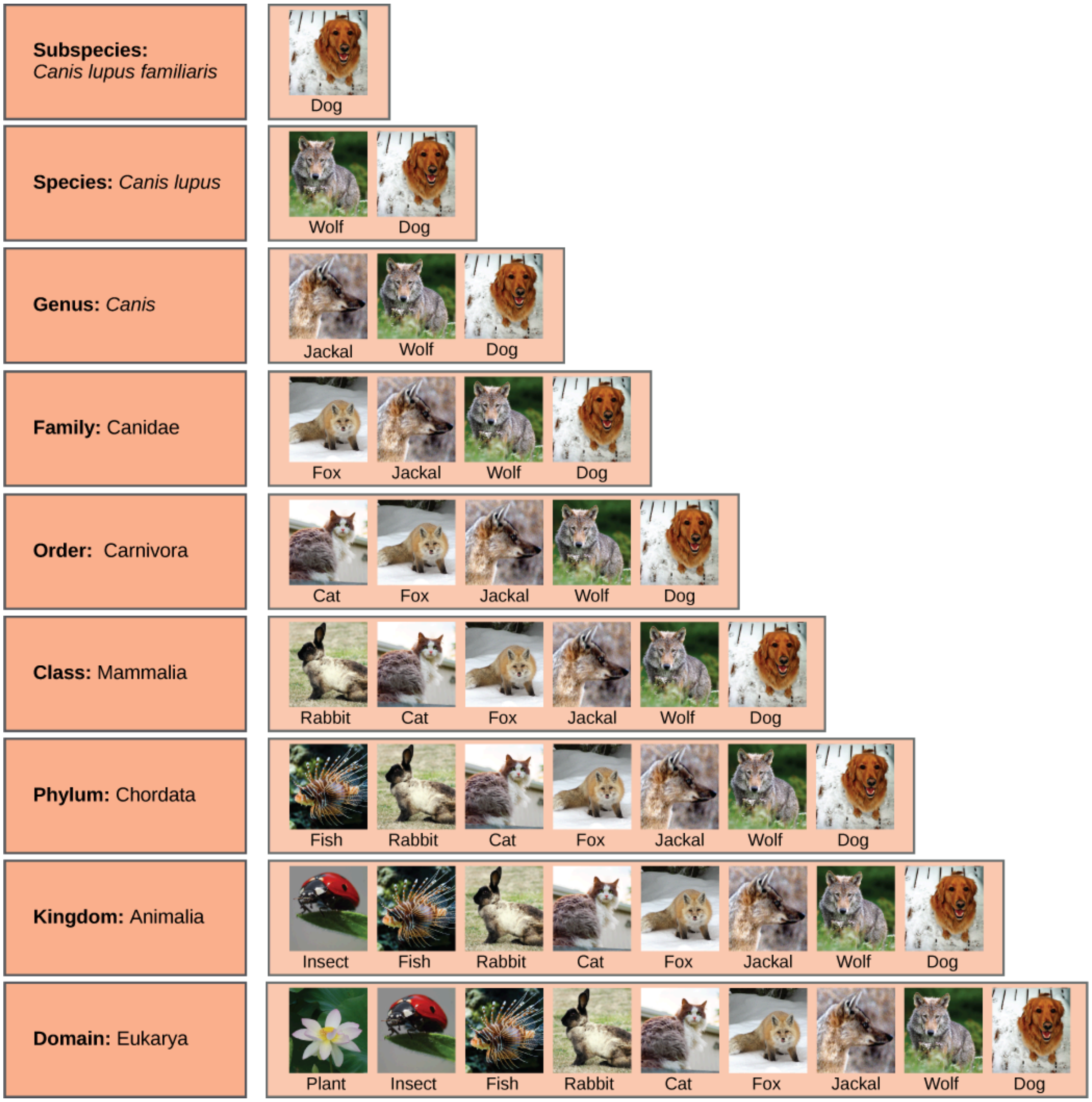


Figure 4. *At each sublevel in the taxonomic classification system, organisms become more similar. Dogs and wolves are the same species because they can breed and produce viable offspring, but they are different enough to be classified as different subspecies. (credit "plant": modification of work by "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)*

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.

Section 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.

Glossary

branch point
: node on a phylogenetic tree where a single lineage splits into distinct new ones

phylogenetic tree
: diagram used to reflect the evolutionary relationships among organisms or groups of organisms

phylogeny
: evolutionary history and relationship of an organism or group of organisms

taxon
: (plural: taxa) group of organisms (could be a species, genus, order, etc.)

taxonomy
: science of classifying organisms

7.

1.7 How Phylogenies are Made

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 are more closely related than those that do not. We refer to such features that overlap both morphologically (in form) and genetically as homologous structures. They stem from developmental similarities that are based on evolution. For example, the bones in bat and bird wings have homologous structures (***Figure 1***).

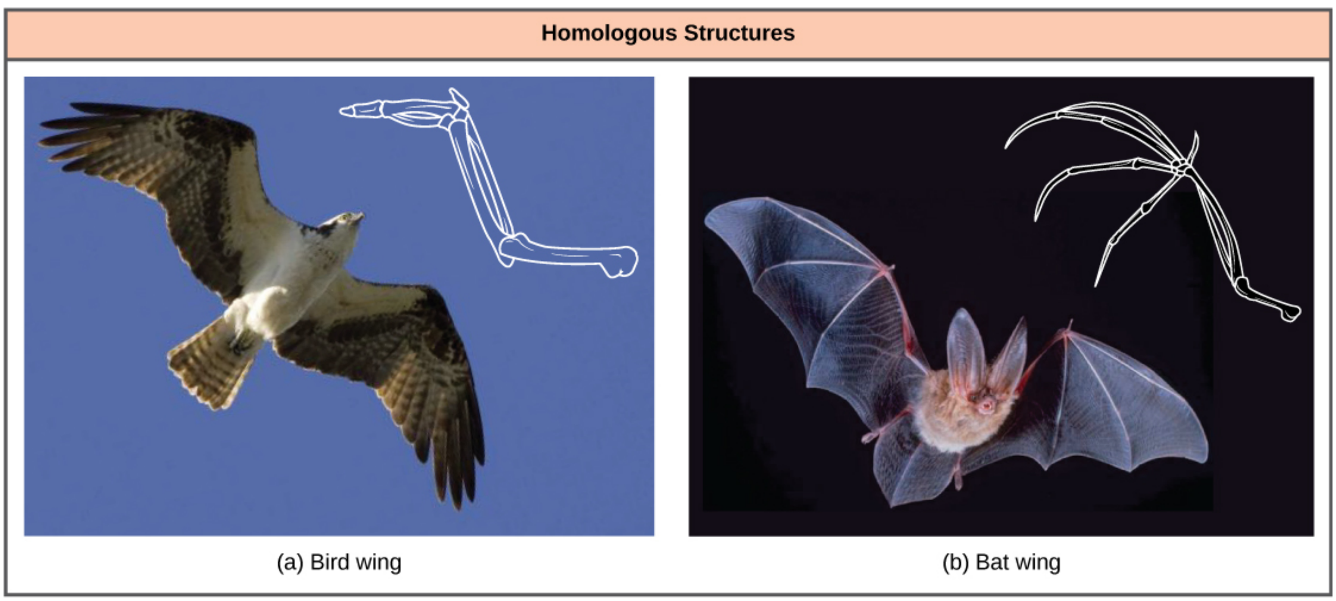


Figure 1. *Bat and bird wings are homologous structures, indicating that bats and birds share a common evolutionary past. (credit a: modification of work by Steve Hillebrand, USFWS; credit b: modification of work by U.S. DOI BLM)*

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, we can reasonably 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 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 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 a whale's front flipper are homologous to the bones in the human arm. These structures are not analogous. A butterfly or bird's wings are analogous but not homologous. Some structures are both analogous and homologous: bird and bat wings are both homologous and analogous. Scientists must determine which type of similarity a feature exhibits to decipher the organisms' phylogeny.

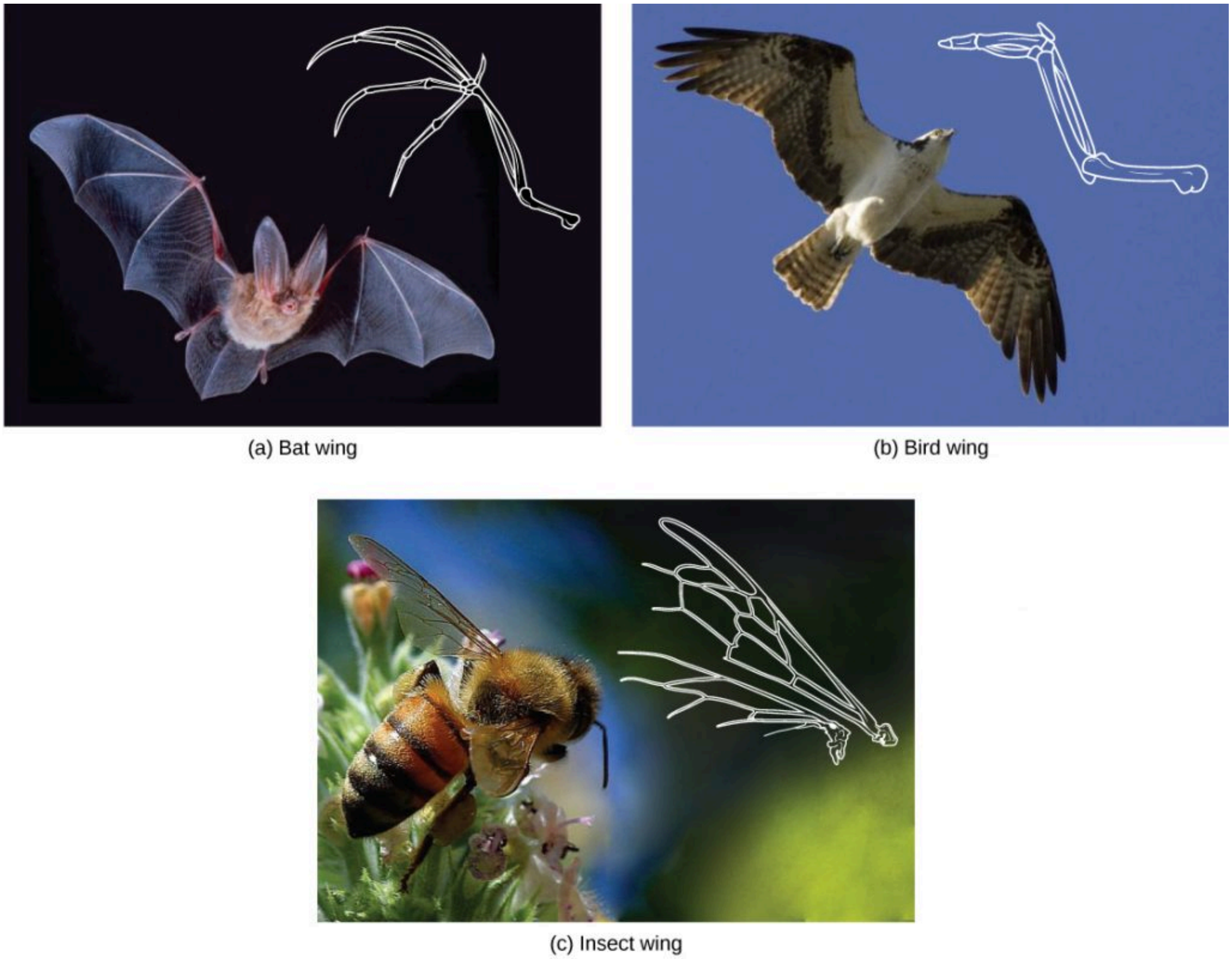


Figure 2. The ***(c)*** *wing of a honeybee is similar in shape to a* ***(b)*** *bird wing and* ***(a)*** *bat wing, and it serves the same function. However, the honeybee wing is not composed of bones and has a distinctly different structure and embryonic origin. These wing types (insect versus bat and bird) illustrate an analogy—similar structures that do not share an evolutionary history. (credit a: modification of work by U.S. DOI BLM; credit b: modification of work by Steve Hillebrand, USFWS; credit c: modification of work by Jon Sullivan)*

Molecular Comparisons

The advancement of DNA technology has given rise to **molecular systematics**, which is use of molecular data in taxonomy and biological geography (biogeography). 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. Inserting or deleting a 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 help identify the actual relationships, and, ultimately, the coupled use of both morphologic and molecular information is more effective in determining phylogeny.

EVOLUTION CONNECTION

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 compounds for other useful drugs.

A research team in China identified a DNA segment that they thought to be common to some medicinal plants in the family Fabaceae (the legume family). They 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.

XXIV.

W.H.Fitch del et lith.

M&N Hanhart imp.

Dalbergia Sissoo, *Roxb.*

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, researchers found that D. sissoo had fungicidal activity, supporting the idea that DNA markers are useful to screen plants with potential medicinal properties.*

Building Phylogenetic Trees

How do scientists construct phylogenetic trees? After they sort the homologous and analogous traits, scientists often organize the homologous traits using **cladistics**. This system sorts organisms into clades: groups of organisms that descended from a single ancestor. For example, in ***Figure 4***, all the organisms in the orange region evolved from a single ancestor that had amniotic eggs. Consequently, these organisms also have amniotic eggs and make a single clade, or a **monophyletic group**. Clades must include all descendants from a branch point.

VISUAL CONNECTION

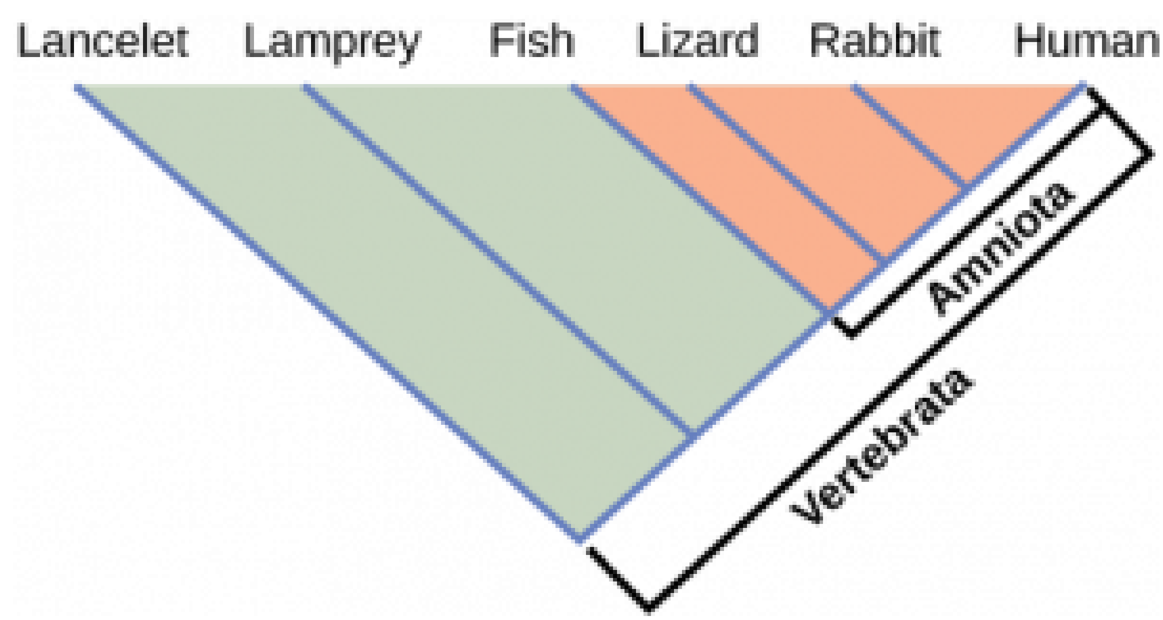


Figure 4. *Lizards, rabbits, and humans all descend from a common ancestor that had an amniotic egg. Thus, lizards, rabbits, and humans all belong to the clade Amniota. Vertebrata is a larger clade that also includes fish and lamprey.*

Which animals in this figure belong to a clade that includes animals with hair? Which evolved first, hair or the amniotic egg?

Shared Traits

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 as one goes through the phylogenetic tree of life:

1. A change in an organism's genetic makeup 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).

The tricky aspect to shared ancestral and shared derived traits is that these terms are relative. We can consider the same trait one or the other depending on the particular diagram that we use. 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 building phylogenetic trees.

Choosing the Right Relationships

To aid in the tremendous task of describing phylogenies accurately, scientists often use the concept of **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 hike, based on the principle of maximum parsimony, one could predict that most 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.

Section 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). Scientists can use newer technologies to help distinguish homologies from analogies. After identifying homologous information, scientists use cladistics to organize these events as a means to determine an evolutionary timeline. They then 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.

Glossary

analogy
: (also, homoplasy) characteristic that is similar between organisms by convergent evolution, not due to the same evolutionary path

cladistics
: system to organize homologous traits to describe phylogenies

maximum parsimony
: applying the simplest, most obvious way with the least number of steps

molecular systematics
: technique using molecular evidence to identify phylogenetic relationships

monophyletic group
: (also, clade) organisms that share a single ancestor

shared ancestral trait
: describes a characteristic on a phylogenetic tree that all organisms on the tree share

shared derived trait
: describes a characteristic on a phylogenetic tree that only a certain clade of organisms share

8.

1.8 The Evolution of Primates

Order **Primates** of class Mammalia includes lemurs, tarsiers, monkeys, apes, and humans. Non-human primates live primarily in the tropical or subtropical regions of South America, Africa, and Asia. They range in size from the mouse lemur at 30 grams (1 ounce) to the mountain gorilla at 200 kilograms (441 pounds). The characteristics and evolution of primates is of particular interest to us as it allows us to understand the evolution of our own species.

Characteristics of Primates

All primate species possess adaptations for climbing trees, as they all descended from tree-dwellers. This arboreal heritage of primates has resulted in hands and feet that are adapted for **brachiation**, or climbing and swinging through trees. These adaptations include, but are not limited to: 1) a rotating shoulder joint, 2) a big toe that is widely separated from the other toes and thumbs, which are widely separated from fingers (except humans), which allow for gripping branches, 3) **stereoscopic vision**, two overlapping fields of vision from the eyes, which allows for the perception of depth and gauging distance. Other characteristics of primates are brains that are larger than those of most other mammals, claws that have been modified into flattened nails, typically only one offspring per pregnancy, and a trend toward holding the body upright.

Order Primates is divided into two groups: prosimians and anthropoids. **Prosimians** include the bush babies of Africa, the lemurs of Madagascar, and the lorises, pottos, and tarsiers of Southeast Asia. **Anthropoids** include monkeys, apes, and humans. In general, prosimians tend to be nocturnal (in contrast to diurnal anthropoids) and exhibit a smaller size and smaller brain than anthropoids.

Evolution of Primates

The first primate-like mammals are referred to as proto-primates. They were roughly similar to squirrels and tree shrews in size and appearance. The existing fossil evidence (mostly from North Africa) is very fragmented. These proto-primates remain largely mysterious creatures until more fossil evidence becomes available. The oldest known primate-like mammals with a relatively robust fossil record is ***Plesiadapis*** (although some researchers do not agree that *Plesiadapis* was a proto-primate). Fossils of this primate have been dated to approximately 55 million years ago. Plesiadapiforms were proto-primates that had some features of the teeth and skeleton in common with true primates. They were found in North America and Europe in the Cenozoic and went extinct by the end of the Eocene.

The first true primates were found in North America, Europe, Asia, and Africa in the Eocene Epoch. These early primates resembled present-day prosimians such as lemurs. Evolutionary changes continued in these early primates, with larger brains and eyes, and smaller muzzles being the trend. By the end of the Eocene Epoch, many of the early prosimian species went extinct due either to cooler temperatures or competition from the first monkeys.

Anthropoid monkeys evolved from prosimians during the Oligocene Epoch. By 40 million years ago, evidence indicates that monkeys were present in the New World (South America) and the Old World (Africa and Asia). New World monkeys are also called **Platyrrhini**—a reference to their broad noses (***Figure 1***). Old World monkeys are called **Catarrhini**—a reference to their narrow noses. There is still quite a bit of uncertainty about the origins of the New World monkeys. At the time the platyrrhines arose, the continents of South American and Africa had drifted apart. Therefore, it is thought that monkeys arose in the Old World and reached the New World either by drifting on log rafts or by crossing land bridges. Due to this reproductive isolation, New World monkeys and Old World monkeys underwent separate adaptive radiations over millions of years. The New World monkeys are all arboreal, whereas Old World monkeys include arboreal and ground-dwelling species.

Figure 1. *The howler monkey is native to Central and South America. It makes a call that sounds like a lion roaring. (credit: Xavi Talleda)*

Apes evolved from the catarrhines in Africa midway through the Cenozoic, approximately 25 million years ago. Apes are generally larger than monkeys and they do not possess a tail. All apes are capable of moving through trees, although many species spend most their time on the ground. Apes are more intelligent than monkeys, and they have relatively larger brains proportionate to body size. The apes are divided into two groups. The lesser apes comprise the family **Hylobatidae**, including gibbons and siamangs. The great apes include the genera ***Pan*** (chimpanzees and bonobos) (***Figure 2a***), ***Gorilla*** (gorillas), ***Pongo*** (orangutans), and ***Homo*** (humans) (***Figure 2b***). The very arboreal gibbons are smaller than the great apes; they have low sexual dimorphism (that is, the sexes are not markedly different in size); and they have relatively longer arms used for swinging through trees.

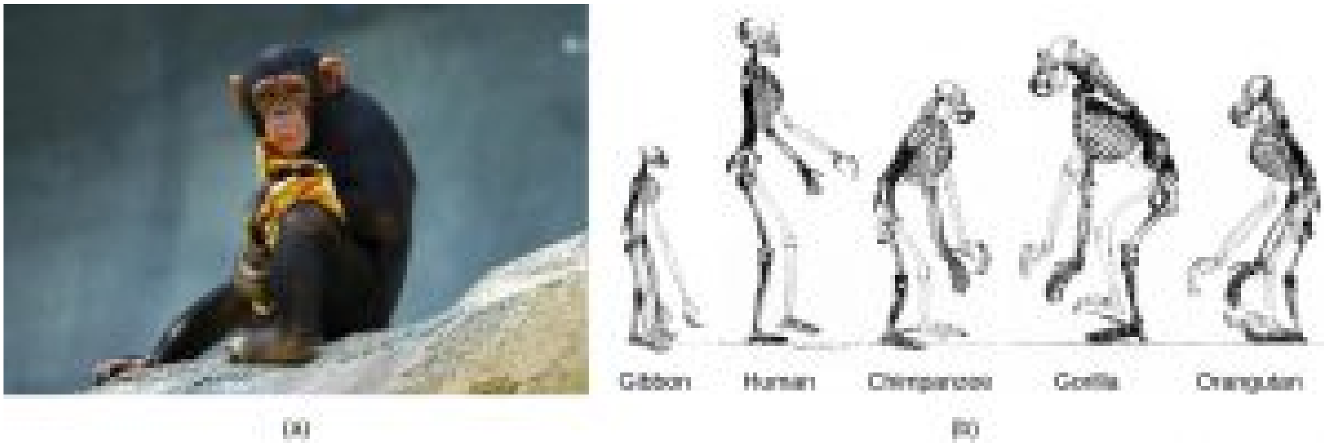


Figure 2. *The (a) chimpanzee is one of the great apes. It possesses a relatively large brain and has no tail. (b) All great apes have a similar skeletal structure. (credit a: modification of work by Aaron Logan; credit b: modification of work by Tim Vickers)*

Human Evolution

The family Hominidae of order Primates includes the **hominoids**: the great apes (***Figure 3***). Evidence from the fossil record and from a comparison of human and chimpanzee DNA suggests that humans and chimpanzees diverged from a common hominoid ancestor approximately 6 million years ago. Several species evolved from the evolutionary branch that includes humans, although our species is the only surviving member. The term **hominin** is used to refer to those species that evolved after this split of the primate line, thereby designating species that are more closely related to humans than to chimpanzees. Hominins were predominantly bipedal and include those groups that likely gave rise to our species—including *Australopithecus*, *Homo habilis*, and *Homo erectus*—and those non-ancestral groups that can be considered "cousins" of modern humans, such as Neanderthals. Determining the true lines of descent in hominins is difficult. In years past, when relatively few hominin fossils had been recovered, some scientists believed that considering them in order, from oldest to youngest, would demonstrate the course of evolution from early hominins to modern humans. In the past several years, however, many new fossils have been found, and it is clear that there was often more than one species alive at any one time and that many of the fossils found (and species named) represent hominin species that died out and are not ancestral to modern humans.

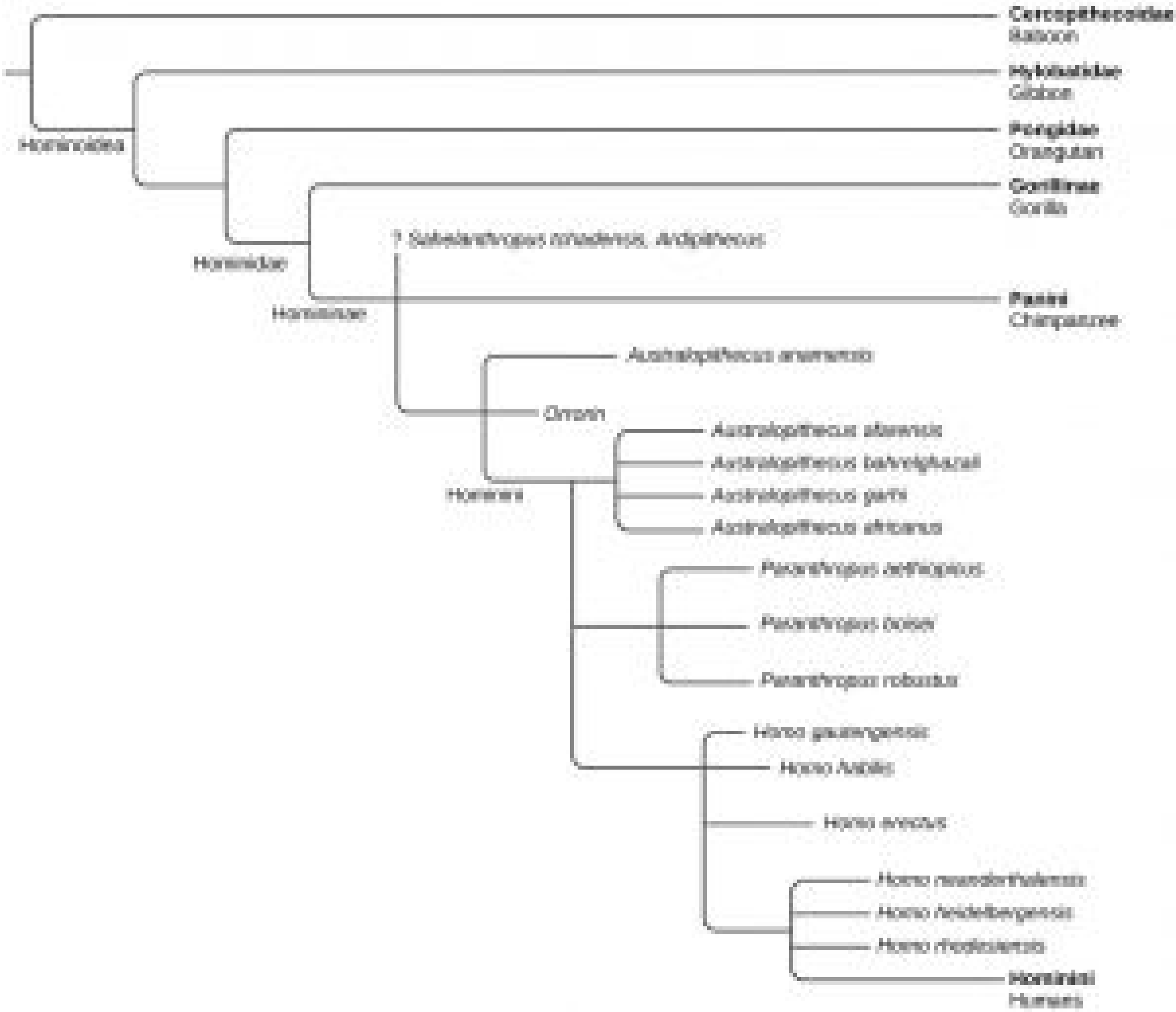


Figure 3. *This chart shows the evolution of modern humans.*

Very Early Hominins

Three species of very early hominids have made news in the past few years. The oldest of these, *Sahelanthropus tchadensis*, has been dated to nearly 7 million years ago. There is a single specimen of this genus, a skull that was a surface find in Chad. The fossil, informally called "Toumai," is a mosaic of primitive and evolved characteristics, and it is unclear how this fossil fits with the picture given by molecular data, namely that the line leading to modern humans and modern chimpanzees apparently bifurcated about 6 million years ago. It is not thought at this time that this species was an ancestor of modern humans.

A second, younger species, *Orrorin tugenensis*, is also a relatively recent discovery, found in 2000. There are several specimens of *Orrorin*. It is not known whether *Orrorin* was a human ancestor, but this possibility has not been ruled out. Some features of *Orrorin* are more similar to those of modern humans than are the australopiths, although *Orrorin* is much older.

A third genus, *Ardipithecus*, was discovered in the 1990s, and the scientists who discovered the first fossil found that some other scientists did not believe the organism to be a biped (thus, it would not be considered a hominid). In the intervening years, several more specimens of *Ardipithecus*, classified as two different species, demonstrated that the organism was bipedal. Again, the status of this genus as a human ancestor is uncertain.

Early Hominins: Genus *Australopithecus*

Australopithecus ("southern ape") is a genus of hominin that evolved in eastern Africa approximately 4 million years ago and went extinct about 2 million years ago. This genus is of particular interest to us as it is thought that our genus, genus *Homo*, evolved from a common ancestor shared with *Australopithecus* about 2 million years ago (after likely passing through some transitional states). *Australopithecus* had a number of characteristics that were more similar to the great apes than to modern humans. For example, sexual dimorphism was more exaggerated than in modern humans. Males were up to 50 percent larger than females, a ratio that is similar to that seen in modern gorillas and orangutans. In contrast, modern human males are approximately 15 to 20 percent larger than females. The brain size of *Australopithecus* relative to its body mass was also smaller than modern humans and more similar to that seen in the great apes. A key feature that *Australopithecus* had in common with modern humans was bipedalism, although it is likely that *Australopithecus* also spent time in trees. Hominin footprints, similar to those of modern humans, were found in Laetoli, Tanzania and dated to 3.6 million years ago. They showed that hominins at the time of *Australopithecus* were walking upright.

There were a number of *Australopithecus* species, which are often referred to as *australopiths*. *Australopithecus anamensis* lived about 4.2 million years ago. More is known about another early species, *Australopithecus afarensis*, which lived between 3.9 and 2.9 million years ago. This species demonstrates a trend in human evolution: the reduction of the dentition and jaw in size. *A. afarensis* (***Figure 4***) had smaller canines and molars compared to apes, but these were larger than those of modern humans. Its brain size was 380–450 cubic centimeters, approximately the size of a modern chimpanzee brain. It also had **prognathic jaws**, which is a relatively longer jaw than that of modern humans. In the mid-1970s, the fossil of an adult female *A. afarensis* was found in the Afar region of Ethiopia and dated to 3.24 million years ago (***Figure 5***). The fossil, which is informally called "Lucy," is significant because it was the most complete australopith fossil found, with 40 percent of the skeleton recovered.

(a) (b)

Figure 4. *The skull of (a) Australopithecus afarensis, an early hominid that lived between two and three million years ago, resembled that of (b) modern humans but was smaller with a sloped forehead and prominent jaw.*

Figure 5. *This adult female Australopithecus afarensis skeleton, nicknamed Lucy, was discovered in the mid 1970s. (credit: "120"/Wikimedia Commons)*

Australopithecus africanus lived between 2 and 3 million years ago. It had a slender build and was bipedal, but had robust arm bones and, like other early hominids, may have spent significant time in trees. Its brain was larger than that of *A.afarensis* at 500 cubic centimeters, which is slightly less than one-third the size of modern human brains. Two other species, *Australopithecus bahrelghazali* and*Australopithecus garhi*, have been added to the roster of australopiths in recent years.

A Dead End: Genus *Paranthropus*

The australopiths had a relatively slender build and teeth that were suited for soft food. In the past several years, fossils of hominids of a different body type have been found and dated to approximately 2.5 million years ago. These hominids, of the genus *Paranthropus*, were muscular, stood 1.3-1.4 meters tall, and had large grinding teeth. Their molars showed heavy wear, suggesting that they had a coarse and fibrous vegetarian diet as opposed to the partially carnivorous diet of the australopiths. *Paranthropus* includes *Paranthropus robustus* of South Africa, and *Paranthropus aethiopicus* and *Paranthropus boisei* of East Africa. The hominids in this genus went extinct more than 1 million years ago and are not thought to be ancestral to modern humans, but rather members of an evolutionary branch on the hominin tree that left no descendants.

Early Hominins: Genus *Homo*

The human genus, *Homo*, first appeared between 2.5 and 3 million years ago. For many years, fossils of a species called *H. habilis* were the oldest examples in the genus *Homo*, but in 2010, a new species called *Homo gautengensis* was discovered and may be older. Compared to *A. africanus*, *H. habilis* had a number of features more similar to modern humans. *H. habilis* had a jaw that was less prognathic than the australopiths and a larger brain, at 600–750 cubic centimeters. However, *H. habilis* retained some features of older hominin species, such as long arms. The name *H. habilis* means "handy man," which is a reference to the stone tools that have been found with its remains.

H. erectus appeared approximately 1.8 million years ago (***Figure 6***). It is believed to have originated in East Africa and was the first hominin species to migrate out of Africa. Fossils of *H. erectus* have been found in India, China, Java, and Europe, and were known in the past as "Java Man" or "Peking Man." *H. erectus* had a number of features that were more similar to modern humans than those of *H. habilis*. *H.erectus* was larger in size than earlier hominins, reaching heights up to 1.85 meters and weighing up to 65 kilograms, which are sizes similar to those of modern humans. Its degree of sexual dimorphism was less than earlier species, with males being 20 to 30 percent larger than females, which is close to the size difference seen in our species. *H. erectus* had a larger brain than earlier species at 775–1,100 cubic centimeters, which compares to the 1,130–1,260 cubic centimeters seen in modern human brains. *H. erectus* also had a nose with downward-facing nostrils similar to modern humans, rather than the forward facing nostrils found in other primates. Longer, downward-facing nostrils allow for the warming of cold air before it enters the lungs and may have been an adaptation to colder climates. Artifacts found with fossils of *H. erectus* suggest that it was the first hominin to use fire, hunt, and have a home base. *H. erectus* is generally thought to have lived until about 50,000 years ago.

Figure 6. *Homo erectus had a prominent brow and a nose that pointed downward rather than forward.*

Humans: *Homo sapiens*

A number of species, sometimes called archaic *Homo sapiens*, apparently evolved from *H. erectus* starting about 500,000 years ago. These species include *Homo heidelbergensis*, *Homo rhodesiensis*, and *Homo neanderthalensis*. These archaic *H.sapiens* had a brain size similar to that of modern humans, averaging 1,200–1,400 cubic centimeters. They differed from modern humans by having a thick skull, a prominent brow ridge, and a receding chin. Some of these species survived until 30,000–10,000 years ago, overlapping with modern humans (***Figure 7***).

Figure 7. *The Homo neanderthalensis used tools and may have worn clothing.*

There is considerable debate about the origins of anatomically modern humans or ***Homo sapiens sapiens***. As discussed earlier, *H. erectus* migrated out of Africa and into Asia and Europe in the first major wave of migration about 1.5 million years ago. It is thought that modern humans arose in Africa from *H. erectus* and migrated out of Africa about 100,000 years ago in a second major migration wave. Then, modern humans replaced *H. erectus* species that had migrated into Asia and Europe in the first wave.

This evolutionary timeline is supported by molecular evidence. One approach to studying the origins of modern humans is to examine mitochondrial DNA (mtDNA) from populations around the world. Because a fetus develops from an egg containing its mother's mitochondria (which have their own, non-nuclear DNA), mtDNA is passed entirely through the maternal line. Mutations in mtDNA can now be used to estimate the timeline of genetic divergence. The resulting evidence suggests that all modern humans have mtDNA inherited from a common ancestor that lived in Africa about 160,000 years ago. Another approach to the molecular understanding of human evolution is to examine the Y chromosome, which is passed from father to son. This evidence suggests that all men today inherited a Y chromosome from a male that lived in Africa about 140,000 years ago.

Section Summary

All primate species possess adaptations for climbing trees, as they all probably descended from tree-dwellers, although not all species are arboreal. Other characteristics of primates are brains that are larger than those of other mammals, claws that have been modified into flattened nails, typically only one young per pregnancy, stereoscopic vision, and a trend toward holding the body upright. Primates are divided into two groups: prosimians and anthropoids. Monkeys evolved from prosimians during the Oligocene Epoch. Apes evolved from catarrhines in Africa during the Miocene Epoch. Apes are divided into the lesser apes and the greater apes. Hominins include those groups that gave rise to our species, such as *Australopithecus* and *H. erectus*, and those groups that can be considered “cousins” of humans, such as Neanderthals. Fossil evidence shows that hominins at the time of *Australopithecus* were walking upright, the first evidence of bipedal hominins. A number of species, sometimes called archaic *H.sapiens*, evolved from *H. erectus* approximately 500,000 years ago. There is considerable debate about the origins of anatomically modern humans or *H.sapiens sapiens*.

Glossary

anthropoid
: monkeys, apes, and humans

Australopithecus
: genus of hominins that evolved in eastern Africa approximately 4 million years ago

brachiation
: movement through trees branches via suspension from the arms

Catarrhini
: clade of Old World monkeys

Gorilla
: genus of gorillas

hominin
: species that are more closely related to humans than chimpanzees

hominoid
: pertaining to great apes and humans

Homo
: genus of humans

Homo sapiens sapiens
: anatomically modern humans

Hylobatidae
: family of gibbons

Pan
: genus of chimpanzees and bonobos

Platyrrhini
: clade of New World monkeys

Plesiadapis
: oldest known primate-like mammal

Pongo
: genus of orangutans

Primates
: order of lemurs, tarsiers, monkeys, apes, and humans

prognathic jaw
: long jaw

prosimian
: division of primates that includes bush babies of Africa, lemurs of Madagascar, and lorises, pottos, and tarsiers of Southeast Asia

stereoscopic vision
: two overlapping fields of vision from the eyes that produces depth perception

9.

1.9 Cells of the Nervous System

Psychologists striving to understand the human mind may study the nervous system. Learning how the cells and organs (like the brain) function, help us understand the biological basis behind human psychology. The **nervous system** is composed of two basic cell types: glial cells (also known as glia) and neurons. **Glial cells**, which outnumber neurons ten to one, are traditionally thought to play a supportive role to neurons, both physically and metabolically. Glial cells provide scaffolding on which the nervous system is built, help neurons line up closely with each other to allow neuronal communication, provide insulation to neurons, transport nutrients and waste products, and mediate immune responses. **Neurons**, on the other hand, serve as interconnected information processors that are essential for all of the tasks of the nervous system. This section briefly describes the structure and function of neurons.

Neuron Structure

Neurons are the central building blocks of the nervous system, 100 billion strong at birth. Like all cells, neurons consist of several different parts, each serving a specialized function (***Figure 1***). A neuron's outer surface is made up of a **semipermeable membrane**. This membrane allows smaller molecules and molecules without an electrical charge to pass through it, while stopping larger or highly charged molecules.

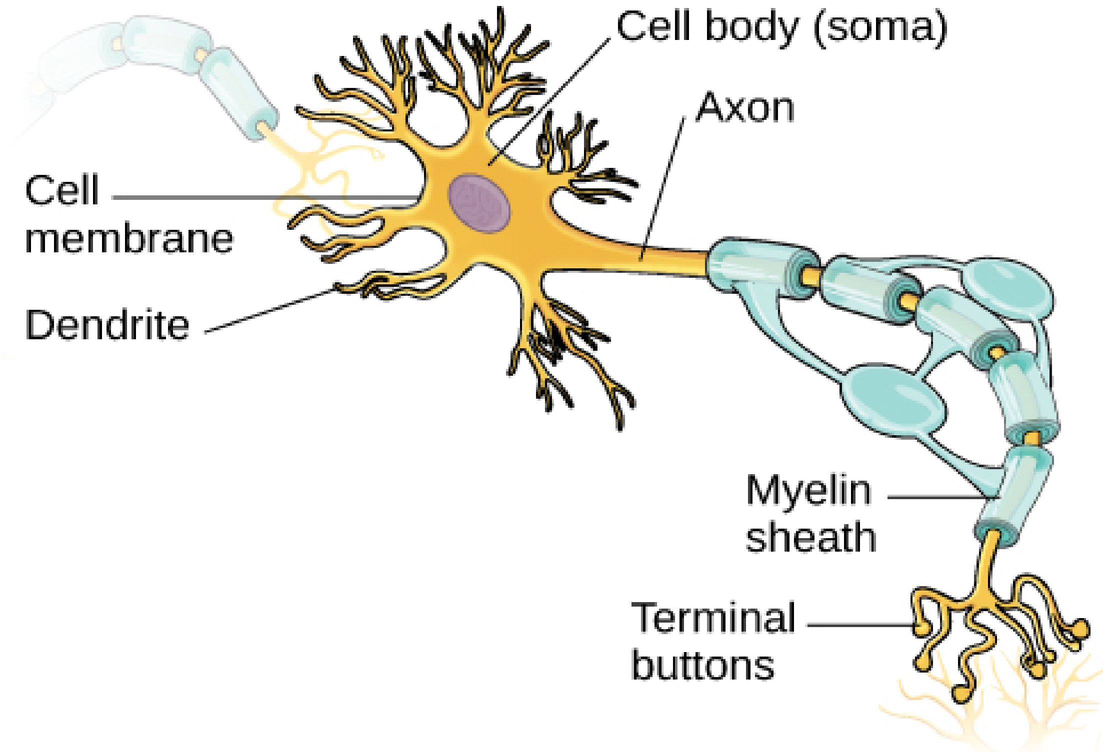


Figure 1. *This illustration shows a prototypical neuron, which is being myelinated.*

The nucleus of the neuron is located in the **soma**, or cell body. The soma has branching extensions known as **dendrites**. The neuron is a small information processor, and dendrites serve as input sites where signals are received from other neurons. These signals are transmitted electrically across the soma and down a major extension from the soma known as the **axon**, which ends at multiple **terminal buttons**. The terminal buttons contain **synaptic vesicles** that house **neurotransmitters**, the chemical messengers of the nervous system.

Axons range in length from a fraction of an inch to several feet. In some axons, glial cells form a fatty substance known as the **myelin sheath**, which coats the axon and acts as an insulator, increasing the speed at which the signal travels. The myelin sheath is crucial for the normal operation of the neurons within the nervous system: the loss of the insulation it provides can be detrimental to normal function. To understand how this works, let's consider an example. Multiple sclerosis (MS), an autoimmune disorder, involves a large-scale loss of the myelin sheath on axons throughout the nervous system. The resulting interference in the electrical signal prevents the quick transmittal of information by neurons and can lead to a number of symptoms, such as dizziness, fatigue, loss of motor control, and sexual dysfunction. While some treatments may help to modify the course of the disease and manage certain symptoms, there is currently no known cure for multiple sclerosis.

In healthy individuals, the neuronal signal moves rapidly down the axon to the terminal buttons, where synaptic vesicles release neurotransmitters into the synapse (***Figure 2***). The **synapse** is a very small space between two neurons and is an important site where communication between neurons occurs. Once neurotransmitters are released into the synapse, they travel across the small space and bind with corresponding receptors on the dendrite of an adjacent neuron. **Receptors**, proteins on the cell surface where neurotransmitters attach, vary in shape, with different shapes "matching" different neurotransmitters.

How does a neurotransmitter "know" which receptor to bind to? The neurotransmitter and the receptor have what is referred to as a lock-and-key relationship—specific neurotransmitters fit specific receptors similar to how a key fits a lock. The neurotransmitter binds to any receptor that it fits.

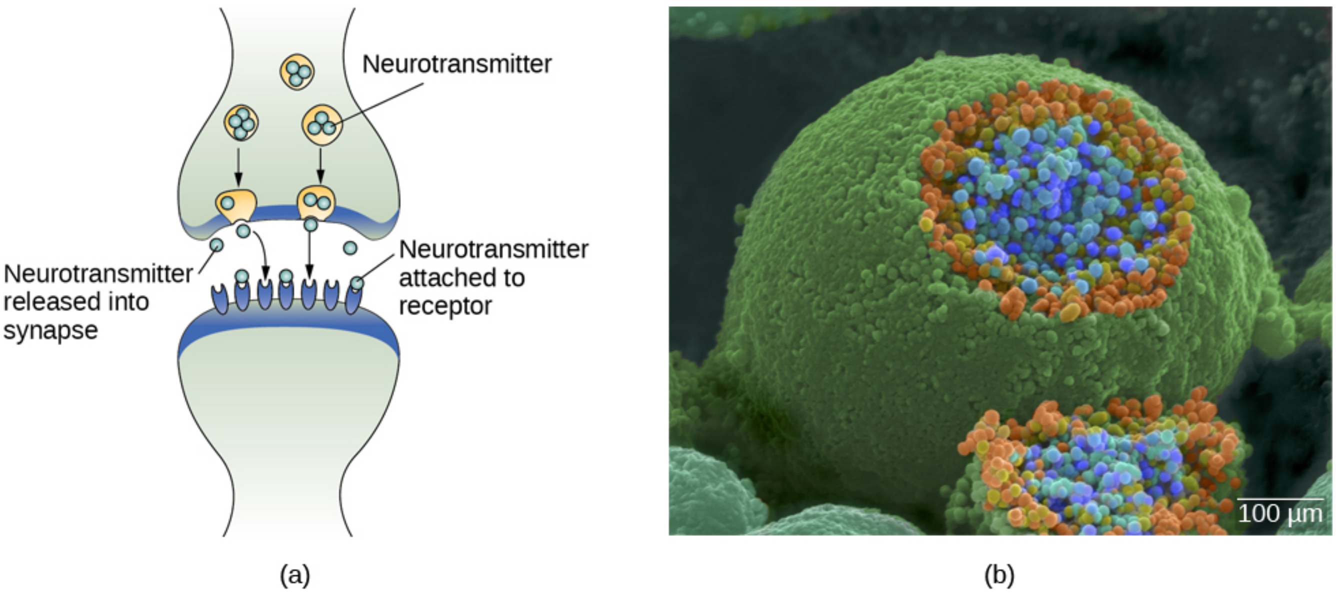


Figure 2. ***(a)*** *The synapse is the space between the terminal button of one neuron and the dendrite of another neuron.* ***(b)*** *In this pseudo-colored image from a scanning electron microscope, a terminal button (green) has been opened to reveal the synaptic vesicles (orange and blue) inside. Each vesicle contains about 10,000 neurotransmitter molecules. (credit b: modification of work by Tina Carvalho, NIH-NIGMS; scale-bar data from Matt Russell)*

Neuronal Communication

Now that we have learned about the basic structures of the neuron and the role that these structures play in neuronal communication, let's take a closer look at the signal itself—how it moves through the neuron and then jumps to the next neuron, where the process is repeated.

We begin at the neuronal membrane. The **neuron** exists in a fluid environment—it is surrounded by extracellular fluid and contains intracellular fluid (i.e., cytoplasm). The neuronal membrane keeps these two fluids separate—a critical role because the electrical signal that passes through the neuron depends on the intra- and extracellular fluids being electrically different. This difference in charge across the membrane, called the **membrane potential**, provides energy for the signal.

The electrical charge of the fluids is caused by charged molecules (ions) dissolved in the fluid. The semipermeable nature of the neuronal membrane somewhat restricts the movement of these charged molecules, and, as a result, some of the charged particles tend to become more concentrated either inside or outside the cell.

Between signals, the neuron membrane's potential is held in a state of readiness, called the **resting potential**. Like a rubber band stretched out and waiting to spring into action, ions line up on either side of the cell membrane, ready to rush across the membrane when the neuron goes active and the membrane opens its gates (i.e., a **sodium-potassium pump** that allows movement of ions across the membrane).

Ions in high-concentration areas are ready to move to low-concentration areas, and positive ions are ready to move to areas with a negative charge.

In the resting state, sodium (Na^+) is at higher concentrations outside the cell, so it will tend to move into the cell. Potassium (K^+), on the other hand, is more concentrated inside the cell, and will tend to move out of the cell (***Figure 3***). In addition, the inside of the cell is slightly negatively charged compared to the outside. This provides an additional force on sodium, causing it to move into the cell.

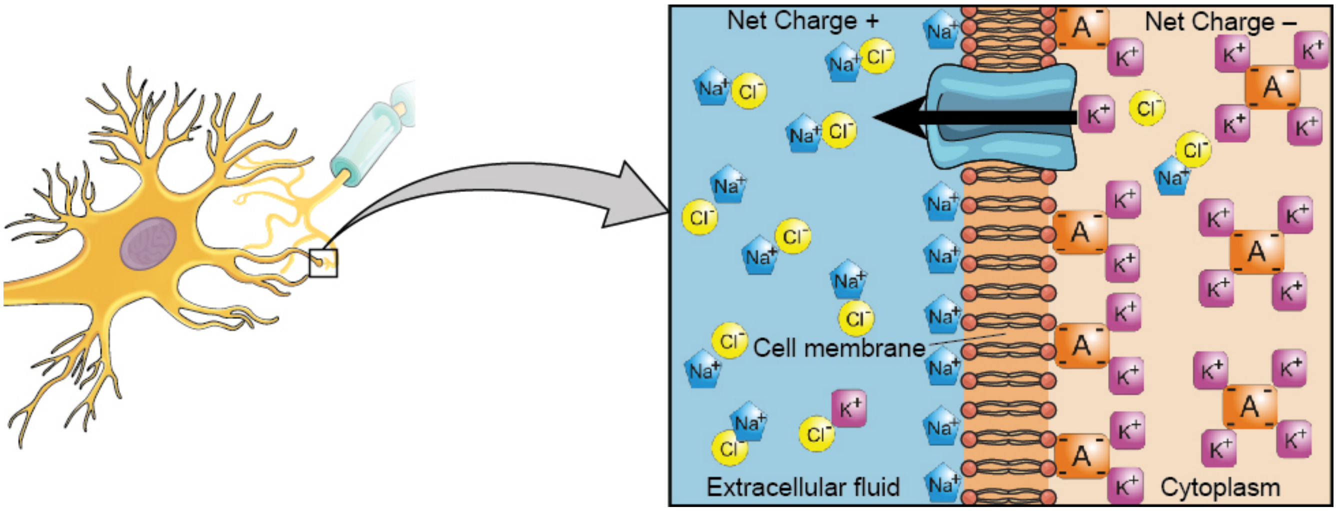


Figure 3. *At resting potential, Na+ (blue pentagons) is more highly concentrated outside the cell in the extracellular fluid (shown in blue), whereas K+ (purple squares) is more highly concentrated near the membrane in the cytoplasm or intracellular fluid. Other molecules, such as chloride ions (yellow circles) and negatively charged proteins (brown squares), help contribute to a positive net charge in the extracellular fluid and a negative net charge in the intracellular fluid.*

From this resting potential state, the neuron receives a signal and its state changes abruptly (***Figure 4***). When a neuron receives signals at the dendrites—due to neurotransmitters from an adjacent neuron binding to its receptors—small pores, or gates, open on the neuronal membrane, allowing Na^+ ions, propelled by both charge and concentration differences, to move into the cell. With this influx of positive ions, the internal charge of the cell becomes more positive. If that charge reaches a certain level, called the **threshold of excitation**, the neuron becomes active and the action potential begins.

Many additional pores open, causing a massive influx of Na^+ ions and a huge positive spike in the membrane potential, the peak action potential. At the peak of the spike, the sodium gates close and the potassium gates open. As positively charged potassium ions leave, the cell quickly begins repolarization. At first, it hyperpolarizes, becoming slightly more negative than the resting potential, and then it levels off, returning to the resting potential.

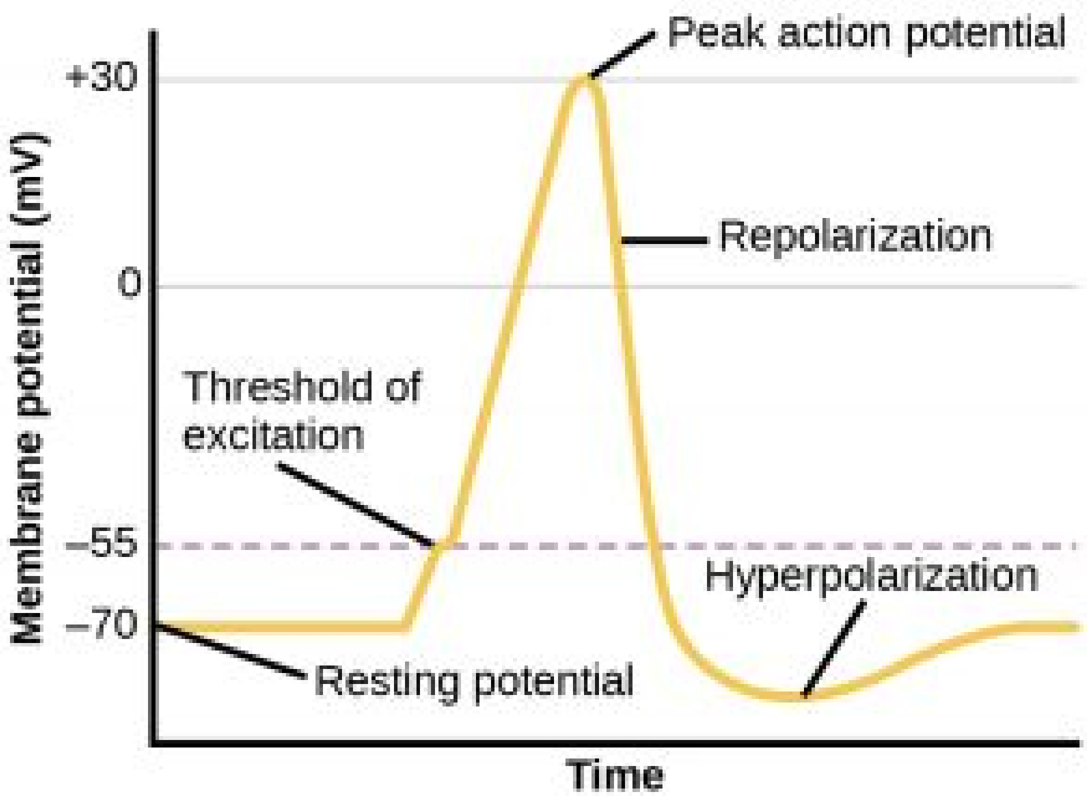


Figure 4. *During the action potential, the electrical charge across the membrane changes dramatically.*

This positive spike constitutes the **action potential**: the electrical signal that typically moves from the cell body down the axon to the axon terminals. The electrical signal moves down the axon like a wave; at each point, some of the sodium ions that enter the cell diffuse to the next section of the axon, raising the charge past the threshold of excitation and triggering a new influx of sodium ions. The action potential moves all the way down the axon to the terminal buttons.

The action potential is an **all-or-none** phenomenon. In simple terms, this means that an incoming signal from another neuron is either sufficient or insufficient to reach the threshold of excitation. There is no in-between, and there is no turning off an action potential once it starts. Think of it like sending an email or a text message. You can think about sending it all you want, but the message is not sent until you hit the send button. Furthermore, once you send the message, there is no stopping it.

Because it is all or none, the **action potential** is recreated, or propagated, at its full strength at every point along the axon. Much like the lit fuse of a firecracker, it does not fade away as it travels down the axon. It is this all-or-none property that explains the fact that your brain perceives an injury to a distant body part like your toe as equally painful as one to your nose.

As noted earlier, when the action potential arrives at the terminal button, the synaptic vesicles release their neurotransmitters into the synapse. The neurotransmitters travel across the synapse and bind to receptors on the dendrites of the adjacent neuron, and the process repeats itself in the new neuron (assuming the signal is sufficiently strong to trigger an action potential). Once the signal is delivered, excess neurotransmitters in the synapse drift away, are broken down into inactive fragments, or are reabsorbed in a process known as **reuptake**. Reuptake involves the neurotransmitter being pumped back into the neuron that released it, in order to clear the synapse (***Figure 5***). Clearing the synapse serves both to provide a clear "on" and "off" state between signals and to regulate the production of neurotransmitter (full synaptic vesicles provide signals that no additional neurotransmitters need to be produced).

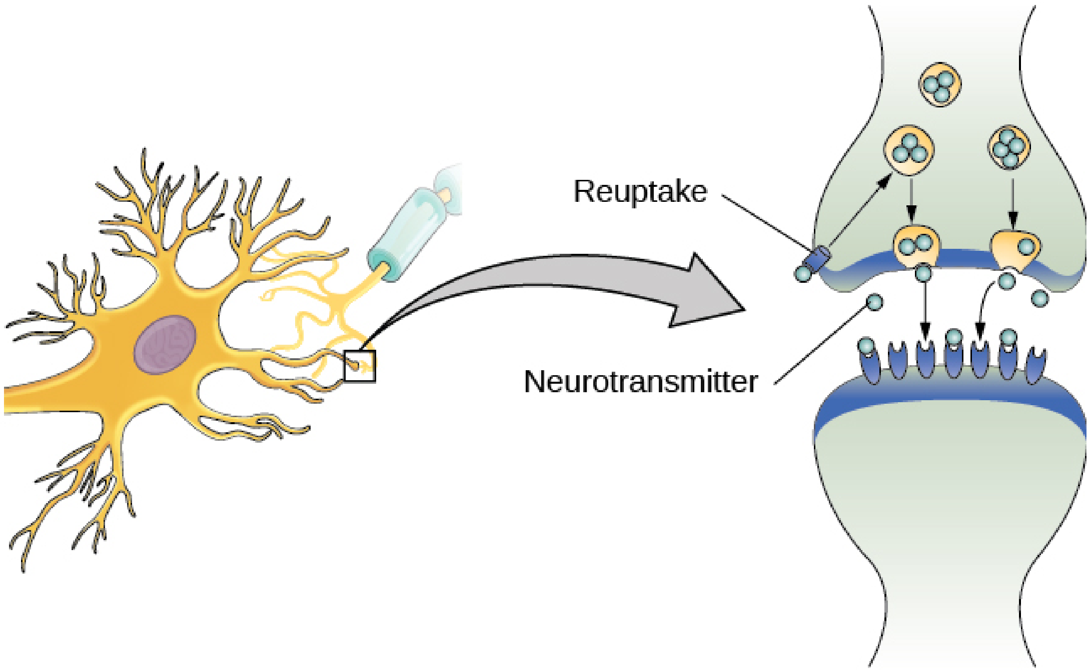


Figure 5. *Reuptake involves moving a neurotransmitter from the synapse back into the axon terminal from which it was released.*

Neuronal communication is often referred to as an electrochemical event. The movement of the action potential down the length of the axon is an electrical event, and movement of the neurotransmitter across the synaptic space represents the chemical portion of the process.

Link to Learning

Click through this interactive simulation for a closer look at neuronal communication.

Neurotransmitters and Drugs

There are several different types of **neurotransmitters** released by different neurons, and we can speak in broad terms about the kinds of functions associated with different neurotransmitters (***Table 1***). Much of what psychologists know about the functions of neurotransmitters comes from research on the effects of drugs in psychological disorders. Psychologists who take a **biological perspective** and focus on the physiological causes of behavior assert that psychological disorders like depression and schizophrenia are associated with imbalances in one or more neurotransmitter systems. In this perspective, psychotropic medications can help improve the symptoms associated with these disorders. **Psychotropic medications** are drugs that treat psychiatric symptoms by restoring neurotransmitter balance.

***Table 1.* Major Neurotransmitters and How They Affect Behavior**

Neurotransmitter	Involved in	Potential Effect on Behavior
Acetylcholine	Muscle action, memory	Increased arousal, enhanced cognition
Beta-endorphin	Pain, pleasure	Decreased anxiety, decreased tension
Dopamine	Mood, motivation	Increased pleasure, suppressed appetite
Norepinephrine	Involuntary muscle, alertness	Increased arousal, suppressed appetite
Serotonin	Mood, sleep	Modulated mood, suppressed appetite

Psychoactive drugs can act as agonists or antagonists for a given neurotransmitter system. **Agonists** are chemicals that mimic a neurotransmitter at the receptor site and, thus, strengthen its effects. An **antagonist**, on the other hand, blocks or impedes the normal activity of a neurotransmitter at the receptor. Agonist and antagonist drugs are prescribed to correct the specific neurotransmitter imbalances underlying a person's condition. For example, Parkinson's disease, a progressive nervous system disorder, is associated with low levels of dopamine. Therefore dopamine agonists, which mimic the effects of dopamine by binding to dopamine receptors, are one treatment strategy.

Certain symptoms of schizophrenia are associated with overactive dopamine neurotransmission. The antipsychotics used to treat these symptoms are antagonists for dopamine—they block dopamine's effects by binding its receptors without activating them. Thus, they prevent dopamine released by one neuron from signaling information to adjacent neurons.

In contrast to agonists and antagonists, which both operate by binding to receptor sites, reuptake inhibitors prevent unused neurotransmitters from being transported back to the neuron. This leaves more neurotransmitters in the synapse for a longer time, increasing its effects. Depression, which has been consistently linked with reduced serotonin levels, is commonly treated with selective serotonin reuptake inhibitors (SSRIs). By preventing reuptake, SSRIs strengthen the effect of serotonin, giving it more time to interact with serotonin receptors on dendrites. Common SSRIs on the market today include Prozac, Paxil, and Zoloft. The drug LSD is structurally very similar to serotonin, and it affects the same neurons and receptors as serotonin. Psychotropic drugs are not instant solutions for people suffering from psychological disorders. Often, an individual must take a drug for several weeks before seeing improvement, and many psychoactive drugs have significant negative side effects. Furthermore, individuals vary dramatically in how they respond to the drugs. To improve chances for success, it is not uncommon for people receiving pharmacotherapy to undergo psychological and/or behavioral therapies as well. Some research suggests that combining drug therapy with other forms of therapy tends to be more effective than any one treatment alone (for one such example, see March et al., 2007).

Section Summary

Glia and neurons are the two cell types that make up the nervous system. While glia generally play supporting roles, the communication between neurons is fundamental to all of the functions associated with the nervous system. Neuronal communication is made possible by the neuron's specialized structures. The soma contains the cell nucleus, and the dendrites extend from the soma in tree-like branches. The axon is another major extension of the cell body; axons are often covered by a myelin sheath, which increases the speed of transmission of neural impulses. At the end of the axon are terminal buttons that contain synaptic vesicles filled with neurotransmitters.

Neuronal communication is an electrochemical event. The dendrites contain receptors for neurotransmitters released by nearby neurons. If the signals received from other neurons are sufficiently strong, an action potential will travel down the length of the axon to the terminal buttons, resulting in the release of neurotransmitters into the synapse. Action potentials operate on the all-or-none principle and involve the movement of Na^+ and K^+across the neuronal membrane.

Different neurotransmitters are associated with different functions. Often, psychological disorders involve imbalances in a given neurotransmitter system. Therefore, psychotropic drugs are prescribed in an attempt to bring the neurotransmitters back into balance. Drugs can act either as agonists or as antagonists for a given neurotransmitter system.

Glossary

action potential
: electrical signal that moves down the neuron's axon

agonist
: drug that mimics or strengthens the effects of a neurotransmitter

all-or-none
: phenomenon that incoming signal from another neuron is either sufficient or insufficient to reach the threshold of excitation

antagonist
: drug that blocks or impedes the normal activity of a given neurotransmitter

axon
: major extension of the soma

biological perspective
: view that psychological disorders like depression and schizophrenia are associated with imbalances in one or more neurotransmitter systems

dendrite
: branch-like extension of the soma that receives incoming signals from other neurons

glial cell
: nervous system cell that provides physical and metabolic support to neurons, including neuronal insulation and communication, and nutrient and waste transport

membrane potential
: difference in charge across the neuronal membrane

myelin sheath
: fatty substance that insulates axons

neuron
: cells in the nervous system that act as interconnected information processors, which are essential for all of the tasks of the nervous system

neurotransmitter
: chemical messenger of the nervous system

psychotropic medication
: drugs that treat psychiatric symptoms by restoring neurotransmitter balance

receptor
: protein on the cell surface where neurotransmitters attach

resting potential
: the state of readiness of a neuron membrane’s potential between signals

reuptake
: neurotransmitter is pumped back into the neuron that released it

semipermeable membrane
: cell membrane that allows smaller molecules or molecules without an electrical charge to pass through it, while stopping larger or highly charged molecules

soma
: cell body

synapse
: small gap between two neurons where communication occurs

synaptic vesicle
: storage site for neurotransmitters

terminal button
: axon terminal containing synaptic vesicles

threshold of excitation
: level of charge in the membrane that causes the neuron to become active

10.

1.10 The Brain and Spinal Cord

The nervous system is divided into two main parts – the central nervous system, made up of the brain and spinal cord, and the peripheral nervous system, which includes all other parts of the nervous system. Here we will focus on the central nervous system, with particular emphasis on the brain and its role in complex thought.

The brain is a remarkably complex organ comprised of billions of interconnected neurons and glia. It is a bilateral, or two-sided, structure that can be separated into distinct lobes. Each lobe is associated with certain types of functions, but, ultimately, all of the areas of the brain interact with one another to provide the foundation for our thoughts and behaviors. In this section, we discuss the overall organization of the brain and the functions associated with different brain areas, beginning with what can be seen as an extension of the brain, the spinal cord.

The Spinal Cord

It can be said that the **spinal cord** is what connects the brain to the outside world. Because of it, the brain can act. The spinal cord is like a relay station, but a very smart one. It not only routes messages to and from the brain, but it also has its own system of automatic processes, called reflexes.

The top of the spinal cord merges with the brain stem, where the basic processes of life are controlled, such as breathing and digestion. In the opposite direction, the spinal cord ends just below the ribs—contrary to what we might expect, it does not extend all the way to the base of the spine.

The Two Hemispheres

The surface of the brain, known as the **cerebral cortex**, is very uneven, characterized by a distinctive pattern of folds or bumps, known as **gyri** (singular: gyrus), and grooves, known as **sulci** (singular: sulcus), shown in ***Figure 1***. These gyri and sulci form important landmarks that allow us to separate the brain into functional centers. The most prominent sulcus, known as the **longitudinal fissure**, is the deep groove that separates the brain into two halves or **hemispheres**: the left hemisphere and the right hemisphere.

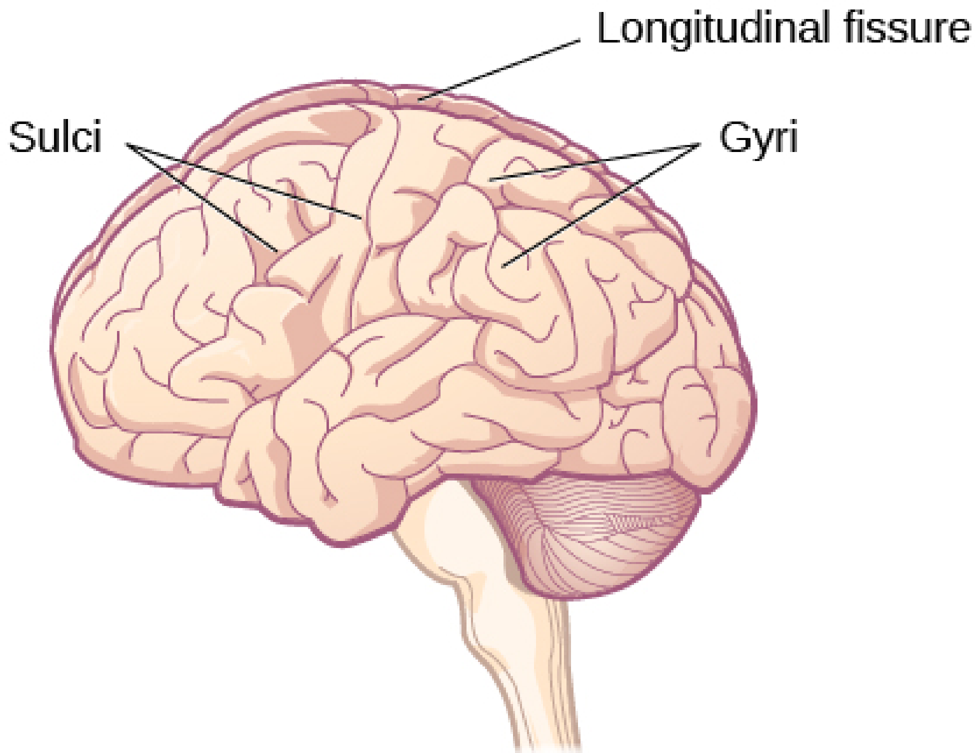


Figure 1. *The surface of the brain is covered with gyri and sulci. A deep sulcus is called a fissure, such as the longitudinal fissure that divides the brain into left and right hemispheres. (credit: modification of work by Bruce Blaus)*

There is evidence of some specialization of function—referred to as **lateralization**—in each hemisphere, mainly regarding differences in language ability. Beyond that, however, the differences that have been found have been minor. What we do know is that the left hemisphere controls the right half of the body, and the right hemisphere controls the left half of the body.

The two hemispheres are connected by a thick band of neural fibers known as the **corpus callosum**, consisting of about 200 million axons. The corpus callosum allows the two hemispheres to communicate with each other and allows for information being processed on one side of the brain to be shared with the other side.

Normally, we are not aware of the different roles that our two hemispheres play in day-to-day functions, but there are people who come to know the capabilities and functions of their two hemispheres quite well. In some cases of severe epilepsy, doctors elect to sever the corpus callosum as a means of controlling the spread of seizures (***Figure 2***). While this is an effective treatment option, it results in individuals who have split brains. After surgery, these split-brain patients show a variety of interesting behaviors. For instance, a split-brain patient is unable to name a picture that is shown in the patient's left visual field because the information is only available in the largely nonverbal right hemisphere. However, they are able to recreate the picture with their left hand, which is also controlled by the right hemisphere. When the more verbal left hemisphere sees the picture that the hand drew, the patient is able to name it (assuming the left hemisphere can interpret what was drawn by the left hand).

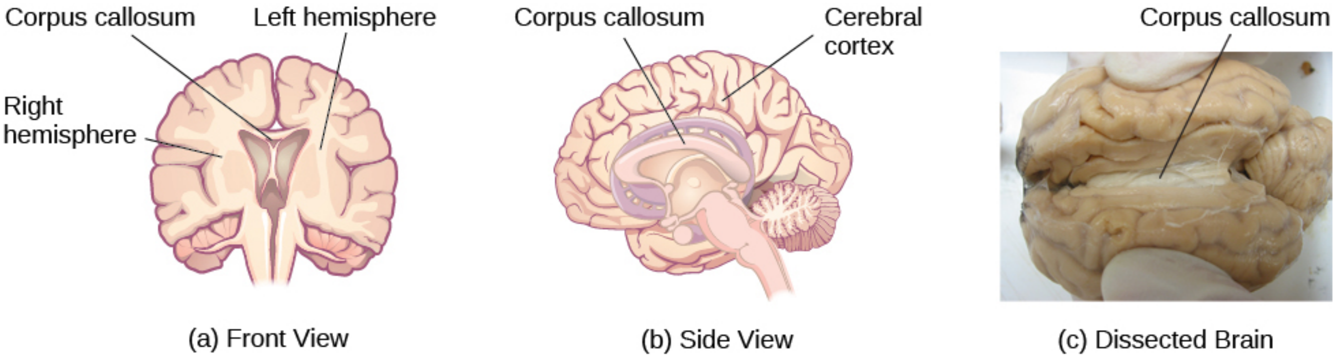


Figure 2. (a, b) *The corpus callosum connects the left and right hemispheres of the brain.* ***(c)*** *A scientist spreads this dissected sheep brain apart to show the corpus callosum between the hemispheres. (credit* ***c****: modification of work by Aaron Bornstein)*

Link to Learning

This interactive animation on the Nobel Prize website walks users through the hemispheres of the brain.

Much of what we know about the functions of different areas of the brain comes from studying changes in the behavior and ability of individuals who have suffered damage to the brain. For example, researchers study the behavioral changes caused by strokes to learn about the functions of specific brain areas. A stroke, caused by an interruption of blood flow to a region in the brain, causes a loss of brain function in the affected region. The damage can be in a small area, and, if it is, this gives researchers the opportunity to link any resulting behavioral changes to a specific area. The types of deficits displayed after a stroke will be largely dependent on where in the brain the damage occurred.

Consider Theona, an intelligent, self-sufficient woman, who is 62 years old. Recently, she suffered a stroke in the front portion of her right hemisphere. As a result, she has great difficulty moving her left leg. (As you learned earlier, the right hemisphere controls the left side of the body; also, the brain's main motor centers are located at the front of the head, in the frontal lobe.) Theona has also experienced behavioral changes. For example, while in the produce section of the grocery store, she sometimes eats grapes, strawberries, and apples directly from their bins before paying for them. This behavior—which would have been very embarrassing to her before the stroke—is consistent with damage in another region in the frontal lobe—the prefrontal cortex, which is associated with judgment, reasoning, and impulse control.

Forebrain Structures

The two hemispheres of the cerebral cortex are part of the **forebrain** (***Figure 3***), which is the largest part of the brain. The forebrain contains the cerebral cortex and a number of other structures that lie beneath the cortex (called subcortical structures): thalamus, hypothalamus, pituitary gland, and the limbic system (collection of structures). The cerebral cortex, which is the outer surface of the brain, is associated with higher level processes such as consciousness, thought, emotion, reasoning, language, and memory. Each cerebral hemisphere can be subdivided into four lobes, each associated with different functions.

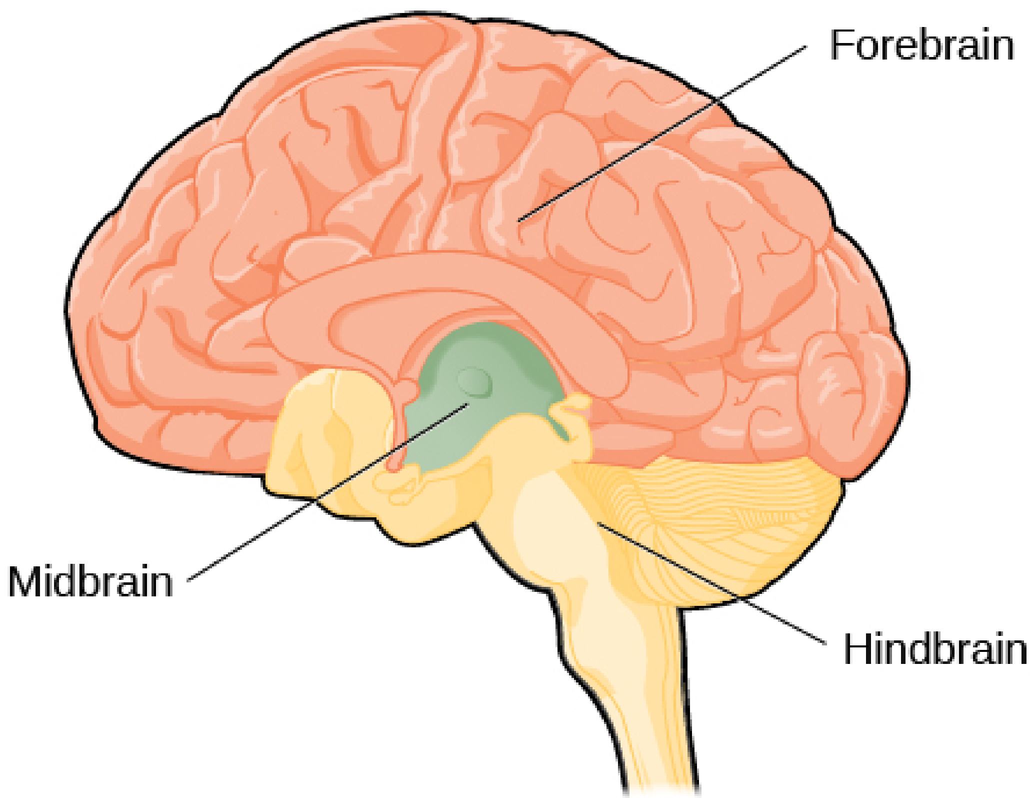


Figure 3. *The brain and its parts can be divided into three main categories: the forebrain, midbrain, and hindbrain.*

Lobes of the Brain

The four lobes of the brain are the frontal, parietal, temporal, and occipital lobes (***Figure 4***). The **frontal lobe** is located in the forward part of the brain, extending back to a fissure known as the central sulcus. The frontal lobe is involved in reasoning, motor control, emotion, and language. It contains the **motor cortex**, which is involved in planning and coordinating movement; the **prefrontal cortex**, which is responsible for higher-level cognitive functioning; and **Broca's area**, which is essential for language production.

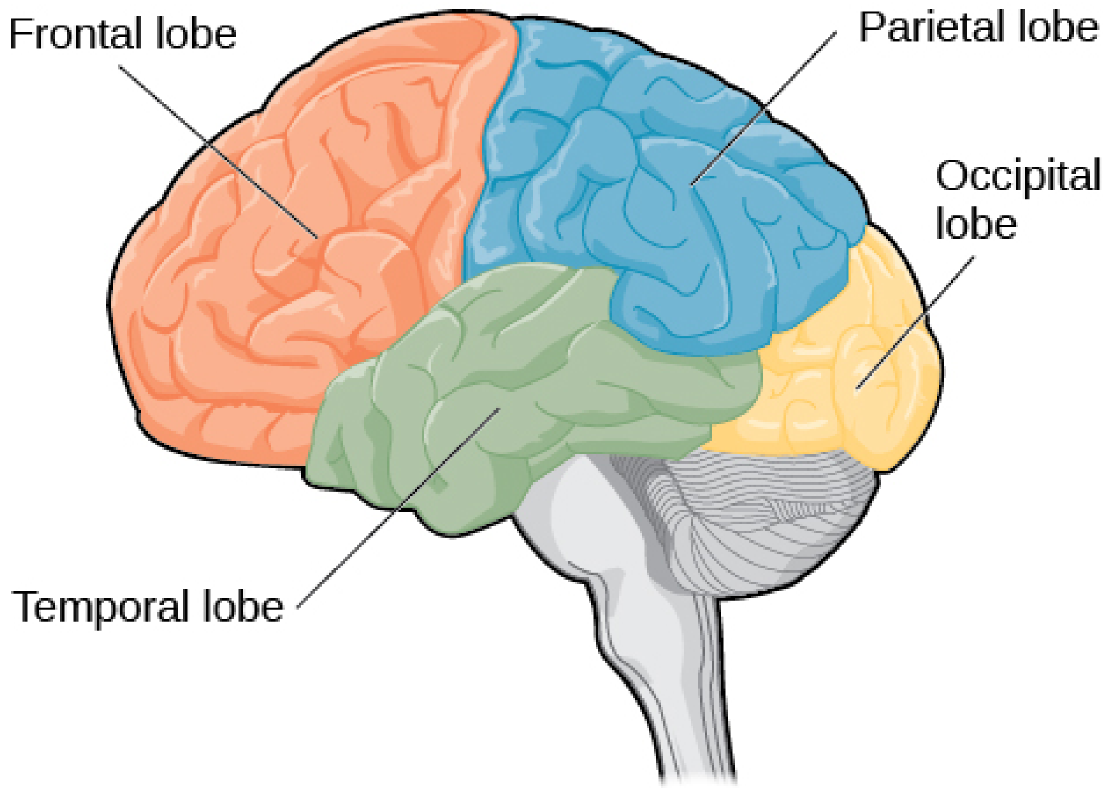


Figure 4. *The lobes of the brain are shown.*

People who suffer damage to Broca's area have great difficulty producing language of any form (***Figure 4***). For example, Padma was an electrical engineer who was socially active and a caring, involved mother. About twenty years ago, she was in a car accident and suffered damage to her Broca's area. She completely lost the ability to speak and form any kind of meaningful language. There is nothing wrong with her mouth or her vocal cords, but she is unable to produce words. She can follow directions but can't respond verbally, and she can read but no longer write. She can do routine tasks like running to the market to buy milk, but she could not communicate verbally if a situation called for it.

Probably the most famous case of frontal lobe damage is that of a man by the name of Phineas **Gage**. On September 13, 1848, Gage (age 25) was working as a railroad foreman in Vermont. He and his crew were using an iron rod to tamp explosives down into a blasting hole to remove rock along the railway's path. Unfortunately, the iron rod created a spark and caused the rod to explode out of the blasting hole, into Gage's face, and through his skull (***Figure 5***). Although lying in a pool of his own blood with brain matter emerging from his head, Gage was conscious and able to get up, walk, and speak. But in the months following his accident, people noticed that his personality had changed. Many of his friends described him as no longer being himself. Before the accident, it was said that Gage was a well-mannered, soft-spoken man, but he began to behave in odd and inappropriate ways after the accident. Such changes in personality would be consistent with loss of impulse control—a frontal lobe function.

Beyond the damage to the frontal lobe itself, subsequent investigations into the rod's path also identified probable damage to pathways between the frontal lobe and other brain structures, including the limbic system. With connections between the planning functions of the frontal lobe and the emotional processes of the limbic system severed, Gage had difficulty controlling his emotional impulses.

However, there is some evidence suggesting that the dramatic changes in Gage's personality were exaggerated and embellished. Gage's case occurred in the midst of a 19th century debate over localization—regarding whether certain areas of the brain are associated with particular functions. On the basis of extremely limited information about Gage, the extent of his injury, and his life before and after the accident, scientists tended to find support for their own views, on whichever side of the debate they fell (Macmillan, 1999).

(a) (b)

Figure 5. ***(a)*** *Phineas Gage holds the iron rod that penetrated his skull in an 1848 railroad construction accident.* ***(b)*** *Gage's prefrontal cortex was severely damaged in the left hemisphere. The rod entered Gage's face on the left side, passed behind his eye, and exited through the top of his skull, before landing about 80 feet away. (credit a: modification of work by Jack and Beverly Wilgus)*

The brain's **parietal lobe** is located immediately behind the frontal lobe, and is involved in processing information from the body's senses. It contains the **somatosensory cortex**, which is essential for processing sensory information from across the body, such as touch, temperature, and pain. The somatosensory cortex is organized topographically, which means that spatial relationships that exist in the body are maintained on the surface of the somatosensory cortex (***Figure 6***). For example, the portion of the cortex that processes sensory information from the hand is adjacent to the portion that processes information from the wrist.

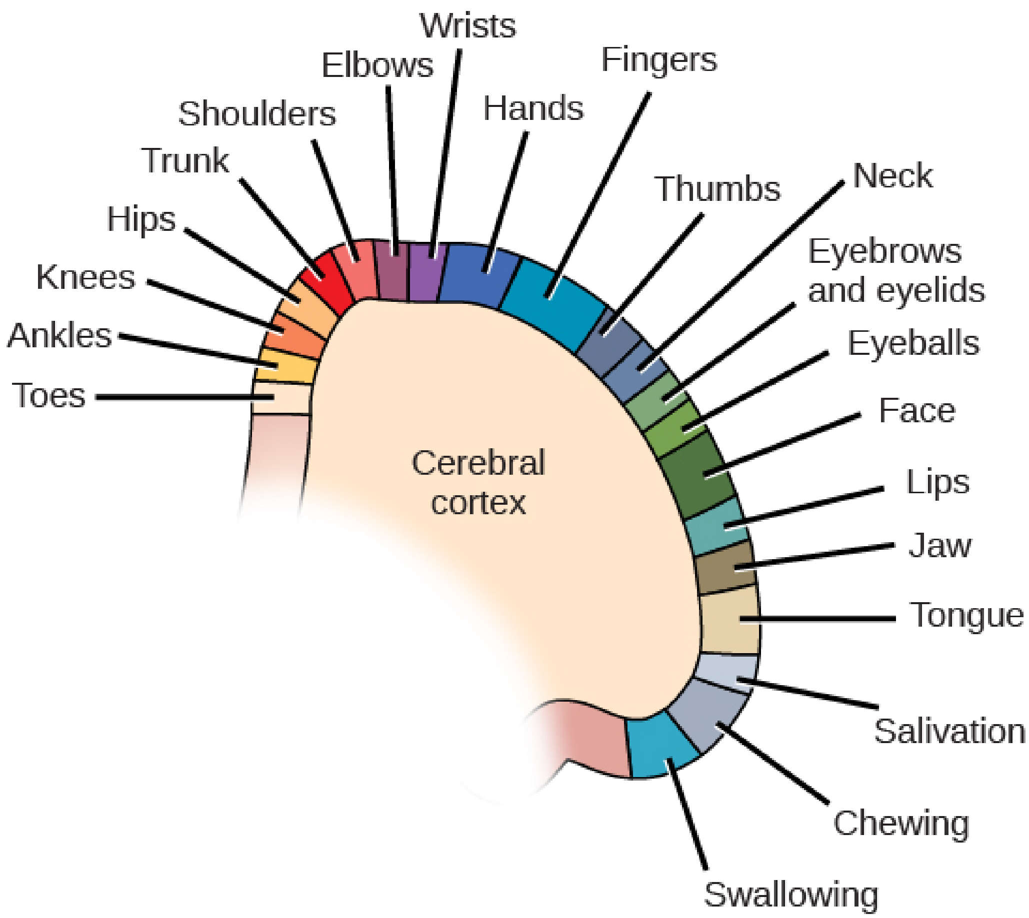


Figure 6. *Spatial relationships in the body are mirrored in the organization of the somatosensory cortex.*

The **temporal lobe** is located on the side of the head (temporal means "near the temples"), and is associated with hearing, memory, emotion, and some aspects of language. The **auditory cortex**, the main area responsible for processing auditory information, is located within the temporal lobe. **Wernicke's area**, important for speech comprehension, is also located here. Whereas individuals with damage to Broca's area have difficulty producing language, those with damage to Wernicke's area can produce sensible language, but they are unable to understand it (***Figure 7***).

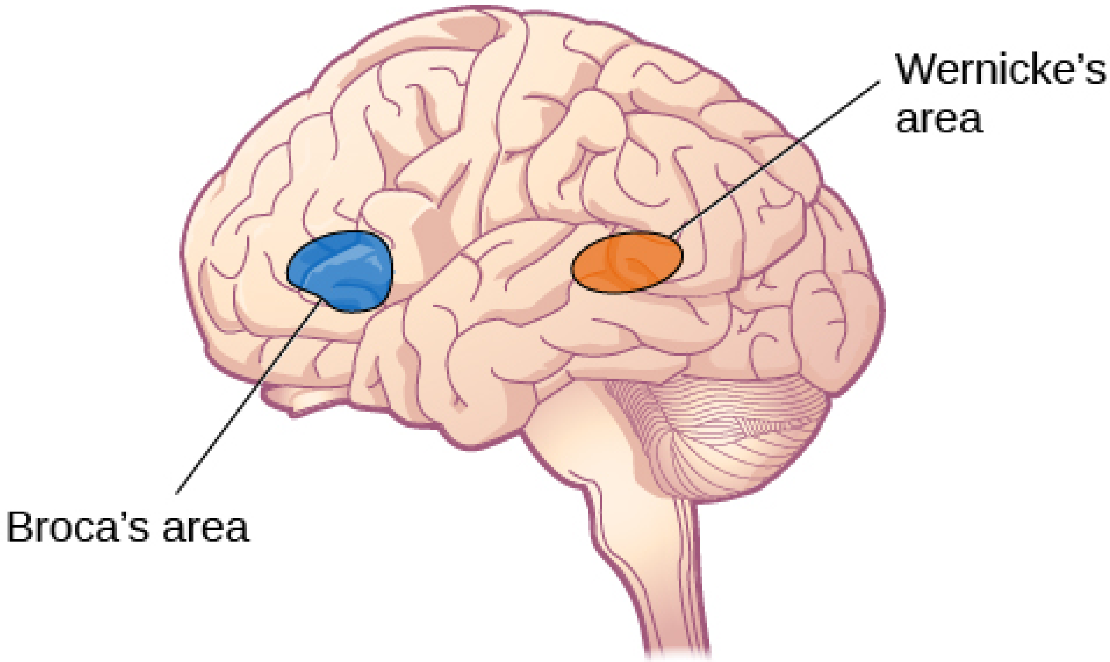


Figure 7. *Damage to either Broca's area or Wernicke's area can result in language deficits. The types of deficits are very different, however, depending on which area is affected.*

The **occipital lobe** is located at the very back of the brain, and contains the primary visual cortex, which is responsible for interpreting incoming visual information. The occipital cortex is organized retinotopically, which means there is a close relationship between the position of an object in a person's visual field and the position of that object's representation on the cortex. You will learn much more about how visual information is processed in the occipital lobe when you study sensation and perception.

Other Areas of the Forebrain

Other areas of the **forebrain**, located beneath the cerebral cortex, include the thalamus and the limbic system. The **thalamus** is a sensory relay for the brain. All of our senses, with the exception of smell, are routed through the thalamus before being directed to other areas of the brain for processing (***Figure 8***).

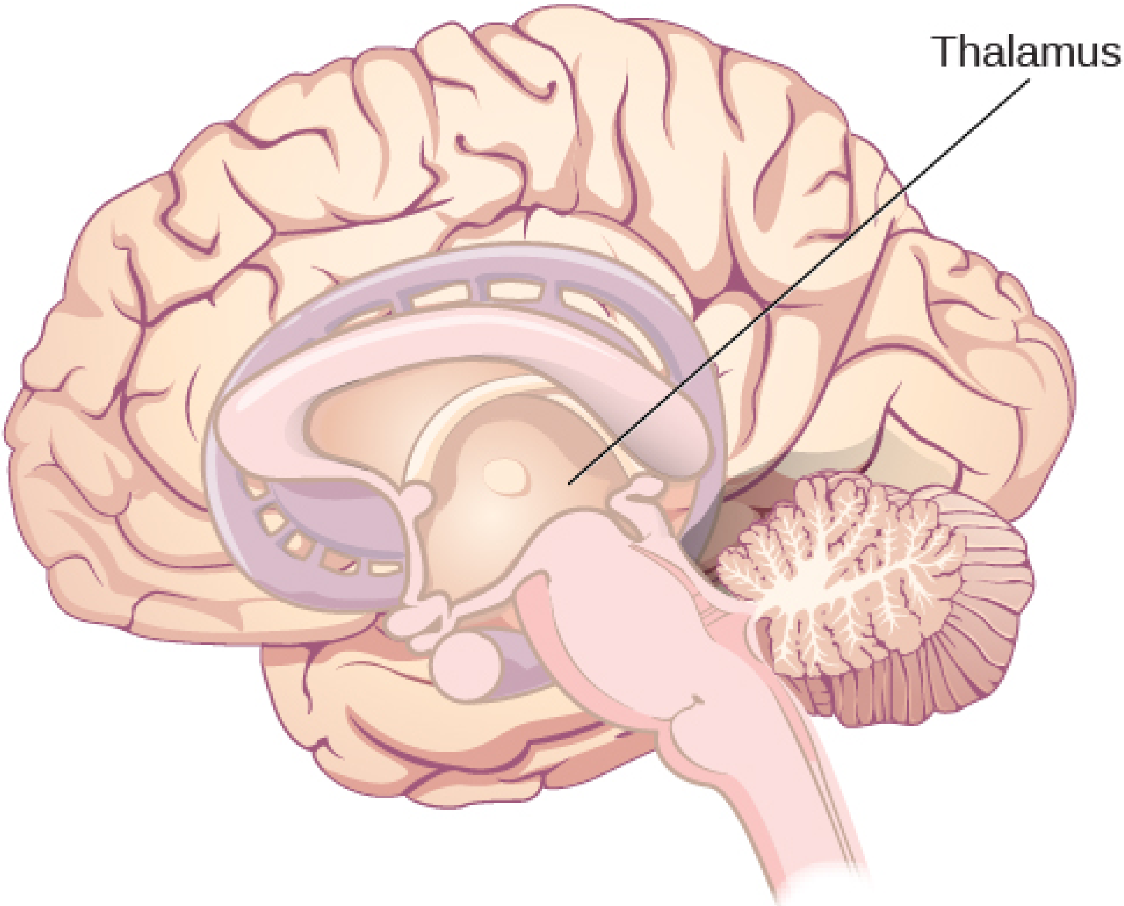


Figure 8. The thalamus serves as the relay center of the brain where most senses are routed for processing.

The **limbic system** is involved in processing both emotion and memory. Interestingly, the sense of smell projects directly to the limbic system; therefore, not surprisingly, smell can evoke emotional responses in ways that other sensory modalities cannot. The limbic system is made up of a number of different structures, but three of the most important are the hippocampus, the amygdala, and the hypothalamus (***Figure 9***). The **hippocampus** is an essential structure for learning and memory. The **amygdala** is involved in our experience of emotion and in tying emotional meaning to our memories. The **hypothalamus** regulates a number of homeostatic processes, including the regulation of body temperature, appetite, and blood pressure. The hypothalamus also serves as an interface between the nervous system and the endocrine system and in the regulation of sexual motivation and behavior.

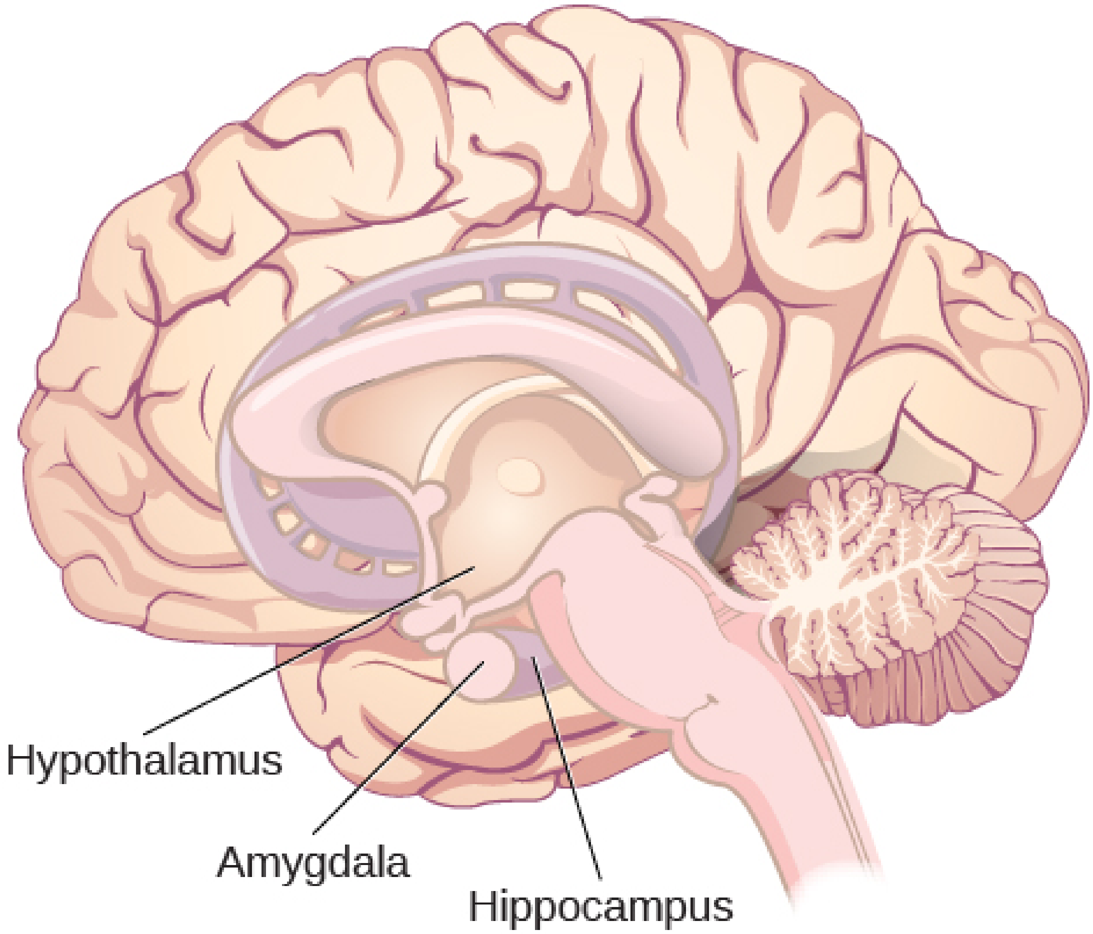


Figure 9. *The limbic system is involved in mediating emotional response and memory.*

The Case of Henry Molaison (H.M.)

In 1953, Henry Gustav **Molaison** (H. M.) was a 27-year-old man who experienced severe seizures. In an attempt to control his seizures, H. M. underwent brain surgery to remove his hippocampus and amygdala. Following the surgery, H.M's seizures became much less severe, but he also suffered some unexpected—and devastating—consequences of the surgery: he lost his ability to form many types of new memories. For example, he was unable to learn new facts, such as who was president of the United States. He was able to learn new skills, but afterward he had no recollection of learning them. For example, while he might learn to use a computer, he would have no conscious memory of ever having used one. He could not remember new faces, and he was unable to remember events, even immediately after they occurred. Researchers were fascinated by his experience, and he is considered one of the most studied cases in medical and psychological history (Hardt, Einarsson, & Nader, 2010; Squire, 2009). Indeed, his case has provided tremendous insight into the role that the hippocampus plays in the consolidation of new learning into explicit memory.

Link to Learning

Clive Wearing, an accomplished musician, lost the ability to form new memories when his hippocampus was damaged through illness. Check out the first few minutes of this documentary video for an introduction to this man and his condition.

Midbrain and Hindbrain Structures

The **midbrain** is comprised of structures located deep within the brain, between the forebrain and the hindbrain. The **reticular formation** is centered in the midbrain, but it actually extends up into the forebrain and down into the hindbrain. The reticular formation is important in regulating the sleep/wake cycle, arousal, alertness, and motor activity.

The **substantia nigra** (Latin for “black substance”) and the **ventral tegmental area (VTA)** are also located in the midbrain (***Figure 10***). Both regions contain cell bodies that produce the neurotransmitter dopamine, and both are critical for movement. Degeneration of the substantia nigra and VTA is involved in Parkinson’s disease. In addition, these structures are involved in mood, reward, and addiction (Berridge & Robinson, 1998; Gardner, 2011; George, Le Moal, & Koob, 2012).

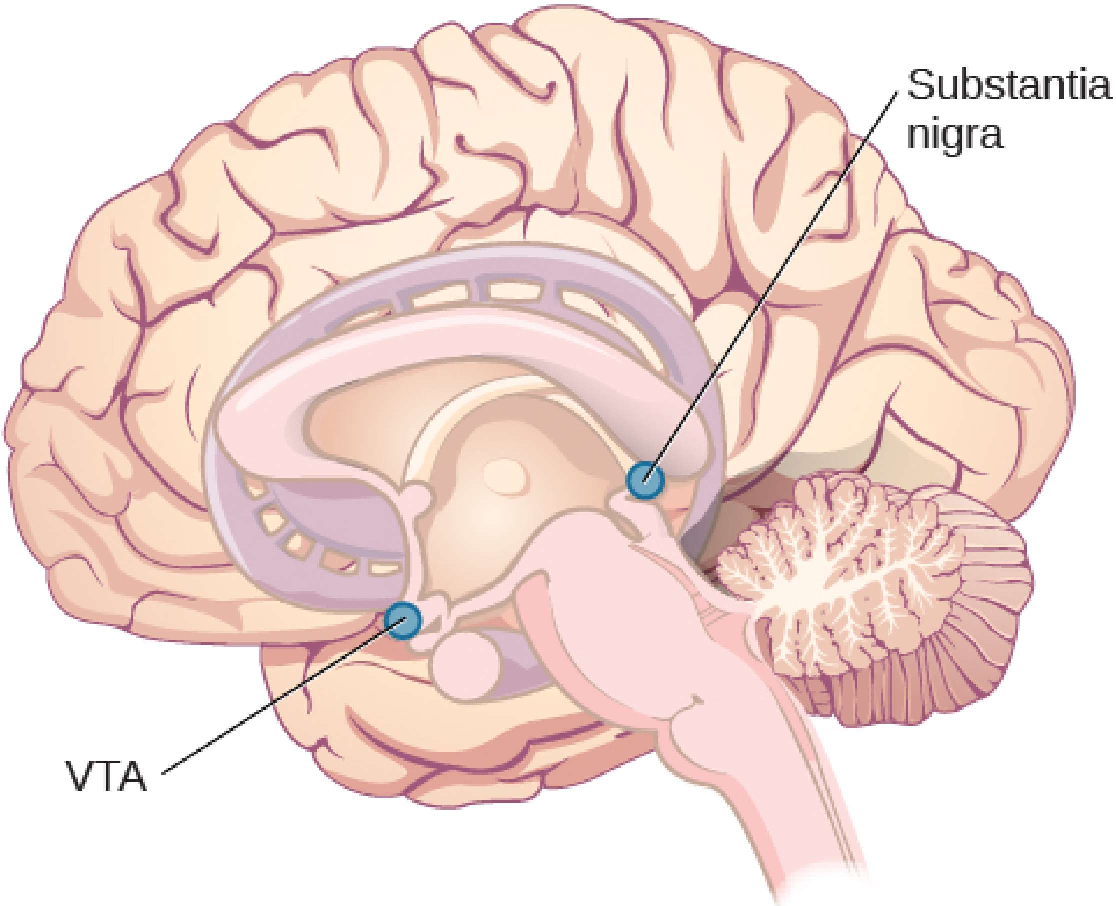


Figure 10. *The substantia nigra and ventral tegmental area (VTA) are located in the midbrain.*

The **hindbrain** is located at the back of the head and looks like an extension of the spinal cord. It contains the medulla, pons, and cerebellum (***Figure 11***). The **medulla** controls the automatic processes of the autonomic nervous system, such as breathing, blood pressure, and heart rate. The word **pons** literally means "bridge," and as the name suggests, the pons serves to connect the brain and spinal cord. It also is involved in regulating brain activity during sleep. The medulla, pons, and midbrain together are known as the brainstem.

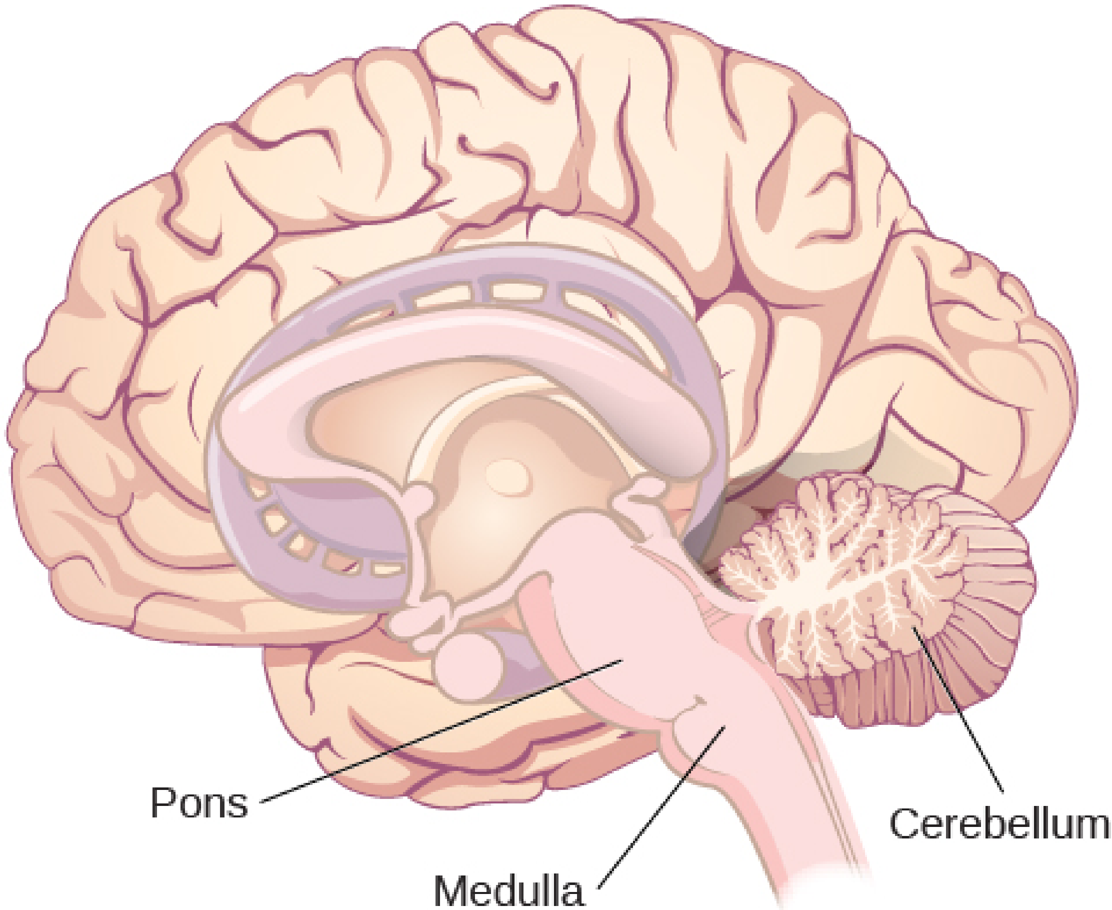


Figure 11. *The pons, medulla, and cerebellum make up the hindbrain.*

The **cerebellum** (Latin for "little brain") receives messages from muscles, tendons, joints, and structures in our ear to control balance, coordination, movement, and motor skills. The cerebellum is also thought to be an important area for processing some types of memories. In particular, procedural memory, or memory involved in learning and remembering how to perform tasks, is thought to be associated with the cerebellum. Recall that H. M. was unable to form new explicit memories, but he could learn new tasks. This is likely due to the fact that H. M.'s cerebellum remained intact.

BRAIN DEAD AND ON LIFE SUPPORT

What would you do if your spouse or loved one was declared brain dead but his or her body was being kept alive by medical equipment? Whose decision should it be to remove a feeding tube? Should medical care costs be a factor?

On February 25, 1990, a Florida woman named Terri **Schiavo** went into cardiac arrest, apparently triggered by a bulimic episode. She was eventually revived, but her brain had been deprived of oxygen for a long time. Brain scans indicated that there was no activity in her cerebral cortex, and she suffered from severe and permanent cerebral atrophy. Basically, Schiavo was in a vegetative state. Medical professionals determined that she would never again be able to move, talk, or respond in any way. To remain alive, she required a feeding tube, and there was no chance that her situation would ever improve.

On occasion, Schiavo's eyes would move, and sometimes she would groan. Despite the doctors' insistence to the contrary, her parents believed that these were signs that she was trying to communicate with them.

After 12 years, Schiavo's husband argued that his wife would not have wanted to be kept alive with no feelings, sensations, or brain activity. Her parents, however, were very much against removing her feeding tube. Eventually, the case made its way to the courts, both in the state of Florida and at the federal level. By 2005, the courts found in favor of Schiavo's husband, and the feeding tube was removed on March 18, 2005. Schiavo died 13 days later.

Why did Schiavo's eyes sometimes move, and why did she groan? Although the parts of her brain that control thought, voluntary movement, and feeling were completely damaged, her brainstem was still intact. Her medulla and pons maintained her breathing and caused involuntary movements of her eyes and the occasional groans. Over the 15-year period that she was on a feeding tube, Schiavo's medical costs may have topped $7 million (Arnst, 2003).

These questions were brought to popular conscience 25 years ago in the case of Terri Schiavo, and they persist today. In 2013, a 13-year-old girl who suffered complications after tonsil surgery was declared brain dead. There was a battle between her family, who wanted her to remain on life support, and the hospital's policies regarding persons declared brain dead. In another complicated 2013–14 case in Texas, a pregnant EMT professional declared brain dead was kept alive for weeks, despite her spouse's directives, which were based on her wishes should this situation arise. In this case, state laws designed to protect an unborn fetus came into consideration until doctors determined the fetus unviable.

Decisions surrounding the medical response to patients declared brain dead are complex. What do you think about these issues?

Brain Imaging

You have learned how brain injury can provide information about the functions of different parts of the brain. Increasingly, however, we are able to obtain that information using **brain imaging** techniques on individuals who have not suffered brain injury. In this section, we take a more in-depth look at some of the techniques that are available for imaging the brain, including techniques that rely on radiation, magnetic fields, or electrical activity within the brain.

Techniques Involving Radiation

A **computerized tomography (CT) scan**involves taking a number of x-rays of a particular section of a person's body or brain (***Figure 12***). The x-rays pass through tissues of different densities at different rates, allowing a computer to construct an overall image of the area of the body being scanned. A CT scan is often used to determine whether someone has a tumor, or significant brain atrophy.

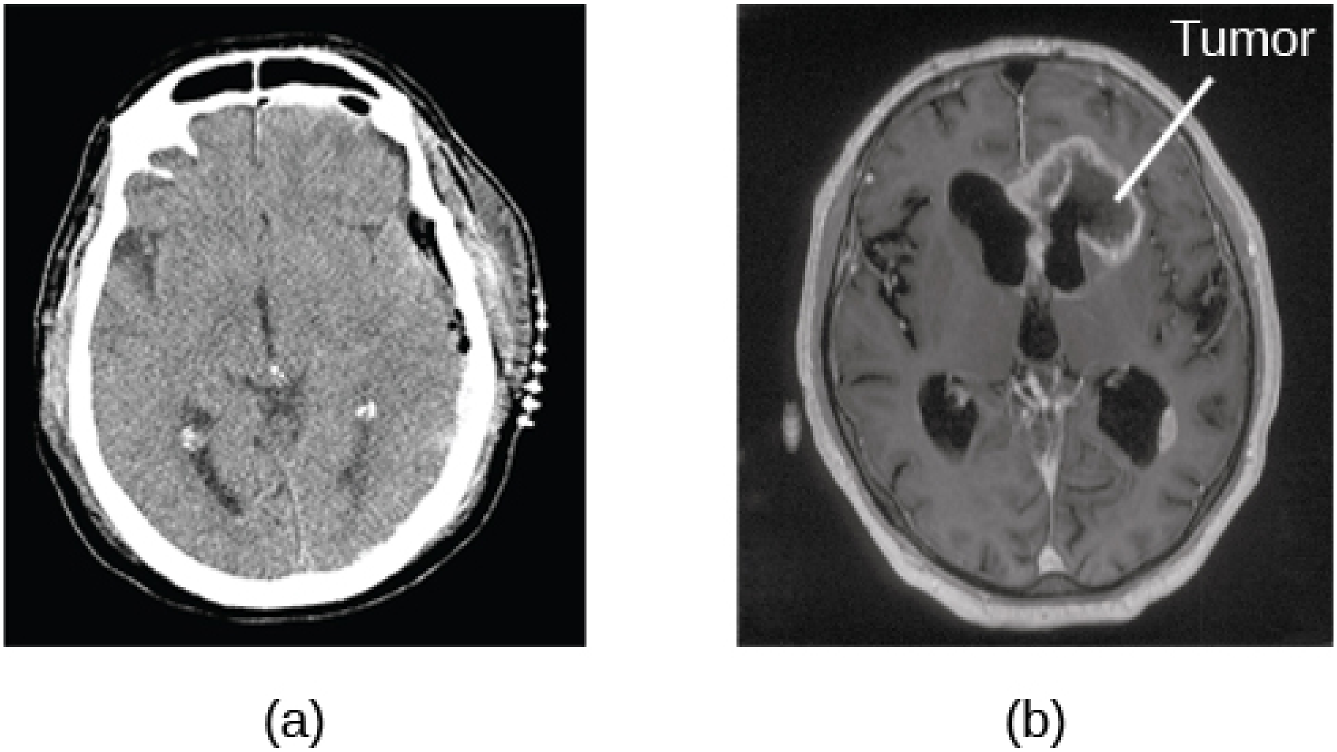


Figure 12. *A CT scan can be used to show brain tumors. (a) The image on the left shows a healthy brain, whereas (b) the image on the right indicates a brain tumor in the left frontal lobe. (credit a: modification of work by "Aceofhearts1968"/Wikimedia Commons; credit b: modification of work by Roland Schmitt et al)*

Positron emission tomography (PET) scans create pictures of the living, active brain (***Figure 13***). An individual receiving a PET scan drinks or is injected with a mildly radioactive substance, called a tracer. Once in the bloodstream, the amount of tracer in any given region of the brain can be monitored. As brain areas become more active, more blood flows to that area. A computer monitors the movement of the tracer and creates a rough map of active and inactive areas of the brain during a given behavior. PET scans show little detail, are unable to pinpoint events precisely in time, and require that the brain be exposed to radiation; therefore, this technique has been replaced by the fMRI as an alternative diagnostic tool. However, combined with CT, PET technology is still being used in certain contexts. For example, CT/PET scans allow better imaging of the activity of neurotransmitter receptors and open new avenues in schizophrenia research. In this hybrid CT/PET technology, CT contributes clear images of brain structures, while PET shows the brain's activity.

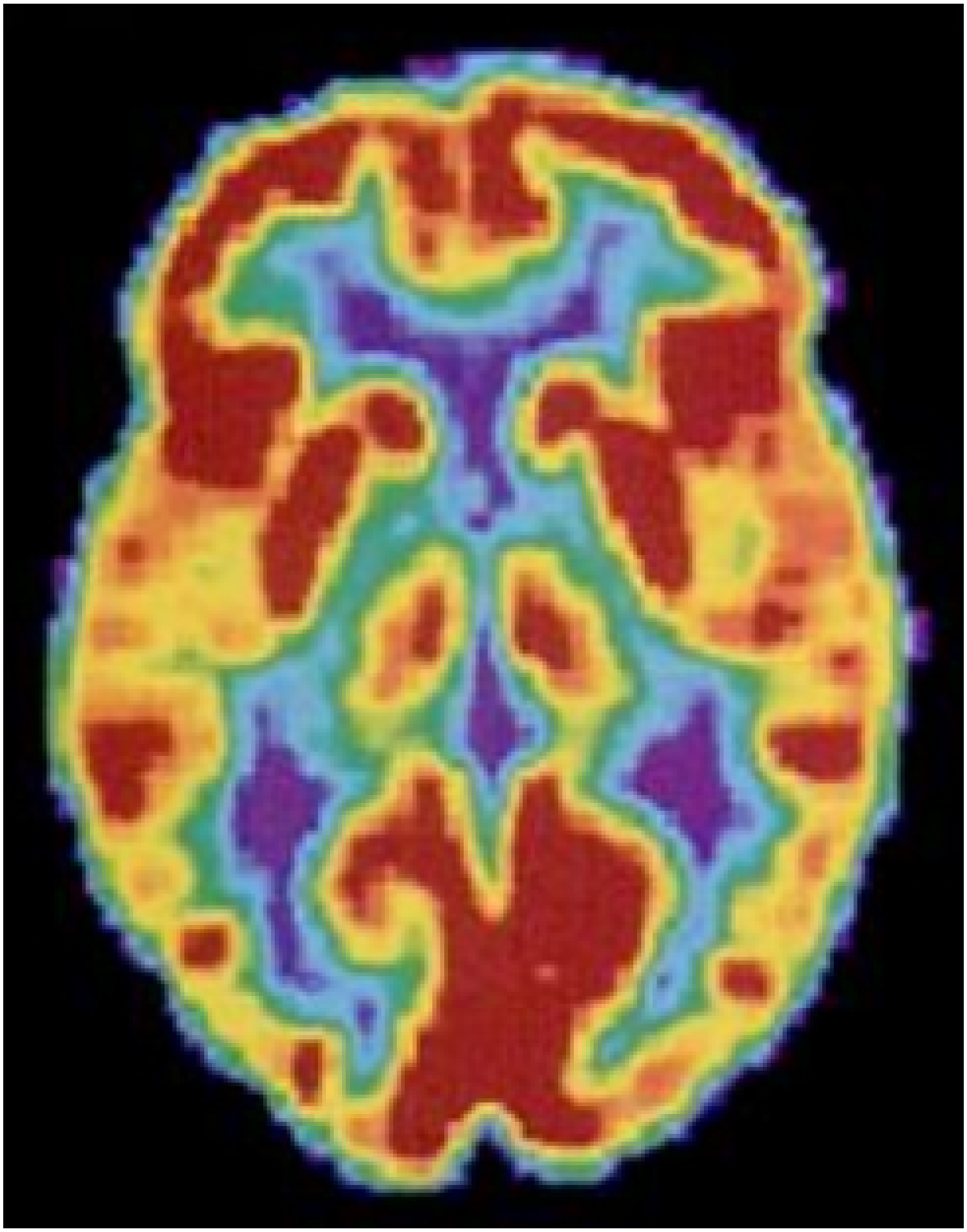


Figure 13. *A PET scan is helpful for showing activity in different parts of the brain. (credit: Health and Human Services Department, National Institutes of Health)*

Techniques Involving Magnetic Fields

In **magnetic resonance imaging (MRI)**, a person is placed inside a machine that generates a strong magnetic field. The magnetic field causes the hydrogen atoms in the body's cells to move. When the magnetic field is turned off, the hydrogen atoms emit electromagnetic signals as they return to their original positions. Tissues of different densities give off different signals, which a computer interprets and displays on a monitor. **Functional magnetic resonance imaging (fMRI)** operates on the same principles, but it shows changes in brain activity over time by tracking blood flow and oxygen levels. The fMRI provides more detailed images of the brain's structure, as well as better accuracy in time, than is possible in PET scans (***Figure 14***). With their high level of detail, MRI and fMRI are often used to compare the brains of healthy individuals to the brains of individuals diagnosed with psychological disorders. This comparison helps determine what structural and functional differences exist between these populations.

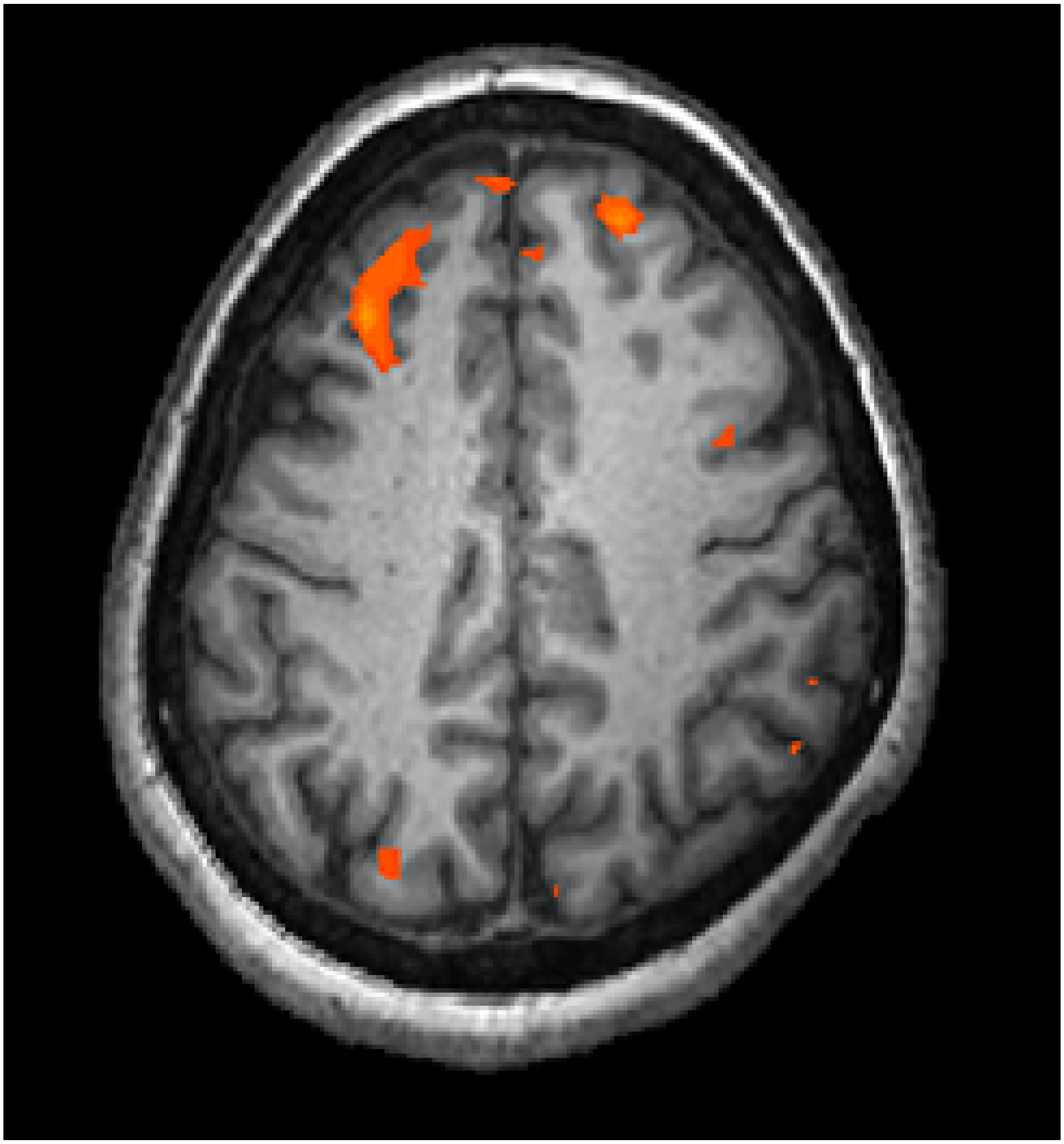


Figure 14. *An fMRI shows activity in the brain over time. This image represents a single frame from an fMRI. (credit: modification of work by Kim J, Matthews NL, Park S.)*

Link to Learning

Visit this virtual lab to learn more about MRI and fMRI.

Techniques Involving Electrical Activity

In some situations, it is helpful to gain an understanding of the overall activity of a person's brain, without needing information on the actual location of the activity. **Electroencephalography (EEG)** serves this purpose by providing a measure of a brain's electrical activity. An array of electrodes is placed around a person's head (***Figure 15***). The signals received by the electrodes result in a printout of the electrical activity of his or her brain, or brainwaves, showing both the frequency (number of waves per second) and amplitude (height) of the recorded brainwaves, with an accuracy within milliseconds. Such information is especially helpful to researchers studying sleep patterns among individuals with sleep disorders.

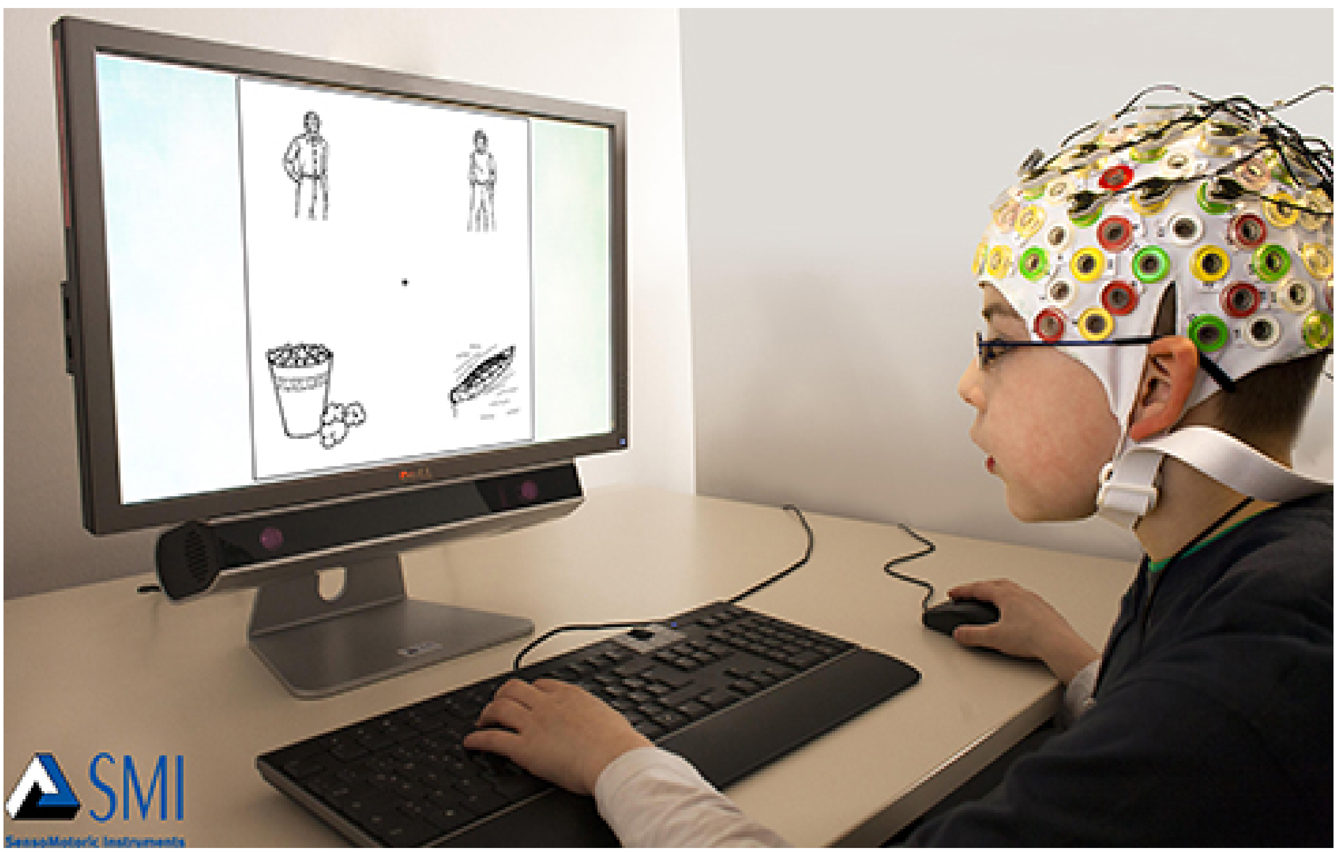


Figure 15. *Using caps with electrodes, modern EEG research can study the precise timing of overall brain activities. (credit: SMI Eye Tracking)*

Section Summary

The brain consists of two hemispheres, each controlling the opposite side of the body. Each hemisphere can be subdivided into different lobes: frontal, parietal, temporal, and occipital. In addition to the lobes of the cerebral cortex, the forebrain includes the thalamus (sensory relay) and limbic system (emotion and memory circuit). The midbrain contains the reticular formation, which is important for sleep and arousal, as well as the substantia nigra and ventral tegmental area. These structures are important for movement, reward, and addictive processes. The hindbrain contains the structures of the brainstem (medulla, pons, and midbrain), which control automatic functions like breathing and blood pressure. The hindbrain also contains the cerebellum, which helps coordinate movement and certain types of memories.

Individuals with brain damage have been studied extensively to provide information about the role of different areas of the brain, and recent advances in technology allow us to glean similar information by imaging brain structure and function. These techniques include CT, PET, MRI, fMRI, and EEG.

Glossary

amygdala
: structure in the limbic system involved in our experience of emotion and tying emotional meaning to our memories

auditory cortex
: strip of cortex in the temporal lobe that is responsible for processing auditory information

Broca's area
: region in the left hemisphere that is essential for language production

cerebellum
: hindbrain structure that controls our balance, coordination, movement, and motor skills, and it is thought to be important in processing some types of memory

cerebral cortex
: surface of the brain that is associated with our highest mental capabilities

computerized tomography (CT) scan
: imaging technique in which a computer coordinates and integrates multiple x-rays of a given area

corpus callosum
: thick band of neural fibers connecting the brain's two hemispheres

electroencephalography (EEG)
: recording the electrical activity of the brain via electrodes on the scalp

forebrain
: largest part of the brain, containing the cerebral cortex, the thalamus, and the limbic system, among other structures

frontal lobe
: part of the cerebral cortex involved in reasoning, motor control, emotion, and language; contains motor cortex

functional magnetic resonance imaging (fMRI)
MRI that shows changes in metabolic activity over time

gyrus
(plural: gyri) bump or ridge on the cerebral cortex

hemisphere
left or right half of the brain

hindbrain
division of the brain containing the medulla, pons, and cerebellum

hippocampus
structure in the temporal lobe associated with learning and memory

hypothalamus
forebrain structure that regulates sexual motivation and behavior and a number of homeostatic processes; serves as an interface between the nervous system and the endocrine system

lateralization
concept that each hemisphere of the brain is associated with specialized functions

limbic system
collection of structures involved in processing emotion and memory

longitudinal fissure
deep groove in the brain's cortex

magnetic resonance imaging (MRI)
magnetic fields used to produce a picture of the tissue being imaged

medulla
hindbrain structure that controls automated processes like breathing, blood pressure, and heart rate

midbrain
division of the brain located between the forebrain and the hindbrain; contains the reticular formation

motor cortex
strip of cortex involved in planning and coordinating movement

occipital lobe
part of the cerebral cortex associated with visual processing; contains the primary visual cortex

parietal lobe
part of the cerebral cortex involved in processing various sensory and perceptual information; contains the primary somatosensory cortex

pons
hindbrain structure that connects the brain and spinal cord; involved in regulating brain activity during sleep

positron emission tomography (PET) scan
involves injecting individuals with a mildly radioactive substance and monitoring changes in blood flow to different regions of the brain

prefrontal cortex
area in the frontal lobe responsible for higher-level cognitive functioning

reticular formation
midbrain structure important in regulating the sleep/wake cycle, arousal, alertness, and motor activity

somatosensory cortex
essential for processing sensory information from across the body, such as touch, temperature, and pain

substantia nigra
midbrain structure where dopamine is produced; involved in control of movement

sulcus
(plural: sulci) depressions or grooves in the cerebral cortex

temporal lobe
part of cerebral cortex associated with hearing, memory, emotion, and some aspects of language; contains primary auditory cortex

thalamus
sensory relay for the brain

ventral tegmental area (VTA)
midbrain structure where dopamine is produced: associated with mood, reward, and addiction

Wernicke's area
important for speech comprehension

11.

1.11 How Memory Functions

Memory is an information processing system; therefore, we often compare it to a computer. **Memory** is the set of processes used to encode, store, and retrieve information over different periods of time (***Figure 1***).

Figure 1. *Encoding involves the input of information into the memory system. Storage is the retention of the encoded information. Retrieval, or getting the information out of memory and back into awareness, is the third function.*

Encoding

We get information into our brains through a process called **encoding**, which is the input of information into the memory system. Once we receive sensory information from the environment, our brains label or code it. We organize the information with other similar information and connect new concepts to existing concepts.

What are the most effective ways to ensure that important memories are well encoded? Even a simple sentence is easier to recall when it is meaningful (Anderson, 1984). Read the following sentences (Bransford & McCarrell, 1974), then look away and count backwards from 30 by threes to zero, and then try to write down the sentences (no peeking back at this page!).

1. The notes were sour because the seams split.
2. The voyage wasn't delayed because the bottle shattered.
3. The haystack was important because the cloth ripped.

How well did you do? By themselves, the statements that you wrote down were most likely confusing and difficult for you to recall. Now, try writing them again, using the following prompts: bagpipe, ship christening, and parachutist. Next count backwards from 40 by fours, then check yourself to see how well you recalled the sentences this time. You can see that the sentences are now much more memorable because each of the sentences was placed in context. Material is far better encoded when you make it meaningful.

Words that had been encoded semantically were better remembered than those encoded visually or acoustically. Semantic encoding involves a deeper level of processing than the shallower visual or acoustic encoding. Craik and Tulving concluded that we process verbal information best through semantic encoding, especially if we apply what is called the **self-reference effect**. The self-reference effect is the tendency for an individual to have better memory for information that relates to oneself in comparison to material that has less personal relevance (Rogers, Kuiper & Kirker, 1977). How could semantic encoding be beneficial to you as you attempt to memorize the concepts in this chapter?

Storage

Once the information has been encoded, we have to somehow retain it. Our brains take the encoded information and place it in storage. **Storage** is the creation of a permanent record of information.

In order for a memory to go into storage (i.e., long-term memory), it has to pass through three distinct stages: **Sensory Memory**, **Short-Term Memory**, and finally **Long-Term Memory**. These stages were first proposed by Richard **Atkinson** and Richard **Shiffrin** (1968). Their model of human memory (***Figure 2***), called Atkinson-Shiffrin (A-S), is based on the belief that we process memories in the same way that a computer processes information.

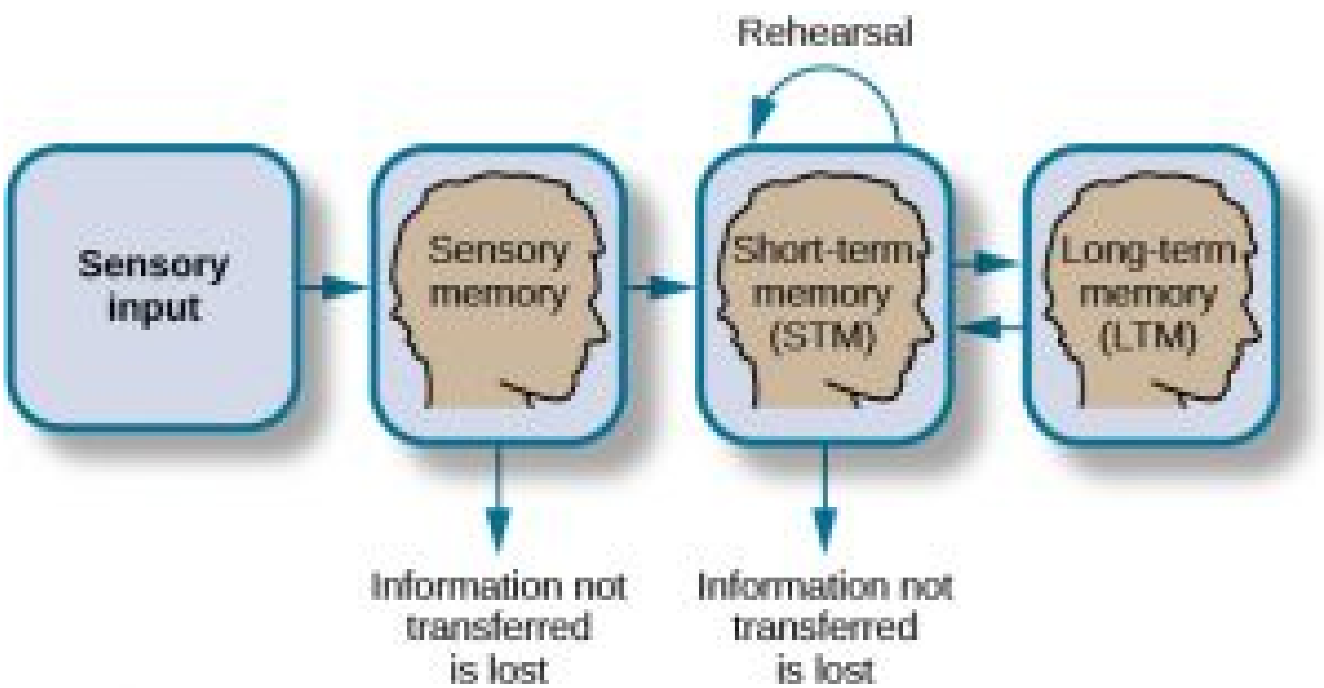


Figure 2. *According to the Atkinson-Shiffrin model of memory, information passes through three distinct stages in order for it to be stored in long-term memory.*

Sensory Memory

In the Atkinson-Shiffrin model, stimuli from the environment are processed first in **sensory memory**: storage of brief sensory events, such as sights, sounds, and tastes. It is very brief storage—up to a couple of seconds. We are constantly bombarded with sensory information. We cannot absorb all of it, or even most of it. And most of it has no impact on our lives. For example, what was your professor wearing the last class period? As long as the professor was dressed appropriately, it does not really matter what she was wearing. Sensory information about sights, sounds, smells, and even textures, which we do not view as valuable information, we discard. If we view something as valuable, the information will move into our short-term memory system.

Short-Term Memory

Short-term memory (STM) is a temporary storage system that processes incoming sensory memory; sometimes it is called working memory. Short-term memory takes information from sensory memory and sometimes connects that memory to something already in long-term memory. Short-term memory storage lasts about 20 seconds. George Miller (1956), in his research on the capacity of memory, found that most people can retain about 7 items in STM. Some remember 5, some 9, so he called the capacity of STM 7 plus or minus 2.

Think of short-term memory as the information you have displayed on your computer screen—a document, a spreadsheet, or a web page. Then, information in short-term memory goes to long-term memory (you save it to your hard drive), or it is discarded (you delete a document or close a web browser). This step of **rehearsal**, the conscious repetition of information to be remembered, to move STM into long-term memory is called **memory consolidation**.

You may find yourself asking, "How much information can our memory handle at once?" To explore the capacity and duration of your short-term memory, have a partner read the strings of random numbers (***Figure 3***) out loud to you, beginning each string by saying, "Ready?" and ending each by saying, "Recall," at which point you should try to write down the string of numbers from memory.

9754	68259	913825	5316842	86951372	719384273
6419	67148	648327	5963827	51739826	163875942

Figure 3. *Work through this series of numbers using the recall exercise explained above to determine the longest string of digits that you can store.*

Note the longest string at which you got the series correct. For most people, this will be close to 7, Miller's famous 7 plus or minus 2. Recall is somewhat better for random numbers than for random letters (Jacobs, 1887), and also often slightly better for information we hear (acoustic encoding) rather than see (visual encoding) (Anderson, 1969).

Long-term Memory

Long-term memory (LTM) is the continuous storage of information. Unlike short-term memory, the storage capacity of LTM has no limits. It encompasses all the things you can remember that happened more than just a few minutes ago to all of the things that you can remember that happened days, weeks, and years ago. In keeping with the computer analogy, the information in your LTM would be like the information you have saved on the hard drive. It isn't there on your desktop (your short-term memory), but you can pull up this information when you want it, at least most of the time. Not all long-term memories are strong memories. Some memories can only be recalled through cues. For example, you might easily recall a fact— "What is the capital of the United States?"—or a procedure—"How do you ride a bike?"—but you might struggle to recall the name of the restaurant you had dinner when you were on vacation in France last summer. A cue, such as that the restaurant was named after its owner, who spoke to you about your shared interest in soccer, may help you recall the name of the restaurant.

Long-term memory is divided into two types: explicit and implicit (***Figure 4***). Understanding the different types is important because a person's age or particular types of brain trauma or disorders can leave certain types of LTM intact while having disastrous consequences for other types. **Explicit memories** are those we consciously try to remember and recall. For example, if you are studying for your chemistry exam, the material you are learning will be part of your explicit memory. (Note: Sometimes, but not always, the terms explicit memory and declarative memory are used interchangeably.)

Implicit memories are memories that are not part of our consciousness. They are memories formed from behaviors. Implicit memory is also called non-declarative memory.

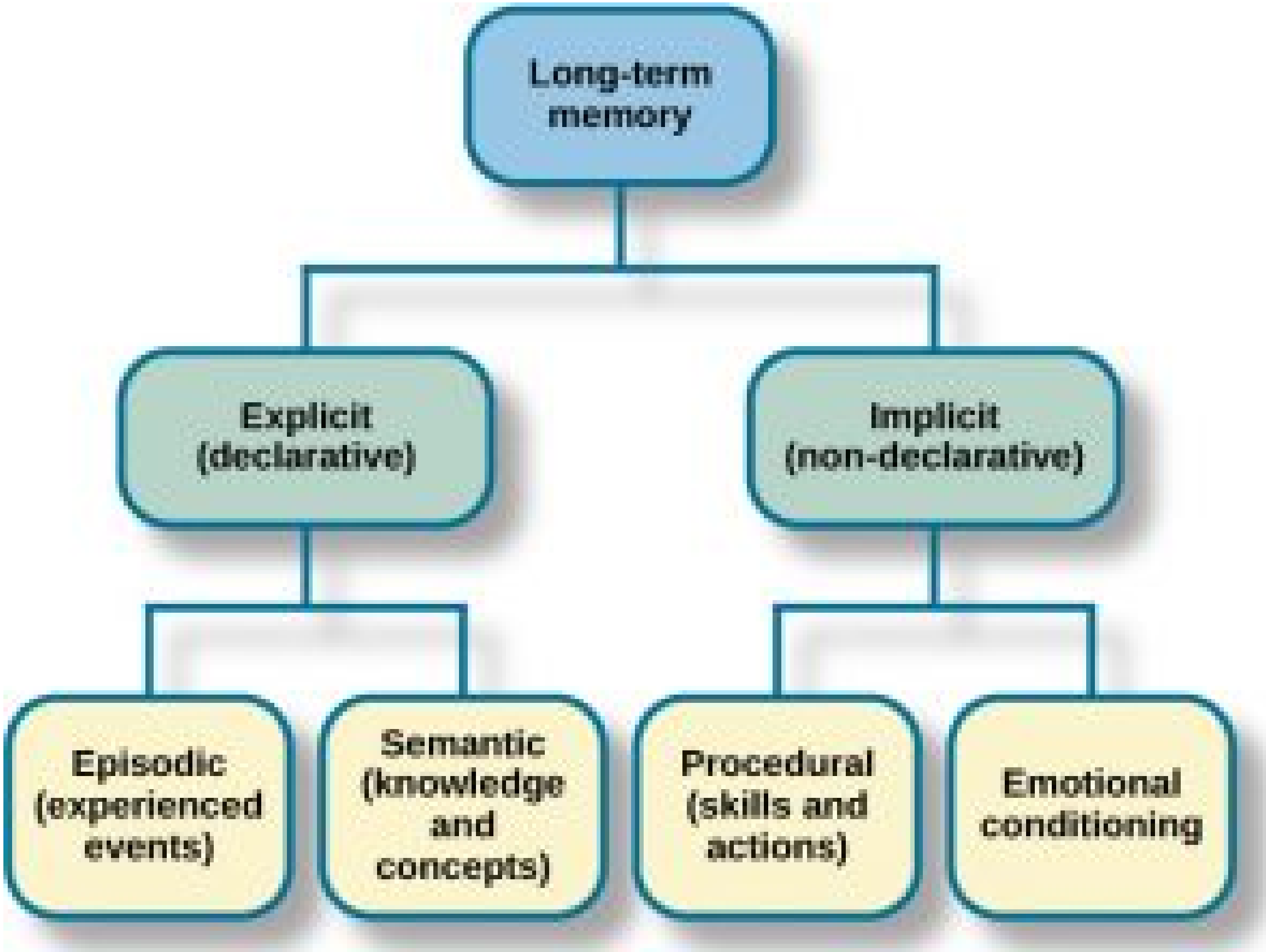


Figure 4. *There are two components of long-term memory: explicit and implicit. Explicit memory includes episodic and semantic memory. Implicit memory includes procedural memory and things learned through conditioning.*

Procedural memory is a type of implicit memory: it stores information about how to do things. It is the memory for skilled actions, such as how to brush your teeth, how to drive a car, how to swim the crawl (freestyle) stroke. If you are learning how to swim freestyle, you practice the stroke: how to move your arms, how to turn your head to alternate breathing from side to side, and how to kick your legs. You would practice this many times until you become good at it. Once you learn how to swim freestyle and your body knows how to move through the water, you will never forget how to swim freestyle, even if you do not swim for a couple of decades. Similarly, if you present an accomplished guitarist with a guitar, even if he has not played in a long time, he will still be able to play quite well.

Declarative memory has to do with the storage of facts and events we personally experienced. Explicit (declarative) memory has two parts: semantic memory and episodic memory. Semantic means having to do with language and knowledge about language. An example would be the question "what does *argumentative* mean?" Stored in our **semantic memory** is knowledge about words, concepts, and language-based knowledge and facts. For example, answers to the following questions are stored in your semantic memory:

1. Who was the first President of the United States?

2. What is democracy?
3. What is the longest river in the world?

Episodic memory is information about events we have personally experienced. The concept of episodic memory was first proposed about 40 years ago (Tulving, 1972). Since then, Tulving and others have looked at scientific evidence and reformulated the theory. Currently, scientists believe that episodic memory is memory about happenings in particular places at particular times, the what, where, and when of an event (Tulving, 2002). It involves recollection of visual imagery as well as the feeling of familiarity (Hassabis & Maguire, 2007).

CAN YOU REMEMBER EVERYTHING YOU EVER DID OR SAID?

Episodic memories are also called autobiographical memories. Let's quickly test your autobiographical memory. What were you wearing exactly five years ago today? What did you eat for lunch on April 10, 2009? You probably find it difficult, if not impossible, to answer these questions. Can you remember every event you have experienced over the course of your life—meals, conversations, clothing choices, weather conditions, and so on? Most likely none of us could even come close to answering these questions; however, American actress Marilu Henner, best known for the television show *Taxi*, can remember. She has an amazing and highly superior autobiographical memory (***Figure 6***).

Figure 6. *Marilu Henner's super autobiographical memory is known as hyperthymesia. (credit: Mark Richardson)*

Very few people can recall events in this way; right now, only 12 known individuals have this ability, and only a few have been studied (Parker, Cahill & McGaugh 2006). And although **hyperthymesia** normally appears in adolescence, two children in the United States appear to have memories from well before their tenth birthdays.

Retrieval

So you have worked hard to encode and store some important information for your upcoming final exam. How do you get that information back out of storage when you need it? The act of getting information out of memory storage and back into conscious awareness is known as **retrieval**. This would be similar to finding and opening a paper you had previously saved on your computer's hard drive. Now it's back on your desktop, and you can work with it again. Our ability to retrieve information from long-term memory is vital to our everyday functioning. You must be able to retrieve information from memory in order to do everything from knowing how to brush your hair and teeth, to driving to work, to knowing how to perform your job once you get there.

There are three ways you can retrieve information out of your long-term memory storage system: recall, recognition, and relearning. **Recall** is what we most often think about when we talk about memory retrieval: it means you can access information without cues. For example, you would use recall for an essay test. **Recognition** happens when you identify information that you have previously learned after encountering it again. It involves a process of comparison. When you take a multiple-choice test, you are relying on recognition to help you choose the correct answer. Here is another example. Let's say you graduated from high school 10 years ago, and you have returned to your hometown for your 10-year reunion. You may not be able to recall all of your classmates, but you recognize many of them based on their yearbook photos.

The third form of retrieval is **relearning**, and it's just what it sounds like. It involves learning information that you previously learned. Whitney took Spanish in high school, but after high school she did not have the opportunity to speak Spanish. Whitney is now 31, and her company has offered her an opportunity to work in their Mexico City office. In order to prepare herself, she enrolls in a Spanish course at the local community center. She's surprised at how quickly she's able to pick up the language after not speaking it for 13 years; this is an example of relearning.

Section Summary

Memory is a system or process that stores what we learn for future use.

Our memory has three basic functions: encoding, storing, and retrieving information. Encoding is the act of getting information into our memory system through automatic or effortful processing. Storage is retention of the information, and retrieval is the act of getting information out of storage and into conscious awareness through recall, recognition, and relearning. The idea that information is processed through three memory systems is called the Atkinson-Shiffrin (A-S) model of memory. First, environmental stimuli enter our sensory memory for a period of less than a second to a few seconds. Those stimuli that we notice and pay attention to then move into short-term memory (also called working memory). According to the A-S model, if we rehearse this information, then it moves into long-term memory for permanent storage. Other models like that of Baddeley and Hitch suggest there is more of a feedback loop between short-term memory and long-term memory. Long-term memory has a practically limitless storage capacity and is divided into implicit and explicit memory. Finally, retrieval is the act of getting memories out of storage and back into conscious awareness. This is done through recall, recognition, and relearning.

Glossary

acoustic encoding
: input of sounds, words, and music

Atkinson-Shiffrin model (A-S)
: memory model that states we process information through three systems: sensory memory, short-term memory, and long-term memory

automatic processing
: encoding of informational details like time, space, frequency, and the meaning of words

declarative memory
type of long-term memory of facts and events we personally experience
effortful processing
encoding of information that takes effort and attention
encoding
input of information into the memory system
episodic memory
type of declarative memory that contains information about events we have personally experienced, also known as autobiographical memory
explicit memory
memories we consciously try to remember and recall
implicit memory
memories that are not part of our consciousness
long-term memory (LTM)
continuous storage of information
memory
system or process that stores what we learn for future use
memory consolidation
active rehearsal to move information from short-term memory into long-term memory
procedural memory
type of long-term memory for making skilled actions, such as how to brush your teeth, how to drive a car, and how to swim
recall
accessing information without cues
recognition
identifying previously learned information after encountering it again, usually in response to a cue
rehearsal
conscious repetition of information to be remembered
relearning
learning information that was previously learned
retrieval
act of getting information out of long-term memory storage and back into conscious awareness
self-reference effect
tendency for an individual to have better memory for information that relates to oneself in comparison to material that has less personal relevance
semantic encoding
input of words and their meaning
semantic memory
type of declarative memory about words, concepts, and language-based knowledge and facts
sensory memory
storage of brief sensory events, such as sights, sounds, and tastes
short-term memory (STM)
(also, working memory) holds about seven bits of information before it is forgotten or stored, as well as information that has been retrieved and is being used
storage
creation of a permanent record of information

visual encoding
input of images

12.

1.12 Parts of the Brain Involved with Memory

Are memories stored in just one part of the brain, or are they stored in many different parts of the brain? Karl Lashley began exploring this problem, about 100 years ago, by making lesions in the brains of animals such as rats and monkeys. He was searching for evidence of the **engram**: the group of neurons that serve as the "physical representation of memory" (Josselyn, 2010). First, Lashley (1950) trained rats to find their way through a maze. Then, he used the tools available at the time—in this case a soldering iron—to create lesions in the rats' brains, specifically in the cerebral cortex. He did this because he was trying to erase the engram, or the original memory trace that the rats had of the maze.

Lashley did not find evidence of the engram, and the rats were still able to find their way through the maze, regardless of the size or location of the lesion. Based on his creation of lesions and the animals' reaction, he formulated the **equipotentiality hypothesis**: if part of one area of the brain involved in memory is damaged, another part of the same area can take over that memory function (Lashley, 1950). Although Lashley's early work did not confirm the existence of the engram, modern psychologists are making progress locating it. Eric Kandel, for example, spent decades working on the synapse, the basic structure of the brain, and its role in controlling the flow of information through neural circuits needed to store memories (Mayford, Siegelbaum, & Kandel, 2012).

Many scientists believe that the entire brain is involved with memory. However, since Lashley's research, other scientists have been able to look more closely at the brain and memory. They have argued that memory is located in specific parts of the brain, and specific neurons can be recognized for their involvement in forming memories. The main parts of the brain involved with memory are the amygdala, the hippocampus, the cerebellum, and the prefrontal cortex (***Figure 1***).

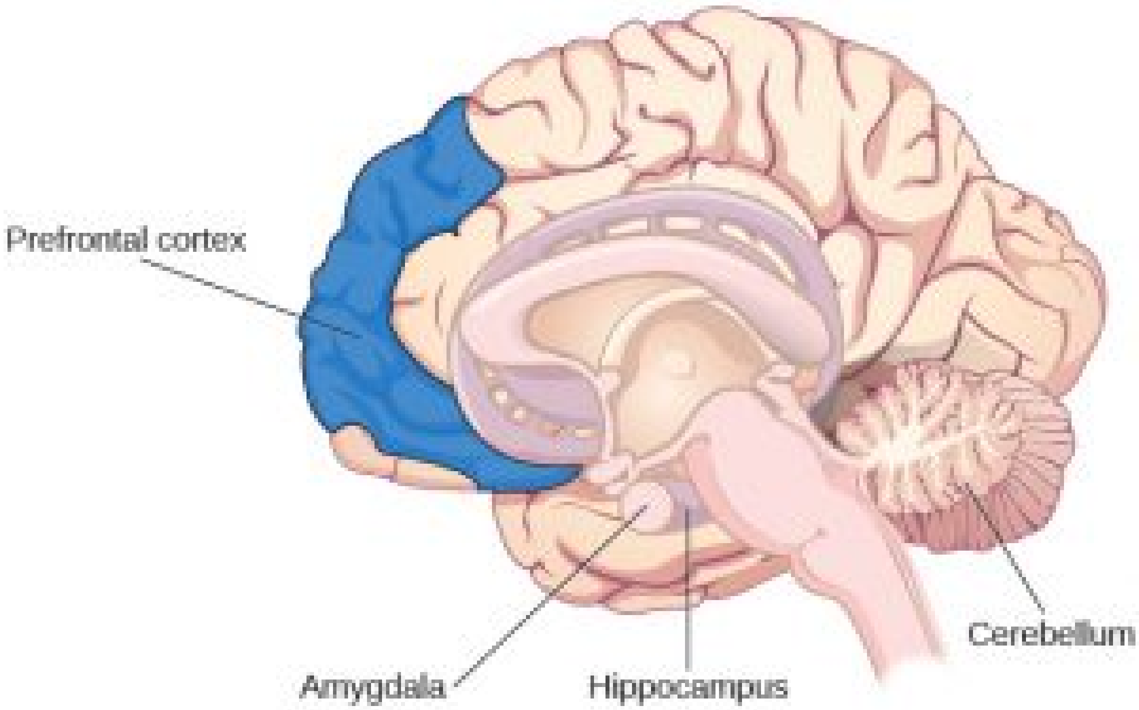


Figure 1. The amygdala is involved in fear and fear memories. The hippocampus is associated with declarative and episodic memory as well as recognition memory. The cerebellum plays a role in processing procedural memories, such as how to play the piano. The prefrontal cortex appears to be involved in remembering semantic tasks.

The Amygdala

First, let's look at the role of the **amygdala** in memory formation. The main job of the amygdala is to regulate emotions, such as fear and aggression (***Figure 1***). The amygdala plays a part in how memories are stored because storage is influenced by stress hormones. For example, one researcher experimented with rats and the fear response (Josselyn, 2010). Using Pavlovian conditioning, a neutral tone was paired with a foot shock to the rats. This produced a fear memory in the rats. After being conditioned, each time they heard the tone, they would freeze (a defense response in rats), indicating a memory for the impending shock. Then the researchers induced cell death in neurons in the lateral amygdala, which is the specific area of the brain responsible for fear memories. They found the fear memory faded (became extinct). Because of its role in processing emotional information, the amygdala is also involved in memory consolidation: the process of transferring new learning into long-term memory. The amygdala seems to facilitate encoding memories at a deeper level when the event is emotionally arousing.

Link to Learning

In this TED Talk called “A Mouse. A Laser Beam. A Manipulated Memory,” Steve Ramirez and Xu Liu from MIT talk about using laser beams to manipulate fear memory in rats. Find out why their work caused a media frenzy once it was published in *Science*.

The Hippocampus

Another group of researchers also experimented with rats to learn how the **hippocampus** functions in memory processing (***Figure 1***). They created lesions in the hippocampi of the rats, and found that the rats demonstrated memory impairment on various tasks, such as object recognition and maze running. They concluded that the hippocampus is involved in memory, specifically normal recognition memory as well as spatial memory (when the memory tasks are like recall tests) (Clark, Zola, & Squire, 2000). Another job of the hippocampus is to project information to cortical regions that give memories meaning and connect them with other connected memories. It also plays a part in memory consolidation: the process of transferring new learning into long-term memory.

Injury to this area leaves us unable to process new declarative memories. One famous patient, known for years only as H. M., had both his left and right temporal lobes (hippocampi) removed in an attempt to help control the seizures he had been suffering from for years (Corkin, Amaral, González, Johnson, & Hyman, 1997). As a result, his declarative memory was significantly affected, and he could not form new semantic knowledge. He lost the ability to form new memories, yet he could still remember information and events that had occurred prior to the surgery.

The Cerrebellum and Prefrontal Cortex

Although the hippocampus seems to be more of a processing area for explicit memories, you could still lose it and be able to create implicit memories (procedural memory, motor learning, and classical conditioning), thanks to your **cerebellum** (***Figure 1***). For example, one classical conditioning experiment is to accustom subjects to blink when they are given a puff of air. When researchers damaged the cerebellums of rabbits, they discovered that the rabbits were not able to learn the conditioned eye-blink response (Steinmetz, 1999; Green & Woodruff-Pak, 2000).

Other researchers have used brain scans, including positron emission tomography (PET) scans, to learn how people process and retain information. From these studies, it seems the prefrontal cortex is involved. In one study, participants had to complete two different tasks: either looking for the letter *a* in words (considered a perceptual task) or categorizing a noun as either living or non-living (considered a semantic task) (Kapur et al., 1994). Participants were then asked which words they had previously seen. Recall was much better for the semantic task than for the perceptual task. According to PET scans, there was much more activation in the left inferior prefrontal cortex in the semantic task. In another study, encoding was associated with left frontal activity, while retrieval of information was associated with the right frontal region (Craik et al., 1999).

Neurotransmitters

There also appear to be specific neurotransmitters involved with the process of memory, such as epinephrine, dopamine, serotonin, glutamate, and acetylcholine (Myhrer, 2003). There continues to be discussion and debate among researchers as to which **neurotransmitter** plays which specific role (Blockland, 1996). Although we don't yet know which role each neurotransmitter plays in memory, we do know that communication among neurons via neurotransmitters is critical for developing new memories. Repeated activity by neurons leads to increased neurotransmitters in the synapses and more efficient and more synaptic connections. This is how memory consolidation occurs.

It is also believed that strong emotions trigger the formation of strong memories, and weaker emotional experiences form weaker memories; this is called **arousal theory** (Christianson, 1992). For example, strong emotional experiences can trigger the release of neurotransmitters, as well as hormones, which strengthen memory; therefore, our memory for an emotional event is usually better than our memory for a non-emotional event. When humans and animals are stressed, the brain secretes more of the neurotransmitter glutamate, which helps them remember the stressful event (McGaugh, 2003). This is clearly evidenced by what is known as the flashbulb memory phenomenon.

A **flashbulb memory** is an exceptionally clear recollection of an important event (***Figure 2***). Where were you when you first heard about the 9/11 terrorist attacks? Most likely you can remember where you were and what you were doing. In fact, a Pew Research Center (2011) survey found that for those Americans who were age 8 or older at the time of the event, 97% can recall the moment they learned of this event, even a decade after it happened.

Figure 2. *Most people can remember where they were when they first heard about the 9/11 terrorist attacks. This is an example of a flashbulb memory: a record of an atypical and unusual event that has very strong emotional associations. (credit: Michael Foran)*

INACCURATE AND FALSE MEMORIES

Even flashbulb memories can have decreased accuracy with the passage of time, even with very important events. For example, on at least three occasions, when asked how he heard about the terrorist attacks of 9/11, President George W. Bush responded inaccurately. In January 2002, less than 4 months after the attacks, the then sitting President Bush was asked how he heard about the attacks. He responded:

"I was sitting there, and my Chief of Staff—well, first of all, when we walked into the classroom, I had seen this plane fly into the first building. There was a TV set on. And you know, I thought it was pilot error and I was amazed that anybody could make such a terrible mistake. (Greenberg, 2004, p. 2)"

Contrary to what President Bush recalled, no one saw the first plane hit, except people on the ground near the twin towers. The first plane was not videotaped because it was a normal Tuesday morning in New York City, until the first plane hit.

Some people attributed Bush's wrong recall of the event to conspiracy theories. However, there is a much more benign explanation: human memory, even flashbulb memories, can be frail. In fact, memory can be so frail that we can convince a person an event happened to them, even when it did not. In studies, research participants will recall hearing a word, even though they never heard the word. For example, participants were given a list of 15 sleep-related words, but the word "sleep" was not on the list. Participants recalled hearing the word "sleep" even though they did not actually hear it (Roediger & McDermott, 2000).

Section Summary

Beginning with Karl Lashley, researchers and psychologists have been searching for the engram, which is the physical trace of memory. Lashley did not find the engram, but he did suggest that memories are distributed throughout the entire brain rather than stored in one specific area. Now we know that three brain areas do play significant roles in the processing and storage of different types of memories: cerebellum, hippocampus, and amygdala. The cerebellum's job is to process procedural memories; the hippocampus is where new memories are encoded; the amygdala helps determine what memories to store, and it plays a part in determining where the memories are stored based on whether we have a strong or weak emotional response to the event. Strong emotional experiences can trigger the release of neurotransmitters, as well as hormones, which strengthen memory, so that memory for an emotional event is usually stronger than memory for a non-emotional event. This is shown by what is known as the flashbulb memory phenomenon: our ability to remember significant life events. However, our memory for life events (autobiographical memory) is not always accurate.

Glossary

arousal theory
: strong emotions trigger the formation of strong memories and weaker emotional experiences form weaker memories

engram
: physical trace of memory

equipotentiality hypothesis
: some parts of the brain can take over for damaged parts in forming and storing memories

flashbulb memory
: exceptionally clear recollection of an important event

13.

1.13 Problems with Memory: Eyewitness Testimony

Memory Construction and Reconstruction

The formulation of new memories is sometimes called **consolidation**, and the process of bringing up old memories is called **reconsolidation**. Yet as we retrieve our memories, we also tend to alter and modify them. A memory pulled from long-term storage into short-term memory is flexible. New events can be added and we can change what we think we remember about past events, resulting in inaccuracies and distortions. People may not intend to distort facts, but it can happen in the process of retrieving old memories and combining them with new memories (Roediger and DeSoto, in press).

Suggestibility

When someone witnesses a crime, that person's memory of the details of the crime is very important in catching the suspect. Because memory is so fragile, witnesses can be easily (and often accidentally) misled due to the problem of suggestibility. **Suggestibility** describes the effects of misinformation from external sources that leads to the creation of false memories. In the fall of 2002, a sniper in the DC area shot people at a gas station, leaving Home Depot, and walking down the street. These attacks went on in a variety of places for over three weeks and resulted in the deaths of ten people. During this time, as you can imagine, people were terrified to leave their homes, go shopping, or even walk through their neighborhoods. Police officers and the FBI worked frantically to solve the crimes, and a tip hotline was set up. Law enforcement received over 140,000 tips, which resulted in approximately 35,000 possible suspects (Newseum, n.d.).

Most of the tips were dead ends, until a white van was spotted at the site of one of the shootings. The police chief went on national television with a picture of the white van. After the news conference, several other eyewitnesses called to say that they too had seen a white van fleeing from the scene of the shooting. At the time, there were more than 70,000 white vans in the area. Police officers, as well as the general public, focused almost exclusively on white vans because they believed the eyewitnesses. Other tips were ignored. When the suspects were finally caught, they were driving a blue sedan.

As illustrated by this example, we are vulnerable to the power of suggestion, simply based on something we see on the news. Or we can claim to remember something that in fact is only a suggestion someone made. It is the suggestion that is the cause of the false memory.

Eyewitness Misidentification

Even though memory and the process of reconstruction can be fragile, police officers, prosecutors, and the courts often rely on eyewitness identification and testimony in the prosecution of criminals. However, faulty eyewitness identification and testimony can lead to wrongful convictions (***Figure 1***).

Figure 1. *In studying cases where DNA evidence has exonerated people from crimes, the* ***Innocence Project*** *discovered that eyewitness misidentification is the leading cause of wrongful convictions (Benjamin N. Cardozo School of Law, Yeshiva University, 2009).*

How does this happen? In 1984, Jennifer Thompson, then a 22-year-old college student in North Carolina, was brutally raped at knifepoint. As she was being raped, she tried to memorize every detail of her rapist's face and physical characteristics, vowing that if she survived, she would help get him convicted. After the police were contacted, a composite sketch was made of the suspect, and Jennifer was shown six photos. She chose two, one of which was of Ronald Cotton. After looking at the photos for 4–5 minutes, she said, "Yeah. This is the one," and then she added, "I think this is the guy." When questioned about this by the detective who asked, "You're sure? Positive?" She said that it was him. Then she asked the detective if she did OK, and he reinforced her choice by telling her she did great. These kinds of unintended cues and suggestions by police officers can lead witnesses to identify the wrong suspect. The district attorney was concerned about her lack of certainty the first time, so she viewed a lineup of seven men. She said she was trying to decide between numbers 4 and 5, finally deciding that Cotton, number 5, "Looks most like him." He was 22 years old.

By the time the trial began, Jennifer Thompson had absolutely no doubt that she was raped by Ronald Cotton. She testified at the court hearing, and her testimony was compelling enough that it helped convict him. How did she go from, "I think it's the guy" and it "Looks most like him," to such certainty? Gary Wells and Deah Quinlivan (2009) assert it's suggestive police identification procedures, such as stacking lineups to make the defendant stand out, telling the witness which person to identify, and confirming witnesses choices by telling them "Good choice," or "You picked the guy."

After Cotton was convicted of the rape, he was sent to prison for life plus 50 years. After 4 years in prison, he was able to get a new trial. Jennifer Thompson once again testified against him. This time Ronald Cotton was given two life sentences. After serving 11 years in prison, DNA evidence finally demonstrated that Ronald Cotton did not commit the rape, was innocent, and had served over a decade in prison for a crime he did not commit.

Ronald Cotton's story, unfortunately, is not unique. There are also people who were convicted and placed on death row, who were later exonerated. The Innocence Project is a non-profit group that works to exonerate falsely convicted people, including those convicted by eyewitness testimony. To learn more, you can visit http://www.innocenceproject.org.

PRESERVING EYEWITNESS MEMORY: THE ELIZABETH SMART CASE

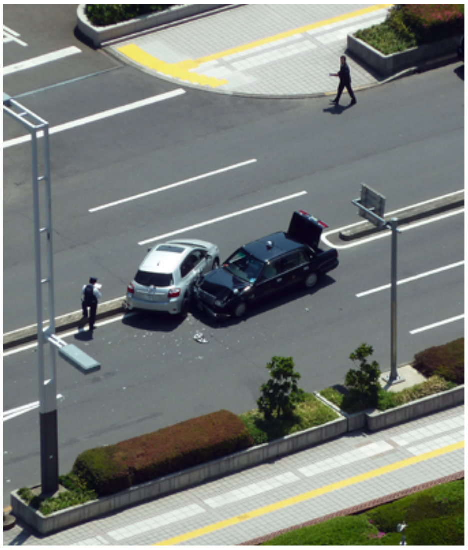
Contrast the Cotton case with what happened in the Elizabeth Smart case. When Elizabeth was 14 years old and fast asleep in her bed at home, she was abducted at knifepoint. Her nine-year-old sister, Mary Katherine, was sleeping in the same bed and watched, terrified, as her beloved older sister was abducted. Mary Katherine was the sole eyewitness to this crime and was very fearful. In the coming weeks, the Salt Lake City police and the FBI proceeded with caution with Mary Katherine. They did not want to implant any false memories or mislead her in any way. They did not show her police line-ups or push her to do a composite sketch of the abductor. They knew if they corrupted her memory, Elizabeth might never be found. For several months, there was little or no progress on the case. Then, about 4 months after the kidnapping, Mary Katherine first recalled that she had heard the abductor's voice prior to that night (he had worked one time as a handyman at the family's home) and then she was able to name the person whose voice it was. The family contacted the press and others recognized him—after a total of nine months, the suspect was caught and Elizabeth Smart was returned to her family.

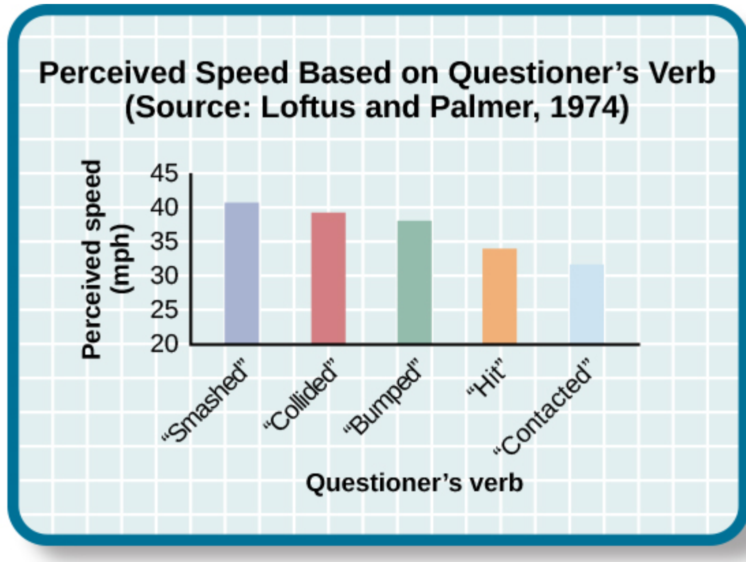
The Misinformation Effect

Cognitive psychologist Elizabeth Loftus has conducted extensive research on memory. She has studied false memories as well as recovered memories of childhood sexual abuse. Loftus also developed the **misinformation effect paradigm**, which holds that after exposure to incorrect information, a person may misremember the original event.

According to Loftus, an eyewitness's memory of an event is very flexible due to the misinformation effect. To test this theory, Loftus and John Palmer (1974) asked 45 U.S. college students to estimate the speed of cars using different forms of questions (***Figure 2***). The participants were shown films of car accidents and were asked to play the role of the eyewitness and describe what happened. They were asked, "About how fast were the cars going when they (smashed, collided, bumped, hit, contacted) each other?" The participants estimated the speed of the cars based on the verb used.

Participants who heard the word "smashed" estimated that the cars were traveling at a much higher speed than participants who heard the word "contacted." The implied information about speed, based on the verb they heard, had an effect on the participants' memory of the accident. In a follow-up one week later, participants were asked if they saw any broken glass (none was shown in the accident pictures). Participants who had been in the “smashed” group were more than twice as likely to indicate that they did remember seeing glass. Loftus and Palmer demonstrated that a leading question encouraged them to not only remember the cars were going faster, but to also falsely remember that they saw broken glass.





(a) (b)

Figure 2. *When people are asked leading questions about an event, their memory of the event may be altered. (credit a: modification of work by Rob Young)*

Controversies over Repressed and Recovered Memories

Other researchers have described how whole events, not just words, can be falsely recalled, even when they did not happen. The idea that memories of traumatic events could be repressed has been a theme in the field of psychology, beginning with Sigmund Freud, and the controversy surrounding the idea continues today.

Recall of false autobiographical memories is called **false memory syndrome**. This syndrome has received a lot of publicity, particularly as it relates to memories of events that do not have independent witnesses—often the only witnesses to the abuse are the perpetrator and the victim (e.g., sexual abuse).

On one side of the debate are those who have recovered memories of childhood abuse years after it occurred. These researchers argue that some children’s experiences have been so traumatizing and distressing that they must lock those memories away in order to lead some semblance of a normal life. They believe that repressed memories can be locked away for decades and later recalled intact through hypnosis and guided imagery techniques (Devilly, 2007).

Research suggests that having no memory of childhood sexual abuse is quite common in adults. For instance, one large-scale study conducted by John Briere and Jon Conte (1993) revealed that 59% of 450 men and women who were receiving treatment for sexual abuse that had occurred before age 18 had forgotten their experiences. Ross Cheit (2007) suggested that repressing these memories created psychological distress in adulthood. The Recovered Memory Project was created so that victims of childhood sexual abuse can recall these memories and allow the healing process to begin (Cheit, 2007; Devilly, 2007).

On the other side, Loftus has challenged the idea that individuals can repress memories of traumatic events from childhood, including sexual abuse, and then recover those memories years later through therapeutic techniques such as hypnosis, guided visualization, and age regression.

Loftus is not saying that childhood sexual abuse doesn't happen, but she does question whether or not those memories are accurate, and she is skeptical of the questioning process used to access these memories, given that even the slightest suggestion from the therapist can lead to misinformation effects. For example, researchers Stephen Ceci and Maggie Brucks (1993, 1995) asked three-year-old children to use an anatomically correct doll to show where their pediatricians had touched them during an exam. Fifty-five percent of the children pointed to the genital/anal area on the dolls, even when they had not received any form of genital exam.

Ever since Loftus published her first studies on the suggestibility of eyewitness testimony in the 1970s, social scientists, police officers, therapists, and legal practitioners have been aware of the flaws in interview practices. Consequently, steps have been taken to decrease suggestibility of witnesses. One way is to modify how witnesses are questioned. When interviewers use neutral and less leading language, children more accurately recall what happened and who was involved (Goodman, 2006; Pipe, 1996; Pipe, Lamb, Orbach, & Esplin, 2004). Another change is in how police lineups are conducted. It's recommended that a blind photo lineup be used. This way the person administering the lineup doesn't know which photo belongs to the suspect, minimizing the possibility of giving leading cues. Additionally, judges in some states now inform jurors about the possibility of misidentification. Judges can also suppress eyewitness testimony if they deem it unreliable.

Section Summary

All of us at times have felt dismayed, frustrated, and even embarrassed when our memories have failed us. Our memory is flexible and prone to many errors, which is why eyewitness testimony has been found to be largely unreliable. Understanding the factors that contribute to memory distortion may help investigators, therapists, and us to lessen the influence of these factors on how we remember past events.

Glossary

absentmindedness
: lapses in memory that are caused by breaks in attention or our focus being somewhere else

amnesia
: loss of long-term memory that occurs as the result of disease, physical trauma, or psychological trauma

bias
: how feelings and view of the world distort memory of past events

blocking
: memory error in which you cannot access stored information

consolidation
: formulation of new memories

false memory syndrome
recall of false autobiographical memories

misattribution
memory error in which you confuse the source of your information

misinformation effect paradigm
after exposure to incorrect information, a person may misremember the original event

persistence
failure of the memory system that involves the involuntary recall of unwanted memories, particularly unpleasant ones

reconsolidation
process of bringing up old memories that might be distorted by new information

suggestibility
effects of misinformation from external sources that leads to the creation of false memories

transience
memory error in which unused memories fade with the passage of time

II

Theme 2: How Does Blood and Organ Donation Work?

For more than a century, physicians have been administering blood from donors to recipients in need of blood, and for more than 50 years modern medical techniques have allowed patients with non-functional organs to extend their lives for decades through transplantation. In a transplant, an organ or tissue is removed from a donor (either a living person, or one who is very recently deceased) and surgically implanted into the body of a recipient whose non-functional organ or tissue has first been removed. Organ and tissue transplantation are not always successful, however, and almost all early attempts at organ transplantation failed because of incompatability between the donor's tissues or organs and recipient's immune system. The human immune system attacks foreign particles in the body to prevent microbial infection, but can also attack transplanted tissues and organs, preventing them from functioning in a recipient's body. Before the role of the immune system in organ rejection was understood, tissue and organ donation was rarely successful, and often resulted in severe and sometimes fatal immune reactions in the organ recipient. The first successful organ transplant occurred in 1954 Dr. Joseph Murray in Boston, Massachusetts. Dr. Murray removed a kidney from a healthy young man and transplanted it into his identical twin brother, who then survived for more than 8 years. Dr. Murray won a Nobel prize for his work on the role of the immune system in organ transplantation and rejection.

In this section of the course, we will be focusing on the scientific and ethical issues surrounding organ and tissue transplantation. First, we will learn about blood, and the cardiovascular and respiratory systems. Then, we will learn about the immune system, the body's defense against microbial invaders, and how our understanding of its function is crucial for successful organ and tissue transplantation. In both lecture and lab, we will explore the processes through which the genetic information in human DNA is decoded by cells to produce actual physical differences in cells, in the processes of transcription and translation. We will then uncover the mechanisms of inheriting DNA from our parents, and passing it on to our children, and how these patterns of inheritance influence our physical traits. Additionally, we will address issues related to ethical design of research studies involving humans and animals, in order to prepare for some of the lab work coming up in the next part of the course.

14.

2.1 Ethics of Research

Research Involving Human Participants

Any experiment involving the participation of human subjects is governed by extensive, strict guidelines designed to ensure that the experiment does not result in harm. Any research institution that receives federal support for research involving human participants must have access to an **institutional review board (IRB)**. The IRB is a committee of individuals often made up of members of the institution's administration, scientists, and community members (***Figure 1***). The purpose of the IRB is to review proposals for research that involves human participants. The IRB reviews these proposals with the principles mentioned above in mind, and generally, approval from the IRB is required in order for the experiment to proceed.

Figure 1. *An institution's IRB meets regularly to review experimental proposals that involve human participants. (credit: modification of work by Lowndes Area Knowledge Exchange (LAKE)/Flickr)*

An institution's IRB requires several components in any experiment it approves. For one, each participant must sign an informed consent form before they can participate in the experiment. An **informed consent** form provides a written description of what participants can expect during the experiment, including potential risks and implications of the research. It also lets participants know that their involvement is completely voluntary and can be discontinued without penalty at any time. Furthermore, the informed consent guarantees that any data collected in the experiment will remain completely confidential. In cases where research participants are under the age of 18, the parents or legal guardians are required to sign the informed consent form.

Link to Learning

Visit this website to see an example of a consent form.

While the informed consent form should be as honest as possible in describing exactly what participants will be doing, sometimes deception is necessary to prevent participants' knowledge of the exact research question from affecting the results of the study. **Deception** involves purposely misleading experiment participants in order to maintain the integrity of the experiment, but not to the point where the deception could be considered harmful. For example, if we are interested in how our opinion of someone is affected by their attire, we might use deception in describing the experiment to prevent that knowledge from affecting participants' responses. In cases where deception is involved, participants must receive a full **debriefing** upon conclusion of the study—complete, honest information about the purpose of the experiment, how the data collected will be used, the reasons why deception was necessary, and information about how to obtain additional information about the study.

ETHICS AND THE TUSKEGEE SYPHILIS STUDY

Unfortunately, the ethical guidelines that exist for research today were not always applied in the past. In 1932, poor, rural, black, male sharecroppers from Tuskegee, Alabama, were recruited to participate in an experiment conducted by the U.S. Public Health Service, with the aim of studying syphilis in black men (***Figure 2***). In exchange for free medical care, meals, and burial insurance, 600 men agreed to participate in the study. A little more than half of the men tested positive for syphilis, and they served as the experimental group (given that the researchers could not randomly assign participants to groups, this represents a quasi-experiment). The remaining syphilis-free individuals served as the control group. However, those individuals that tested positive for syphilis were never informed that they had the disease.

While there was no treatment for syphilis when the study began, by 1947 penicillin was recognized as an effective treatment for the disease. Despite this, no penicillin was administered to the participants in this study, and the participants were not allowed to seek treatment at any other facilities if they continued in the study. Over the course of 40 years, many of the participants unknowingly spread syphilis to their wives (and subsequently their children born from their wives) and eventually died because they never received treatment for the disease. This study was discontinued in 1972 when the experiment was discovered by the national press (Tuskegee University, n.d.). The resulting outrage over the experiment led directly to the National Research Act of 1974 and the strict ethical guidelines for research on humans described in this chapter. Why is this study unethical? How were the men who participated and their families harmed as a function of this research?

Figure 2. *A participant in the Tuskegee Syphilis Study receives an injection.*

Link to Learning

Visit this [website] to learn more about the Tuskegee Syphilis Study.

Research Involving Animal Subjects

Many psychologists, biologists and health care researchers conduct research involving animal subjects. Often, these researchers use rodents (***Figure 3***) or birds as the subjects of their experiments. Because many basic processes in animals are sufficiently similar to those in humans, these animals are acceptable substitutes for research that would be considered unethical in human participants.

Figure 3. *Rats, like the one shown here, often serve as the subjects of animal research.*

This does not mean that animal researchers are immune to ethical concerns. Indeed, the humane and ethical treatment of animal research subjects is a critical aspect of this type of research. Researchers must design their experiments to minimize any pain or distress experienced by animals serving as research subjects.

Whereas IRBs review research proposals that involve human participants, animal experimental proposals are reviewed by an **Institutional Animal Care and Use Committee (IACUC)**. An IACUC consists of institutional administrators, scientists, veterinarians, and community members. This committee is charged with ensuring that all experimental proposals require the humane treatment of animal research subjects. It also conducts semi-annual inspections of all animal facilities to ensure that the research protocols are being followed. No animal research project can proceed without the committee's approval.

Section Summary

Ethics in research is an evolving field, and some practices that were accepted or tolerated in the past would be considered unethical today. Researchers are expected to adhere to basic ethical guidelines when conducting experiments that involve human participants. Any experiment involving human participants must be approved by an IRB. Participation in experiments is voluntary and requires informed consent of the participants. If any deception is involved in the experiment, each participant must be fully debriefed upon the conclusion of the study.

Animal research is also held to a high ethical standard. Researchers who use animals as experimental subjects must design their projects so that pain and distress are minimized. Animal research requires the approval of an IACUC, and all animal facilities are subject to regular inspections to ensure that animals are being treated humanely.

Glossary

debriefing
: when an experiment involved deception, participants are told complete and truthful information about the experiment at its conclusion

deception
: purposely misleading experiment participants in order to maintain the integrity of the experiment

informed consent
: process of informing a research participant about what to expect during an experiment, any risks involved, and the implications of the research, and then obtaining the person’s consent to participate

Institutional Animal Care and Use Committee (IACUC)
: committee of administrators, scientists, veterinarians, and community members that reviews proposals for research involving non-human animals

Institutional Review Board (IRB)
: committee of administrators, scientists, and community members that reviews proposals for research involving human participants

15.

2.2 Circulatory and Respiratory Systems

The medium for transport of gases and other molecules is the blood, which continually circulates through the system. Pressure differences within the system cause the movement of the blood and are created by the pumping of the heart.

Gas exchange between tissues and the blood is an essential function of the circulatory system. In humans, other mammals, and birds, blood absorbs oxygen and releases carbon dioxide in the lungs. Thus the circulatory and respiratory system, whose function is to obtain oxygen and discharge carbon dioxide, work in tandem.

The Respiratory System

Take a breath in and hold it. Wait several seconds and then let it out. Humans, when they are not exerting themselves, breathe approximately 15 times per minute on average. This equates to about 900 breaths an hour or 21,600 breaths per day. With every inhalation, air fills the lungs, and with every exhalation, it rushes back out. That air is doing more than just inflating and deflating the lungs in the chest cavity. The air contains oxygen that crosses the lung tissue, enters the bloodstream, and travels to organs and tissues. There, oxygen is exchanged for carbon dioxide, which is a cellular waste material. Carbon dioxide exits the cells, enters the bloodstream, travels back to the lungs, and is expired out of the body during exhalation.

Breathing is both a voluntary and an involuntary event. How often a breath is taken and how much air is inhaled or exhaled is regulated by the respiratory center in the brain in response to signals it receives about the carbon dioxide content of the blood. However, it is possible to override this automatic regulation for activities such as speaking, singing and swimming under water.

During inhalation the **diaphragm** descends creating a negative pressure around the lungs and they begin to inflate, drawing in air from outside the body. The air enters the body through the nasal cavity located just inside the nose (***Figure 1***). As the air passes through the **nasal cavity**, the air is warmed to body temperature and humidified by moisture from mucous membranes. These processes help equilibrate the air to the body conditions, reducing any damage that cold, dry air can cause. Particulate matter that is floating in the air is removed in the nasal passages by hairs, mucus, and cilia. Air is also chemically sampled by the sense of smell.

From the nasal cavity, air passes through the **pharynx** (throat) and the **larynx** (voice box) as it makes its way to the **trachea** (***Figure 1***). The main function of the trachea is to funnel the inhaled air to the lungs and the exhaled air back out of the body. The human trachea is a cylinder, about 25 to 30 cm (9.8–11.8 in) long, which sits in front of the esophagus and extends from the pharynx into the chest cavity to the lungs. It is made of incomplete rings of cartilage and smooth muscle. The cartilage provides strength and support to the trachea to keep the passage open. The trachea is lined with cells that have cilia and secrete mucus. The mucus catches particles that have been inhaled, and the cilia move the particles toward the pharynx.

The end of the trachea divides into two bronchi that enter the right and left lung. Air enters the lungs through the **primary bronchi**. The primary bronchus divides, creating smaller and smaller diameter **bronchi** until the passages are under 1 mm (.03 in) in diameter when they are called **bronchioles** as they split and spread through the lung. Like the trachea, the bronchus and bronchioles are made of cartilage and smooth muscle. Bronchi are innervated by nerves of both the parasympathetic and sympathetic nervous systems that control muscle contraction (parasympathetic) or relaxation (sympathetic) in the bronchi and bronchioles, depending on the nervous system's cues. The final bronchioles are the respiratory bronchioles. Alveolar ducts are attached to the end of each respiratory bronchiole. At the end of each duct are alveolar sacs, each containing 20 to 30 **alveoli**. Gas exchange occurs only in the alveoli. The alveoli are thin-walled and look like tiny bubbles within the sacs. The alveoli are in direct contact with capillaries of the circulatory system. Such intimate contact ensures that oxygen will diffuse from the alveoli into the blood. In addition, carbon dioxide will diffuse from the blood into the alveoli to be exhaled. The anatomical arrangement of capillaries and alveoli emphasizes the structural and functional relationship of the respiratory and circulatory systems. Estimates for the surface area of alveoli in the lungs vary around 100 m^2. This large area is about the area of half a tennis court. This large surface area, combined with the thin-walled nature of the alveolar cells, allows gases to easily diffuse across the cells.

ART CONNECTION

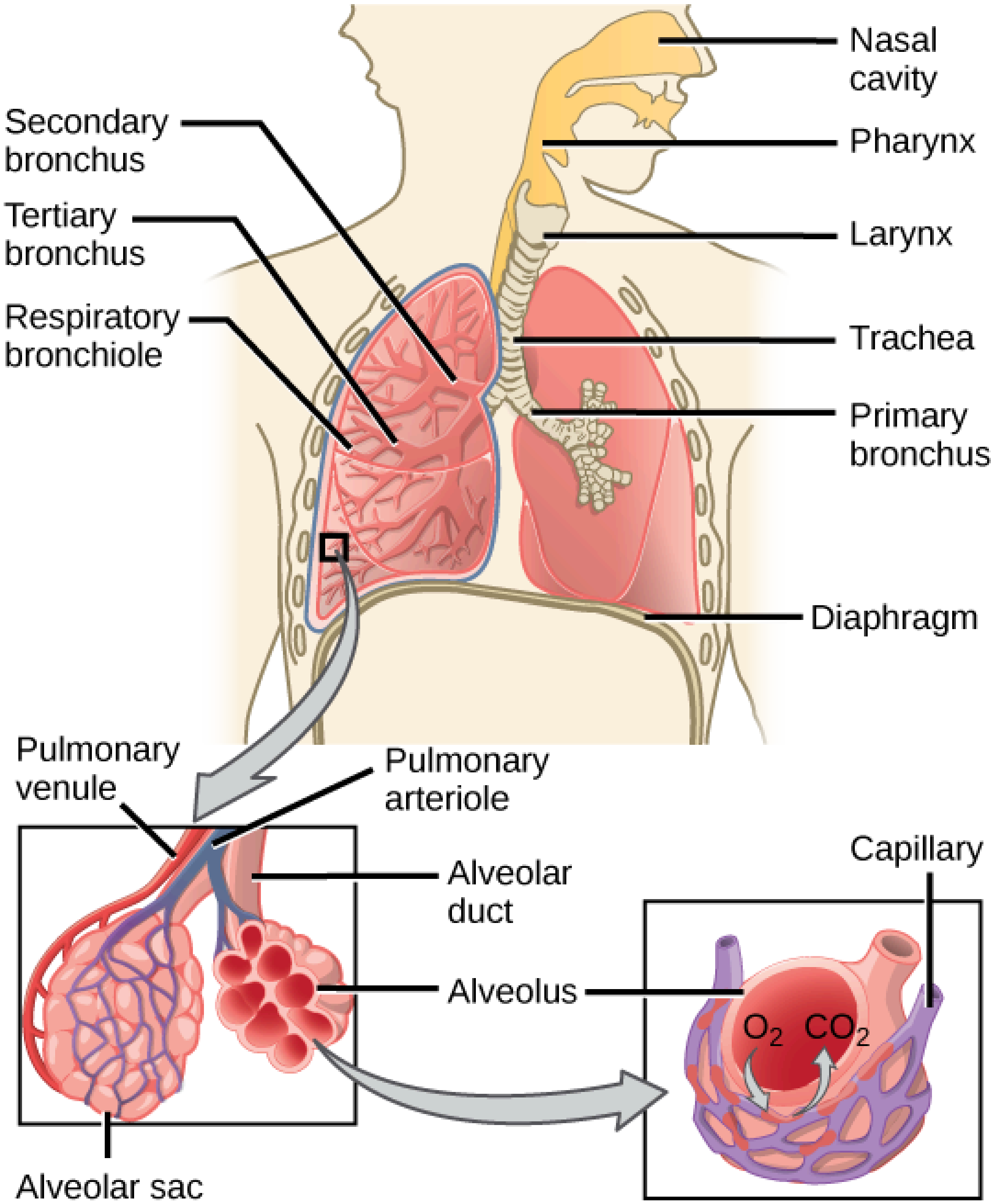


Figure 1. *Air enters the respiratory system through the nasal cavity, and then passes through the pharynx and the trachea into the lungs. (credit: modification of work by NCI)*

Which of the following statements about the human respiratory system is false?

a. When we breathe in, air travels from the pharynx to the trachea.

b. The bronchioles branch into bronchi.

c. Alveolar ducts connect to alveolar sacs.

d. Gas exchange between the lungs and blood takes place in the alveolus.

Concept in Action

Watch this video for a review of the respiratory system.

The Circulatory System

The circulatory system is a network of vessels—the arteries, veins, and capillaries—and a pump, the heart. In all vertebrate organisms this is a closed-loop system, in which the blood is largely separated from the body's other extracellular fluid compartment, the interstitial fluid, which is the fluid bathing the cells. Blood circulates inside blood vessels and circulates unidirectionally from the heart around one of two circulatory routes, then returns to the heart again; this is a **closed circulatory system**. **Open circulatory systems** are found in invertebrate animals in which the circulatory fluid bathes the internal organs directly even though it may be moved about with a pumping heart.

The Heart

The heart is a complex muscle that consists of two pumps: one that pumps blood through **pulmonary circulation** to the lungs, and the other that pumps blood through **systemic circulation** to the rest of the body's tissues (and the heart itself).

The heart is asymmetrical, with the left side being larger than the right side, correlating with the different sizes of the pulmonary and systemic circuits (***Figure 2***). In humans, the heart is about the size of a clenched fist; it is divided into four chambers: two atria and two ventricles. There is one **atrium** and one **ventricle** on the right side and one atrium and one ventricle on the left side. The right atrium receives deoxygenated blood from the systemic circulation through the major veins: the **superior vena cava**, which drains blood from the head and from the veins that come from the arms, as well as the **inferior vena cava**, which drains blood from the veins that come from the lower organs and the legs. This deoxygenated blood then passes to the right ventricle through the **tricuspid valve**, which prevents the backflow of blood. After it is filled, the right ventricle contracts, pumping the blood to the lungs for reoxygenation. The left atrium receives the oxygen-rich blood from the lungs. This blood passes through the **bicuspid valve** to the left ventricle where the blood is pumped into the **aorta**. The aorta is the major artery of the body, taking oxygenated blood to the organs and muscles of the body. This pattern of pumping is referred to as double circulation and is found in all mammals. (***Figure 2***).

ART CONNECTION

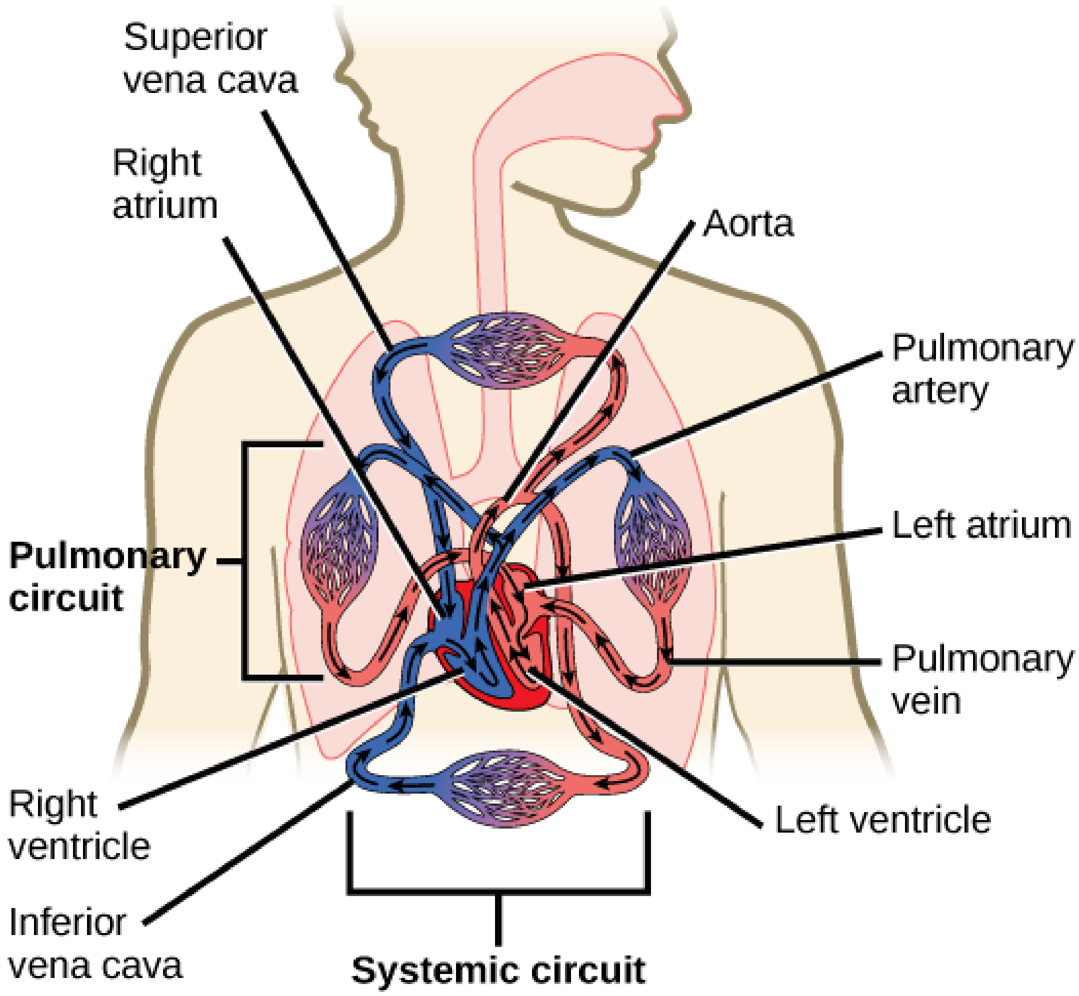


Figure 2. *The heart is divided into four chambers, two atria, and two ventricles. Each chamber is separated by one-way valves. The right side of the heart receives deoxygenated blood from the body and pumps it to the lungs. The left side of the heart pumps blood to the rest of the body.*

Which of the following statements about the circulatory system is false?

a. Blood in the pulmonary vein is deoxygenated.

b. Blood in the inferior vena cava is deoxygenated.

c. Blood in the pulmonary artery is deoxygenated.

d. Blood in the aorta is oxygenated.

The Cardiac Cycle

The main purpose of the heart is to pump blood through the body; it does so in a repeating sequence called the cardiac cycle. The **cardiac cycle** is the flow of blood through the heart coordinated by electrochemical signals that cause the heart muscle to contract and relax. In each cardiac cycle, a sequence of contractions pushes out the blood, pumping it through the body; this is followed by a relaxation phase, where the heart fills with blood. These two phases are called the **systole** (contraction) and **diastole** (relaxation), respectively (***Figure 3***). The signal for contraction begins at a location on the outside of the right atrium. The electrochemical signal moves from there across the atria causing them to contract. The contraction of the atria forces blood through the valves into the ventricles. Closing of these valves caused by the contraction of the ventricles produces a "lub" sound. The signal has, by this time, passed down the walls of the heart, through a point between the right atrium and right ventricle. The signal then causes the ventricles to contract. The ventricles contract together forcing blood into the aorta and the pulmonary arteries. Closing of the valves to these arteries caused by blood being drawn back toward the heart during ventricular relaxation produces a monosyllabic "dub" sound.

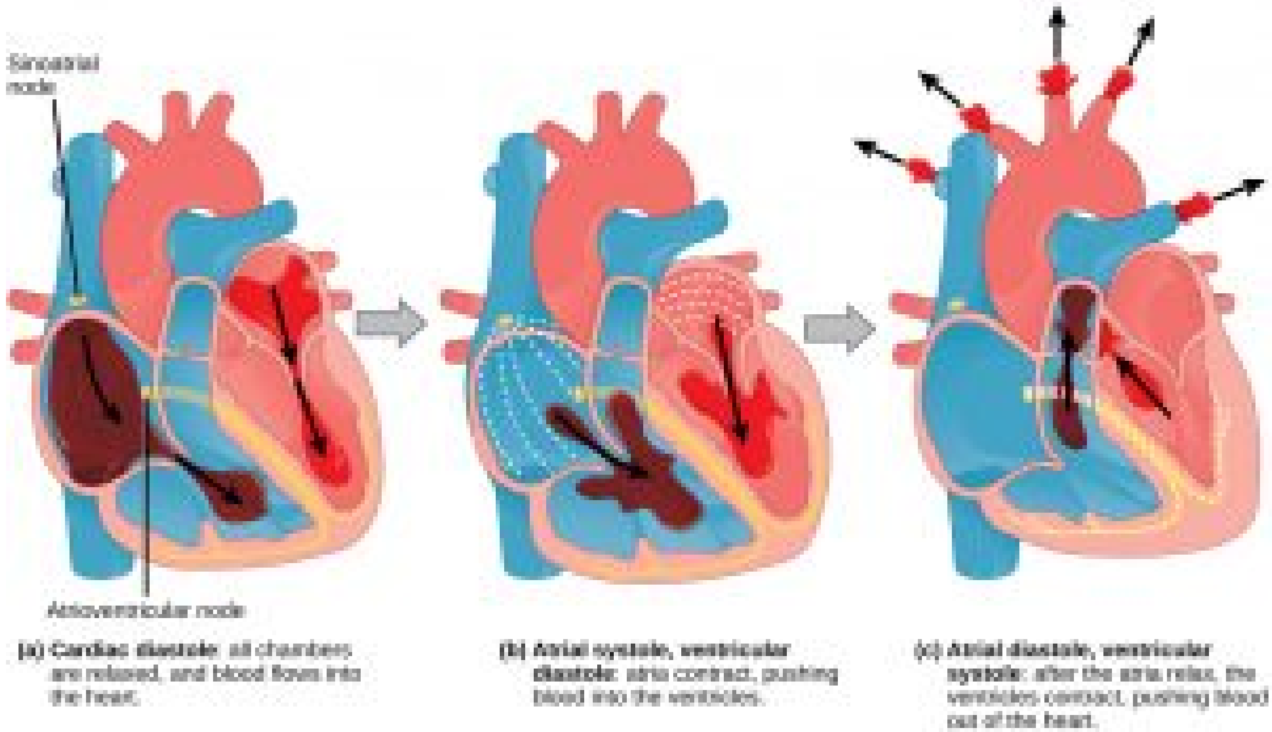


Figure 3. *In each cardiac cycle, a series of contractions (systoles) and relaxations (diastoles) pumps blood through the heart and through the body. (a) During cardiac diastole, blood flows into the heart while all chambers are relaxed. (b) Then the ventricles remain relaxed while atrial systole pushes blood into the ventricles. (c) Once the atria relax again, ventricle systole pushes blood out of the heart.*

The pumping of the heart is a function of the cardiac muscle cells, or cardiomyocytes, that make up the heart muscle. Cardiomyocytes are distinctive muscle cells that are striated like skeletal muscle but pump rhythmically and involuntarily like smooth muscle; adjacent cells are connected by intercalated disks found only in cardiac muscle. These connections allow the electrical signal to travel directly to neighboring muscle cells.

The electrical impulses in the heart produce electrical currents that flow through the body and can be measured on the skin using electrodes. This information can be observed as an **electrocardiogram (ECG)** a recording of the electrical impulses of the cardiac muscle.

Concept in Action

Visit the following website to see the heart's pacemaker, or electrocardiogram system, in action.

Blood Vessels

The blood from the heart is carried through the body by a complex network of blood vessels (***Figure 4***). **Arteries** take blood away from the heart. The main artery of the systemic circulation is the aorta; it branches into major arteries that take blood to different limbs and organs. The aorta and arteries near the heart have heavy but elastic walls that respond to and smooth out the pressure differences caused by the beating heart. Arteries farther away from the heart have more muscle tissue in their walls that can constrict to affect flow rates of blood. The major arteries diverge into minor arteries, and then smaller vessels called arterioles, to reach more deeply into the muscles and organs of the body.

Arterioles diverge into capillary beds. Capillary beds contain a large number, 10’s to 100’s of **capillaries** that branch among the cells of the body. Capillaries are narrow-diameter tubes that can fit single red blood cells and are the sites for the exchange of nutrients, waste, and oxygen with tissues at the cellular level. Fluid also leaks from the blood into the interstitial space from the capillaries. The capillaries converge again into venules that connect to minor veins that finally connect to major veins. **Veins** are blood vessels that bring blood high in carbon dioxide back to the heart. Veins are not as thick-walled as arteries, since pressure is lower, and they have valves along their length that prevent backflow of blood away from the heart. The major veins drain blood from the same organs and limbs that the major arteries supply.

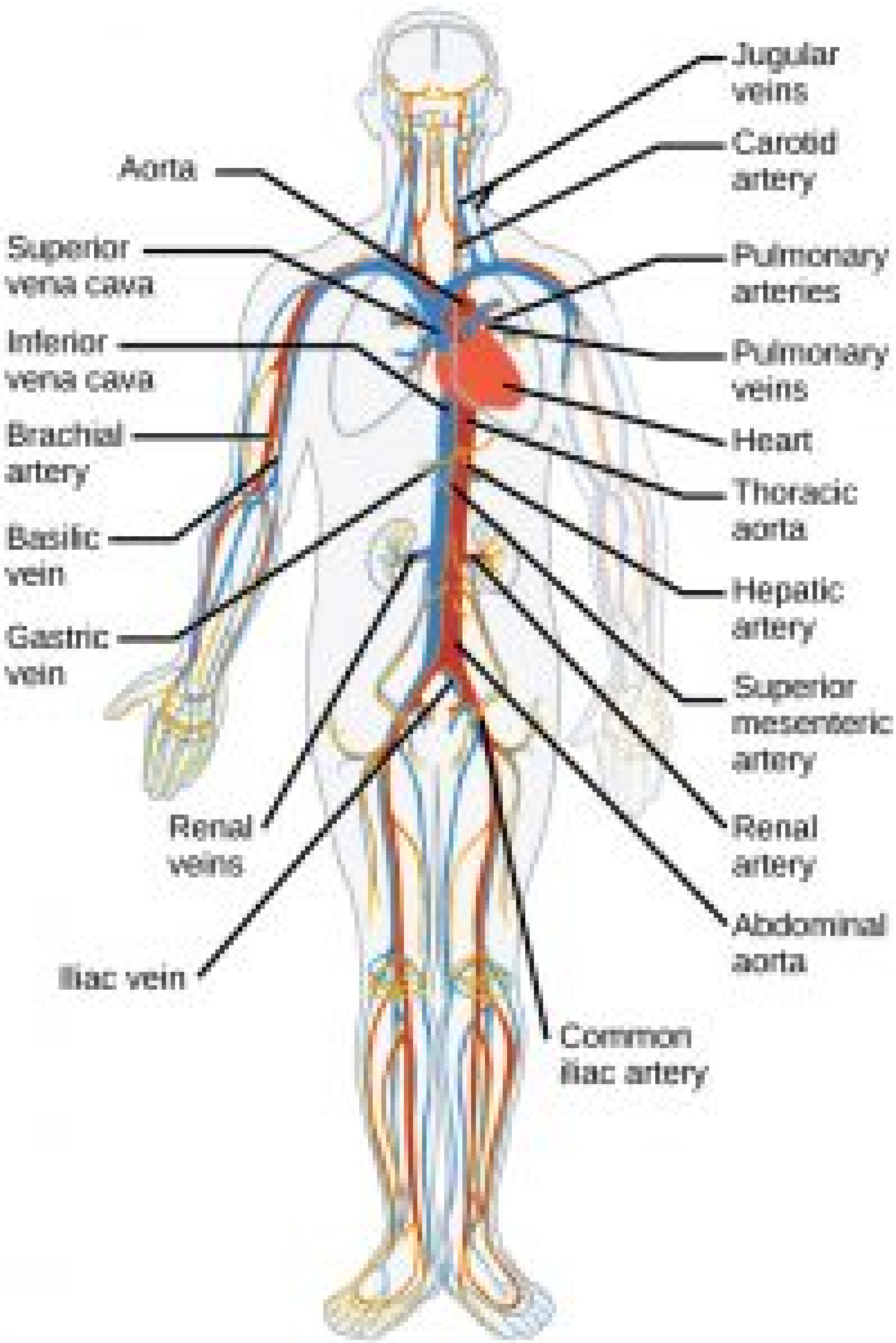


Figure 4. *The arteries of the body, indicated in red, start at the aortic arch and branch to supply the organs and muscles of the body with oxygenated blood. The veins of the body, indicated in blue,*

return blood to the heart. The pulmonary arteries are blue to reflect the fact that they are deoxygenated, and the pulmonary veins are red to reflect that they are oxygenated. (credit: modification of work by Mariana Ruiz Villareal)

Section Summary

Animal respiratory systems are designed to facilitate gas exchange. In mammals, air is warmed and humidified in the nasal cavity. Air then travels down the pharynx and larynx, through the trachea, and into the lungs. In the lungs, air passes through the branching bronchi, reaching the respiratory bronchioles. The respiratory bronchioles open up into the alveolar ducts, alveolar sacs, and alveoli. Because there are so many alveoli and alveolar sacs in the lung, the surface area for gas exchange is very large.

The mammalian circulatory system is a closed system with double circulation passing through the lungs and the body. It consists of a network of vessels containing blood that circulates because of pressure differences generated by the heart.

The heart contains two pumps that move blood through the pulmonary and systemic circulations. There is one atrium and one ventricle on the right side and one atrium and one ventricle on the left side. The pumping of the heart is a function of cardiomyocytes, distinctive muscle cells that are striated like skeletal muscle but pump rhythmically and involuntarily like smooth muscle. The signal for contraction begins in the wall of the right atrium. The electrochemical signal causes the two atria to contract in unison; then the signal causes the ventricles to contract. The blood from the heart is carried through the body by a complex network of blood vessels; arteries take blood away from the heart, and veins bring blood back to the heart.

Glossary

alveolus
: (plural: alveoli) (also, air sacs) the terminal structure of the lung passage where gas exchange occurs

aorta
: the major artery that takes blood away from the heart to the systemic circulatory system

artery
: a blood vessel that takes blood away from the heart

atrium
: (plural: atria) a chamber of the heart that receives blood from the veins

bicuspid valve
: a one-way opening between the atrium and the ventricle in the left side of the heart

bronchi
: (singular: bronchus) smaller branches of cartilaginous tissue that stem off of the trachea; air is funneled through the bronchi to the region where gas exchange occurs in the alveoli

bronchiole
: an airway that extends from the main bronchus to the alveolar sac

capillary
: the smallest blood vessel that allows the passage of individual blood cells and the site of diffusion of oxygen and nutrient exchange

cardiac cycle
: the filling and emptying the heart of blood caused by electrical signals that cause the heart muscles to contract and relax

closed circulatory system
: a system that has the blood separated from the bodily interstitial fluid and contained in blood vessels

diaphragm
: a skeletal muscle located under lungs that encloses the lungs in the thorax

diastole
: the relaxation phase of the cardiac cycle when the heart is relaxed and the ventricles are filling with blood

electrocardiogram (ECG)
: a recording of the electrical impulses of the cardiac muscle

inferior vena cava
: the major vein of the body returning blood from the lower parts of the body to the right atrium

larynx
: the voice box, located within the throat

nasal cavity
: an opening of the respiratory system to the outside environment

open circulatory system
: a circulatory system that has the blood mixed with interstitial fluid in the body cavity and directly bathes the organs

pharynx
: the throat

primary bronchus
: (also, main bronchus) a region of the airway within the lung that attaches to the trachea and bifurcates to form the bronchioles

pulmonary circulation
: the flow of blood away from the heart through the lungs where oxygenation occurs and then back to the heart

superior vena cava
: the major vein of the body returning blood from the upper part of the body to the right atrium

systemic circulation
: the flow of blood away from the heart to the brain, liver, kidneys, stomach, and other organs, the limbs, and the muscles of the body, and then back to the heart

systole
: the contraction phase of cardiac cycle when the ventricles are pumping blood into the arteries

trachea
: the cartilaginous tube that transports air from the throat to the lungs

tricuspid valve
: a one-way opening between the atrium and the ventricle in the right side of the heart

vein
: a blood vessel that brings blood back to the heart

ventricle

(of the heart) a large chamber of the heart that pumps blood into arteries

16.

2.3 Components of the Blood

The Role of Blood in the Body

Most of us have suffered a cut or a scraped knee and have seen our own blood. What exactly is blood, and what does it do? Blood, like the human blood illustrated in Figure 1 is important for regulation of the body's systems and homeostasis. Blood helps maintain the body's systems in working order by stabilizing pH, temperature, proper amounts of water, salts and nutrients, and by eliminating excess heat. Blood supports growth by distributing nutrients and hormones, and by removing waste. Blood plays a protective role by transporting clotting factors and platelets to prevent blood loss and transporting disease-fighting agents to sites of infection.

Blood is actually a term used to describe the liquid that moves through the vessels and includes plasma (the liquid portion, which contains water, proteins, salts, lipids, and glucose) and the different types of cells found in the plasma. The three main types of cells in human blood are red blood cells, also called erythrocytes, white blood cells, or leukocytes, and platelets, or thrombocytes. Each of these blood components plays specific roles in maintaining the health of the body.

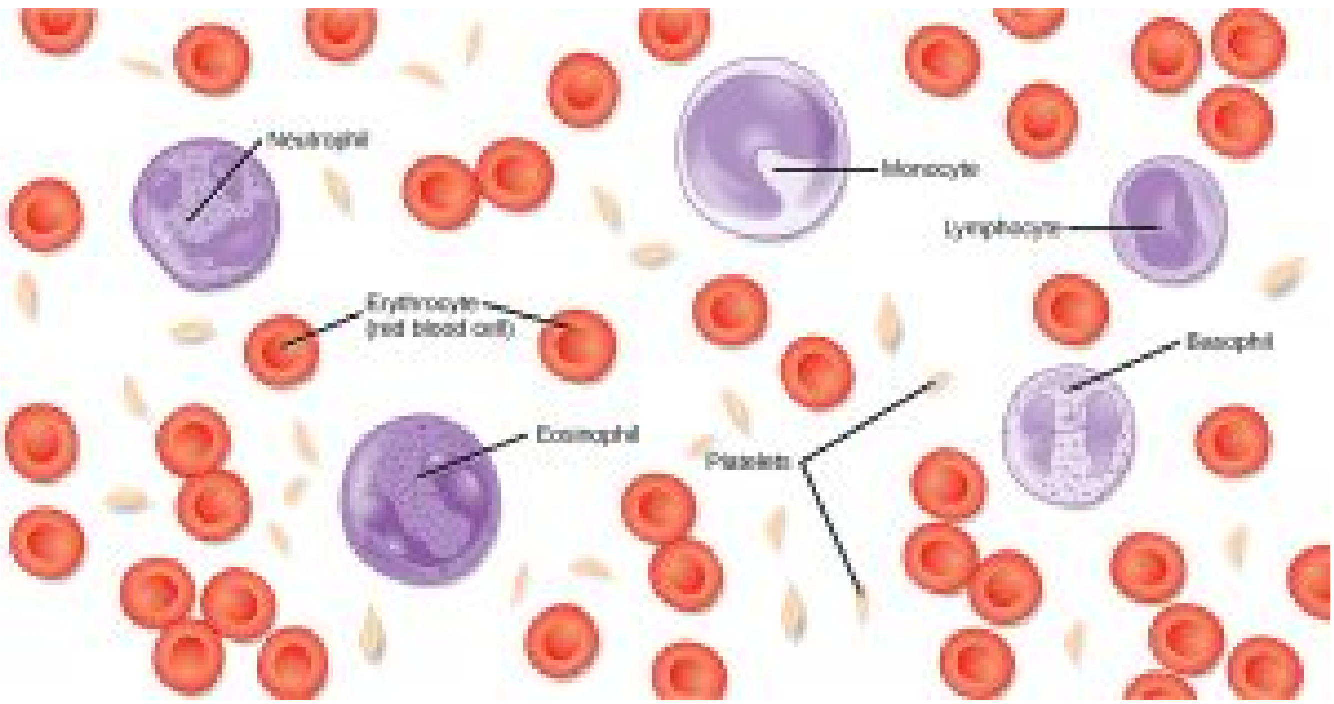


Figure 1. *The cells and cellular components of human blood are shown. Red blood cells deliver oxygen to the cells and remove carbon dioxide. White blood cells—including neutrophils, monocytes, lymphocytes, eosinophils, and basophils—are involved in the immune response. Platelets form clots that prevent blood loss after injury.*

Red Blood Cells

Red blood cells, or erythrocytes (erythro- = “red”; -cyte = “cell”), are specialized cells that circulate through the body delivering oxygen to cells; they are formed from stem cells in the bone marrow. In mammals, red blood cells are small biconcave cells that at maturity do not contain a nucleus or mitochondria and are only 7–8 µm in size.

The red coloring of blood comes from the iron-containing protein hemoglobin. The principal job of this protein is to carry oxygen, but it also transports carbon dioxide as well. Hemoglobin is packed into red blood cells at a rate of about 250 million molecules of hemoglobin per cell. Each hemoglobin molecule binds four oxygen molecules so that each red blood cell carries one billion molecules of oxygen. There are approximately 25 trillion red blood cells in the five liters of blood in the human body, which could carry up to 25 sextillion (25 × 1021) molecules of oxygen in the body at any time. In mammals, the lack of organelles in erythrocytes leaves more room for the hemoglobin molecules, and the lack of mitochondria also prevents use of the oxygen for metabolic respiration. Variants of hemoglobin help humans adapt to different environments. For example, Hgb-S causes sickle-cell anemia; although this variant of hemoglobin is not as efficient at transporting O2, it does provide some protection against malaria, thus providing an advantage to heterozygous individuals. Another variant is Hgb-F or fetal hemoglobin, which transports O2 efficiently in the low oxygen conditions found in the developing fetus. Red blood cells develop and mature in the bone marrow.

The small size and large surface area of red blood cells allows for rapid diffusion of oxygen and carbon dioxide across the plasma membrane. In the lungs, carbon dioxide is released and oxygen is taken in by the blood. In the tissues, oxygen is released from the blood and carbon dioxide is bound for transport back to the lungs.

A characteristic of red blood cells is their glycolipid and glycoprotein coating; these are lipids and proteins that have carbohydrate molecules attached. In humans, the surface glycoproteins and glycolipids on red blood cells vary between individuals, producing the different blood types, such as A, B, and O. Red blood cells have an average life span of 120 days, at which time they are broken down. We will take a deeper dive into blood typing, antigens, and antibodies when we explore the immune system in a later section, and we also will learn that white blood cells play important roles in immunity.

White Blood Cells

White blood cells, also called leukocytes (leuko = white), make up approximately one percent by volume of the cells in blood. The role of white blood cells is very different than that of red blood cells: they are primarily involved in the immune response to identify and target pathogens, such as invading bacteria, viruses, and other foreign organisms. White blood cells are formed continually; some only live for hours or days, but some live for years.

The morphology of white blood cells differs significantly from red blood cells. They have nuclei and do not contain hemoglobin. The different types of white blood cells are identified by their microscopic appearance after histologic staining, and each has a different specialized function. The two main groups, both illustrated in Figure 2 are the granulocytes, which include the neutrophils, eosinophils, and basophils, and the agranulocytes, which include the monocytes and lymphocytes.

Figure 2. (a) *Granulocytes—including neutrophils, eosinophils and basophils—are characterized by a lobed nucleus and granular inclusions in the cytoplasm. Granulocytes are typically first-responders during injury or infection.* ***(b)*** *Agranulocytes include lymphocytes and monocytes. Lymphocytes, including B and T cells, are responsible for adaptive immune response. Monocytes differentiate into macrophages and dendritic cells, which in turn respond to infection or injury.*

Granulocytes contain granules in their cytoplasm; the agranulocytes are so named because of the lack of granules in their cytoplasm. While they are named based on their appearances, each type of white blood cell also has unique functions. We will discuss the different types of leukocytes and their functions in a later section of this book.

Platelets and Coagulation Factors

Blood must clot to heal wounds and prevent excess blood loss. Small cell fragments called platelets (thrombocytes) are attracted to the wound site where they adhere by extending many projections and releasing their contents. These contents activate other platelets and also interact with other coagulation factors, which convert fibrinogen, a water-soluble protein present in blood serum into fibrin (a non-water soluble protein), causing the blood to clot. Many of the clotting factors require vitamin K to work, and vitamin K deficiency can lead to problems with blood clotting. Many platelets converge and stick together at the wound site forming a platelet plug (also called a fibrin clot), as illustrated in Figure 3b. The plug or clot lasts for a number of days and stops the loss of blood. Platelets are formed from the disintegration of larger cells called megakaryocytes, like that shown in Figure 3a. For each megakaryocyte, 2000–3000 platelets are formed with 150,000 to 400,000 platelets present in each cubic millimeter of blood. Each platelet is disc shaped and 2–4 μm in diameter. They contain many small vesicles but do not contain a nucleus.

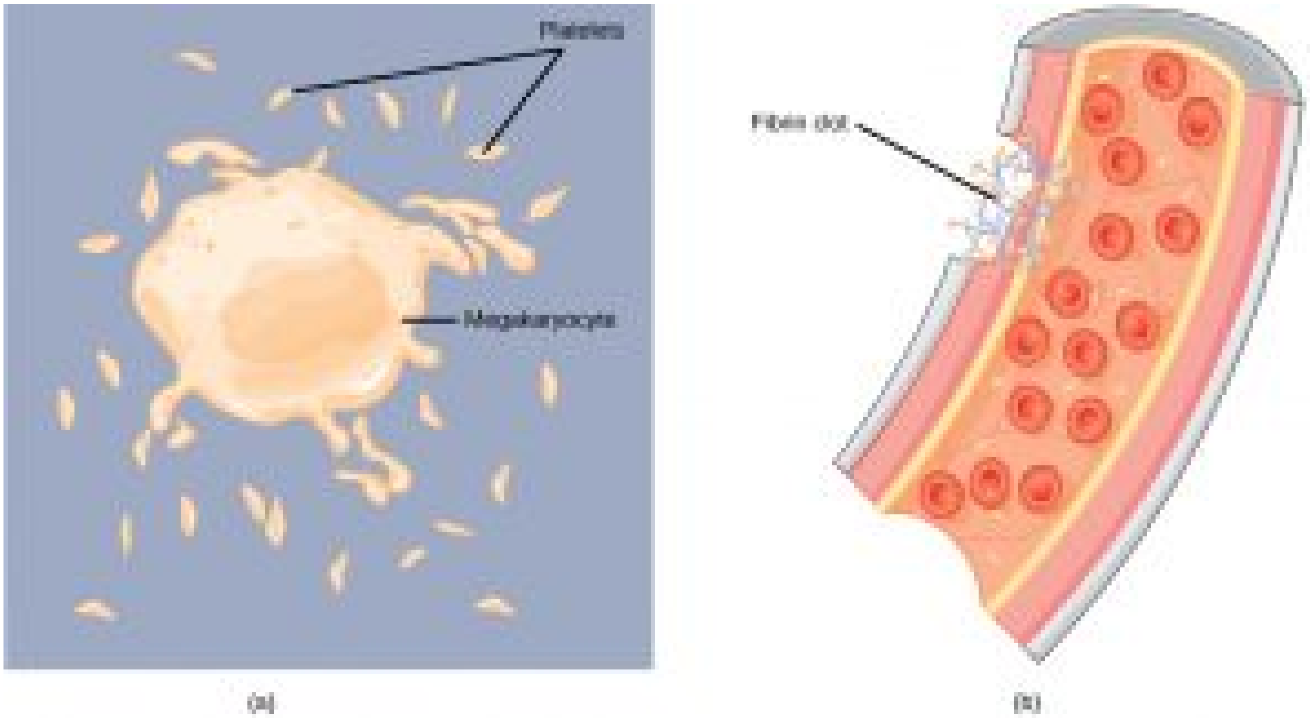


Figure 3. ***(a)*** *Platelets are formed from large cells called megakaryocytes. The megakaryocyte breaks up into thousands of fragments that become platelets.* ***(b)*** *Platelets are required for clotting of the blood. The platelets collect at a wound site in conjunction with other clotting factors, such as fibrinogen, to form a fibrin clot that prevents blood loss and allows the wound to heal.*

Plasma and Serum

The liquid component of blood is called plasma, and it is separated by spinning or centrifuging the blood at high rotations (3000 rpm or higher). The blood cells and platelets are separated by centrifugal forces to the bottom of a specimen tube. The upper liquid layer, the plasma, consists of 90 percent water along with dissolved substances, including the coagulation factors mentioned above.

The plasma component of blood without the coagulation factors is called the serum. Serum is similar to the fluid surrounding the cells in other tissues ,and contains precise concentrations of important ions like calcium and sodium necessary for cellular functions. Other components in the serum include proteins that assist with maintaining pH and water balance while giving viscosity to the blood. The serum also contains antibodies, specialized proteins that are important for defense against viruses and bacteria. Lipids, including cholesterol, are also transported in the serum, along with various other substances including nutrients, hormones, metabolic waste, plus external substances, such as, drugs, viruses, and bacteria.

Human serum albumin is the most abundant protein in human blood plasma and is synthesized in the liver. Albumin, which constitutes about half of the blood serum protein, transports hormones and fatty acids, buffers pH, and maintains water balance in the body.

EVOLUTION CONNECTION

Blood Types Related to Proteins on the Surface of the Red Blood Cells

Red blood cells are coated in antigens made of glycolipids and glycoproteins. The composition of these molecules is determined by genetics, which have evolved over time. In humans, the different surface antigens are grouped into 24 different blood groups with more than 100 different antigens on each red blood cell. The two most well known blood groups are the ABO, shown in ***Figure 4***, and Rh systems. The surface antigens in the ABO blood group are glycolipids, called antigen A and antigen B. People with blood type A have antigen A, those with blood type B have antigen B, those with blood type AB have both antigens, and people with blood type O have neither antigen. Antibodies called agglutinogens are found in the blood plasma and react with the A or B antigens, if the two are mixed. When type A and type B blood are combined, agglutination (clumping) of the blood occurs because of antibodies in the plasma that bind with the opposing antigen; this causes clots that coagulate in the kidney causing kidney failure. Type O blood has neither A or B antigens, and therefore, type O blood can be given to all blood types. Type O negative blood is the universal donor. Type AB positive blood is the universal acceptor because it has both A and B antigen. The ABO blood groups were discovered in 1900 and 1901 by Karl Landsteiner at the University of Vienna.

The Rh blood group was first discovered in Rhesus monkeys. Most people have the Rh antigen (Rh+) and do not have anti-Rh antibodies in their blood. The few people who do not have the Rh antigen and are Rh– can develop anti-Rh antibodies if exposed to Rh+ blood. This can happen after a blood transfusion or after an Rh– woman has an Rh+ baby. The first exposure does not usually cause a reaction; however, at the second exposure, enough antibodies have built up in the blood to produce a reaction that causes agglutination and breakdown of red blood cells. An injection can prevent this reaction.

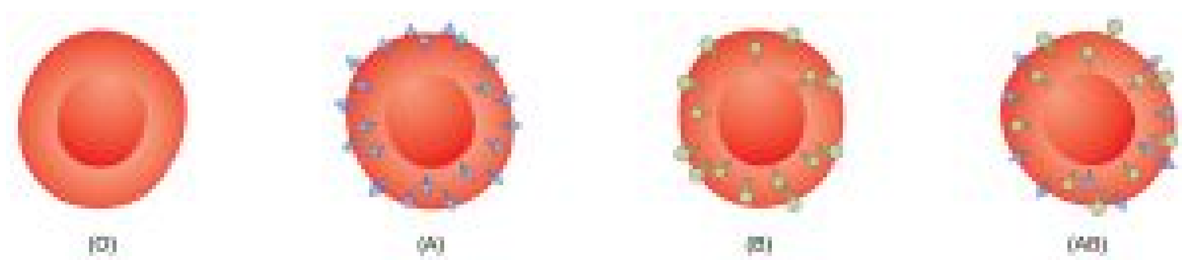


Figure 4. *Human red blood cells may have either type A or B glycoproteins on their surface, both glycoproteins combined (AB), or neither (O). The glycoproteins serve as antigens and can elicit an immune response in a person who receives a transfusion containing unfamiliar antigens. Type O blood, which has no A or B antigens, does not elicit an immune response when injected into a person of any blood type. Thus, O is considered the universal donor. Persons with type AB blood can accept blood from any blood type, and type AB is considered the universal acceptor.*

Link to Learning

Play a blood typing game on the Nobel Prize website to solidify your understanding of blood types.

Section Summary

Specific components of the blood include red blood cells, white blood cells, platelets, and the plasma, which contains coagulation factors and serum. Blood is important for regulation of the body's pH, temperature, osmotic pressure, the circulation of nutrients and removal of waste, the distribution of hormones from endocrine glands, and the elimination of excess heat; it also contains components for blood clotting. Red blood cells are specialized cells that contain hemoglobin and circulate through the body delivering oxygen to cells. White blood cells are involved in the immune response to identify and target invading bacteria, viruses, and other foreign organisms; they also recycle waste components, such as old red blood cells. Platelets and blood clotting factors cause the change of the soluble protein fibrinogen to the insoluble protein fibrin at a wound site forming a plug. Plasma consists of 90 percent water along with various substances, such as coagulation factors and antibodies. The serum is the plasma component of the blood without the coagulation factors.

Glossary

plasma
: liquid component of blood that is left after the cells are removed

platelet
: (also, thrombocyte) small cellular fragment that collects at wounds, cross-reacts with clotting factors, and forms a plug to prevent blood loss

red blood cell
: small (7–8 µm) biconcave cell without mitochondria (and in mammals without nuclei) that is packed with hemoglobin, giving the cell its red color; transports oxygen through the body

serum
: plasma without the coagulation factors

white blood cell
: large (30 µm) cell with nuclei of which there are many types with different roles including the protection of the body from viruses and bacteria, and cleaning up dead cells and other waste

17.

2.4 Innate Immunity

The immune system in vertebrates, including humans, is a complex multilayered system for defending against external and internal threats to the integrity of the body. Many of these threats are caused by infectious microbes, called **pathogens**, which include viruses, bacteria, fungi, and parasites. The system can be divided into two types of defense systems: the innate immune system, which is nonspecific toward a particular kind of pathogen, and the adaptive immune system, which is specific (***Figure 1***). **Innate immunity** is not caused by an infection or vaccination and depends initially on physical and chemical barriers that work on all pathogens, sometimes called the first line of defense. The second line of defense of the innate system includes chemical signals that produce inflammation and fever responses as well as mobilizing protective cells and other chemical defenses. Together, the two parts of the innate immune response to a pathogen produce a fast response that is the same every time a pathogen enters the body. This response does not recognize a specific pathogen, but rather recognizes broad patterns of molecules that are not found in healthy human cells or tissues. In contrast, the adaptive immune system mounts a highly specific response to substances and organisms that do not belong in the body, recognizing unique molecular "signatures" found on each pathogen. The adaptive system takes longer to respond and has a memory system that allows it to respond with greater intensity should the body reencounter a the same pathogen again, even years later.

Vertebrate Immunity		
Innate Immune System		Adaptive Immune System
Physical Barriers	Internal Defenses	
• Skin, hair, cilia • Mucus membranes • Mucus and chemical secretions • Digestive enzymes in mouth • Stomach acid	• Inflammatory response • Complement proteins • Phagocytic cells • Natural killer (NK) cells	• Antibodies and the humoral immune response • Cell-mediated immune response • Memory response

Figure 1. *There are two main parts to the vertebrate immune system. The innate immune system, which is made up of physical barriers and internal defenses, responds to all pathogens. The adaptive immune system is highly specific.*

External and Chemical Barriers

The body has significant physical barriers to potential pathogens. The skin contains the protein keratin, which resists physical entry into cells. Other body surfaces, particularly those associated with body openings, are protected by the mucous membranes. The sticky mucus provides a physical trap for pathogens, preventing their movement deeper into the body. The openings of the body, such as the nose and ears, are protected by hairs that catch pathogens, and the mucous membranes of the upper respiratory tract have cilia that constantly move pathogens trapped in the mucus coat up to the mouth.

The skin and mucous membranes also create a chemical environment that is hostile to many microorganisms. The surface of the skin is acidic, which prevents bacterial growth. Saliva, mucus, and the tears of the eye contain an enzyme that breaks down bacterial cell walls. The stomach secretions create a highly acidic environment, which kills many pathogens entering the digestive system.

Finally, the surface of the body and the lower digestive system have a community of microorganisms such as bacteria, archaea, and fungi that coexist without harming the body. There is evidence that these organisms are highly beneficial to their host, combating disease-causing organisms and outcompeting them for nutritional resources provided by the host body. Despite these defenses, pathogens may enter the body through skin abrasions or punctures, or by collecting on mucosal surfaces in large numbers that overcome the protections of mucus or cilia.

Internal Defenses

When pathogens enter the body, the innate immune system responds with a variety of internal defenses. These include the inflammatory response, phagocytosis, natural killer cells, and the complement system. White blood cells in the blood and lymph recognize pathogens as foreign to the body, and carry out these defense responses, often with the help of chemical signaling molecules they release. A **white blood cell,** also called a **leukocyte,** is larger than a red blood cell, is nucleated, and is typically able to move using amoeboid locomotion. Because they can move on their own, white blood cells can leave the blood to go to infected tissues. For example, a **monocyte** is a type of white blood cell that circulates in the blood and lymph and develops into a macrophage after it moves into infected tissue. A **macrophage** is a large cell that engulfs foreign particles and pathogens. **Mast cells** are produced in the same way as white blood cells, but unlike circulating white blood cells, mast cells take up residence in the tissues other than blood. They are responsible for releasing chemicals in response to physical injury.

When a pathogen is recognized by white blood cells, chemicals called cytokines are released. A **cytokine** is a chemical messenger that regulates many different cellular processes, including cell division and gene expression, to produce a variety of immune responses. Approximately 40 types of cytokines exist in humans. In addition to being released from white blood cells after pathogen recognition, cytokines are also released by the infected cells and bind to nearby uninfected cells, inducing those cells to release cytokines. This positive feedback loop results in a burst of cytokine production.

One class of early-acting cytokines is the interferons, which are released by infected cells as a warning to nearby uninfected cells. An **interferon** is a small protein that signals a viral infection to other cells. The interferons stimulate uninfected cells to produce compounds that interfere with viral replication. Interferons also activate macrophages and other cells.

The Inflammatory Response and Phagocytosis

The first cytokines to be produced encourage **inflammation**, a localized redness, swelling, heat, and pain. Inflammation is a response to physical trauma, such as a cut or a blow, chemical irritation, and infection by pathogens (viruses, bacteria, or fungi). The chemical signals that trigger an inflammatory response enter the extracellular fluid and cause capillaries to dilate (expand) and capillary walls to become more permeable, or leaky. The serum and other compounds leaking from capillaries cause swelling of the area, which in turn causes pain. Various kinds of white blood cells are attracted to the area of inflammation. The types of white blood cells that arrive at an inflamed site depend on the nature of the injury or infecting pathogen. For example, a **neutrophil** is an early arriving white blood cell that engulfs and digests pathogens. Neutrophils are the most abundant white blood cells of the immune system (***Figure 2***). Macrophages follow neutrophils and take over the phagocytosis function and are involved in the resolution of an inflamed site, cleaning up cell debris and pathogens. The process of ingesting and digesting up pathogens is called phagocytosis, and the cells that carry this process out are called phagocytes.

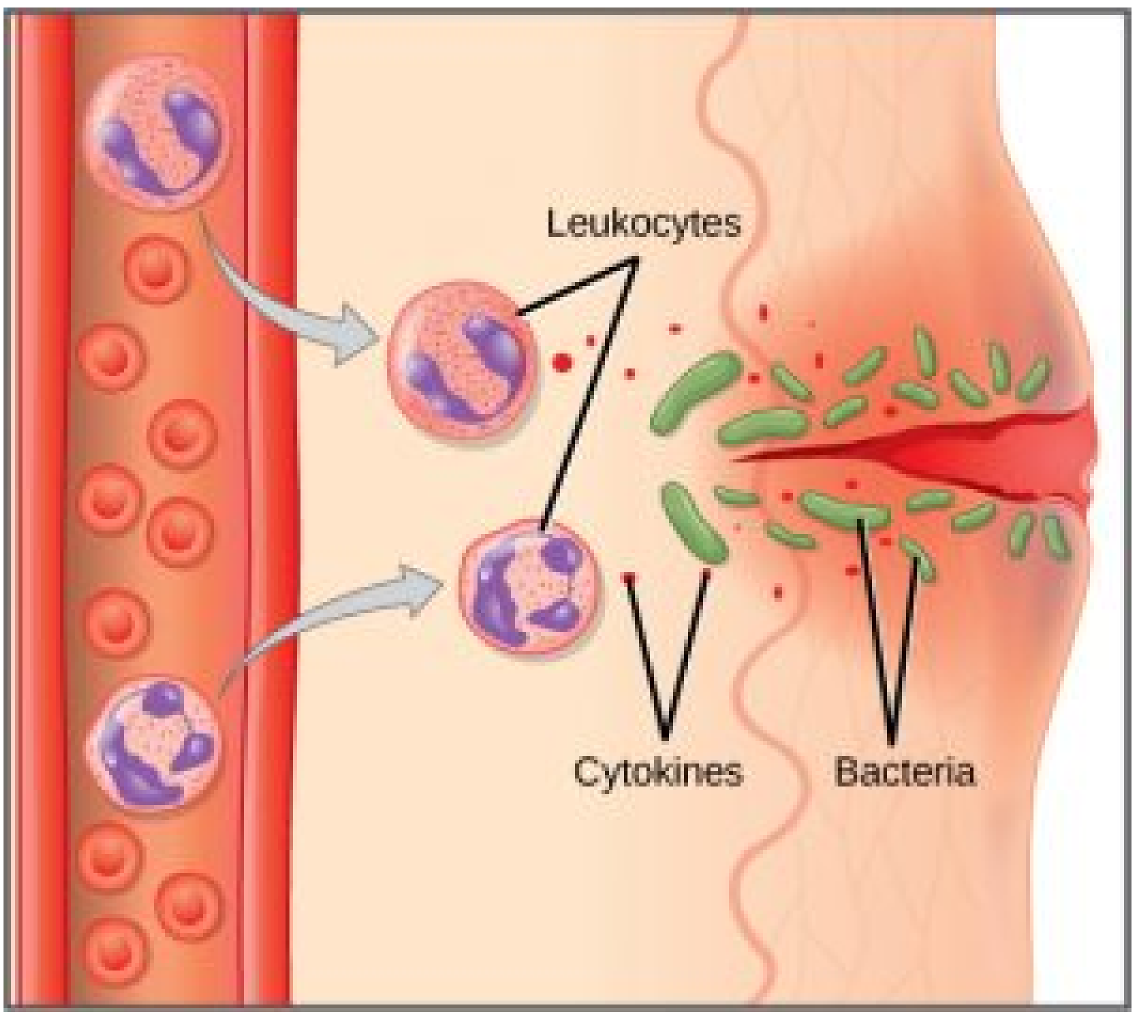


Figure 2. *White blood cells (leukocytes) release chemicals to stimulate the inflammatory response following a cut in the skin.*

Cytokines also send feedback to cells of the nervous system to bring about the overall symptoms of feeling sick, which include lethargy, muscle pain, and nausea. Cytokines also increase the core body temperature, causing a fever. The elevated temperatures of a fever inhibit the growth of pathogens and speed up cellular repair processes. For these reasons, suppression of fevers should be limited to those that are dangerously high.

Concept in Action

Check out this 23-second, stop-motion video showing a neutrophil that searches and engulfs fungus spores during an elapsed time of 79 minutes.

Natural Killer Cells

A **lymphocyte** is a white blood cell that contains a large nucleus (***Figure 3***). Most lymphocytes are associated with the adaptive immune response, but infected cells are identified and destroyed by natural killer cells, the only lymphocytes of the innate immune system. A **natural killer (NK) cell** is a lymphocyte that can kill cells infected with viruses (or cancerous cells). NK cells identify intracellular infections, especially from viruses, by the altered expression of particular types of proteins on the surface of infected cells. Unhealthy cells, whether infected or cancerous, display an altered set of proteins on their cell surfaces.

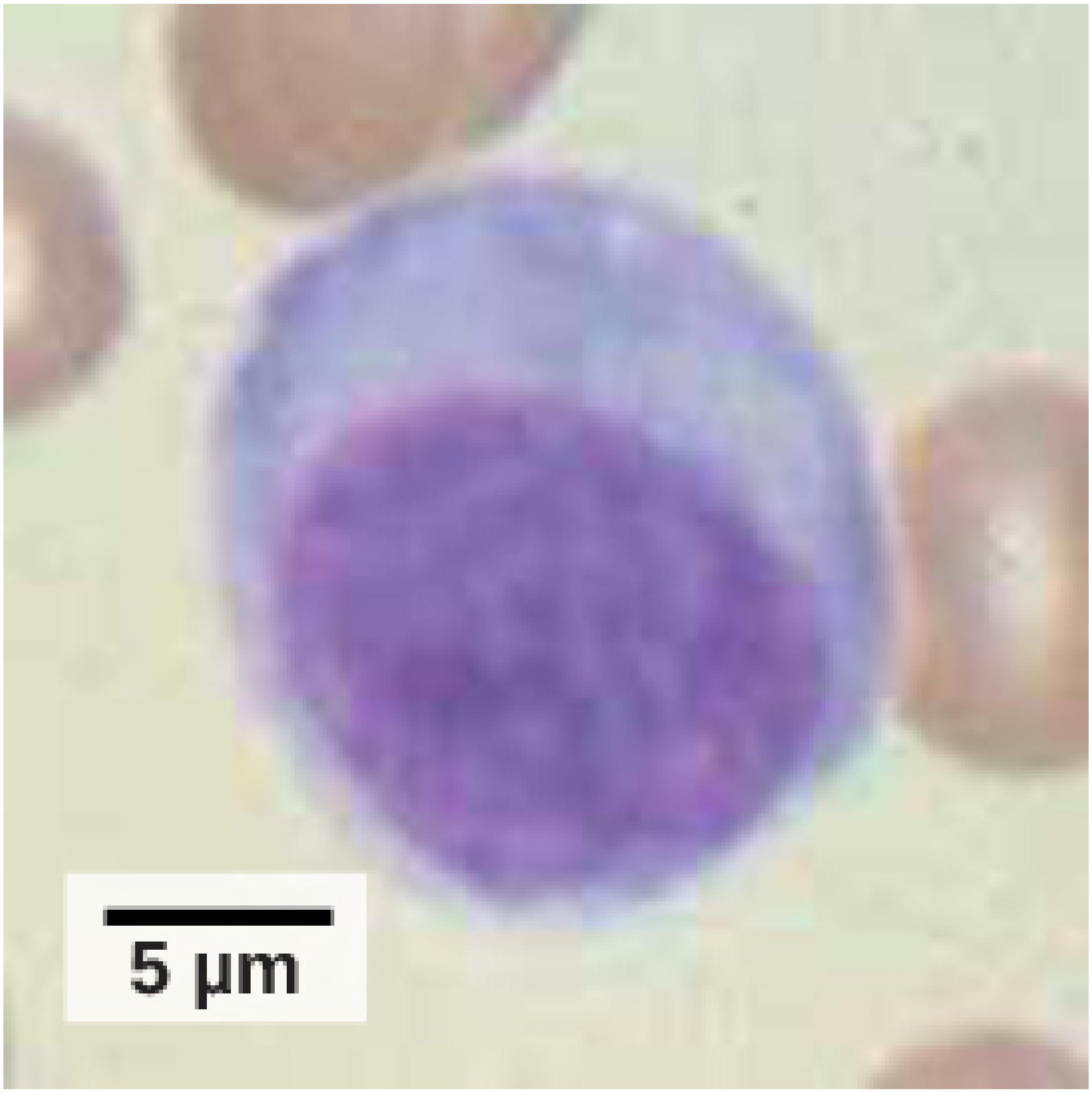


Figure 3. *Lymphocytes, such as NK cells, are characterized by their large nuclei that actively absorb Wright stain and therefore appear dark colored under a microscope. (credit: scale-bar data from Matt Russell)*

After the NK cell detects an infected or tumor cell, it induces programmed cell death, or apoptosis. Phagocytic cells then come along and digest the cell debris left behind. NK cells are constantly patrolling the body and are an effective mechanism for controlling potential infections and preventing cancer progression. The various types of immune cells are shown in ***Figure 4***.

Figure 4. *Cells involved in the innate immune response include mast cells, natural killer cells, and white blood cells, such as monocytes, macrophages and neutrophils.*

Complement

An array of approximately 20 types of proteins, called a **complement system**, is also activated by infection or the activity of the cells of the adaptive immune system and functions to destroy extracellular pathogens. Liver cells and macrophages synthesize inactive forms of complement proteins continuously; these proteins are abundant in the blood serum and are capable of responding immediately to infecting microorganisms. The complement system is so named because it is complementary to the innate and adaptive immune system. Complement proteins bind to the surfaces of microorganisms and are particularly attracted to pathogens that are already tagged by the adaptive immune system. This “tagging” involves the attachment of specific proteins called antibodies (discussed in detail later) to the pathogen. When they attach, the antibodies change shape providing a binding site for one of the complement proteins. After the first few complement proteins bind, a cascade of binding in a specific sequence of proteins follows in which the pathogen rapidly becomes coated in complement proteins.

Complement proteins perform several functions, one of which is to serve as a marker to indicate the presence of a pathogen to phagocytic cells and enhance engulfment. Certain complement proteins can combine to open pores in microbial cell membranes and cause lysis of the cells.

Section Summary

The innate immune system consists first of physical and chemical barriers to infection including the skin and mucous membranes and their secretions, ciliated surfaces, and body hairs. The second line of defense is an internal defense system designed to counter pathogenic threats that bypass the physical and chemical barriers of the body. Using a combination of cellular and molecular responses, the innate immune system identifies the nature of a pathogen and responds with inflammation, phagocytosis, cytokine release, destruction by NK cells, or the complement system.

Glossary

complement system
: an array of approximately 20 soluble proteins of the innate immune system that enhance phagocytosis, bore holes in pathogens, and recruit lymphocytes

cytokine
: a chemical messenger that regulates cell differentiation, proliferation, and gene expression to effect immune responses

inflammation
: the localized redness, swelling, heat, and pain that results from the movement of leukocytes through opened capillaries to a site of infection

innate immunity
: an immunity that occurs naturally because of genetic factors or physiology, and is not caused by infection or vaccination

interferon
: a cytokine that inhibits viral replication

lymphocyte
: a type of white blood cell that includes natural killer cells of the innate immune system and B and T cells of the adaptive immune system

macrophage
: a large phagocytic cell that engulfs foreign particles and pathogens

mast cell
: a leukocyte that produces inflammatory molecules, such as histamine, in response to large pathogens

monocyte
: a type of white blood cell that circulates in the blood and lymph and differentiates into a macrophage after it moves into infected tissue

natural killer (NK) cell
: a lymphocyte that can kill cells infected with viruses or tumor cells

neutrophil
: a phagocytic leukocyte that engulfs and digests pathogens

phagocytosis
: the process by which a cell engulfs and digests a small object, such as a viral particle, bacterium or piece of cellular debris. Cells that can carry out phagocytosis are called phagocytes.

pathogen a microbe (virus, bacteria, fungus, parasite) capable of causing disease

white blood cell
: a nucleated cell found in the blood that is a part of the immune system; also called leukocytes

18.

2.5 Adaptive Immunity

Antigens and the Adaptive Immune Response

Adaptive immunity occurs after exposure to an antigen either from a pathogen or a vaccination.The adaptive, or acquired, immune response takes days or even weeks to become established—much longer than the innate response; however, adaptive immunity is more specific to an invading pathogen. This part of the immune system works in tandem with the innate immune response to neutralize pathogens. In fact, without information from the innate immune system, the adaptive response could not be mobilized.

An **antigen** is a small, specific molecule on a particular pathogen that stimulates a response in the immune system. One example of an antigen is a specific sequence of 8 amino acids in a protein found only in an influenza virus, the virus responsible for causing "the flu.". Another example is a short chain of carbohydrates found on the cell wall of *Neisseria meningitidis*, the bacteria that causes meningitis. There are millions of potential sequences of amino acids, carbohydrates, and other small molecules that can act as antigens.   The adaptive immune system works because the immune cells responsible for it are each able to recognize and respond to one specific antigen, or a few very similar ones.

The adaptive immune responses depends on the function of two types of lymphocytes, called **B cells** and **T cells**. In adaptive immunity, activated T and B cells whose surface binding sites are specific to the antigen molecules on a pathogen greatly increase in numbers and attack the invading pathogen. Their attack can kill pathogens directly or they can secrete antibodies that enhance the phagocytosis of pathogens and disrupt the infection. Adaptive immunity also involves a memory to give the host long-term protection from reinfection with the same type of pathogen carrying the same antigens; on reexposure, this host memory will facilitate a rapid and powerful response.

B and T Cells

Lymphocytes, which are a subclass of white blood cells, are formed with other blood cells in the red bone marrow found in many flat bones, such as the shoulder or pelvic bones. The two types of lymphocytes of the adaptive immune response are B and T cells (***Figure 1***). Whether an immature lymphocyte becomes a B cell or T cell depends on where in the body it matures. The B cells remain in the bone marrow to mature (hence the name "B" for "bone marrow"), while T cells migrate to the thymus, where they mature (hence the name "T" for "thymus"). During the maturation process, each B or T cell develops unique surface proteins that are able to recognize a unique set of very specific molecules on antigens (discussed below). In other words, each B or T cell can recognize only a very few different molecules, but together the entire lymphocyte population in a healthy person should be able to recognize molecules from most pathogens. The specificity of these unique surface proteins, or receptors, on the lymphocytes is determined by the genetics of the individual and is present before a foreign molecule is introduced to the body or encountered. Except in certain immune system diseases called autoimmune diseases, no mature B or T cells are able to recognize and bind to molecules that are found on healthy human cells, but only to molecules found on pathogens or on unhealthy human cells.

B cells are involved in the **humoral immune response**, which targets pathogens loose in blood and lymph, and B cells carry out this response by secreting antibodies.T cells are involved in the **cell-mediated immune response**, which targets infected cells in the body. T cells include the Helper T cells and the Cytotoxic, or Killer, T cells. Cytotoxic T cells directly kill human cells that are infected or unhealthy. Helper T cells do not directly kill infected cells, but secrete molecules that are crucial for the function of all other cells in the immune response to a pathogen.

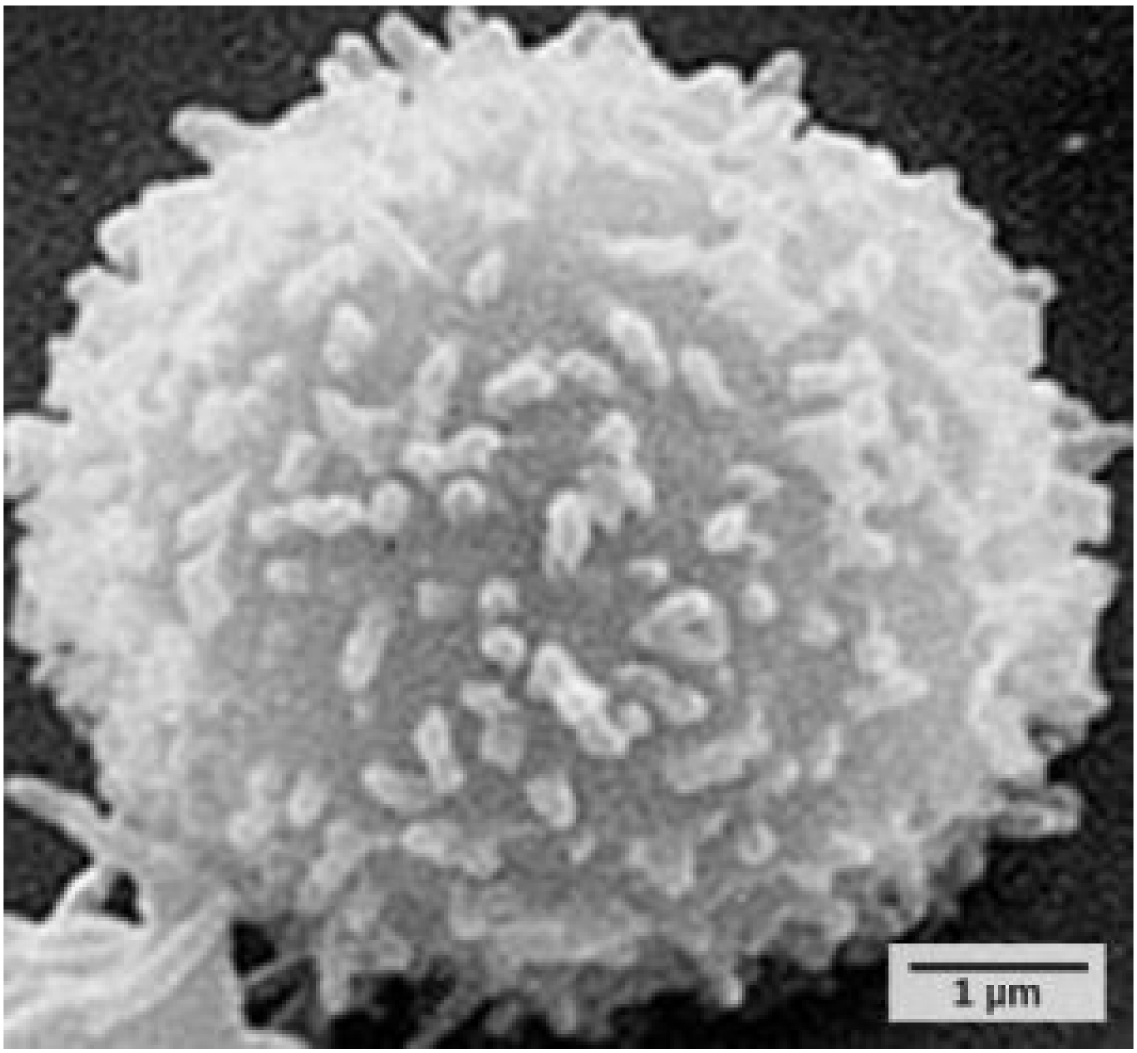


Figure 1. *This scanning electron micrograph shows a T lymphocyte. T and B cells are indistinguishable by light microscopy but can be differentiated experimentally by probing their surface receptors. (credit: modification of work by NCI; scale-bar data from Matt Russell)*

Four Stages of the Adaptive Immune Response

B cells, Helper T cells, and Cytotoxic T cells all respond to antigens in a similar pattern; subsequent sections of this chapter will address the specifics of the immune response in each cell type. In order for their immune functions to be elicited, the cells must first *encounter* antigens by binding specifically to them using specialized membrane proteins. This binding elicits changes in the activity of the immune cells, termed *activation*, which is the second step in the adaptive immune response. Activation responses vary between the three types of cells, but in general all involve both changes in gene expression and in the initiation of cell division. Third, the immune cells *attack* invading pathogens or infected cells. Depending on the type of lymphocyte, the specific methods used to neutralize pathogens can vary. In most infections, the attacks from many different lymphocytes, and from cells of the innate immune system, occur simultaneously and the atttacking cells often stimulate each other through chemical messenger such as cytokines. Finally, long-lived, pre-activated immune cells that wait for a subsequent infection are formed in the *memory* phase of the adaptive immune response. These cells are identical to the initial cells that first encountered the pathogen except that they have already undergone the activation step so are able to attack right away when activated again by an antigen.

Humoral Immune Response: B cells

As mentioned, an antigen is a molecule that stimulates a response in the immune system. B cells participate in a chemical response to antigens present in the body by producing specific **antibodies** that circulate throughout the body and bind with an antigen whenever it is encountered. This is known as the humoral immune response because it involves molecules secreted into the blood plasma, an acellular fluid or "humor." As discussed, during maturation of B cells, a set of highly specific B cells are produced, each with antigen receptor molecules in their membrane (***Figure 2***).

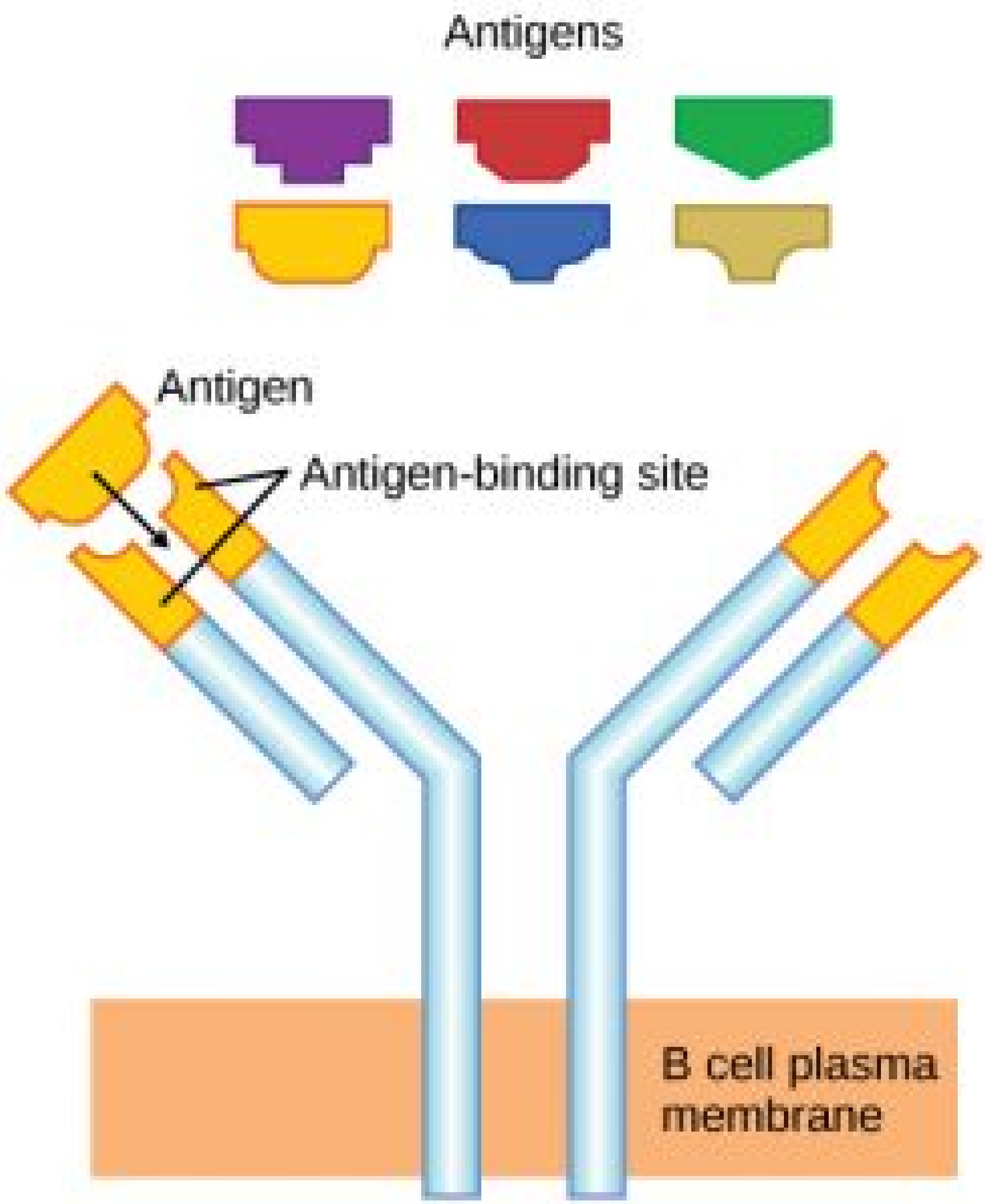


Figure 2. *B cell receptors are embedded in the membranes of B cells and bind a variety of antigens through their variable regions.*

Each B cell has only one kind of antigen receptor, which makes every B cell different. Once the B cells mature in the bone marrow, they migrate to lymph nodes or other lymphatic organs, where they wait to encounter potential antigens. When a B cell encounters the antigen that binds to its receptor, the antigen molecule is brought into the cell, where it is processed by the cell, and reappears on the surface of the cell bound to B cell protein. When this process is complete, the B cell is considered to be activated.

Activation induces the B cell to divide rapidly, which makes thousands of identical (clonal) cells.

These cells become either plasma cells or memory B cells. The memory B cells remain inactive at this point, until another later encounter with the antigen, caused by a reinfection by the same bacteria or virus, results in them dividing into a new population of plasma cells to carry out the memory phase of the B cell response. The plasma cells, on the other hand, produce and secrete large quantities, up to 100 million molecules per hour, of antibody molecules. An **antibody**, also known as an immunoglobulin (Ig), is a protein that is produced by plasma cells after stimulation by an antigen. Antibodies are the agents of humoral immunity; they are the weapons the B cells use in their attacks on pathogens. Antibodies occur in the blood, in gastric and mucus secretions, and in breast milk. Antibodies in these bodily fluids can bind pathogens and mark them for destruction by phagocytes before they can infect cells.

These antibodies circulate in the blood stream and lymphatic system and bind with the antigen whenever it is encountered. The binding can fight infection in several ways. Antibodies can bind to viruses or bacteria and interfere with the chemical interactions required for them to infect or bind to other cells. The antibodies may create bridges between different particles containing antigenic sites clumping them all together and preventing their proper functioning. The antigen-antibody complex stimulates the complement system described previously, destroying the cell bearing the antigen. Phagocytic cells of the innate immune system are attracted by the antigen-antibody complexes, and phagocytosis is enhanced when the complexes are present. Finally, antibodies stimulate inflammation, and their presence in mucus and on the skin prevents pathogen attack.

The production of antibodies by plasma cells in response to an antigen is called **active immunity** and describes the host's active response of the immune system to an infection or to a vaccination. There is also a **passive immune response** where antibodies come from an outside source, instead of the individual's own plasma cells, and are introduced into the host. For example, antibodies circulating in a pregnant female's body move across the placenta into the developing fetus. Antibodies are also passed to the child in breast milk. The child benefits from the presence of these antibodies for up to several months after birth. In addition, a passive immune response is possible by injecting antibodies into an individual in the form of an antivenom to a snake-bite toxin or antibodies in blood serum to help fight a hepatitis infection. This gives immediate protection since the body does not need the time required to mount its own response.

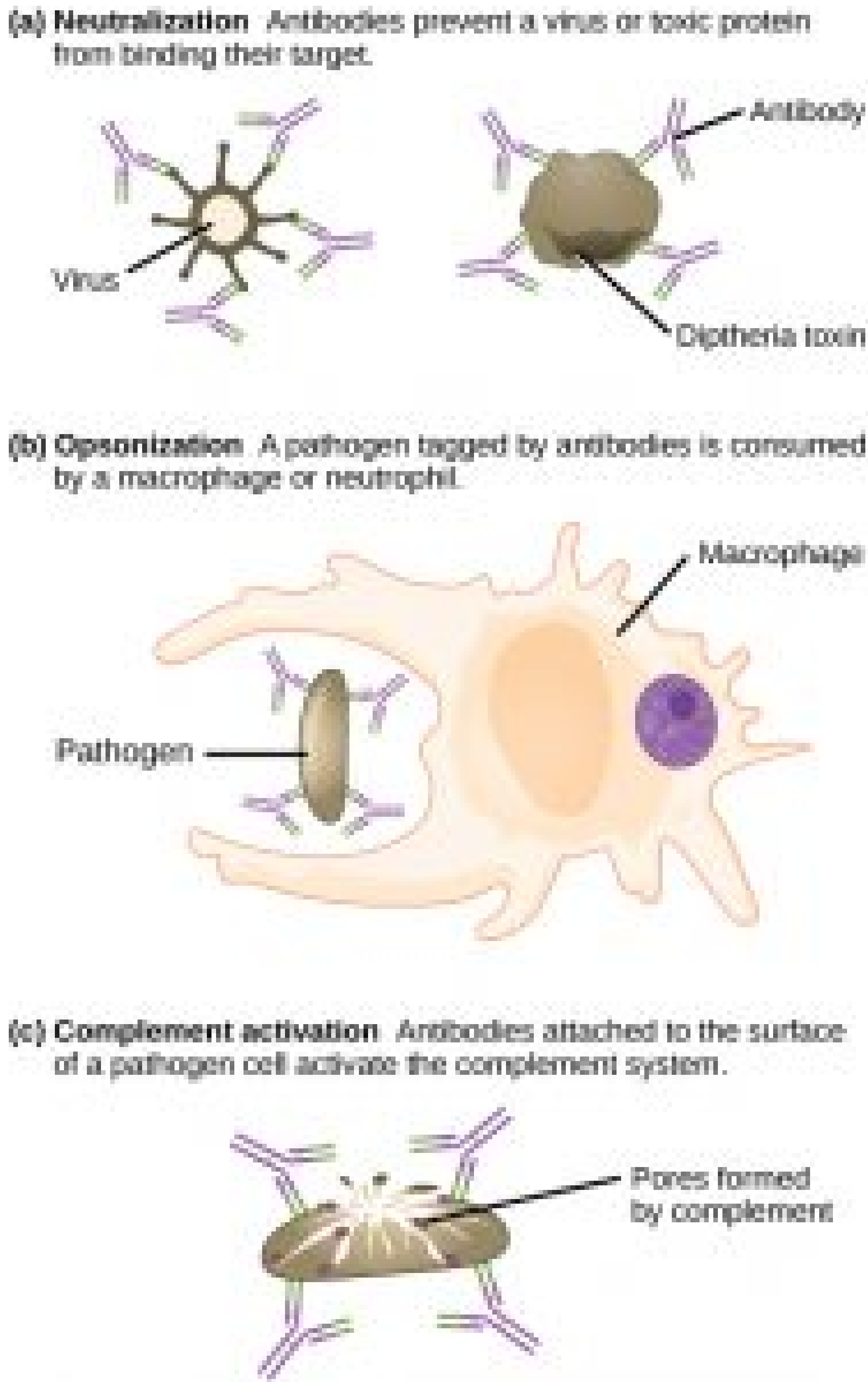

(a) Neutralization Antibodies prevent a virus or toxic protein from binding their target.
Antibody
Virus
Diptheria toxin
(b) Opsonization A pathogen tagged by antibodies is consumed by a macrophage or neutrophil.
Macrophage
Pathogen
(c) Complement activation Antibodies attached to the surface of a pathogen cell activate the complement system.
Pores formed by complement

Figure 3. *Antibodies may inhibit infection by (a) preventing the antigen from binding its target, (b) tagging a pathogen for destruction by macrophages or neutrophils, or (c) activating the complement cascade.*

Cell-Mediated Immunity: T cells

Activation of T cells also begins when T cells encounter antigens and bind to them with specific proteins on their cell surfaces, called T cell receptors. Each T cell's receptor proteins are able to bind to only one or a few very similar antigens, allowing each one to respond to different pathogens. Unlike B cells, T lymphocytes are unable to recognize pathogens without assistance. They rely on cells of the innate immune system to help them encounter antigens, and to start the responses required for an immune attack, in a complex process that is beyond the scope of this textbook. T cells that have encountered antigens that bind to their specific receptors in this process, and received important signals from the innate immune cells are considered to be "activated.

There are two main types of T cells: Helper T cells (sometimes called T_H) and Cytotoxic T cells (T_C, also known as killer T cells.) The Helper T cells function indirectly to tell other immune cells about potential pathogens. After encountering an antigen, a Helper T cell's activation causes many rounds of cell division to occur, producing thousands of identical cells that all have the same T cell receptor and all have initiated the expression of genes required for producing and secreting cytokines. As with B cells, the clone includes active T_H cells and inactive memory T_H cells. Helper T cells carry out their attacks on pathogens through the secretion of these cytokines. These chemical messengers greatly enhance the activities of macrophages, innate immune cells and Cytotoxic T cells, and also stimulate naïve B cells to secrete antibodies. A summary of how the humoral and cell-mediated immune responses are activated appears in ***Figure 4***. Similar to memory B cells, memory Helper T cells are long-lived, and can carry out a fast response to the same pathogen after decades.

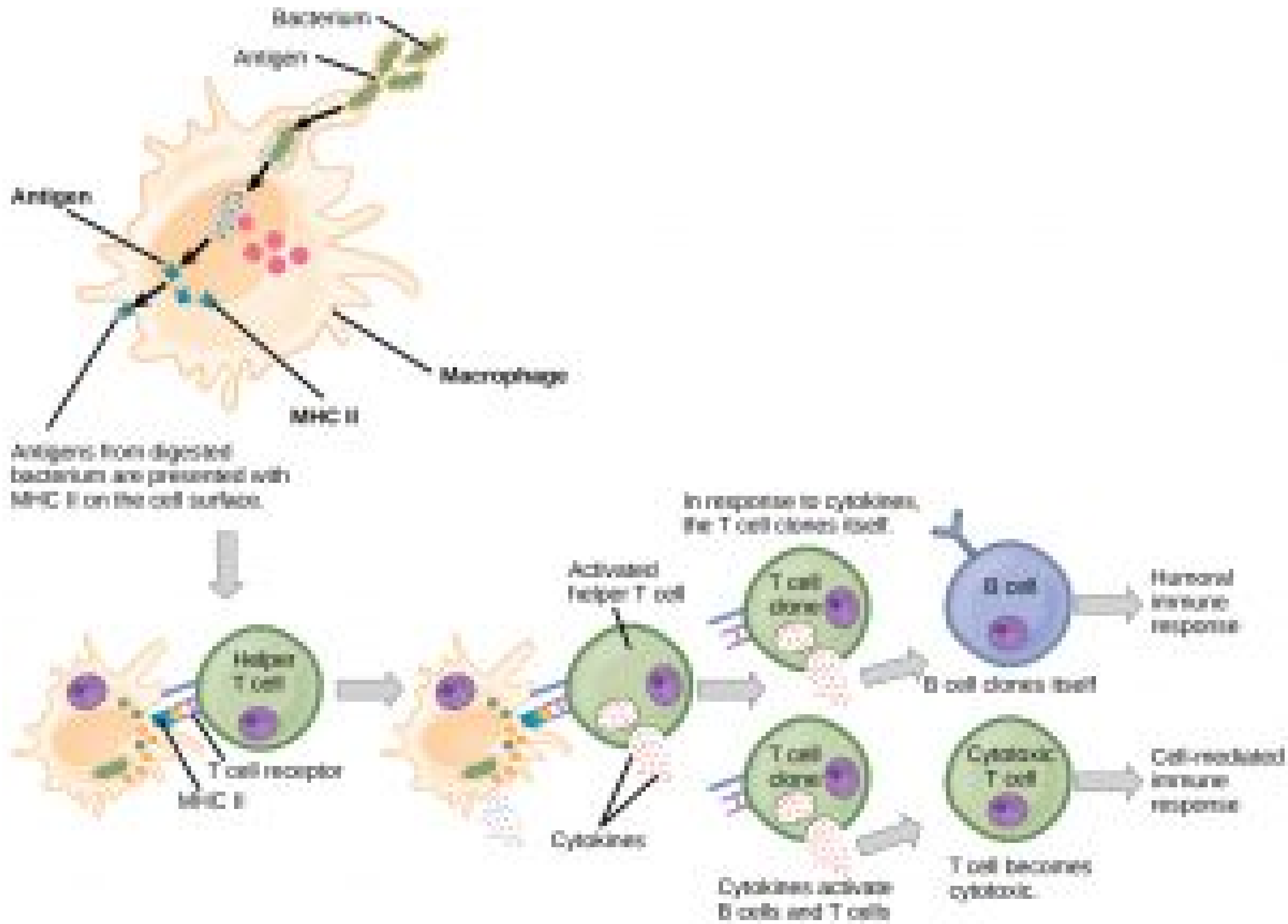


Figure 4. *A helper T cell becomes activated by binding to an antigen presented by an innate immune cell, causing it to release cytokines. Depending on the cytokines released, this activates either the humoral or the cell-mediated immune response.*

Although they do not directly kill infected cells, the function of the Helper T cells is crucial for fighting off pathogens. This important function is lost in untreated infection with the Human Immunodeficiency Virus, which causes AIDS (Acquired Immunodeficiency Syndrome.) HIV infects Helper T cells cells using their CD4 surface molecules, gradually depleting the number of Helper T cells in the body; this inhibits the adaptive immune system's capacity to generate sufficient responses to infection or tumors. As a result, HIV-infected individuals often suffer from infections that would not cause illness in people with healthy immune systems but which can cause devastating illness to immune-compromised individuals.Cytotoxic T cells are the key component of the cell-mediated part of the adaptive immune system and attack and destroy infected cells. Cytotoxic cells are particularly important in protecting against viral infections; this is because viruses replicate within cells where they are shielded from extracellular contact with circulating antibodies. Once activated, a cytotoxic T cell creates a large clone of cells with one specific set of cell-surface T cell receptors, as in the case with proliferation of activated B cells and Helper T cells. As with B cells, the clone includes active T_C cells and inactive memory T_C cells. The resulting active Cytotoxic T cells then identify infected host cells, bind to them and kill them. Phagocytes from the innate immune system then clean up the cellular debris, and ingest and destroy any pathogens that were inside of the infected cells. Cytotoxic T cells attempt to identify and destroy infected cells before the pathogen can replicate and escape, thereby halting the progression of intracellular infections. They also support NK lymphocytes to destroy early cancers. Cytokines secreted by activated Helper T cells stimulate cytotoxic T cells and enhance their ability to identify and destroy infected cells and tumors.

B plasma cells and Cytotoxic T cells are collectively called **effector cells** because they are involved in "effecting" (bringing about) the immune response of killing pathogens and infected host cells. Because of the time required to generate a population of clonal T and B cells during the first response to a pathogen, there is a delay in the adaptive immune response compared to the innate immune response. Memory cells of all three types of adaptive lymphocytes are, however, able to respond quickly upon subsequent infection.

Immunological Memory

The adaptive immune system has a memory component that allows for a rapid and large response upon reinvasion of the same pathogen. During the adaptive immune response to a pathogen that has not been encountered before, known as the **primary immune response**, plasma cells secreting antibodies and differentiated T cells increase, then plateau over time. As B and T cells mature into effector cells, a subset of the naïve populations differentiates into B and T memory cells with the same antigen specificities (***Figure 5***). A **memory cell** is an antigen-specific B or T lymphocyte that does not differentiate into an effector cell during the primary immune response, but that can immediately become an effector cell on re-exposure to the same pathogen. As the infection is cleared and pathogenic stimuli subside, the effectors are no longer needed and they undergo apoptosis. In contrast, the memory cells persist in the circulation.

ART CONNECTION

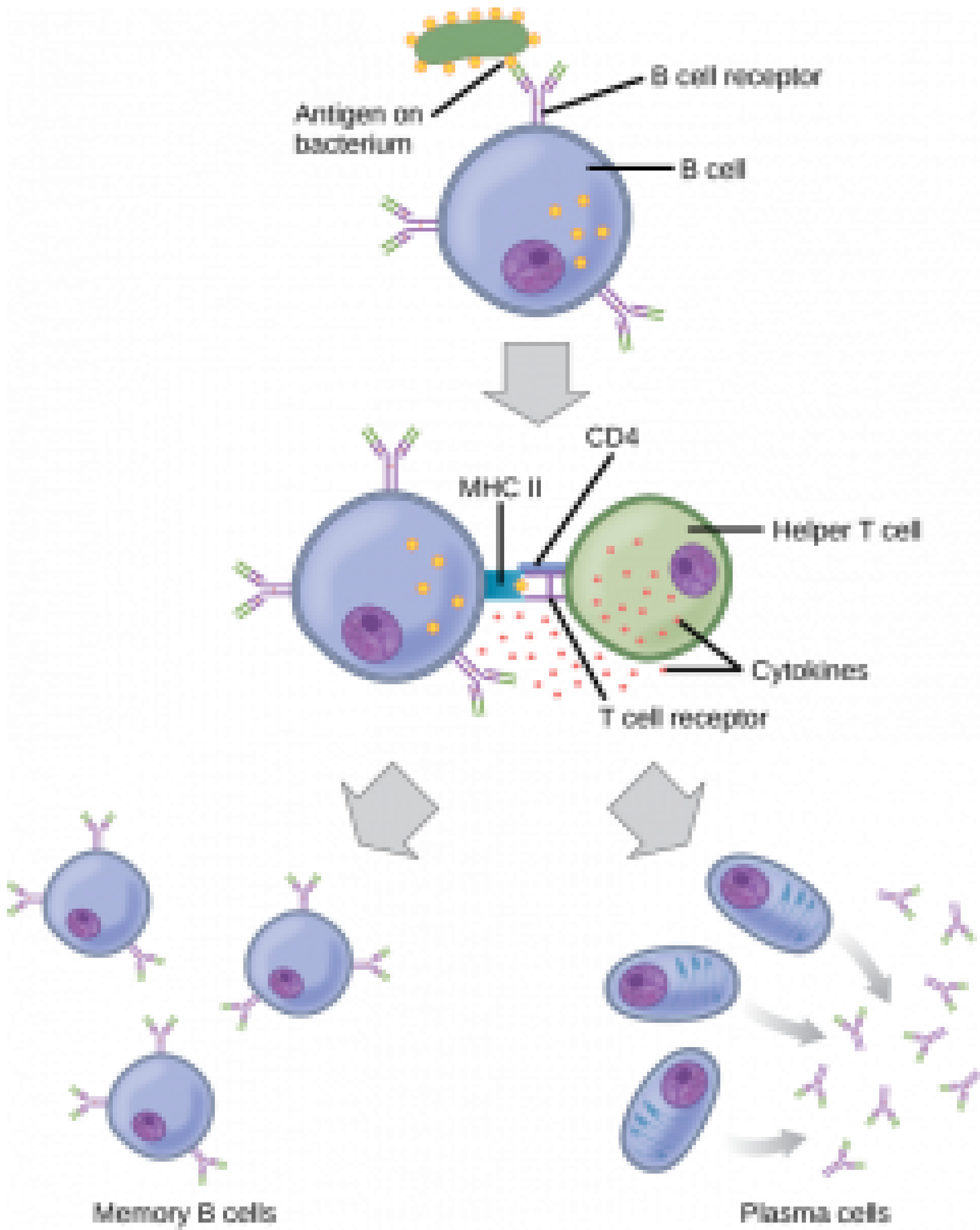


Figure 5. *After initially binding an antigen to the B cell receptor, a B cell is activated, and this response is enhanced by cytokines secreted by an activated Helper T cell. As a result, memory B cells and plasma cells are made. Similar processes generate memory Helper T cells and memory Cytotoxic T cells.*

If the pathogen is never encountered again during the individual's lifetime, B and T memory cells will circulate for a few years or even several decades and will gradually die off, having never functioned as effector cells. However, if the host is re-exposed to the same pathogen type, circulating memory cells will immediately differentiate into plasma cells and active Helper T cells without requiring the help of innate cells, or cytokines from activated Helper T cells. This is known as the **secondary immune response**. One reason why the adaptive immune response is delayed is because it takes time for naïve B and T cells with the appropriate antigen specificities to be identified, activated, and proliferate. On reinfection, this step is skipped, and the result is a more rapid production of immune defenses. Each of these newly reactivated effector cells also produces a stronger response than the first set of effector cells. For example, memory B cells that differentiate into plasma cells in a secondary immune response output tens to hundreds-fold greater antibody amounts than were secreted during the primary response (***Figure 6***). This rapid and dramatic antibody response may stop the infection before it can even become established, and before the innate immune system can initiate the inflammatory response that causes symptoms of infection such as fever, redness, swelling and body aches. The individual may not even realize they had been exposed to the pathogen.

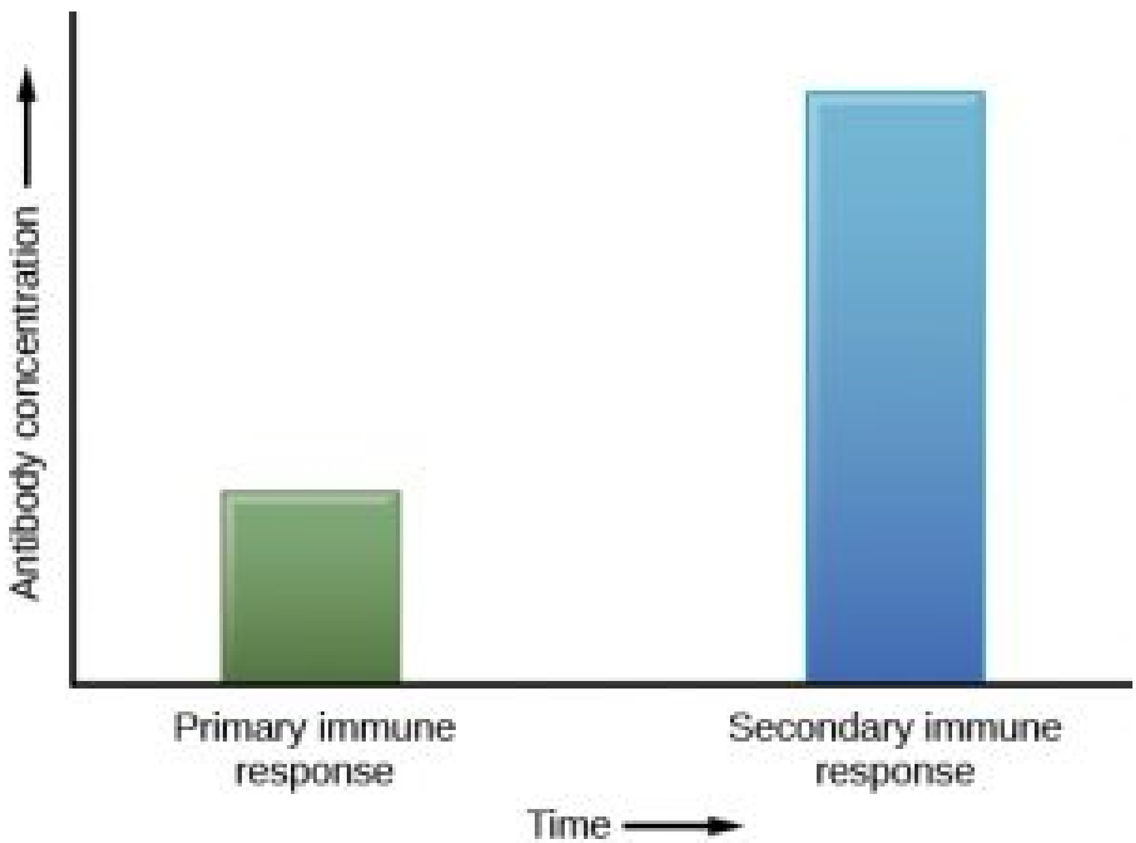


Figure 6. *In the primary response to infection, antibodies are secreted first from plasma cells. Upon re-exposure to the same pathogen, memory cells differentiate into antibody-secreting plasma cells that output a greater amount of antibody for a longer period of time.*

Vaccination is based on the knowledge that exposure to noninfectious antigens, derived from known pathogens, generates a mild primary immune response. Vaccines usually contain either weakened or dead microbes, or fragments of molecules found on pathogens that are known to contain antigens. The also usually contain molecules that stimulate the innate response, since innate immune responses are important in the generation of strong adaptive responses. The immune response to vaccination may not be perceived by the host as illness but still confers immune memory. When exposed to the corresponding pathogen to which an individual was vaccinated, the reaction is similar to a secondary exposure. Because each reinfection generates more memory cells and increased resistance to the pathogen, some vaccine courses involve one or more booster vaccinations to mimic repeat exposures.

Section Summary

The adaptive immune response is a slower-acting, longer-lasting, and more specific response than the innate response. However, the adaptive response requires information from the innate immune system to function. The adaptive immune response in B cells, Helper T cells and Cytotoxic T cells involved four phases: encounter, activation, attack, and memory. in this response, activated T cells differentiate and proliferate, becoming Helper (T_H) cells or Cytotoxic (T_C) cells. Helper T cells cells stimulate B cells, Cytotoxic T cells, and innate immune cells to enhance their attacks on pathogens. B cells differentiate into plasma cells that secrete antibodies, whereas Cytotoxic T cells destroy infected or cancerous cells. Memory cells are produced by activated and proliferating B and T cells and persist after a primary exposure to a pathogen. If re-exposure occurs, memory cells differentiate into effector cells without input from the innate immune system. Vaccination works by eliciting a primary immune response, without exposure to pathogens that could cause disease.

Glossary

active immunity
: an immunity that occurs as a result of the activity of the body’s own cells rather than from antibodies acquired from an external source

adaptive immunity
: a specific immune response that occurs after exposure to an antigen either from a pathogen or a vaccination

antibody
: a protein that is produced by plasma cells after stimulation by an antigen; also known as an immunoglobulin

antigen
: a macromolecule that reacts with cells of the immune system and which may or may not have a stimulatory effect

B cell
: a lymphocyte that matures in the bone marrow

cell-mediated immune response
: an adaptive immune response that is controlled by T cells

cytotoxic T lymphocyte (T_C)
an adaptive immune cell that directly kills infected cells via enzymes, and that releases cytokines to enhance the immune response
effector cell
a lymphocyte that has differentiated, such as a B cell, plasma cell, or cytotoxic T cell
helper T lymphocyte (T_H)
a cell of the adaptive immune system that binds APCs via MHC II molecules and stimulates B cells or secretes cytokines to initiate the immune response
humoral immune response
the adaptive immune response that is controlled by activated B cells and antibodies
lymph
the watery fluid present in the lymphatic circulatory system that bathes tissues and organs with protective white blood cells and does not contain erythrocytes
memory cell
an antigen-specific B or T lymphocyte that does not differentiate into an effector cell during the primary immune response but that can immediately become an effector cell on reexposure to the same pathogen
passive immunity
an immunity that does not result from the activity of the body's own immune cells but by transfer of antibodies from one individual to another
primary immune response
the response of the adaptive immune system to the first exposure to an antigen
secondary immune response
the response of the adaptive immune system to a second or later exposure to an antigen mediated by memory cells
T cell
a lymphocyte that matures in the thymus gland

19.

2.6 Human Genetics

Understanding Genetics

Biological researchers study genetics in order to better understand why individuals develop different physical traits, and psychological researchers study genetics in order to better understand the biological basis that contributes to certain behaviors. While all humans share certain biological mechanisms, we are each unique. And while our bodies have many of the same parts—brains and hormones and cells with genetic codes—these are expressed in a wide variety of traits, characteristics, behaviors, thoughts, and reactions.

Why do two people infected by the same disease have different outcomes: one surviving and one succumbing to the ailment? How are genetic diseases passed through family lines? Are there genetic components to psychological disorders, such as depression or schizophrenia? To what extent might there be a psychological basis to health conditions such as childhood obesity?

To explore these questions, let's start by focusing on a specific disease, **sickle-cell anemia**, and how it might affect two infected sisters. Sickle-cell anemia is a genetic condition in which red blood cells, which are normally round, take on a crescent-like shape (***Figure 1***). The changed shape of these cells affects how they function: sickle-shaped cells can clog blood vessels and block blood flow, leading to high fever, severe pain, swelling, and tissue damage.

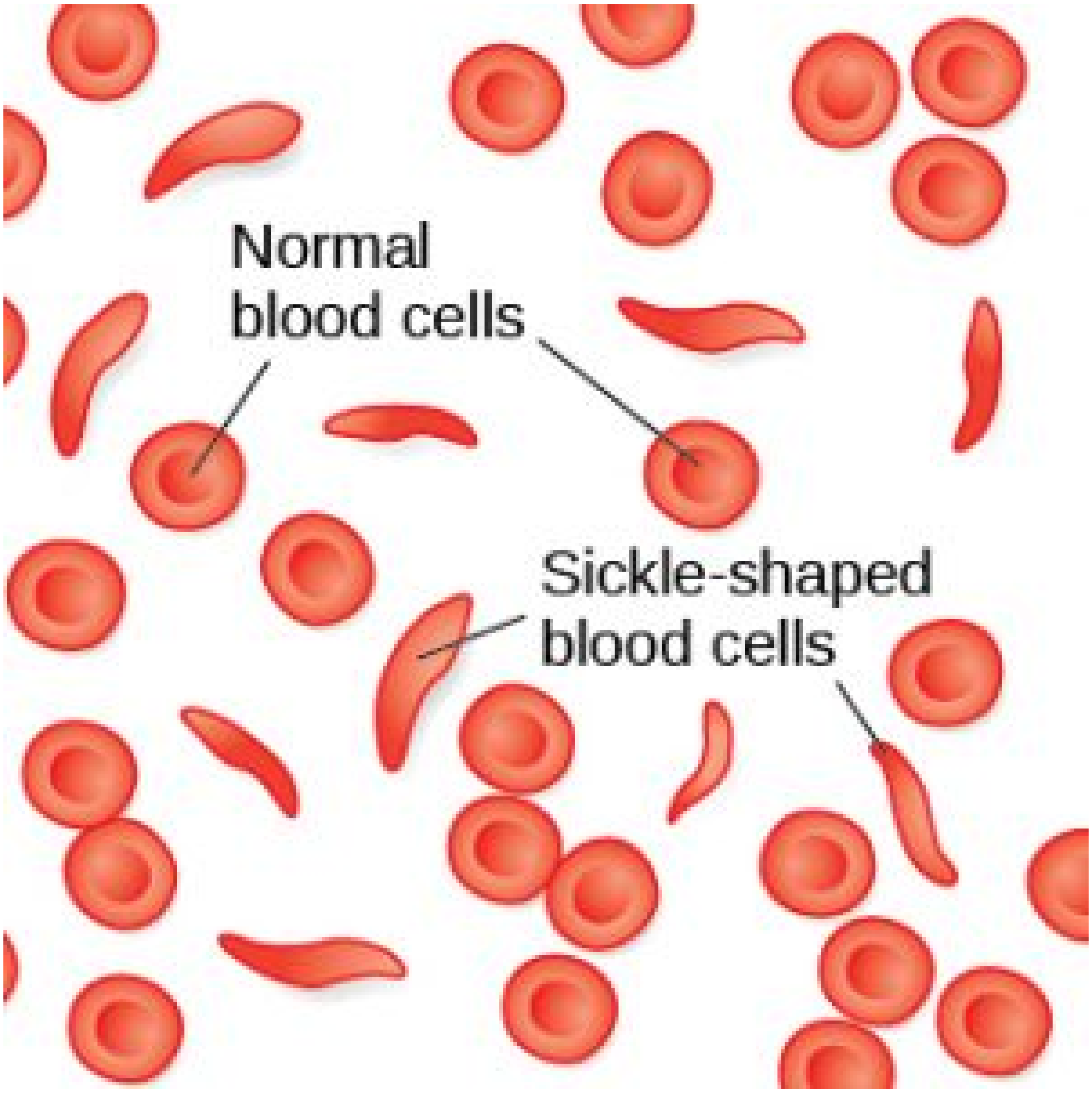


Figure 1. *Normal blood cells travel freely through the blood vessels, while sickle-shaped cells form blockages preventing blood flow.*

Many people with sickle-cell anemia—and the particular genetic mutation that causes it—die at an early age. While the notion of "survival of the fittest" may suggest that people suffering from this disease have a low survival rate and therefore the disease will become less common, this is not the case. Despite the negative evolutionary effects associated with this genetic mutation, the sickle-cell gene remains relatively common among people of African descent. Why is this? The explanation is illustrated with the following scenario.

Imagine two young women—Luwi and Sena—sisters in rural Zambia, Africa. Luwi carries the gene for sickle-cell anemia; Sena does not carry the gene. Sickle-cell carriers have one copy of the sickle-cell gene but do not have full-blown sickle-cell anemia. They experience symptoms only if they are severely dehydrated or are deprived of oxygen (as in mountain climbing). Carriers are thought to be immune from malaria (an often deadly disease that is widespread in tropical climates) because changes in their blood chemistry and immune functioning prevent the malaria parasite from having its effects. However, full-blown sickle-cell anemia, with two copies of the sickle-cell gene, does not provide immunity to malaria.

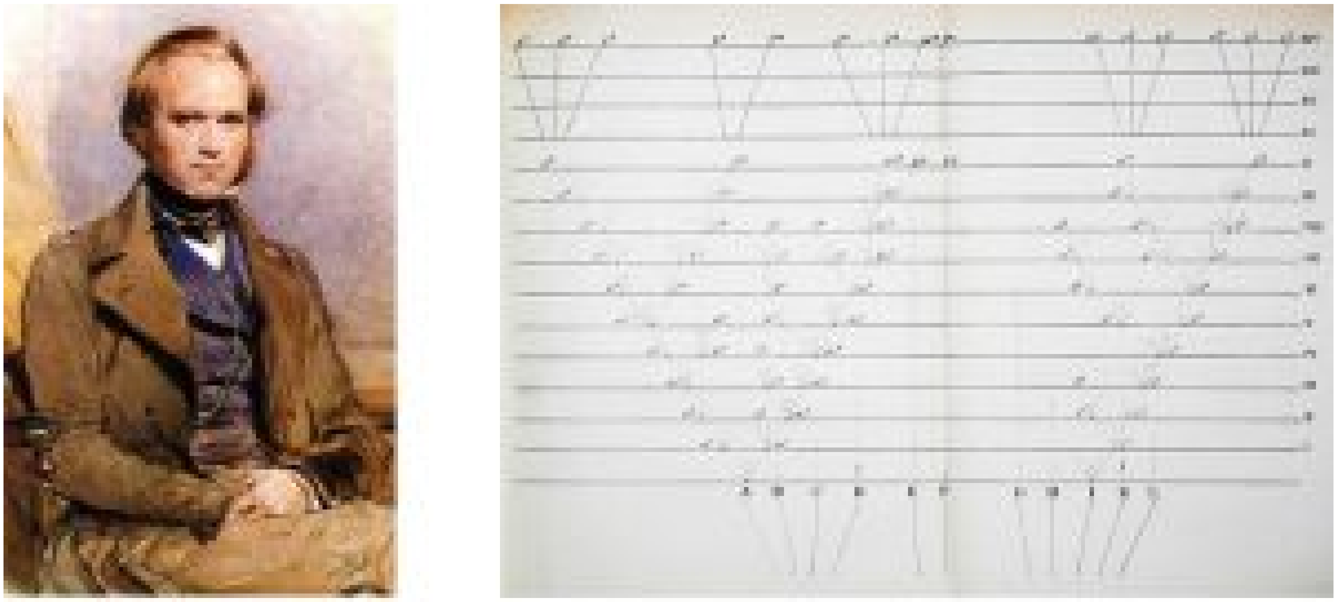
While walking home from school, both sisters are bitten by mosquitos carrying the malaria parasite. Luwi does not get malaria because she carries the sickle-cell mutation. Sena, on the other hand, develops malaria and dies just two weeks later. Luwi survives and eventually has children, to whom she may pass on the sickle-cell mutation.

Link to Learning

Visit this website to learn more about how a mutation in DNA leads to sickle-cell anemia.

Malaria is rare in the United States, so the sickle-cell gene benefits nobody: the gene manifests primarily in health problems—minor in carriers, severe in the full-blown disease—with no health benefits for carriers. However, the situation is quite different in other parts of the world. In parts of Africa where malaria is prevalent, having the sickle-cell mutation does provide health benefits for carriers (protection from malaria).

This is precisely the situation that Charles **Darwin** describes in the **theory of evolution by natural selection** (***Figure 2***), which you learned about in the previous section of this book. In simple terms, the theory states that organisms that are better suited for their environment will survive and reproduce, while those that are poorly suited for their environment will die off. In our example, we can see that as a carrier, Luwi's mutation is highly adaptive in her African homeland; however, if she resided in the United States (where malaria is much less common), her mutation could prove costly—with a high probability of the disease in her descendants and minor health problems of her own.



(a) (b)

Figure 2. ***(a)*** *In 1859, Charles Darwin proposed his theory of evolution by natural selection in his book, On the Origin of Species.* ***(b)*** *The book contains just one illustration: this diagram that shows how species evolve over time through natural selection.*

Genetic Variation

Genetic variation, the genetic difference between individuals, is what contributes to a species' adaptation to its environment. In humans, genetic variation begins with an egg, about 100 million sperm, and fertilization. Fertile women ovulate roughly once per month, releasing an egg from follicles in the ovary. During the egg's journey from the ovary through the fallopian tubes, to the uterus, a sperm may fertilize an egg.

The egg and the sperm each contain 23 chromosomes. **Chromosomes** are long strings of genetic material known as **deoxyribonucleic acid (DNA)**. DNA is a helix-shaped molecule made up of nucleotide base pairs. In each chromosome, sequences of DNA make up **genes** that control or partially control a number of visible characteristics, known as traits, such as eye color, hair color, and so on. A single gene may have multiple possible variations, or alleles. An **allele** is a specific version of a gene. So, a given gene may code for the trait of hair color, and the different alleles of that gene affect which hair color an individual has. The sickle-cell allele is one version of the hemoglobin gene, and this version of the gene has a different DNA sequence from the normal version of the hemoglobin.

When a sperm and egg fuse, their 23 chromosomes pair up and create a zygote with 23 pairs of chromosomes. Therefore, each parent contributes half the genetic information carried by the offspring; the resulting physical characteristics of the offspring (called the phenotype) are determined by the interaction of genetic material supplied by the parents (called the genotype). A person's **genotype** is the genetic makeup of that individual. **Phenotype**, on the other hand, refers to the individual's inherited physical characteristics, which are a combination of genetic and environmental influences (***Figure 3***).

(a) (b)

Figure 3. (a) *Genotype refers to the genetic makeup of an individual based on the genetic material (DNA) inherited from one's parents.* ***(b)*** *Phenotype describes an individual's observable characteristics, such as hair color, skin color, height, and build. (credit a: modification of work by Caroline Davis; credit b: modification of work by Cory Zanker)*

Most traits are controlled by multiple genes, but some traits are controlled by one gene. A characteristic like **cleft chin**, for example, is influenced by a single gene from each parent. In this example, we will call the gene for cleft chin "B," and the gene for smooth chin "b." Cleft chin is a dominant trait, which means that having the **dominant allele** either from one parent (Bb) or both parents (BB) will always result in the phenotype associated with the dominant allele. When someone has two copies of the same allele, they are said to be **homozygous** for that allele. When someone has a combination of alleles for a given gene, they are said to be **heterozygous**. For example, smooth chin is a recessive trait, which means that an individual will only display the smooth chin phenotype if they are homozygous for that **recessive allele** (bb).

Imagine that a woman with a cleft chin mates with a man with a smooth chin. What type of chin will their child have? The answer to that depends on which alleles each parent carries. If the woman is homozygous for cleft chin (BB), her offspring will always have cleft chin. It gets a little more complicated, however, if the mother is heterozygous for this gene (Bb). Since the father has a smooth chin—therefore homozygous for the recessive allele (bb)—we can expect the offspring to have a 50% chance of having a cleft chin and a 50% chance of having a smooth chin (***Figure 4***).

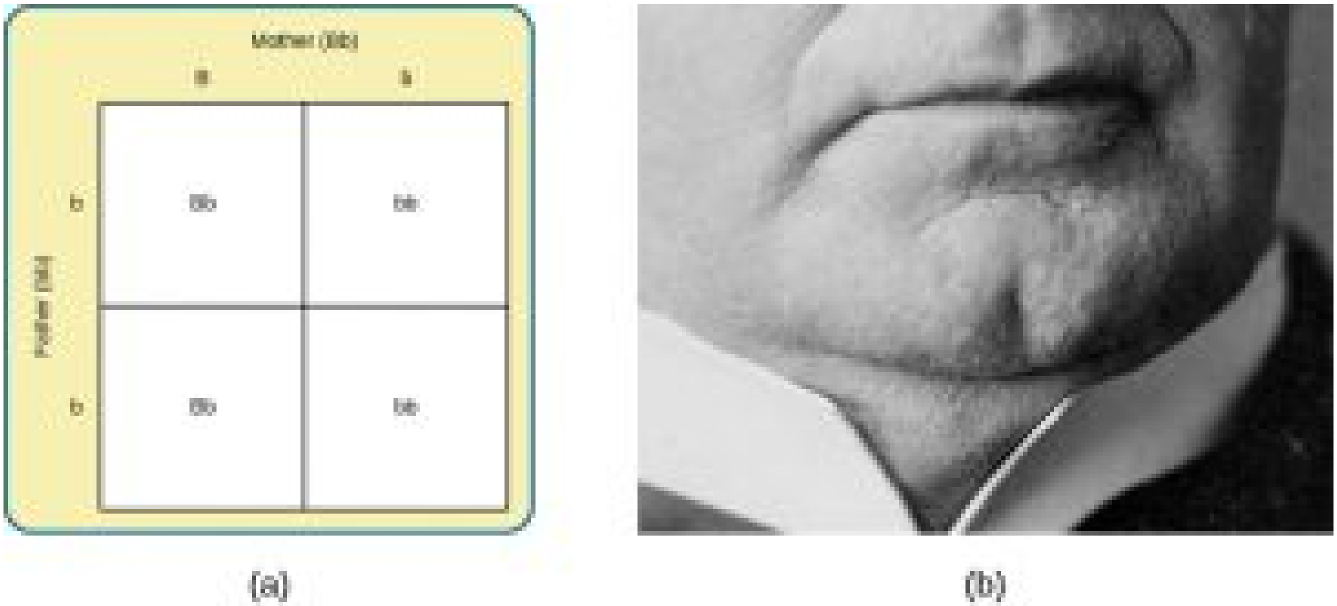


Figure 4. (a) *A Punnett square is a tool used to predict how genes will interact in the production of offspring. The capital B represents the dominant allele, and the lowercase b represents the recessive allele. In the example of the cleft chin, where B is cleft chin (dominant allele), wherever a pair contains the dominant allele, B, you can expect a cleft chin phenotype. You can expect a smooth chin phenotype only when there are two copies of the recessive allele, bb.* ***(b****) A cleft chin, shown here, is an inherited trait.*

Sickle-cell anemia is just one of many genetic disorders caused by the pairing of two recessive genes. For example, **phenylketonuria** (PKU) is a condition in which individuals lack an enzyme that normally converts harmful amino acids into harmless byproducts. If someone with this condition goes untreated, he or she will experience significant deficits in cognitive function, seizures, and increased risk of various psychiatric disorders. Because PKU is a recessive trait, each parent must have at least one copy of the recessive allele in order to produce a child with the condition (***Figure 5***).

So far, we have discussed traits that involve just one gene, but few human characteristics are controlled by a single gene. Most traits are **polygenic**: controlled by more than one gene. Height is one example of a polygenic trait, as are skin color and weight.

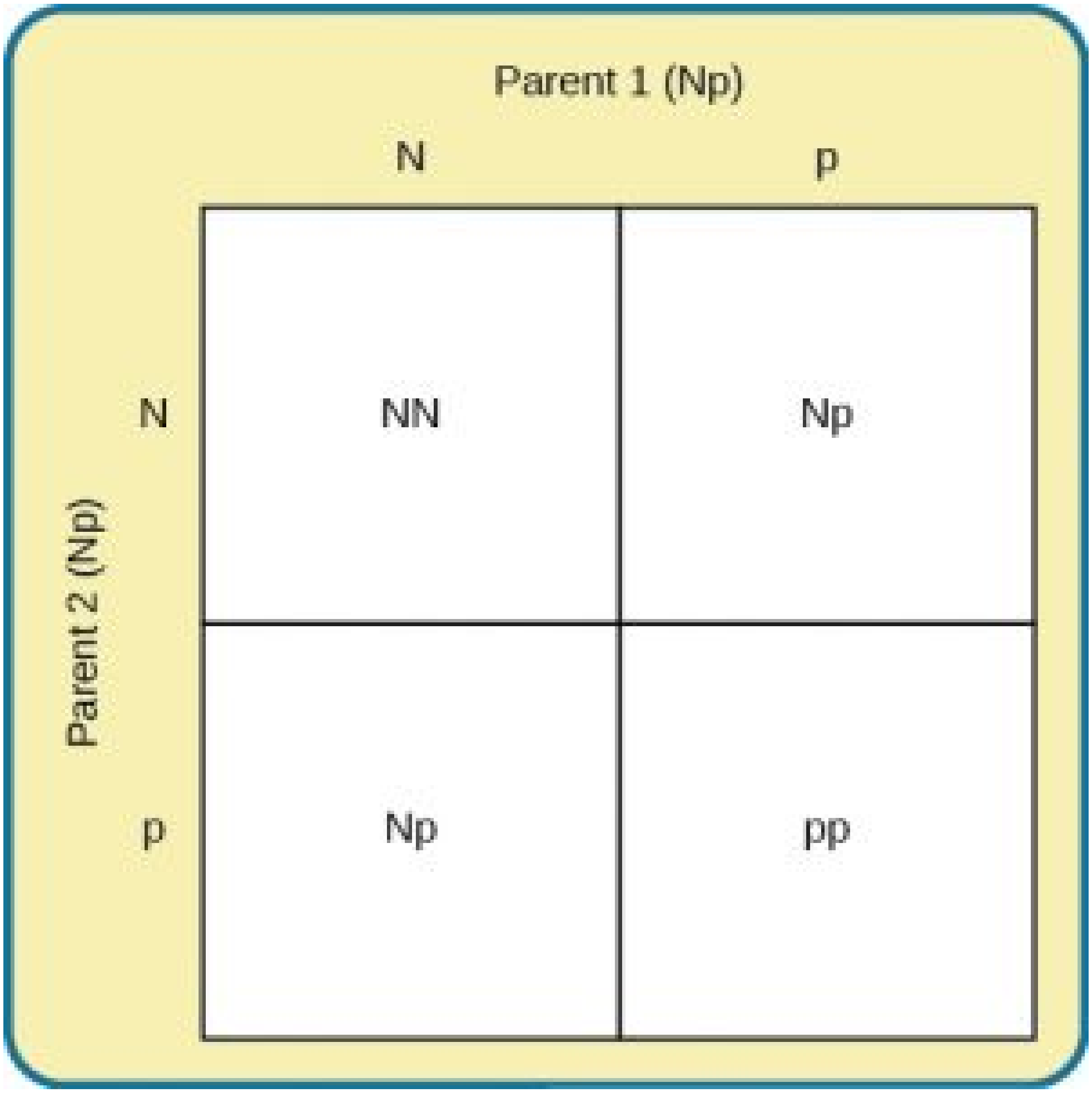


Figure 5. *In this* ***Punnett square,*** *N represents the normal allele, and p represents the recessive allele that is associated with PKU. If two individuals mate who are both heterozygous for the allele associated with PKU, their offspring have a 25% chance of expressing the PKU phenotype.*

Where do harmful genes that contribute to diseases like PKU come from? Gene mutations provide one source of harmful genes. A **mutation** is a sudden, permanent change in a gene. While many mutations can be harmful or lethal, once in a while, a mutation benefits an individual by giving that person an advantage over those who do not have the mutation. Recall that the theory of evolution asserts that individuals best adapted to their particular environments are more likely to reproduce and pass on their genes to future generations. In order for this process to occur, there must be competition—more technically, there must be variability in genes (and resultant traits) that allow for variation in adaptability to the environment. If a population consisted of identical individuals, then any dramatic changes in the environment would affect everyone in the same way, and there would be no variation in selection. In contrast, diversity in genes and associated traits allows some individuals to perform slightly better than others when faced with environmental change. This creates a distinct advantage for individuals best suited for their environments in terms of successful reproduction and genetic transmission.

Gene-Environment Interactions

Genes do not exist in a vacuum. Although we are all biological organisms, we also exist in an environment that is incredibly important in determining not only when and how our genes express themselves, but also in what combination. Each of us represents a unique interaction between our genetic makeup and our environment; range of reaction is one way to describe this interaction. **Range of reaction** asserts that our genes set the boundaries within which we can operate, and our environment interacts with the genes to determine where in that range we will fall. For example, if an individual's genetic makeup predisposes her to high levels of intellectual potential and she is reared in a rich, stimulating environment, then she will be more likely to achieve her full potential than if she were raised under conditions of significant deprivation. According to the concept of range of reaction, genes set definite limits on potential, and environment determines how much of that potential is achieved. Some disagree with this theory and argue that genes do not set a limit on a person's potential.

Another perspective on the interaction between genes and the environment is the concept of **genetic environmental correlation**. Stated simply, our genes influence our environment, and our environment influences the expression of our genes (***Figure 6***). Not only do our genes and environment interact, as in range of reaction, but they also influence one another bidirectionally. For example, the child of an NBA player would probably be exposed to basketball from an early age. Such exposure might allow the child to realize his or her full genetic, athletic potential. Thus, the parents' genes, which the child shares, influence the child's environment, and that environment, in turn, is well suited to support the child's genetic potential.

Figure 6. *Nature and nurture work together like complex pieces of a human puzzle. The interaction of our environment and genes makes us the individuals we are. (credit "puzzle": modification of work by Cory Zanker; credit "houses": modification of work by Ben Salter; credit "DNA": modification of work by NHGRI)*

In another approach to gene-environment interactions, the field of **epigenetics** looks beyond the genotype itself and studies how the same genotype can be expressed in different ways. In other words, researchers study how the same genotype can lead to very different phenotypes. As mentioned earlier, gene expression is often influenced by environmental context in ways that are not entirely obvious. For instance, identical twins share the same genetic information (**identical twins** develop from a single fertilized egg that split, so the genetic material is exactly the same in each; in contrast, **fraternal twins** develop from two different eggs fertilized by different sperm, so the genetic material varies as with non-twin siblings). But even with identical genes, there remains an incredible amount of variability in how gene expression can unfold over the course of each twin's life. Sometimes, one twin will develop a disease and the other will not. In one example, Tiffany, an identical twin, died from cancer at age 7, but her twin, now 19 years old, has never had cancer. Although these individuals share an identical genotype, their phenotypes differ as a result of how that genetic information is expressed over time. The epigenetic perspective is very different from range of reaction, because here the genotype is not fixed and limited.

Link to Learning

Visit this site for an engaging video primer on the **epigenetics** of twin studies.

Genes affect more than our physical characteristics. Indeed, scientists have found genetic linkages to a number of behavioral characteristics, ranging from basic personality traits to sexual orientation to spirituality (for examples, see Mustanski et al., 2005; Comings, Gonzales, Saucier, Johnson, & MacMurray, 2000). Genes are also associated with temperament and a number of psychological disorders, such as depression and schizophrenia. So while it is true that genes provide the biological blueprints for our cells, tissues, organs, and body, they also have significant impact on our experiences and our behaviors.

Let's look at the following findings regarding **schizophrenia** in light of our three views of gene-environment interactions. Which view do you think best explains this evidence?

In a study of people who were given up for adoption, adoptees whose biological mothers had schizophrenia *and* who had been raised in a disturbed family environment were much more likely to develop schizophrenia or another psychotic disorder than were any of the other groups in the study:

- Of adoptees whose biological mothers had schizophrenia (high genetic risk) and who were raised in disturbed family environments, 36.8% were likely to develop schizophrenia.
- Of adoptees whose biological mothers had schizophrenia (high genetic risk) and who were raised in healthy family environments, 5.8% were likely to develop schizophrenia.
- Of adoptees with a low genetic risk (whose mothers did not have schizophrenia) and who were raised in disturbed family environments, 5.3% were likely to develop schizophrenia.
- Of adoptees with a low genetic risk (whose mothers did not have schizophrenia) and who were raised in healthy family environments, 4.8% were likely to develop schizophrenia (Tienari et al., 2004).

The study shows that adoptees with high genetic risk were especially likely to develop schizophrenia only if they were raised in disturbed home environments. This research lends credibility to the notion that both genetic vulnerability and environmental stress are necessary for schizophrenia to develop, and that genes alone do not tell the full tale.

Section Summary

Genes are sequences of DNA that code for a particular trait. Different versions of a gene are called alleles—sometimes alleles can be classified as dominant or recessive. A dominant allele always results in the dominant phenotype. In order to exhibit a recessive phenotype, an individual must be homozygous for the recessive allele. Genes affect both physical and psychological characteristics. Ultimately, how and when a gene is expressed, and what the outcome will be—in terms of both physical and psychological characteristics—is a function of the interaction between our genes and our environments.

Glossary

allele
: specific version of a gene

chromosome
: long strand of genetic information

deoxyribonucleic acid (DNA)
: helix-shaped molecule made of nucleotide base pairs

dominant allele
: allele whose phenotype will be expressed in an individual that possesses that allele

epigenetics
: study of gene-environment interactions, such as how the same genotype leads to different phenotypes

fraternal twins
: twins who develop from two different eggs fertilized by different sperm, so their genetic material varies the same as in non-twin siblings

gene
: sequence of DNA that controls or partially controls physical characteristics

genetic environmental correlation
: view of gene-environment interaction that asserts our genes affect our environment, and our environment influences the expression of our genes

genotype
: genetic makeup of an individual

heterozygous
: consisting of two different alleles

homozygous
: consisting of two identical alleles

identical twins
: twins that develop from the same sperm and egg

mutation
: sudden, permanent change in a gene

phenotype
: individual's inheritable physical characteristics

polygenic
: multiple genes affecting a given trait

range of reaction
: asserts our genes set the boundaries within which we can operate, and our environment interacts with the genes to determine where in that range we will fall

recessive allele
: allele whose phenotype will be expressed only if an individual is homozygous for that allele

theory of evolution by natural selection
: states that organisms that are better suited for their environments will survive and reproduce compared to those that are poorly suited for their environments

20.

2.7 Transcription

In all cells, the second function of DNA (the first was replication) is to provide the information needed to construct the proteins necessary so that the cell can perform all of its functions. To do this, the DNA is “read” or transcribed into an **mRNA** molecule. The mRNA then provides the code to form a protein by a process called translation. Through the processes of transcription and translation, a protein is built with a specific sequence of amino acids that was originally encoded in the DNA. This module discusses the details of transcription.

The Central Dogma: DNA Encodes RNA; RNA Encodes Protein

The 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 sequences of mRNAs, which in turn specify the sequences of proteins.

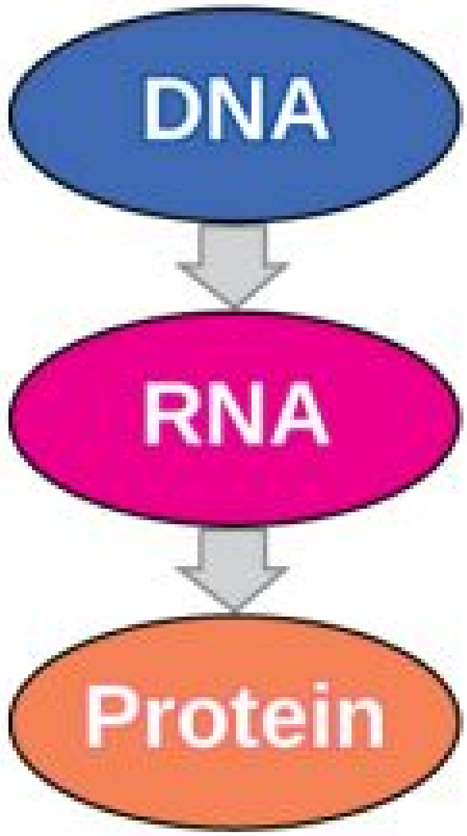


Figure 1. *The central dogma states that DNA encodes RNA, which in turn encodes protein.*

The copying of DNA to mRNA is relatively straightforward, with one nucleotide being added to the mRNA strand for every complementary nucleotide read in the DNA strand. The translation to protein is more complex because groups of three mRNA nucleotides correspond to one amino acid of the protein sequence. However, as we shall see in the next module, the translation to protein is still systematic, such that nucleotides 1 to 3 correspond to amino acid 1, nucleotides 4 to 6 correspond to amino acid 2, and so on.

Transcription: from DNA to mRNA

Transcription occurs in three main stages: initiation, elongation and termination. Because genes in animal cells are found in the nucleus, transcription occurs in the nucleus of the cell and the mRNA transcript must be transported to the cytoplasm.

Initiation

Transcription 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 (***Figure 2***).

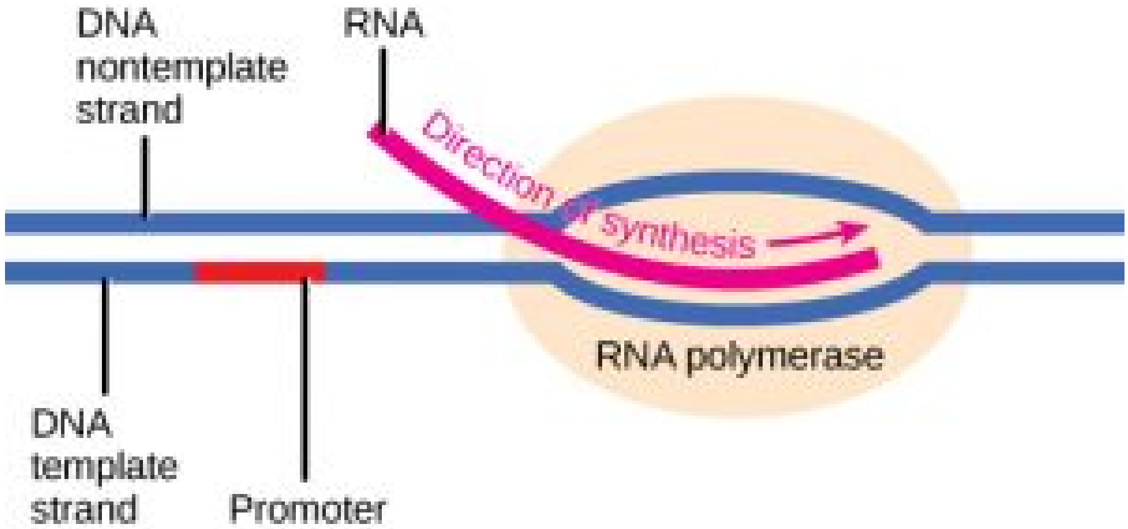


Figure 2. *The initiation of transcription begins when DNA is unwound, forming a transcription bubble. Enzymes and other proteins involved in transcription bind at the promoter.*

Elongation

Transcription always proceeds from one of the two DNA strands, which is called the **template strand**. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the **nontemplate strand**, with the exception that RNA contains a uracil (U) in place of the thymine (T) found in DNA. During elongation, an enzyme called **RNA polymerase** proceeds along the DNA template adding nucleotides by base pairing with the DNA template in a manner similar to DNA replication, with the difference that an RNA strand is being synthesized that does not remain bound to the DNA template. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it (***Figure 3***).

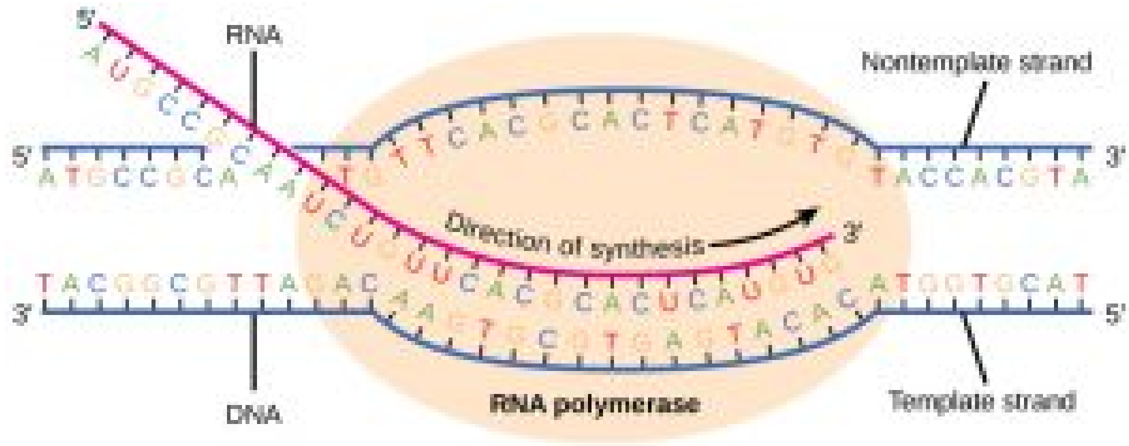


Figure 3. *During elongation, RNA polymerase tracks along the DNA template, synthesizes mRNA in the 5' to 3' direction, and unwinds then rewinds the DNA as it is read.*

Termination

Once a gene is transcribed, the 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, but both involve repeated nucleotide sequences in the DNA template that result in RNA polymerase stalling, leaving the DNA template, and freeing the mRNA transcript. On termination, the process of transcription is complete. After termination, the transcript can be transported out of the nucleus, where translation can occur.

Section Summary

In animal cells, mRNA synthesis is initiated at a promoter sequence on the DNA template. Elongation synthesizes new mRNA. Termination liberates the mRNA and occurs by mechanisms that stall the RNA polymerase and cause it to fall off the DNA template. Only finished mRNAs are exported from the nucleus to the cytoplasm.

Glossary

mRNA
: messenger RNA; a form of RNA that carries the nucleotide sequence code for a protein sequence that is translated into a polypeptide sequence

nontemplate strand
: the strand of DNA that is not used to transcribe mRNA; this strand is identical to the mRNA except that T nucleotides in the DNA are replaced by U nucleotides in the mRNA

promoter
a sequence on DNA to which RNA polymerase and associated factors bind and initiate transcription

RNA polymerase
an enzyme that synthesizes an RNA strand from a DNA template strand

template strand
the strand of DNA that specifies the complementary mRNA molecule

transcription bubble
the region of locally unwound DNA that allows for transcription of mRNA

21.

2.8 Translation

The synthesis of proteins is one of a cell's most energy-consuming metabolic processes. In turn, proteins account for more mass than any other component of living organisms (with the exception of water), and proteins perform a wide variety of the functions of a cell. The process of translation, or protein synthesis, involves decoding an mRNA message into a protein product. Amino acids are strung together in lengths ranging from approximately 50 amino acids to more than 1,000.

The Protein Synthesis Machinery

In addition to the mRNA template, many other molecules contribute to the process of translation. The composition of each component may vary across species; for instance, ribosomes may consist of different numbers of ribosomal RNAs (**rRNA**) and polypeptides depending on the organism. However, 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***).

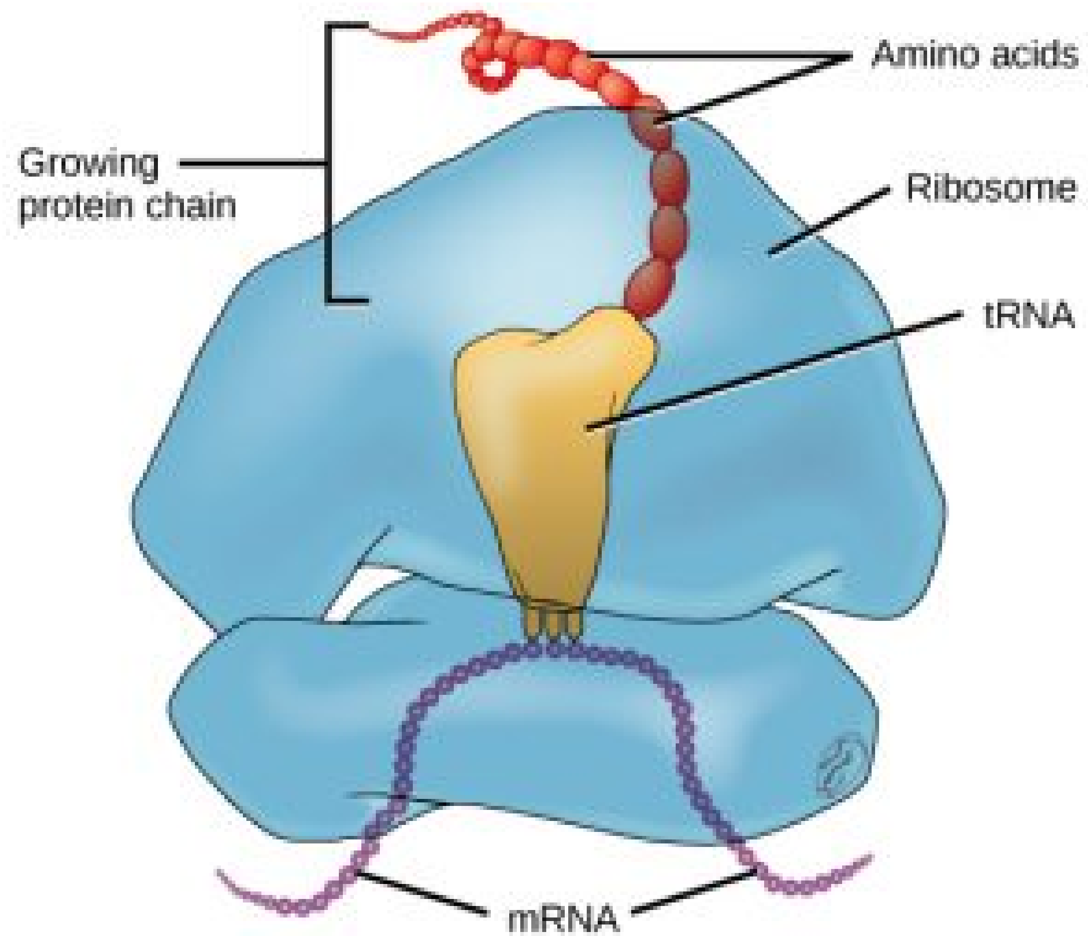


Figure 1. *The protein synthesis machinery includes the large and small subunits of the ribosome, mRNA, and tRNA. (credit: modification of work by NIGMS, NIH)*

A ribosome is a complex macromolecule composed of structural and catalytic rRNAs, and many distinct proteins. Ribosomes are located in the cytoplasm and endoplasmic reticulum of animal cells. Each mRNA molecule is simultaneously translated by many ribosomes, all synthesizing protein in the same direction. Ribosomes are made up of a large and a small subunit that come together for translation. 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.

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. 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.

The Genetic Code

To summarize the last chapter, 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 this mRNA template converts nucleotide-based genetic information into a protein product. Protein sequences consist of 20 commonly occurring amino acids; therefore, it can be said that the protein alphabet consists of 20 letters. 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 and protein "alphabets," combinations of nucleotides corresponded to single amino acids. Using a three-nucleotide code means that there are a total of 64 (4 × 4 × 4) possible combinations; therefore, a given amino acid is encoded by more than one nucleotide triplet. A table of the amino acids encoded by each triplet, called a Codon Table, is shown in ***Figure 2***.

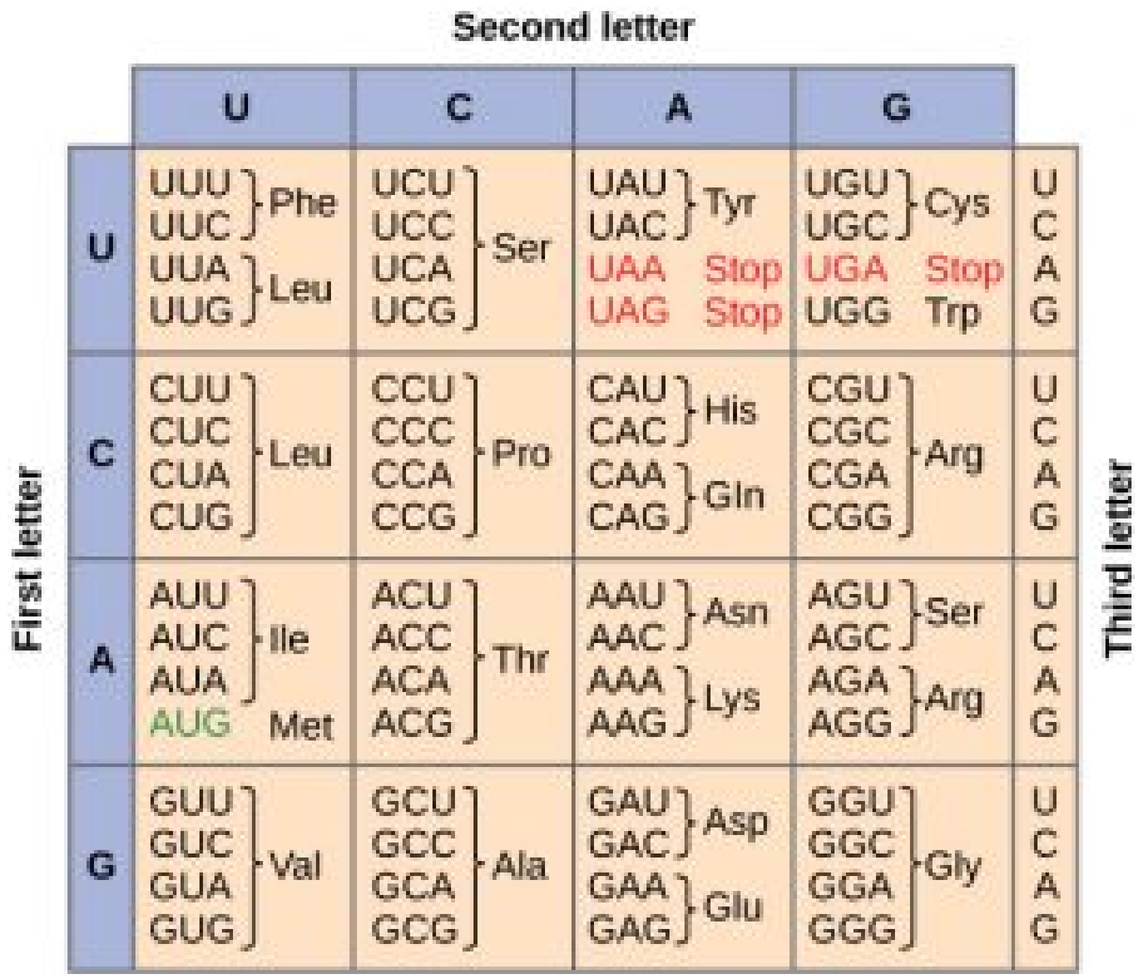


Figure 2. *This figure shows the genetic code for translating each nucleotide triplet, or codon, in mRNA into an amino acid or a termination signal in a nascent protein. (credit: modification of work by NIH)*

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.

This codon signals to the the parts of the translation machinery to begin adding amino acids, and also recruits a tRNA carrying the amino acid methionine, abbreviated Met on the codon table above. Because of this, all newly . synthesized proteins start with the amino acid methionine. 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.

The Mechanism of Protein Synthesis

Just as with mRNA synthesis, protein synthesis can be divided into three phases: initiation, elongation, and termination. The process of translation is similar in most cells, in most species. Protein synthesis begins with the formation of an initiation complex. In *E. coli*, this complex involves the small ribosome subunit, the mRNA template, three initiation factors, and a special initiator tRNA. The initiator tRNA interacts with the AUG start codon, and links to a special form of the amino acid methionine that is typically removed from the polypeptide after translation is complete.

The elongation phase of translation involves the large ribosomal subunit. The large ribosomal subunit of *E. coli* consists of three compartments: the A site binds *Arriving* charged tRNAs (tRNAs with their attached specific amino acids). The P site binds charged tRNAs carrying amino acids that have formed bonds with the growing *Protein* chain but have not yet dissociated from their corresponding tRNA. The E site releases dissociated tRNAs so they can *exit* and be recharged with free amino acids. The ribosome shifts one codon at a time, catalyzing each process that occurs in the three sites. With each step, a charged tRNA enters the complex, the polypeptide becomes one amino acid longer, and an uncharged tRNA departs(***Figure 3***).

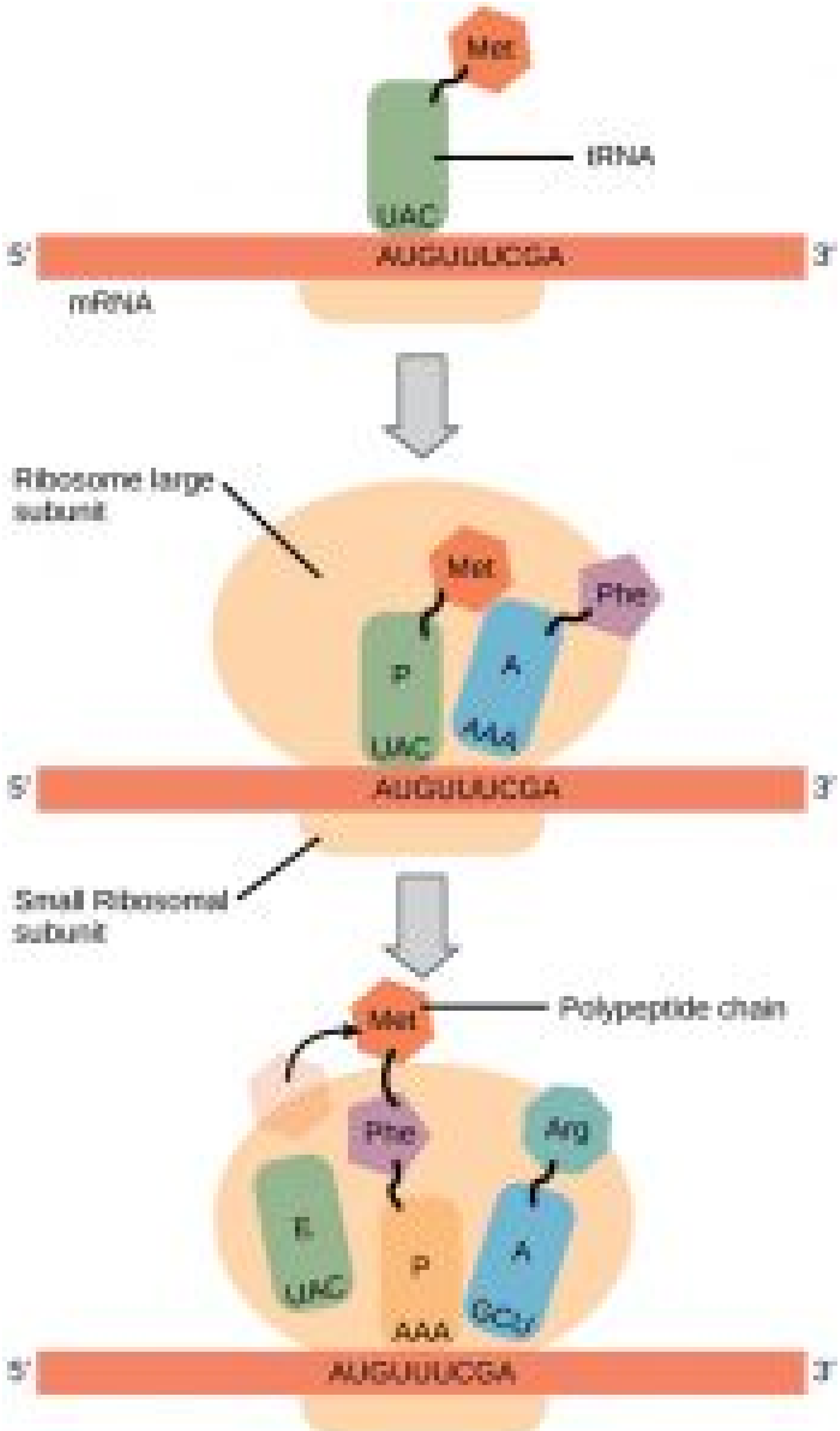

Met
tRNA
UAC
5'
AUGUUUCGA
3'
mRNA
Ribosome large subunit
Met
Phe
P
A
UAC
AAA
5'
AUGUUUCGA
3'
Small Ribosomal subunit
Met
Polypeptide chain
Phe
Arg
E
P
A
UAC
AAA
GCU
5'
AUGUUUCGA
3'

Figure 3. *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.*

Termination of translation occurs when a stop codon (UAA, UAG, or UGA) is encountered. When the ribosome encounters the stop codon, the growing polypeptide is released and the ribosome subunits dissociate and leave the mRNA. After many ribosomes have completed translation, the mRNA is degraded so the nucleotides can be reused in another transcription reaction.

Concept in Action

Transcribe a gene and translate it to protein using complementary pairing and the genetic code at this site.

Section Summary

The central dogma describes the flow of genetic information in the cell from genes to mRNA to proteins. Genes are used to make mRNA by the process of transcription; mRNA is used to synthesize proteins by the process of translation. The genetic code is the correspondence between the three-nucleotide mRNA codon and an amino acid. The genetic code is “translated” by the tRNA molecules, which associate a specific codon with a specific amino acid. The genetic code is degenerate because 64 triplet codons in mRNA specify only 20 amino acids and three stop codons. This means that more than one codon corresponds to an amino acid. Almost every species on the planet uses the same genetic code.

The players in translation include the mRNA template, ribosomes, tRNAs, and various enzymatic factors. The small ribosomal subunit binds to the mRNA template. Translation begins at the initiating AUG on the mRNA. The formation of bonds occurs between sequential amino acids specified by the mRNA template according to the genetic code. The ribosome accepts charged tRNAs, and as it steps along the mRNA, it catalyzes bonding between the new amino acid and the end of the growing polypeptide. The entire mRNA is translated in three-nucleotide “steps” of the ribosome. When a stop codon is encountered, a release factor binds and dissociates the components and frees the new protein.

Glossary

codon
: three consecutive nucleotides in mRNA that specify the addition of a specific amino acid or the release of a polypeptide chain during translation

genetic code
: the amino acids that correspond to three-nucleotide codons of mRNA

rRNA
: ribosomal RNA; molecules of RNA that combine to form part of the ribosome

stop codon
: one of the three mRNA codons that specifies termination of translation

start codon
: the AUG (or, rarely GUG) on an mRNA from which translation begins; always specifies methionine

tRNA
: transfer RNA; an RNA molecule that contains a specific three-nucleotide anticodon sequence to pair with the mRNA codon and also binds to a specific amino acid

22.

2.9 How Genes are Regulated

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 an individual cell of 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. Each cell also has many genes that are not expressed, and expresses many that are not expressed by other cells, such that it can carry out its 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 also turn on and off genes in response to the demands of their environment so that they can respond to special conditions.

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.

Gene Expression

To understand how gene expression is regulated, we must first understand how a gene becomes a functional protein in a cell. In simple organisms like bacteria, the primary method to control what type and how much protein is expressed is through the regulation of DNA transcription into RNA. Animal cells, in contrast, have intracellular organelles and are much more complex. Recall that in complex cells, the DNA is contained inside the cell's nucleus and it is transcribed into mRNA there. The newly synthesized mRNA is then transported out of the nucleus into the cytoplasm, where ribosomes translate the mRNA into 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 all stages of the process (***Figure 1***). Regulation may occur when the DNA is uncoiled and loosened from nucleosomes to bind transcription factors (**epigenetic** level), when the RNA is transcribed (transcriptional level), when RNA is processed and exported to the cytoplasm after it is transcribed (**post-transcriptional** level), when the RNA is translated into protein (translational level), or after the protein has been made (**post-translational** level).

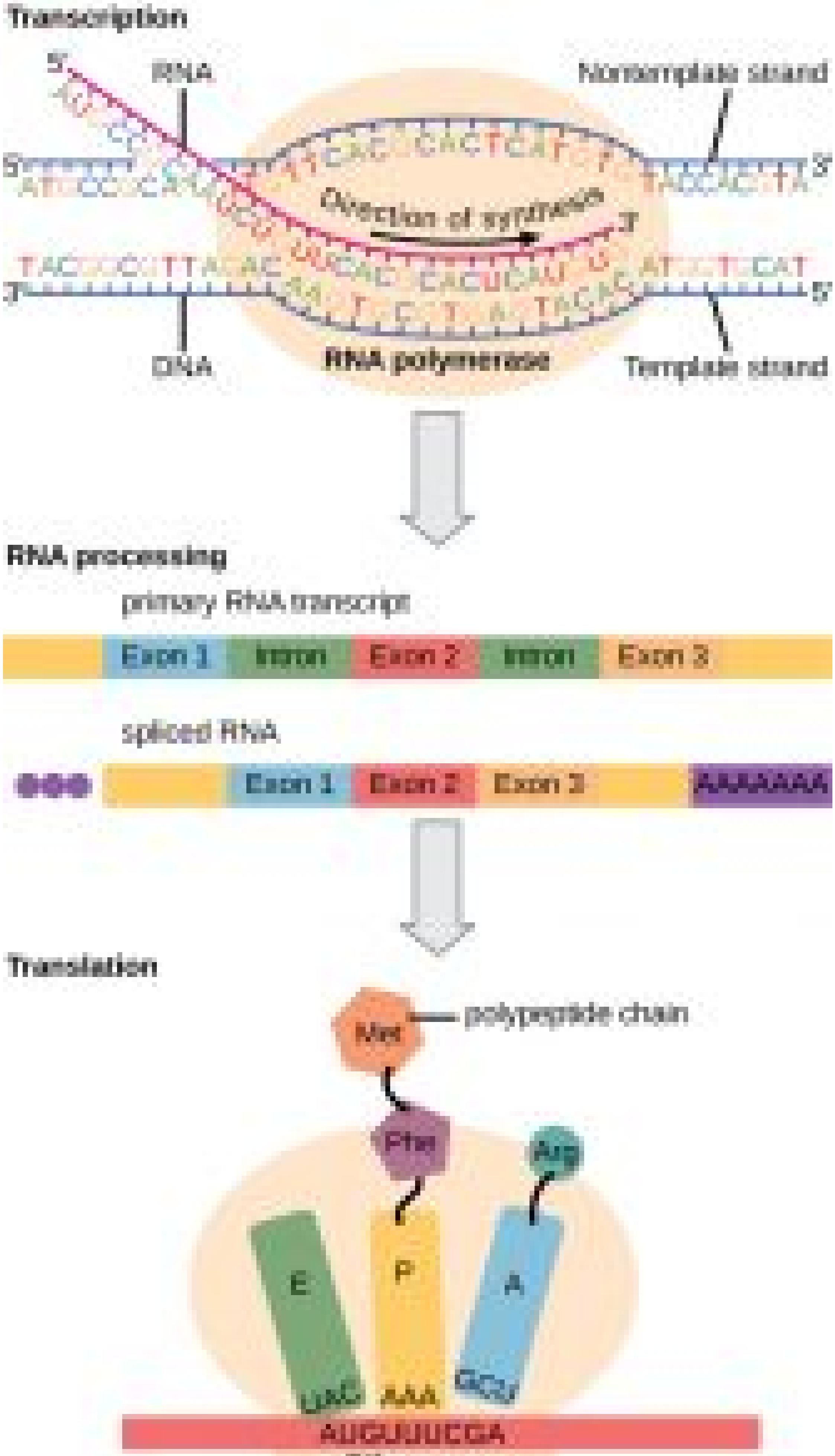

Transcription
5'
RNA
Nontemplate strand
3'
Direction of synthesis
3'
5'
DNA
RNA polymerase
Template strand
RNA processing
primary RNA transcript
Exon 1
Intron
Exon 2
Intron
Exon 3
spliced RNA
Exon 1
Exon 2
Exon 3
AAAAAAA
Translation
Met
polypeptide chain
Phe
Arg
E
P
A
UAC
AAA
GCU
AUGUUUCGA

Figure 1. *Human 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.*

Section Summary

While all somatic cells within an organism contain the same DNA, not all cells within that organism express the same proteins. Simple organisms like bacteria express the entire DNA they encode in every cell, but not necessarily all at the same time. Proteins are expressed only when they are needed. Complex organisms express a subset of the DNA that is encoded in any given cell. In each cell type, the type and amount of protein is regulated by controlling gene expression. To express a protein, the DNA is first transcribed into RNA, which is then translated into proteins. In simple cells, these processes occur almost simultaneously. In animal cells, transcription occurs in the nucleus and is separate from the translation that occurs in the cytoplasm. In animal cells, gene expression is regulated at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels.

Glossary

epigenetic
: describing non-genetic regulatory factors, such as changes in modifications to histone proteins and DNA that control accessibility to genes in chromosomes

gene expression
: processes that control whether a gene is expressed

post-transcriptional
: control of gene expression after the RNA molecule has been created but before it is translated into protein

post-translational
: control of gene expression after a protein has been created

III

Theme 3: How Do Diet, Exercise and Stress Affect Health?

In this part of the course, we will explore how stress affects health, and also how our bodies gain, store and use energy. First, we will examine the role of the cardiovascular system, nervous system and endocrine system help our bodies prepare for stressful situations. Then, we will look at the types of molecules required in our diets in order to maintain and build healthy cells, tissues and organs. We will also learn about how those nutrients are incorporated into our cells and tissues in the processes of digestion and metabolism. Finally, we will cover how the balance of energy use and storage in the body is related to exercise and weight, and discuss some of the consequences of having too much or too little energy storage, both for the individual and for society as a whole.

Of course, diet and exercise alone determine health. Among many other components of health, we will consider stress and its impact on health. In lab, you will explore physical and psychological measurements of stress. In lecture, we will discuss the processes used by the body to deal with stress, some ways that stress affects health, and some of the disparities in stress-related diseases based on income, race and gender.

23.

3.1 What is Stress?

Learning Objectives

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

- Describe a scenario in which stress has a positive effect on performance.
- What is the "fight or flight" response? With which part of the nervous system is this response associated?
- Differentiate between sympathetic and parasympathetic responses of the autonomic nervous system.
- What bodily functions change when the parasympathetic division is activated? Think about your daily routine. Which division is most active when you are running to class? Which is most active when you are dozing in class?
- Describe how the hypothalamus, pituitary gland and adrenal glands work together to respond to stress during a "sympathetic" response. Be able to match structure to function.
- Describe the short term effects of adrenaline and cortisol on the body. How does this help us respond to a stressor? (alarm response)
- Explain how short and long term effects of stress differ. (interpret tables shown in class)
- List some of the negative health consequences of chronic stress (interpret tables shown in class)

The term **stress** as it relates to the human condition first emerged in scientific literature in the 1930s, but it did not enter the popular vernacular until the 1970s (Lyon, 2012). Today, we often use the term loosely in describing a variety of unpleasant feeling states; for example, we often say we are stressed out when we feel frustrated, angry, conflicted, overwhelmed, or fatigued. Despite the widespread use of the term, stress is a fairly vague concept that is difficult to define with precision.

Researchers have had a difficult time agreeing on an acceptable definition of stress. Some have conceptualized stress as a demanding or threatening event or situation (e.g., a high-stress job, overcrowding, and long commutes to work), however, for the purpose of this course, we will describe stress as a **response to environmental conditions**. For example, the endocrinologist Hans **Selye**, a famous stress researcher, once defined stress as the "response of the body to any demand, whether it is caused by, or results in, pleasant or unpleasant conditions" (Selye, 1976, p. 74). Selye's definition of stress is response-based in that it conceptualizes stress chiefly in terms of the body's physiological reaction to any demand that is placed on it. Many of the physiological reactions that occur when faced with demanding situations (e.g., accelerated heart rate) can also occur in response to things that most people would not consider to be genuinely stressful, such as receiving unanticipated good news: an unexpected promotion or raise.

A useful way to conceptualize **stress** is to view it as a process whereby an individual perceives and responds to events that he appraises as overwhelming or threatening to his well-being (Lazarus & Folkman, 1984). A critical element of this definition is that it emphasizes the importance of how we appraise—that is, judge—demanding or threatening events (often referred to as **stressors**); these appraisals, in turn, influence our reactions to such events. Two kinds of appraisals of a stressor are especially important in this regard: primary and secondary appraisals. A **primary appraisal** involves judgment about the degree of potential harm or threat to well-being that a stressor might entail. A stressor would likely be appraised as a threat if one anticipates that it could lead to some kind of harm, loss, or other negative consequence; conversely, a stressor would likely be appraised as a challenge if one believes that it carries the potential for gain or personal growth. For example, an employee who is promoted to a leadership position would likely perceive the promotion as a much greater threat if she believed the promotion would lead to excessive work demands than if she viewed it as an opportunity to gain new skills and grow professionally. Similarly, a college student on the cusp of graduation may face the change as a threat or a challenge (***Figure 1***).

Figure 1. *Graduating from college and entering the workforce can be viewed as either a threat (loss of financial support) or a challenge (opportunity for independence and growth). (credit: Timothy Zanker)*

The perception of a threat triggers a **secondary appraisal**: judgment of the options available to cope with a stressor, as well as perceptions of how effective such options will be (Lyon, 2012) (***Figure 2***). As you may recall from what you learned about self-efficacy, an individual's belief in his ability to complete a task is important (Bandura, 1994). A threat tends to be viewed as less catastrophic if one believes something can be done about it (Lazarus & Folkman, 1984). Imagine that two middle-aged women, Robin and Maria, perform breast self-examinations one morning and each woman notices a lump on the lower region of her left breast. Although both women view the breast lump as a potential threat (primary appraisal), their secondary appraisals differ considerably. In considering the breast lump, some of the thoughts racing through Robin's mind are, "Oh my God, I could have breast cancer! What if the cancer has spread to the rest of my body and I cannot recover? What if I have to go through chemotherapy? I've heard that experience is awful! What if I have to quit my job? My husband and I won't have enough money to pay the mortgage. Oh, this is just horrible…I can't deal with it!" On the other hand, Maria thinks, "Hmm, this may not be good. Although most times these things turn out to be benign, I need to have it checked out. If it turns out to be breast cancer, there are doctors who can take care of it because the medical technology today is quite advanced. I'll have a lot of different options, and I'll be just fine." Clearly, Robin and Maria have different outlooks on what might turn out to be a very serious situation: Robin seems to think that little could be done about it, whereas Maria believes that, worst case scenario, a number of options that are likely to be effective would be available. As such, Robin would clearly experience greater stress than would Maria.

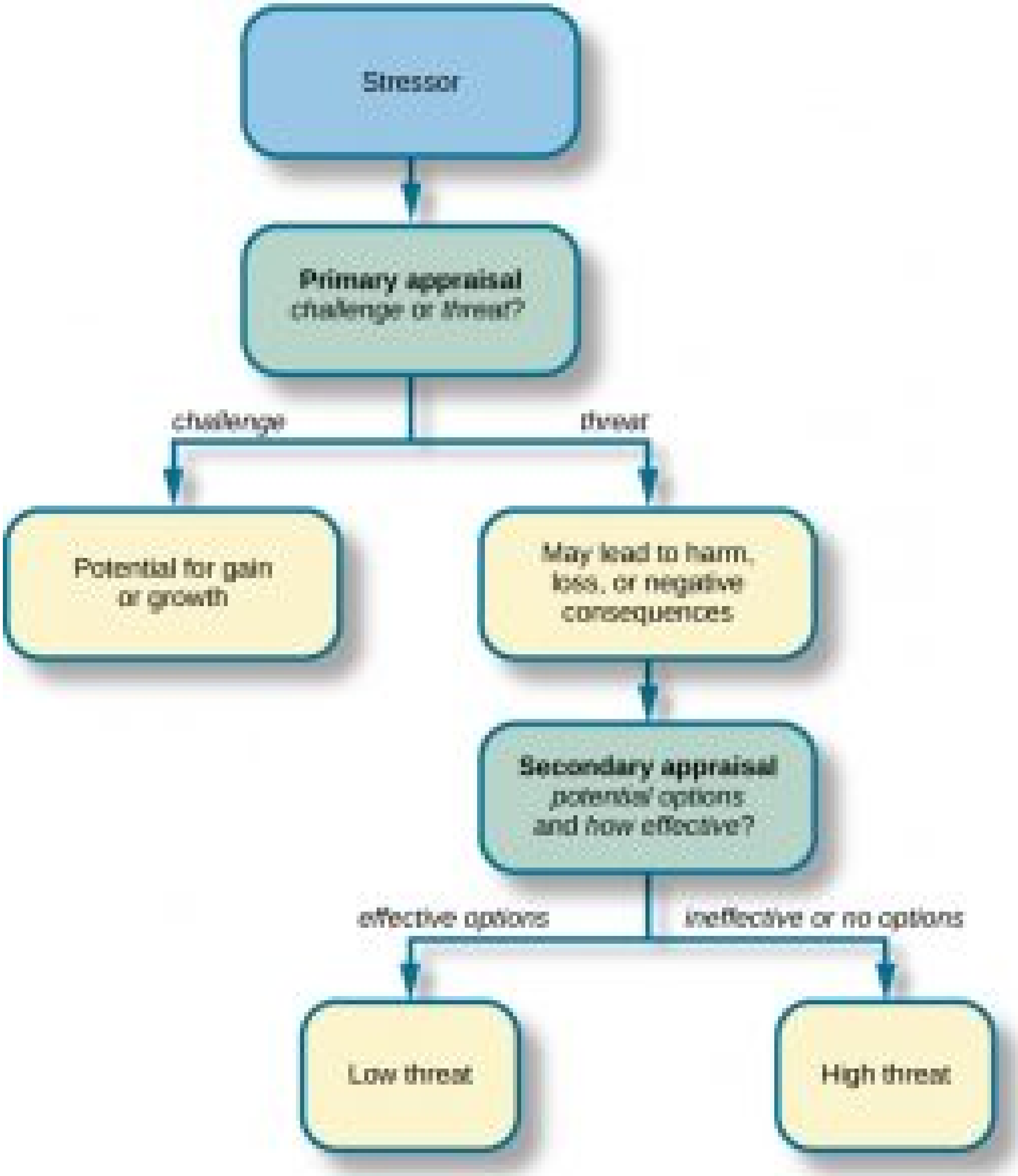


Figure 2. *When encountering a stressor, a person judges its potential threat (primary appraisal) and then determines if effective options are available to manage the situation. Stress is likely to result if a stressor is perceived as extremely threatening or threatening with few or no effective coping options available.*

To be sure, some stressors are inherently more stressful than others in that they are more threatening and leave less potential for variation in cognitive appraisals (e.g., objective threats to one's health or safety). Nevertheless, appraisal will still play a role in augmenting or diminishing our reactions to such events (Everly & Lating, 2002).

If a person appraises an event as harmful and believes that the demands imposed by the event exceed the available resources to manage or adapt to it, the person will subjectively experience a state of stress. In contrast, if one does not appraise the same event as harmful or threatening, she is unlikely to experience stress. According to this definition, environmental events trigger stress reactions by the way they are interpreted and the meanings they are assigned. In short, stress is largely in the eye of the beholder: it's not so much what happens to you as it is how you respond (Selye, 1976).

Good Stress?

Although stress carries a negative connotation, at times it may be of some benefit. Stress can motivate us to do things in our best interests, such as study for exams, visit the doctor regularly, exercise, and perform to the best of our ability at work. Indeed, Selye (1974) pointed out that not all stress is harmful. He argued that stress can sometimes be a positive, motivating force that can improve the quality of our lives. This kind of stress, which Selye called **eustress** (from the Greek *eu* = "good"), is a good kind of stress associated with positive feelings, optimal health, and performance. A moderate amount of stress can be beneficial in challenging situations. For example, athletes may be motivated and energized by pregame stress, and students may experience similar beneficial stress before a major exam. Indeed, research shows that moderate stress can enhance both immediate and delayed recall of educational material. Male participants in one study who memorized a scientific text passage showed improved memory of the passage immediately after exposure to a mild stressor as well as one day following exposure to the stressor (Hupbach & Fieman, 2012).

Increasing one's level of stress will cause performance to change in a predictable way. As shown in ***Figure 3***, as stress increases, so do performance and general well-being (eustress); when stress levels reach an optimal level (the highest point of the curve), performance reaches its peak. A person at this stress level is colloquially at the top of his game, meaning he feels fully energized, focused, and can work with minimal effort and maximum efficiency. But when stress exceeds this optimal level, it is no longer a positive force—it becomes excessive and debilitating, or what Selye termed **distress** (from the Latin *dis* = "bad"). People who reach this level of stress feel burned out; they are fatigued, exhausted, and their performance begins to decline. If the stress remains excessive, health may begin to erode as well (Everly & Lating, 2002).

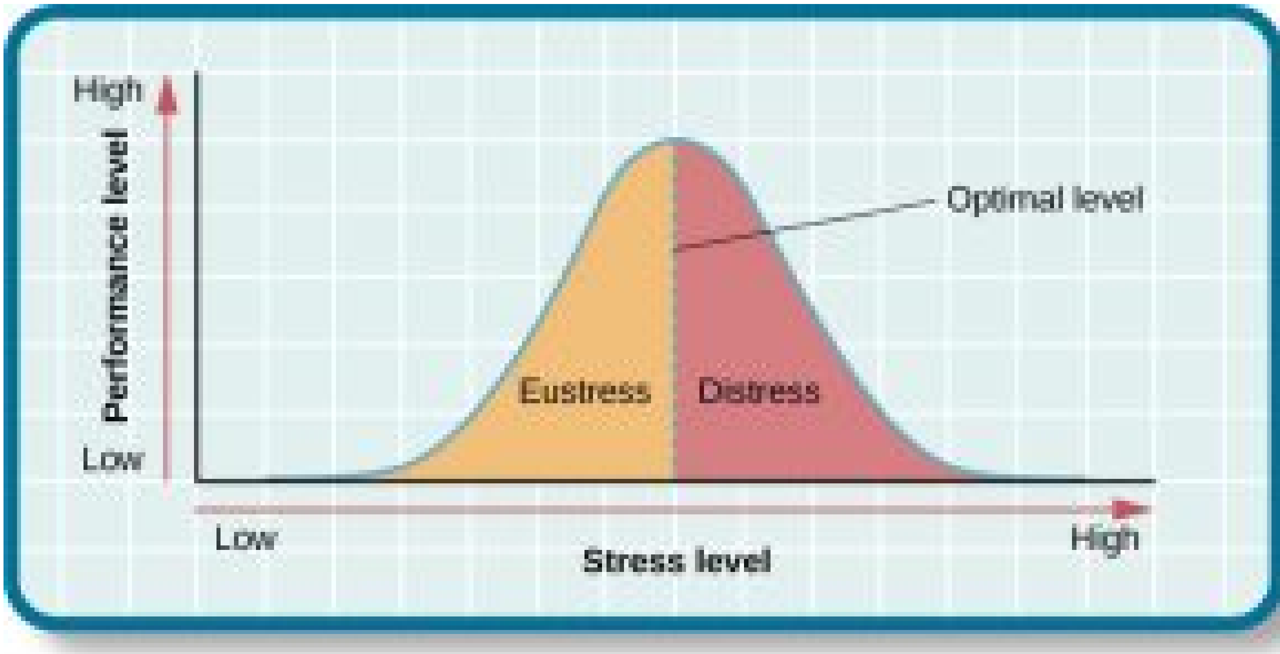


Figure 3. As the stress level increases from low to moderate, so does performance (eustress). At the optimal level (the peak of the curve), performance has reached its peak. If stress exceeds the optimal level, it will reach the distress region, where it will become excessive and debilitating, and performance will decline (Everly & Lating, 2002).

The Prevalence of Stress

Stress is an experience that evokes a variety of responses, including those that are physiological (e.g., accelerated heart rate, headaches, or gastrointestinal problems), cognitive (e.g., difficulty concentrating or making decisions), and behavioral (e.g., drinking alcohol, smoking, or taking actions directed at eliminating the cause of the stress). Although stress can be positive at times, it can have deleterious health implications, contributing to the onset and progression of a variety of physical illnesses and diseases (Cohen & Herbert, 1996).

The scientific study of how stress and other psychological factors impact health falls within the realm of **health psychology**, a subfield of psychology devoted to understanding the importance of psychological influences on health, illness, and how people respond when they become ill (Taylor, 1999). Health psychologists identify which groups of people are especially at risk for negative health outcomes, based on psychological or behavior factors, and also design and monitor interventions aimed at changing unhealthy behaviors. Measuring differences in stress levels among demographic groups and how these levels change over time can help identify populations who may have an increased risk for illness or disease.

Figure 4 depicts the results of three national surveys in which several thousand individuals from different demographic groups completed a brief stress questionnaire; the surveys were administered in 1983, 2006, and 2009 (Cohen & Janicki-Deverts, 2012). All three surveys demonstrated higher stress in women than in men. Unemployed individuals reported high levels of stress in all three surveys, as did those with less education and income; retired persons reported the lowest stress levels. However, from 2006 to 2009 the greatest increase in stress levels occurred among men, Whites, people aged 45–64, college graduates, and those with full-time employment. One interpretation of these findings is that concerns surrounding the 2008–2009 economic downturn (e.g., threat of or actual job loss and substantial loss of retirement savings) may have been especially stressful to White, college-educated, employed men with limited time remaining in their working careers.

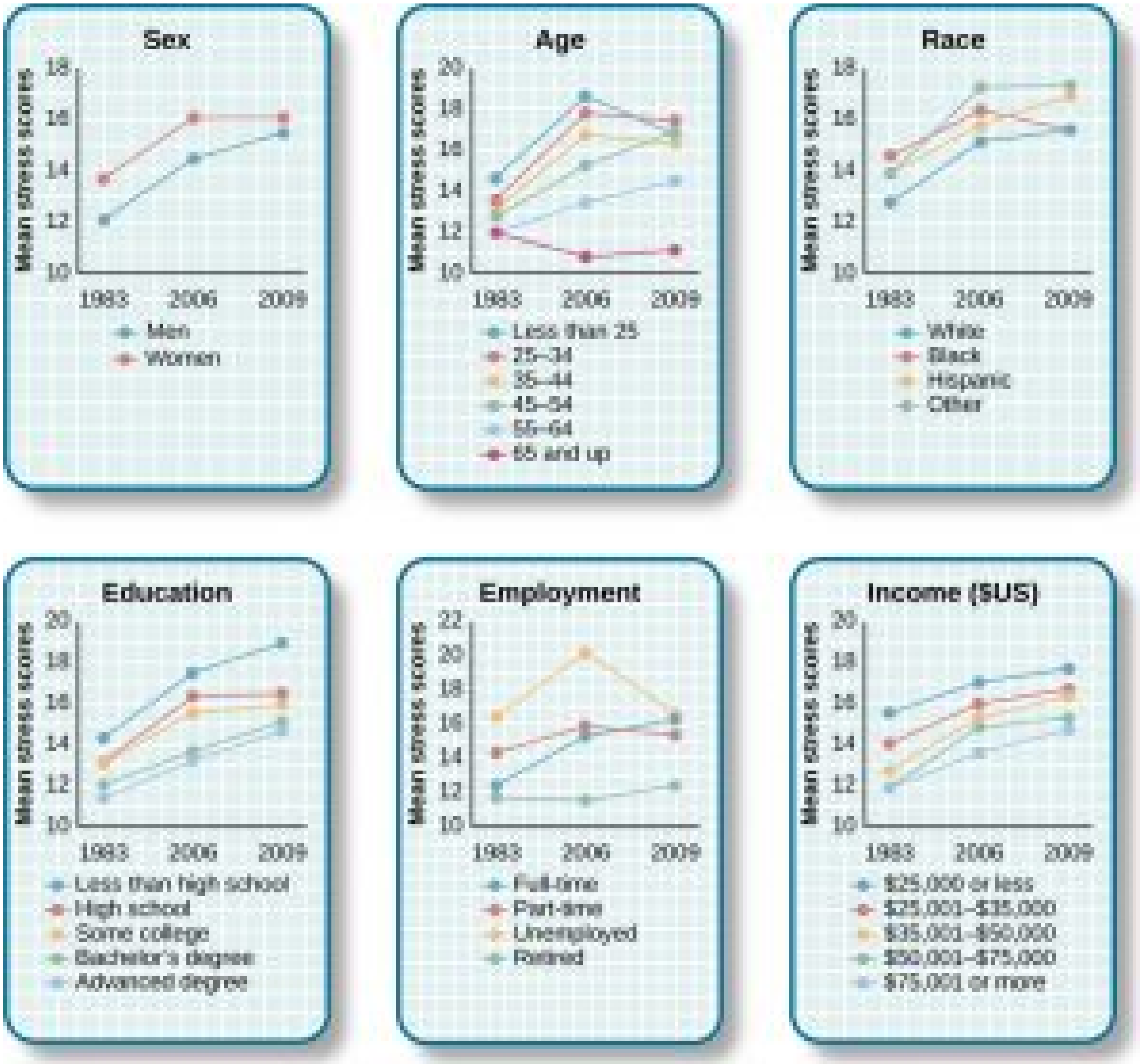


Figure 4. *The charts above, adapted from Cohen & Janicki-Deverts (2012), depict the mean stress level scores among different demographic groups during the years 1983, 2006, and 2009. Across categories of sex, age, race, education level, employment status, and income, stress levels generally show a marked increase over this quarter-century time span.*

Early Contributions to the Study of Stress

In the early 20th century, Walter **Cannon**, an eminent American physiologist at Harvard Medical School (***Figure 5***), was the first to identify the body's physiological reactions to stress.

Figure 5. *Harvard physiologist Walter Cannon first articulated and named the fight-or-flight response, the nervous system's sympathetic response to a significant stressor.*

Cannon and the Fight-or-Flight Response

Imagine that you are hiking in the beautiful mountains of Colorado on a warm and sunny spring day. At one point during your hike, a large, frightening-looking black bear appears from behind a stand of trees and sits about 50 yards from you. The bear notices you, sits up, and begins to lumber in your direction. In addition to thinking, “This is definitely not good,” a constellation of physiological reactions begins to take place inside you. Prompted by a deluge of epinephrine (adrenaline) and norepinephrine (noradrenaline) from your adrenal glands, your pupils begin to dilate. Your heart starts to pound and speeds up, you begin to breathe heavily and perspire, you get butterflies in your stomach, and your muscles become tense, preparing you to take some kind of direct action. Cannon proposed that this reaction, which he called the **fight-or-flight response**, occurs when a person experiences very strong emotions—especially those associated with a perceived threat (Cannon, 1932). During the fight-or-flight response, the body is rapidly aroused by activation of both the sympathetic nervous system and the endocrine system (***Figure 6***). This arousal helps prepare the person to either fight or flee from a perceived threat.

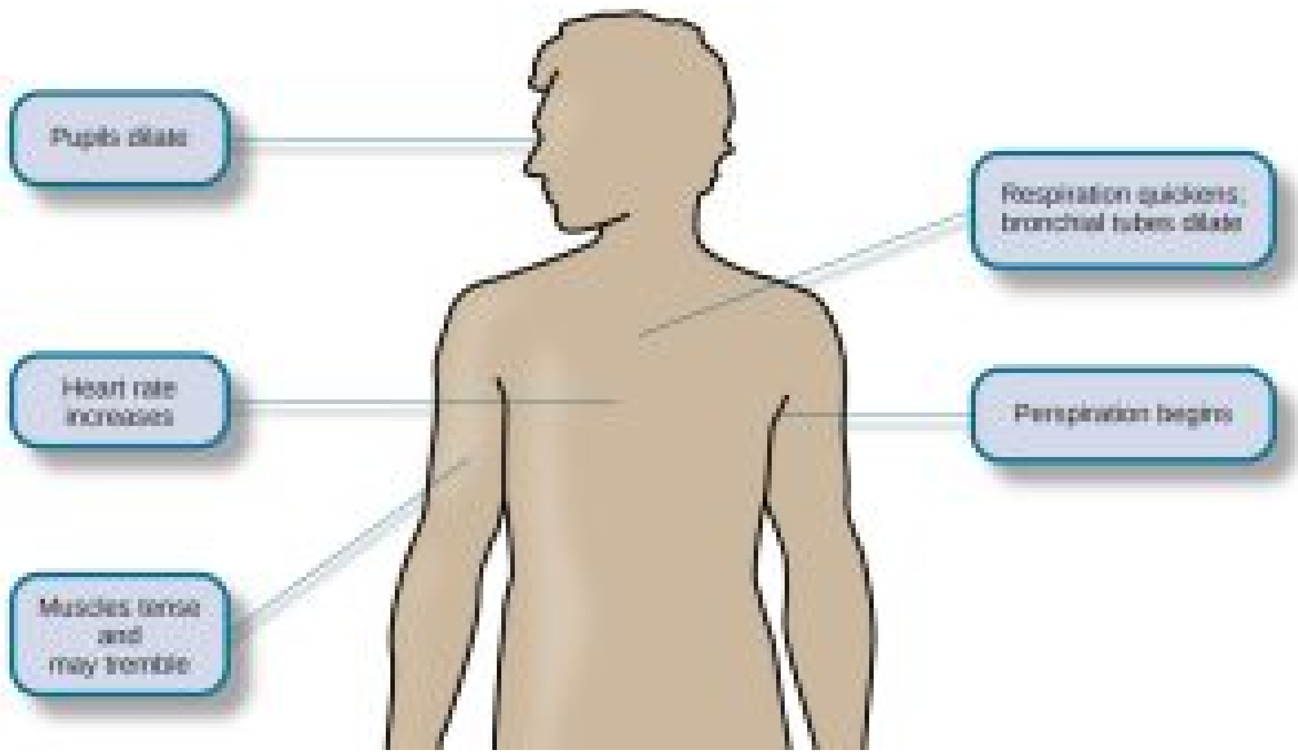


Figure 6. *Fight or flight is a physiological response to a stressor.*

According to Cannon, the fight-or-flight response is a built-in mechanism that assists in maintaining homeostasis—an internal environment in which physiological variables such as blood pressure, respiration, digestion, and temperature are stabilized at levels optimal for survival. Thus, Cannon viewed the fight-or-flight response as adaptive because it enables us to adjust internally and externally to changes in our surroundings, which is helpful in species survival.

Selye and the General Adaptation Syndrome

Another important early contributor to the stress field was Hans **Selye**, mentioned earlier. He would eventually become one of the world’s foremost experts in the study of stress (***Figure 7***). As a young assistant in the biochemistry department at McGill University in the 1930s, Selye was engaged in research involving sex hormones in rats. Although he was unable to find an answer for what he was initially researching, he incidentally discovered that when exposed to prolonged negative stimulation (stressors)—such as extreme cold, surgical injury, excessive muscular exercise, and shock—the rats showed signs of adrenal enlargement, thymus and lymph node shrinkage, and stomach ulceration. Selye realized that these responses were triggered by a coordinated series of physiological reactions that unfold over time during continued exposure to a stressor. These physiological reactions were nonspecific, which means that regardless of the type of stressor, the same pattern of reactions would occur. What Selye discovered was the **general adaptation syndrome**, the body's nonspecific physiological response to stress.

Figure 7. *Hans Selye specialized in research about stress. In 2009, his native Hungary honored his work with this stamp, released in conjunction with the 2nd annual World Conference on Stress.*

The general adaptation syndrome, shown in ***Figure 8***, consists of three stages: (1) alarm reaction, (2) stage of resistance, and (3) stage of exhaustion (Selye, 1936; 1976). **Alarm reaction** describes the body's immediate reaction upon facing a threatening situation or emergency, and it is roughly analogous to the fight-or-flight response described by Cannon. During an alarm reaction, you are alerted to a stressor, and your body alarms you with a cascade of physiological reactions that provide you with the energy to manage the situation. A person who wakes up in the middle of the night to discover her house is on fire, for example, is experiencing an alarm reaction.

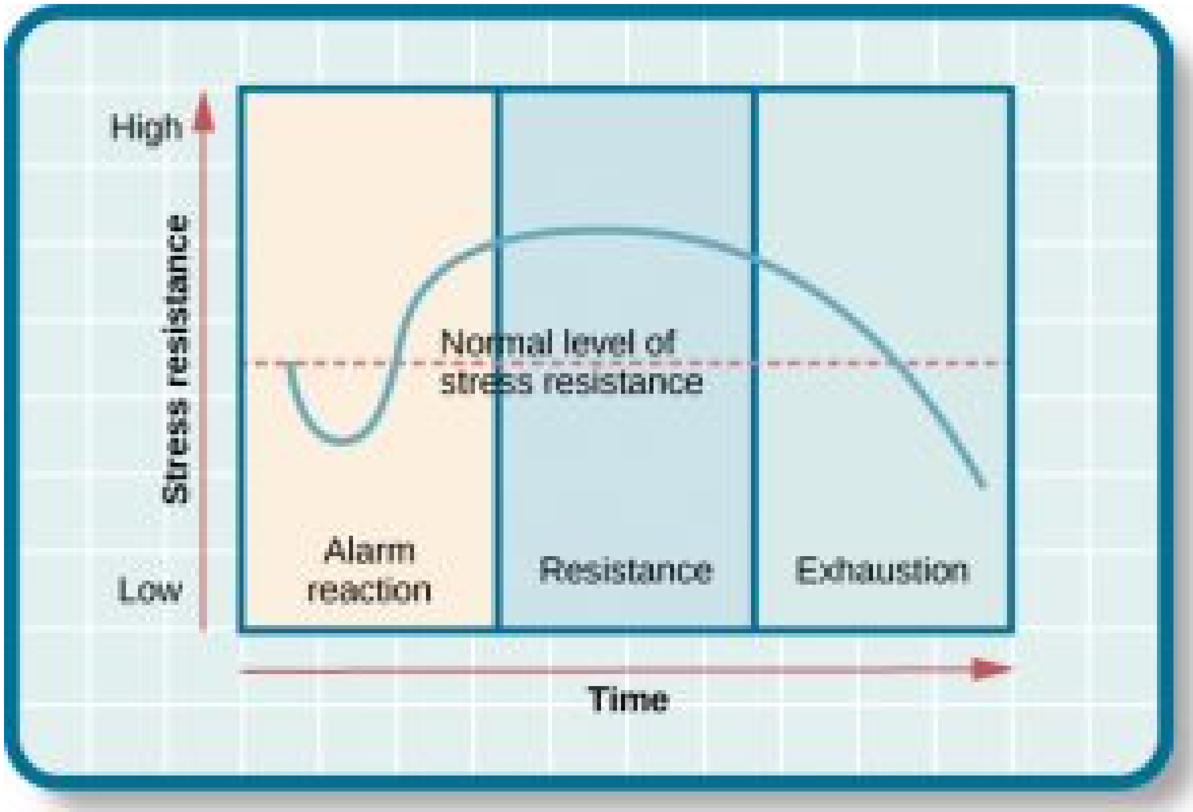


Figure 8. *The three stages of Selye's general adaptation syndrome are shown in this graph. Prolonged stress ultimately results in exhaustion.*

If exposure to a stressor is prolonged, the organism will enter the **stage of resistance**. During this stage, the initial shock of alarm reaction has worn off and the body has adapted to the stressor. Nevertheless, the body also remains on alert and is prepared to respond as it did during the alarm reaction, although with less intensity. For example, suppose a child who went missing is still missing 72 hours later. Although the parents would obviously remain extremely disturbed, the magnitude of physiological reactions would likely have diminished over the 72 intervening hours due to some adaptation to this event.

If exposure to a stressor continues over a longer period of time, the **stage of exhaustion** ensues. At this stage, the person is no longer able to adapt to the stressor: the body's ability to resist becomes depleted as physical wear takes its toll on the body's tissues and organs. As a result, illness, disease, and other permanent damage to the body—even death—may occur. If a missing child still remained missing after three months, the long-term stress associated with this situation may cause a parent to literally faint with exhaustion at some point or even to develop a serious and irreversible illness.

In short, Selye's general adaptation syndrome suggests that stressors tax the body via a three-phase process—an initial jolt, subsequent readjustment, and a later depletion of all physical resources—that ultimately lays the groundwork for serious health problems and even death. It should be pointed out, however, that this model is a response-based conceptualization of stress, focusing exclusively on the body's physical responses while largely ignoring psychological factors such as appraisal and interpretation of threats. Nevertheless, Selye's model has had an enormous impact on the field of stress because it offers a general explanation for how stress can lead to physical damage and, thus, disease. As we shall discuss later, prolonged or repeated stress has been implicated in development of a number of disorders such as hypertension and coronary artery disease.

The Physiological Basis of Stress

What goes on inside our bodies when we experience stress? The physiological mechanisms of stress are extremely complex, but they generally involve the work of two systems—the **sympathetic nervous system** and the **hypothalamic-pituitary-adrenal (HPA) axis**. When a person first perceives something as stressful (Selye's alarm reaction), the sympathetic nervous system triggers arousal via the release of adrenaline from the adrenal glands. Release of these hormones activates the fight-or-flight responses to stress, such as accelerated heart rate and respiration. At the same time, the HPA axis, which is primarily endocrine in nature, becomes especially active, although it works much more slowly than the sympathetic nervous system. In response to stress, the hypothalamus (one of the limbic structures in the brain) releases corticotrophin-releasing factor, a hormone that causes the pituitary gland to release adrenocorticotropic hormone (ACTH) (***Figure 9***). The ACTH then activates the adrenal glands to secrete a number of hormones into the bloodstream; an important one is cortisol, which can affect virtually every organ within the body. Cortisol is commonly known as a stress hormone and helps provide that boost of energy when we first encounter a stressor, preparing us to run away or fight. However, sustained elevated levels of cortisol weaken the immune system.

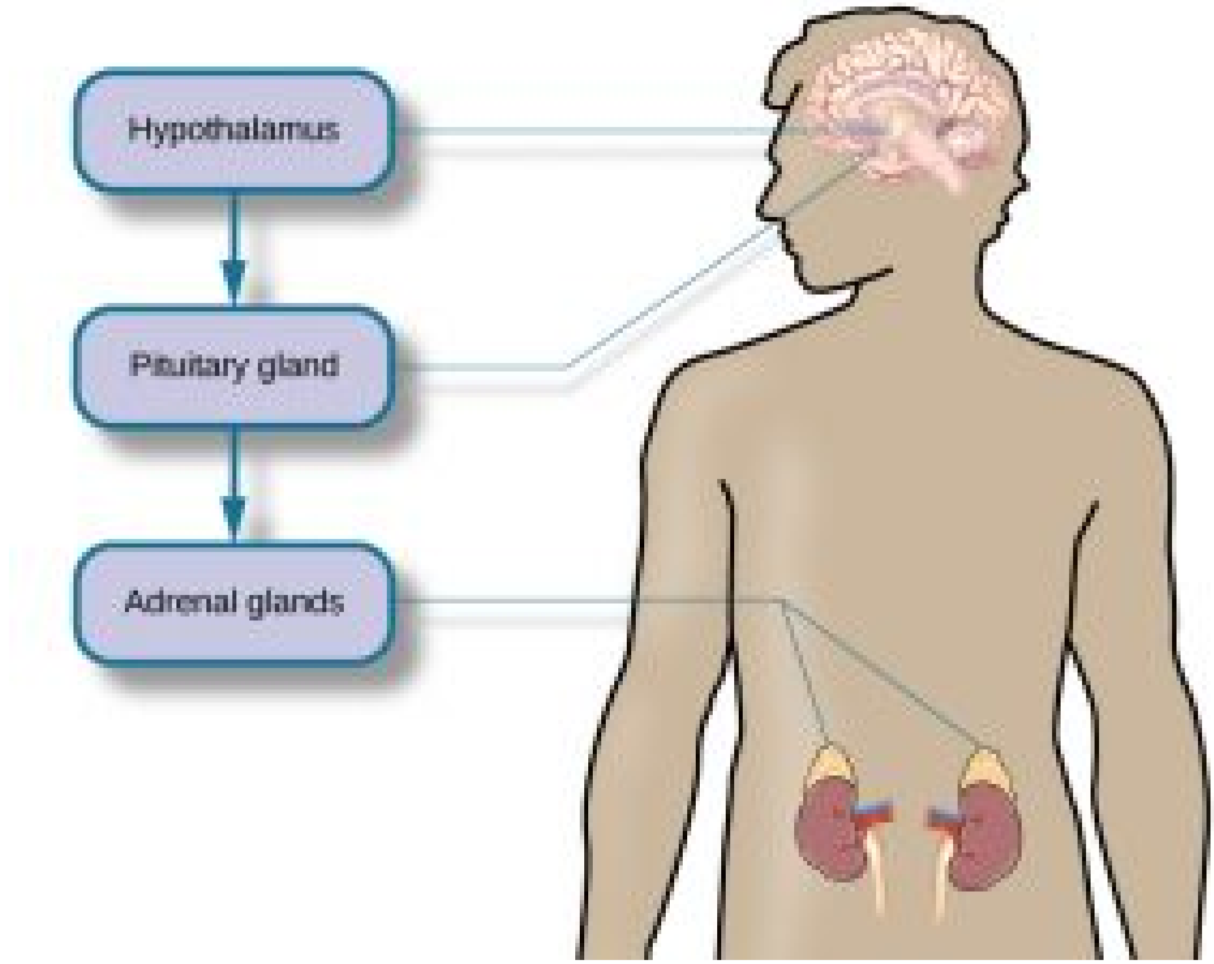


Figure 9. This diagram shows the functioning of the hypothalamic-pituitary-adrenal (HPA) axis. The hypothalamus activates the pituitary gland, which in turn activates the adrenal glands, increasing their secretion of cortisol.

In short bursts, this process can have some favorable effects, such as providing extra energy, improving **immune system** functioning temporarily, and decreasing pain sensitivity. However, extended release of cortisol—as would happen with prolonged or chronic stress—often comes at a high price. High levels of cortisol have been shown to produce a number of harmful effects. For example, increases in cortisol can significantly weaken our immune system (Glaser & Kiecolt-Glaser, 2005), and high levels are frequently observed among depressed individuals (Geoffroy, Hertzman, Li, & Power, 2013). In summary, a stressful event causes a variety of physiological reactions that activate the adrenal glands, which in turn release epinephrine, norepinephrine, and cortisol. These hormones affect a number of bodily processes in ways that prepare the stressed person to take direct action, but also in ways that may heighten the potential for illness.

When stress is extreme or chronic, it can have profoundly negative consequences. For example, stress often contributes to the development of certain psychological disorders, including post-traumatic stress disorder, major depressive disorder, and other serious psychiatric conditions. Additionally, we noted earlier that stress is linked to the development and progression of a variety of physical illnesses and diseases. For example, researchers in one study found that people injured during the September 11, 2001, World Trade Center disaster or who developed post-traumatic stress symptoms afterward later suffered significantly elevated rates of heart disease (Jordan, Miller-Archie, Cone, Morabia, & Stellman, 2011). Another investigation yielded that self-reported stress symptoms among aging and retired Finnish food industry workers were associated with morbidity 11 years later. This study also predicted the onset of musculoskeletal, nervous system, and endocrine and metabolic disorders (Salonen, Arola, Nygård, & Huhtala, 2008). Another study reported that male South Korean manufacturing employees who reported high levels of work-related stress were more likely to catch the common cold over the next several months than were those employees who reported lower work-related stress levels (Park et al., 2011). Later, you will explore the mechanisms through which stress can produce physical illness and disease.

Section Summary

Stress is a process whereby an individual perceives and responds to events appraised as overwhelming or threatening to one's well-being. The scientific study of how stress and emotional factors impact health and well-being is called health psychology, a field devoted to studying the general impact of psychological factors on health. The body's primary physiological response during stress, the fight-or-flight response, was first identified in the early 20th century by Walter Cannon. The fight-or-flight response involves the coordinated activity of both the sympathetic nervous system and the hypothalamic-pituitary-adrenal (HPA) axis. Hans Selye, a noted endocrinologist, referred to these physiological reactions to stress as part of general adaptation syndrome, which occurs in three stages: alarm reaction (fight-or-flight reactions begin), resistance (the body begins to adapt to continuing stress), and exhaustion (adaptive energy is depleted, and stress begins to take a physical toll).

Glossary

alarm reaction
: first stage of the general adaptation syndrome; characterized as the body's immediate physiological reaction to a threatening situation or some other emergency; analogous to the fight-or-flight response

cortisol
: stress hormone released by the adrenal glands when encountering a stressor; helps to provide a boost of energy, thereby preparing the individual to take action

distress
: bad form of stress; usually high in intensity; often leads to exhaustion, fatigue, feeling burned out; associated with erosions in performance and health

eustress
: good form of stress; low to moderate in intensity; associated with positive feelings, as well as optimal health and performance

fight-or-flight response
: set of physiological reactions (increases in blood pressure, heart rate, respiration rate, and sweat) that occur when an individual encounters a perceived threat; these reactions are produced by activation of the sympathetic nervous system and the endocrine system

general adaptation syndrome
Hans Selye's three-stage model of the body's physiological reactions to stress and the process of stress adaptation: alarm reaction, stage of resistance, and stage of exhaustion

health psychology
subfield of psychology devoted to studying psychological influences on health, illness, and how people respond when they become ill

hypothalamic-pituitary-adrenal (HPA) axis
set of structures found in both the limbic system (hypothalamus) and the endocrine system (pituitary gland and adrenal glands) that regulate many of the body's physiological reactions to stress through the release of hormones

primary appraisal
judgment about the degree of potential harm or threat to well-being that a stressor might entail

secondary appraisal
judgment of options available to cope with a stressor and their potential effectiveness

stage of exhaustion
third stage of the general adaptation syndrome; the body's ability to resist stress becomes depleted; illness, disease, and even death may occur

stage of resistance
second stage of the general adaptation syndrome; the body adapts to a stressor for a period of time

stress
process whereby an individual perceives and responds to events that one appraises as overwhelming or threatening to one's well-being

stressors
environmental events that may be judged as threatening or demanding; stimuli that initiate the stress process

24.

3.2 Parts of the Nervous System

Learning Objectives

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

- Describe the difference between the central and peripheral nervous systems
- Explain the difference between the somatic and autonomic nervous systems
- Differentiate between the sympathetic and parasympathetic divisions of the autonomic nervous system

The **nervous system** can be divided into two major subdivisions: the **central nervous system (CNS)** and the **peripheral nervous system (PNS)**, shown in ***Figure 1***. The CNS is comprised of the brain and spinal cord; the PNS connects the CNS to the rest of the body. In this section, we focus on the peripheral nervous system; later, we look at the brain and spinal cord.

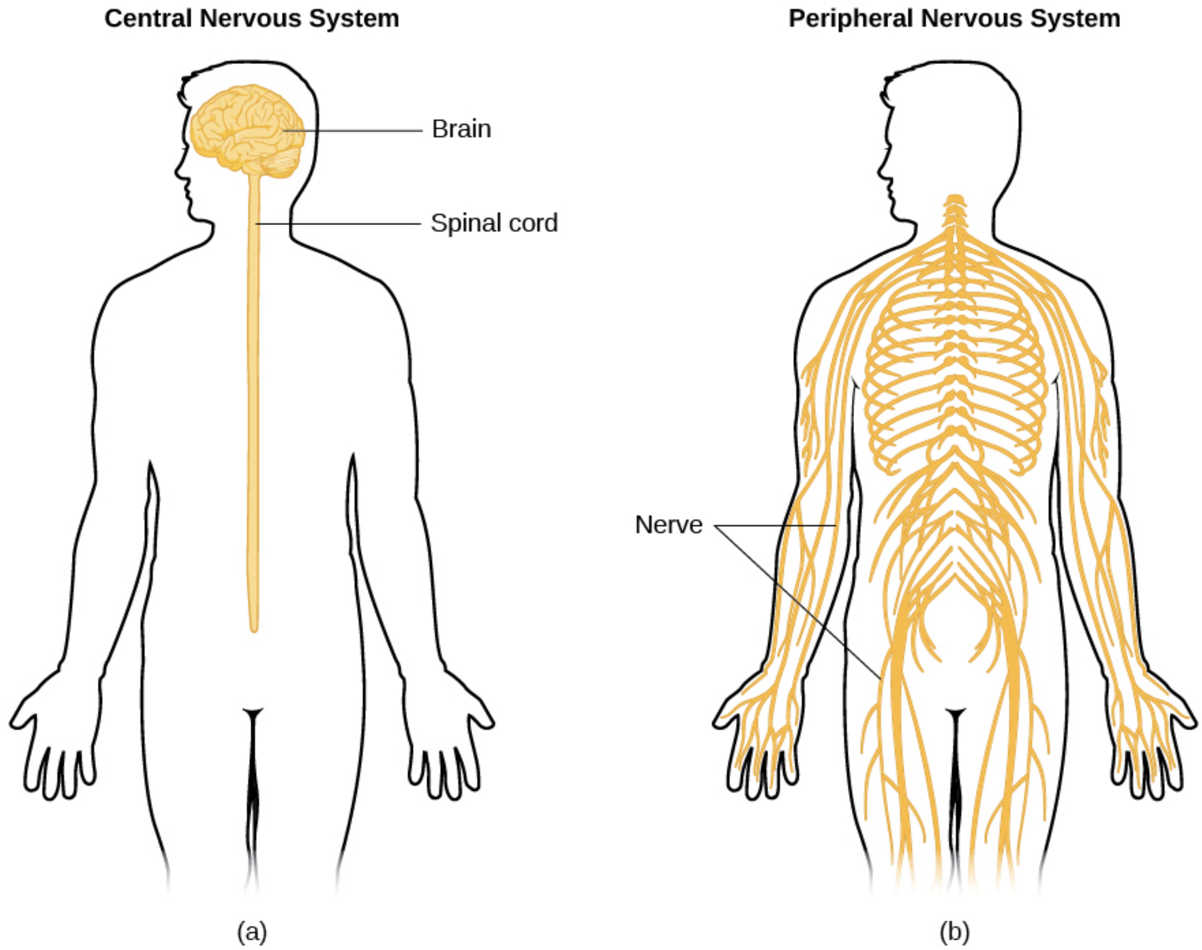


Figure 1. *The nervous system is divided into two major parts:* ***(a)*** *the Central Nervous System and* ***(b)*** *the Peripheral Nervous System.*

Peripheral Nervous System

The peripheral nervous system is made up of thick bundles of axons, called nerves, carrying messages back and forth between the CNS and the muscles, organs, and senses in the periphery of the body (i.e., everything outside the CNS). The PNS has two major subdivisions: the somatic nervous system and the autonomic nervous system.

The **somatic nervous system** is associated with activities traditionally thought of as conscious or voluntary. It is involved in the relay of sensory and motor information to and from the CNS; therefore, it consists of motor neurons and sensory neurons. Motor neurons, carrying instructions from the CNS to the muscles, are efferent fibers (efferent means “moving away from”). Sensory neurons, carrying sensory information to the CNS, are afferent fibers (afferent means “moving toward”). Each nerve is basically a two-way superhighway, containing thousands of axons, both efferent and afferent.

The **autonomic nervous system** controls our internal organs and glands and is generally considered to be outside the realm of voluntary control. It can be further subdivided into the sympathetic and parasympathetic divisions (***Figure 2***). The **sympathetic nervous system** is involved in preparing the body for stress-related activities; the **parasympathetic nervous system** is associated with returning the body to routine, day-to-day operations. The two systems have complementary functions, operating in tandem to maintain the body's homeostasis. **Homeostasis** is a state of equilibrium, in which biological conditions (such as body temperature) are maintained at optimal levels.

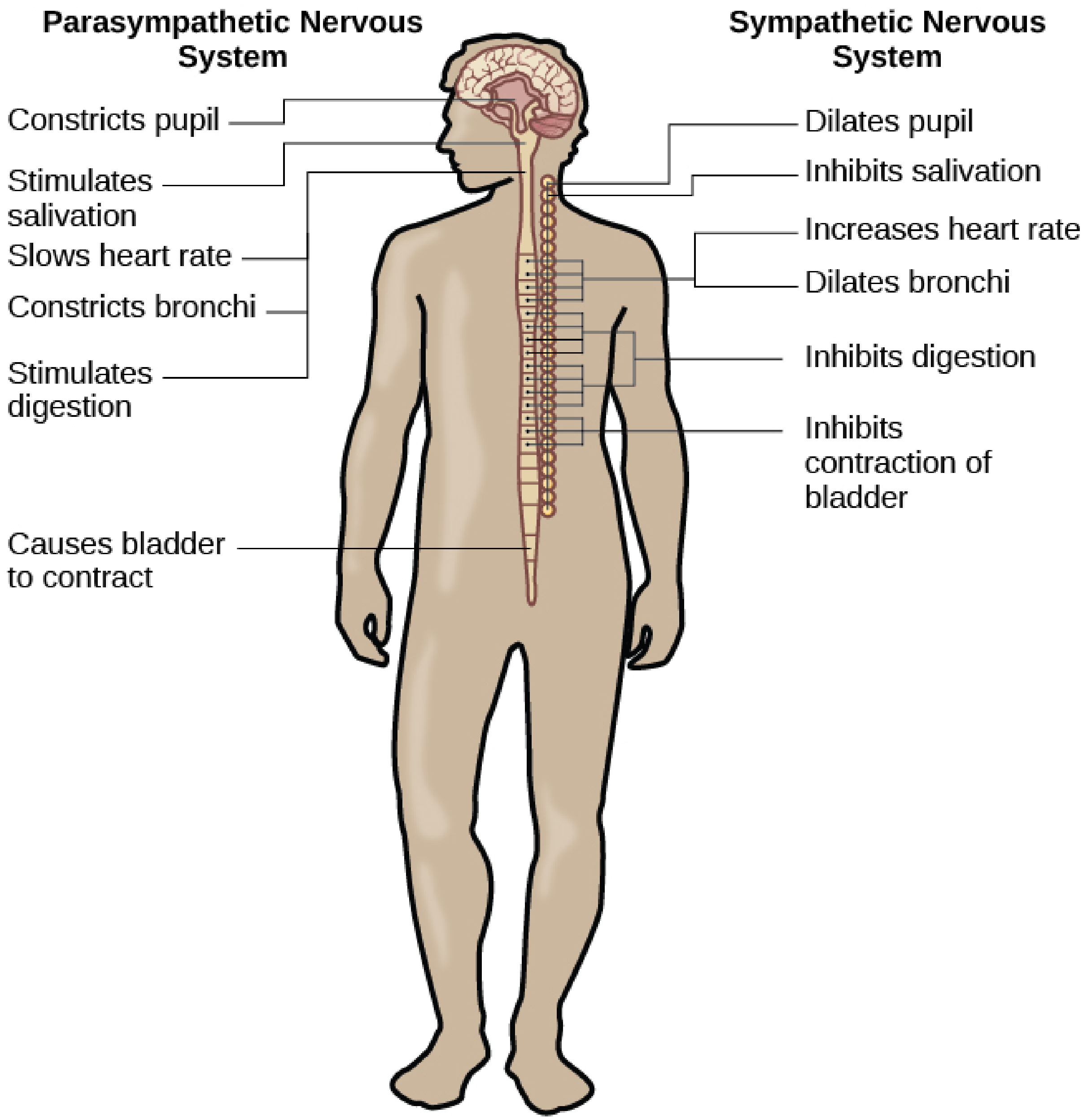


Figure 2. *The sympathetic and parasympathetic divisions of the autonomic nervous system have the opposite effects on various systems.*

The sympathetic nervous system is activated when we are faced with stressful or high-arousal situations. The activity of this system was adaptive for our ancestors, increasing their chances of survival. Imagine, for example, that one of our early ancestors, out hunting small game, suddenly disturbs a large bear with her cubs. At that moment, his body undergoes a series of changes—a direct function of sympathetic activation—preparing him to face the threat. His pupils dilate, his heart rate and blood pressure increase, his bladder relaxes, his liver releases glucose, and adrenaline surges into his bloodstream. This constellation of physiological changes, known as the **fight or flight response**, allows the body access to energy reserves and heightened sensory capacity so that it might fight off a threat or run away to safety.

While it is clear that such a response would be critical for survival for our ancestors, who lived in a world full of real physical threats, many of the high-arousal situations we face in the modern world are more psychological in nature. For example, think about how you feel when you have to stand up and give a presentation in front of a roomful of people, or right before taking a big test. You are in no real physical danger in those situations, and yet you have evolved to respond to any perceived threat with the **fight or flight response**. This kind of response is not nearly as adaptive in the modern world; in fact, we suffer negative health consequences when faced constantly with psychological threats that we can neither fight nor flee. Recent research suggests that an increase in susceptibility to heart disease (Chandola, Brunner, & Marmot, 2006) and impaired function of the immune system (Glaser & Kiecolt-Glaser, 2005) are among the many negative consequences of persistent and repeated exposure to stressful situations.

Once the threat has been resolved, the parasympathetic nervous system takes over and returns bodily functions to a relaxed state. Our hunter's heart rate and blood pressure return to normal, his pupils constrict, he regains control of his bladder, and the liver begins to store glucose in the form of glycogen for future use. These processes are associated with activation of the parasympathetic nervous system.

Section Summary

The brain and spinal cord make up the central nervous system. The peripheral nervous system is comprised of the somatic and autonomic nervous systems. The somatic nervous system transmits sensory and motor signals to and from the central nervous system. The autonomic nervous system controls the function of our organs and glands, and can be divided into the sympathetic and parasympathetic divisions. Sympathetic activation prepares us for fight or flight, while parasympathetic activation is associated with normal functioning under relaxed conditions.

Glossary

autonomic nervous system
: controls our internal organs and glands

central nervous system (CNS)
: brain and spinal cord

fight or flight response
: activation of the sympathetic division of the autonomic nervous system, allowing access to energy reserves and heightened sensory capacity so that we might fight off a given threat or run away to safety

homeostasis
: state of equilibrium—biological conditions, such as body temperature, are maintained at optimal levels

parasympathetic nervous system
: associated with routine, day-to-day operations of the body

peripheral nervous system (PNS)
: connects the brain and spinal cord to the muscles, organs and senses in the periphery of the body

somatic nervous system
: relays sensory and motor information to and from the CNS

sympathetic nervous system
: involved in stress-related activities and functions

25.

3.3 Atherosclerosis, blood lipids, and stress

Learning Objectives

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

- Differentiate between HDL and LDL. Which is the “good” cholesterol?
- Understand what atherosclerosis is and how it relates to HDL, LDL and triglycerides? (See youtube video link below)
- Recommend lifestyle changes to reduce the risk of atherosclerosis (genetics also play a big role in this disease).

Heart Disease is a Leading Cause of Death in the United States

According to the Center for Disease control and prevention (CDC.gov), heart disease is responsible for 1 in 4 deaths in theUnited States and is the leading cause of death for both men and women. ***Table 1*** shows the percentage of deaths due to heart disease grouped by race/ethnicity. Hispanics are excluded from this table, but heart disease is the number one cause of death for this group.

Race or Ethnic Group	% of deaths caused by heart disease
American Indians or Alaska Natives	18.4
Asians or Pacific Islanders	22.2
Non-Hispanic Blacks	23.8
Non-Hispanic Whites	23.8

Table 1. *Percent of deaths caused by heart disease, grouped into race or ethnic groups.*

Heart Disease Risk Factors

Age (over 45 for men, 55 for women) is an important risk factor for heart disease. People who have diabetes, are overweight, have limited physical activity or use alcohol excessively are at particular risk. While leading a healthy lifestyle can certainly reduce a person's heart disease risk, it's also important to realize that genetics play a role, and that a family history of heart disease is a major risk factor. Geographic patterns of heart disease in the United States are affected by these risk factors, as well as access to regular health screenings and treatment. The map in ***Figure 1*** shows heart disease death rates by region for 2014-2016, with dark red indicating the highest rates of death, light pink, the lowest rates of death.

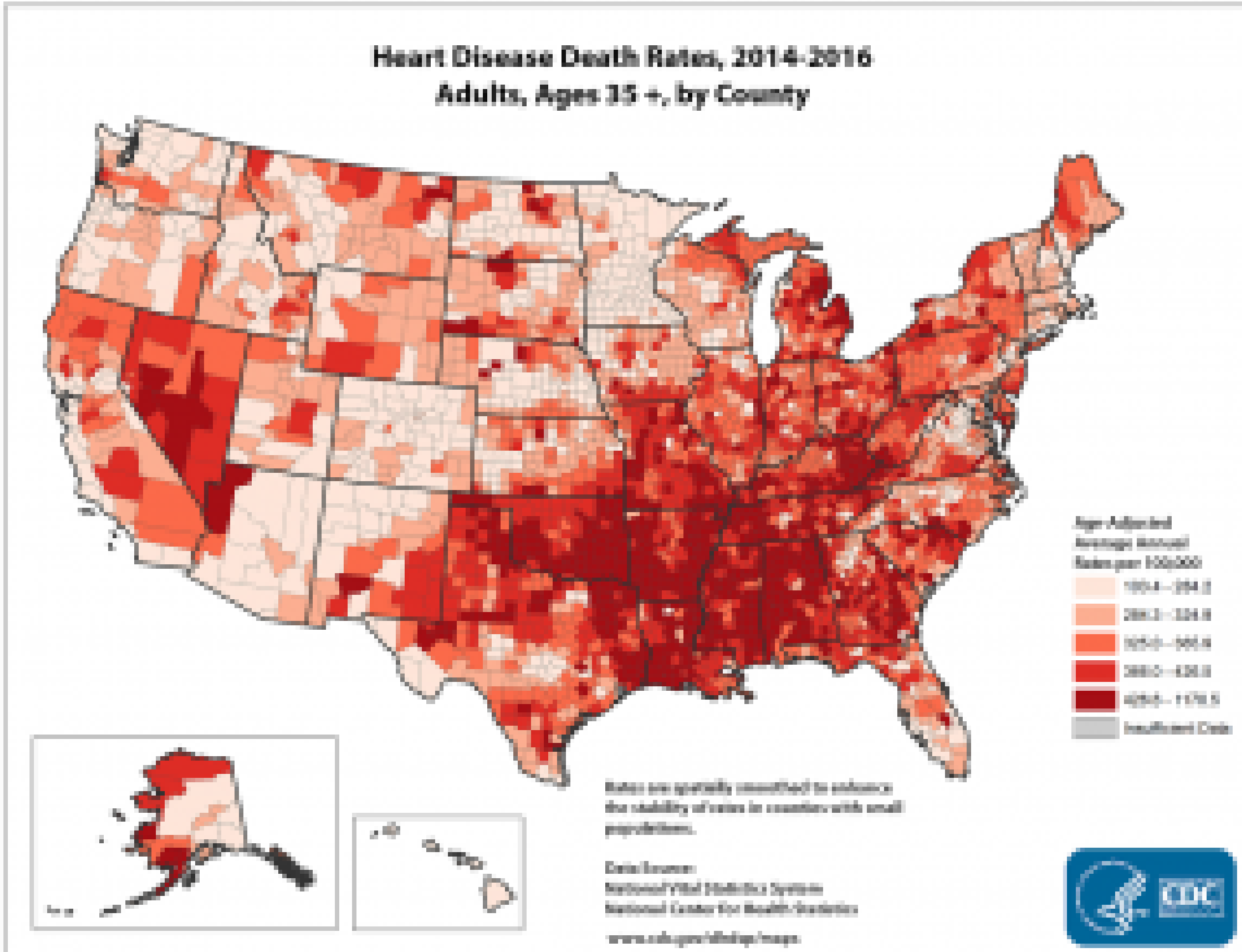


Figure 1. *Heart disease death rates by country.*

Heart Disease and Blood Lipids

Many people with heart disease likely don't know they have it. However, for people who are able to get regular, physical exams, some screening does take place. For example, it is typical for a medical practitioner to take your blood pressure at all medical exams. High systolic blood pressure can be a sign of heart disease, because it may suggest that arteries are narrowing or hardening. It is also typical to listen to the heart for abnormal heart sounds. Adults may also have their blood tested for the presence of certain lipids, in particular high density lipoprotein (HDL), low density lipoprotein (LDL) and triglycerides (***Table 2***). When people talk about "cholesterol" they are talking about these lipoproteins.

	Total cholesterol	LDL cholesterol	HDL cholesterol	Triglycerides
Children, 0-9				
Ideal	< 170	< 110	> 45	< 75
At risk	≥ 200	≥ 130	< 40	≥ 100
Children, 10-19				
Ideal	< 170	< 110	> 45	< 90
At risk	≥ 200	≥ 130	< 40	≥ 130
Adults				
Ideal	< 200	< 100	≥ 60	< 149
At risk	≥ 240	≥ 160	< 40	≥ 200

Table 2. *Current guidelines on blood lipid levels by The American College of Cardiology, summarized by Fairview.org.*

HDL, LDL, Triglycerides and Atherosclerosis

Watch this videoclip for a really good overview of these lipoproteins, how they work, and why they are important. You do not need to memorize information about the chemical structure of these lipoproteins, but should instead focus on their impact on atherosclerosis and cardiovascular health. After watching, you should be able to do the following:

- Differentiate between HDL and LDL. Which is the "good" cholesterol?
- What is atherosclerosis and how does it relate to HDL and LDL?
- Recommend lifestyle changes to reduce the risk of atherosclerosis

In lecture, we will look at data and discuss how stress may affect these lipoproteins.

26.

3.4 Biological Molecules

The large molecules necessary for life that are built from smaller organic molecules are called biological **macromolecules**. 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 (with some exceptions, like carbon dioxide). In addition, they may contain hydrogen, oxygen, nitrogen, phosphorus, sulfur, and additional minor elements.

Carbon

It is often said that life is "carbon-based." This means that carbon atoms, bonded to other carbon atoms or other elements, form the fundamental components of many, if not most, of the molecules found uniquely in living things. Other elements play important roles in biological molecules, but carbon certainly qualifies as the "foundation" element for molecules in living things. It is the bonding properties of carbon atoms that are responsible for its important role.

Carbon Bonding

Carbon contains four electrons in its outer shell. Therefore, it can form four covalent bonds with other atoms or molecules. The simplest organic carbon molecule is methane (CH_4), in which four hydrogen atoms bind to a carbon atom (***Figure 1***).

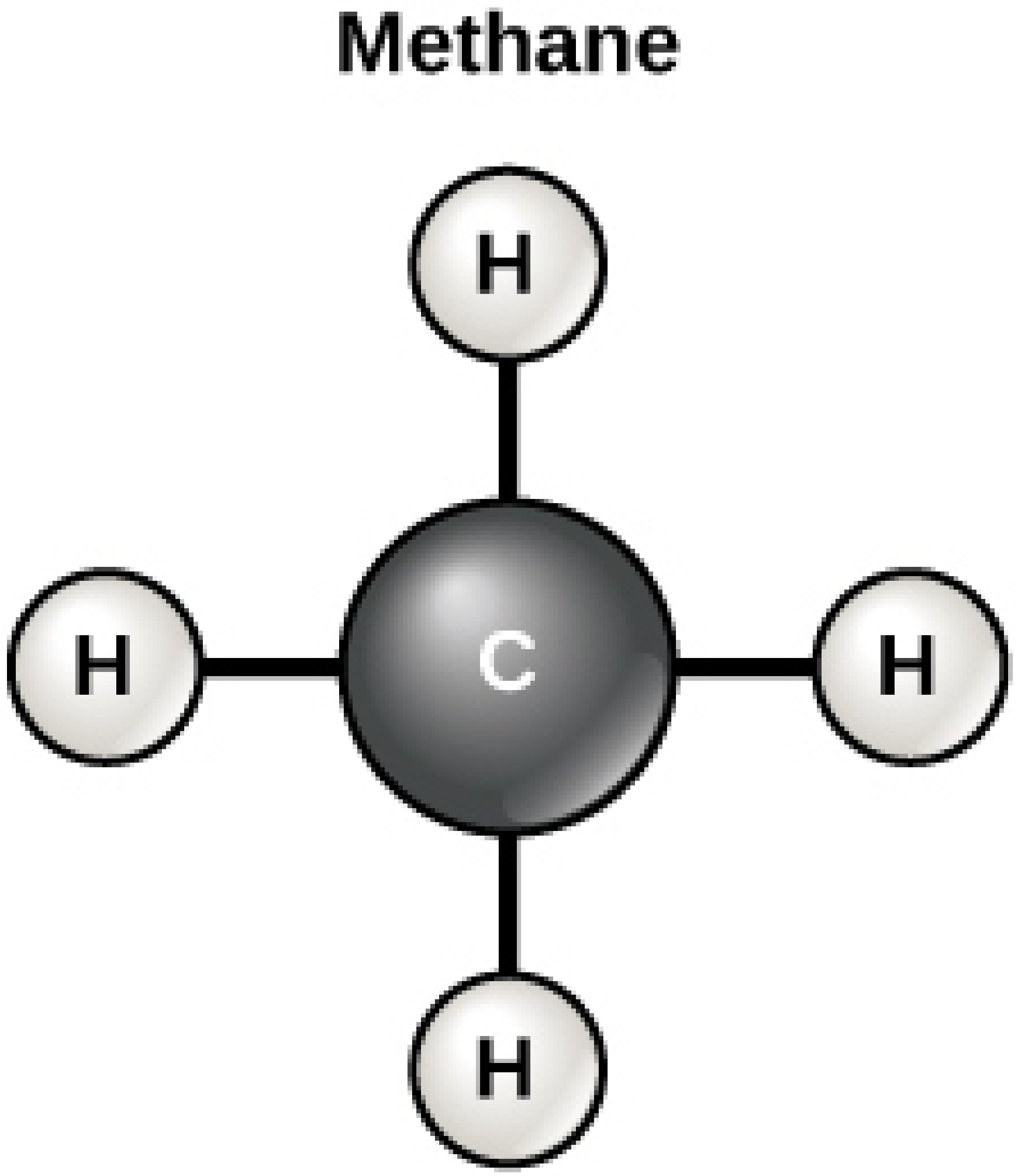


Figure 1. *Carbon can form four covalent bonds to create an organic molecule. The simplest carbon molecule is methane (CH4), depicted here.*

However, structures that are more complex are made using carbon. Any of the hydrogen atoms can be replaced with another carbon atom covalently bonded to the first carbon atom. In this way, long and branching chains of carbon compounds can be made (***Figure 2a***). The carbon atoms may bond with atoms of other elements, such as nitrogen, oxygen, and phosphorus (***Figure 2b***). The molecules may also form rings, which themselves can link with other rings (***Figure 2c***). This diversity of molecular forms accounts for the diversity of functions of the biological macromolecules and is based to a large degree on the ability of carbon to form multiple bonds with itself and other atoms.

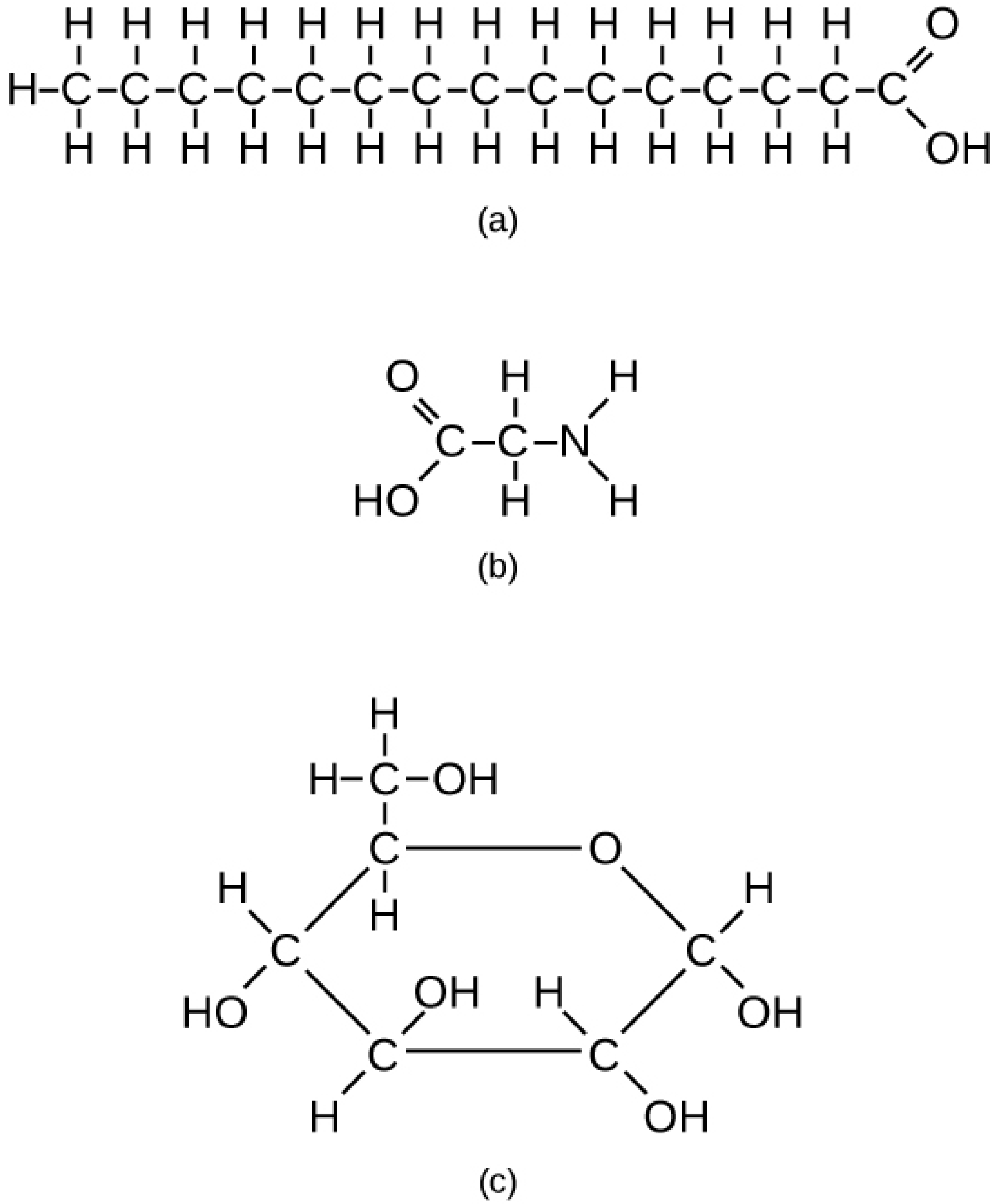


Figure 2. *These examples show three molecules (found in living organisms) that contain carbon atoms bonded in various ways to other carbon atoms and the atoms of other elements. (a) This molecule of stearic acid has a long chain of carbon atoms. (b) Glycine, a component of proteins, contains carbon, nitrogen, oxygen, and hydrogen atoms. (c) Glucose, a sugar, has a ring of carbon atoms and one oxygen atom.*

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.

Carbohydrates can be represented by the formula $(CH_2O)_n$, where *n* is the number of carbon atoms in the molecule. In other words, the ratio of carbon to hydrogen to oxygen is 1:2:1 in carbohydrate molecules. Carbohydrates are classified into three subtypes: monosaccharides, disaccharides, and polysaccharides.

Monosaccharides (mono- = “one”; sacchar- = “sweet”) are simple sugars, the most common of which is glucose. In monosaccharides, the number of carbon atoms usually ranges from three to six. Most monosaccharide names end with the suffix -ose. Depending on the number of carbon atoms in the sugar, they may be known as trioses (three carbon atoms), pentoses (five carbon atoms), and hexoses (six carbon atoms).

Monosaccharides may exist as a linear chain or as ring-shaped molecules; in aqueous solutions, they are usually found in the ring form.

The chemical formula for glucose is $C_6H_{12}O_6$. In most living species, 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 by the process of photosynthesis, and the glucose, in turn, is used for the energy requirements of the plant. The excess synthesized glucose is often stored as starch that is broken down by other organisms that feed on plants.

Galactose (part of lactose, or milk sugar) and fructose (found 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 differing arrangements of atoms in the carbon chain (***Figure 3***).

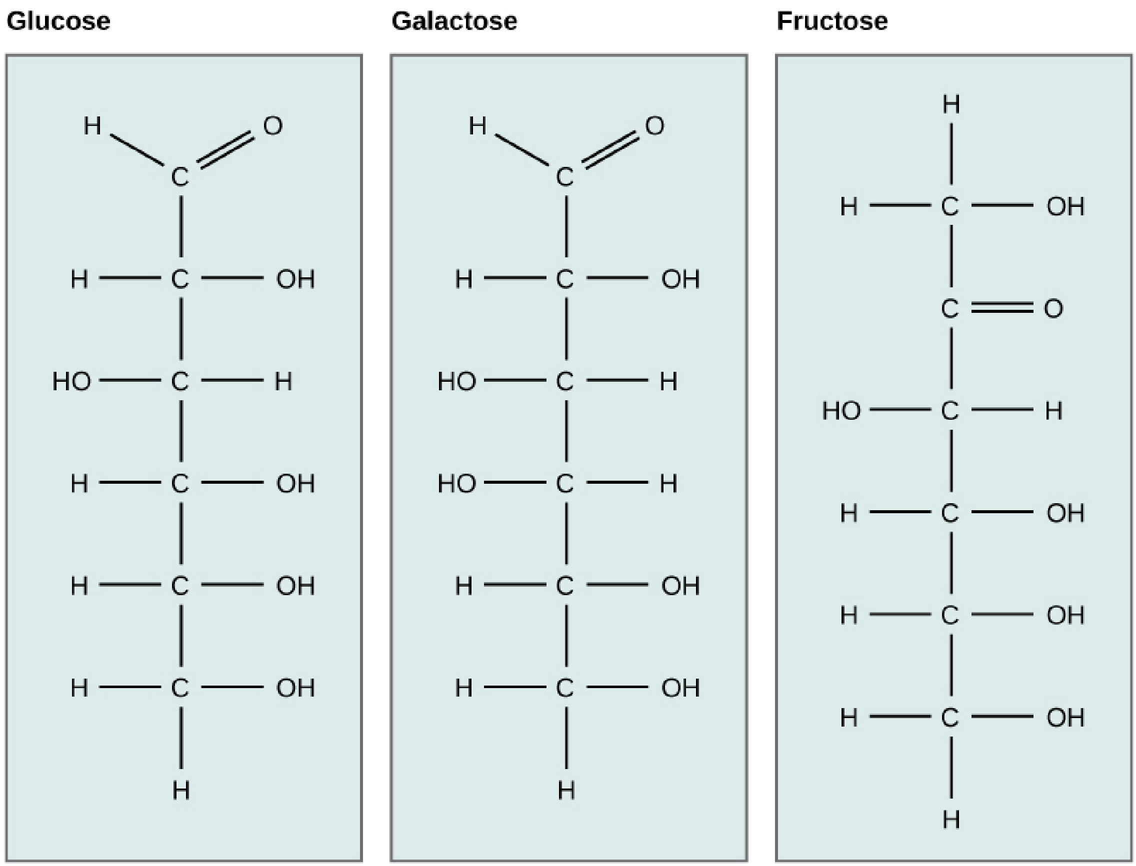


Figure 3. *Glucose, galactose, and fructose are isomeric monosaccharides, meaning that they have the same chemical formula but slightly different structures.*

Disaccharides (di- = “two”) form when two monosaccharides undergo a dehydration reaction (a reaction in which the removal of a water molecule occurs). During this process, the hydroxyl group (–OH) of one monosaccharide combines with a hydrogen atom of another monosaccharide, releasing a molecule of water (H_2O) and forming a covalent bond between atoms in the two sugar molecules.

Common disaccharides include lactose, maltose, and sucrose. Lactose is a disaccharide consisting of the monomers glucose and galactose. It is found naturally in milk. Maltose, or malt sugar, is a disaccharide formed from a dehydration reaction between two glucose molecules. The most common disaccharide is sucrose, or table sugar, which is composed of the monomers glucose and fructose.

A long chain of monosaccharides linked by covalent bonds is known as a **polysaccharide** (poly- = “many”). The chain may be branched or unbranched, and it may contain different types of monosaccharides. Polysaccharides may be very large molecules. Starch, glycogen, cellulose, and chitin are examples of polysaccharides.

Starch is the stored form of sugars in plants and is made up of amylose and amylopectin (both polymers of glucose). Plants are able to synthesize glucose, and the excess glucose is stored as starch in different plant parts, including roots and seeds. The starch that is consumed by animals is broken down into smaller molecules, such as 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 glucose levels decrease, glycogen is broken down to release glucose.

Cellulose is one of the most abundant natural biopolymers. The cell walls of plants are mostly made of cellulose, which provides structural support to the cell. Wood and paper are mostly cellulosic in nature. Cellulose is made up of glucose monomers that are linked by bonds between particular carbon atoms in the glucose molecule.

Every other glucose monomer in cellulose is flipped over and packed tightly as extended long chains. This gives cellulose its rigidity and high tensile strength—which is so important to plant cells. Cellulose passing through our digestive system is called dietary fiber. While the glucose-glucose bonds in cellulose cannot be broken down by human digestive enzymes, herbivores such as cows, buffalos, and horses are able to digest grass that is rich in cellulose and use it as a food source. In these animals, certain species of bacteria reside in the rumen (part of the digestive system of herbivores) and secrete the enzyme cellulase. The appendix also contains bacteria that break down cellulose, giving it an important role in the digestive systems of ruminants. Cellulases can break down cellulose into glucose monomers that can be used as an energy source by the animal.

Carbohydrates serve other functions in different animals. Arthropods, such as insects, spiders, and crabs, have an outer skeleton, called the exoskeleton, which protects their internal body parts. This exoskeleton is made of the biological macromolecule **chitin**, which is a nitrogenous carbohydrate. It is made of repeating units of a modified sugar containing nitrogen.

Thus, through differences in molecular structure, carbohydrates are able to serve the very different functions of energy storage (starch and glycogen) and structural support and protection (cellulose and chitin) (***Figure 4***).

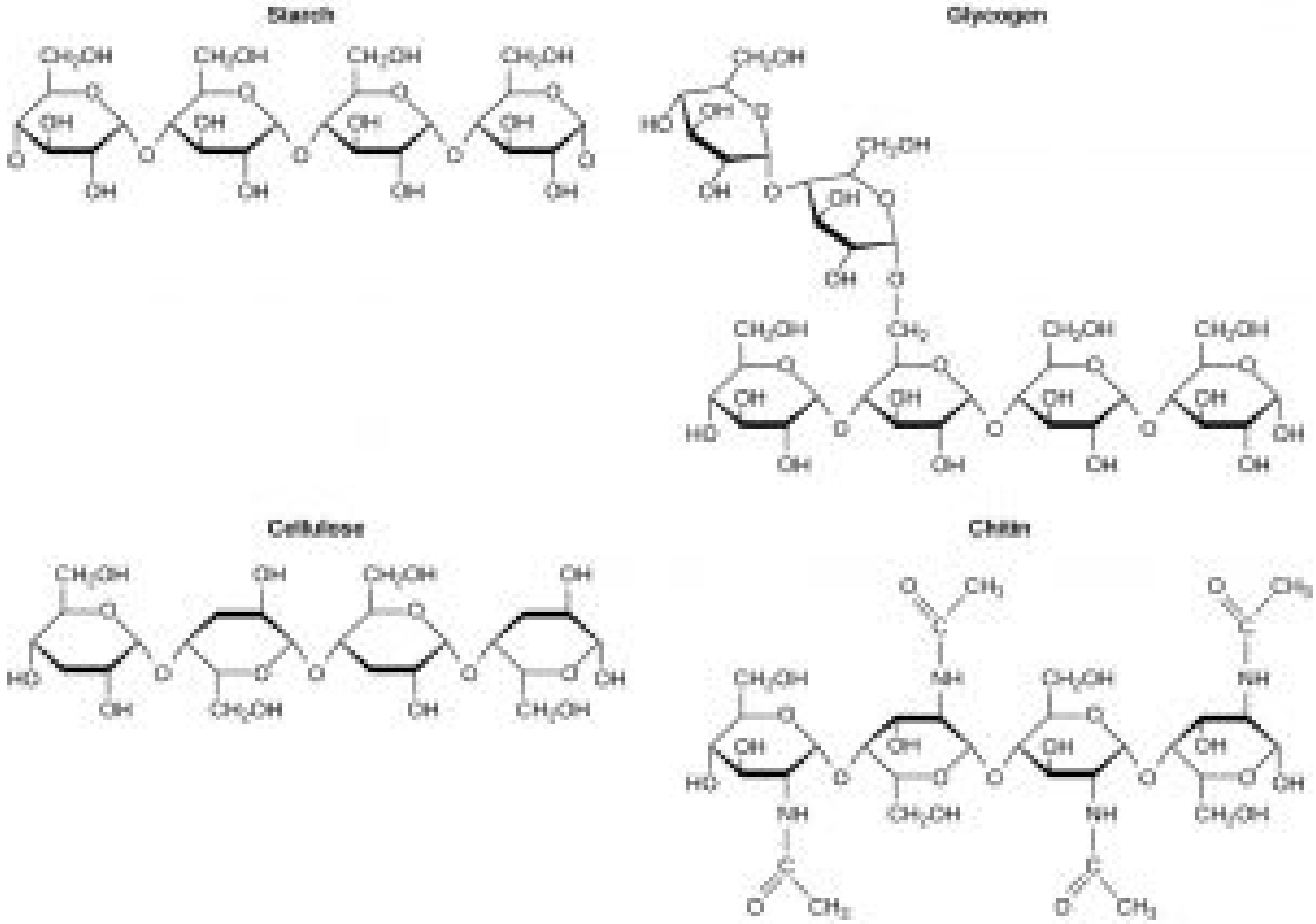


Figure 4. *Although their structures and functions differ, all polysaccharide carbohydrates are made up of monosaccharides and have the chemical formula (CH2O)n.*

CAREERS IN ACTION

Registered Dietitian

Obesity is a worldwide health concern, and many diseases, such as diabetes and heart disease, are becoming more prevalent because of obesity. This is one of the reasons why registered dietitians are increasingly sought after for advice. Registered dietitians help plan food and nutrition programs for individuals in various settings. They often work with patients in health-care facilities, designing nutrition plans to prevent and treat diseases. For example, dietitians may teach a patient with diabetes how to manage blood-sugar levels by eating the correct types and amounts of carbohydrates. Dietitians may also work in nursing homes, schools, and private practices.

To become a registered dietitian, one needs to earn at least a bachelor's degree in dietetics, nutrition, food technology, or a related field. In addition, registered dietitians must complete a supervised internship program and pass a national exam. Those who pursue careers in dietetics take courses in nutrition, chemistry, biochemistry, biology, microbiology, and human physiology. Dietitians must become experts in the chemistry and functions of food (proteins, carbohydrates, and fats).

Lipids

Lipids include a diverse group of compounds that are united by a common feature. **Lipids** are hydrophobic ("water-fearing"), or insoluble in water, because they are nonpolar molecules. This is because they are hydrocarbons that include only nonpolar carbon-carbon or carbon-hydrogen bonds. 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 (***Figure 5***). 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.

Figure 5. *Hydrophobic lipids in the fur of aquatic mammals, such as this river otter, protect them from the elements. (credit: Ken Bosma)*

A **fat** molecule, such as a triglyceride, consists of two main components—glycerol and fatty acids. Glycerol is an organic compound with three carbon atoms, five hydrogen atoms, and three hydroxyl (–OH) groups. Fatty acids have a long chain of hydrocarbons to which an acidic 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, a fatty acid is attached to each of the three oxygen atoms in the –OH groups of the glycerol molecule with a covalent bond (***Figure 6***).

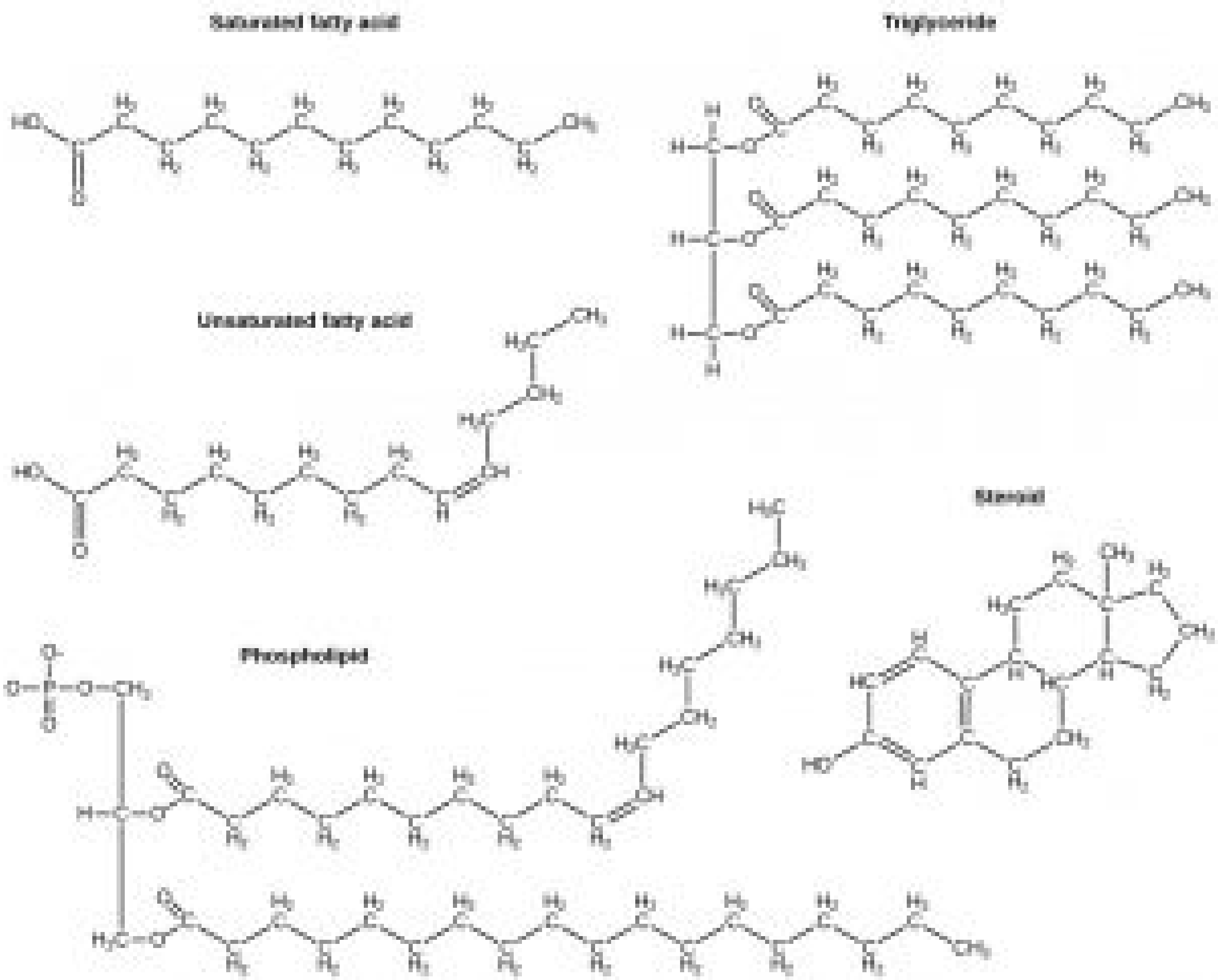


Figure 6. *Lipids include fats, such as triglycerides, which are made up of fatty acids and glycerol, phospholipids, and steroids.*

During this covalent bond formation, three water molecules are released. The three fatty acids in the fat may be similar or dissimilar. These fats are also called **triglycerides** because they have three fatty acids. Some fatty acids have common names that specify their origin. For example, palmitic acid, a saturated fatty acid, is derived from the palm tree. Arachidic acid is derived from *Arachis hypogaea*, the scientific name for peanuts.

Fatty acids may be saturated or unsaturated. In a fatty acid chain, if there are only single bonds between neighboring carbons in the hydrocarbon chain, the fatty acid is saturated. **Saturated fatty acids** are saturated with hydrogen; in other words, the number of hydrogen atoms attached to the carbon skeleton is maximized.

When the hydrocarbon chain contains a double bond, the fatty acid is an **unsaturated fatty acid**.

Most unsaturated fats are liquid at room temperature and are called oils. If there is one double bond in the molecule, then it is known as a monounsaturated fat (e.g., olive oil), and if there is more than one double bond, then it is known as a polyunsaturated fat (e.g., canola oil).

Saturated fats tend to get packed tightly and are solid at room temperature. Animal fats with stearic acid and palmitic acid contained in meat, and the fat with butyric acid contained in butter, are examples of saturated fats. Mammals store fats in specialized cells called adipocytes, where globules of fat occupy most of the cell. In plants, fat or oil is stored in seeds and is used as a source of energy during embryonic development.

Unsaturated fats or oils are usually of plant origin and contain unsaturated fatty acids. The double bond causes a bend or a "kink" that prevents the fatty acids from packing tightly, keeping them liquid at room temperature. Olive oil, corn oil, canola oil, and cod liver oil are examples of unsaturated fats. Unsaturated fats help to improve blood cholesterol levels, whereas saturated fats contribute to plaque formation in the arteries, which increases the risk of a heart attack.

In the food industry, oils are artificially hydrogenated to make them semi-solid, leading to less spoilage and increased shelf life. Simply speaking, hydrogen gas is bubbled through oils to solidify them. During this hydrogenation process, double bonds of the *cis*-conformation in the hydrocarbon chain may be converted to double bonds in the *trans*-conformation. This forms a ***trans*-fat** from a *cis*-fat. The orientation of the double bonds affects the chemical properties of the fat (***Figure 7***).

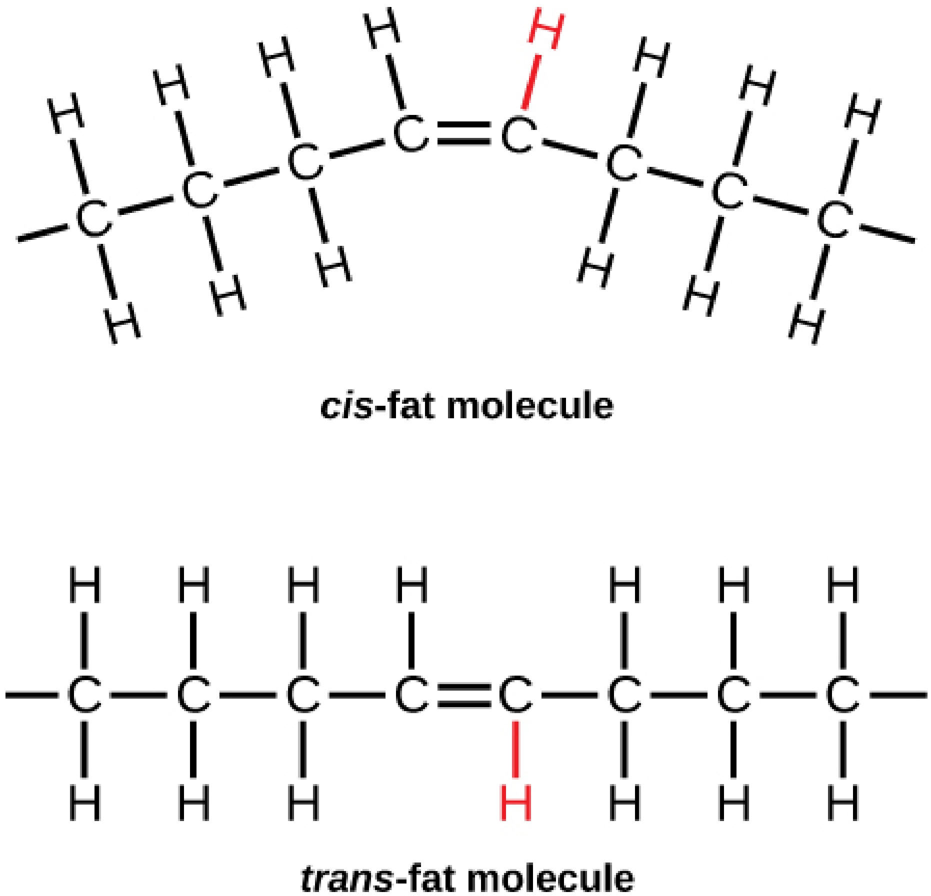


Figure 7. *During the hydrogenation process, the orientation around the double bonds is changed, making a trans-fat from a cis-fat. This changes the chemical properties of the molecule.*

Margarine, some types of peanut butter, and shortening are examples of artificially hydrogenated *trans*-fats. Recent studies have shown that an increase in *trans*-fats in the human diet may lead to an increase in levels of low-density lipoprotein (LDL), or "bad" cholesterol, which, in turn, may lead to plaque deposition in the arteries, resulting in heart disease. Many fast food restaurants have recently eliminated the use of *trans*-fats, and U.S. food labels are now required to list their *trans*-fat content.

Essential fatty acids are fatty acids that are required but not synthesized by the human body. Consequently, they must be supplemented through the diet. Omega-3 fatty acids fall into this category and are one of only two known essential fatty acids for humans (the other being omega-6 fatty acids). They are a type of polyunsaturated fat and are called omega-3 fatty acids because the third carbon from the end of the fatty acid participates in a double bond.

Salmon, trout, and tuna are good sources of omega-3 fatty acids. Omega-3 fatty acids are important in brain function and normal growth and development. They may also prevent heart disease and reduce the risk of cancer.

Like carbohydrates, fats have received a lot of bad publicity. It is true that eating an excess of fried foods and other "fatty" foods leads to weight gain. However, fats do have important functions. Fats serve as long-term energy storage. They also provide insulation for the body. Therefore, "healthy" unsaturated fats in moderate amounts should be consumed on a regular basis.

Phospholipids are the major constituent of the plasma membrane. Like fats, they are composed of fatty acid chains attached to a glycerol or similar backbone. Instead of three fatty acids attached, however, there are two fatty acids and the third carbon of the glycerol backbone is bound to a phosphate group. The phosphate group is modified by the addition of an alcohol.

A phospholipid has both hydrophobic and hydrophilic regions. The fatty acid chains are hydrophobic and exclude themselves from water, whereas the phosphate is hydrophilic and interacts with water.

Cells are surrounded by a membrane, which has a bilayer of phospholipids. The fatty acids of phospholipids face inside, away from water, whereas the phosphate group can face either the outside environment or the inside of the cell, which are both aqueous.

Steroids and Waxes

Unlike the phospholipids and fats discussed earlier, **steroids** have a ring structure. Although they do not resemble other lipids, they are grouped with them because they are also hydrophobic. All steroids have four, linked carbon rings and several of them, like cholesterol, have a short tail.

Cholesterol is a steroid. Cholesterol is mainly synthesized in the liver and is the precursor of many steroid hormones, such as testosterone and estradiol. It is also the precursor of vitamins E and K. Cholesterol is the precursor of bile salts, which help in the breakdown of fats and their subsequent absorption by cells. Although cholesterol is often spoken of in negative terms, it is necessary for the proper functioning of the body. It is a key component of the plasma membranes of animal cells.

Waxes are made up of a hydrocarbon chain with an alcohol (–OH) group and a fatty acid. Examples of animal waxes include beeswax and lanolin. Plants also have waxes, such as the coating on their leaves, that helps prevent them from drying out.

CONCEPT IN ACTION

For an additional perspective on lipids, explore “Biomolecules: The Lipids” through this interactive animation.

Proteins

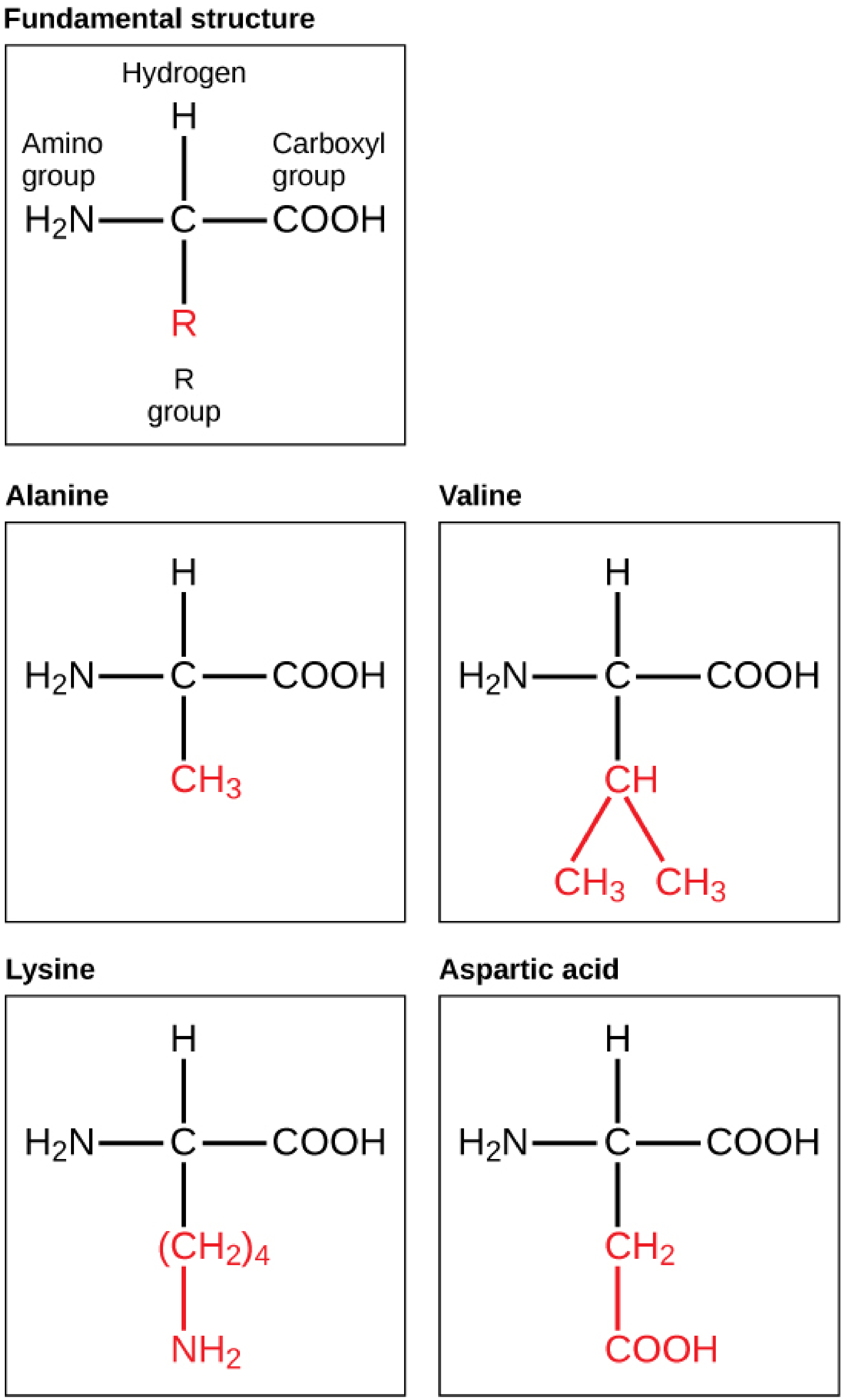
Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective; they may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each 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.

The functions of proteins are very diverse because there are 20 different chemically distinct amino acids that form long chains, and the amino acids can be in any order. For example, proteins can function as enzymes or hormones. **Enzymes**, which are produced by living cells, are catalysts in biochemical reactions (like digestion) and are usually proteins. Each enzyme is specific for the substrate (a reactant that binds to an enzyme) upon which it acts. Enzymes can function to break molecular bonds, to rearrange bonds, or to form new bonds. An example of an enzyme is salivary amylase, which breaks down amylose, a component of starch.

Hormones are chemical signaling molecules, usually proteins or steroids, secreted by an endocrine gland or group of endocrine cells that act to control or regulate specific physiological processes, including growth, development, metabolism, and reproduction. For example, insulin is a protein hormone that maintains blood glucose levels.

Proteins have different shapes and molecular weights; some proteins are globular in shape whereas others are fibrous in nature. For example, hemoglobin is a globular protein, but collagen, found in our skin, is a fibrous protein. Protein shape is critical to its function. Changes in temperature, pH, and exposure to chemicals may lead to permanent changes in the shape of the protein, leading to a loss of function or **denaturation** (to be discussed in more detail later). All proteins are made up of different arrangements of the same 20 kinds of amino acids.

Amino acids are the monomers that make up proteins. Each amino acid has the same fundamental structure, which consists of a central carbon atom bonded to an amino group ($–NH_2$), a carboxyl group ($–COOH$), and a hydrogen atom. Every amino acid also has another variable atom or group of atoms bonded to the central carbon atom known as the R group. The R group is the only difference in structure between the 20 amino acids; otherwise, the amino acids are identical (***Figure 8***).



Fundamental structure
Hydrogen
H
Amino group
Carboxyl group
H_2N—C—COOH
R
R group
Alanine
H
H_2N—C—COOH
CH_3
Valine
H
H_2N—C—COOH
CH
CH_3 CH_3
Lysine
H
H_2N—C—COOH
$(CH_2)_4$
NH_2
Aspartic acid
H
H_2N—C—COOH
CH_2
COOH

Figure 8. *Amino acids are made up of a central carbon bonded to an amino group (–NH2), a carboxyl group (–COOH), and a hydrogen atom. The central carbon's fourth bond varies among the different amino acids, as seen in these examples of alanine, valine, lysine, and aspartic acid.*

The chemical nature of the R group determines the chemical nature of the amino acid within its protein (that is, whether it is acidic, basic, polar, or nonpolar).

The sequence and number of amino acids ultimately determine a 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 a second amino acid combine, releasing a water molecule. The resulting bond is the peptide bond.

The products formed by such a linkage are called polypeptides. While the terms polypeptide and protein are sometimes used interchangeably, a **polypeptide** is technically a polymer of amino acids, whereas the term protein is used for a polypeptide or polypeptides that have combined together, have a distinct shape, and have a unique function.

EVOLUTION IN ACTION

The Evolutionary Significance of Cytochrome c

Cytochrome c is an important component of the molecular machinery that harvests energy from glucose. Because this protein's role in producing cellular energy is crucial, it has changed very little over millions of years. Protein sequencing has shown that there is a considerable amount of sequence similarity among cytochrome c molecules of different species; evolutionary relationships can be assessed by measuring the similarities or differences among various species' protein sequences.

For example, scientists have determined that human cytochrome c contains 104 amino acids. For each cytochrome c molecule that has been sequenced to date from different organisms, 37 of these amino acids appear in the same position in each cytochrome c. This indicates that all of these organisms are descended from a common ancestor. On comparing the human and chimpanzee protein sequences, no sequence difference was found. When human and rhesus monkey sequences were compared, a single difference was found in one amino acid. In contrast, human-to-yeast comparisons show a difference in 44 amino acids, suggesting that humans and chimpanzees have a more recent common ancestor than humans and the rhesus monkey, or humans and yeast.

Protein Structure

As discussed earlier, the shape of a protein is critical to its function. 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 9***).

The unique sequence and number of amino acids in a polypeptide chain is its primary structure. The unique sequence 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. In sickle cell anemia, the hemoglobin β chain has a single amino acid substitution, causing a change in both the structure and function of the protein. What is most remarkable to consider is that a hemoglobin molecule is made up of two alpha chains and two beta chains that each consist of about 150 amino acids. The molecule, therefore, has about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule—that dramatically decreases life expectancy in the affected individuals—is a single amino acid of the 600.

Because of this change of one amino acid in the chain, the normally biconcave, or disc-shaped, red blood cells assume a crescent or “sickle” shape, which clogs arteries. This can lead to a myriad of serious health problems, such as breathlessness, dizziness, headaches, and abdominal pain for those who have this disease.

Folding patterns resulting from interactions between the non-R group portions of amino acids give rise to the secondary structure of the protein. The most common are the alpha (α)-helix and beta (β)-pleated sheet structures. Both structures are held in shape by hydrogen bonds. In the alpha helix, the bonds form between every fourth amino acid and cause a twist in the amino acid chain.

In the β-pleated sheet, the “pleats” are formed by hydrogen bonding between atoms on the backbone of the polypeptide chain. The R groups are attached to the carbons, and extend above and below the folds of the pleat. The pleated segments align parallel to each other, and hydrogen bonds form between the same pairs of atoms on each of the aligned amino acids. The α-helix and β-pleated sheet structures are found in many globular and fibrous proteins.

The unique three-dimensional structure of a polypeptide is known as its tertiary structure. This structure is caused by chemical interactions between various amino acids and regions of the polypeptide. Primarily, the interactions among R groups create the complex three-dimensional tertiary structure of a protein. There may be ionic bonds formed between R groups on different amino acids, or hydrogen bonding beyond that involved in the secondary structure. 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.

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, hemoglobin is a combination of four polypeptide subunits.

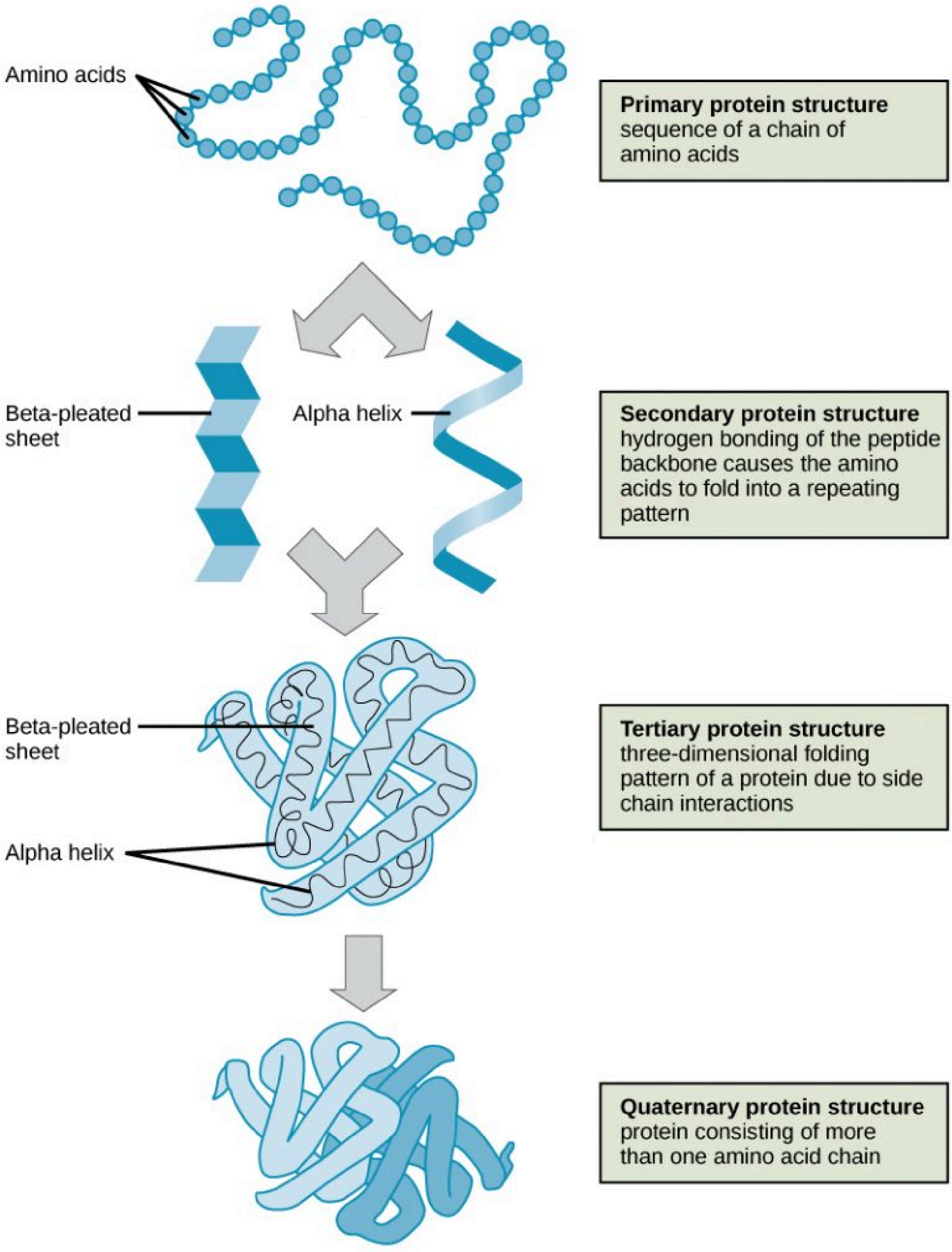


Figure 9*. The four levels of protein structure can be observed in these illustrations. (credit: modification of work by National Human Genome Research Institute)*

Each protein has its own unique sequence and shape 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 in what is known as denaturation as discussed earlier. Denaturation is often reversible because the primary structure is preserved if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to a loss of function. One example of protein denaturation can be seen when an egg is fried or boiled. The albumin protein in the liquid egg white is denatured when placed in a hot pan, changing from a clear substance to an opaque white substance. Not all proteins are denatured at high temperatures; for instance, bacteria that survive in hot springs have proteins that are adapted to function at those temperatures.

Concept in Action

For an additional perspective on proteins, explore "Biomolecules: The Proteins" through this interactive animation.

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.

DNA and RNA are made up of monomers known as nucleotides. The **nucleotides** combine with each other to form a polynucleotide, DNA or RNA. Each nucleotide is made up of three components: a nitrogenous base, a pentose (five-carbon) sugar, and a phosphate group (***Figure 10***). Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to a phosphate group.

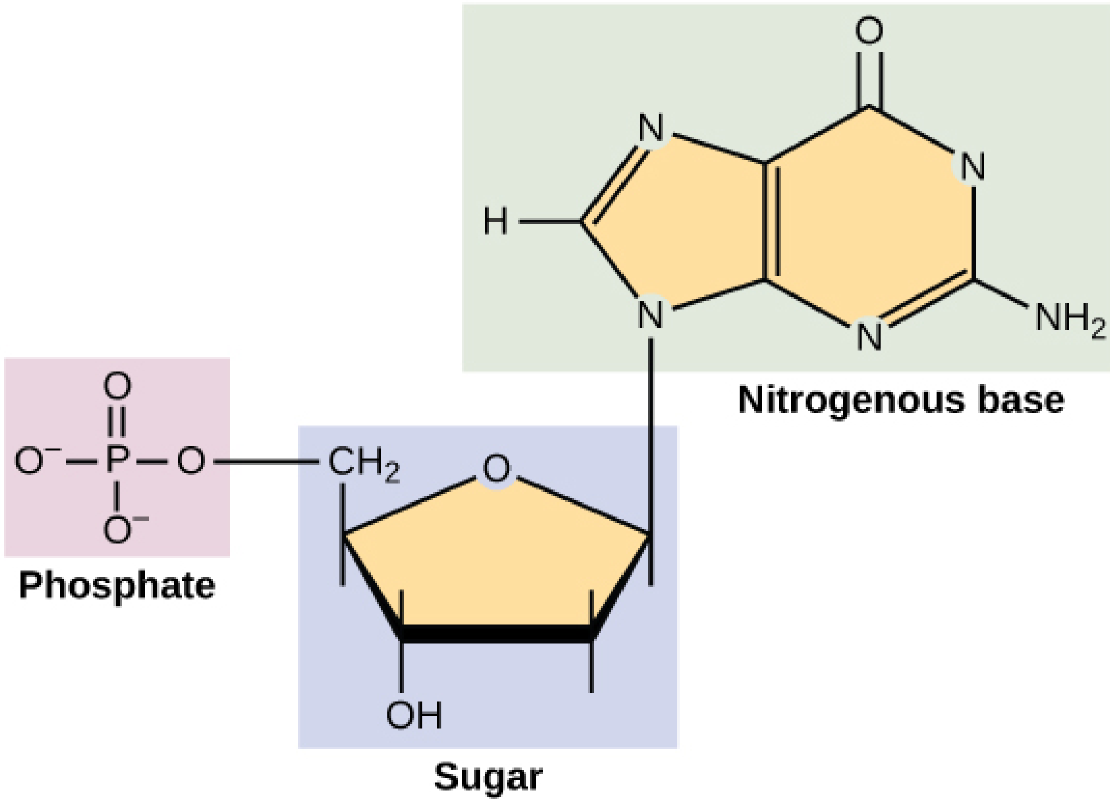


Figure 10. *A nucleotide is made up of three components: a nitrogenous base, a pentose sugar, and a phosphate group.*

DNA Double-Helical Structure

DNA has a double-helical structure (***Figure 11***). It is composed of two strands, or polymers, of nucleotides. The strands are formed with bonds between phosphate and sugar groups of adjacent nucleotides. The strands are bonded to each other at their bases with hydrogen bonds, and the strands coil about each other along their length, hence the “double helix” description, which means a double spiral.

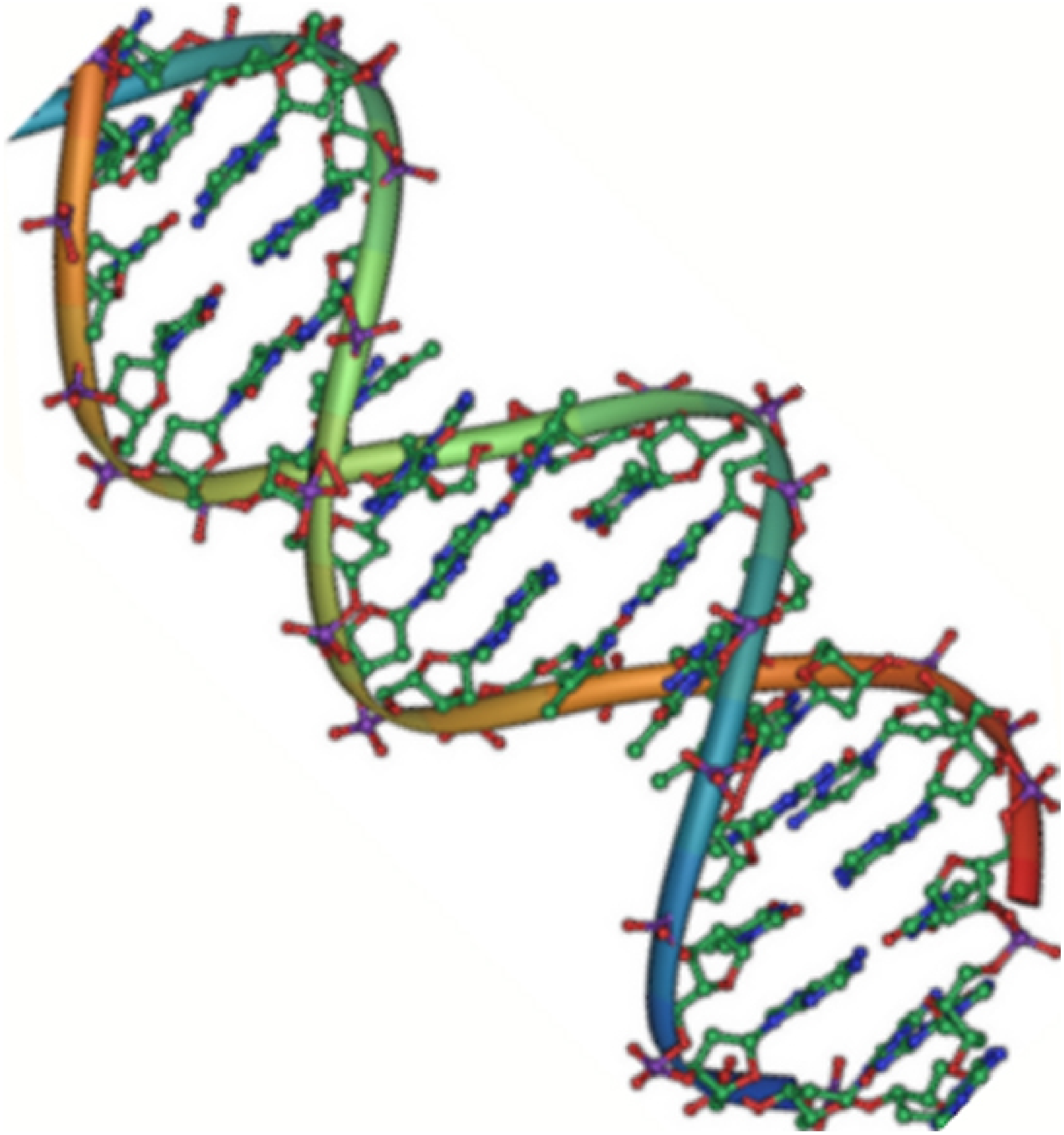


Figure 11. *The double-helix model shows DNA as two parallel strands of intertwining molecules. (credit: Jerome Walker, Dennis Myts)*

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 bases pair; 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.

Section Summary

Living things are carbon-based because carbon plays such a prominent role in the chemistry of living things. The four covalent bonding positions of the carbon atom can give rise to a wide diversity of compounds with many functions, accounting for the importance of carbon in living things. Carbohydrates are a group of macromolecules that are a vital energy source for the cell, provide structural support to many organisms, and can be found on the surface of the cell as receptors or for cell recognition. Carbohydrates are classified as monosaccharides, disaccharides, and polysaccharides, depending on the number of monomers in the molecule.

Lipids are a class of macromolecules that are nonpolar and hydrophobic in nature. Major types include fats and oils, waxes, phospholipids, and steroids. Fats and oils are a stored form of energy and can include triglycerides. Fats and oils are usually made up of fatty acids and glycerol.

Proteins are a class of macromolecules that can perform a diverse range of functions for the cell. They help in metabolism by providing structural support and by acting as enzymes, carriers or as hormones. The building blocks of proteins are amino acids. Proteins are organized at four levels: primary, secondary, tertiary, and quaternary. Protein shape and function are intricately linked; any change in shape caused by changes in temperature, pH, or chemical exposure may lead to protein denaturation and a loss of function.

Nucleic acids are molecules made up of repeating units of nucleotides that direct cellular activities such as cell division and protein synthesis. Each nucleotide is made up of a pentose sugar, a nitrogenous base, and a phosphate group. There are two types of nucleic acids: DNA and RNA.

Glossary

amino acid
: a monomer of a protein

carbohydrate
: a biological macromolecule in which the ratio of carbon to hydrogen to oxygen is 1:2:1; carbohydrates serve as energy sources and structural support in cells

cellulose
: a polysaccharide that makes up the cell walls of plants and provides structural support to the cell

chitin
: a type of carbohydrate that forms the outer skeleton of arthropods, such as insects and crustaceans, and the cell walls of fungi

denaturation
: the loss of shape in a protein as a result of changes in temperature, pH, or exposure to chemicals

deoxyribonucleic acid (DNA)
: a double-stranded polymer of nucleotides that carries the hereditary information of the cell

disaccharide
: two sugar monomers that are linked together by a peptide bond

enzyme
: a catalyst in a biochemical reaction that is usually a complex or conjugated protein

fat
: a lipid molecule composed of three fatty acids and a glycerol (triglyceride) that typically exists in a solid form at room temperature

glycogen
: a storage carbohydrate in animals

hormone
: a chemical signaling molecule, usually a protein or steroid, secreted by an endocrine gland or

group of endocrine cells; acts to control or regulate specific physiological processes

lipids
: a class of macromolecules that are nonpolar and insoluble in water

macromolecule
: a large molecule, often formed by polymerization of smaller monomers

monosaccharide
: a single unit or monomer of carbohydrates

nucleic acid
: a biological macromolecule that carries the genetic information of a cell and carries instructions for the functioning of the cell

nucleotide
: a monomer of nucleic acids; contains a pentose sugar, a phosphate group, and a nitrogenous base

oil
: an unsaturated fat that is a liquid at room temperature

phospholipid
: a major constituent of the membranes of cells; composed of two fatty acids and a phosphate group attached to the glycerol backbone

polypeptide
: a long chain of amino acids linked by peptide bonds

polysaccharide
: a long chain of monosaccharides; may be branched or unbranched

protein
: a biological macromolecule composed of one or more chains of amino acids

ribonucleic acid (RNA)
: a single-stranded polymer of nucleotides that is involved in protein synthesis

saturated fatty acid
: a long-chain hydrocarbon with single covalent bonds in the carbon chain; the number of hydrogen atoms attached to the carbon skeleton is maximized

starch
: a storage carbohydrate in plants

steroid
: a type of lipid composed of four fused hydrocarbon rings

***trans*-fat**
: a form of unsaturated fat with the hydrogen atoms neighboring the double bond across from each other rather than on the same side of the double bond

triglyceride
: a fat molecule; consists of three fatty acids linked to a glycerol molecule

unsaturated fatty acid
: a long-chain hydrocarbon that has one or more than one double bonds in the hydrocarbon chain

27.

3.5 Nutrition and Diet

The carbohydrates, lipids, and proteins in the foods you eat are used for energy to power molecular, cellular, and organ system activities. Importantly, the energy is stored primarily as fats. The quantity and quality of food that is ingested, digested, and absorbed affects the amount of fat that is stored as excess calories. Diet—both what you eat and how much you eat—has a dramatic impact on your health. Eating too much or too little food can lead to serious medical issues, including cardiovascular disease, cancer, anorexia, and diabetes, among others. Combine an unhealthy diet with unhealthy environmental conditions, such as smoking, and the potential medical complications increase significantly.

Food and Metabolism

The amount of energy that is needed or ingested per day is measured in calories. The nutritional **Calorie (C)** is the amount of heat it takes to raise 1 kg (1000 g) of water by 1 °C. This is different from the calorie (c) used in the physical sciences, which is the amount of heat it takes to raise 1 g of water by 1 °C. When we refer to "calorie," we are referring to the nutritional Calorie.

On average, a person needs 1500 to 2000 calories per day to sustain (or carry out) daily activities. The total number of calories needed by one person is dependent on their body mass, age, height, gender, activity level, and the amount of exercise per day. If exercise is regular part of one's day, more calories are required. As a rule, people underestimate the number of calories ingested and overestimate the amount they burn through exercise. This can lead to ingestion of too many calories per day. The accumulation of an extra 3500 calories adds one pound of weight. If an excess of 200 calories per day is ingested, one extra pound of body weight will be gained every 18 days. At that rate, an extra 20 pounds can be gained over the course of a year. Of course, this increase in calories could be offset by increased exercise. Running or jogging one mile burns almost 100 calories.

The type of food ingested also affects the body's metabolic rate. Processing of carbohydrates requires less energy than processing of proteins. In fact, the breakdown of carbohydrates requires the least amount of energy, whereas the processing of proteins demands the most energy. In general, the amount of calories ingested and the amount of calories burned determines the overall weight. To lose weight, the number of calories burned per day must exceed the number ingested. Calories are in almost everything you ingest, so when considering calorie intake, beverages must also be considered.

To help provide guidelines regarding the types and quantities of food that should be eaten every day, the USDA has updated their food guidelines from MyPyramid to MyPlate. They have put the recommended elements of a healthy meal into the context of a place setting of food. MyPlate categorizes food into the standard six food groups: fruits, vegetables, grains, protein foods, dairy, and oils. The accompanying website gives clear recommendations regarding quantity and type of each food that you should consume each day, as well as identifying which foods belong in each category. The accompanying graphic (***Figure 1***) gives a clear visual with general recommendations for a healthy and balanced meal. The guidelines recommend to "Make half your plate fruits and vegetables." The other half is grains and protein, with a slightly higher quantity of grains than protein. Dairy products are represented by a drink, but the quantity can be applied to other dairy products as well.

MyPlate

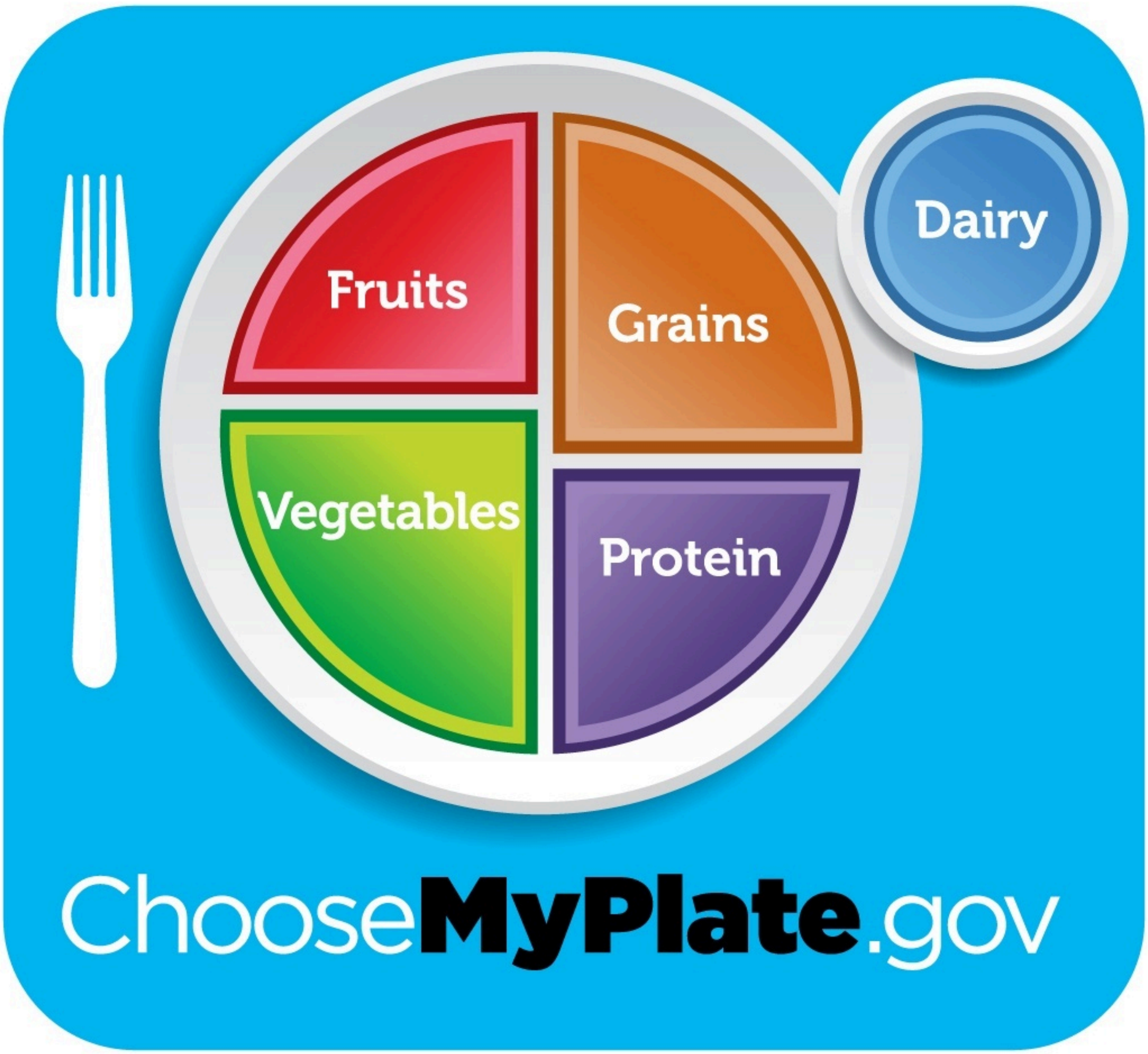


Figure 1. The U.S. Department of Agriculture developed food guidelines called MyPlate to help demonstrate how to maintain a healthy lifestyle.

ChooseMyPlate.gov provides extensive online resources for planning a healthy diet and lifestyle, including offering weight management tips and recommendations for physical activity. It also includes the SuperTracker, a web-based application to help you analyze your own diet and physical activity.

Nutrition

The organic molecules required for building cellular material and tissues must come from food. During digestion, digestible carbohydrates are ultimately broken down into glucose and used to provide energy within the cells of the body. Complex carbohydrates, including polysaccharides, can be broken down into glucose through biochemical modification; however, humans do not produce the enzyme necessary to digest cellulose (fiber). The intestinal flora in the human gut are able to extract some nutrition from these plant fibers. These plant fibers are known as dietary fiber and are an important component of the diet. The excess sugars in the body are converted into glycogen and stored for later use in the liver and muscle tissue. Glycogen stores are used to fuel prolonged exertions, such as long-distance running, and to provide energy during food shortage. Fats are stored under the skin of mammals for insulation and energy reserves.

Proteins in food are broken down during digestion and the resulting amino acids are absorbed. All of the proteins in the body must be formed from these amino-acid constituents; no proteins are obtained directly from food.

Fats add flavor to food and promote a sense of satiety or fullness. Fatty foods are also significant sources of energy, and fatty acids are required for the construction of lipid membranes. Fats are also required in the diet to aid the absorption of fat-soluble vitamins and the production of fat-soluble hormones.

While the animal body can synthesize many of the molecules required for function from precursors, there are some nutrients that must be obtained from food. These nutrients are termed **essential nutrients**, meaning they must be eaten, because the body cannot produce them.

The fatty acids omega-3 alpha-linolenic acid and omega-6 linoleic acid are essential fatty acids needed to make some membrane phospholipids. **Vitamins** are another class of essential organic molecules that are required in small quantities. Many of these assist enzymes in their function and, for this reason, are called coenzymes. Absence or low levels of vitamins can have a dramatic effect on health. **Minerals** are another set of inorganic essential nutrients that must be obtained from food. Minerals perform many functions, from muscle and nerve function, to acting as enzyme cofactors. Certain amino acids also must be procured from food and cannot be synthesized by the body. These amino acids are the "essential" amino acids. The human body can synthesize only 11 of the 20 required amino acids; the rest must be obtained from food.

Vitamins

Vitamins are organic compounds found in foods and are a necessary part of the biochemical reactions in the body. They are involved in a number of processes, including mineral and bone metabolism, and cell and tissue growth, and they act as cofactors for energy metabolism. The B vitamins play the largest role of any vitamins in metabolism ***(Table 1*** and ***Table 2***).

You get most of your vitamins through your diet, although some can be formed from the precursors absorbed during digestion. For example, the body synthesizes vitamin A from the β-carotene in orange vegetables like carrots and sweet potatoes. Vitamins are either fat-soluble or water-soluble. Fat-soluble vitamins A, D, E, and K, are absorbed through the intestinal tract with lipids in chylomicrons. Vitamin D is also synthesized in the skin through exposure to sunlight. Because they are carried in lipids, fat-soluble vitamins can accumulate in the lipids stored in the body. If excess vitamins are retained in the lipid stores in the body, hypervitaminosis can result.

Water-soluble vitamins, including the eight B vitamins and vitamin C, are absorbed with water in the gastrointestinal tract. These vitamins move easily through bodily fluids, which are water based, so they are not stored in the body. Excess water-soluble vitamins are excreted in the urine. Therefore, hypervitaminosis of water-soluble vitamins rarely occurs, except with an excess of vitamin supplements.

Table 1. Fat-soluble Vitamins

Vitamin and alternative name	**Sources**	**Recommended daily allowance**	**Function**	**Problems associated with deficiency**
A retinal or β-carotene	Yellow and orange fruits and vegetables, dark green leafy vegetables, eggs, milk, liver	700–900 μg	Eye and bone development, immune function	Night blindness, epithelial changes, immune system deficiency
D cholecalciferol	Dairy products, egg yolks; also synthesized in the skin from exposure to sunlight	5–15 μg	Aids in calcium absorption, promoting bone growth	Rickets, bone pain, muscle weakness, increased risk of death from cardiovascular disease, cognitive impairment, asthma in children, cancer
E tocopherols	Seeds, nuts, vegetable oils, avocados, wheat germ	15 mg	Antioxidant	Anemia
K phylloquinone	Dark green leafy vegetables, broccoli, Brussels sprouts, cabbage	90–120 μg	Blood clotting, bone health	Hemorrhagic disease of newborn in infants; uncommon in adults

Table 2. Water-soluble Vitamins

Vitamin and alternative name	Sources	Recommended daily allowance	Function	Problems associated with deficiency
B_1 thiamine	Whole grains, enriched bread and cereals, milk, meat	1.1–1.2 mg	Carbohydrate metabolism	Beriberi, Wernicke-Korsikoff syndrome
B_2 riboflavin	Brewer's yeast, almonds, milk, organ meats, legumes, enriched breads and cereals, broccoli, asparagus	1.1–1.3 mg	Synthesis of FAD for metabolism, production of red blood cells	Fatigue, slowed growth, digestive problems, light sensitivity, epithelial problems like cracks in the corners of the mouth
B_3 niacin	Meat, fish, poultry, enriched breads and cereals, peanuts	14–16 mg	Synthesis of NAD, nerve function, cholesterol production	Cracked, scaly skin; dementia; diarrhea; also known as pellagra
B_5 pantothenic acid	Meat, poultry, potatoes, oats, enriched breads and cereals, tomatoes	5 mg	Synthesis of coenzyme A in fatty acid metabolism	Rare: symptoms may include fatigue, insomnia, depression, irritability
B_6 pyridoxine	Potatoes, bananas, beans, seeds, nuts, meat, poultry, fish, eggs, dark green leafy vegetables, soy, organ meats	1.3–1.5 mg	Sodium and potassium balance, red blood cell synthesis, protein metabolism	Confusion, irritability, depression, mouth and tongue sores
B_7 biotin	Liver, fruits, meats	30 μg	Cell growth, metabolism of fatty acids, production of blood cells	Rare in developed countries; symptoms include dermatitis, hair loss, loss of muscular coordination
B_9 folic acid	Liver, legumes, dark green leafy vegetables, enriched breads and cereals, citrus fruits	400 μg	DNA/protein synthesis	Poor growth, gingivitis, appetite loss, shortness of breath, gastrointestinal problems, mental deficits
B_{12} cyanocobalamin	Fish, meat, poultry, dairy products, eggs	2.4 μg	Fatty acid oxidation, nerve cell function, red blood cell production	Pernicious anemia, leading to nerve cell damage
C ascorbic acid	Citrus fruits, red berries, peppers, tomatoes, broccoli, dark green leafy vegetables	75–90 mg	Necessary to produce collagen for formation of connective tissue and teeth, and for wound healing	Dry hair, gingivitis, bleeding gums, dry and scaly skin, slow wound healing, easy bruising, compromised immunity; can lead to scurvy

Minerals

Minerals in food are inorganic compounds that work with other nutrients to ensure the body functions properly. Minerals cannot be made in the body; they come from the diet. The amount of minerals in the body is small—only 4 percent of the total body mass—and most of that consists of the minerals that the body requires in moderate quantities: potassium, sodium, calcium, phosphorus, magnesium, and chloride.

The most common minerals in the body are calcium and phosphorous, both of which are stored in the skeleton and necessary for the hardening of bones. Most minerals are ionized, and their ionic forms are used in physiological processes throughout the body. Sodium and chloride ions are electrolytes in the blood and extracellular tissues, and iron ions are critical to the formation of hemoglobin. There are additional trace minerals that are still important to the body's functions, but their required quantities are much lower.

Like vitamins, minerals can be consumed in toxic quantities (although it is rare). A healthy diet includes most of the minerals your body requires, so supplements and processed foods can add potentially toxic levels of minerals. ***Table 3*** and ***Table 4*** provide a summary of minerals and their function in the body.

Table 3. Major Minerals

Mineral	**Sources**	**Recommended daily allowance**	**Function**	**Problems associated with deficiency**
Potassium	Meats, some fish, fruits, vegetables, legumes, dairy products	4700 mg	Nerve and muscle function; acts as an electrolyte	Hypokalemia: weakness, fatigue, muscle cramping, gastrointestinal problems, cardiac problems
Sodium	Table salt, milk, beets, celery, processed foods	2300 mg	Blood pressure, blood volume, muscle and nerve function	Rare
Calcium	Dairy products, dark green leafy vegetables, blackstrap molasses, nuts, brewer's yeast, some fish	1000 mg	Bone structure and health; nerve and muscle functions, especially cardiac function	Slow growth, weak and brittle bones
Phosphorous	Meat, milk	700 mg	Bone formation, metabolism, ATP production	Rare
Magnesium	Whole grains, nuts, leafy green vegetables	310–420 mg	Enzyme activation, production of energy, regulation of other nutrients	Agitation, anxiety, sleep problems, nausea and vomiting, abnormal heart rhythms, low blood pressure, muscular problems
Chloride	Most foods, salt, vegetables, especially seaweed, tomatoes, lettuce, celery, olives	2300 mg	Balance of body fluids, digestion	Loss of appetite, muscle cramps

Table 4. Trace Minerals

Mineral	Sources	Recommended daily allowance	Function	Problems associated with deficiency
Iron	Meat, poultry, fish, shellfish, legumes, nuts, seeds, whole grains, dark leafy green vegetables	8–18 mg	Transport of oxygen in blood, production of ATP	Anemia, weakness, fatigue
Zinc	Meat, fish, poultry, cheese, shellfish	8–11 mg	Immunity, reproduction, growth, blood clotting, insulin and thyroid function	Loss of appetite, poor growth, weight loss, skin problems, hair loss, vision problems, lack of taste or smell
Copper	Seafood, organ meats, nuts, legumes, chocolate, enriched breads and cereals, some fruits and vegetables	900 μg	Red blood cell production, nerve and immune system function, collagen formation, acts as an antioxidant	Anemia, low body temperature, bone fractures, low white blood cell concentration, irregular heartbeat, thyroid problems
Iodine	Fish, shellfish, garlic, lima beans, sesame seeds, soybeans, dark leafy green vegetables	150 μg	Thyroid function	Hypothyroidism: fatigue, weight gain, dry skin, temperature sensitivity
Sulfur	Eggs, meat, poultry, fish, legumes	None	Component of amino acids	Protein deficiency
Fluoride	Fluoridated water	3–4 mg	Maintenance of bone and tooth structure	Increased cavities, weak bones and teeth
Manganese	Nuts, seeds, whole grains, legumes	1.8–2.3 mg	Formation of connective tissue and bones, blood clotting, sex hormone development, metabolism, brain and nerve function	Infertility, bone malformation, weakness, seizures
Cobalt	Fish, nuts, leafy green vegetables, whole grains	None	Component of B_{12}	None
Selenium	Brewer’s yeast, wheat germ, liver, butter, fish, shellfish, whole grains	55 μg	Antioxidant, thyroid function, immune system function	Muscle pain
Chromium	Whole grains, lean meats, cheese, black pepper, thyme, brewer’s yeast	25–35 μg	Insulin function	High blood sugar, triglyceride, and cholesterol levels

Mineral	Sources	Recommended daily allowance	Function	Problems associated with deficiency
Molybdenum	Legumes, whole grains, nuts	45 μg	Cofactor for enzymes	Rare

Section Summary

Nutrition and diet affect your metabolism. More energy is required to break down fats and proteins than carbohydrates; however, all excess calories that are ingested will be stored as fat in the body. On average, a person requires 1500 to 2000 calories for normal daily activity, although routine exercise will increase that amount. If you ingest more than that, the remainder is stored for later use. Conversely, if you ingest less than that, the energy stores in your body will be depleted. Both the quantity and quality of the food you eat affect your metabolism and can affect your overall health. Eating too much or too little can result in serious medical conditions, including cardiovascular disease, cancer, and diabetes.

Vitamins and minerals are essential parts of the diet. They are needed for the proper function of metabolic pathways in the body. Vitamins are not stored in the body, so they must be obtained from the diet or synthesized from precursors available in the diet. Minerals are also obtained from the diet, but they are also stored, primarily in skeletal tissues.

Glossary

calorie
: amount of heat it takes to raise 1 kg (1000 g) of water by 1 °C

essential nutrient
: a nutrient that cannot be synthesized by the body; it must be obtained from food

minerals
: inorganic compounds required by the body to ensure proper function of the body

vitamins
: organic compounds required by the body to perform biochemical reactions like metabolism and bone, cell, and tissue growth

28.

3.6 The Digestive System

All living organisms need nutrients to survive. While plants can obtain nutrients from their roots and the energy molecules required for cellular function through the process of photosynthesis, animals obtain their nutrients by the consumption of other organisms. At the cellular level, the biological molecules necessary for animal function are amino acids, lipid molecules, nucleotides, and simple sugars. However, the food consumed consists of protein, fat, and complex carbohydrates. Animals must convert these macromolecules into the simple molecules required for maintaining cellular function. The conversion of the food consumed to the nutrients required is a multistep process involving digestion and absorption. During digestion, food particles are broken down to smaller components, which are later absorbed by the body. This happens by both physical means, such as chewing, and by chemical means.

One of the challenges in human nutrition is maintaining a balance between food intake, storage, and energy expenditure. Taking in more food energy than is used in activity leads to storage of the excess in the form of fat deposits. The rise in obesity and the resulting diseases like type 2 diabetes makes understanding the role of diet and nutrition in maintaining good health all the more important.

The Human Digestive System

The process of digestion begins in the mouth with the intake of food (***Figure 1***). The teeth play an important role in masticating (chewing) or physically breaking food into smaller particles. The enzymes present in saliva also begin to chemically break down food. The food is then swallowed and enters the **esophagus**—a long tube that connects the mouth to the stomach. Using **peristalsis**, or wave-like smooth-muscle contractions, the muscles of the esophagus push the food toward the stomach. The stomach contents are extremely acidic, with a pH between 1.5 and 2.5. This acidity kills microorganisms, breaks down food tissues, and activates digestive enzymes. Further breakdown of food takes place in the small intestine where bile produced by the liver, and enzymes produced by the small intestine and the pancreas, continue the process of digestion. The smaller molecules are absorbed into the blood stream through the epithelial cells lining the walls of the small intestine. The waste material travels on to the large intestine where water is absorbed and the drier waste material is compacted into feces; it is stored until it is excreted through the anus.

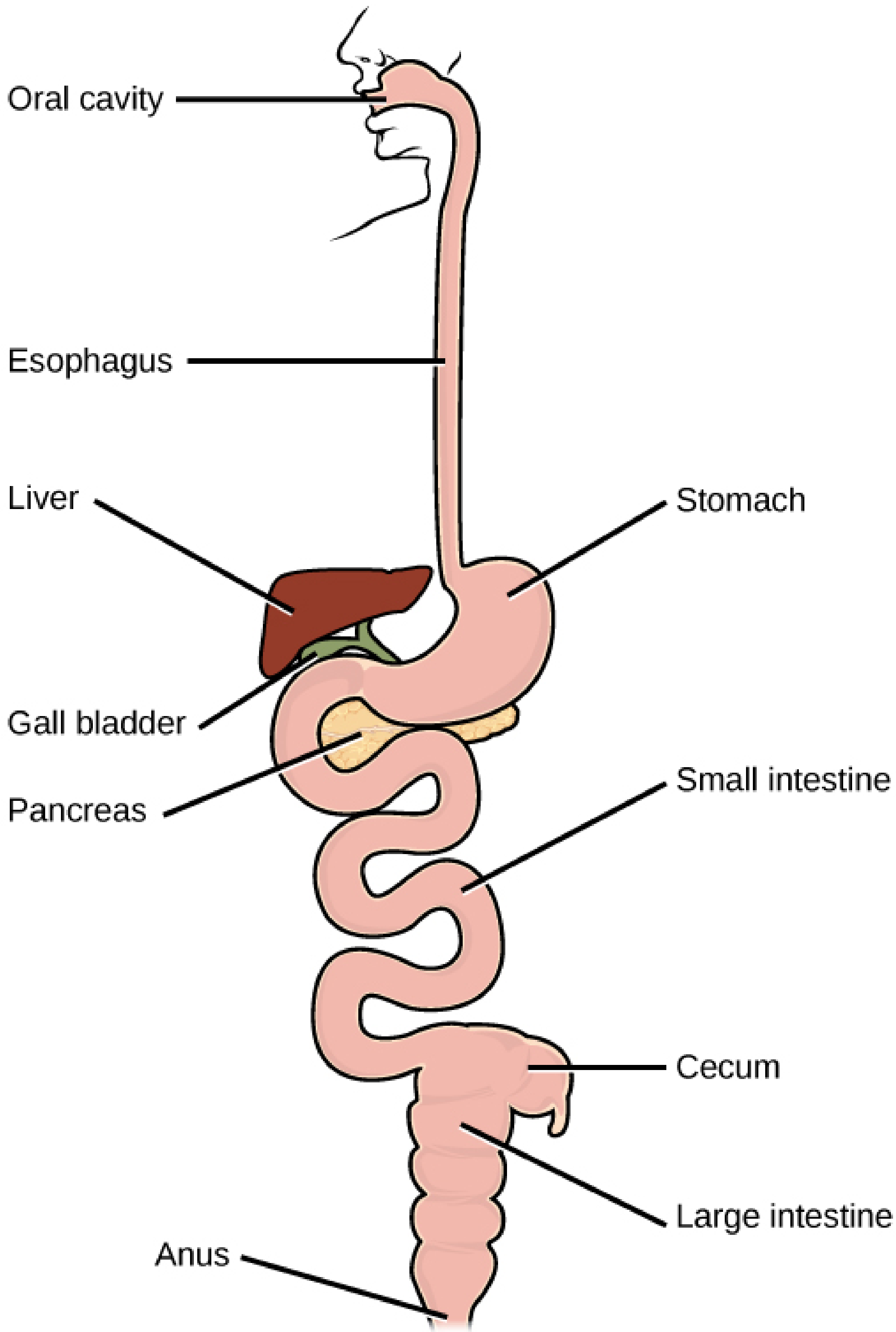


Figure 1. *The components of the human digestive system are shown.*

Oral Cavity

Both physical and chemical digestion begin in the mouth or **oral cavity**, which is the point of entry of food into the digestive system. The food is broken into smaller particles by mastication, the chewing action of the teeth. All mammals have teeth and can chew their food to begin the process of physically breaking it down into smaller particles.

The chemical process of digestion begins during chewing as food mixes with saliva, produced by the **salivary glands** (***Figure 2***). Saliva contains mucus that moistens food and buffers the pH of the food. Saliva also contains lysozyme, which has antibacterial action. It also contains an enzyme called **salivary amylase** that begins the process of converting starches in the food into a disaccharide called maltose. Another enzyme called lipase is produced by cells in the tongue to break down fats. The chewing and wetting action provided by the teeth and saliva prepare the food into a mass called the **bolus** for swallowing. The tongue helps in swallowing—moving the bolus from the mouth into the pharynx. The pharynx opens to two passageways: the esophagus and the trachea. The esophagus leads to the stomach and the trachea leads to the lungs. The epiglottis is a flap of tissue that covers the tracheal opening during swallowing to prevent food from entering the lungs.

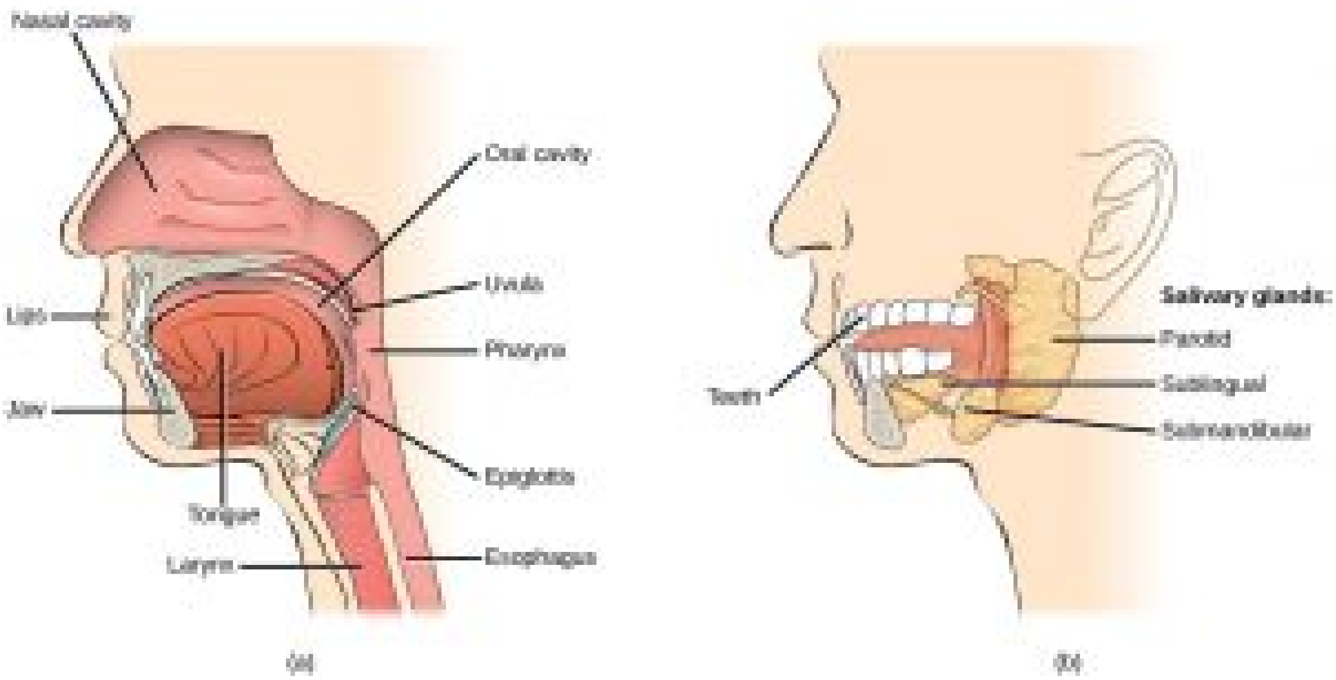


Figure 2. ***(a)*** *Digestion of food begins in the mouth.* ***(b)*** *Food is masticated by teeth and moistened by saliva secreted from the salivary glands. Enzymes in the saliva begin to digest starches and fats. With the help of the tongue, the resulting bolus is moved into the esophagus by swallowing. (credit: modification of work by Mariana Ruiz Villareal)*

Esophagus

The esophagus is a tubular organ that connects the mouth to the stomach. The chewed and softened food passes through the esophagus after being swallowed. The smooth muscles of the esophagus undergo peristalsis that pushes the food toward the stomach. The peristaltic wave is unidirectional—it moves food from the mouth the stomach, and reverse movement is not possible, except in the case of the vomit reflex. The peristaltic movement of the esophagus is an involuntary reflex; it takes place in response to the act of swallowing.

Ring-like muscles called sphincters form valves in the digestive system. The gastro-esophageal sphincter (or cardiac sphincter) is located at the stomach end of the esophagus. In response to swallowing and the pressure exerted by the bolus of food, this sphincter opens, and the bolus enters the stomach. When there is no swallowing action, this sphincter is shut and prevents the contents of the stomach from traveling up the esophagus. Acid reflux or "heartburn" occurs when the acidic digestive juices escape into the esophagus.

Stomach

A large part of protein digestion occurs in the stomach (***Figure 4***). The **stomach** is a saclike organ that secretes gastric digestive juices.

Protein digestion is carried out by an enzyme called **pepsin** in the stomach chamber. The highly acidic environment kills many microorganisms in the food and, combined with the action of the enzyme pepsin, results in the catabolism of protein in the food. Chemical digestion is facilitated by the churning action of the stomach caused by contraction and relaxation of smooth muscles. The partially digested food and gastric juice mixture is called **chyme**. Gastric emptying occurs within two to six hours after a meal. Only a small amount of chyme is released into the small intestine at a time. The movement of chyme from the stomach into the small intestine is regulated by hormones, stomach distension and muscular reflexes that influence the pyloric sphincter.

The stomach lining is unaffected by pepsin and the acidity because pepsin is released in an inactive form and the stomach has a thick mucus lining that protects the underlying tissue.

Small Intestine

Chyme moves from the stomach to the small intestine. The **small intestine** is the organ where the digestion of protein, fats, and carbohydrates is completed. The small intestine is a long tube-like organ with a highly folded surface containing finger-like projections called the villi. The top surface of each villus has many microscopic projections called microvilli. The epithelial cells of these structures absorb nutrients from the digested food and release them to the bloodstream on the other side. The villi and microvilli, with their many folds, increase the surface area of the small intestine and increase absorption efficiency of the nutrients.

The human small intestine is over 6 m (19.6 ft) long and is divided into three parts: the duodenum, the jejunum and the ileum. The duodenum is separated from the stomach by the pyloric sphincter. The chyme is mixed with pancreatic juices, an alkaline solution rich in bicarbonate that neutralizes the acidity of chyme from the stomach. Pancreatic juices contain several digestive enzymes that break down starches, disaccharides, proteins, and fats. **Bile** is produced in the liver and stored and concentrated in the gallbladder; it enters the duodenum through the bile duct. Bile contains bile salts, which make lipids accessible to the water-soluble enzymes. The monosaccharides, amino acids, bile salts, vitamins, and other nutrients are absorbed by the cells of the intestinal lining.

The undigested food is sent to the colon from the ileum via peristaltic movements. The ileum ends and the large intestine begins at the ileocecal valve. The vermiform, "worm-like," appendix is located at the ileocecal valve. The appendix of humans has a minor role in immunity.

Large Intestine

The **large intestine** reabsorbs the water from indigestible food material and processes the waste material (***Figure 3***). The human large intestine is much smaller in length compared to the small intestine but larger in diameter. It has three parts: the cecum, the colon, and the rectum. The cecum joins the ileum to the colon and is the receiving pouch for the waste matter. The colon is home to many bacteria or "intestinal flora" that aid in the digestive processes. The **colon** has four regions, the ascending colon, the transverse colon, the descending colon and the sigmoid colon. The main functions of the colon are to extract the water and mineral salts from undigested food, and to store waste material.

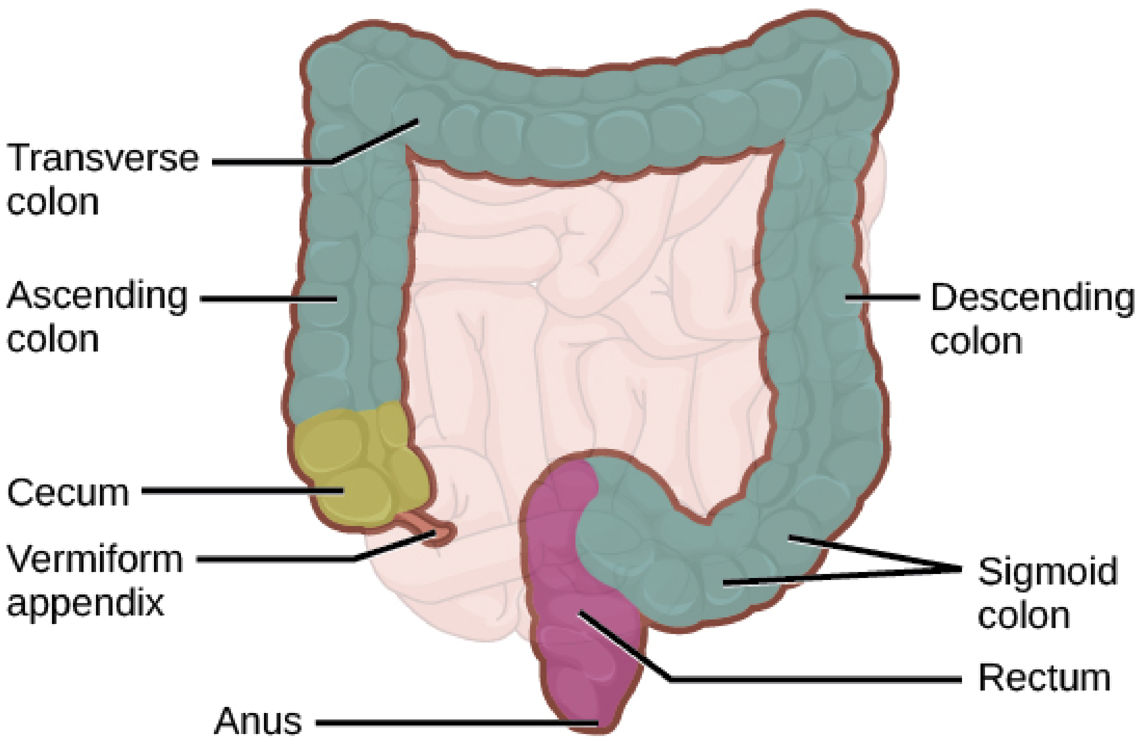


Figure 3. *The large intestine reabsorbs water from undigested food and stores waste until it is eliminated. (credit: modification of work by Mariana Ruiz Villareal)*

The **rectum** (***Figure 3***) stores feces until defecation. The feces are propelled using peristaltic movements during elimination. The **anus** is an opening at the far-end of the digestive tract and is the exit point for the waste material. Two sphincters regulate the exit of feces, the inner sphincter is involuntary and the outer sphincter is voluntary.

Accessory Organs

The organs discussed above are the organs of the digestive tract through which food passes. Accessory organs add secretions and enzymes that break down food into nutrients. Accessory organs include the salivary glands, the liver, the pancreas, and the gall bladder. The secretions of the liver, pancreas, and gallbladder are regulated by hormones in response to food consumption.

The **liver** is the largest internal organ in humans and it plays an important role in digestion of fats and detoxifying blood. The liver produces bile, a digestive juice that is required for the breakdown of fats in the duodenum. The liver also processes the absorbed vitamins and fatty acids and synthesizes many plasma proteins. The **gallbladder** is a small organ that aids the liver by storing bile and concentrating bile salts.

The **pancreas** secretes bicarbonate that neutralizes the acidic chyme and a variety of enzymes for the digestion of protein and carbohydrates.

ART CONNECTION

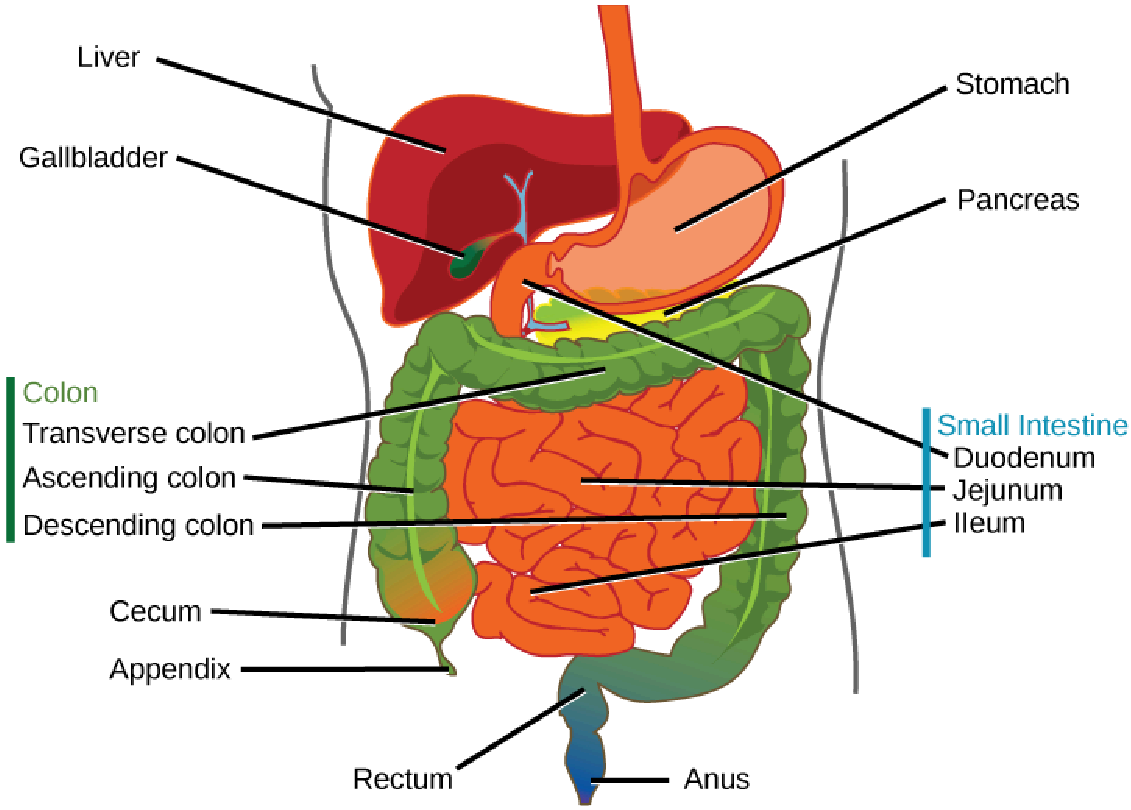


Figure 4. *The stomach has an extremely acidic environment where most of the protein gets digested. (credit: modification of work by Mariana Ruiz Villareal)*

BIOLOGY IN ACTION

Obesity

With obesity at high rates in the United States, there is a public health focus on reducing obesity and associated health risks, which include diabetes, colon and breast cancer, and cardiovascular disease. How does the food consumed contribute to obesity?

Fatty foods are calorie-dense, meaning that they have more calories per unit mass than carbohydrates or proteins. One gram of carbohydrates has four calories, one gram of protein has four calories, and one gram of fat has nine calories. Animals tend to seek lipid-rich food for their higher energy content. Greater amounts of food energy taken in than the body's requirements will result in storage of the excess in fat deposits.

Excess carbohydrate is used by the liver to synthesize glycogen. When glycogen stores are full, additional glucose is converted into fatty acids. These fatty acids are stored in adipose tissue cells—the fat cells in the mammalian body whose primary role is to store fat for later use.

The rate of obesity among children is rapidly rising in the United States. To combat childhood obesity and ensure that children get a healthy start in life, in 2010 First Lady Michelle Obama launched the Let's Move! campaign. The goal of this campaign is to educate parents and caregivers on providing healthy nutrition and encouraging active lifestyles in future generations. This program aims to involve the entire community, including parents, teachers, and healthcare providers to ensure that children have access to healthy foods—more fruits, vegetables, and whole grains—and consume fewer calories from processed foods. Another goal is to ensure that children get physical activity. With the increase in television viewing and stationary pursuits such as video games, sedentary lifestyles have become the norm. Visit www.letsmove.gov to learn more.

Section Summary

There are many organs that work together to digest food and absorb nutrients. The mouth is the point of ingestion and the location where both mechanical and chemical breakdown of food begins. Saliva contains an enzyme called amylase that breaks down carbohydrates. The food bolus travels through the esophagus by peristaltic movements to the stomach. The stomach has an extremely acidic environment. The enzyme pepsin digests protein in the stomach. Further digestion and absorption take place in the small intestine. The large intestine reabsorbs water from the undigested food and stores waste until elimination.

Carbohydrates, proteins, and fats are the primary components of food. Some essential nutrients are required for cellular function but cannot be produced by the animal body. These include vitamins, minerals, some fatty acids, and some amino acids. Food intake in more than necessary amounts is stored as glycogen in the liver and muscle cells, and in adipose tissue. Excess adipose storage can lead to obesity and serious health problems.

Glossary

amylase

an enzyme found in saliva and secreted by the pancreas that converts carbohydrates to maltose

anus
the exit point of the digestive system for waste material

bile
a digestive juice produced by the liver; important for digestion of lipids

bolus
a mass of food resulting from chewing action and wetting by saliva

colon
the largest portion of the large intestine consisting of the ascending colon, transverse colon, and descending colon

chyme
a mixture of partially digested food and stomach juices

esophagus
a tubular organ that connects the mouth to the stomach

gallbladder
the organ that stores and concentrates bile

large intestine
a digestive system organ that reabsorbs water from undigested material and processes waste matter

liver
an organ that produces bile for digestion and processes vitamins and lipids

oral cavity
the point of entry of food into the digestive system

pancreas
a gland that secretes digestive juices

pepsin
an enzyme found in the stomach whose main role is protein digestion

peristalsis
wave-like movements of muscle tissue

rectum
the area of the body where feces is stored until elimination

salivary gland
one of three pairs of exocrine glands in the mammalian mouth that secretes saliva, a mix of watery mucus and enzymes

small intestine
the organ where digestion of protein, fats, and carbohydrates is completed

stomach
a saclike organ containing acidic digestive juices

29.

3.7 ATP: Adenosine Triphosphate

Almost all chemical reactions in human cells require energy. Within the cell, from where does energy to power such reactions come? The answer lies with an energy-supplying molecule scientists call **adenosine triphosphate**, or **ATP**. This is a small, relatively simple molecule (***Figure 1***), but within some of its bonds, it contains the potential for a quick burst of energy that can be harnessed to perform cellular work. Think of this molecule as the cells' primary energy currency in much the same way that money is the currency that people exchange for things they need. ATP powers the majority of energy-requiring cellular reactions.

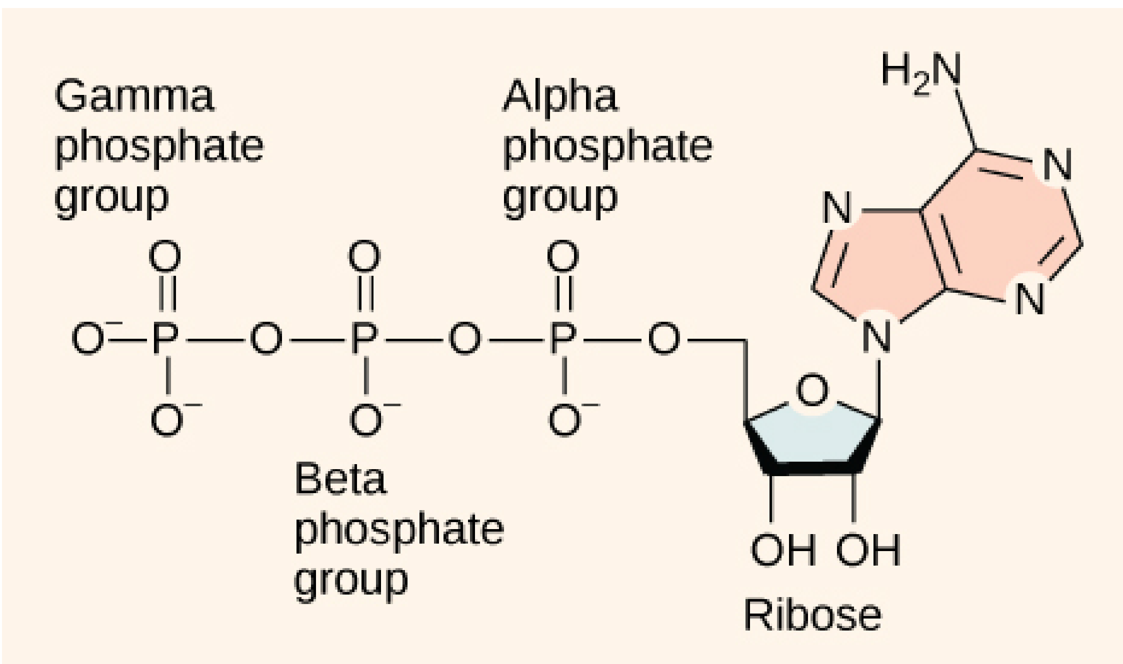


Figure 1. *ATP is the cell's primary energy currency. It has an adenosine backbone with three phosphate groups attached.*

As its name suggests, adenosine triphosphate is comprised of adenosine bound to three phosphate groups (***Figure 1***). Adenosine is a nucleoside consisting of the nitrogenous base adenine and a five-carbon sugar, ribose. The three phosphate groups, in order of closest to furthest from the ribose sugar, are alpha, beta, and gamma. Together, these chemical groups constitute an energy powerhouse. However, not all bonds within this molecule exist in a particularly high-energy state. Both bonds that link the phosphates are equally high-energy bonds (**phosphoanhydride bonds**) that, when broken, release sufficient energy to power a variety of cellular reactions and processes. These high-energy bonds are the bonds between the second and third (or beta and gamma) phosphate groups and between the first and second phosphate groups. These bonds are “high-energy” because the products of such bond breaking—adenosine diphosphate (ADP) and one inorganic phosphate group (P_i)—have considerably lower free energy than the reactants: ATP and a water molecule. Because this reaction takes place using a water molecule, it is a hydrolysis reaction. In other words, ATP hydrolyzes into ADP in the following reaction:

ATP + H_2O ATP+H2O→ADP+Pi+free energyATP+H2O→ADP+Pi+free energysize 12{{ATP} + H rSub { size 8{2} } O ADP + P rSub { size 8{i}} + {free energy} } {}">→ ADP + Pi + free energy

Like most chemical reactions, ATP to ADP hydrolysis is reversible. The reverse reaction regenerates ATP from ADP + P_i. Cells rely on ATP regeneration just as people rely on regenerating spent money through some sort of income. Since ATP hydrolysis releases energy, ATP regeneration must require an input of free energy. This equation expresses ATP formation:

ADP + Pi + free energy ATP+H2O→ADP+Pi+free energyATP+H2O→ADP+Pi+free energysize 12{{ATP} + H rSub { size 8{2} } O ADP + P rSub { size 8{i}} + {free energy} } {}">→ ATP + H_2O

ATP is a highly unstable molecule. Unless quickly used to perform work, ATP spontaneously dissociates into ADP + P_i, and the free energy released during this process is lost as heat. Cells can harness the energy released during ATP hydrolysis by using energy coupling, where the process of ATP hydrolysis is linked to other processes in the cell. One example of energy coupling using ATP involves a transmembrane ion pump that is extremely important for cellular function. This sodium-potassium pump (Na^+/K^+ pump) drives sodium out of the cell and potassium into the cell (***Figure 2***). A large percentage of a cell’s ATP powers this pump, because cellular processes bring considerable sodium into the cell and potassium out of it. The pump works constantly to stabilize cellular concentrations of sodium and potassium. In order for the pump to turn one cycle (exporting three Na+ ions and importing two K^+ ions), one ATP molecule must hydrolyze. When ATP hydrolyzes, its gamma phosphate does not simply float away, but it actually transfers onto the pump protein. Scientists call this process of a phosphate group binding to a molecule phosphorylation. As with most ATP hydrolysis cases, a phosphate from ATP transfers onto another molecule. In a phosphorylated state, the Na^+/K^+ pump has more free energy and is triggered to undergo a conformational change (a change in the shape of the protein.) This change allows it to release $Na^+$to the cell’s outside. It then binds extracellular K^+, which, through another conformational change, causes the phosphate to detach from the pump. This phosphate release triggers the K^+ to release to the cell’s inside. Essentially, the energy released from the ATP hydrolysis couples with the energy required to power the pump and transport Na^+ and K^+ ions. ATP performs cellular work using this basic form of energy coupling through phosphorylation.

VISUAL CONNECTION

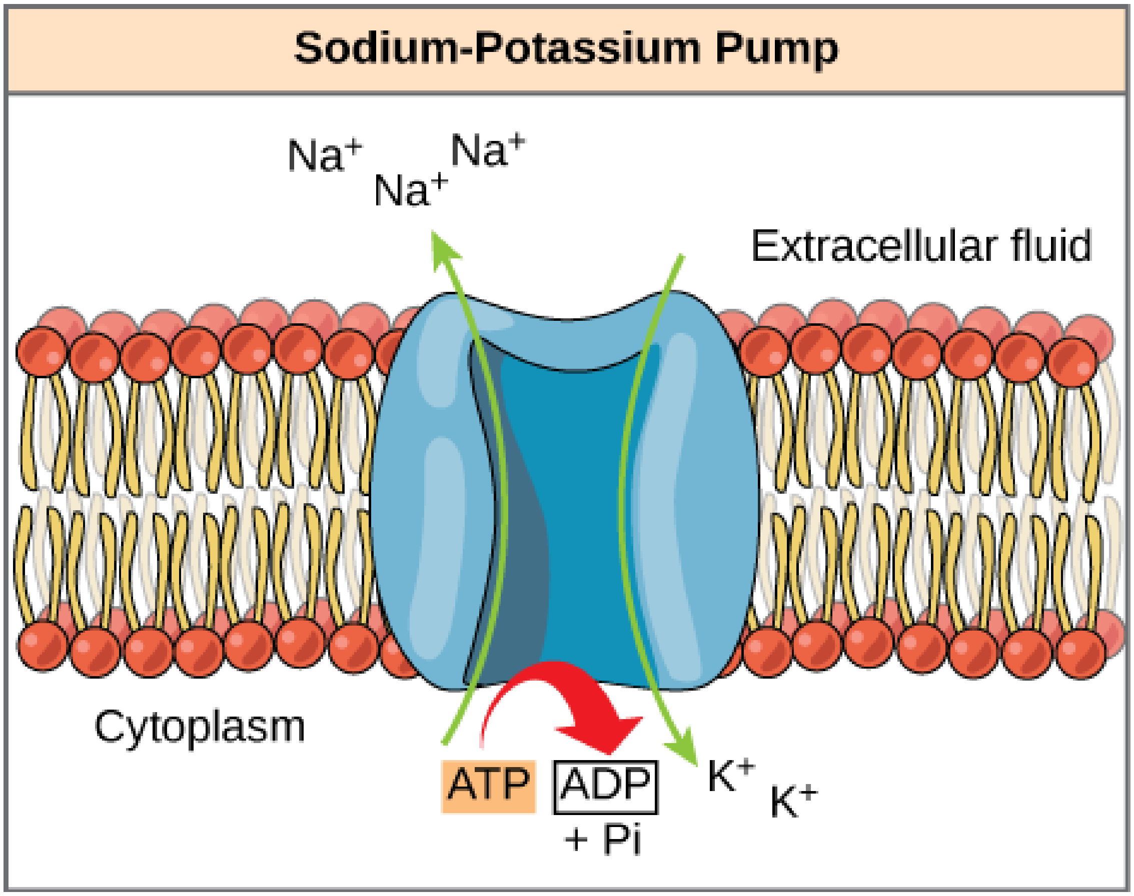


Figure 2. *The sodium-potassium pump is an example of energy coupling. The energy derived from exergonic ATP hydrolysis pumps sodium and potassium ions across the cell membrane.*

Often during cellular metabolic reactions, such as nutrient synthesis and breakdown, certain molecules must alter slightly in their conformation to become substrates for the next step in the reaction series. One example is during the very first steps of cellular respiration, when a sugar glucose molecule breaks down in the process of glycolysis. In the first step, ATP is required to phosphorylze glucose, creating a high-energy but unstable intermediate. This phosphorylation reaction powers a conformational change that allows the phosphorylated glucose molecule to convert to the phosphorylated sugar fructose. Fructose is a necessary intermediate for glycolysis to move forward. Here, ATP hydrolysis' exergonic reaction couples with the endergonic reaction of converting glucose into a phosphorylated intermediate in the pathway. Once again, the energy released by breaking a phosphate bond within ATP was used for phosphorylyzing another molecule, creating an unstable intermediate and powering an important conformational change.

Link to Learning

See an interactive animation of the ATP-producing glycolysis process at this site.

Section Summary

ATP is the primary energy-supplying molecule for living cells. ATP is comprised of a nucleotide, a five-carbon sugar, and three phosphate groups. The bonds that connect the phosphates (phosphoanhydride bonds) have high-energy content. The energy released from ATP hydrolysis into ADP + P_i performs cellular work. Cells use ATP to perform work by coupling ATP hydrolysis' exergonic reaction with endergonic reactions. ATP donates its phosphate group to another molecule via phosphorylation. The phosphorylated molecule is at a higher-energy state and is less stable than its unphosphorylated form, and this added energy from phosphate allows the molecule to undergo its endergonic reaction.

Glossary

ATP
adenosine triphosphate, the cell's energy currency

phosphoanhydride bond
bond that connects phosphates in an ATP molecule

30.

3.8 Energy and Metabolism

Scientists use the term **bioenergetics** to discuss the concept of energy flow (***Figure 1***) 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 what has been used, cells must continually produce more energy to replenish that used by the many energy-requiring chemical reactions that constantly take place. All of the chemical reactions that take place inside cells, including those that use energy and those that release energy, are the cell's **metabolism**.

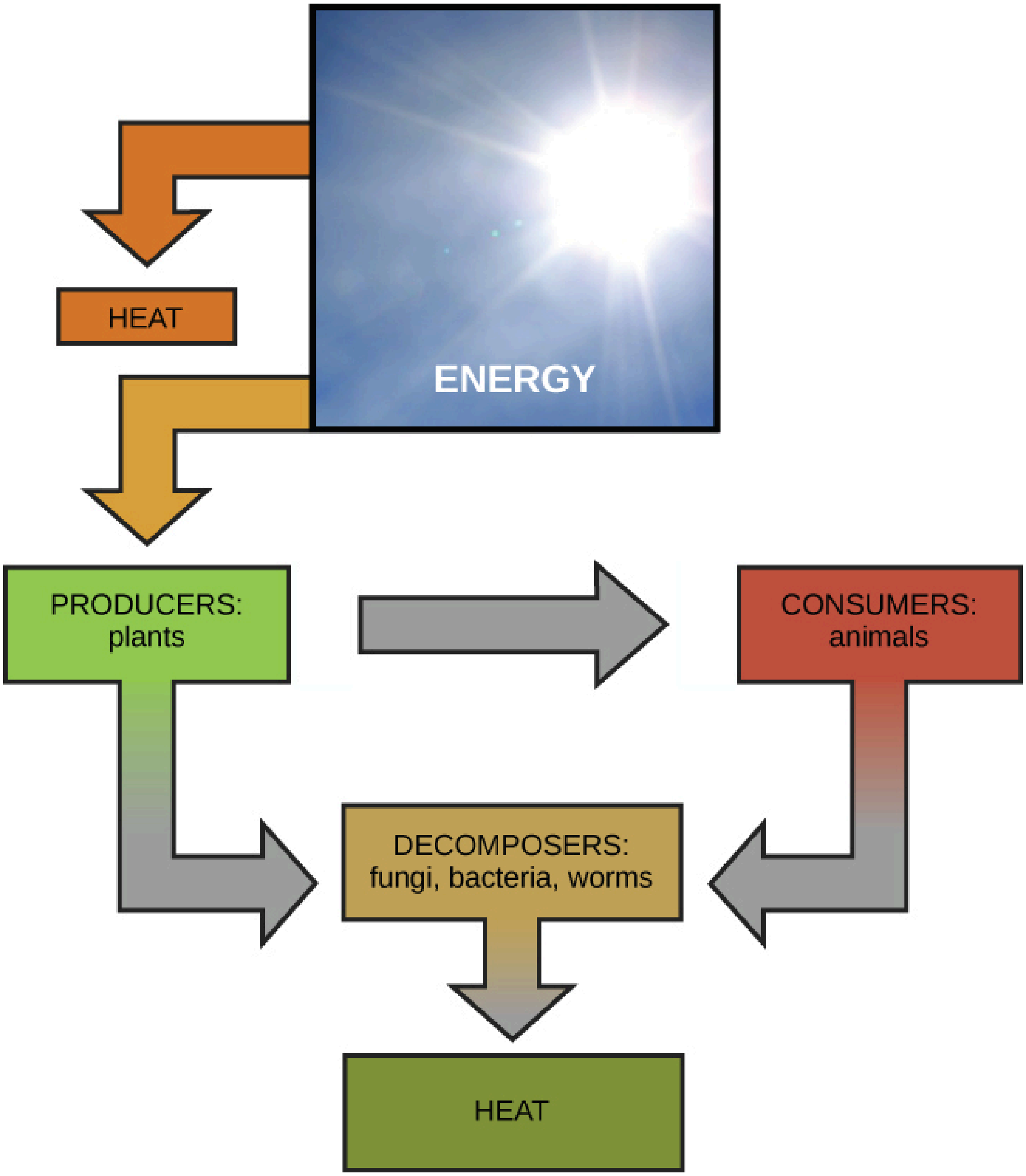


Figure 1. *Most life forms on earth get their energy from the sun. Plants use photosynthesis to capture sunlight, and herbivores eat those plants to obtain energy. Carnivores eat the herbivores, and decomposers digest plant and animal matter.*

Metabolism of Carbohydrates

The metabolism of sugar (a simple carbohydrate) is a classic example of the many cellular processes that use and produce energy. Living things consume sugar as a major energy source, because sugar molecules have a great deal of energy stored within their bonds. The breakdown of glucose, a simple sugar, is described by the equation:

$C_6H_{12}O_6$ + $6O_2$ C6H12O6+6O2→6CO2+6H2O+energyC6H12O6+6O2→6CO2+6H2O+energy size 12{C

rSub { size 8{6} } H rSub { size 8{12} } O rSub { size 8{2} } } {}">→ $6CO_2 + 6H_2O +$ energyC6H12O6+6O2→6CO2+6H2O+energyC6H12O6+6O2→6CO2+6H2O+energy size 12{C rSub { size 8{6} } H rSub { size 8{12} } O rSub { size 8{2} } } {}">

Carbohydrates that are consumed have their origins in photosynthesizing organisms like plants (***Figure 2***). During photosynthesis, plants use the energy of sunlight to convert carbon dioxide gas (CO_2) into sugar molecules, like glucose ($C_6H_{12}O_6$). Because this process involves synthesizing a larger, energy-storing molecule, it requires an input of energy to proceed. The synthesis of glucose is described by this equation (notice that it is the reverse of the previous equation):

$6CO_2 + 6H_2O +$ energy C6H12O6+6O2→6CO2+6H2O+energyC6H12O6+6O2→6CO2+6H2O+energy size 12{C rSub { size 8{6} } H rSub { size 8{12} } O rSub { size 8{2} } } {}">→ $C_6H_{12}O_6 + 6O_2$

During the chemical reactions of photosynthesis, energy is provided in the form of a very high-energy molecule called ATP, or adenosine triphosphate, 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. The sugar (glucose) is stored as starch or glycogen. Energy-storing polymers like these are broken down into glucose to supply molecules of ATP. Plant cells use solar energy, energy from the sun, to synthesize the ATP they need to power the reactions of photosynthesis.

Figure 2. *Plants, like this oak tree and acorn, use energy from sunlight to make sugar and other organic molecules. Both plants and animals (like this squirrel) use cellular respiration to derive energy from the organic molecules originally produced by plants. (credit "acorn": modification of work by Noel Reynolds; credit "squirrel": modification of work by Dawn Huczek)*

Metabolic Pathways

The processes of making and breaking down sugar molecules illustrate two types of metabolic pathways. A metabolic pathway is a series of interconnected biochemical reactions that convert a substrate molecule or molecules, step-by-step, through a series of metabolic intermediates, eventually yielding a final product or products. In the case 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 (building) and catabolic (breaking down) pathways, respectively. Consequently, metabolism is composed of building (anabolism) and degradation (catabolism).

EVOLUTION CONNECTION

Evolution of Metabolic Pathways

There is more to the complexity of metabolism than understanding the metabolic pathways alone. Metabolic complexity varies from organism to organism. Photosynthesis is the primary pathway in which photosynthetic organisms like plants (the majority of global synthesis is done by planktonic algae) harvest the sun's energy and convert it into carbohydrates. The by-product of photosynthesis is oxygen, required by some cells to carry out cellular respiration. During cellular respiration, oxygen aids in the catabolic breakdown of carbon compounds, like carbohydrates. Among the products of this catabolism are CO_2 and ATP. In addition, some eukaryotes perform catabolic processes without oxygen (fermentation); that is, they perform or use anaerobic metabolism.

Organisms probably evolved anaerobic metabolism to survive (living organisms came into existence about 3.8 billion years ago, when the atmosphere lacked oxygen). Despite the differences between organisms and the complexity of metabolism, researchers have found that all branches of life share some of the same metabolic pathways, suggesting that all organisms evolved from the same ancient common ancestor (***Figure 3***). Evidence indicates that over time, the pathways diverged, adding specialized enzymes to allow organisms to better adapt to their environment, thus increasing their chance to survive. However, the underlying principle remains that all organisms must harvest energy from their environment and convert it to ATP to carry out cellular functions.

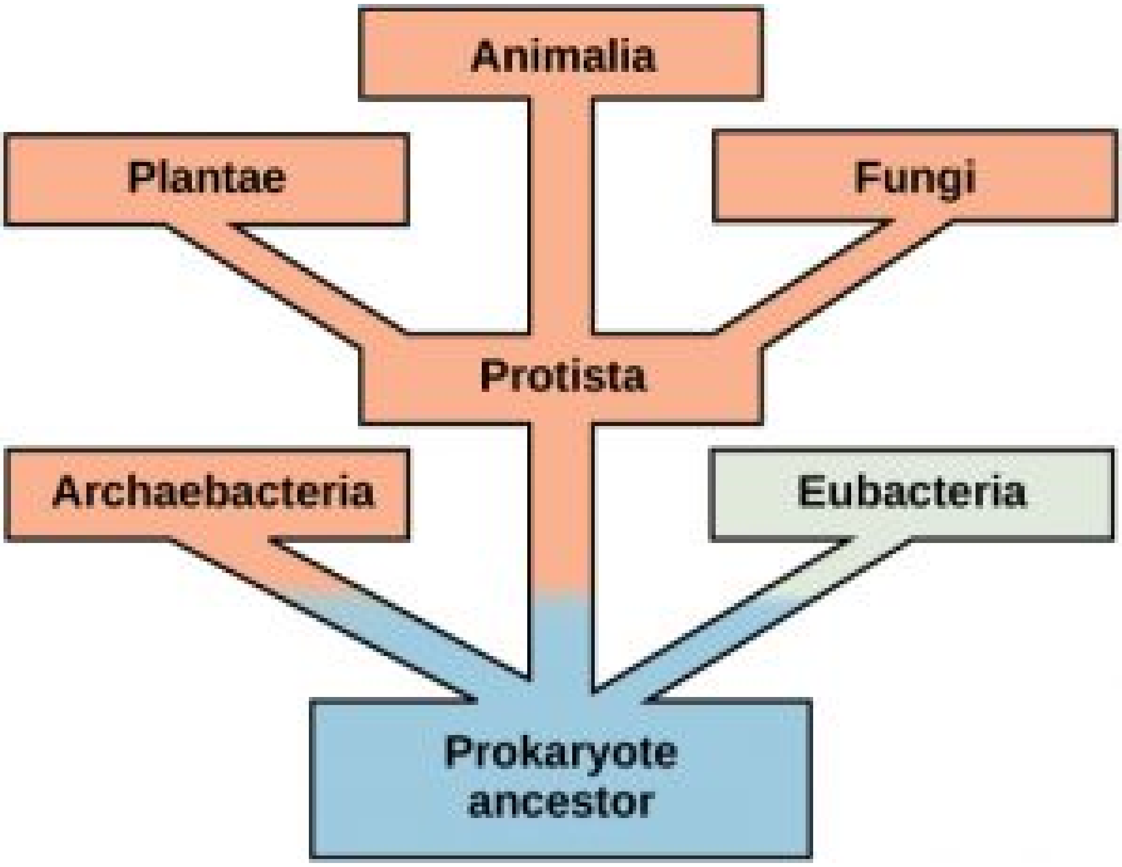


Figure 3. *This tree shows the evolution of the various branches of life. The vertical dimension is time. Early life forms, in blue, used anaerobic metabolism to obtain energy from their surroundings.*

Anabolic and Catabolic Pathways

Anabolic pathways require an input of energy to synthesize complex molecules from simpler ones. Synthesizing sugar from CO_2 is one example. Other examples are the synthesis of large proteins from amino acid building blocks, and the synthesis of new DNA strands from nucleic acid building blocks. These biosynthetic processes are critical to the life of the cell, take place constantly, and demand energy provided by ATP and other high-energy molecules (***Figure 4***).

ATP is an important molecule for cells to have in sufficient supply at all times. The breakdown of sugars illustrates how a single molecule of glucose can store enough energy to make a great deal of ATP, 36 to 38 molecules. This is a **catabolic** pathway. Catabolic pathways involve the degradation (or breakdown) of complex molecules into simpler ones. Molecular energy stored in the bonds of complex molecules is released in catabolic pathways and harvested in such a way that it can be used to produce ATP. Other energy-storing molecules, such as fats, are also broken down through similar catabolic reactions to release energy and make ATP (***Figure 4***).

It is important to know that the chemical reactions of metabolic pathways don't take place spontaneously. 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.

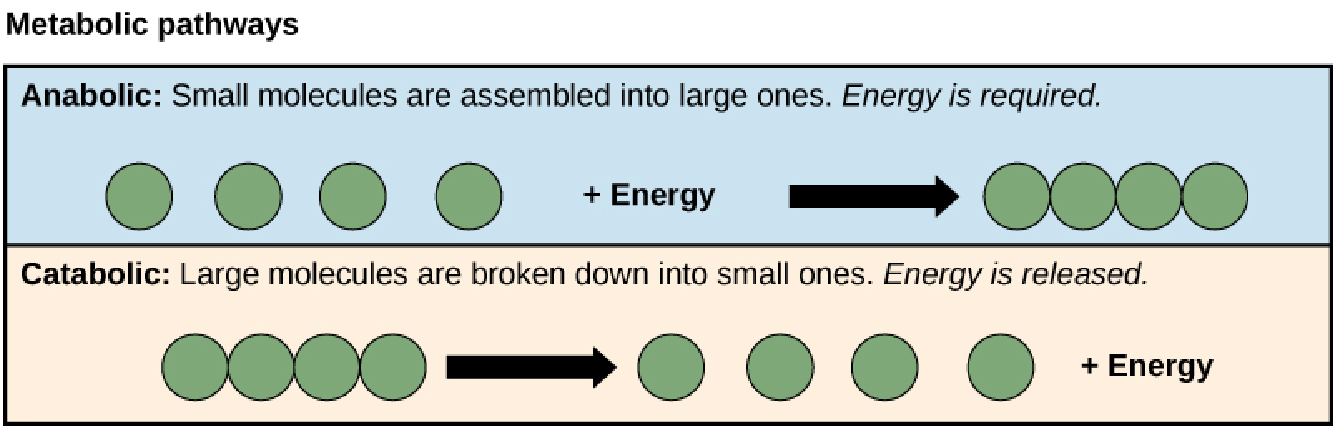


Figure 4. *Anabolic pathways are those that require energy to synthesize larger molecules. Catabolic pathways are those that generate energy by breaking down larger molecules. Both types of pathways are required for maintaining the cell's energy balance.*

Section Summary

Cells perform the functions of life through various chemical reactions. A cell's metabolism refers to the chemical reactions that take place within it. There are metabolic reactions that involve the breaking down of complex chemicals into simpler ones, such as the breakdown of large macromolecules. This process is referred to as catabolism, and such reactions are associated with a release of energy. On the other end of the spectrum, anabolism refers to metabolic processes that build complex molecules out of simpler ones, such as the synthesis of macromolecules. Anabolic processes require energy. Glucose synthesis and glucose breakdown are examples of anabolic and catabolic pathways, respectively.

Glossary

anabolic
: (also, anabolism) pathways that require an input of energy to synthesize complex molecules from simpler ones

bioenergetics
: study of energy flowing through living systems

catabolic
: (also, catabolism) pathways in which complex molecules are broken down into simpler ones

metabolism
: all the chemical reactions that take place inside cells, including anabolism and catabolism

31.

3.9 Enzymes

A substance that helps a chemical reaction to occur is a catalyst, and the special molecules that catalyze biochemical reactions are enzymes. Almost all enzymes are proteins, comprised of amino acid chains. Enzymes facilitate chemical reactions by binding to the reactant molecules, and holding them in such a way as to make the chemical bond-breaking and bond-forming processes take place more readily.

Enzyme Active Site and Substrate Specificity

The chemical reactants to which an enzyme binds are the enzyme's **substrates**. There may be one or more substrates, depending on the particular chemical reaction. In some reactions, a single-reactant substrate breaks down into multiple products. In others, two substrates may come together to create one larger molecule. Two reactants might also enter a reaction, both become modified, and leave the reaction as two products. The location within the enzyme where the substrate binds is the enzyme's **active site**. This is where the "action" happens. Since enzymes are proteins, there is a unique combination of amino acid residues (also side chains, or R groups) within the active site. Different properties characterize each residue. These can be large or small, weakly acidic or basic, hydrophilic or hydrophobic, positively or negatively charged, or neutral. The unique combination of amino acid residues, their positions, sequences, structures, and properties, creates a very specific chemical environment within the active site. This specific environment is suited to bind, albeit briefly, to a specific chemical substrate (or substrates). Due to this jigsaw puzzle-like match between an enzyme and its substrates (which adapts to find the best fit between the transition state and the active site), enzymes are known for their specificity. The "best fit" results from the shape and the amino acid functional group's attraction to the substrate. There is a specifically matched enzyme for each substrate and, thus, for each chemical reaction; however, there is flexibility as well.

The fact that active sites are so perfectly suited to provide specific environmental conditions also means that they are subject to local environmental influences. It is true that increasing the environmental temperature generally increases reaction rates, enzyme-catalyzed or otherwise. However, increasing or decreasing the temperature outside of an optimal range can affect chemical bonds within the active site in such a way that they are less well suited to bind substrates. High temperatures will eventually cause enzymes, like other biological molecules, to **denature**, a process that changes the substance's natural properties. Likewise, the local environment's pH can also affect enzyme function. Active site amino acid residues have their own acidic or basic properties that are optimal for catalysis. These residues are sensitive to changes in pH that can impair the way substrate molecules bind. Enzymes are suited to function best within a certain pH range, and, as with temperature, extreme environmental pH values (acidic or basic) can cause enzymes to denature.

Induced Fit and Enzyme Function

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 more refined view scientists call **induced fit** (***Figure 1***). This model expands upon the lock-and-key model by describing a more dynamic interaction between enzyme and substrate. As the enzyme and substrate come together, their interaction causes a mild shift in the enzyme’s structure that confirms an ideal binding arrangement between the enzyme and the substrate’s transition state. This ideal binding maximizes the enzyme’s ability to catalyze its reaction.

Link to Learning

View an induced fit animation at this website.

When an enzyme binds its substrate, it forms an enzyme-substrate complex. This complex promotes the reaction’s rapid progression in one of many ways. On a basic level, enzymes promote chemical reactions that involve more than one substrate by bringing the substrates together in an optimal orientation. The appropriate region (atoms and bonds) of one molecule is juxtaposed to the other molecule’s appropriate region with which it must react. Another way in which enzymes promote substrate reaction is by creating an optimal environment within the active site for the reaction to occur. Certain chemical reactions might proceed best in a slightly acidic or non-polar environment. The chemical properties that emerge from the particular arrangement of amino acid residues within an active site create the perfect environment for an enzyme’s specific substrates to react.

The enzyme-substrate complex can also facilitate reactions by contorting substrate molecules in such a way as to facilitate bond-breaking, helping to reach the reaction to proceed. Finally, enzymes can also facilitate reactions by taking part in the chemical reaction itself. The amino acid residues can provide certain ions or chemical groups that actually form covalent bonds with substrate molecules as a necessary step of the reaction process. In these cases, it is important to remember that the enzyme will always return to its original state at the reaction’s completion. One of enzymes’ hallmark properties is that they remain ultimately unchanged by the reactions they catalyze. After an enzyme catalyzes a reaction, it releases its product(s).

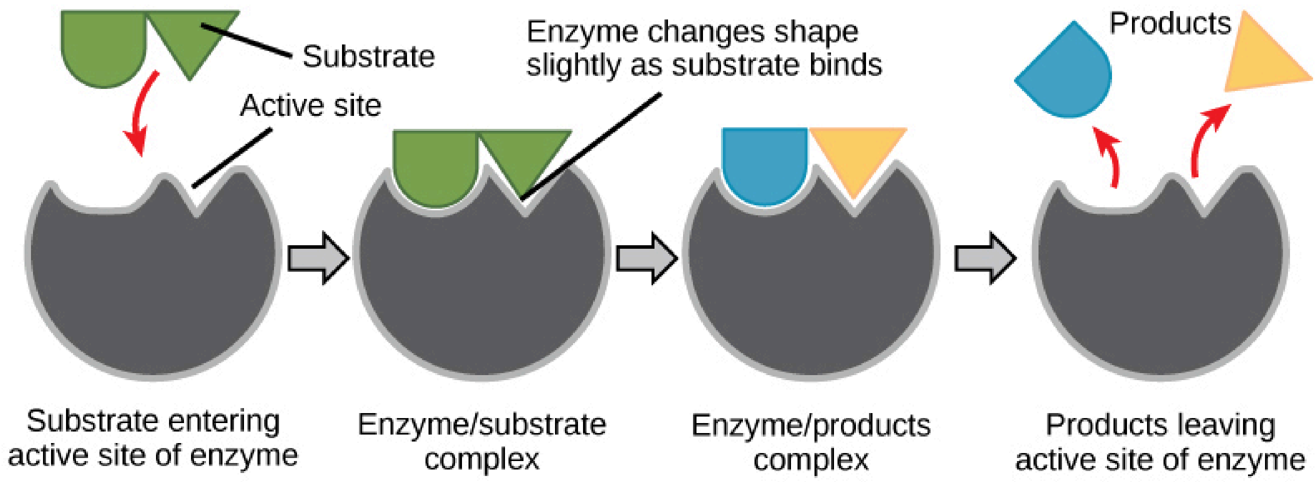


Figure 1. *According to the induced-fit model, both enzyme and substrate undergo dynamic conformational changes upon binding. The enzyme contorts the substrate into its transition state, thereby increasing the reaction's rate.*

Metabolism Control Through Enzyme Regulation

It would seem ideal to have a scenario in which all the encoded enzymes in an organism's genome existed in abundant supply and functioned optimally under all cellular conditions, in all cells, at all times. In reality, this is far from the case. A variety of mechanisms ensure that this does not happen. Cellular needs and conditions vary from cell to cell, and change within individual cells over time. The required enzymes and energetic demands of stomach cells are different from those of fat storage cells, skin cells, blood cells, and nerve cells. Furthermore, a digestive 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 do the amounts and functionality of different enzymes.

The relative amounts and functioning of the variety of enzymes within a cell ultimately determine which reactions will proceed and at which rates in that cell. This determination is tightly controlled. In certain cellular environments, environmental factors like pH and temperature partly control enzyme activity. There are other mechanisms through which cells control enzyme activity and determine the rates at which various biochemical reactions will occur.

Molecular Regulation of Enzymes

Enzymes can be regulated in ways that either promote or reduce their activity. There are many different kinds of molecules that inhibit or promote enzyme function, and various mechanisms exist for doing so. For example, 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. Alternatively, in noncompetitive inhibition, an inhibitor molecule binds to the enzyme in a location other than an active site, a binding site away from the active site, and still manages to block substrate binding to the active site.

EVERYDAY CONNECTION

Drug Discovery by Looking for Inhibitors of Key Enzymes in Specific Pathways

Enzymes are key components of metabolic pathways. Understanding how enzymes work and how they can be regulated is a key principle behind developing many pharmaceutical drugs (***Figure 2***) on the market today. Biologists working in this field collaborate with other scientists, usually chemists, to design drugs.

Consider statins for example—which is a class of drugs that reduces cholesterol levels. These compounds are essentially inhibitors of the enzyme HMG-CoA reductase. HMG-CoA reductase is the enzyme that synthesizes cholesterol from lipids in the body. By inhibiting this enzyme, the drug reduces cholesterol levels synthesized in the body. Similarly, acetaminophen, popularly marketed under the brand name Tylenol, is an inhibitor of the enzyme cyclooxygenase. While it is effective in providing relief from fever and inflammation (pain), scientists still do not completely understand its mechanism of action.

How are drugs developed? One of the first challenges in drug development is identifying the specific molecule that the drug is intended to target. In the case of statins, HMG-CoA reductase is the drug target. Researchers identify targets through painstaking research in the laboratory. Identifying the target alone is not sufficient. Scientists also need to know how the target acts inside the cell and which reactions go awry in the case of disease. Once researchers identify the target and the pathway, then the actual drug design process begins. During this stage, chemists and biologists work together to design and synthesize molecules that can either block or activate a particular reaction. However, this is only the beginning: both if and when a drug prototype is successful in performing its function, then it must undergo many tests from *in vitro* experiments to clinical trials before it can obtain FDA approval to be on the market.

Figure 2. *Have you ever wondered how pharmaceutical drugs are developed? (credit: Deborah Austin)*

Many enzymes don't work optimally, or even at all, unless bound to other specific non-protein helper molecules, either temporarily through ionic or hydrogen bonds or permanently through stronger covalent bonds. Two types of helper molecules are **cofactors** and **coenzymes**. Binding to these molecules promotes optimal conformation and function for their respective enzymes. Cofactors are inorganic ions such as iron (Fe^{++}) and magnesium (Mg^{++}). One example of an enzyme that requires a metal ion as a cofactor is the enzyme that builds DNA molecules, DNA polymerase, which requires a bound zinc ion (Zn^{++}) to function. Coenzymes are organic helper molecules, with a basic atomic structure comprised of carbon and hydrogen, which are required for enzyme action. The most common sources of coenzymes are dietary vitamins (***Figure 3***). Some vitamins are precursors to coenzymes and others act directly as coenzymes. Vitamin C is a coenzyme for multiple enzymes that take part in building the important connective tissue component, collagen. An important step in breaking down glucose to yield energy is catalysis by a multi-enzyme complex scientists call pyruvate dehydrogenase. Pyruvate dehydrogenase is a complex of several enzymes that actually requires one cofactor (a magnesium ion) and five different organic coenzymes to catalyze its specific chemical reaction. Therefore, enzyme function is, in part, regulated by an abundance of various cofactors and coenzymes, which the diets of most organisms supply.

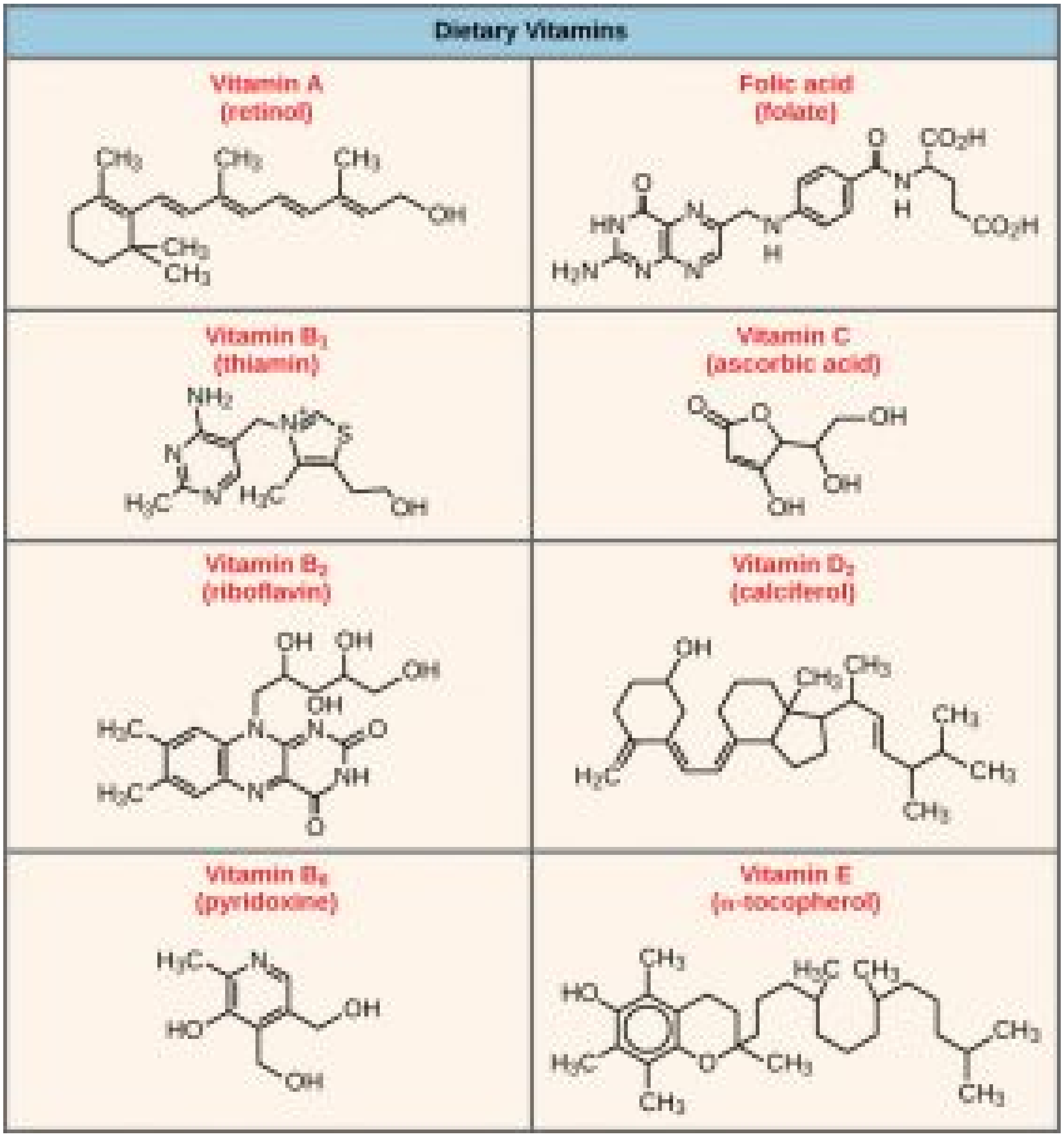


Figure 3. *Vitamins are important coenzymes or precursors of coenzymes, and are required for enzymes to function properly. Multivitamin capsules usually contain mixtures of all the vitamins at different percentages.*

Enzyme Compartmentalization

In animal cells, molecules such as enzymes are usually compartmentalized into different organelles. This allows for yet another level of regulation of enzyme activity. Enzymes required only for certain cellular processes are sometimes housed separately along with their substrates, allowing for more efficient chemical reactions. Examples of this sort of enzyme regulation based on location and proximity include the enzymes involved in the latter stages of cellular respiration, which take place exclusively in the mitochondria, and the enzymes involved in digesting cellular debris and foreign materials, located within lysosomes.

Section Summary

Enzymes are chemical catalysts that accelerate chemical reactions at physiological temperatures by lowering their activation energy. Enzymes are usually proteins consisting of one or more polypeptide chains. Enzymes have an active site that provides a unique chemical environment, comprised of certain amino acid R groups (residues). This unique environment is perfectly suited to convert particular chemical reactants for that enzyme, scientists call substrates, into unstable intermediates that they call transition states. Enzymes and substrates bind with an induced fit, which means that enzymes undergo slight conformational adjustments upon substrate contact, leading to full, optimal binding. Enzymes bind to substrates and catalyze reactions in four different ways: bringing substrates together in an optimal orientation, compromising the bond structures of substrates so that bonds can break down more easily, providing optimal environmental conditions for a reaction to occur, or participating directly in their chemical reaction by forming transient covalent bonds with the substrates.

Enzyme action must be regulated so that in a given cell at a given time, the desired reactions catalyze and the undesired reactions are not. Enzymes are regulated by cellular conditions, such as temperature and pH. They are also regulated through their location within a cell, sometimes compartmentalized so that they can only catalyze reactions under certain circumstances. Enzyme inhibition and activation via other molecules are other important ways that enzymes are regulated.

Glossary

active site
: enzyme's specific region to which the substrate binds

coenzyme
: small organic molecule, such as a vitamin or its derivative, which is required to enhance an enzyme's activity

cofactor
: inorganic ion, such as iron and magnesium ions, required for optimal enzyme activity regulation

denature
: process that changes a subtance's natural properties

induced fit
: dynamic fit between the enzyme and its substrate, in which both components modify their structures to allow for ideal binding

substrate
: molecule on which the enzyme acts

32.

3.10 Musculoskeletal System

The muscular and skeletal systems provide support to the body and allow for movement. The bones of the skeleton protect the body's internal organs and support the weight of the body. The muscles of the muscular system contract and pull on the bones, allowing for movements as diverse as standing, walking, running, and grasping items. Muscle contraction is an energy intensive process, requiring large amounts of ATP hydrolysis.

Injury or disease affecting the musculoskeletal system can be very debilitating. The most common musculoskeletal diseases worldwide are caused by malnutrition, which can negatively affect development and maintenance of bones and muscles. Other diseases affect the joints, such as arthritis, which can make movement difficult and, in advanced cases, completely impair mobility.

Progress in the science of prosthesis design has resulted in the development of artificial joints, with joint replacement surgery in the hips and knees being the most common. Replacement joints for shoulders, elbows, and fingers are also available.

Skeletal System

The human skeleton consists of 206 bones in the adult and has five main functions: providing support to the body, storing minerals and lipids, producing blood cells, protecting internal organs, and allowing for movement. The skeletal system divided into the axial skeleton (which consists of the skull, vertebral column, and rib cage), and the appendicular skeleton (which consists of limb bones, the pectoral or shoulder girdle, and the pelvic girdle).

Concept in Action

Explore the human skeleton by viewing the following video with digital 3D sculpturing.

The **axial skeleton** forms the central axis of the body and includes the bones of the skull, ossicles of the middle ear, hyoid bone of the throat, vertebral column, and the thoracic cage (rib cage) (***Figure 1***).

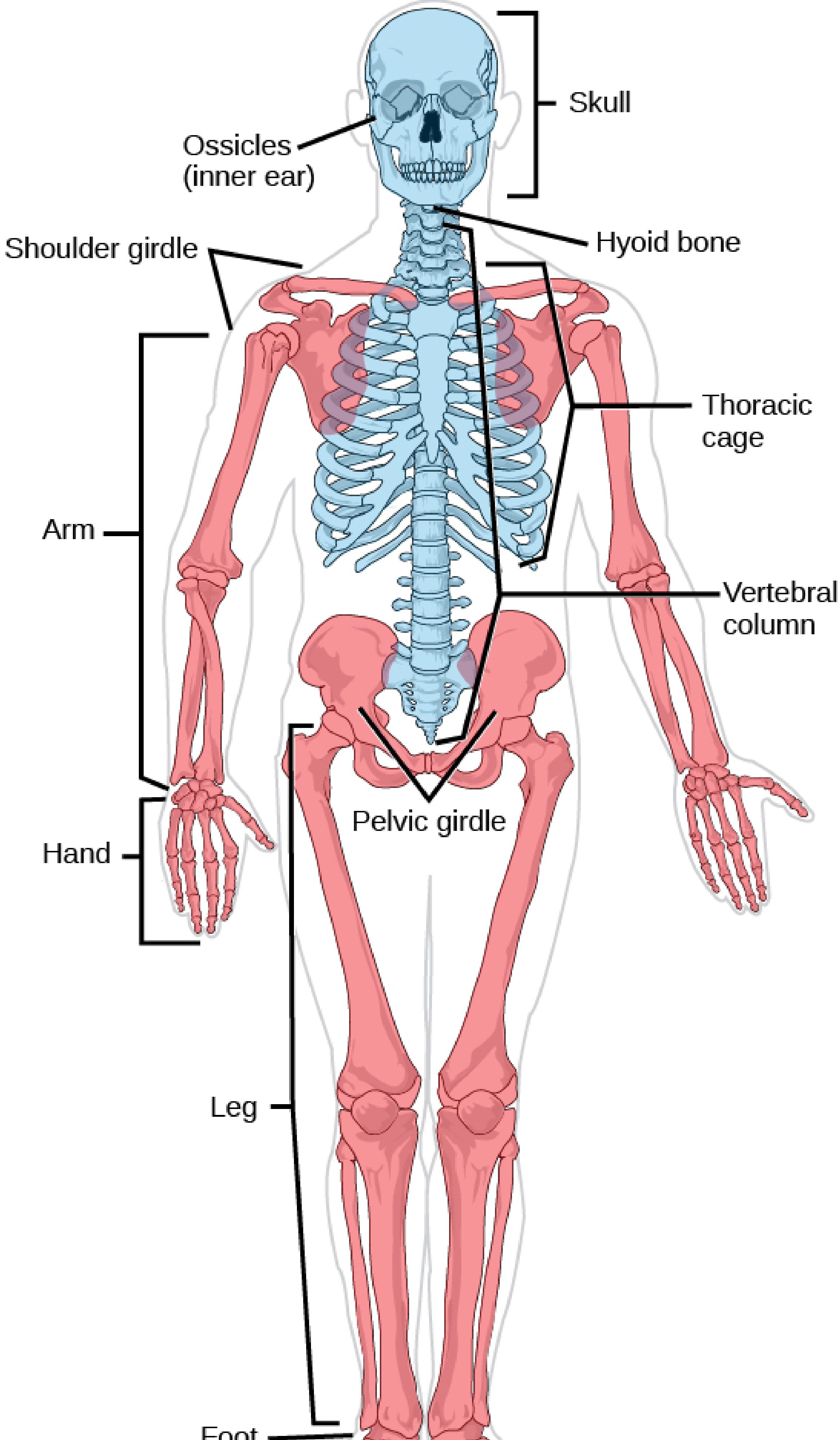

Skull
Ossicles (inner ear)
Hyoid bone
Shoulder girdle
Thoracic cage
Arm
Vertebral column
Pelvic girdle
Hand
Leg
Foot

Figure 1. *The axial skeleton, shown in blue, consists of the bones of the skull, ossicles of the middle ear, hyoid bone, vertebral column, and thoracic cage. The appendicular skeleton, shown in red, consists of the bones of the pectoral limbs, pectoral girdle, pelvic limb, and pelvic girdle. (credit: modification of work by Mariana Ruiz Villareal)*

The bones of the **skull** support the structures of the face and protect the brain. The skull consists of cranial bones and facial bones. The cranial bones form the cranial cavity, which encloses the brain and serves as an attachment site for muscles of the head and neck. In the adult they are tightly jointed with connective tissue and adjoining bones do not move.

The **auditory ossicles** of the middle ear transmit sounds from the air as vibrations to the fluid-filled cochlea. The auditory ossicles consist of two malleus (hammer) bones, two incus (anvil) bones, and two stapes (stirrups), one on each side. Facial bones provide cavities for the sense organs (eyes, mouth, and nose), and serve as attachment points for facial muscles.

The **hyoid bone** lies below the mandible in the front of the neck. It acts as a movable base for the tongue and is connected to muscles of the jaw, larynx, and tongue. The mandible forms a joint with the base of the skull. The mandible controls the opening to the mouth and hence, the airway and gut.

The **vertebral column**, or spinal column, surrounds and protects the spinal cord, supports the head, and acts as an attachment point for ribs and muscles of the back and neck. It consists of 26 bones: the 24 vertebrae, the sacrum, and the coccyx. Each vertebral body has a large hole in the center through which the spinal cord passes down to the level of the first lumbar vertebra. Below this level, the hole contains spinal nerves which exit between the vertebrae. There is a notch on each side of the hole through which the spinal nerves, can exit from the spinal cord to serve different regions of the body. The vertebral column is approximately 70 cm (28 in) in adults and is curved, which can be seen from a side view.

Intervertebral discs composed of fibrous cartilage lie between adjacent vertebrae from the second cervical vertebra to the sacrum. Each disc helps form a slightly moveable joint and acts as a cushion to absorb shocks from movements such as walking and running.

The **thoracic cage**, also known as the rib cage consists of the ribs, sternum, thoracic vertebrae, and costal cartilages. The thoracic cage encloses and protects the organs of the thoracic cavity including the heart and lungs. It also provides support for the shoulder girdles and upper limbs and serves as the attachment point for the diaphragm, muscles of the back, chest, neck, and shoulders. Changes in the volume of the thorax enable breathing. The sternum, or breastbone, is a long flat bone located at the anterior of the chest. Like the skull, it is formed from many bones in the embryo, which fuse in the adult. The ribs are 12 pairs of long curved bones that attach to the thoracic vertebrae and curve toward the front of the body, forming the ribcage. Costal cartilages connect the anterior ends of most ribs to the sternum.

The **appendicular skeleton** is composed of the bones of the upper and lower limbs. It also includes the pectoral, or shoulder girdle, which attaches the upper limbs to the body, and the pelvic girdle, which attaches the lower limbs to the body (***Figure 1***).

The **pectoral girdle** bones transfer force generated by muscles acting on the upper limb to the thorax. It consists of the clavicles (or collarbones) in the anterior, and the scapulae (or shoulder blades) in the posterior.

The upper limb contains bones of the arm (shoulder to elbow), the forearm, and the hand. The humerus is the largest and longest bone of the upper limb. It forms a joint with the shoulder and with the forearm at the elbow. The forearm extends from the elbow to the wrist and consists of two bones. The hand includes the bones of the wrist, the palm, and the bones of the fingers.

The **pelvic girdle** attaches to the lower limbs of the axial skeleton. Since it is responsible for bearing the weight of the body and for locomotion, the pelvic girdle is securely attached to the axial skeleton by strong ligaments. It also has deep sockets with robust ligaments that securely attach to the femur. The pelvic girdle is mainly composed of two large hip bones. The hip bones join together in the anterior of the body at a joint called the pubic symphysis and with the bones of the sacrum at the posterior of the body.

The lower limb consists of the thigh, the leg, and the foot. The bones of the lower limbs are thicker and stronger than the bones of the upper limbs to support the entire weight of the body and the forces from locomotion. The femur, or thighbone, is the longest, heaviest, and strongest bone in the body. The femur and pelvis form the hip joint. At its other end, the femur, along with the shinbone and kneecap, form the knee joint.

Joints and Skeletal Movement

The point at which two or more bones meet is called a **joint**, or articulation. Joints are responsible for movement, such as the movement of limbs, and stability, such as the stability found in the bones of the skull.

Joints can be classified based on their structure or functions. The structural classification divides joints into fibrous, cartilaginous, and synovial joints depending on the material composing the joint and the presence or absence of a cavity in the joint, as shown in ***Figure 2a***. Fibrous joints do not allow much movement and function to tightly connect bones. Cartilaginous joints are slighlty more flexible, but still primarily function to provide connection and stability to bones. Synovial joints are capable of the greatest movement of the joint types, and therefore have the widest variety of functions. Knees, elbows, and shoulders are examples of synovial joints.

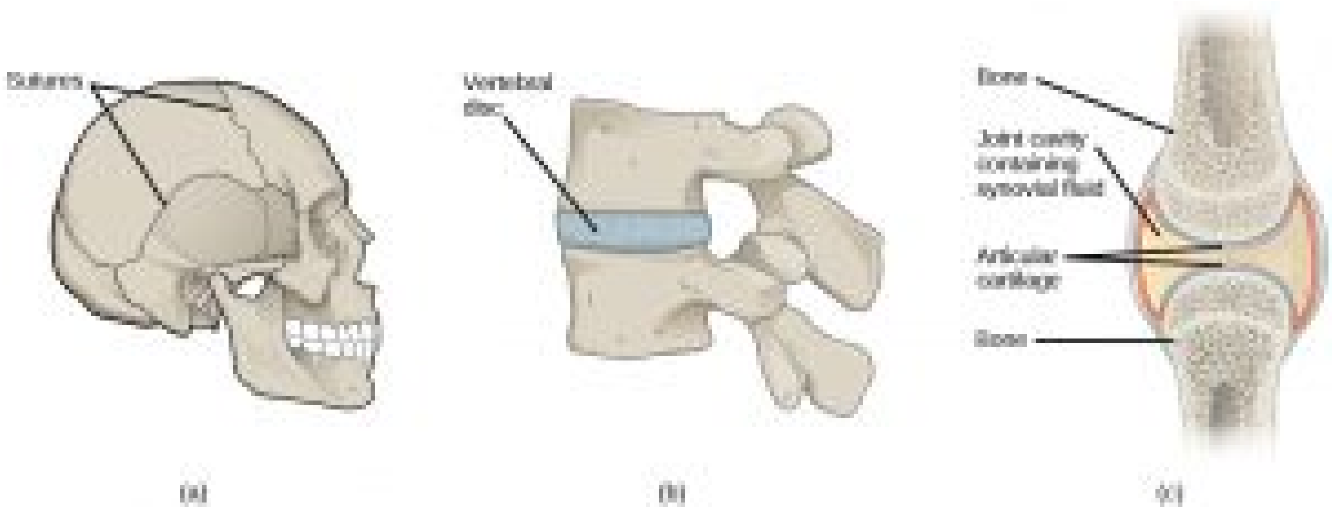


Figure 2. (a) *Sutures are fibrous joints found only in the skull.* ***(b)*** *Cartilaginous joints are bones connected by cartilage, such as between vertebrae.* ***(c)*** *Synovial joints are the only joints that have a space or "synovial cavity" in the joint.*

CAREER IN ACTION

Rheumatologist

Rheumatologists are medical doctors who specialize in the diagnosis and treatment of disorders of the joints, muscles, and bones. They diagnose and treat diseases such as arthritis, musculoskeletal disorders, osteoporosis, plus autoimmune diseases like ankylosing spondylitis, a chronic spinal inflammatory disease and rheumatoid arthritis.

Rheumatoid arthritis (RA) is an inflammatory disorder that primarily affects synovial joints of the hands, feet, and cervical spine. Affected joints become swollen, stiff, and painful. Although it is known that RA is an autoimmune disease in which the body's immune system mistakenly attacks healthy tissue, the exact cause of RA remains unknown. Immune cells from the blood enter joints and the joint capsule causing cartilage breakdown and swelling of the joint lining. Breakdown of cartilage causes bones to rub against each other causing pain. RA is more common in women than men and the age of onset is usually between 40 to 50 years.

Rheumatologists can diagnose RA based on symptoms such as joint inflammation and pain, x-ray and MRI imaging, and blood tests. Arthrography is a type of medical imaging of joints that uses a contrast agent, such as a dye that is opaque to x-rays. This allows the soft tissue structures of joints—such as cartilage, tendons, and ligaments—to be visualized. An arthrogram differs from a regular x-ray by showing the surface of soft tissues lining the joint in addition to joint bones. An arthrogram allows early degenerative changes in joint cartilage to be detected before bones become affected.

There is currently no cure for RA; however, rheumatologists have a number of treatment options available. Treatments are divided into those that reduce the symptoms of the disease and those that reduce the damage to bone and cartilage caused by the disease. Early stages can be treated with rest of the affected joints through the use of a cane, or with joint splints that minimize inflammation. When inflammation has decreased, exercise can be used to strengthen muscles that surround the joint and to maintain joint flexibility. If joint damage is more extensive, medications can be used to relieve pain and decrease inflammation. Anti-inflammatory drugs that may be used include aspirin, topical pain relievers, and corticosteroid injections. Surgery may be required in cases where joint damage is severe. Physicians are now using drugs that reduce the damage to bones and cartilage caused by the disease to slow its development. These drugs are diverse in their mechanisms but they all act to reduce the impact of the autoimmune response, for example by inhibiting the inflammatory response or reducing the number of T lymphocytes, a cell of the immune system.

Muscles

Muscles allow for movement such as walking, and they also facilitate bodily processes such as respiration and digestion. The body contains three types of muscle tissue: skeletal muscle, cardiac muscle, and smooth muscle (***Figure 3***).

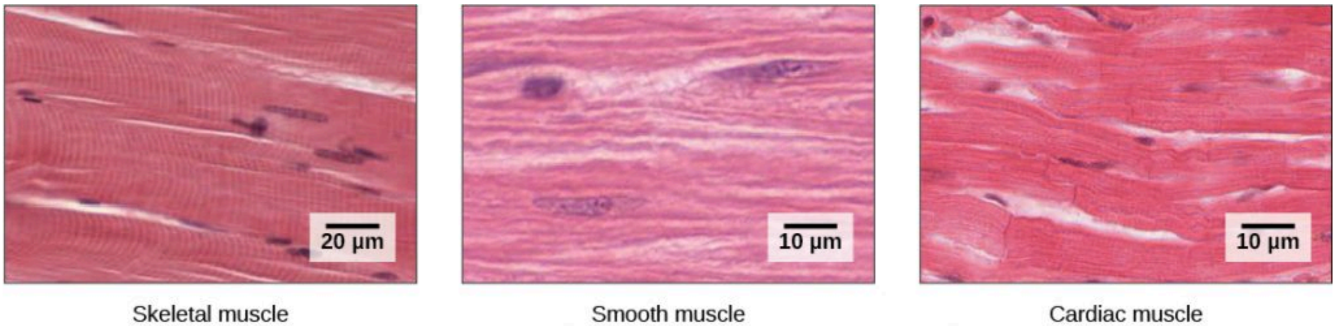


Figure 3. *The body contains three types of muscle tissue: skeletal muscle, smooth muscle, and cardiac muscle. Notice that skeletal muscle cells are long and cylindrical, they have multiple nuclei, and the small, dark nuclei are pushed to the periphery of the cell. Smooth muscle cells are short, tapered at each end, and have only one nucleus each. Cardiac muscle cells are also cylindrical, but short. The cytoplasm may branch, and they have one or two nuclei in the center of the cell. (credit: modification of work by NCI, NIH; scale-bar data from Matt Russell)*

Skeletal muscle tissue forms skeletal muscles, which attach to bones and sometimes the skin and control locomotion and any other movement that can be consciously controlled. Because it can be controlled intentionally, skeletal muscle is also called voluntary muscle. When viewed under a microscope, skeletal muscle tissue has a striped or striated appearance. This appearance results from the arrangement of the proteins inside the cell that are responsible for contraction. The cells of skeletal muscle are long and tapered and have multiple nuclei on the periphery of each cell.

Smooth muscle tissue occurs in the walls of hollow organs such as the intestines, stomach, and urinary bladder, and around passages such as in the respiratory tract and blood vessels. Smooth muscle has no striations, is not under voluntary control, and is called involuntary muscle. Smooth muscle cells have a single nucleus.

Cardiac muscle tissue is only found in the heart. The contractions of cardiac muscle tissue pump blood throughout the body and maintain blood pressure. Like skeletal muscle, cardiac muscle is striated, but unlike skeletal muscle, cardiac muscle cannot be consciously controlled and is called involuntary muscle. The cells of cardiac muscle tissue are connected to each other through intercalated disks and usually have just one nucleus per cell.

Skeletal Muscle Fiber Structure and Function

Each skeletal muscle fiber is a skeletal muscle cell. Within each muscle fiber are **myofibrils**, long cylindrical structures that lie parallel to the muscle fiber. Myofibrils run the entire length of the muscle fiber. They attach to the plasma membrane, called the **sarcolemma**, at their ends, so that as myofibrils shorten, the entire muscle cell contracts (***Figure 4***).

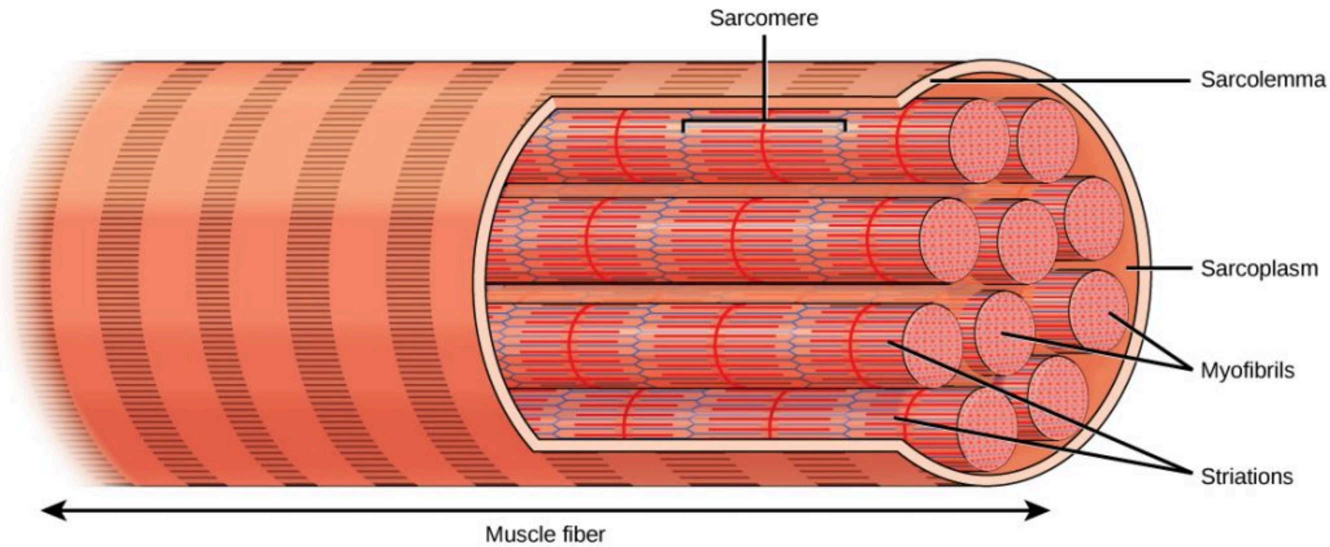


Figure 4. *A skeletal muscle fiber is surrounded by a plasma membrane called the sarcolemma, with a cytoplasm called the sarcoplasm. A muscle fiber is composed of many fibrils packaged into orderly units. The orderly arrangement of the proteins in each unit, shown as red and blue lines, gives the cell its striated appearance.*

The striated appearance of skeletal muscle tissue is a result of repeating bands of the proteins actin and myosin that occur along the length of myofibrils.

Myofibrils are composed of smaller structures called **myofilaments**. There are two main types of myofilaments: thick filaments and thin filaments. Thick filaments are composed of the protein myosin. The primary component of thin filaments is the protein actin.

The thick and thin filaments alternate with each other in a structure called a **sarcomere**. The sarcomere is the unit of contraction in a muscle cell. Contraction is stimulated by an electrochemical signal from a nerve cell associated with the muscle fiber. For a muscle cell to contract, the sarcomere must shorten. However, thick and thin filaments do not shorten. Instead, they slide by one another, causing the sarcomere to shorten while the filaments remain the same length. The sliding is accomplished when a molecular extension of myosin, called the myosin head, temporarily binds to an actin filament next to it and through a change in conformation, bends, dragging the two filaments in opposite directions. The myosin head then releases its actin filament, relaxes, and then repeats the process, dragging the two filaments further along each other. The combined activity of many binding sites and repeated movements within the sarcomere causes it to contract. The coordinated contractions of many sarcomeres in a myofibril leads to contraction of the entire muscle cell and ultimately the muscle itself. The movement of the myosin head requires ATP, which provides the energy for the contraction.

Millions of myofibrils can be found in each large muscle of the body, such as the biceps muscle in the upper arm. Even small muscles, such as those used to move the eye, contain hundreds of thousands of myofibrils. Each myofibril itself contains contains thousands of sarcomeres, and each of sarcomere usually contains hundreds of myosin heads. Complete contraction of a muscle requires each myosin head to hydrolyze many ATP molecules. Hydrolysis of billions of ATP molecules is therefore required for contraction of a large muscle, and ATP stores can be rapidly depleted when repeated muscle contractions occur during exercise.

Concept in Action

View this animation to see how muscle fibers are organized.

Section Summary

The human skeleton is an endoskeleton that is composed of the axial and appendicular skeleton. The axial skeleton is composed of the bones of the skull, ossicles of the ear, hyoid bone, vertebral column, and ribcage. The appendicular skeleton is made up of the upper and lower limbs.

The structural classification of joints divides them into fibrous, cartilaginous, and synovial joints. The bones of fibrous joints are held together by fibrous connective tissue. Cartilaginous joints are joints in which the bones are connected by cartilage. Synovial joints are joints that have a space between the adjoining bones.

The body contains three types of muscle tissue: skeletal muscle, cardiac muscle, and smooth muscle. Muscles are composed of individual cells called muscle fibers. Muscle fibers consist of myofilaments composed of the proteins actin and myosin arranged in units called sarcomeres. Contraction of the muscle occurs by the combined action of myosin and actin fibers sliding past each other when the myosin heads bind to the actin fiber, bend, disengage, and then repeat the process. Muscle contraction requires ATP hydrolysis.

Glossary

appendicular skeleton
: the skeleton composed of the bones of the upper limbs, which function to grasp and manipulate objects, and the lower limbs, which permit locomotion

auditory ossicles
: (also, middle ear bones) the bones that transduce sounds from the air into vibrations in the fluid-filled cochlea

axial skeleton
: skeleton that forms the central axis of the body and includes the bones of the skull, the ossicles of the middle ear, the hyoid bone of the throat, the vertebral column, and the thoracic cage (ribcage)

cardiac muscle tissue
: the muscle tissue found only in the heart; cardiac contractions pump blood throughout the body and maintain blood pressure

cartilaginous joint
: a joint in which the bones are connected by cartilage

fibrous joint
: a joint held together by fibrous connective tissue

hyoid bone
: the bone that lies below the mandible in the front of the neck

joint
: the point at which two or more bones meet

myofibril
: the long cylindrical structures that lie parallel to the muscle fiber

myofilament
: the small structures that make up myofibrils

pectoral girdle
: the bones that transmit the force generated by the upper limbs to the axial skeleton

pelvic girdle
: the bones that transmit the force generated by the lower limbs to the axial skeleton

sarcolemma
: the plasma membrane of a skeletal muscle fiber

sarcomere
: the functional unit of skeletal muscle

skeletal muscle tissue
: forms skeletal muscles, which attach to bones and control locomotion and any movement that can be consciously controlled

skull
: the bone that supports the structures of the face and protects the brain

smooth muscle tissue
: the muscle that occurs in the walls of hollow organs such as the intestines, stomach, and urinary bladder, and around passages such as the respiratory tract and blood vessels

synovial joints
: the only joints that have a space between the adjoining bones

thoracic cage
: (also, ribcage) the skeleton of the chest, which consists of the ribs, thoracic vertebrae, sternum, and costal cartilages

vertebral column
: (also, spine) the column that surrounds and protects the spinal cord, supports the head, and acts as an attachment point for ribs and muscles of the back and neck

33.

3.11 Hunger, Eating, and Weight

Eating is essential for survival, and it is no surprise that a drive like hunger exists to ensure that we seek out sustenance. While this chapter will focus primarily on the physiological mechanisms that regulate hunger and eating, powerful social, cultural, and economic influences also play important roles. This section will explain the regulation of hunger, eating, and body weight, and we will discuss the adverse consequences of disordered eating.

Physiological Mechanisms

There are a number of physiological mechanisms that serve as the basis for hunger. When our stomachs are empty, they contract. Typically, a person then experiences hunger pangs. Chemical messages travel to the brain, and serve as a signal to initiate feeding behavior, or eating. When our blood glucose levels drop, the pancreas and liver generate a number of chemical signals that induce hunger (and thus initiate feeding behavior.

For most people, once they have eaten, they feel **satiation**, or fullness and satisfaction, and their eating behavior stops. Like the initiation of eating, satiation is also regulated by several physiological mechanisms. As blood glucose levels increase, the pancreas and liver send signals to shut off hunger and eating . The food's passage through the gastrointestinal tract also provides important satiety signals to the brain (Woods, 2004), and fat cells release **leptin**, a satiety hormone.

The various **hunger** and satiety signals that are involved in the regulation of eating are integrated in the brain. Research suggests that several areas of the hypothalamus and hindbrain are especially important sites where this integration occurs. Ultimately, activity in the brain determines whether or not we engage in feeding behavior (***Figure 1***).

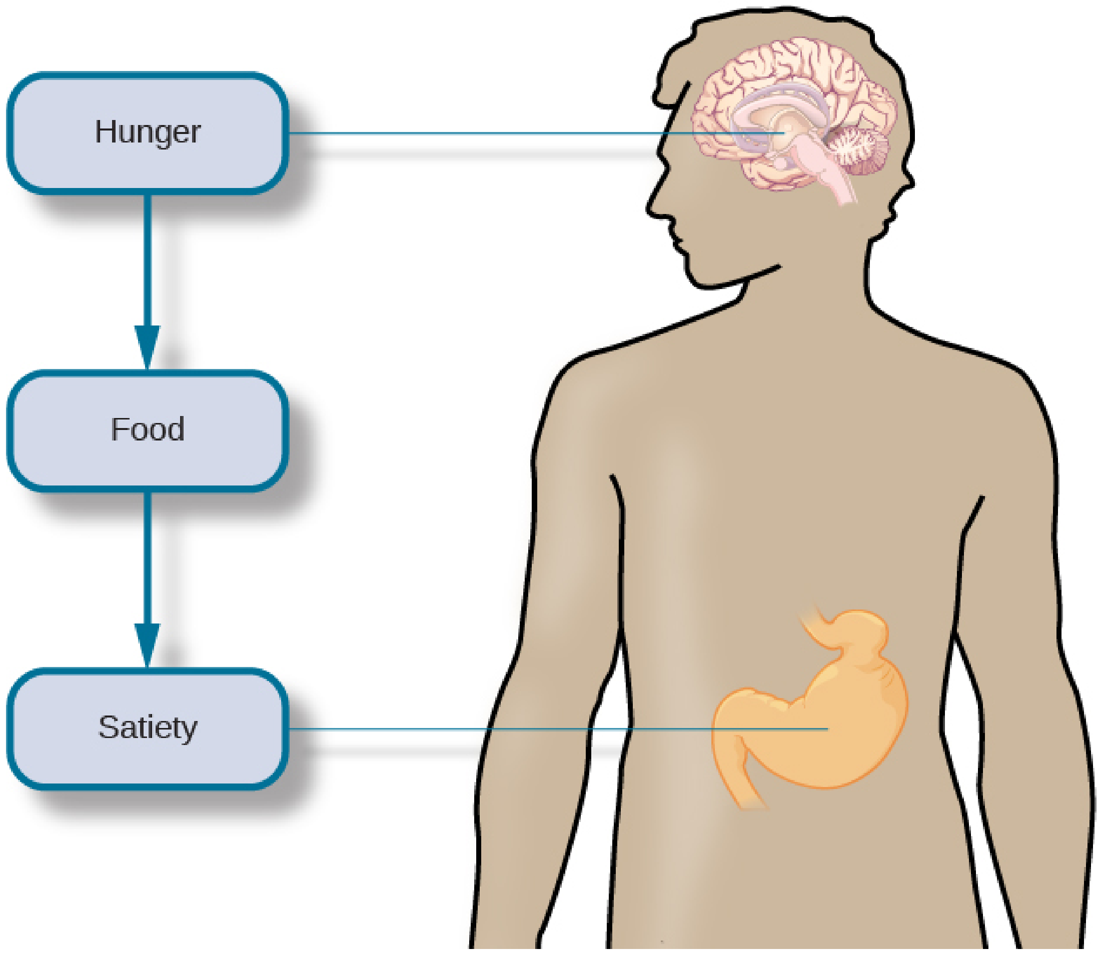


Figure 1. *Hunger and eating are regulated by a complex interplay of hunger and satiety signals that are integrated in the brain.*

Metabolism and Body Weight

Our body weight is affected by a number of factors, including gene-environment interactions, and the number of calories we consume versus the number of calories we burn in daily activity. If our caloric intake exceeds our caloric use, our bodies store excess energy in the form of fat. If we consume fewer calories than we burn off, then stored fat will be converted to energy. Our energy expenditure is obviously affected by our levels of activity, but our body's metabolic rate also comes into play. A person's **metabolic rate** is the amount of energy that is expended in a given period of time, and there is tremendous individual variability in our metabolic rates. People with high rates of metabolism are able to burn off calories more easily than those with lower rates of metabolism.

We all experience fluctuations in our weight from time to time, but generally, most people's weights fluctuate within a narrow margin, in the absence of extreme changes in diet and/or physical activity. This observation led some to propose a **set-point theory** of body weight regulation. The set-point theory asserts that each individual has an ideal body weight, or set point, which is resistant to change. This set-point is genetically predetermined and efforts to move our weight significantly from the set-point are resisted by compensatory changes in energy intake and/or expenditure.

Some of the predictions generated from this particular theory have not received empirical support. For example, there are no changes in metabolic rate between individuals who had recently lost significant amounts of weight and a control group (Weinsier et al., 2000). In addition, the set-point theory fails to account for the influence of social and environmental factors in the regulation of body weight . Despite these limitations, set-point theory is still often used as a simple, intuitive explanation of how body weight is regulated.

Obesity

When someone weighs more than what is generally accepted as healthy for a given height, they are considered overweight or obese. According to the Centers for Disease Control and Prevention (CDC), an adult with a **body mass index (BMI)** between 25 and 29.9 is considered **overweight** (***Figure 2***). An adult with a BMI of 30 or higher is considered **obese**. People who are so overweight that they are at risk for death are classified as morbidly obese. **Morbid obesity** is defined as having a BMI over 40. Note that although BMI has been used as a healthy weight indicator by the World Health Organization (WHO), the CDC, and other groups, its value as an assessment tool has been questioned. The BMI is most useful for studying populations, which is the work of these organizations. It is less useful in assessing an individual since height and weight measurements fail to account for important factors like fitness level. An athlete, for example, may have a high BMI because the tool doesn’t distinguish between the body’s percentage of fat and muscle in a person’s weight.

<img class=”wp-image-223 ” src=”https://open.lib.umn.edu/app/uploads/sites/272/2020/07/Figure-2-chart-BMI.png” alt=”A chart has an x-axis labeled “weight” (pounds/kilograms) and a y-axis labeled “height” (meters and feet/inches). Four areas are shaded different colors indicating the BMI for ranges of weight and height. The “underweight BMI 30” area covers approximately 140–350 pounds at height 4’11” and extends to approximately 265–350 pounds at height 6’6.”” width=”566″ height=”462″> **Figure 2.** This chart shows how adult BMI is calculated. Individuals find their height on the y-axis and their weight on the x-axis to determine their BMI.

Being extremely overweight or obese is a risk factor for several negative health consequences. These include, but are not limited to, an increased risk for cardiovascular disease, stroke, Type 2 diabetes, liver disease, sleep apnea, colon cancer, breast cancer, infertility, and arthritis. Given that it is estimated that in the United States around one-third of the adult population is obese and that nearly two-thirds of adults and one in six children qualify as overweight, there is substantial interest in trying to understand how to combat this important public health concern. The emphasis placed on weight and weight loss can also have profound psychological effects on individuals and contribute to increased stress and anxiety, which are themselves important public health concerns.

What causes someone to be overweight or obese? You have already read that both genes and environment are important factors for determining body weight, and if more calories are consumed than expended, excess energy is stored as fat. However, socioeconomic status and the physical environment must also be considered as contributing factors. For example, an individual who lives in an impoverished neighborhood that is overrun with crime may never feel comfortable walking or biking to work or to the local market. This might limit the amount of physical activity in which he engages and result in an increased body weight. Similarly, some people may not be able to afford healthy food options from their market, or these options may be unavailable (especially in urban areas or poorer neighborhoods); therefore, some people rely primarily on available, inexpensive, high fat, and high calorie fast food as their primary source of nutrition.

Generally, overweight and obese individuals are encouraged to try to reduce their weights through a combination of both diet and exercise. While some people are very successful with these approaches, many struggle to lose excess weight. In cases in which a person has had no success with repeated attempts to reduce weight or is at risk for death because of obesity, bariatric surgery may be recommended. **Bariatric surgery** is a type of surgery specifically aimed at weight reduction, and it involves modifying the gastrointestinal system to reduce the amount of food that can be eaten and/or limiting how much of the digested food can be absorbed (***Figure 3***).

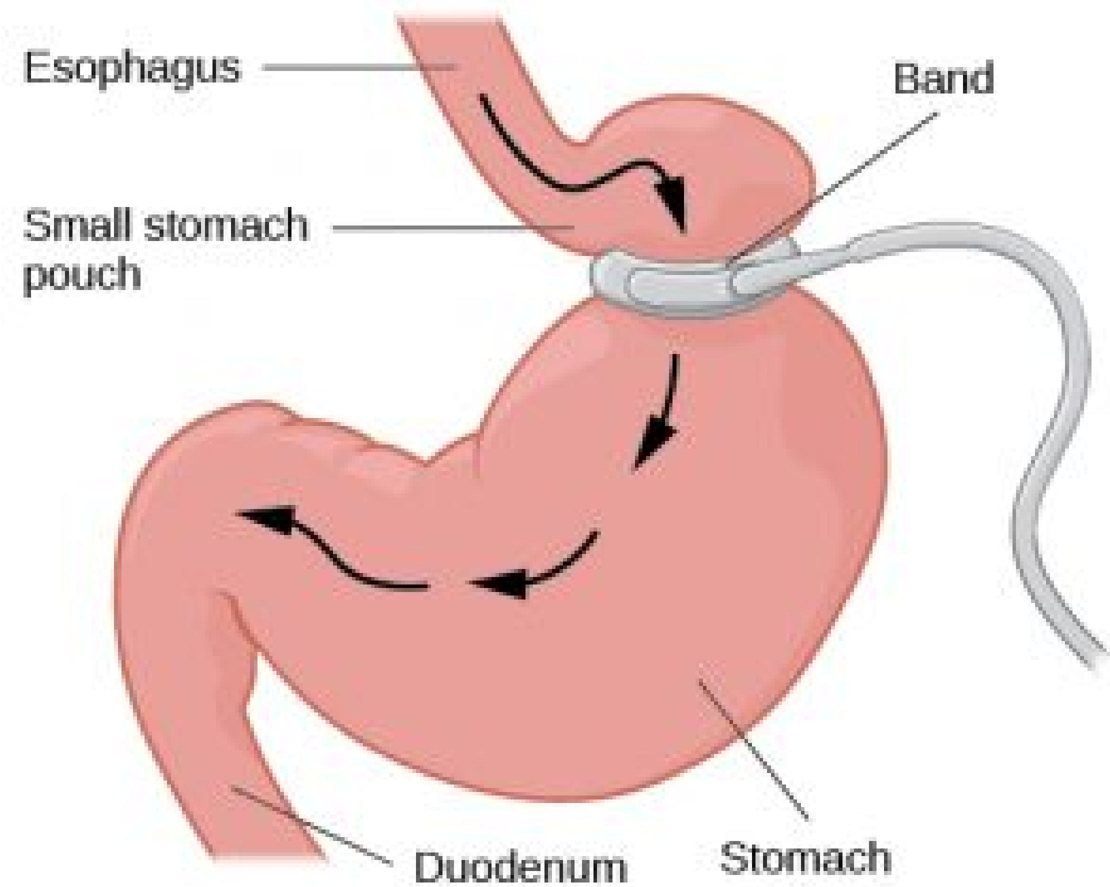


Figure 3. *Gastric banding surgery creates a small pouch of stomach, reducing the size of the stomach that can be used for digestion.*

Link to Learning

Watch this video that describes two different types of bariatric surgeries.

PRADER-WILLI SYNDROME

Prader-Willi Syndrome (PWS) is a genetic disorder that results in persistent feelings of intense hunger and reduced rates of metabolism. Typically, affected children have to be supervised around the clock to ensure that they do not engage in excessive eating. Currently, PWS is the leading genetic cause of morbid obesity in children, and it is associated with a number of cognitive deficits and emotional problems (***Figure 4***).

Figure 4. *Eugenia Martínez Vallejo, depicted in this 1680 painting, may have had Prader-Willi syndrome. At just eight years old, she weighed approximately 120 pounds, and she was nicknamed "La Monstrua" (the monster).*

While genetic testing can be used to make a diagnosis, there are a number of behavioral diagnostic criteria associated with PWS. From birth to 2 years of age, lack of muscle tone and poor sucking behavior may serve as early signs of PWS. Developmental delays are seen between the ages of 6 and 12, and excessive eating and cognitive deficits associated with PWS usually onset a little later.

While the exact mechanisms of PWS are not fully understood, there is evidence that affected individuals have hypothalamic abnormalities. This is not surprising, given the hypothalamus's role in regulating hunger and eating. However, as you will learn in the next section of this chapter, the hypothalamus is also involved in the regulation of sexual behavior. Consequently, many individuals suffering from PWS fail to reach sexual maturity during adolescence.

There is no current treatment or cure for PWS. However, if weight can be controlled in these individuals, then their life expectancies are significantly increased (historically, sufferers of PWS often died in adolescence or early adulthood). Advances in the use of various psychoactive medications and growth hormones continue to enhance the quality of life for individuals with PWS (Cassidy & Driscoll, 2009; Prader-Willi Syndrome Association, 2012).

Eating Disorders

While nearly two out of three US adults struggle with issues related to being overweight, a smaller, but significant, portion of the population has eating disorders that typically result in being normal weight or underweight. Often, these individuals are fearful of gaining weight. Individuals who suffer from bulimia nervosa and anorexia nervosa face many adverse health consequences.

People suffering from **bulimia nervosa** engage in binge eating behavior (eating a very large amount of food in a short time) that is followed by an attempt to compensate for the large amount of food consumed. Purging the food by inducing vomiting or through the use of laxatives are two common compensatory behaviors. Some affected individuals engage in excessive amounts of exercise to compensate for their binges. Bulimia is associated with many adverse health consequences that can include kidney failure, heart failure, and tooth decay. In addition, these individuals often suffer from anxiety and depression, and they are at an increased risk for substance abuse. The lifetime prevalence rate for bulimia nervosa is estimated at around 1% for women and less than 0.5% for men.

Binge eating disorder is similar to bulimia nervosa in that people suffering from it engage in binge eating behavior. Unlike with bulimia, eating binges are not followed by inappropriate behavior, such as purging, but they are followed by distress, including feelings of guilt and embarrassment. The resulting psychological distress distinguishes binge eating disorder from overeating.

Anorexia nervosa is an eating disorder characterized by the maintenance of a body weight well below average through starvation and/or excessive exercise. Individuals suffering from anorexia nervosa often have a **distorted body image**, referenced in literature as a type of body dysmorphia, meaning that they view themselves as overweight even though they are not. Like bulimia nervosa, anorexia nervosa is associated with a number of significant negative health outcomes: bone loss, heart failure, kidney failure, amenorrhea (cessation of the menstrual period), reduced function of the gonads, and in extreme cases, death. Furthermore, there is an increased risk for a number of psychological problems, which include anxiety disorders, mood disorders, and substance abuse. Estimates of the prevalence of anorexia nervosa vary from study to study but generally range from just under one percent to just over four percent in women. Generally, prevalence rates are considerably lower for men.

Link to Learning

Watch this news story about an Italian advertising campaign to raise public awareness of anorexia nervosa.

While both anorexia and bulimia nervosa occur in men and women of many different cultures, Caucasian females from Western societies tend to be the most at-risk population. Recent research indicates that females between the ages of 15 and 19 are most at risk, and it has long been suspected that these eating disorders are culturally-bound phenomena that are related to messages of a thin ideal often portrayed in popular media and the fashion world (***Figure 5***). While social factors play an important role in the development of eating disorders, there is also evidence that genetic factors may predispose people to these disorders.

Figure 5. *Young women in our society are inundated with images of extremely thin models (sometimes accurately depicted and sometimes digitally altered to make them look even thinner). These images may contribute to eating disorders. (credit: Peter Duhon)*

Section Summary

Hunger and satiety are highly regulated processes that result in a person maintaining a fairly stable weight that is resistant to change. When more calories are consumed than expended, a person will store excess energy as fat. Being significantly overweight adds substantially to a person's health risks and problems, including cardiovascular disease, type 2 diabetes, certain cancers, and other medical issues. Sociocultural factors that emphasize thinness as a beauty ideal and a genetic predisposition contribute to the development of eating disorders in many young females, though eating disorders span ages and genders.

Glossary

anorexia nervosa
: eating disorder characterized by an individual maintaining body weight that is well below average through starvation and/or excessive exercise

bariatric surgery
: type of surgery that modifies the gastrointestinal system to reduce the amount of food that can be eaten and/or limiting how much of the digested food can be absorbed

binge eating disorder
: type of eating disorder characterized by binge eating and associated distress

bulimia nervosa
: type of eating disorder characterized by binge eating followed by purging

distorted body image
: individuals view themselves as overweight even though they are not

leptin
: satiety hormone

metabolic rate
: amount of energy that is expended in a given period of time

morbid obesity
: adult with a BMI over 40

obese
: adult with a BMI of 30 or higher

overweight
: adult with a BMI between 25 and 29.9

satiation
: fullness; satisfaction

set point theory
: assertion that each individual has an ideal body weight, or set point, that is resistant to change

IV

Theme 4: How Does A Unique Individual Develop From A Single Cell?

Genghis Khan, founder and leader of the Mongol empire, is thought to have fathered at least 1000, and possibly as many as 2000 children with his numerous wives. There have been several dozen documented cases of women giving birth to more than 20 children, most of which occurred before the widespread availability of contraceptives, otherwise known as birth control methods. Using contraceptives allows us to plan our fertility in ways that were not possible for most of human history. Today, most people in the United States have a wide variety of contraceptive methods to choose from, and most females in this country use one or more of these methods during their lifetimes. In other areas of the world, contraception is not as widespread.

In this part of the course, we will be investigating human reproduction. First, we will cover the structures, hormones and cycles that are important for reproduction. We will address the processes through which the gametes, sperm and egg cells, develop and fuse to form new individuals. We will also consider pregnancy, childbirth, and the development of humans from a single cell through birth. Finally, we will consider some of the contraceptive choices available to prevent pregnancy, and some ethical issues associated with reproduction, sex and gender.

Understanding the growth and development of a new individual from a single cell is crucial to understanding cancer, a disease in which growth and division of cells in an individual becomes uncontrolled. Unfortunately, most of us have had a personal experience with cancer. In this module, we will explore the molecular and environmental causes of cancer, and compare the behavior of cancerous cells to that of normal cells. We will also examine genetic testing for cancer and consider the limits of such testing. We will also discuss cancer treatments, as well as how to lower the risk of developing cancer.

34.

4.1 Human Reproductive Anatomy

Human Reproductive Anatomy

In general, the reproductive structures in humans can be divided into three main categories: **gonads, internal genitalia and external genitalia.** The gonads are the organs in which **gametes**, the cells that fuse in fertilization to form new individuals, develop and mature. All other reproductive structures are called genitalia, or genitals. Internal genitalia are found inside of the body, while external genitalia are visible from the outside. The structures seen in adult males and females actually come from the same precursors in embryos, so there are many similarities in both structure and function between males and females. There is also a wide spectrum of structures present in any one individual; many people have structures that resemble a combination of male and female structures, or that resemble neither. In this textbook, we will define "male" and "female" based on individuals who have the most typical structures characteristic of those two sexes; other types of structures are also normal and common. We will describe the functions of these structures during vaginal sexual intercourse, since that is the sexual act used in reproduction; keep in mind that other types of sexual activity are also common and normal.

Male Reproductive Anatomy

In the male reproductive system, the **scrotum** houses the testicles or testes (singular: testis), including providing passage for blood vessels, nerves, and muscles related to testicular function. The **testes** are gonads, and they produce sperm (the male gametes) and some reproductive hormones. Each testis is approximately 2.5 by 3.8 cm (1.5 by 1 in) in size and divided into wedge-shaped lobules by connective tissue called septa.

Sperm are immobile at body temperature; therefore, the scrotum and penis are external to the body, as illustrated in ***Figure 1*** so that a proper temperature is maintained for motility.

ART CONNECTION

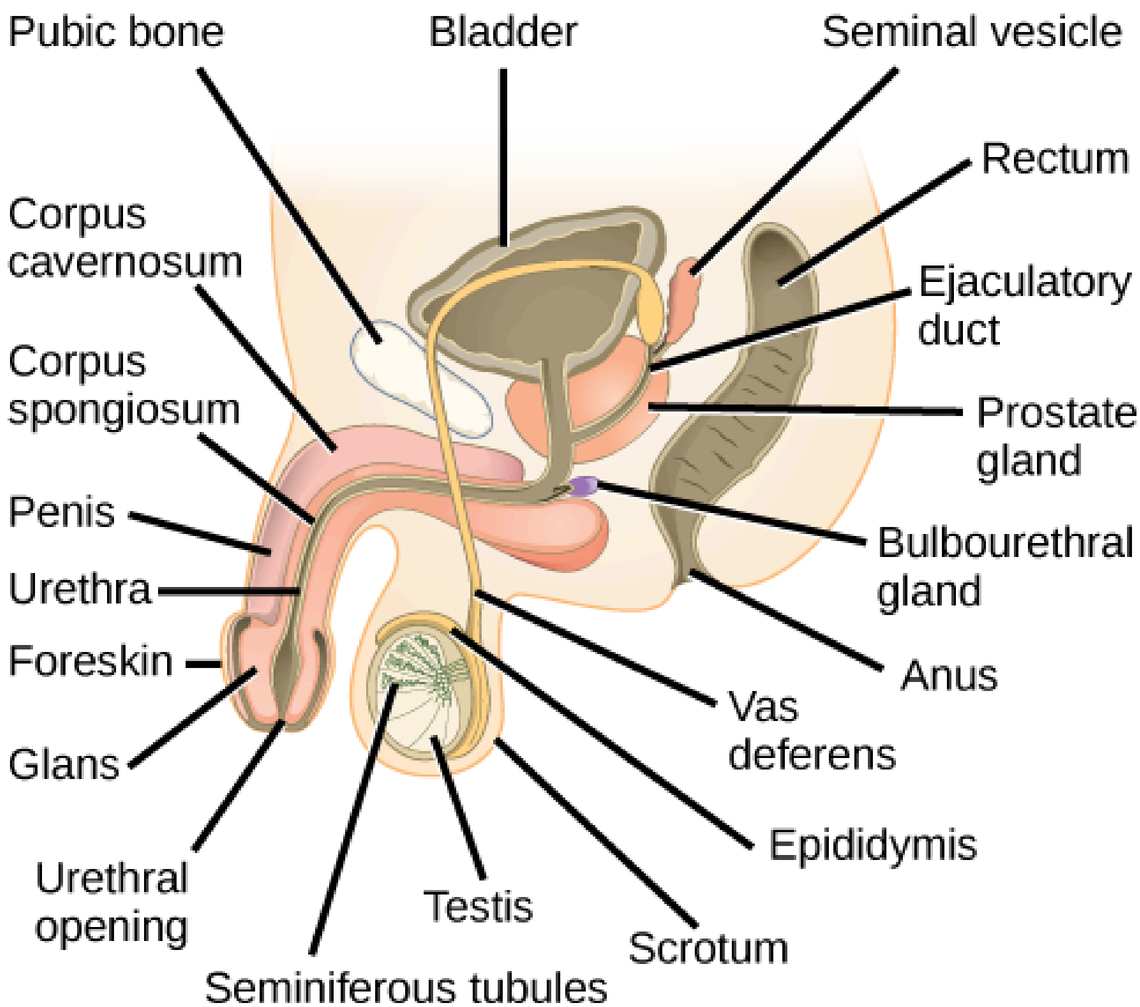


Figure 1. *The reproductive structures of the human male are shown.*

The internal genitalia in males are important for the production of sperm, and of other components of . the semen. Sperm mature in **seminiferous tubules** that are coiled inside the testes, as illustrated in ***Figure 1***. The walls of the seminiferous tubules are made up of the developing sperm cells, with the least developed sperm at the periphery of the tubule and the fully developed sperm in the lumen. The sperm cells are mixed with "nursemaid" cells called **Sertoli cells** which protect the germ cells and promote their development. Other cells mixed in the wall of the tubules are the **interstitial cells of Leydig**. These cells produce high levels of testosterone once the male reaches adolescence.

When the sperm have developed flagella and are nearly mature, they leave the testicles and enter the epididymis, shown in ***Figure 1***. This structure resembles a comma and lies along the top and posterior portion of the testes; it is the site of sperm maturation. The sperm leave the epididymis and enter the vas deferens (or ductus deferens), which carries the sperm, behind the bladder, and forms the ejaculatory duct with the duct from the seminal vesicles. During a vasectomy, a section of the vas deferens is removed, preventing sperm from being passed out of the body during ejaculation and preventing fertilization.

Semen is a mixture of sperm and spermatic duct secretions (about 10 percent of the total) and fluids from accessory glands that contribute most of the semen's volume. Sperm are haploid cells, consisting of a flagellum as a tail, a neck that contains the cell's energy-producing mitochondria, and a head that contains the genetic material. ***Figure 2*** shows a micrograph of human sperm as well as a diagram of the parts of the sperm. An acrosome is found at the top of the head of the sperm. This structure contains enzymes that can digest the protective coverings that surround the egg to help the sperm penetrate and fertilize the egg. An ejaculate (a single emission of sperm) will contain from two to five milliliters of fluid with from 50–120 million sperm per milliliter.

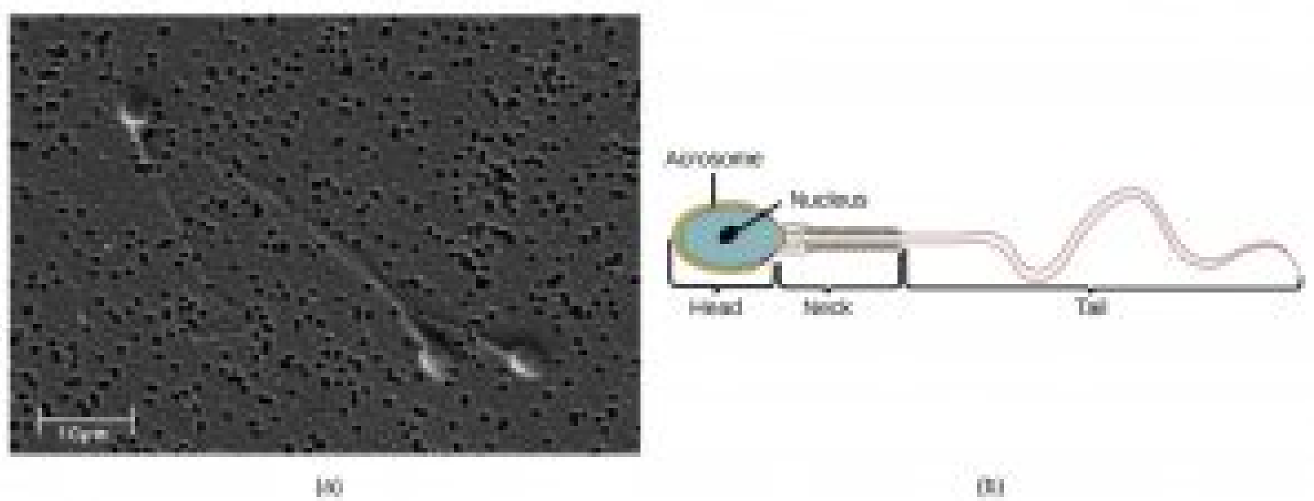


Figure 2. *Human sperm, visualized using scanning electron microscopy, have a flagellum, neck, and head. (credit b: modification of work by Mariana Ruiz Villareal; scale-bar data from Matt Russell)*

The bulk of the semen comes from the accessory glands associated with the male reproductive system. These are the seminal vesicles, the prostate gland, and the bulbourethral gland, all of which are illustrated in ***Figure 1***. The **seminal vesicles** are a pair of glands that lie along the posterior border of the urinary bladder. The glands make a solution that is thick, yellowish, and alkaline. As sperm are only motile in an alkaline environment, a basic pH is important to reverse the acidity of the vaginal environment. The solution also contains mucus, fructose (a sperm mitochondrial nutrient), a coagulating enzyme, ascorbic acid, and local-acting hormones called prostaglandins. The seminal vesicle glands account for 60 percent of the bulk of semen.

The **penis**, illustrated in ***Figure 1***, is an organ that drains urine from the renal bladder and functions as a copulatory organ during intercourse. The penis contains three tubes of erectile tissue running through the length of the organ. These consist of a pair of tubes on the dorsal side, called the corpus cavernosum, and a single tube of tissue on the ventral side, called the corpus spongiosum. This tissue will become engorged with blood, becoming erect and hard, in preparation for sexual intercourse. The organ is inserted into the vagina culminating with an ejaculation. During intercourse, the smooth muscle sphincters at the opening to the renal bladder close and prevent urine from entering the penis. An orgasm is a two-stage process: first, glands and accessory organs connected to the testes contract, then semen (containing sperm) is expelled through the urethra during ejaculation. After intercourse, the blood drains from the erectile tissue and the penis becomes flaccid.

The walnut-shaped **prostate gland** surrounds the urethra, the connection to the urinary bladder. It has a series of short ducts that directly connect to the urethra. The gland is a mixture of smooth muscle and glandular tissue. The muscle provides much of the force needed for ejaculation to occur. The glandular tissue makes a thin, milky fluid that contains citrate (a nutrient), enzymes, and prostate specific antigen (PSA). PSA is a proteolytic enzyme that helps to liquefy the ejaculate several minutes after release from the male. Prostate gland secretions account for about 30 percent of the bulk of semen.

The **bulbourethral gland**, or Cowper's gland, releases its secretion prior to the release of the bulk of the semen. It neutralizes any acid residue in the urethra left over from urine. This usually accounts for a couple of drops of fluid in the total ejaculate and may contain a few sperm. Withdrawal of the penis from the vagina before ejaculation to prevent pregnancy may not work if sperm are present in the bulbourethral gland secretions. The location and functions of the male reproductive organs are summarized in ***Table 1***.

Table 1. Male Reproductive Anatomy

Organ	**Location**	**Function**
Scrotum	External	Carry and support testes
Penis	External	Deliver urine, copulating organ
Seminiferous Tubules	Internal	Site of sperm maturation in testes
Epididymus	Internal	Part of pathway for sperm exit from body
Vas Deferens	Internal	Part of pathway for sperm exit from body
Ejaculatory Duct	Internal	Site of mixing of sperm with semen components, part of pathway for sperm exit from body
Testes	Internal	Produce sperm and male hormones
Seminal Vesicles	Internal	Contribute to semen production
Prostate Gland	Internal	Contribute to semen production
Bulbourethral Glands	Internal	Clean urethra at ejaculation

Female Reproductive Anatomy

A number of reproductive structures are exterior to the female's body. These include the breasts and the vulva, which consists of the mons pubis, clitoris, labia majora, labia minora, and the vestibular glands, all illustrated in ***Figure 3***. The location and functions of the female reproductive organs are summarized in ***Table 2***. The mons pubis is a round, fatty area that overlies the pubic bone. The **clitoris** is a structure with erectile tissue that contains a large number of sensory nerves and serves as a source of stimulation during intercourse. The **labia majora** are a pair of elongated folds of tissue that run posterior from the mons pubis and enclose the other components of the vulva. The labia majora derive from the same tissue that produces the scrotum in a male. The **labia minora** are thin folds of tissue centrally located within the labia majora. These labia protect the openings to the vagina and urethra. The mons pubis and the anterior portion of the labia majora become covered with hair during adolescence; the labia minora is hairless. The greater vestibular glands are found at the sides of the vaginal opening and provide lubrication during intercourse.The vulva is the name for the entire set of external genitalia in the inguinal (groin) area of females; in common language this is sometimes referred to as the vagina, but that is not anatomically accurate; the vagina is an entirely internal structure.

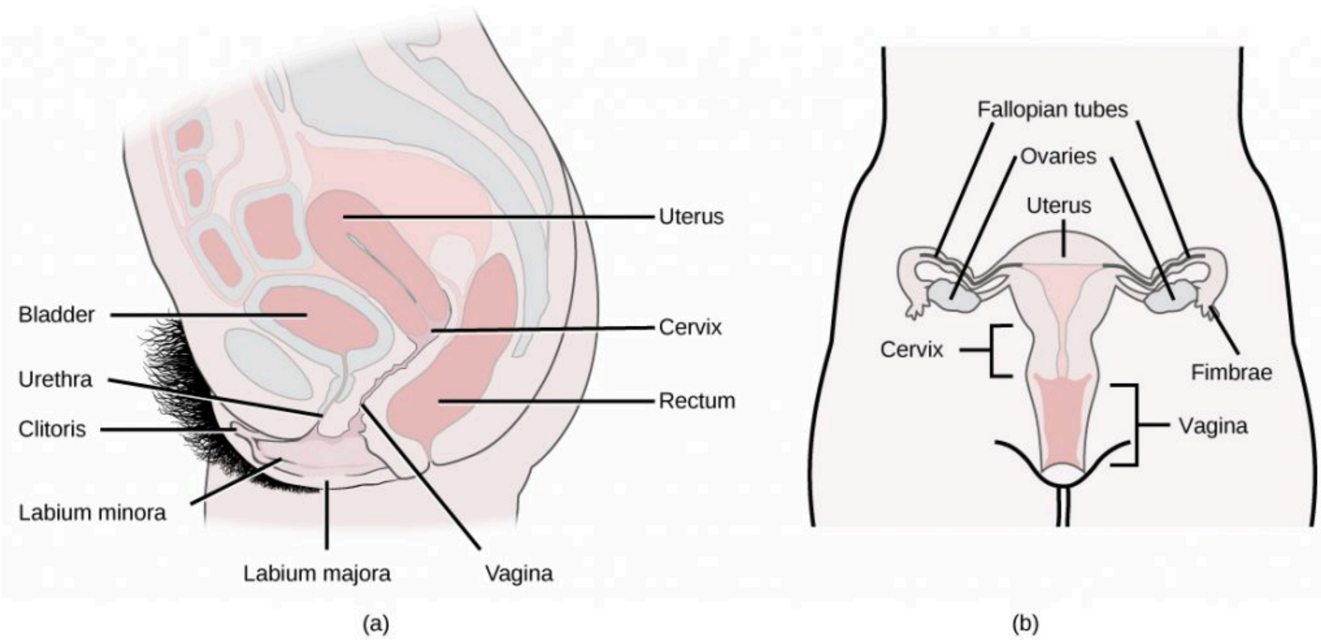


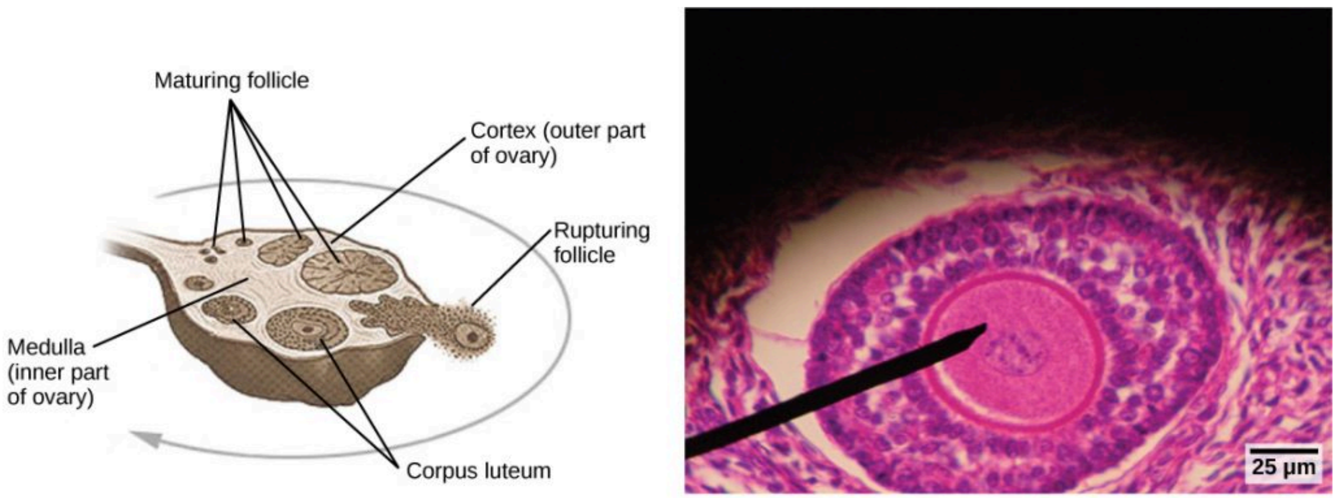
Figure 3. *The reproductive structures of the human female are shown. (credit* ***a****: modification of work by Gray's Anatomy; credit* ***b****: modification of work by CDC)*

Table 2. Female Reproductive Anatomy

Organ	**Location**	**Function**
Clitoris	External	Sensory organ
Mons pubis	External	Fatty area overlying pubic bone
Labia majora	External	Covers labia minora
Labia minora	External	Covers vestibule
Greater vestibular glands	External	Secrete mucus; lubricate vagina
Breast	External	Produce and deliver milk
Ovaries	Internal	Carry and develop eggs
Oviducts (Fallopian tubes)	Internal	Transport egg to uterus
Uterus	Internal	Support developing embryo
Vagina	Internal	Common tube for intercourse, birth canal, passing menstrual flow

The breasts consist of mammary glands and fat. The size of the breast is determined by the amount of fat deposited behind the gland. Each gland consists of 15 to 25 lobes that have ducts that empty at the nipple and that supply the nursing child with nutrient- and antibody-rich milk to aid development and protect the child.

Internal female reproductive structures include ovaries, oviducts, the **uterus**, and the vagina, shown in ***Figure 3***. The two ovaries (the female gonads) are held in place in the abdominal cavity by a system of ligaments. Ovaries consist of a medulla and cortex: the medulla contains nerves and blood vessels to supply the cortex with nutrients and remove waste. The outer layers of cells of the cortex are the functional parts of the ovaries. The cortex is made up of follicular cells that surround eggs that develop during fetal development *in utero*. During the menstrual period, a batch of follicular cells develops and prepares the eggs for release. At ovulation, one follicle ruptures and one egg is released, as illustrated in ***Figure 4a***.



(a) (b)

Figure 4. *Oocytes develop in* ***(a)*** *follicles, located in the ovary. At the beginning of the menstrual cycle, the follicle matures. At ovulation, the follicle ruptures, releasing the egg. The follicle becomes a corpus luteum, which eventually degenerates. The* ***(b)*** *follicle in this light micrograph has an oocyte at its center. (credit a: modification of work by NIH; scale-bar data from Matt Russell)*

The **oviducts**, or fallopian tubes, extend from the uterus in the lower abdominal cavity to the ovaries, but they are not in contact with the ovaries. The lateral ends of the oviducts flare out into a trumpet-like structure and have a fringe of finger-like projections called fimbriae, illustrated in ***Figure 4b***. When an egg is released at ovulation, the fimbrae help the non-motile egg enter into the tube and passage to the uterus. The walls of the oviducts are ciliated and are made up mostly of smooth muscle. The cilia beat toward the middle, and the smooth muscle contracts in the same direction, moving the egg toward the uterus. Fertilization usually takes place within the oviducts and the developing embryo is moved toward the uterus for development. It usually takes the egg or embryo a week to travel through the oviduct. Sterilization in females is called a tubal ligation; it is analogous to a vasectomy in males in that the oviducts are severed and sealed.

The uterus is a structure about the size of a females's fist. This is lined with an endometrium rich in blood vessels and mucus glands. The uterus supports the developing embryo and fetus during gestation.

The thickest portion of the wall of the uterus is made of smooth muscle. Contractions of the smooth muscle in the uterus aid in passing the baby through the vagina during labor. A portion of the lining of the uterus sloughs off during each menstrual period, and then builds up again in preparation for an implantation. Part of the uterus, called the cervix, protrudes into the top of the vagina. A small opening called the cervical orifice allows menstrual fluid out of the cervix into the vagina, and sperm into the uterus. During childbirth the cervical orifice is greatly enlarged.

The **vagina** is a muscular tube that serves several purposes. It allows menstrual flow to leave the body. It is the receptacle for the penis during intercourse and the vessel for the delivery of offspring. It is lined cells that produce acidic secretions that limit the growth of microbes that could potentially travel into the uterus.

Development of Reproductive Organs in Humans.

The reproductive tissues of male and female humans develop similarly *in utero* (in a fetus developing in the mother's uterus) for the first several weeks of gestation. The hormone testosterone is typically only released in embryos that have a male sex chromosome (the Y chromosome, discussed in the next chapter), and this hormone controls the generation of reproductive structures. A low level of the hormone testosterone is released from male gonads in the developing embryos, starting at around the second month of gestation. Testosterone causes the undeveloped tissues to differentiate into male sexual organs. When testosterone is absent, the tissues develop into female sexual tissues. Primitive gonads become testes or ovaries. Tissues that produce a penis in males produce a clitoris in females. The tissue that will become the scrotum in a male becomes the labia in a female; that is, they are homologous structures. Because of this, there are often variations in development resulting in structures that may have characteristics of both sexes, or of neither sex, depending on hormonal levels and other factors present during embryogenesis. These variations in sexual structures are quite common and normal.

Sexual Response During Intercourse

The sexual response in humans is both psychological and physiological. Both sexes experience sexual arousal through psychological and physical stimulation. There are four phases of the sexual response. During phase one, called excitement, vasodilation leads to vasocongestion in erectile tissues in both males and females. The nipples, clitoris, labia, and penis engorge with blood and become enlarged. Vaginal secretions are released to lubricate the vagina to facilitate intercourse. During the second phase, called the plateau, stimulation continues, the outer third of the vaginal wall enlarges with blood, and breathing and heart rate increase.

During phase three, or orgasm, rhythmic, involuntary contractions of muscles occur in both sexes. In the male, the reproductive accessory glands and tubules constrict placing semen in the urethra, then the urethra contracts expelling the semen through the penis. In females, the uterus and vaginal muscles contract in waves that may last slightly less than a second each. During phase four, or resolution, the processes described in the first three phases reverse themselves and return to their normal state. Males experience a refractory period in which they cannot maintain an erection or ejaculate for a period of time ranging from minutes to hours.

Section Summary

The reproductive structures that evolved in humans allow males and females to mate, fertilize internally, and support the growth and development of offspring. Reproductive structures include gonads, internal and external genitalia. Some male and female reproductive structures have analogous functions and are derived from common precursor structures. Both males and females have four stages of the sexual response.

Glossary

bulbourethral gland
: secretion that cleanses the urethra prior to ejaculation

clitoris
: sensory structure in females; stimulated during sexual arousal

external genitalia
: reproductive structures visible on the outside of the body

gametes
: haploid cells that combine in fertilization to produce a new diploid individual; eggs and sperm

gametogenesis
: the process of egg or sperm development

gonads
: structures that produce gametes (eggs and sperm);female gonads are ovaries and male gonads are testes

internal genitalia
: all internal reproductive structures except for gonads

interstitial cells of Leydig
: cells in the testes that produce testosterone, also called Leydig cells

labia majora
: large folds of tissue covering the inguinal area

labia minora
: smaller folds of tissue within the labia majora

oviduct
: (also, fallopian tube) muscular tube connecting the uterus with the ovary area

penis
: male reproductive structure for urine elimination and copulation

prostate gland
: structure that is a mixture of smooth muscle and glandular material and that contributes to semen

scrotum
: sac containing testes; exterior to the body

semen
: fluid mixture of sperm and supporting materials

seminal vesicle
: secretory accessory gland in males; contributes to semen

seminiferous tubule
: site of sperm production in testes

Sertoli cells
: cells inside of the testes that protect sperm and help them to mature

testes
: pair of reproductive organs in males

uterus
: environment for developing embryo and fetus

vagina
: muscular tube for the passage of menstrual flow, copulation, and birth of offspring

35.

4.2 Meiosis and Gametogenesis

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. 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. Meiosis occurs during the process of **gametogenesis**, which is the production of gametes (oocytes and sperm.)

Figure 1 shows a **karyotype** of a human male. A karyotype is an image produced by arranging images of each chromosome in a cell into systematic pairs. As you can see, there are 23 total pairs of chromosomes, and 22 of them contain matched pairs of the same size with similar banding patterns, indicating that the cells in each pair contain the same genes. Each of these 22 pairs are called homologous pairs, and the pairs are numbered in order of size, with the two copies of chromosome 1 being the largest. The final pair of chromosomes are not matched in size or banding patterns; these are the sex chromosomes and are called the X chromosome and the Y chromosome. In most human cells, there are 22 matched pairs of chromosomes and 1 pair of sex chromosomes. Almost all males have one X sex chromosome and 1 Y sex chromosome, and almost all females have 2 X sex chromosomes, in each cell. Gametes are unique in that they have half of the chromosomes of other cells: only one sex chromosome and only 22 other chromosomes, not 22 pairs.

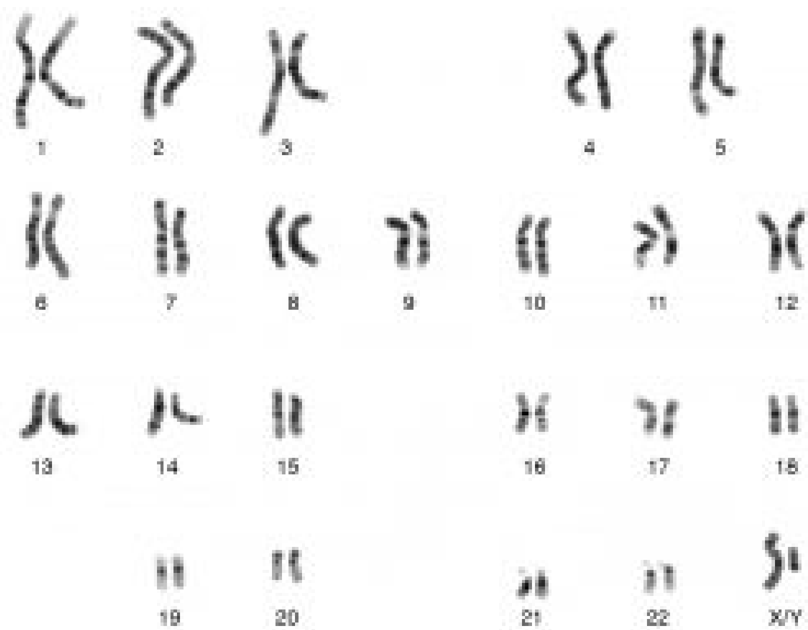


Figure 1: *A karyotype of a human male. Each pair of chromosomes contains hundreds to thousands of genes. The banding patterns are nearly identical for the two chromosomes within each pair, indicating the same organization of genes. As is visible in this karyotype, the only exception to this is the XY sex chromosome pair in males. (credit: National Human Genome Research Institute)*

Most animals and plants are diploid, containing two sets of chromosomes; in each **somatic cell** (the nonreproductive 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. 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.

The nuclear division that forms haploid cells, which is called meiosis, is related to mitosis. 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. 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. 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. Meiosis I reduces the number of chromosome sets from two to one. 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, includes all of the stages of division again. These stages have specific names (Prophase, Metaphase, Anaphase, Telophase), but you are not required to know the specific names of the stages for this course.

Interphase

Meiosis is preceded by a stage called interphase. In this stage, the DNA of the chromosomes is replicated, so that each cell contains two copies of each chromatid. This means that in a human cell, both copies of chromosome 1 are copied to produce 4 chromatids, both copies of chromosome 2 are copied to produce 4 chromatids, and so on. The cell also grows, and produces enough of the enzymes and structures required for meiosis during this interphase period.

Meiosis I

Early in meiosis 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 homologous chromosomes now become arranged in the center of the cell, with the ends of each pair of homologous chromosomes 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. Consider that the homologous chromosomes of a human are originally inherited as two separate sets, one from each parent. 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. These pairs line up at the midway point between the two poles of the cell, and their arrangement in regard to the two poles is random. Any maternally inherited chromosome may face either pole. Any paternally inherited chromosome may also face either pole. The orientation of each pair is independent of the orientation of the other 22 pairs.

In each cell that undergoes meiosis, the arrangement of the chromosomes is different. There are two possibilities for orientation (for each pair); 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. Other mechanisms not discussed in this class can increase the variation in each cell produced as well. Given both independent assortment and these other mechanisms, it is highly unlikely that any two haploid cells resulting from meiosis will have the same genetic composition.

Next, protein fibers in the cell pull the linked chromosomes apart. The fibers pull the chromosome to the opposite poles of the cell. At each pole, there is just one member of each pair of the homologous chromosomes, so 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 there are duplicate copies of the set because each homolog still consists of two “sister: chromatids that are still attached to each other. Cytokinesis, where the cell membrane pinches off in the center of the cell to separate it into two, now occurs, splitting each cell into two cells containing one full set of chromosomes.

Concept in Action

Review the process of meiosis, observing how chromosomes align and migrate, at this site.

Meiosis II

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.

First, the nuclear envelope breaks down and the chromosomes are clearly visible under a microscope. The sister chromatids then line up at the center of the cell. Protein fibers pull one of each pair of sisters toward the poles of the cell.The chromosomes arrive at opposite poles.. Nuclear envelopes form around the chromosomes. Cytokinesis separates the two cells into four genetically unique haploid cells. At this point, the nuclei in the newly produced cells 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 other sources of variation not discussed here.

Comparing Meiosis and Mitosis

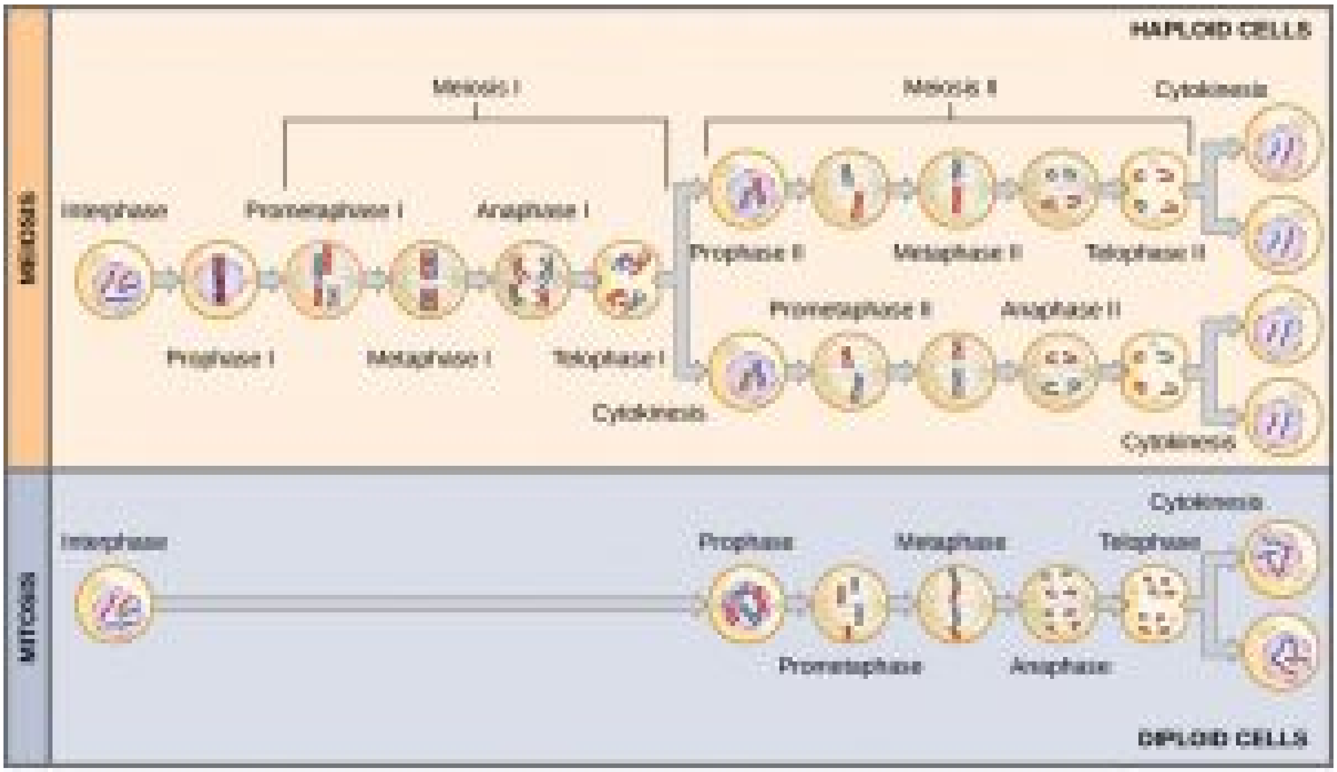
Mitosis and meiosis, which are both forms of division of the nucleus in eukaryotic cells, share some similarities, but also exhibit distinct differences that lead to their very different outcomes. Mitosis is a single nuclear division that results in two nuclei, usually partitioned into two new cells. 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 (***Figure 2***).

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. As in mitotis, one sister chromatid is pulled to one pole and the other sister chromatid is pulled to the other pole during meiosis II. 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, but cells produced by meiosis 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 2. *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.*

Concept in Action

For an animation comparing mitosis and meiosis, go to this website.

Gametogenesis (Spermatogenesis and Oogenesis)

Gametogenesis, the production of sperm and eggs, includes the process of meiosis to produce haploid cells, and growth and maturation of these cells into oocytes and sperm. The production of sperm is called **spermatogenesis** and the production of eggs is called **oogenesis**.

Spermatogenesis

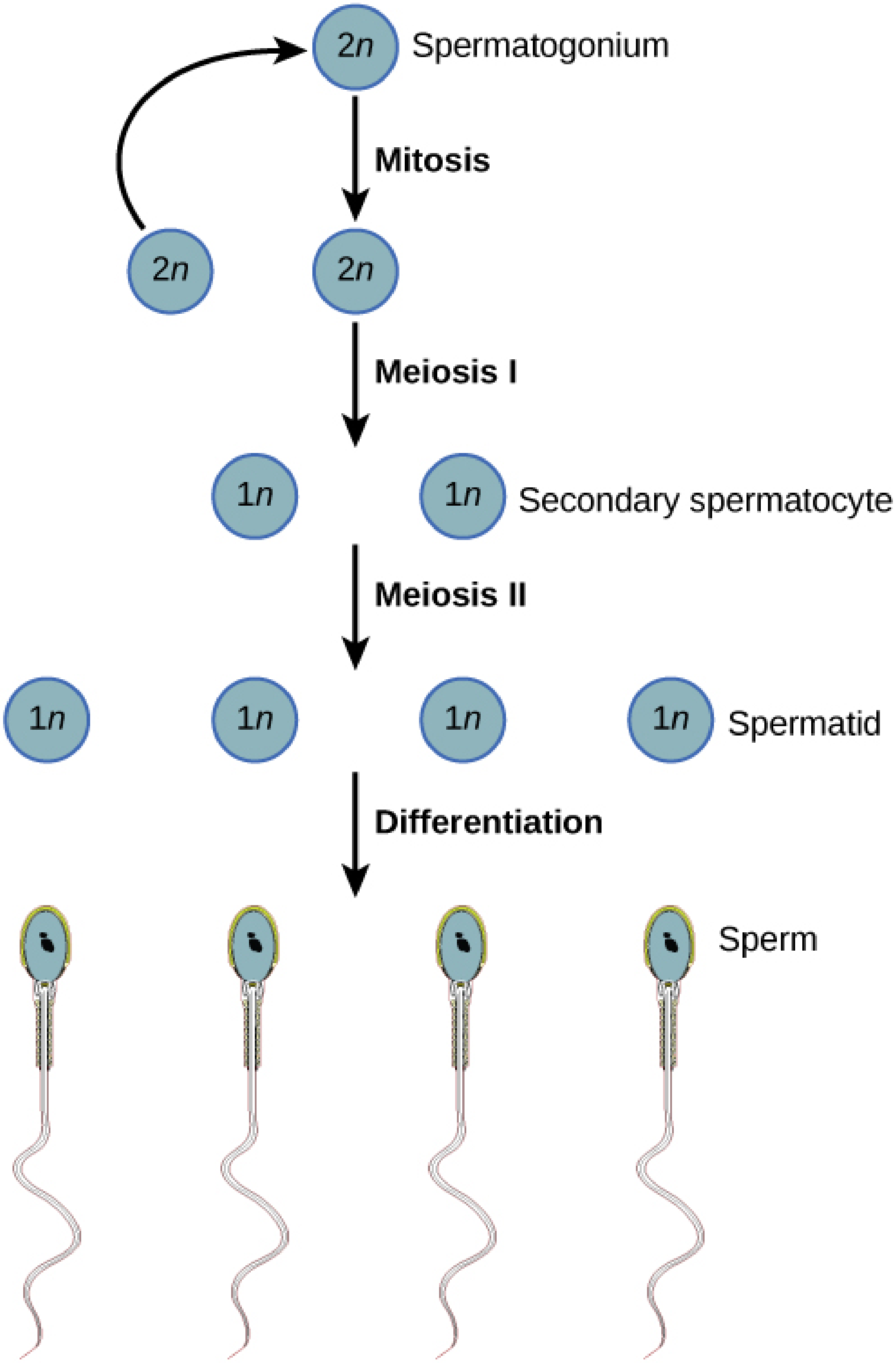


Figure 3. *During spermatogenesis, four sperm result from each primary spermatocyte.*

Spermatogenesis, illustrated in ***Figure 3***, occurs in the wall of the seminiferous tubules, with stem cells at the periphery of the tube and the spermatozoa at the lumen of the tube. Immediately under the capsule of the tubule are diploid, undifferentiated cells. These stem cells, called spermatogonia (singular:

spermatagonium), go through mitosis with one offspring going on to differentiate into a sperm cell and the other giving rise to the next generation of sperm.

Meiosis starts with a cell called a primary spermatocyte. At the end of the first meiotic division, a haploid cell is produced called a secondary spermatocyte. This cell is haploid and must go through another meiotic cell division. The cell produced at the end of meiosis is called a spermatid and when it reaches the lumen of the tubule and grows a flagellum, it is called a sperm cell. Four sperm result from each primary spermatocyte that goes through meiosis.

Stem cells are deposited during gestation and are present at birth through the beginning of adolescence, but in an inactive state. During adolescence, gonadotropic hormones from the anterior pituitary cause the activation of these cells and the production of viable sperm. This continues into old age.

Link to Learning

Visit this site to see the process of spermatogenesis.

Oogenesis

Oogenesis, illustrated in ***Figure 4***, occurs in the outermost layers of the ovaries. As with sperm production, oogenesis starts with a germ cell, called an oogonium (plural: oogonia), but this cell undergoes mitosis to increase in number, eventually resulting in up to about one to two million cells in the embryo.

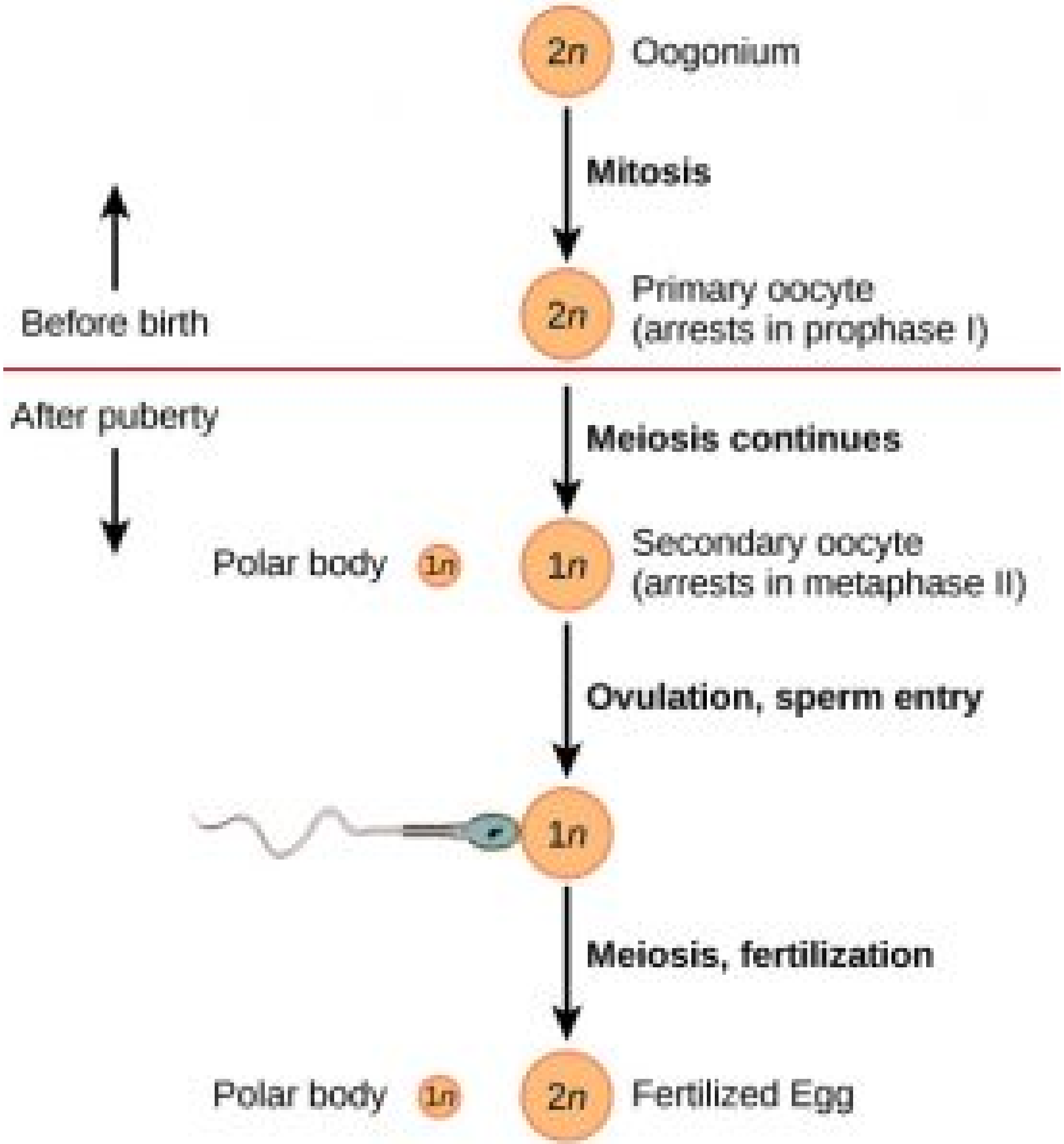


Figure 4 *The process of oogenesis occurs in the ovary's outermost layer.*

The cell starting meiosis is called a primary oocyte, as shown in ***Figure 4***. This cell will start the meiosis I and be arrested in its progress very early on, in a stage called the first prophase stage. At the time of birth, all future oocytes are in the prophase stage; no additional oocytes or precursors are produced after birth. At adolescence, anterior pituitary hormones cause the development of a number of follicles in an ovary. This results in the primary oocyte finishing meiosis I. The cell divides unequally, with most of the cellular material and organelles going to one cell, called a secondary oocyte, and only one set of chromosomes and a small amount of cytoplasm going to the other cell. This second cell is called a polar body and usually dies. A secondary meiotic arrest occurs, about halfway through the meiosis II in a stage called the metaphase II stage. At ovulation, this secondary oocyte will be released and travel toward the uterus through the oviduct. If the secondary oocyte is fertilized, the cell continues through meiosis II, producing a second polar body and a fertilized egg containing all 46 chromosomes of a human being, half of them coming from the sperm. If the oocyte is not fertilized, however, it does not complete meiosis II.

Oocyte production begins before birth, is arrested during meiosis until puberty, and then individual cells continue through at each menstrual cycle. One oocyte is produced from each meiotic process, with the extra chromosomes and chromatids going into polar bodies that degenerate and are reabsorbed by the body.

Section Summary

Sexual reproduction in humans requires that diploid individual cells produce haploid cells that can fuse during fertilization to form diploid offspring. The process that results in haploid cells is called meiosis. Meiosis is a series of events that arrange and separate chromosomes into daughter cells. During the interphase of meiosis, each chromosome is duplicated. In meiosis, there are two rounds of nuclear division resulting in four nuclei and usually four haploid daughter cells, each with half the number of chromosomes as the parent cell. During meiosis, variation in the daughter nuclei is introduced because of random alignment in meiosis I. The cells that are produced by meiosis are genetically unique.

Meiosis and mitosis share similarities, but have distinct outcomes. Mitotic divisions are single nuclear divisions that produce daughter nuclei that are genetically identical and have the same number of chromosome sets as the original cell. Meiotic divisions are two nuclear divisions that produce four daughter nuclei that are genetically different and have one chromosome set rather than the two sets the parent cell had. The main differences between the processes occur in the first division of meiosis. The homologous chromosomes separate into different nuclei during meiosis I causing a reduction of ploidy level. The second division of meiosis is much more similar to a mitotic division.

Important similarities exist between spermatogenesis and oogenesis: during both processes a diploid cell is duplicated a number of times in mitosis to produce precursors that undergo two rounds of meiosis to produce haploid sperm and eggs. Important differences exist in the timing of these divisions and in the symmetry of divisions. In males, each spermatagonium produces four mature sperm, but in oogenesis each oogonium can produce only one fertilized egg.

Glossary

fertilization
: the union of two haploid cells typically from two individual organisms

karyotype
: systematic arrangement of images of chromosomes into homologous pairs

meiosis I
: the first round of meiotic cell division; referred to as reduction division because the resulting cells are haploid

meiosis II
: the second round of meiotic cell division following meiosis I; sister chromatids are separated from each other, and the result is four unique haploid cells

oogenesis
: the process of producing haploid eggs

reduction division
: a nuclear division that produces daughter nuclei each having one-half as many chromosome sets as the parental nucleus; meiosis I is a reduction division

somatic cell
: all the cells of a multicellular organism except the gamete-forming cells

spermatogenesis
: the process of producing haploid sperm

36.

4.3 Hormones and the Endocrine System

The endocrine system produces hormones that function to control and regulate many different body processes. The endocrine system coordinates with the nervous system to control the functions of the other organ systems. Cells of the endocrine system produce molecular signals called hormones. These cells may compose endocrine glands, may be tissues or may be located in organs or tissues that have functions in addition to hormone production. Hormones circulate throughout the body and stimulate a response in cells that have receptors able to bind with them. The changes brought about in the receiving cells affect the functioning of the organ system to which they belong. Many of the hormones are secreted in response to signals from the nervous system, thus the two systems act in concert to effect changes in the body.

Hormones

Maintaining balance within the body requires the coordination of many different systems and organs. One mechanism of communication between neighboring cells, and between cells and tissues in distant parts of the body, occurs through the release of chemicals called **hormones**. Hormones are released into body fluids, usually blood, which carries them to their **target cells** where they elicit a response. The cells that secrete hormones are often located in specific organs, called **endocrine glands**, and the cells, tissues, and organs that secrete hormones make up the endocrine system. Examples of endocrine organs include the pancreas, which produces the hormones insulin and glucagon to regulate blood-glucose levels, the adrenal glands, which produce hormones such as epinephrine and norepinephrine that regulate responses to stress, and the thyroid gland, which produces thyroid hormones that regulate metabolic rates. In addition, many organs in other body systems such as the reproductive system also secrete hormones and are therefore also part of endocrine system.

The level of hormones present in the blood is directly related to the magnitude of the responses, or actions, performed by the target cells. For this reason, the body has evolved mechanisms to tightly regulate the level of hormone production for many hormones. The release of hormones by endocrine cells is usually related to the amount of stimulus the receive. A stimulus is any variable that can change the level of hormone secreted by an endocrine cell; common stimuli include nutrients, neurotransmitters, and even other hormones. Most stimuli have a positive effect on endocrine cells, that is, with increasing amounts of stimuli increasing amounts of hormone are secreted. Some stimuli are negative regulators of endocrine cells, however. When increased negative stimuli are present, hormone secretion decreases. Many endocrine cells integrate inputs from multiple stimuli to regulate their level of hormone secretion.

How Hormones Work

Hormones cause changes in target cells by binding to specific cell-surface or intracellular hormone **receptors**, molecules embedded in the cell membrane or floating in the cytoplasm with a binding site that matches a binding site on the hormone molecule. In this way, even though hormones circulate throughout the body and come into contact with many different cell types, they only affect cells that possess the necessary receptors. Receptors for a specific hormone may be found on or in many different cells or may be limited to a small number of specialized cells. For example, thyroid hormones act on many different tissue types, stimulating metabolic activity throughout the body. Different types of cells can also respond to hormones in different ways. For example, immune cell function is decreased when the hormone cortisol acts upon them, but cortisol causes liver cells to do something completely different, release glucose into the blood. Cells can have many receptors for the same hormone but often also possess receptors for different types of hormones.

Glossary

endocrine gland
: an organ containing endocrine cells, cells that secrete hormones

hormone
: a chemical secreted by an endocrine gland into body fluids (usually blood) that affects the function of target cells

hormone receptor
: a molecule located on the surface of a cell, or inside it, that binds to a hormone; binding of a hormone to its receptor causes changes in the cell with the receptor

stimulus
: a signal that increases or decreases the level of hormone secreted by an endocrine cell

target
: a cell that expresses a hormone receptor, and changes its functions after the hormone binds to the receptor

37.

4.4 Hormonal Control of Human Reproduction

The human male and female reproductive cycles are controlled by the interaction of hormones from the hypothalamus and anterior pituitary with hormones from reproductive tissues and organs. In both sexes, the hypothalamus monitors and causes the release of hormones from the pituitary gland. When the reproductive hormone is required, the hypothalamus sends a **gonadotropin-releasing hormone (GnRH)** to the anterior pituitary. This causes the release of **follicle stimulating hormone (FSH)** and **luteinizing hormone (LH)** from the anterior pituitary into the blood. Note that the body must reach puberty in order for GnRH to be produced. Although FSH and LH are named after their functions in female reproduction, they are produced in both sexes and play important roles in controlling reproduction. Other hormones have specific functions in the male and female reproductive systems. *Remember that we are defining "male" and "female" as individuals with a narrow set of sex characteristics for the purposes of this course; variations in these characteristics, and in hormonal control of reproduction are very common.*

Male Hormones

At the onset of puberty, the hypothalamus causes the release of FSH and LH into the male system for the first time. FSH enters the testes and stimulates the **Sertoli cells** to begin facilitating spermatogenesis using negative feedback, as illustrated in ***Figure 1***. LH also enters the testes and stimulates the **interstitial cells of Leydig** to make and release testosterone into the testes and the blood.

Testosterone, the hormone responsible for the secondary sexual characteristics that develop in the male during adolescence, stimulates spermatogenesis. These secondary sex characteristics include a deepening of the voice, the growth of facial, axillary, and pubic hair, and the beginnings of the sex drive.

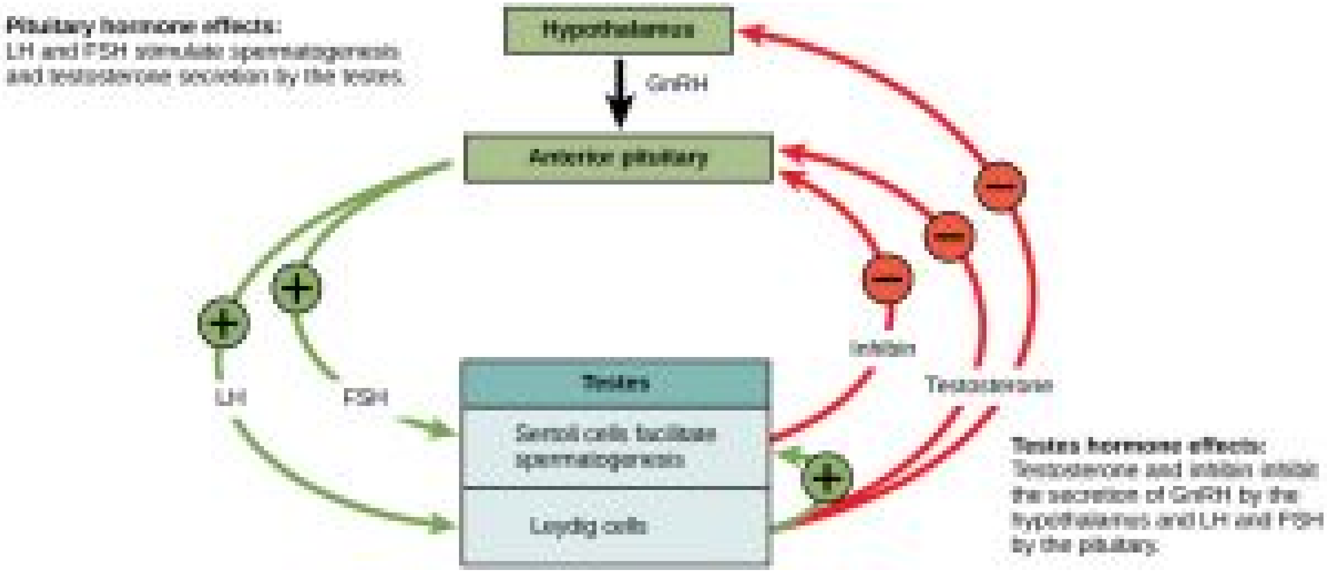


Figure 1. *Hormones control sperm production in a negative feedback system.*

A negative feedback system occurs in the male with rising levels of testosterone acting on the hypothalamus and anterior pituitary to inhibit the release of GnRH, FSH, and LH. This system works to keep the concentration of testosterone in the blood at a relatively constant level in each male individual after puberty. If testosterone levels become elevated above their normal levels (called the set point levels), the testosterone will inhibit the hypothalamus from secreting as much GnRH. This will in turn cause lower levels of FSH and LH to be produced. Lower LH levels will lower the amount of stimulation that the Leydig cells receive, and lower the production of testosterone. The Sertoli cells also produce the hormone **inhibin**, which is released into the blood when the sperm count is too high. This inhibits the release of GnRH and FSH, which will cause spermatogenesis to slow down. If the sperm count reaches 20 million/ml, the Sertoli cells cease the release of inhibin, and the sperm count increases.

Conversely, if testosterone levels are decreased, there will be less inhibition of GnRH production by the hypothalamus. This will cause more stimulation of the anterior pituitary, increased FSH and LH secretion, and increased stimulation of the Leydig cells and Sertoli cells. The end result will be an increase in testosterone levels. As you may imagine, this intricate system of regulation is constantly responding to slight fluctuations in testosterone with concomitant fluctuations in hypothalamic secretion of GnRH, and anterior pituitary secretion of FSH and LH.

Female Hormones

The control of reproduction in females is more complex. As with the male, the anterior pituitary cause the releases the hormones FSH and LH. GnRH controls the secretion of FSH and LH, but in a more complex cyclical manner. In addition, estrogens and progesterone are released from the developing follicles. In females, FSH stimulates development of egg cells, called ova, which develop in structures called follicles. Follicle cells produce the hormone inhibin, which inhibits FSH production. LH also plays a role in the development of ova, induction of ovulation, and stimulation of estrogen and progesterone production by the ovaries. **Estrogen** and **progesterone** are steroid hormones that prepare the body for pregnancy. Estrogen produces secondary sex characteristics in females, including breast development, maturation of the external genitalia, and widening of the hips, while both estrogen and progesterone regulate the menstrual cycle.

The Ovarian Cycle and the Uterine Cycle

Female reproduction occurs in a cyclical fashion in females. The **ovarian cycle** governs the preparation of endocrine tissues and release of eggs, while the **uterine cycle** governs the preparation and maintenance of the uterine lining to receive any potential fertilized eggs. The uterine cycle is sometimes referred to as the menstrual cycle. These cycles occur concurrently and are coordinated over a 22–32 day cycle, with an average length of 28 days. The events of these cycles and the hormone actions during them are summarized in ***Figure 2***.

ART CONNECTION

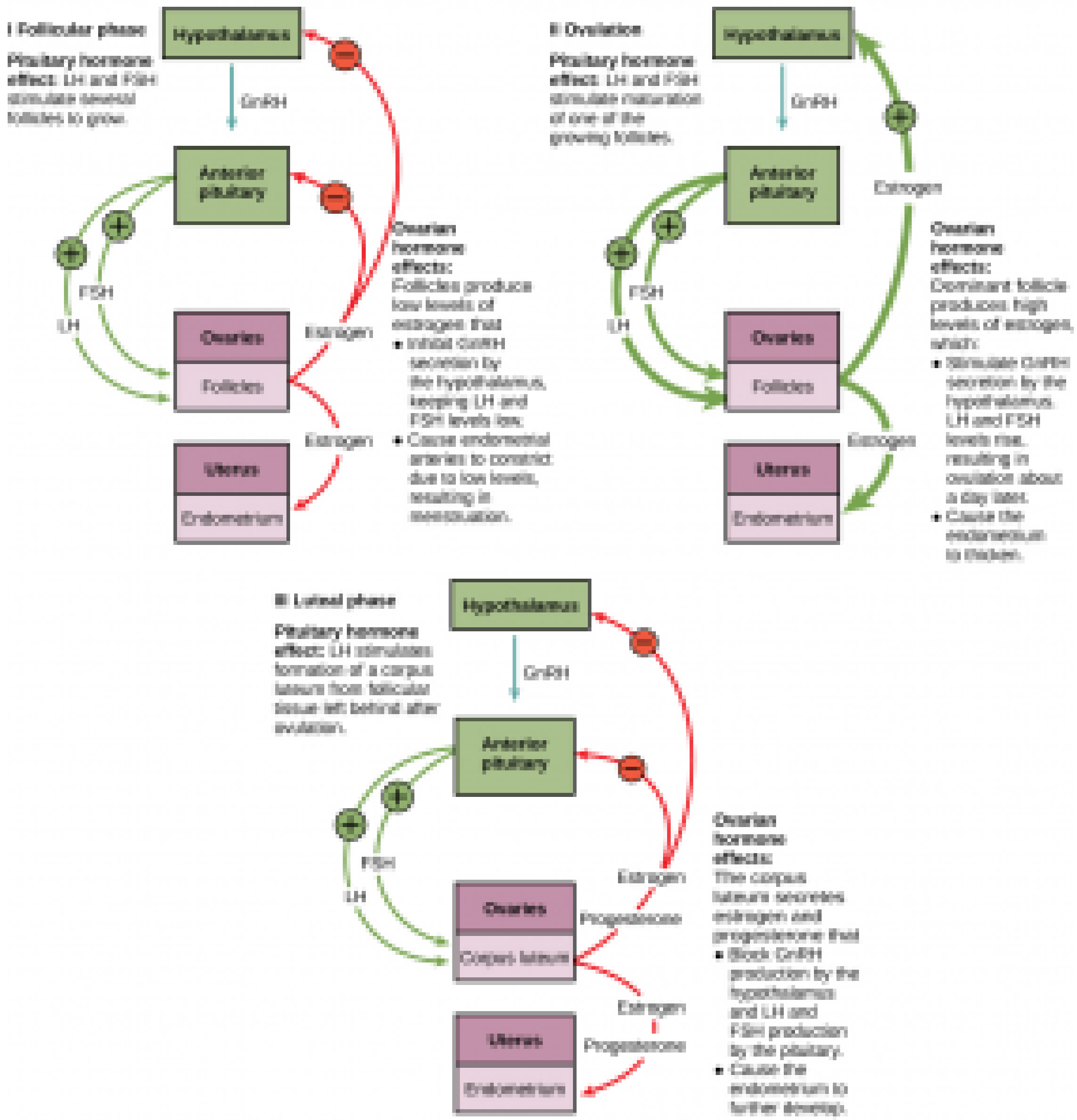


Figure 2. *The ovarian and menstrual cycles of female reproduction are regulated by hormones produced by the hypothalamus, pituitary, and ovaries. Note that estradiol is a type of estrogen.*

The timing of the uterine cycle starts with the first day of menses, referred to as day one of a females's period. Cycle length is determined by counting the days between the onset of bleeding in two subsequent cycles. Because the average length of a female's uterine cycle is 28 days, this is the time period used to identify the timing of events in the cycle. However, the length of the uterine cycle varies among women, and even in the same woman from one cycle to the next. The menses phase of the uterine cycle is the phase during which the lining is shed; that is, the days that the woman menstruates. Although it averages approximately five days, the menses phase can last from 2 to 7 days, or longer. FSH, LH and progesterone levels are low at this time; the drop in these levels causes the endometrial lining to shed.

Once menstrual flow ceases, the endometrium begins to proliferate again, marking the beginning of the proliferative phase of the menstrual cycle. This usually occurs at around day 5-7 of the cycle. The proliferative phase is named after the growth, or proliferation of the endometrium during this phase, which is caused by increasing concentrations of estrogen.

The first half of the ovarian cycle is the follicular phase, which occurs on day 1-14 of the cycle, overlapping with both menses and the proliferative phase of the uterine cycle. Slowly rising levels of FSH and LH cause the growth of follicles on the surface of the ovary. This process prepares the egg for ovulation. As the follicles grow, they begin releasing estrogens and a low level of progesterone. Just prior to the middle of the cycle (approximately day 14), the high level of estrogen causes FSH and especially LH to rise rapidly, then fall. The spike in LH causes **ovulation**: the most mature follicle, like that shown in ***Figure 3***, ruptures and releases its egg. The follicles that did not rupture degenerate and their eggs are lost. The level of estrogen decreases when the extra follicles degenerate.

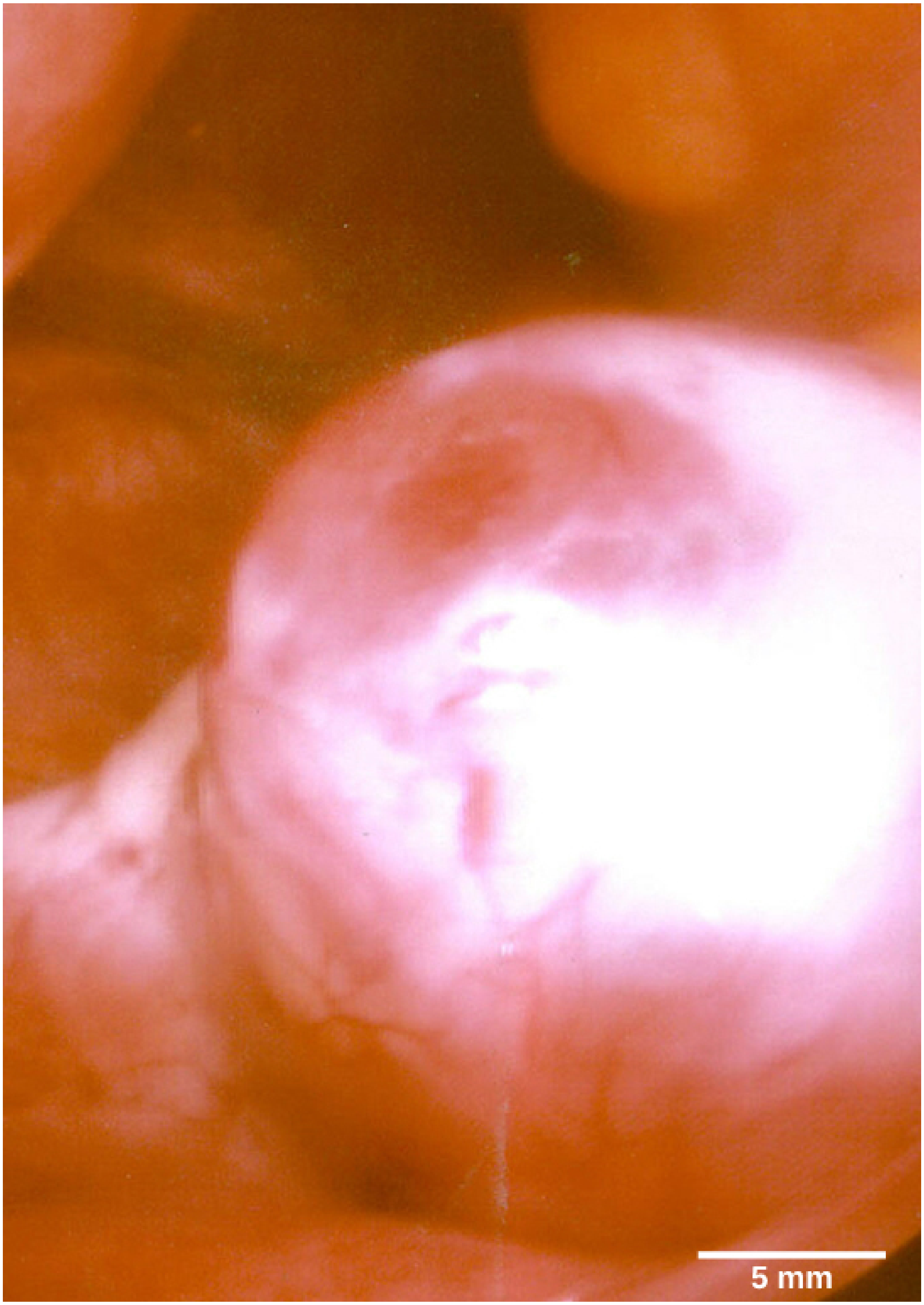


Figure 3. *This mature egg follicle may rupture and release an egg. (credit: scale-bar data from Matt Russell)*

Following ovulation, the ovarian cycle enters its luteal phase, illustrated in ***Figure 2,*** and the menstrual cycle enters its secretory phase, both of which run from about day 15 to 28. The luteal and secretory phases refer to changes in the ruptured follicle. The cells in the follicle undergo physical changes and produce a structure called a corpus luteum. The corpus luteum produces estrogen and progesterone. The progesterone facilitates the maintenance and further development of the uterine lining and inhibits the release of further FSH and LH. The uterus is being prepared to accept a fertilized egg, should it occur during this cycle. The inhibition of FSH and LH prevents any further eggs and follicles from developing, while the progesterone is elevated. The level of estrogen produced by the corpus luteum increases to a steady level for the next few days.

If no fertilized egg is implanted into the uterus, the corpus luteum degenerates and the levels of estrogen and progesterone decrease. The endometrium begins to degenerate as the progesterone levels drop, initiating the next menstrual cycle. The decrease in progesterone also results in increased FSH and LH secretion from the anterior pituitary, and starts the cycles again. ***Figure 4*** visually compares the ovarian and uterine cycles as well as the commensurate hormone levels.

ART CONNECTION

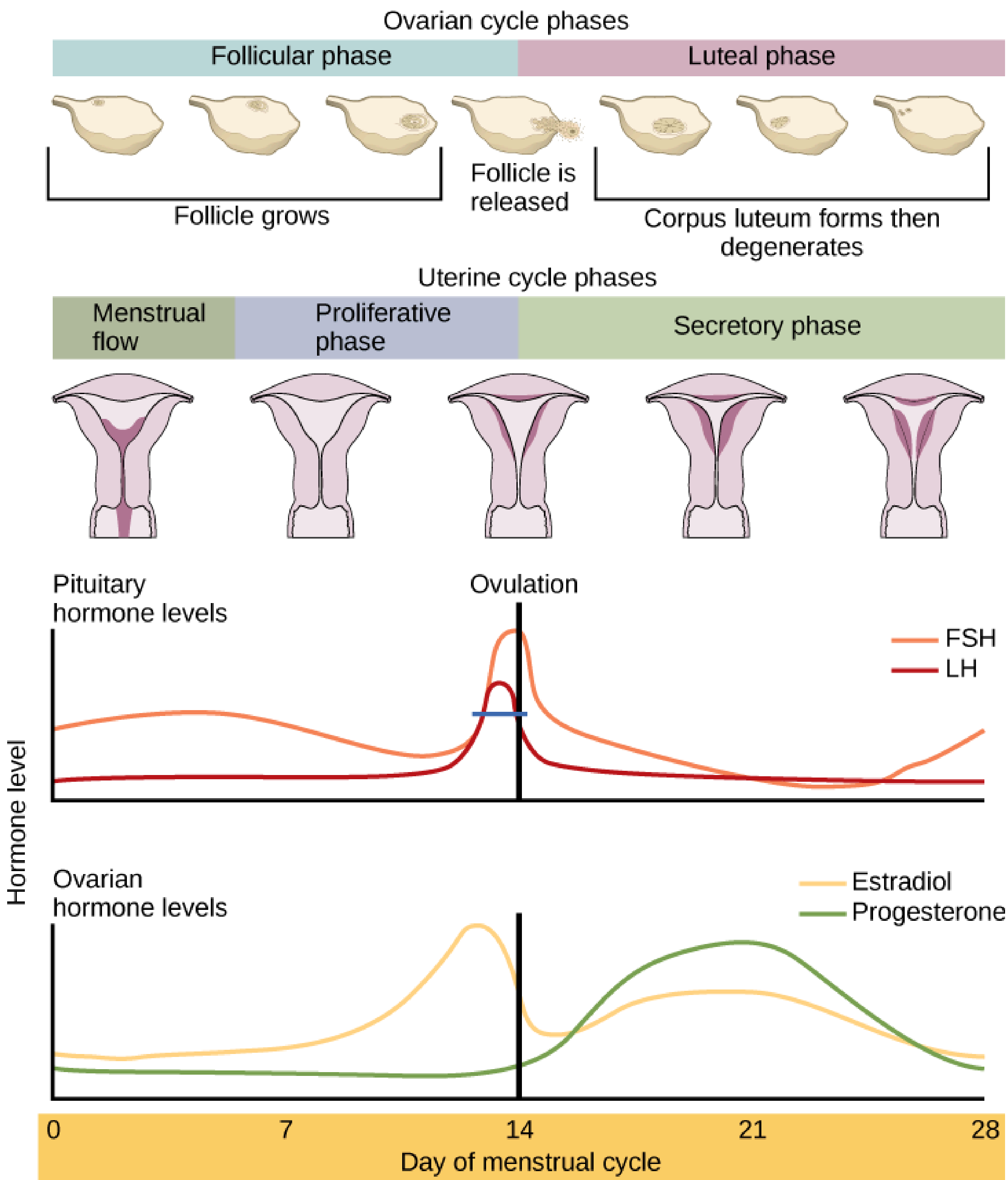


Figure 4. *Rising and falling hormone levels result in progression of the ovarian and menstrual cycles. Note that estradiol is a type of estrogen. (credit: modification of work by Mikael Häggström)*

Menopause

As women approach their mid-40s to mid-50s, their ovaries begin to lose their sensitivity to FSH and LH. Menstrual periods become less frequent and finally cease; this is **menopause**. There are still eggs and potential follicles on the ovaries, but without the stimulation of FSH and LH, they will not produce a viable egg to be released. The outcome of this is the inability to have children.

The side effects of menopause include hot flashes, heavy sweating (especially at night), headaches, some hair loss, muscle pain, vaginal dryness, insomnia, depression, weight gain, and mood swings. Estrogen is involved in calcium metabolism and, without it, blood levels of calcium decrease. To replenish the blood, calcium is lost from bone which may decrease the bone density and lead to osteoporosis. Supplementation of estrogen in the form of hormone replacement therapy (HRT) can prevent bone loss, but the therapy can have negative side effects. While HRT is thought to give some protection from colon cancer, osteoporosis, heart disease, macular degeneration, and possibly depression, its negative side effects include increased risk of: stroke or heart attack, blood clots, breast cancer, ovarian cancer, endometrial cancer, gall bladder disease, and possibly dementia.

CAREER CONNECTION

Reproductive Endocrinologist

A reproductive endocrinologist is a physician who treats a variety of hormonal disorders related to reproduction and infertility in both men and women. The disorders include menstrual problems, infertility, pregnancy loss, sexual dysfunction, and menopause. Doctors may use fertility drugs, surgery, or assisted reproductive techniques (ART) in their therapy. ART involves the use of procedures to manipulate the egg or sperm to facilitate reproduction, such as *in vitro* fertilization.

Reproductive endocrinologists undergo extensive medical training, first in a four-year residency in obstetrics and gynecology, then in a three-year fellowship in reproductive endocrinology. To be board certified in this area, the physician must pass written and oral exams in both areas.

Section Summary

The male and female reproductive cycles are controlled by hormones released from the hypothalamus and anterior pituitary as well as hormones from reproductive tissues and organs. The hypothalamus monitors the need for the FSH and LH hormones made and released from the anterior pituitary. FSH and LH affect reproductive structures to cause the formation of sperm and the preparation of eggs for release and possible fertilization. In the male, FSH and LH stimulate Sertoli cells and interstitial cells of Leydig in the testes to facilitate sperm production. The Leydig cells produce testosterone, which also is responsible for the secondary sexual characteristics of males. In females, FSH and LH cause estrogen and progesterone to be produced. They regulate the female reproductive system which is divided into the ovarian cycle and the menstrual cycle. Menopause occurs when the ovaries lose their sensitivity to FSH and LH and the female reproductive cycles slow to a stop.

Glossary

estrogen
: reproductive hormone in females that assists in endometrial regrowth, ovulation, and calcium absorption

follicle stimulating hormone (FSH)
: reproductive hormone that causes sperm production in men and follicle development in women

gonadotropin-releasing hormone (GnRH)
: hormone from the hypothalamus that causes the release of FSH and LH from the anterior pituitary

inhibin
: hormone made by Sertoli cells; provides negative feedback to hypothalamus in control of FSH and GnRH release

interstitial cell of Leydig
: cell in seminiferous tubules that makes testosterone

luteinizing hormone (LH)
: reproductive hormone in both men and women, causes testosterone production in men and ovulation and lactation in women

menopause
: loss of reproductive capacity in women due to decreased sensitivity of the ovaries to FSH and LH

ovarian cycle
: cycle of preparation of egg for ovulation and the conversion of the follicle to the corpus luteum

ovulation
: release of the egg by the most mature follicle

progesterone
: reproductive hormone in women; assists in endometrial re-growth and inhibition of FSH and LH release

Sertoli cell
: cell in seminiferous tubules that assists developing sperm and makes inhibin

testosterone
: reproductive hormone in men that assists in sperm production and promoting secondary sexual characteristics

uterine cycle
: cycle of the degradation and re-growth of the endometrium; also called the menstrual cycle

38.

4.5 From Fertilization to Birth Part 1: Fertilization and Early Development

The process in which an organism develops from a single-celled zygote to a multi-cellular organism is complex and well-regulated. The early stages of embryonic development are also crucial for ensuring the fitness of the organism. In this section and the next, we will cover the entire process that results in a new human, starting from fertilization, through the various stages of a pregnancy, and finally with the birth of a child.

Fertilization

Fertilization, pictured in ***Figure 1a*** is the process in which gametes (an egg and sperm) fuse to form a zygote. The egg and sperm each contain one set of chromosomes. To ensure that the offspring has only one complete diploid set of chromosomes, only one sperm must fuse with one egg. In mammals, the egg is protected by a layer of extracellular matrix consisting mainly of glycoproteins called the **zona pellucida**. When a sperm binds to the zona pellucida, a series of biochemical events, called the **acrosomal reactions**, take place. In placental mammals, the acrosome contains digestive enzymes that initiate the degradation of the glycoprotein matrix protecting the egg and allowing the sperm plasma membrane to fuse with the egg plasma membrane, as illustrated in ***Figure 1b***. The fusion of these two membranes creates an opening through which the sperm nucleus is transferred into the ovum. The nuclear membranes of the egg and sperm break down and the two haploid genomes condense to form a diploid genome.

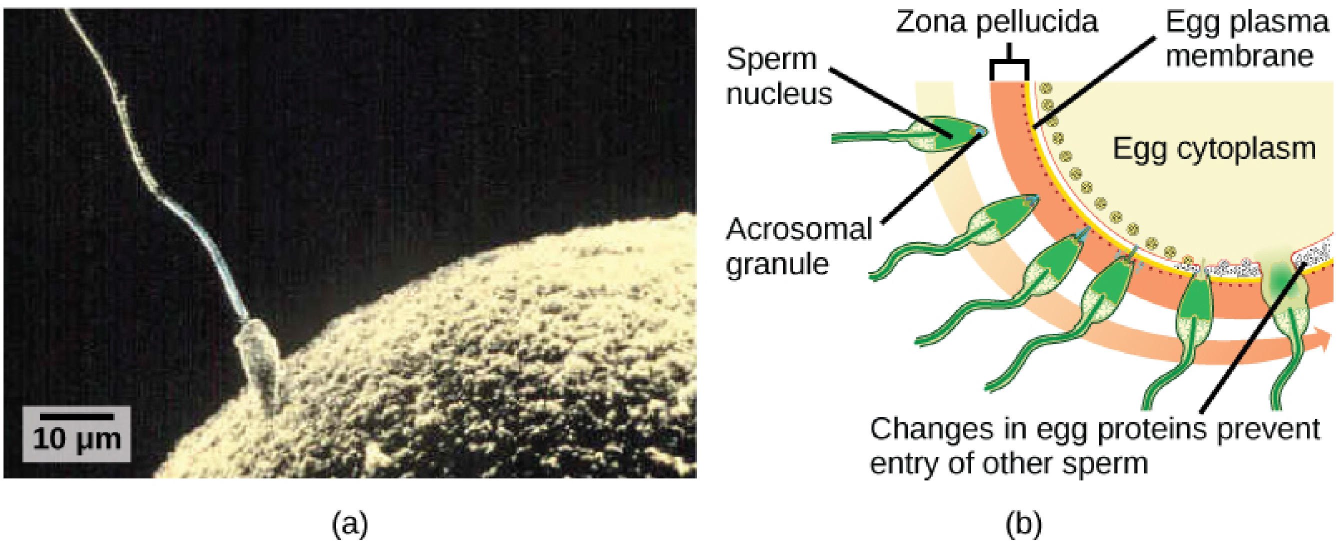


Figure 1. (a) *Fertilization is the process in which sperm and egg fuse to form a zygote.* ***(b)*** *Acrosomal reactions help the sperm degrade the glycoprotein matrix protecting the egg and allow the sperm to transfer its nucleus. (credit: (b) modification of work by Mariana Ruiz Villareal; scale-bar data from Matt Russell)*

To ensure that no more than one sperm fertilizes the egg, once the acrosomal reactions take place at one location of the egg membrane, the egg releases proteins in other locations to prevent other sperm from fusing with the egg. If this mechanism fails, multiple sperm can fuse with the egg, resulting in **polyspermy**. The resulting embryo is not genetically viable and dies within a few days.

Cleavage and Blastula Stage

The development of multi-cellular organisms begins from a single-celled zygote, which undergoes rapid cell division to form the blastula. The rapid, multiple rounds of cell division are termed cleavage. Cleavage is illustrated in ***Figure 2a***. After the cleavage has produced over 100 cells, the embryo is called a blastula. The blastula is usually a spherical layer of cells (the blastoderm) surrounding a fluid-filled or yolk-filled cavity (the blastocoel). Mammals at this stage form a structure called the blastocyst, characterized by an inner cell mass that is distinct from the surrounding blastula, shown in ***Figure 2b***. During cleavage, the cells divide without an increase in mass; that is, one large single-celled zygote divides into multiple smaller cells. Each cell within the blastula is called a blastomere.

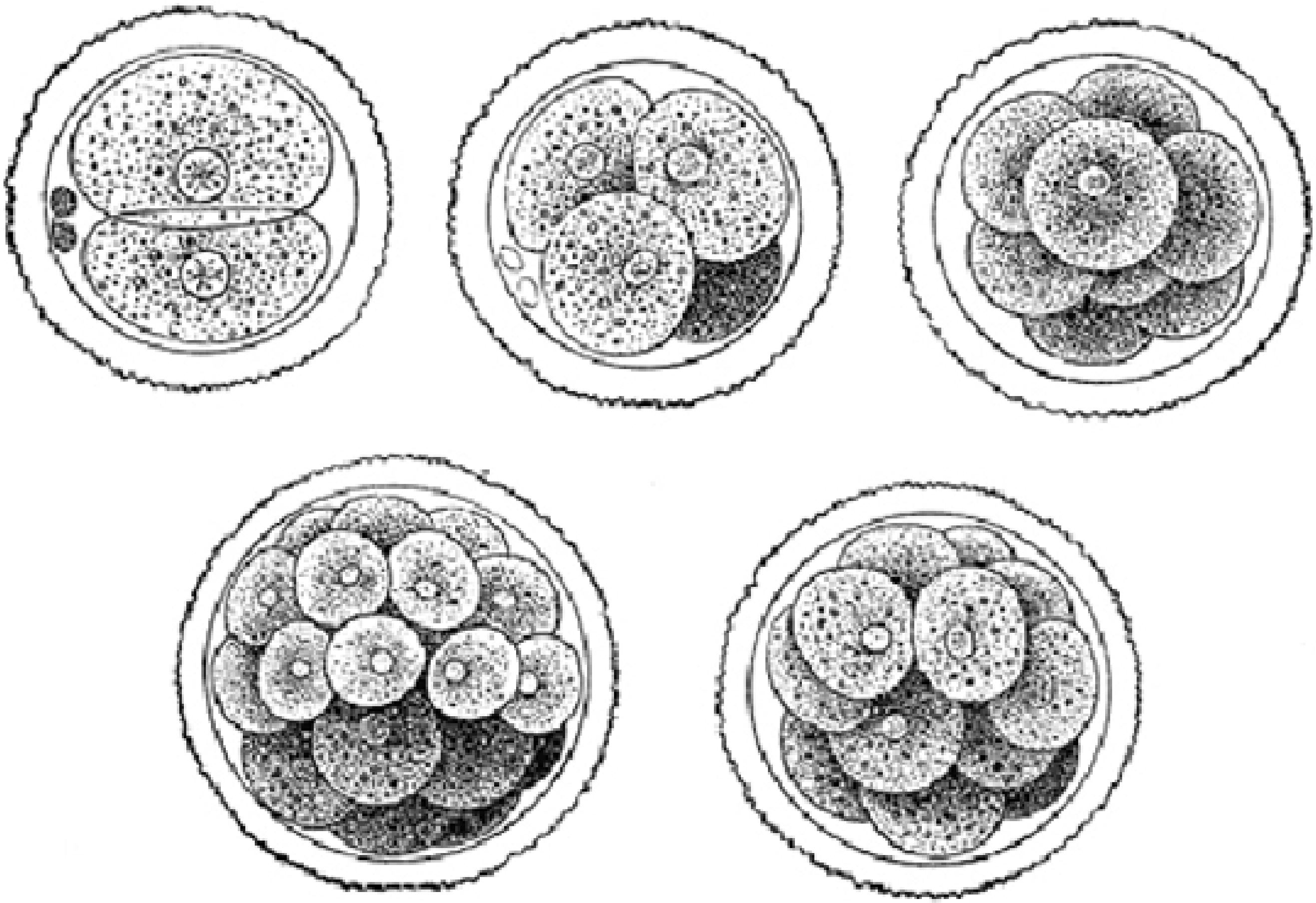


__Figure 2. (a)__ During cleavage, the zygote rapidly divides into multiple cells without increasing in size.

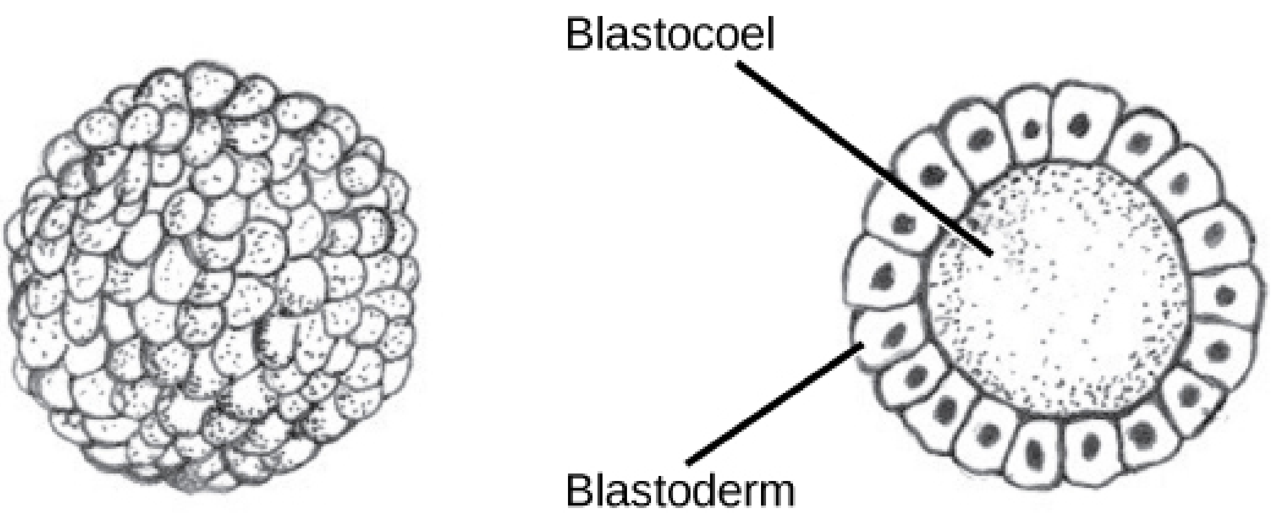


__Figure 2. (b)__ The cells rearrange themselves to form a hollow ball with a fluid-filled or yolk-filled cavity called the blastula. (credit a: modification of work by Gray's Anatomy; credit b: modification of work by Pearson Scott Foresman, donated to the Wikimedia Foundation)

In mammals, the blastula forms the **blastocyst** in the next stage of development. Here the cells in the blastula arrange themselves in two layers: the **inner cell mass**, and an outer layer called the **trophoblast**. The inner cell mass is also known as the embryoblast and this mass of cells will go on to form the embryo. At this stage of development, illustrated in ***Figure 3*** the inner cell mass consists of embryonic stem cells that will differentiate into the different cell types needed by the organism. The trophoblast will contribute to the placenta and nourish the embryo.

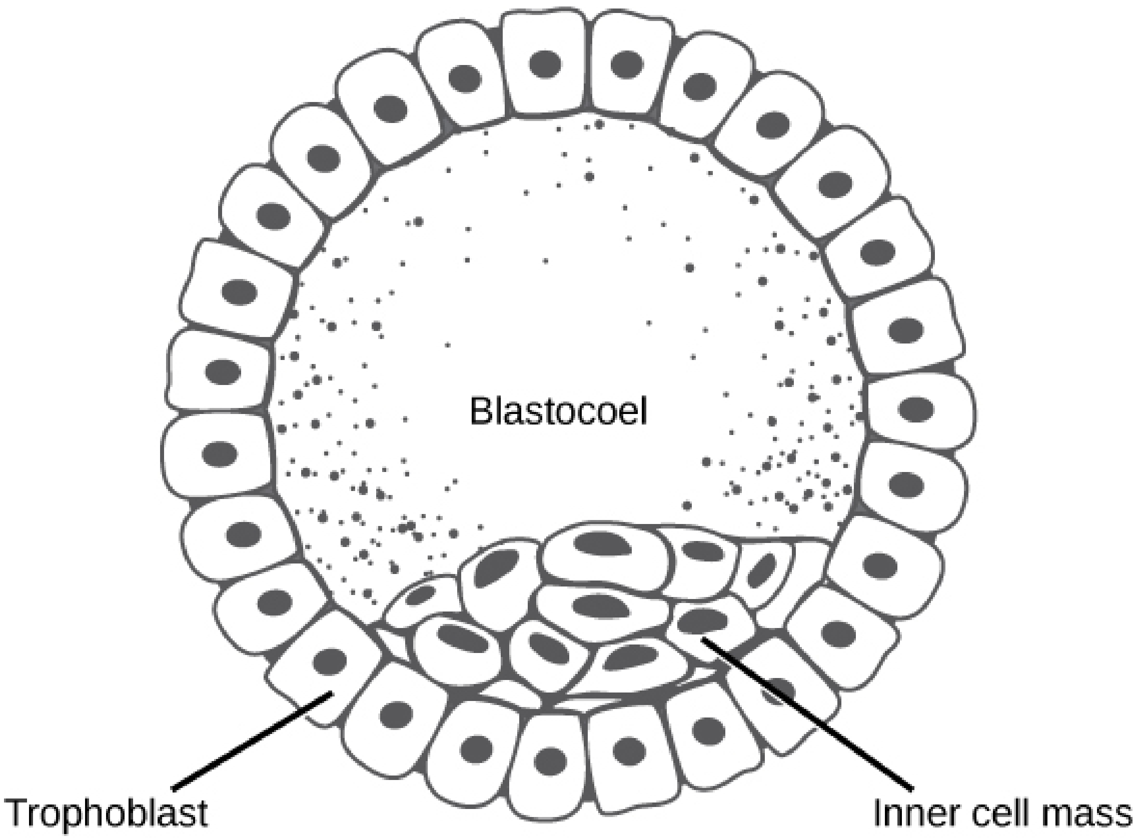


Figure 3. *The rearrangement of the cells in the mammalian blastula to two layers—the inner cell mass and the trophoblast—results in the formation of the blastocyst.*

Link to Learning

Visit the [Virtual Human Embryo project]() at the Endowment for Human Development site to step through an interactive that shows the stages of embryo development, including micrographs and rotating 3-D images.

Rapid, successive cell divisions are crucial for the normal development of a human embryo. In most cases, however, rapid cell divisions cease after the early stages of development of an embryo. Compared to the number of cell divisions in a very early embryo, the rate of cell divisions in a fetus close to birth is much slower. Although infants and children do have considerable cell division as they grow, the rates are again much slower than these early cell divisions in the embryo. Most adult tissues do not undergo cell division, with the exception of some adult stem cells that generate blood, skin, gametes, and the lining of the digestive tract, for example. The process of cell division is under considerable genetic and molecular controls in children and adults. The failure of these controls, and the return to rapid, successive cell divisions, is a hallmark of cancers.

Gastrulation

The typical blastula is a ball of cells. The next stage in embryonic development is the formation of the body plan. The cells in the blastula rearrange themselves spatially to form three layers of cells. This process is called **gastrulation**. During gastrulation, the blastula folds upon itself to form the three layers of cells. Each of these layers is called a germ layer and each germ layer differentiates into different organ systems.

The three germs layers, shown in ***Figure 4***, are the endoderm, the ectoderm, and the mesoderm. The cells in each layer have specified "fates"; in other words, they are programmed to develop further into discrete sets of organs and organ systems. The ectoderm gives rise to the nervous system and the epidermis, or skin. The mesoderm gives rise to the muscle cells and connective tissue in the body. The endoderm gives rise to cells found in the digestive system and many internal organs.

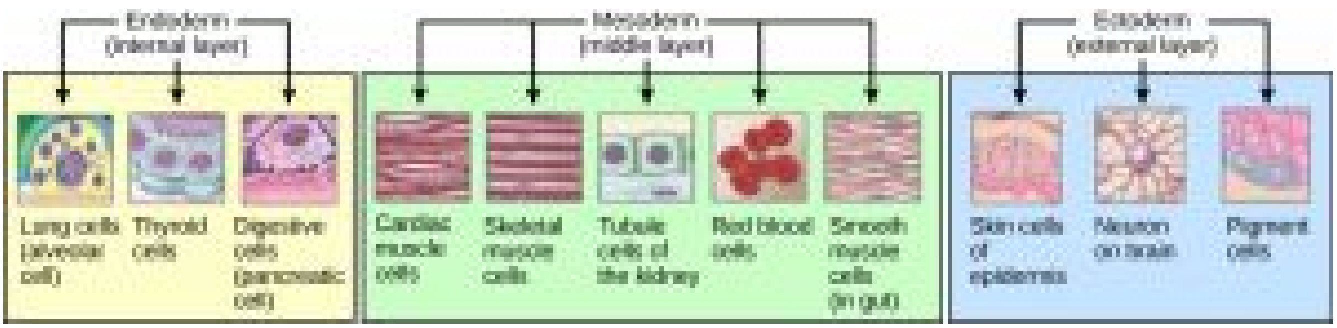


Figure 4. *The three germ layers give rise to different cell types in the animal body. (credit: modification of work by NIH, NCBI)*

EVERYDAY CONNECTION

Are Designer Babies in Our Future?

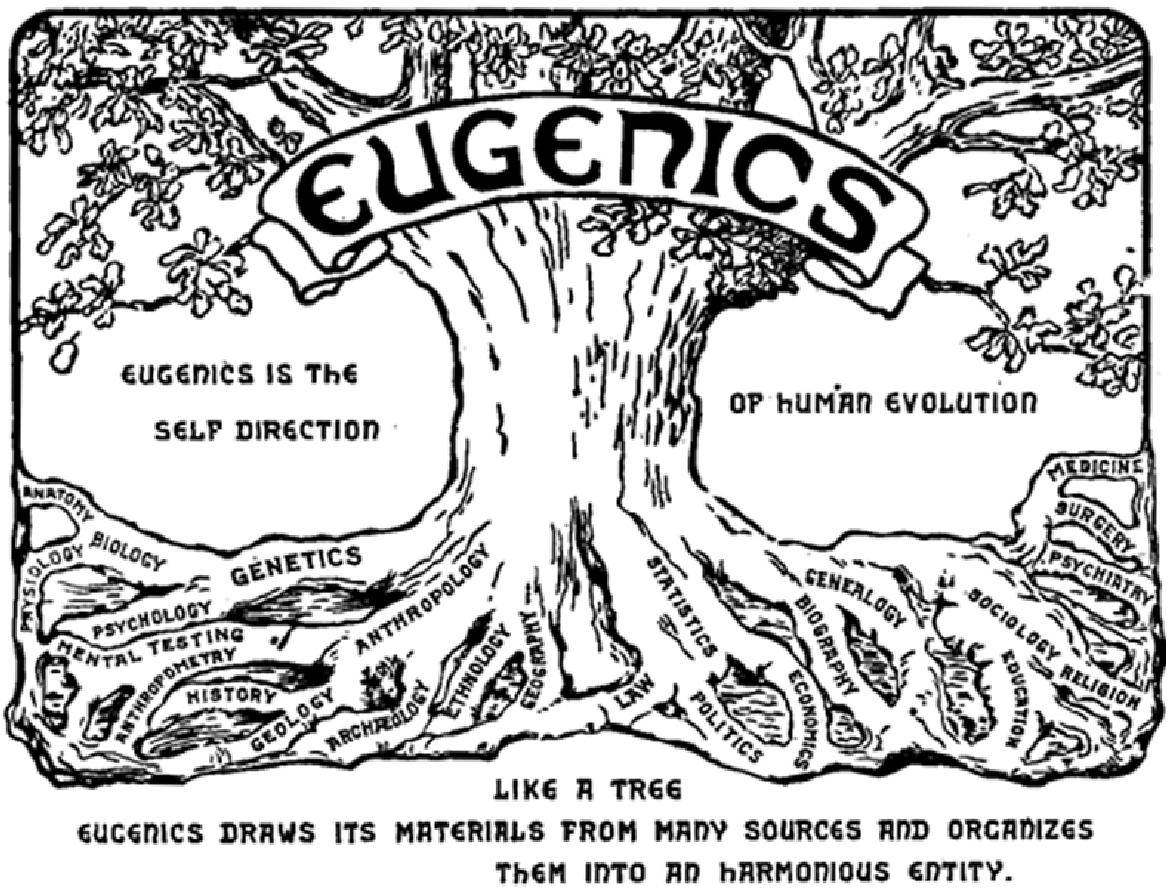


Figure 5. *This logo from the Second International Eugenics Conference in New York City in September of 1921 shows how eugenics attempted to merge several fields of study with the goal of producing a genetically superior human race.*

If you could prevent your child from getting a devastating genetic disease, would you do it? Would you select the sex or gender of your child or select for their attractiveness, strength, or intelligence? How far would you go to maximize the possibility of resistance to disease? The genetic engineering of a human child, the production of "designer babies" with desirable phenotypic characteristics, was once a topic restricted to science fiction. This is the case no longer: science fiction is now overlapping into science fact. Many phenotypic choices for offspring are already available, with many more likely to be possible in the not too distant future. Which traits should be selected and how they should be selected are topics of much debate within the worldwide medical community. The ethical and moral line is not always clear or agreed upon, and some fear that modern reproductive technologies could lead to a new form of eugenics.

Eugenics is the use of information and technology from a variety of sources to improve the genetic makeup of the human race. The goal of creating genetically superior humans was quite prevalent (although controversial) in several countries during the early 20th century (***Figure 5***), but fell into disrepute when Nazi Germany developed an extensive eugenics program in the 1930's and 40's. As part of their program, the Nazis forcibly sterilized hundreds of thousands of the so-called "unfit" and killed tens of thousands of institutionally disabled people as part of a systematic program to develop a genetically superior race of Germans known as Aryans. Ever since, eugenic ideas have not been as publicly expressed, but there are still those who promote them.

Efforts have been made in the past to control traits in human children using donated sperm from men with desired traits. In fact, eugenicist Robert Klark Graham established a sperm bank in 1980 that included samples exclusively from donors with high IQs. The "genius" sperm bank failed to capture the public's imagination and the operation closed in 1999.

In more recent times, the procedure known as prenatal genetic diagnosis (PGD) has been developed. PGD involves the screening of human embryos as part of the process of *in vitro* fertilization, during which embryos are conceived and grown outside the mother's body for some period of time before they are implanted. The term PGD usually refers to both the diagnosis, selection, and the implantation of the selected embryos.

In the least controversial use of PGD, embryos are tested for the presence of alleles which cause genetic diseases such as sickle cell disease, muscular dystrophy, and hemophilia, in which a single disease-causing allele or pair of alleles has been identified. By excluding embryos containing these alleles from implantation into the mother, the disease is prevented, and the unused embryos are either donated to science or discarded. There are relatively few in the worldwide medical community that question the ethics of this type of procedure, which allows individuals scared to have children because of the alleles they carry to do so successfully. The major limitation to this procedure is its expense. Not usually covered by medical insurance and thus out of reach financially for most couples, only a very small percentage of all live births use such complicated methodologies. Yet, even in cases like these where the ethical issues may seem to be clear-cut, not everyone agrees with the morality of these types of procedures. For example, to those who take the position that human life begins at conception, the discarding of unused embryos, a necessary result of PGD, is unacceptable under any circumstances.

A murkier ethical situation is found in the selection of a child's chromosomal sex, which is easily performed by PGD. Currently, countries such as Great Britain have banned the selection of a child's sex for reasons other than preventing sex-linked diseases. Other countries allow the procedure for "family balancing", based on the desire of some parents to have at least one child of each sex. Still others, including the United States, have taken a scattershot approach to regulating these practices, essentially leaving it to the individual practicing physician to decide which practices are acceptable and which are not.

Even murkier are rare instances of disabled parents, such as those with deafness or dwarfism, who select embryos via PGD to ensure that they share their disability. These parents usually cite many positive aspects of their disabilities and associated culture as reasons for their choice, which they see as their moral right. To others, to purposely cause a disability in a child violates the basic medical principle of *Primum non nocere*, "first, do no harm." This procedure, although not illegal in most countries, demonstrates the complexity of ethical issues associated with choosing genetic traits in offspring.

Where could this process lead? Will this technology become more affordable and how should it be used? With the ability of technology to progress rapidly and unpredictably, a lack of definitive guidelines for the use of reproductive technologies before they arise might make it difficult for legislators to keep pace once they are in fact realized, assuming the process needs any government regulation at all. Other bioethicists argue that we should only deal with technologies that exist now, and not in some uncertain future. They argue that these types of procedures will always be expensive and rare, so the fears of eugenics and “master” races are unfounded and overstated. The debate continues.

Section Summary

The early stages of embryonic development begin with fertilization. The process of fertilization is tightly controlled to ensure that only one sperm fuses with one egg. After fertilization, the zygote undergoes cleavage to form the blastula. The blastula, which in some species is a hollow ball of cells, undergoes a process called gastrulation, in which the three germ layers form. The ectoderm gives rise to the nervous system and the epidermal skin cells, the mesoderm gives rise to the muscle cells and connective tissue in the body, and the endoderm gives rise to columnar cells and internal organs.

Glossary

acrosomal reaction
: series of biochemical reactions that the sperm uses to break through the zona pellucida

blastocyst
: structure formed when cells in the mammalian blastula separate into an inner and outer layer

gastrulation
: process in which the blastula folds over itself to form the three germ layers

inner cell mass
: inner layer of cells in the blastocyst

polyspermy
: condition in which one egg is fertilized by multiple sperm

trophoblast
: outer layer of cells in the blastocyst

zona pellucida
: protective layer of glycoproteins on the mammalian egg

39.

4.6 From Fertilization to Birth Part II: Pregnancy and Birth

Pregnancy begins with the fertilization of an egg and continues through to the birth of the individual. The length of time of **gestation** varies among animals, but is very similar among the great apes: human gestation is 266 days, while chimpanzee gestation is 237 days, a gorilla's is 257 days, and orangutan gestation is 260 days long. The fox has a 57-day gestation. Dogs and cats have similar gestations averaging 60 days. The longest gestation for a land mammal is an African elephant at 640 days. The longest gestations among marine mammals are the beluga and sperm whales at 460 days.

Human Gestation

Twenty-four hours before fertilization, the egg has finished meiosis and becomes a mature oocyte. When fertilized (at conception) the egg becomes known as a zygote. The zygote, which at this stage is also known as a blastocyst, travels through the oviduct to the uterus (***Figure 1***). The developing embryo must implant into the wall of the uterus within seven days, or it will deteriorate and die. The outer layers of the zygote (trophoblast) grow into the endometrium by digesting the endometrial cells, and wound healing of the endometrium closes up the blastocyst into the tissue. Another layer of the blastocyst, the chorion, begins releasing a hormone called **human beta chorionic gonadotropin (β-HCG)** which makes its way to the corpus luteum and keeps that structure active. This ensures adequate levels of progesterone that will maintain the endometrium of the uterus for the support of the developing embryo. Pregnancy tests determine the level of β-HCG in urine or serum. If the hormone is present, the test is positive.

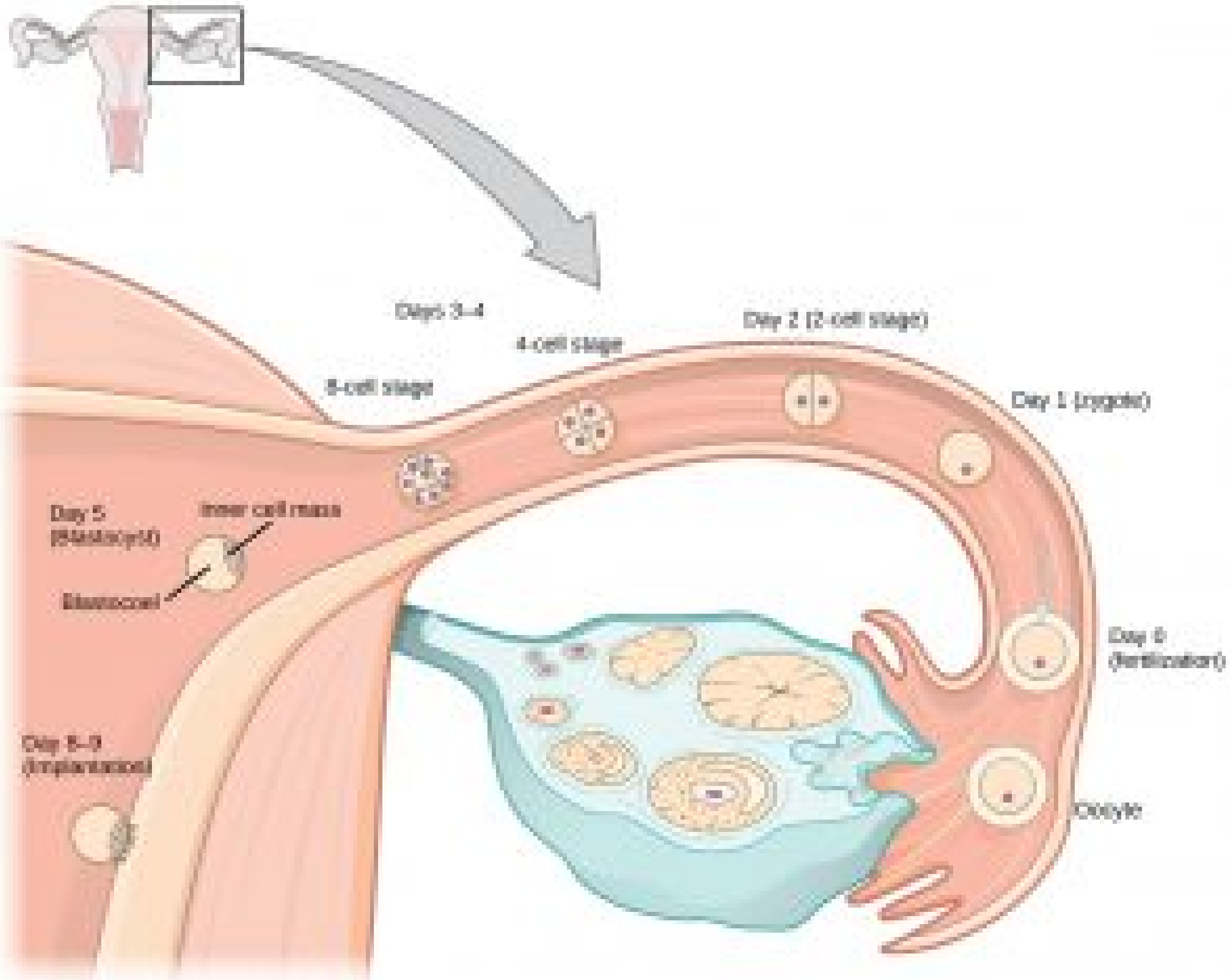


Figure 1. *In humans, fertilization occurs soon after the oocyte leaves the ovary. Implantation occurs eight or nine days later.(credit: Ed Uthman)*

The gestation period is divided into three equal periods or trimesters. During the first two to four weeks of the first trimester, nutrition and waste are handled by the endometrial lining through diffusion. As the trimester progresses, the outer layer of the embryo begins to merge with the endometrium, and the **placenta** forms. This organ takes over the nutrient and waste requirements of the embryo and fetus, with the mother's blood passing nutrients to the placenta and removing waste from it. Chemicals from the fetus, such as bilirubin, are processed by the mother's liver for elimination. Some of the mother's immunoglobulins will pass through the placenta, providing passive immunity against some potential infections.

Internal organs and body structures begin to develop during the first trimester. By five weeks, limb buds, eyes, the heart, and liver have been basically formed. By eight weeks, the term fetus applies, and the body is essentially formed, as shown in ***Figure 2***. The individual is about five centimeters (two inches) in length and many of the organs, such as the lungs and liver, are not yet functioning. Exposure to any toxins is especially dangerous during the first trimester, as all of the body's organs and structures are going through initial development. Anything that affects that development can have a severe effect on the fetus’ survival.

Figure 2. *Fetal development is shown at nine weeks gestation. (credit: Ed Uthman)*

During the second trimester, the fetus grows to about 30 cm (12 inches), as shown in ***Figure 3***. It becomes active and the mother usually feels the first movements. All organs and structures continue to develop. The placenta has taken over the functions of nutrition and waste and the production of estrogen and progesterone from the corpus luteum, which has degenerated. The placenta will continue functioning up through the delivery of the baby.

Figure 3. *This fetus is just entering the second trimester, when the placenta takes over more of the functions performed as the baby develops. (credit: National Museum of Health and Medicine)*

During the third trimester, the fetus grows to 3 to 4 kg (6 ½ -8 ½ lbs.) and about 50 cm (19-20 inches) long, as illustrated in ***Figure 4***. This is the period of the most rapid growth during the pregnancy. Organ development continues to birth (and some systems, such as the nervous system and liver, continue to develop after birth). The mother will be at her most uncomfortable during this trimester. She may urinate frequently due to pressure on the bladder from the fetus. There may also be intestinal blockage and circulatory problems, especially in her legs. Clots may form in her legs due to pressure from the fetus on returning veins as they enter the abdominal cavity.

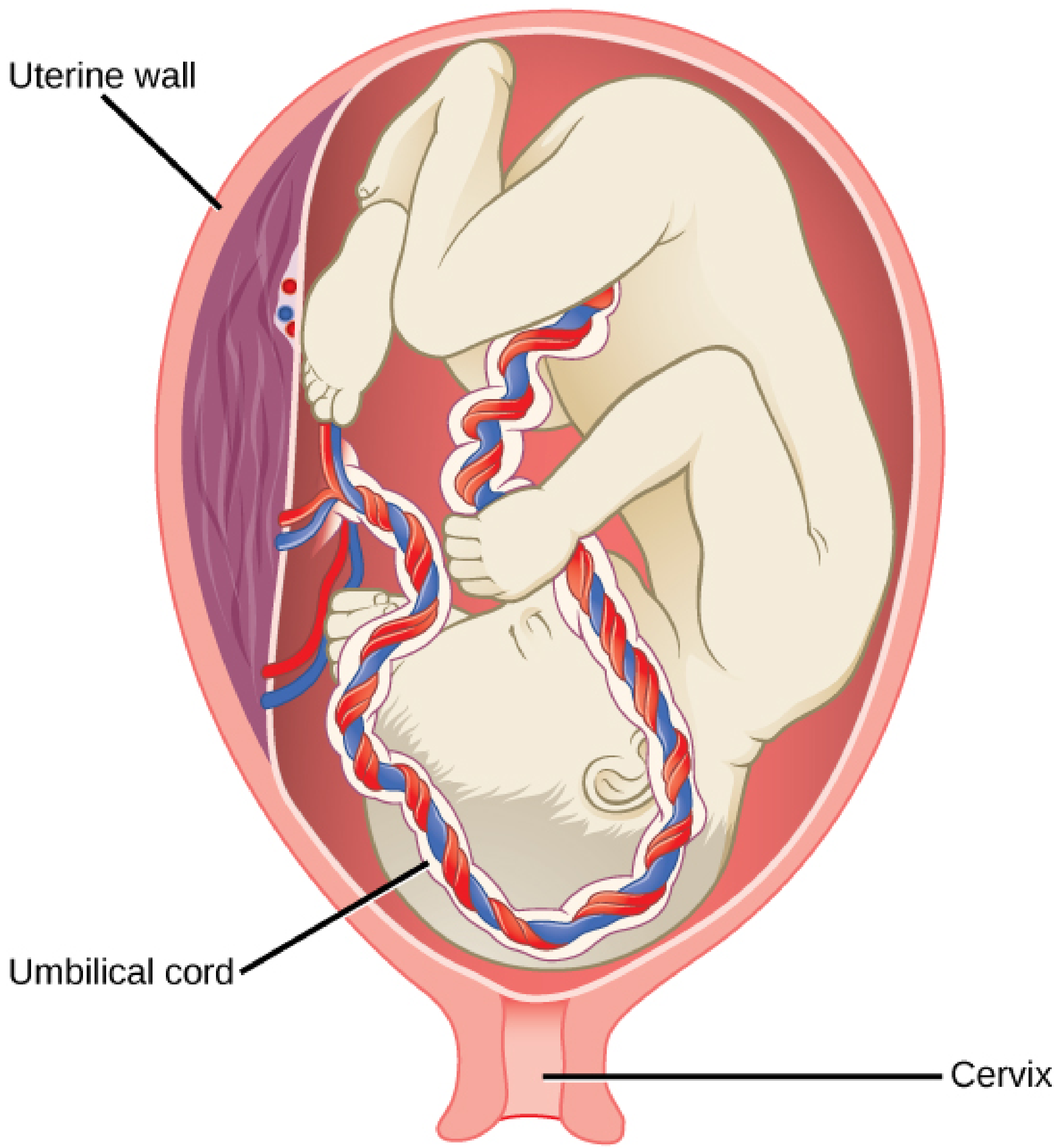


Figure 4. *There is rapid fetal growth during the third trimester. (credit: modification of work by Gray's Anatomy)*

Link to Learning

Visit this site to see the stages of human fetal development.

Labor and Birth

Labor is the physical efforts of expulsion of the fetus and the placenta from the uterus during birth (parturition). Toward the end of the third trimester, estrogen causes receptors on the uterine wall to develop and bind the hormone oxytocin. At this time, the baby reorients, facing forward and down with the back or crown of the head engaging the cervix (uterine opening). This causes the cervix to stretch and nerve impulses are sent to the hypothalamus, which signals for the release of oxytocin from the posterior pituitary. The oxytocin causes the smooth muscle in the uterine wall to contract. At the same time, the placenta releases prostaglandins into the uterus, increasing the contractions. A positive feedback relay occurs between the uterus, hypothalamus, and the posterior pituitary to assure an adequate supply of oxytocin. As more smooth muscle cells are recruited, the contractions increase in intensity and force.

There are three stages to labor. During stage one, the cervix thins and dilates. This is necessary for the baby and placenta to be expelled during birth. The cervix will eventually dilate to about 10 cm. During stage two, the baby is expelled from the uterus. The uterus contracts and the mother pushes as she compresses her abdominal muscles to aid the delivery. The last stage is the passage of the placenta after the baby has been born and the organ has completely disengaged from the uterine wall. If labor should stop before stage two is reached, synthetic oxytocin, known as Pitocin, can be administered to restart and maintain labor.

An alternative to labor and delivery is the surgical delivery of the baby through a procedure called a Caesarian section. This is major abdominal surgery and can lead to post-surgical complications for the mother, but in some cases it may be the only way to safely deliver the baby.

The mother's mammary glands go through changes during the third trimester to prepare for lactation and breastfeeding. When the baby begins suckling at the breast, signals are sent to the hypothalamus causing the release of prolactin from the anterior pituitary. Prolactin causes the mammary glands to produce milk. Oxytocin is also released, promoting the release of the milk. The milk contains nutrients for the baby's development and growth as well as immunoglobulins to protect the child from bacterial and viral infections.

Contraception and Birth Control

The prevention of a pregnancy comes under the terms contraception or birth control. Strictly speaking, **contraception** refers to preventing the sperm and egg from joining. Both terms are, however, frequently used interchangeably.

Table 1. Contraceptive Methods

Method	Examples	Failure Rate in Typical Use Over 12 Months
Barrier	male condom, female condom, sponge, cervical cap, diaphragm, spermicides	12-21%
Hormonal	oral, patch, vaginal ring	7%
	injection	4%
	implant, some intrauterine devices	less than 1%
Other	natural family planning	2 to 23%
	withdrawal	27%
	sterilization, some intrauterine devices	less than 1%

modified from Trussell J et al. eds. Contraceptive technology. 21st ed. New York, NY: Ayer Company Publishers, Inc., 2018.

Table 1 lists common methods of contraception. The failure rates listed are not the ideal rates that could be realized, but the typical rates that occur. A failure rate is the number of pregnancies resulting from the method's use over a twelve-month period. Barrier methods, such as condoms, cervical caps, and diaphragms, block sperm from entering the uterus, preventing fertilization. Spermicides are chemicals that are placed in the vagina that kill sperm. Sponges, which are saturated with spermicides, are placed in the vagina at the cervical opening. Combinations of spermicidal chemicals and barrier methods achieve lower failure rates than do the methods when used separately.

Nearly a quarter of the couples using barrier methods, natural family planning, or withdrawal can expect a failure of the method. Natural family planning is based on the monitoring of the menstrual cycle and having intercourse only during times when the egg is not available. A female's body temperature may rise a degree Celsius at ovulation and the cervical mucus may increase in volume and become more pliable. These changes give a general indication of when intercourse is more or less likely to result in fertilization. Withdrawal involves the removal of the penis from the vagina during intercourse, before ejaculation occurs. This is a risky method with a high failure rate due to the possible presence of sperm in the bulbourethral gland's secretion, which may enter the vagina prior to removing the penis.

Hormonal methods use synthetic progesterone (sometimes in combination with estrogen), to inhibit the hypothalamus from releasing FSH or LH, and thus prevent an egg from being available for fertilization. The method of administering the hormone affects failure rate. The most reliable method, with a failure rate of less than 1 percent, is the implantation of the hormone under the skin. The same rate can be achieved through the sterilization procedures of vasectomy in the male or of tubal ligation in the female, or by using an intrauterine device (IUD). IUDs are inserted into the uterus and establish an inflammatory condition that prevents fertilized eggs from implanting into the uterine wall. Some IUDs also prevent ovulation, or prevent sperm from entering the cervix and uterus.

Compliance with the contraceptive method is a strong contributor to the success or failure rate of any particular method. The only method that is completely effective at preventing conception is abstinence.

The choice of contraceptive method depends on the goals of the female or couple. Tubal ligation and vasectomy are considered permanent prevention, while other methods are reversible and provide short-term contraception.

Termination of an existing pregnancy can be spontaneous or voluntary. Spontaneous termination is also known as miscarriage and usually occurs very early in the pregnancy, usually within the first few weeks. This occurs when the fetus cannot develop properly and the gestation is naturally terminated, and is very common. About one fifth of all clinically recognized pregnancies end in spontaneous termination. Voluntary termination of a pregnancy is referred to as abortion. Laws regulating abortion vary between states and tend to view fetal viability as the criteria for allowing or preventing the procedure.

Infertility

Infertility is the inability to conceive a child or carry a child to birth. About 75 percent of causes of infertility can be identified; these include diseases, such as sexually transmitted diseases that can cause scarring of the reproductive tubes in either males or females, or developmental problems frequently related to abnormal hormone levels in one of the individuals. Inadequate nutrition, especially starvation, can delay menstruation. Stress can also lead to infertility. Short-term stress can affect hormone levels, while long-term stress can delay puberty and cause less frequent menstrual cycles. Other factors that affect fertility include toxins (such as cadmium), tobacco smoking, marijuana use, gonadal injuries, and aging.

If infertility is identified, several assisted reproductive technologies (ART) are available to aid conception. A common type of ART is *in vitro* fertilization (IVF) where an egg and sperm are combined outside the body and then placed in the uterus. Eggs are obtained from the female after extensive hormonal treatments that prepare mature eggs for fertilization and prepare the uterus for implantation of the fertilized egg. Sperm are obtained from the male and they are combined with the eggs and supported through several cell divisions to ensure viability of the zygotes. When the embryos have reached the eight-cell stage, one or more is implanted into the female's uterus. If fertilization is not accomplished by simple IVF, a procedure that injects the sperm into an egg can be used. This is called intracytoplasmic sperm injection (ICSI) and is shown in ***Figure 5***. IVF procedures produce a surplus of fertilized eggs and embryos that can be frozen and stored for future use. The procedures can also result in multiple births.

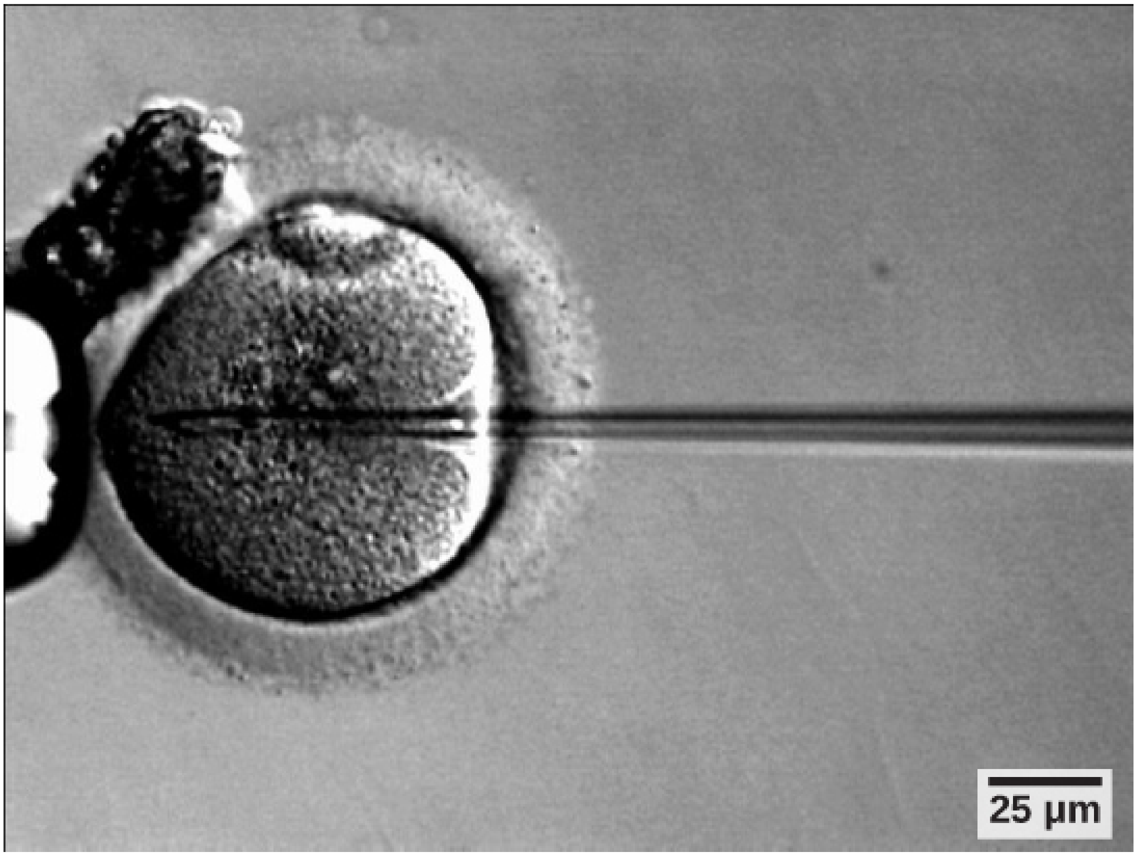


Figure 5. *A sperm is inserted into an egg for fertilization during intracytoplasmic sperm injection (ICSI). (credit: scale-bar data from Matt Russell)*

Section Summary

Human pregnancy begins with fertilization of an egg and proceeds through the three trimesters of gestation. The labor process has three stages (contractions, delivery of the fetus, expulsion of the placenta), each propelled by hormones. The first trimester lays down the basic structures of the body, including the limb buds, heart, eyes, and the liver. The second trimester continues the development of all of the organs and systems. The third trimester exhibits the greatest growth of the fetus and culminates in labor and delivery. Prevention of a pregnancy can be accomplished through a variety of methods including barriers, hormones, or other means. Assisted reproductive technologies may help individuals who have infertility problems.

Glossary

contraception
: (also, birth control) various means used to prevent pregnancy

gestation
: length of time for fetal development to birth

human beta chorionic gonadotropin (*β*-HCG)
: hormone produced by the chorion of the zygote that helps to maintain the corpus luteum and elevated levels of progesterone

infertility
: inability to conceive, carry, and deliver children

morning sickness
: condition in the mother during the first trimester; includes feelings of nausea

placenta
: organ that supports the diffusion of nutrients and waste between the mother’s and fetus’ blood

40.

4.7 The Cell Cycle

In this section, we will cover the cell cycle in more detail. We briefly addressed this in a previous chapter, when covering meiosis and gametogenesis. Now, we will explore in more detail what happens to allow a single cell to divide and produce identical copies of itself, a process that is crucial for the normal development of an embryo, but that can lead to cancers if uncontrolled.

The cell cycle

The **cell cycle** is an ordered series of events involving cell growth and cell division that produces two new daughter cells. 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.

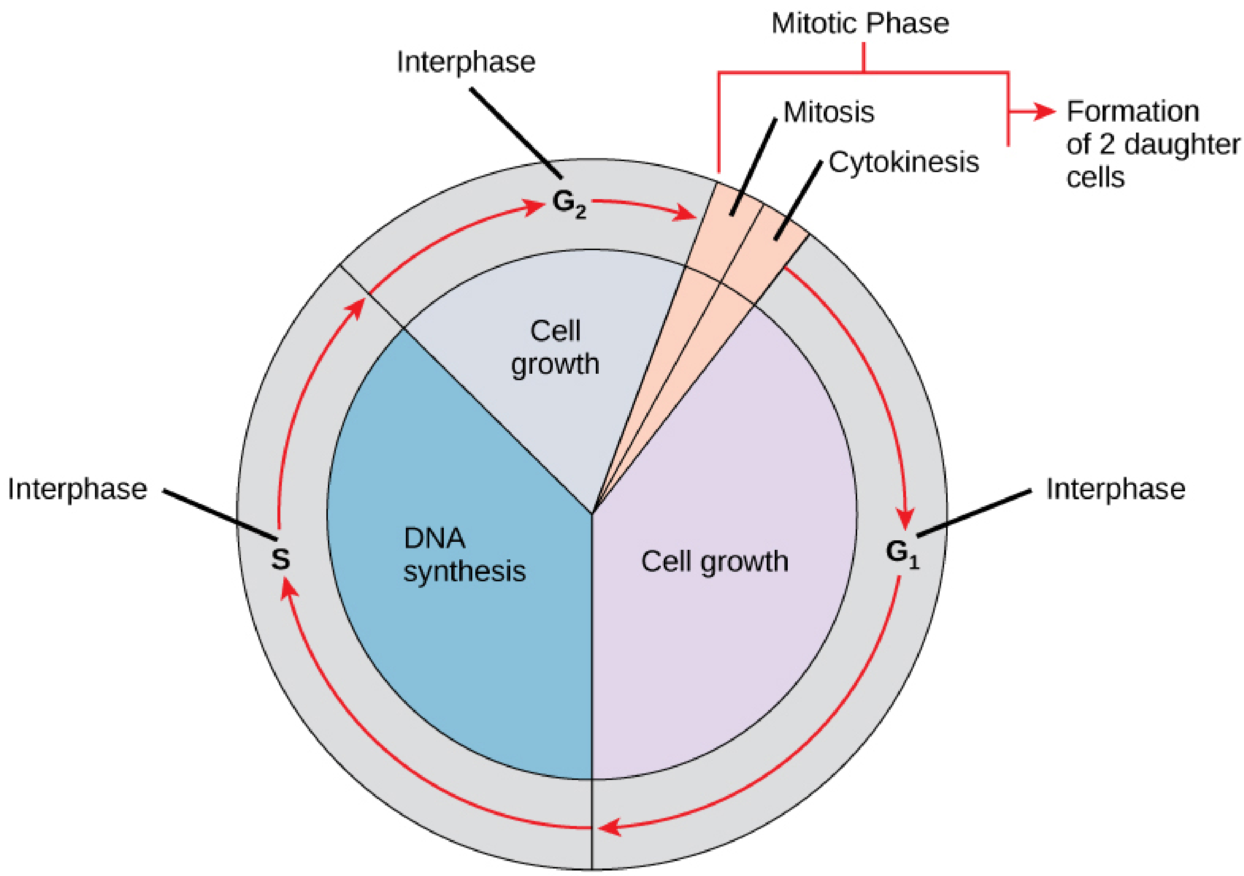


Figure 1. *A cell moves through a series of phases in an orderly manner. During interphase, G1 involves cell growth and protein synthesis, the S phase involves DNA replication and the replication of the centrosome, and G2 involves further growth and protein synthesis. The mitotic phase follows interphase. 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

The first stage of interphase is called the **G_1 phase**, or first gap, because 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 enough energy reserves to complete the task of replicating each chromosome in the nucleus.

S Phase

Throughout interphase, nuclear DNA remains in a semi-condensed chromatin configuration. In the **S phase** (synthesis phase), DNA replication results in the formation of two identical copies of each chromosome—sister chromatids—that are firmly attached at the centromere region. At this stage, each chromosome is made of two sister chromatids and is a duplicated chromosome. The centrosome is 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. The centrosome consists of a pair of rod-like **centrioles** at right angles to each other. Centrioles help organize cell division. Centrioles are not present in the centrosomes of many eukaryotic species, such as plants and most fungi.

G_2 Phase

In the **G_2 phase**, or second gap, the cell replenishes its energy stores and synthesizes the proteins necessary for chromosome manipulation. Some cell organelles are duplicated, and the cytoskeleton is dismantled to provide resources for the mitotic spindle. 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

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. The second portion of the mitotic phase, called cytokinesis, is the physical separation of the cytoplasmic components into two daughter cells.

Mitosis

Mitosis is divided into a series of phases—prophase, prometaphase, metaphase, anaphase, and telophase—that result in the division of the cell nucleus (***Figure 2***). The processes that occur during each of these phases, and microscope images of cells in each phase are shown below for your reference, but **YOU DO NOT NEED TO MEMORIZE THEM FOR THE EXAM.**

ART CONNECTION

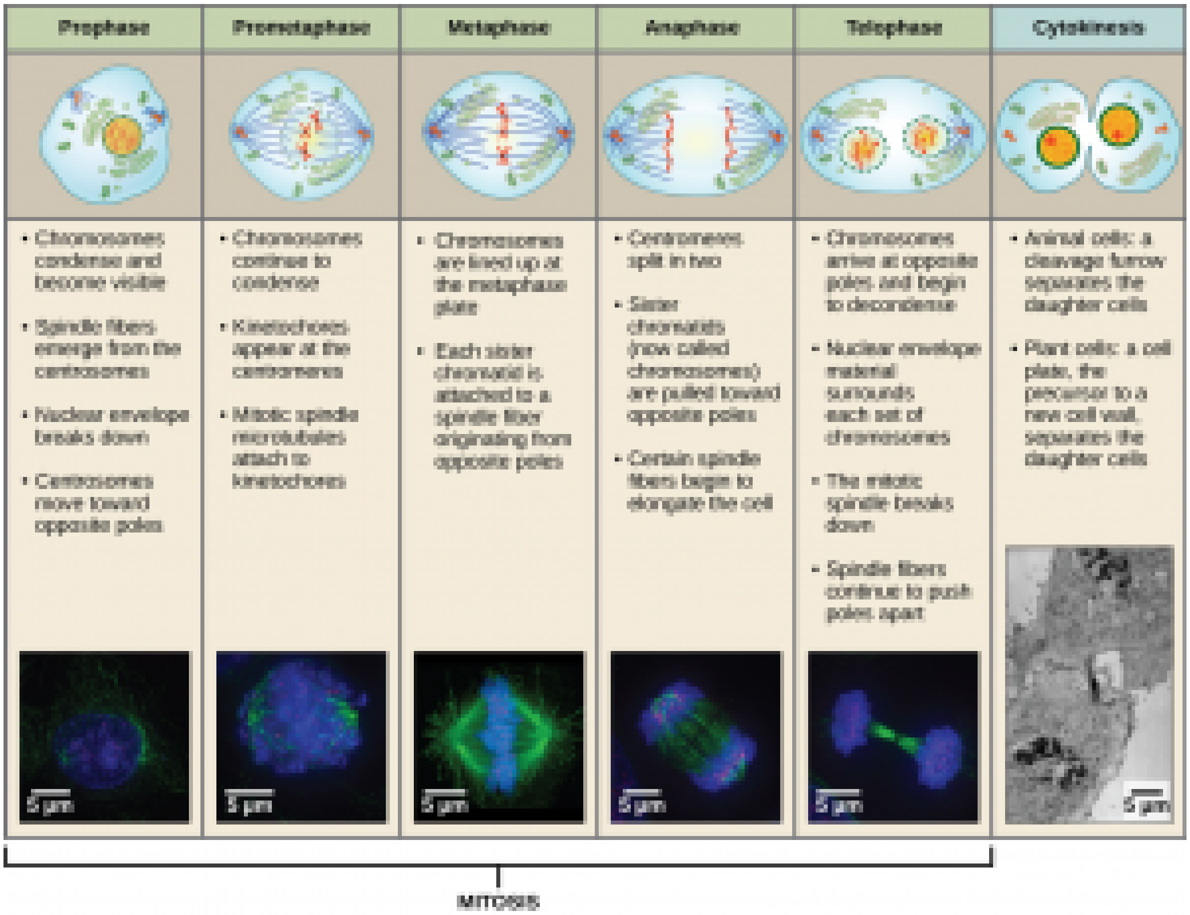


Figure 2. *Animal cell mitosis is divided into five stages—prophase, prometaphase, metaphase, anaphase, and telophase—visualized here by light microscopy with fluorescence. Mitosis is usually accompanied by cytokinesis, shown here by a transmission electron microscope. (credit "diagrams": modification of work by Mariana Ruiz Villareal; credit "mitosis micrographs": modification of work by Roy van Heesbeen; credit "cytokinesis micrograph": modification of work by the Wadsworth Center, NY State Department of Health; donated to the Wikimedia foundation; scale-bar data from Matt Russell)*

G_0 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 **G_0 phase** are not actively preparing to divide. The cell is in a quiescent (inactive) stage, having exited the cell cycle. Some cells enter G_0 temporarily until an external signal triggers the onset of G_1. Other cells that never or rarely divide, such as mature cardiac muscle and nerve cells, remain in G_0 permanently (***Figure 3***).

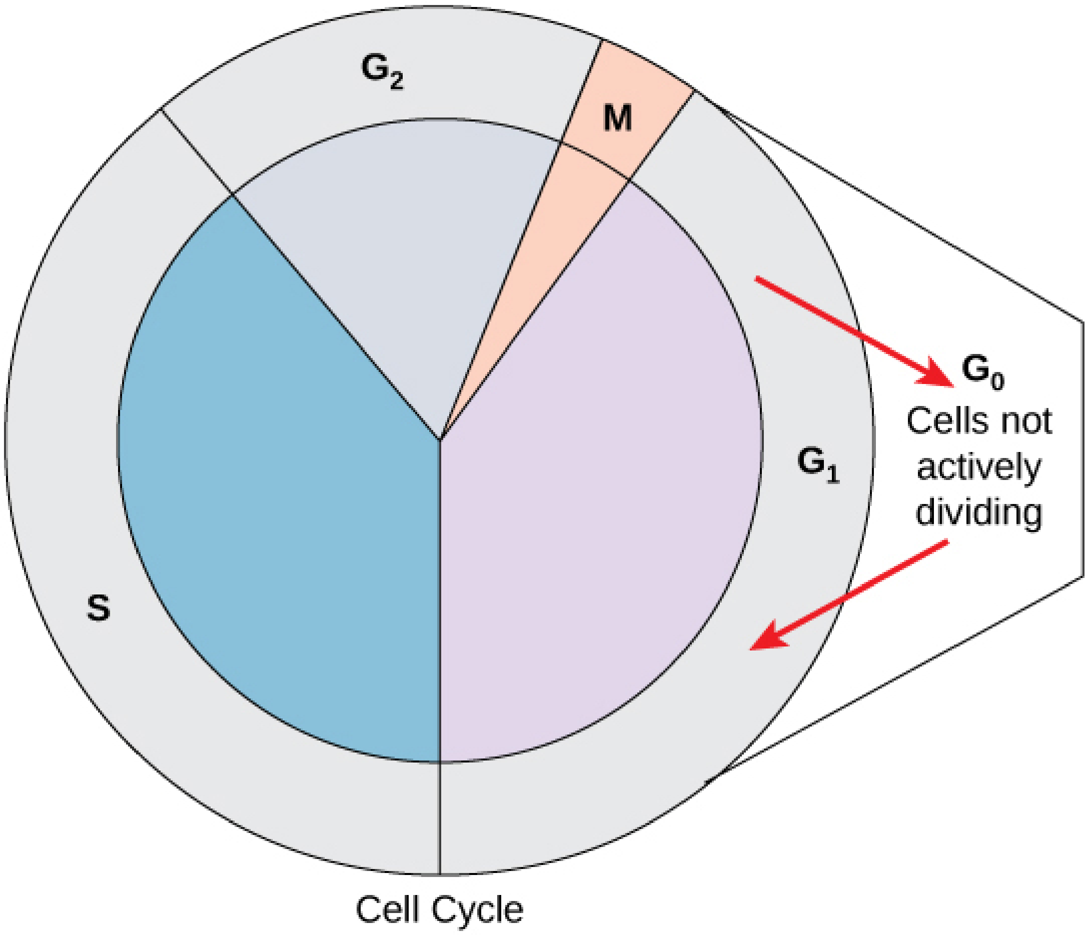


Figure 3. Cells that are not actively preparing to divide enter an alternate phase called G0. In some cases, this is a temporary condition until triggered to enter G1. In other cases, the cell will remain in G0 permanently.

Control of the Cell Cycle

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 in G_0 by 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. In rapidly dividing human cells with a 24-hour cell cycle, the G_1 phase lasts approximately 11 hours. The timing of events in the cell cycle is controlled by mechanisms that are both internal and external to the cell.

Regulation at Internal Checkpoints

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. These checkpoints occur near the end of G_1, at the G_2–M transition, and during metaphase (***Figure 4***).

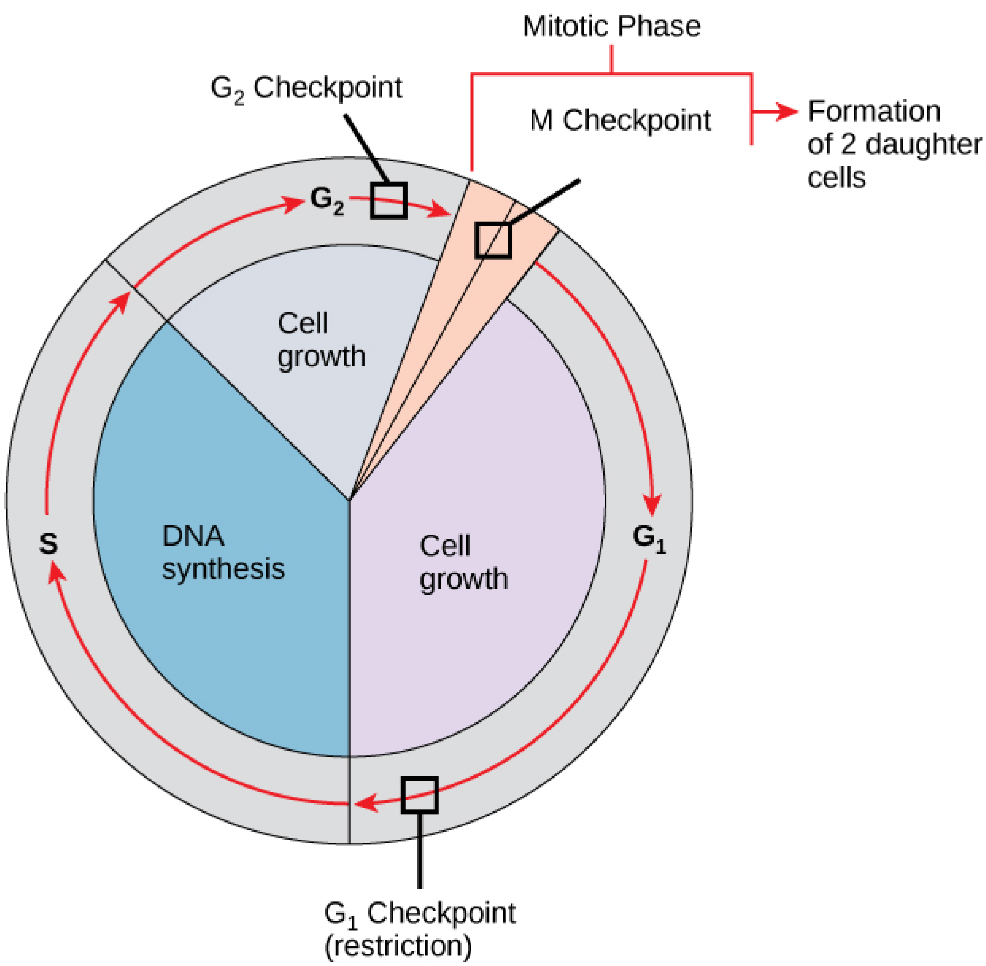


Figure 4. *The cell cycle is controlled at three checkpoints. Integrity of the DNA is assessed at the G1 checkpoint. Proper chromosome duplication is assessed at the G2 checkpoint. Attachment of each kinetochore to a spindle fiber is assessed at the M checkpoint.*

The G_1 Checkpoint

The G_1 checkpoint determines whether all conditions are favorable for cell division to proceed. The G_1 checkpoint, also called the restriction point, 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 at the G_1 checkpoint. A cell that does not meet all the requirements will not be released into the S phase.

The G_2 Checkpoint

The G_2 checkpoint bars the entry to the mitotic phase if certain conditions are not met. As in the G_1 checkpoint, cell size and protein reserves are assessed. However, the most important role of the G_2 checkpoint is to ensure that all of the chromosomes have been replicated and that the replicated DNA is not damaged.

The M Checkpoint

The M checkpoint occurs near the end of the metaphase stage of mitosis. The M checkpoint is also known as the spindle checkpoint because it determines if all the sister chromatids are correctly attached to the spindle microtubules. Because the separation of the sister chromatids during anaphase is an irreversible step, the cycle will not proceed until the kinetochores of each pair of sister chromatids are firmly anchored to spindle fibers arising from opposite poles of the cell.

Concept in Action

Watch what occurs at the G_1, G_2, and M checkpoints by visiting this animation of the cell cycle.

Section Summary

The cell cycle is an orderly sequence of events. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages. In eukaryotes, the cell cycle consists of a long preparatory period, called interphase. Interphase is divided into G_1, S, and G_2 phases. Mitosis and cytokinesis are the steps during which the cell divides into two daughter cells.

Each step of the cell cycle is monitored by internal controls called checkpoints. There are three major checkpoints in the cell cycle: one near the end of G_1, a second at the G_2–M transition, and the third during metaphase.

Glossary

anaphase
: the stage of mitosis during which sister chromatids are separated from each other

cell cycle
: the ordered sequence of events that a cell passes through between one cell division and the next

cell cycle checkpoints
: mechanisms that monitor the preparedness of a eukaryotic cell to advance through the various cell cycle stages

cytokinesis
: the division of the cytoplasm following mitosis to form two daughter cells

G_0 phase
: a cell-cycle phase distinct from the G_1 phase of interphase; a cell in G_0 is not preparing to divide

G_1 phase
: (also, first gap) a cell-cycle phase; first phase of interphase centered on cell growth during mitosis

G_2 phase
: (also, second gap) a cell-cycle phase; third phase of interphase where the cell undergoes the final preparations for mitosis

interphase
: the period of the cell cycle leading up to mitosis; includes G_1, S, and G_2 phases; the interim between two consecutive cell divisions

metaphase
: the stage of mitosis during which chromosomes are lined up at the metaphase plate

mitosis
: the period of the cell cycle at which the duplicated chromosomes are separated into identical nuclei; includes prophase, prometaphase, metaphase, anaphase, and telophase

mitotic phase
: the period of the cell cycle when duplicated chromosomes are distributed into two nuclei and the cytoplasmic contents are divided; includes mitosis and cytokinesis

mitotic spindle
: the microtubule apparatus that orchestrates the movement of chromosomes during mitosis

prophase
: the stage of mitosis during which chromosomes condense and the mitotic spindle begins to form

quiescent
: describes a cell that is performing normal cell functions and has not initiated preparations for cell division

S phase
: the second, or synthesis phase, of interphase during which DNA replication occurs

telophase
: the stage of mitosis during which chromosomes arrive at opposite poles, decondense, and are surrounded by new nuclear envelopes

41.

4.8 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 occur within a coding portion of a gene and are not corrected, a gene mutation results. All cancers start when a gene mutation gives rise to a faulty protein that plays a key role in cell reproduction. The change in the cell that results from the malformed 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 can result. The suffix "-oma" refers to a tumor.

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, the alteration of the DNA sequence will result in a less functional (or non-functional) protein. The result is detrimental to the cell and will likely prevent the cell from completing the cell cycle; however, the organism is not harmed because the mutation will not be carried forward. If a cell cannot reproduce, the mutation is not propagated 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, the type of regulators that, when activated, can 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 1***). 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.

ART CONNECTION

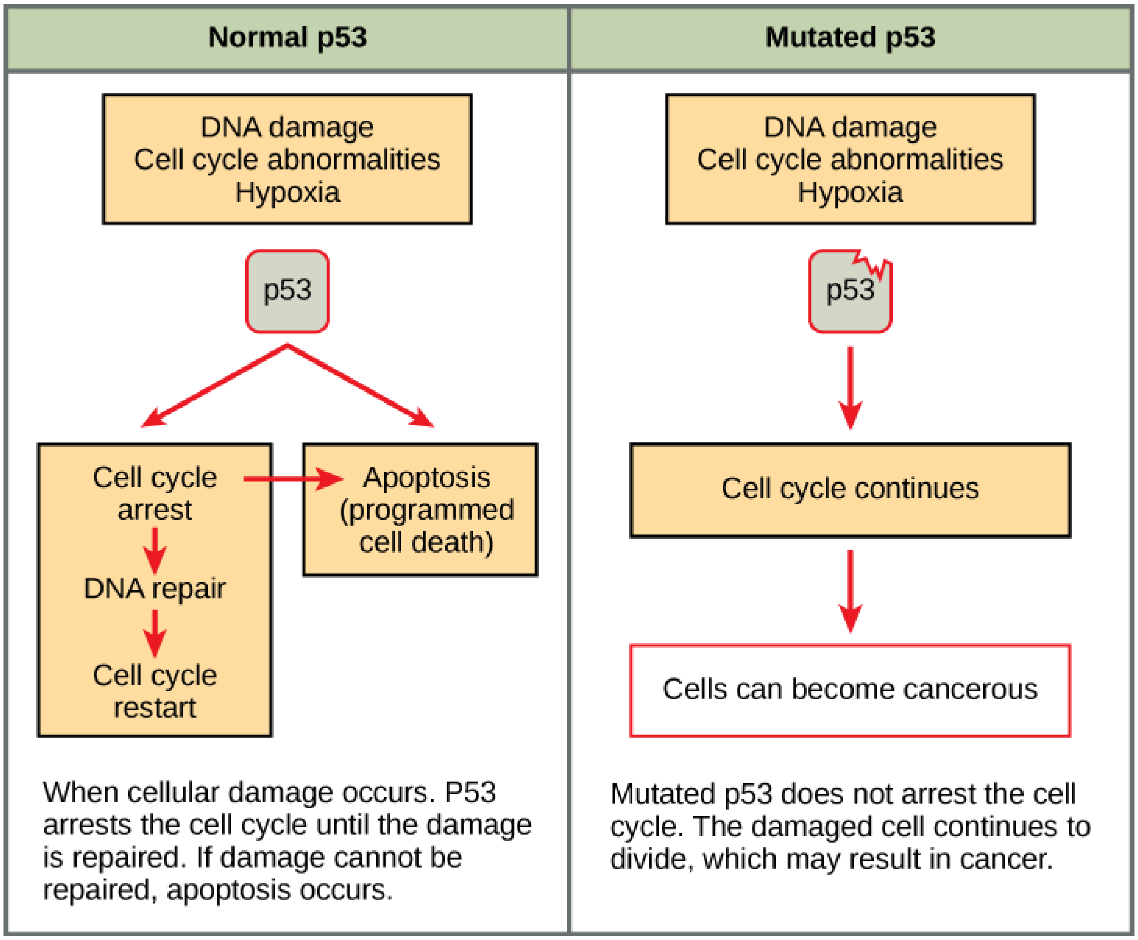


Figure 1. *The role of normal p53 is to monitor DNA and the supply of oxygen (hypoxia is a condition of reduced oxygen supply). If damage is detected, p53 triggers repair mechanisms. If repairs are unsuccessful, p53 signals apoptosis. A cell with an abnormal p53 protein cannot repair damaged DNA and thus cannot signal apoptosis. Cells with abnormal p53 can become cancerous. (credit: modification of work by Thierry Soussi)*

The loss of p53 function has other repercussions for the cell cycle. Mutated p53 might lose its ability to trigger p21 production. Without adequate levels of p21, there is no effective block on Cdk activation. Essentially, without a fully functional p53, the G_1checkpoint is severely compromised and the cell proceeds directly from G_1 to S regardless of internal and external conditions. At the completion of this shortened cell cycle, two daughter cells are produced that have inherited the mutated p53 gene. Given the non-optimal conditions under which the parent cell reproduced, it is likely that the daughter cells will have acquired other mutations in addition to the faulty tumor suppressor gene. Cells such as these daughter cells quickly accumulate both oncogenes and non-functional tumor suppressor genes. Again, the result is tumor growth.

Link to Learning

Watch an animation of how cancer results from errors in the cell cycle. https://www.youtube.com/watch?time_continue=4&v=RZhL7LDPk8w

Cancer Progression

Cancers are often named based on the organ or tissue which originally accumulated mutations in several proto-oncogenes and tumor suppressor genes, and was therefore the site of the cancer origin. For example, a sarcoma is a cancer that developed in a muscle cell; the prefix "sarc-" means muscle and the suffix "-oma" means tumor. Cancer development usually occurs through a typical series of steps, otherwise known as a cancer progression or as cancer stages. The clinical diagnostic criteria for which stage a cancer is in varies by the type of cancer, but they are generally listed as Stage 0 through Stage 4, with Stage 0 being the least progressed and likely the easiest to treat for any given cancer.

In general, most lower stage cancers are marked by tumors that contain rapidly dividing abnormal cells that remain in one discreet location. As cancers progress and additional mutations accumulate during the rapid rounds of cell division coupled with unsuccessful cell cycle checkpoints, tumors can become malignant. A **malignant tumor** has cells that start to migrate and invade neighboring tissues, or enter the blood or lymph. If these tumor cells that circulate in the blood leave the circulation at a different location and enter the tissues there, they may form a secondary tumor, called a metastasis. **Metastatic cancer**, or cancer with tumors that have developed at sites distant from the original location, are usually late stage cancers that become increasingly difficult to treat.

Section Summary

Cancer is the result of unchecked cell division caused by a breakdown of the mechanisms that regulate the cell cycle. The loss of control begins with a change in the DNA sequence of a gene that codes for one of the regulatory molecules. Faulty instructions lead to a protein that does not function as it should. Any disruption of the monitoring system can allow other mistakes to be passed on to the daughter cells. Each successive cell division will give rise to daughter cells with even more accumulated damage. Eventually, all checkpoints become nonfunctional, and rapidly reproducing cells crowd out normal cells, resulting in a tumor or leukemia (blood cancer). Cancers then progress, becoming malignant and sometimes developing metastases.

Glossary

oncogene
: mutated version of a normal gene involved in the positive regulation of the cell cycle

proto-oncogene
: normal gene that when mutated becomes an oncogene

tumor suppressor gene
segment of DNA that codes for regulator proteins that prevent the cell from undergoing uncontrolled division

V

Resources

42.

Geological Time

Geological Time

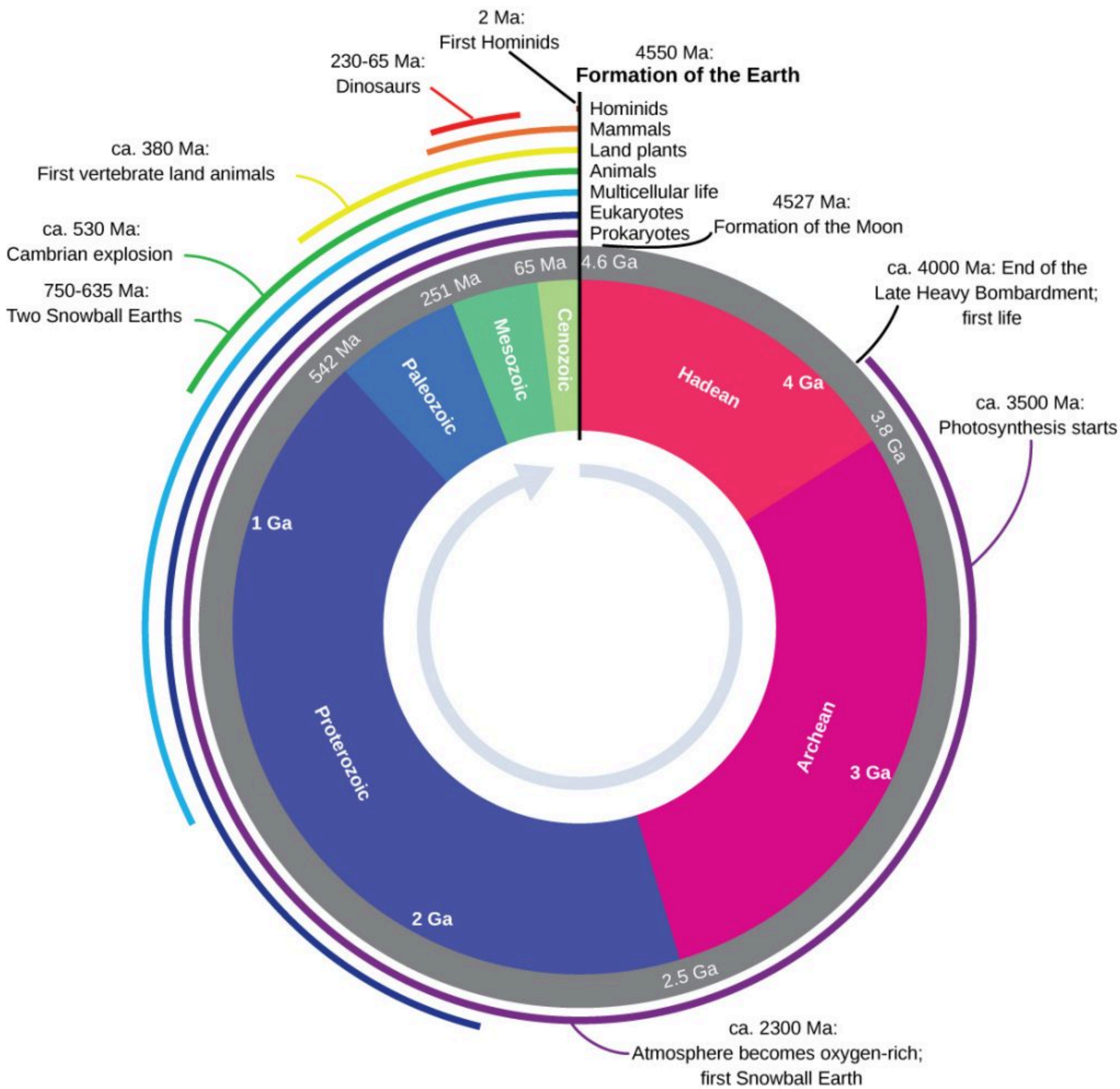


Figure 1. *Geological Time Clock*

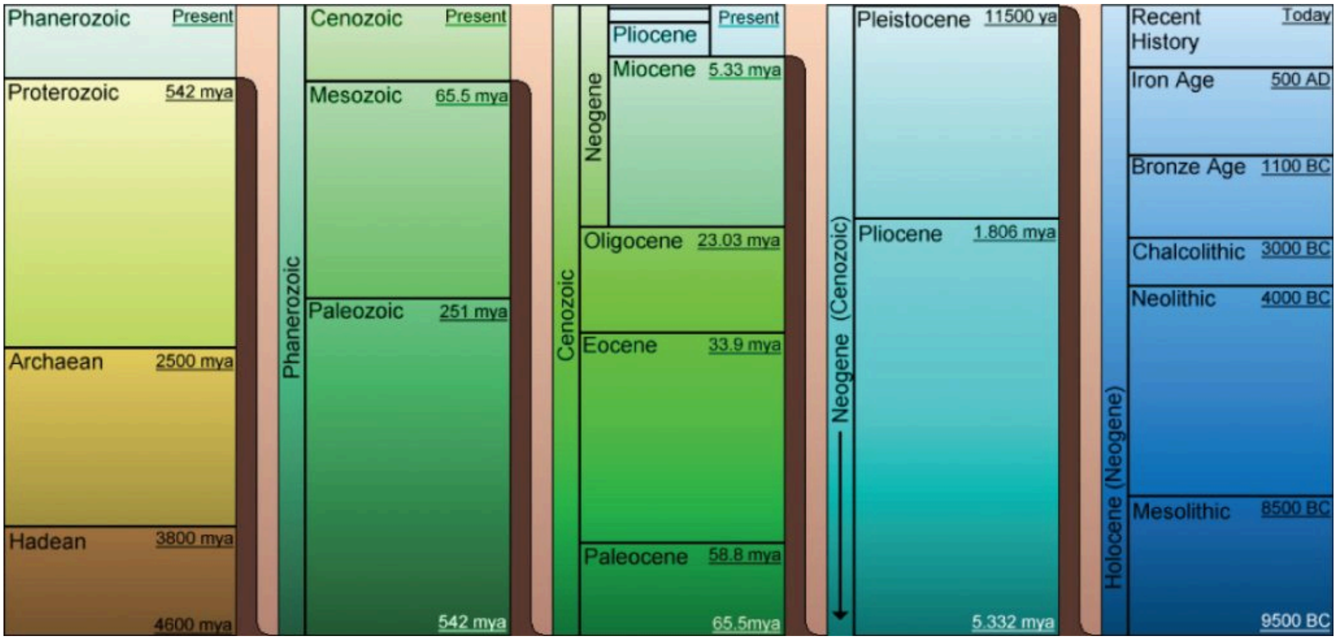


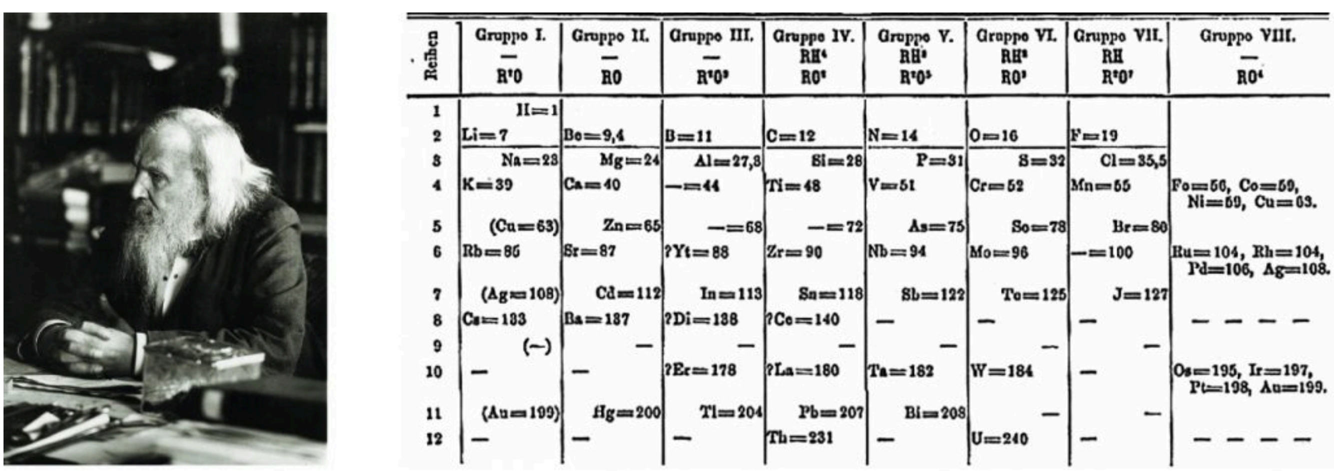
Figure 2. *Geological Time Chart (credit: Richard S. Murphy, Jr.)*

43.

The Periodic Table

As early chemists worked to purify ores and discovered more elements, they realized that various elements could be grouped together by their similar chemical behaviors. One such grouping includes lithium (Li), sodium (Na), and potassium (K): These elements all are shiny, conduct heat and electricity well, and have similar chemical properties. A second grouping includes calcium (Ca), strontium (Sr), and barium (Ba), which also are shiny, good conductors of heat and electricity, and have chemical properties in common. However, the specific properties of these two groupings are notably different from each other. For example: Li, Na, and K are much more reactive than are Ca, Sr, and Ba; Li, Na, and K form compounds with oxygen in a ratio of two of their atoms to one oxygen atom, whereas Ca, Sr, and Ba form compounds with one of their atoms to one oxygen atom. Fluorine (F), chlorine (Cl), bromine (Br), and iodine (I) also exhibit similar properties to each other, but these properties are drastically different from those of any of the elements above.

Dimitri **Mendeleev** in Russia (1869) and Lothar **Meyer** in Germany (1870) independently recognized that there was a periodic relationship among the properties of the elements known at that time. Both published tables with the elements arranged according to increasing atomic mass. But Mendeleev went one step further than Meyer: He used his table to predict the existence of elements that would have the properties similar to aluminum and silicon, but were yet unknown. The discoveries of gallium (1875) and germanium (1886) provided great support for Mendeleev's work. Although Mendeleev and Meyer had a long dispute over priority, Mendeleev's contributions to the development of the periodic table are now more widely recognized (***Figure 1***).



Reihen	Gruppo I. — R^2O	Gruppo II. — RO	Gruppo III. — R^2O^3	Gruppo IV. RH^4 RO^2	Gruppo V. RH^3 R^2O^5	Gruppo VI. RH^2 RO^3	Gruppo VII. RH R^2O^7	Gruppo VIII. — RO^4
1	H=1							
2	Li=7	Be=9,4	B=11	C=12	N=14	O=16	F=19	
3	Na=23	Mg=24	Al=27,3	Si=28	P=31	S=32	Cl=35,5	
4	K=39	Ca=40	—=44	Ti=48	V=51	Cr=52	Mn=55	Fe=56, Co=59, Ni=59, Cu=63.
5	(Cu=63)	Zn=65	—=68	—=72	As=75	Se=78	Br=80	
6	Rb=85	Sr=87	?Yt=88	Zr=90	Nb=94	Mo=96	—=100	Ru=104, Rh=104, Pd=106, Ag=108.
7	(Ag=108)	Cd=112	In=113	Sn=118	Sb=122	Te=125	J=127	
8	Cs=133	Ba=137	?Di=138	?Ce=140	—	—	—	— — — —
9	(—)	—	—	—	—	—	—	
10	—	—	?Er=178	?La=180	Ta=182	W=184	—	Os=195, Ir=197, Pt=198, Au=199.
11	(Au=199)	Hg=200	Tl=204	Pb=207	Bi=208	—	—	
12	—	—	—	Th=231	—	U=240	—	— — — —

(a) (b)

Figure 1. ***(a)*** *Dimitri Mendeleev is widely credited with creating* ***(b)*** *the first periodic table of the elements. (credit* ***a****: modification of work by Serge Lachinov; credit* ***b****: modification of work by "Den fjättrade ankan"/Wikimedia Commons)*

By the twentieth century, it became apparent that the periodic relationship involved atomic numbers rather than atomic masses. The modern statement of this relationship, the **periodic law**, is as follows: *the properties of the elements are periodic functions of their atomic numbers*. A modern **periodic table** arranges the elements in increasing order of their atomic numbers and groups atoms with similar properties in the same vertical column (***Figure 2***). Each box represents an element and contains its atomic number, symbol, average atomic mass, and (sometimes) name. The elements are arranged in seven horizontal rows, called **periods** or **series**, and 18 vertical columns, called **groups**. Groups are labeled at the top of each column. In the United States, the labels traditionally were numerals with capital letters. However, IUPAC recommends that the numbers 1 through 18 be used, and these labels are more common. For the table to fit on a single page, parts of two of the rows, a total of 14 columns, are usually written below the main body of the table.

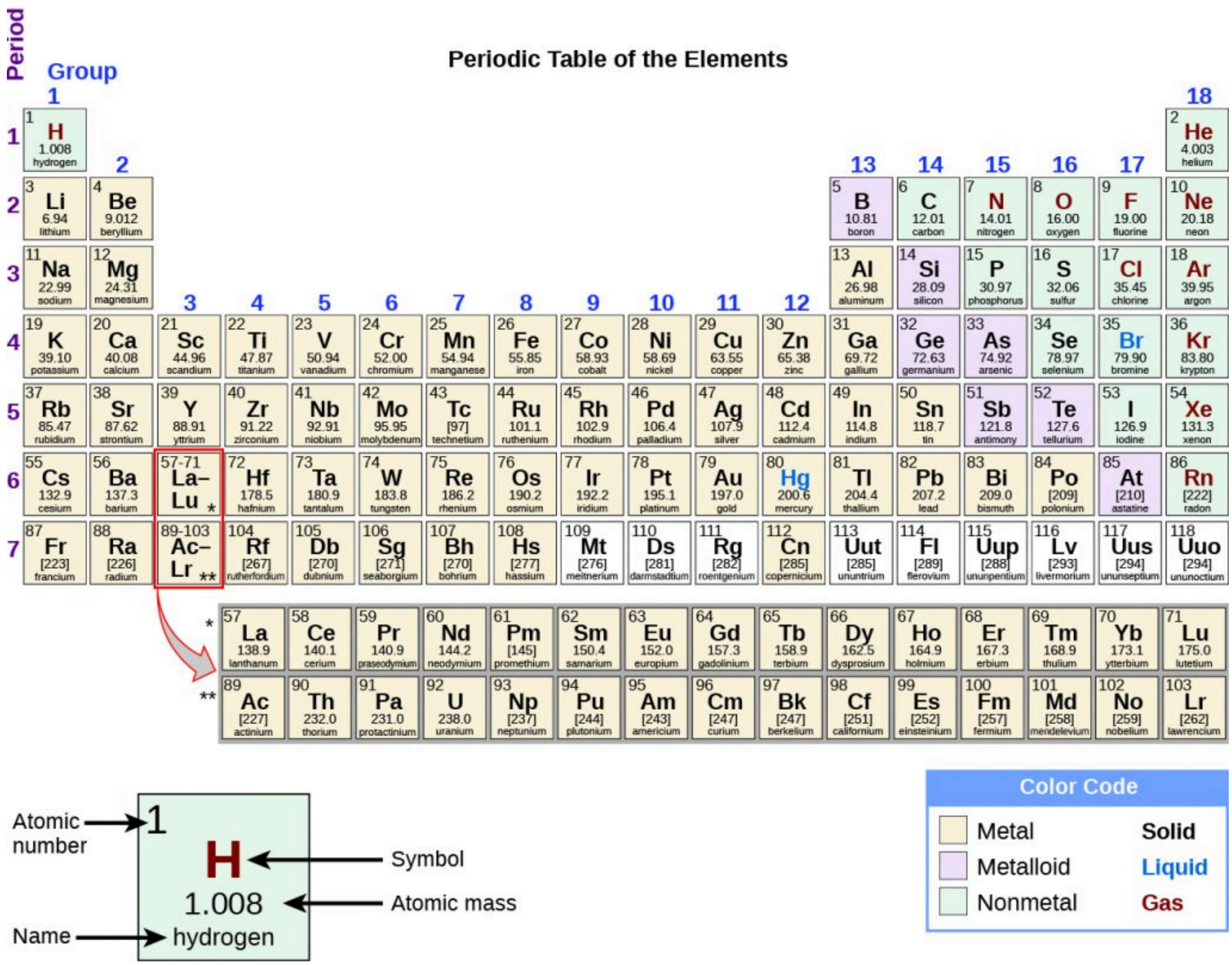


Figure 2. *Elements in the periodic table are organized according to their properties.*

Many elements differ dramatically in their chemical and physical properties, but some elements are similar in their behaviors. For example, many elements appear shiny, are malleable (able to be deformed without breaking) and ductile (can be drawn into wires), and conduct heat and electricity well. Other elements are not shiny, malleable, or ductile, and are poor conductors of heat and electricity. We can sort the elements into large classes with common properties: **metals** (elements that are shiny, malleable, good conductors of heat and electricity—shaded yellow); **nonmetals** (elements that appear dull, poor conductors of heat and electricity—shaded green); and **metalloids** (elements that conduct heat and electricity moderately well, and possess some properties of metals and some properties of nonmetals—shaded purple).

The elements can also be classified into the **main-group elements** (or **representative elements**) in the columns labeled 1, 2, and 13–18; the **transition metals** in the columns labeled 3–12; and **inner transition metals** in the two rows at the bottom of the table (the top-row elements are called **lanthanides** and the bottom-row elements are **actinides**; ***Figure 3***). The elements can be subdivided further by more specific properties, such as the composition of the compounds they form. For example, the elements in group 1 (the first column) form compounds that consist of one atom of the element and one atom of hydrogen. These elements (except hydrogen) are known as **alkali metals**, and they all have similar chemical properties. The elements in group 2 (the second column) form compounds consisting of one atom of the element and two atoms of hydrogen: These are called **alkaline earth metals**, with similar properties among members of that group. Other groups with specific names are the **pnictogens** (group 15), **chalcogens** (group 16), **halogens** (group 17), and the **noble gases** (group 18, also known as **inert gases**). The groups can also be referred to by the first element of the group: For example, the chalcogens can be called the oxygen group or oxygen family. Hydrogen is a unique, nonmetallic element with properties similar to both group 1 and group 17 elements. For that reason, hydrogen may be shown at the top of both groups, or by itself.

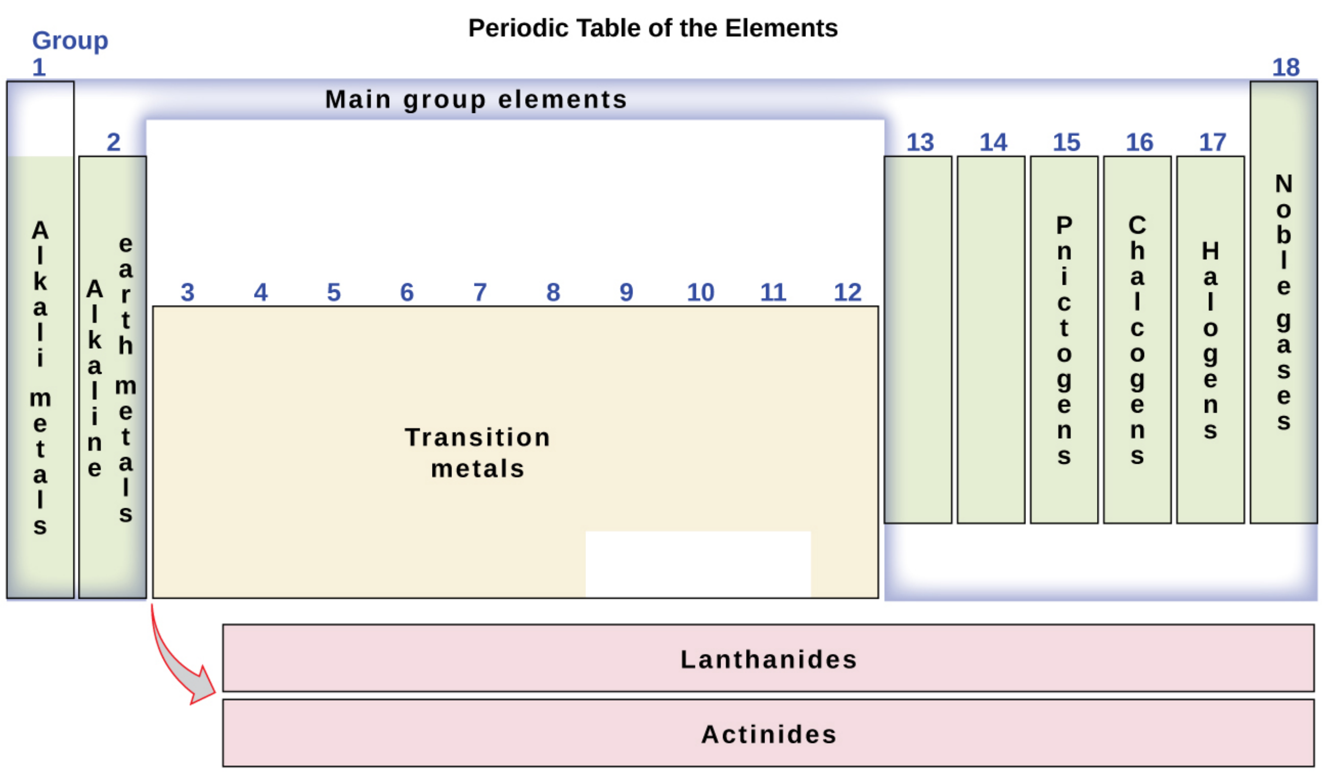


Figure 3. *The periodic table organizes elements with similar properties into groups.*

Link to Learning

Click on this link for an interactive periodic table, which you can use to explore the properties of the elements (includes podcasts and videos of each element). You may also want to refer to ***Figure 1*** above that shows photos of all the elements.

Naming Groups of Elements

Atoms of each of the following elements are essential for life.

1. **Give the group name for the following elements:**

   (a) chlorine

   (b) calcium

   (c) sodium

   (d) sulfur

   ANSWER

2. **Give the group name for each of the following elements:**

   (a) krypton

   (b) selenium

   (c) barium

   (d) lithium

   ANSWER

In studying the periodic table, you might have noticed something about the atomic masses of some of the elements. Element 43 (technetium), element 61 (promethium), and most of the elements with atomic number 84 (polonium) and higher have their atomic mass given in square brackets. This is done for elements that consist entirely of unstable, radioactive isotopes (you will learn more about radioactivity in the nuclear chemistry chapter). An average atomic weight cannot be determined for these elements because their radioisotopes may vary significantly in relative abundance, depending on the source, or may not even exist in nature. The number in square brackets is the atomic mass number (an approximate atomic mass) of the most stable isotope of that element.

Section Summary

The discovery of the periodic recurrence of similar properties among the elements led to the formulation of the periodic table, in which the elements are arranged in order of increasing atomic number in rows known as periods and columns known as groups. Elements in the same group of the periodic table have similar chemical properties. Elements can be classified as metals, metalloids, and nonmetals, or as a main-group elements, transition metals, and inner transition metals. Groups are numbered 1–18 from left to right. The elements in group 1 are known as the alkali metals; those in group 2 are the alkaline earth metals; those in 15 are the pnictogens; those in 16 are the chalcogens; those in 17 are the halogens; and those in 18 are the noble gases.

Glossary

actinide
: inner transition metal in the bottom of the bottom two rows of the periodic table

alkali metal
: element in group 1

alkaline earth metal
: element in group 2

chalcogen
: element in group 16

group
: vertical column of the periodic table

halogen
: element in group 17

inert gas
: (also, noble gas) element in group 18

inner transition metal
: (also, lanthanide or actinide) element in the bottom two rows; if in the first row, also called lanthanide, or if in the second row, also called actinide

lanthanide
: inner transition metal in the top of the bottom two rows of the periodic table

main-group element
: (also, representative element) element in columns 1, 2, and 12–18

metal
: element that is shiny, malleable, good conductor of heat and electricity

metalloid
: element that conducts heat and electricity moderately well, and possesses some properties of

metals and some properties of nonmetals

noble gas
(also, inert gas) element in group 18

nonmetal
element that appears dull, poor conductor of heat and electricity

period
(also, series) horizontal row of the periodic table

periodic law
properties of the elements are periodic function of their atomic numbers.

periodic table
table of the elements that places elements with similar chemical properties close together

pnictogen
element in group 15

representative element
(also, main-group element) element in columns 1, 2, and 12–18

series
(also, period) horizontal row of the period table

transition metal
element in columns 3–11

44.

Measurements and the Metric System

Table 1. Measurements and the Metric System

Measurement	Unit	Abbreviation	Metric Equivalent	Approximate Standard Equivalent
Length	nanometer	nm	1 nm = 10^{-9} m	
	micrometer	µm	1 µm = 10^{-6} m	
	millimeter	mm	1 mm = 0.001 m	1 mm = 0.039 inch
	centimeter	cm	1 cm = 0.01 m	1 cm = 0.394 inch
	meter	m	1 m = 100 cm = 1000 mm	1 m = 39.37 inches = 3.28 feet = 1.093 yards
	kilometer	km	1 km = 1000 m	1 km = 0.621 miles
Mass	microgram	µg	1 µg = 10^{-6}g	
	milligram	mg	1 mg = 10^{-3} g	
	gram	g	1 g = 1000 mg	1 g = 0.035 ounce
	kilogram	kg	1 kg = 1000 g	1 kg = 2.205 pounds
Volume	microliter	µl	1 µl = 10^{-6}l	
	milliliter	ml	1 ml = 10^{-3}l	1 ml = 0.034 fluid ounce
	liter	l	1 l = 1000 ml	1 l = 1.057 quarts
	kiloliter	kl	1 kl = 1000 l	1 kl = 264.172 gallons
Area	square centimeter	cm^2	1 cm^2 = 100 mm^2	1 cm^2 = 0.155 square inch
	square meter	m^2	1 m^2 = 10,000 cm^2	1 m^2 = 10.764 square feet = 1.196 square yards
	hectare	ha	1 ha = 10,000 m^2	1 ha = 2.471 acres
Temperature	Celsius	°C	—	1 °C = 5/9 × (°F − 32)

45.

The Process of Science

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.

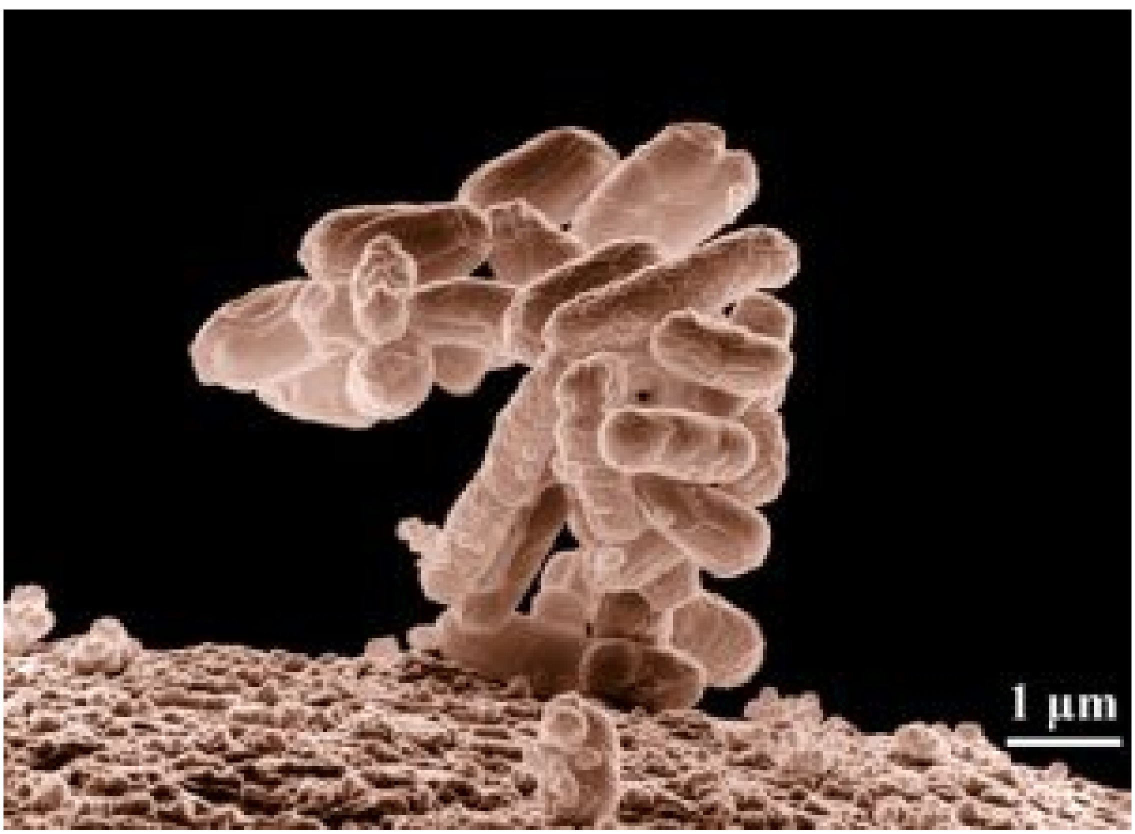


Figure 1. *Biologists may choose to study Escherichia coli (E. coli), a bacterium that is a normal resident of our digestive tracts but which is also sometimes responsible for disease outbreaks. In this micrograph, the bacterium is visualized using a scanning electron microscope and digital colorization. (credit: Eric Erbe; digital colorization by Christopher Pooley, USDA-ARS)*

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 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.

The **scientific method** is a method of research with defined steps that include experiments and careful observation. The steps of the scientific method will be examined in detail later, but 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.

Natural Sciences

What would you expect to see in a museum of natural sciences? Frogs? Plants? Dinosaur skeletons? Exhibits about how the brain functions? A planetarium? Gems and minerals? Or maybe all of the above? Science includes such diverse fields as astronomy, biology, computer sciences, geology, logic, physics, chemistry, and mathematics (***Figure 2***). However, those fields of science related to the physical world and its phenomena and processes are considered **natural sciences**. Thus, a museum of natural sciences might contain any of the items listed above.

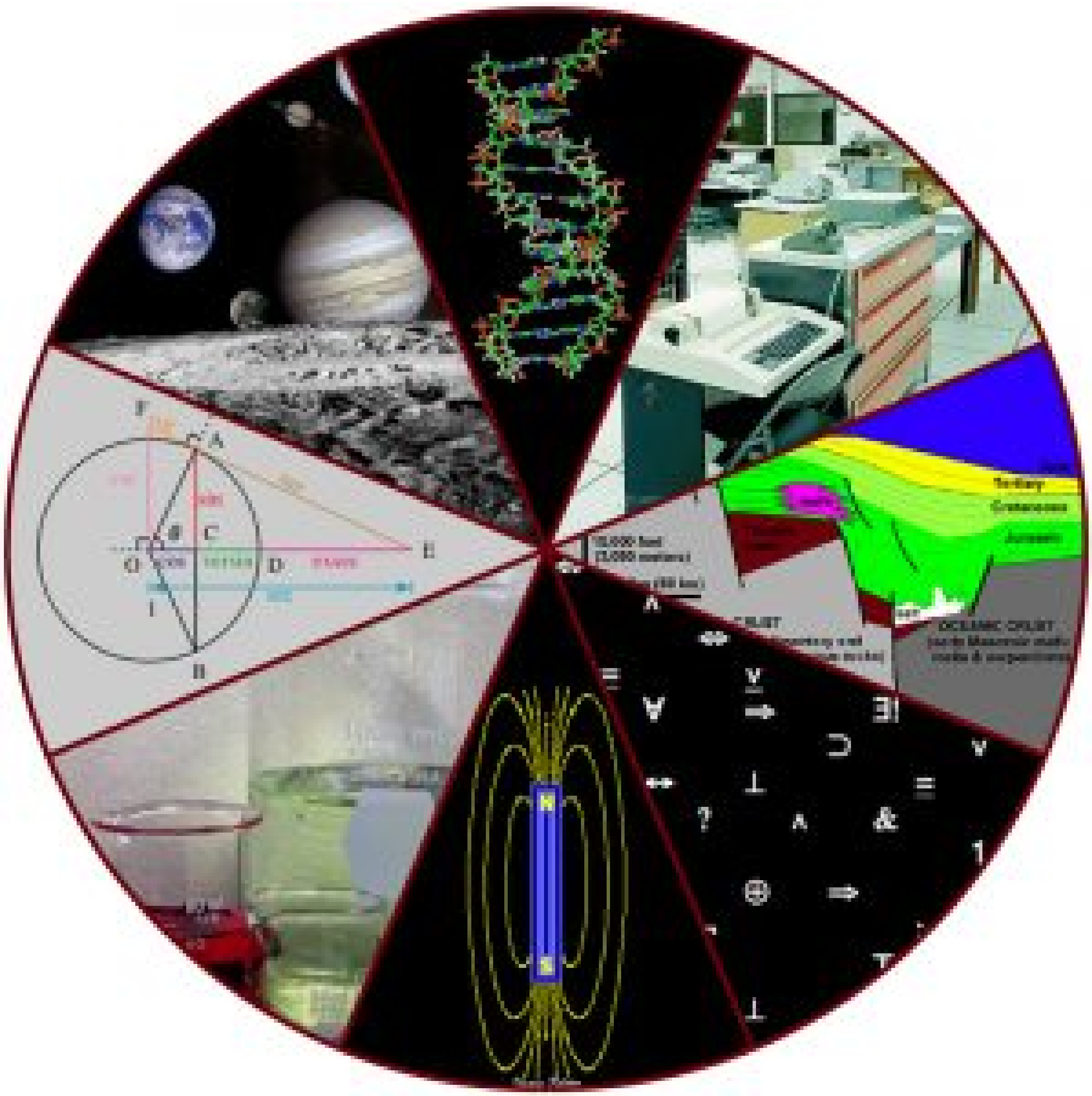


Figure 2. *Some fields of science include astronomy, biology, computer science, geology, logic, physics, chemistry, and mathematics. (credit: "Image Editor"/Flickr)*

There is no complete agreement when it comes to defining what the natural sciences include. For some experts, the natural sciences are astronomy, biology, chemistry, earth science, and physics. Other scholars choose to divide natural sciences into **life sciences**, which study living things and include biology, and **physical sciences**, which study nonliving matter and include astronomy, physics, and chemistry. Some disciplines such as biophysics and biochemistry build on two sciences and are interdisciplinary.

Scientific Inquiry

One thing is common to all forms of science: an ultimate goal "to know." Curiosity and inquiry are the driving forces for the development of science. Scientists seek to understand the world and the way it operates. Two methods of logical thinking are used: inductive reasoning and deductive reasoning.

Inductive reasoning is a form of logical thinking that uses related observations to arrive at a general conclusion. This type of reasoning is common in descriptive science. A life scientist such as a biologist makes observations and records them. These data can be qualitative (descriptive) or quantitative (consisting of numbers), and the raw data can be supplemented with drawings, pictures, photos, or videos. From many observations, the scientist can infer conclusions (inductions) based on evidence. Inductive reasoning involves formulating generalizations inferred from careful observation and the analysis of a large amount of data. Brain studies often work this way. Many brains are observed while people are doing a task. The part of the brain that lights up, indicating activity, is then demonstrated to be the part controlling the response to that task.

Deductive reasoning or deduction is the type of logic used in hypothesis-based science. In deductive reasoning, the pattern of thinking moves in the opposite direction as compared to inductive reasoning. **Deductive reasoning** is a form of logical thinking that uses a general principle or law to forecast specific results. From those general principles, a scientist can extrapolate and predict the specific results that would be valid as long as the general principles are valid. For example, a prediction would be that if the climate is becoming warmer in a region, the distribution of plants and animals should change. Comparisons have been made between distributions in the past and the present, and the many changes that have been found are consistent with a warming climate. Finding the change in distribution is evidence that the climate change conclusion is a valid one.

Both types of logical thinking are related to the two main pathways of scientific study: descriptive science and hypothesis-based science. **Descriptive** (or discovery) **science** aims to observe, explore, and discover, while **hypothesis-based science** begins with a specific question or problem and a potential answer or solution that can be tested. The boundary between these two forms of study is often blurred, because most scientific endeavors combine both approaches. Observations lead to questions, questions lead to forming a hypothesis as a possible answer to those questions, and then the hypothesis is tested. Thus, descriptive science and hypothesis-based science are in continuous dialogue.

Hypothesis Testing

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 method was used even in ancient times, but it was first documented by England's Sir Francis Bacon (1561–1626) (***Figure 3***), 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 3. *Sir Francis Bacon is credited with being the first to document the scientific method.*

The scientific process typically starts with an observation (often a problem to be solved) that leads to a question. Let's think about a simple problem that starts with an observation and apply the scientific method to solve the problem. One Monday morning, a student arrives at class and quickly discovers that the classroom is too warm. That is an observation that also describes a problem: the classroom is too warm. The student then asks a question: "Why is the classroom so warm?"

Recall that a hypothesis is a suggested explanation that can be tested. To solve a problem, several hypotheses may be proposed. For example, one hypothesis might be, "The classroom is warm because no one turned on the air conditioning." But there could be other responses to the question, and therefore other hypotheses may be proposed. A second hypothesis might be, "The classroom is warm because there is a power failure, and so the air conditioning doesn't work."

Once a hypothesis has been selected, a prediction may be made. A prediction is similar to a hypothesis but it typically has the format "If . . . then . . . ." For example, the prediction for the first hypothesis might be, "*If* the student turns on the air conditioning, *then* the classroom will no longer be too warm."

A hypothesis must be testable to ensure that it is valid. For example, a hypothesis that depends on what a bear thinks is not testable, because it can never be known what a bear thinks. It should also be **falsifiable**, meaning that it can be disproven by experimental results. An example of an unfalsifiable hypothesis is "Botticelli's *Birth of Venus* is beautiful." 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 we find support for that explanation, but this is not to say that down the road a better explanation will not be found, or a more carefully designed experiment will be found to falsify the hypothesis.

Each experiment will have one or more variables and one or more controls. A **variable** is any part of the experiment that can vary or change during the experiment. A **control** is a part of the experiment that does not change. Look for the variables and controls in the example that follows. As a simple example, an experiment might be conducted to test the hypothesis that phosphate limits the growth of algae in freshwater ponds. A series of artificial ponds are filled with water and half of them are treated by adding phosphate each week, while the other half are treated by adding a salt that is known not to be used by algae. The variable here is the phosphate (or lack of phosphate), the experimental or treatment cases are the ponds with added phosphate and the control ponds are those with something inert added, such as the salt. Just adding something is also a control against the possibility that adding extra matter to the pond has an effect. If the treated ponds show lesser growth of algae, then we have found support for our hypothesis. If they do not, then we reject our hypothesis. Be aware that rejecting one hypothesis does not determine whether or not the other hypotheses can be accepted; it simply eliminates one hypothesis that is not valid (***Figure 4***). Using the scientific method, the hypotheses that are inconsistent with experimental data are rejected.

ART CONNECTION

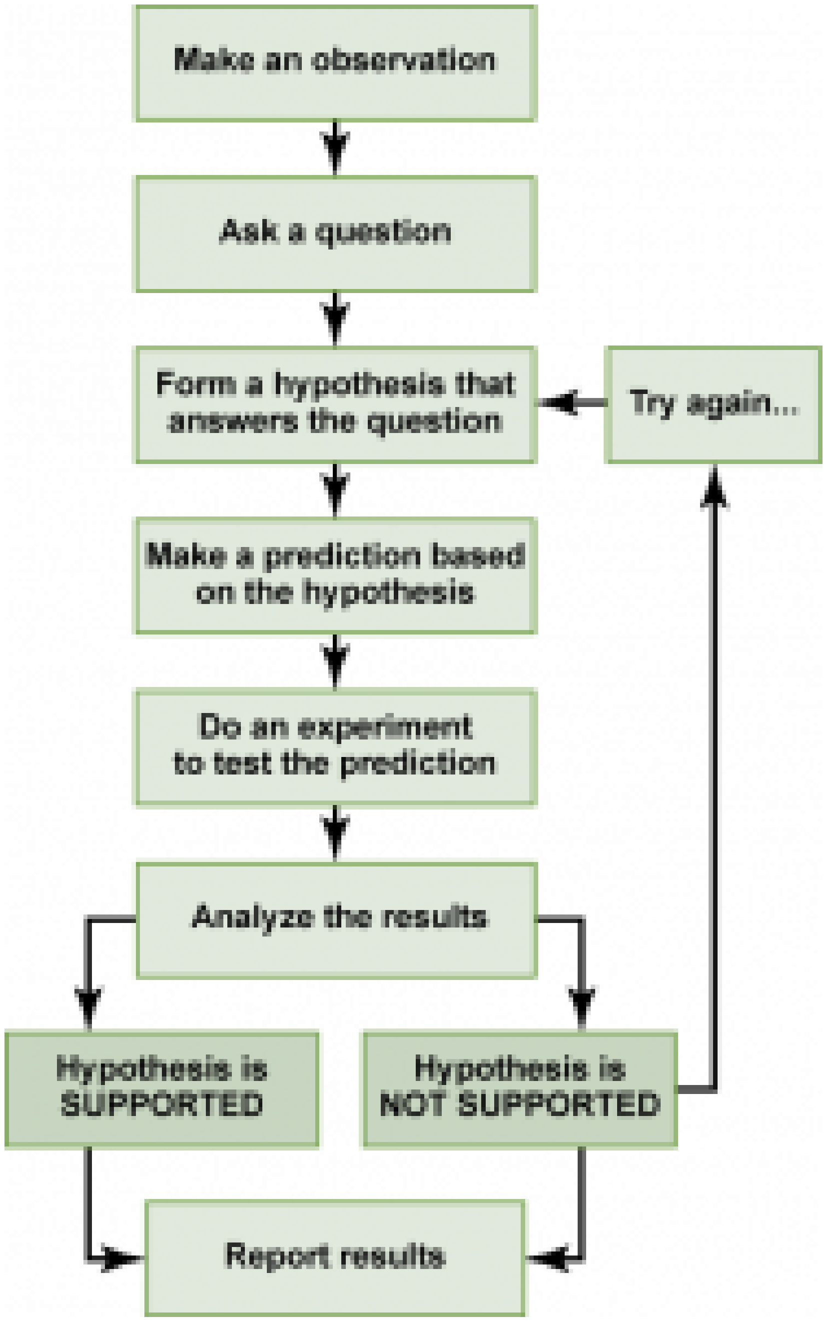


Figure 4. *The scientific method is a series of defined steps that include experiments and careful observation. If a hypothesis is not supported by data, a new hypothesis can be proposed.*

In the example below, the scientific method is used to solve an everyday problem. Which part in the example below is the hypothesis? Which is the prediction? Based on the results of the experiment, is the hypothesis supported? If it is not supported, propose some alternative hypotheses.

1. My toaster doesn’t toast my bread.
2. Why doesn’t my toaster work?
3. There is something wrong with the electrical outlet.
4. If something is wrong with the outlet, my coffeemaker also won’t work when plugged into it.
5. I plug my coffeemaker into the outlet.
6. My coffeemaker works.

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.

Basic and Applied Science

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 5***) 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.

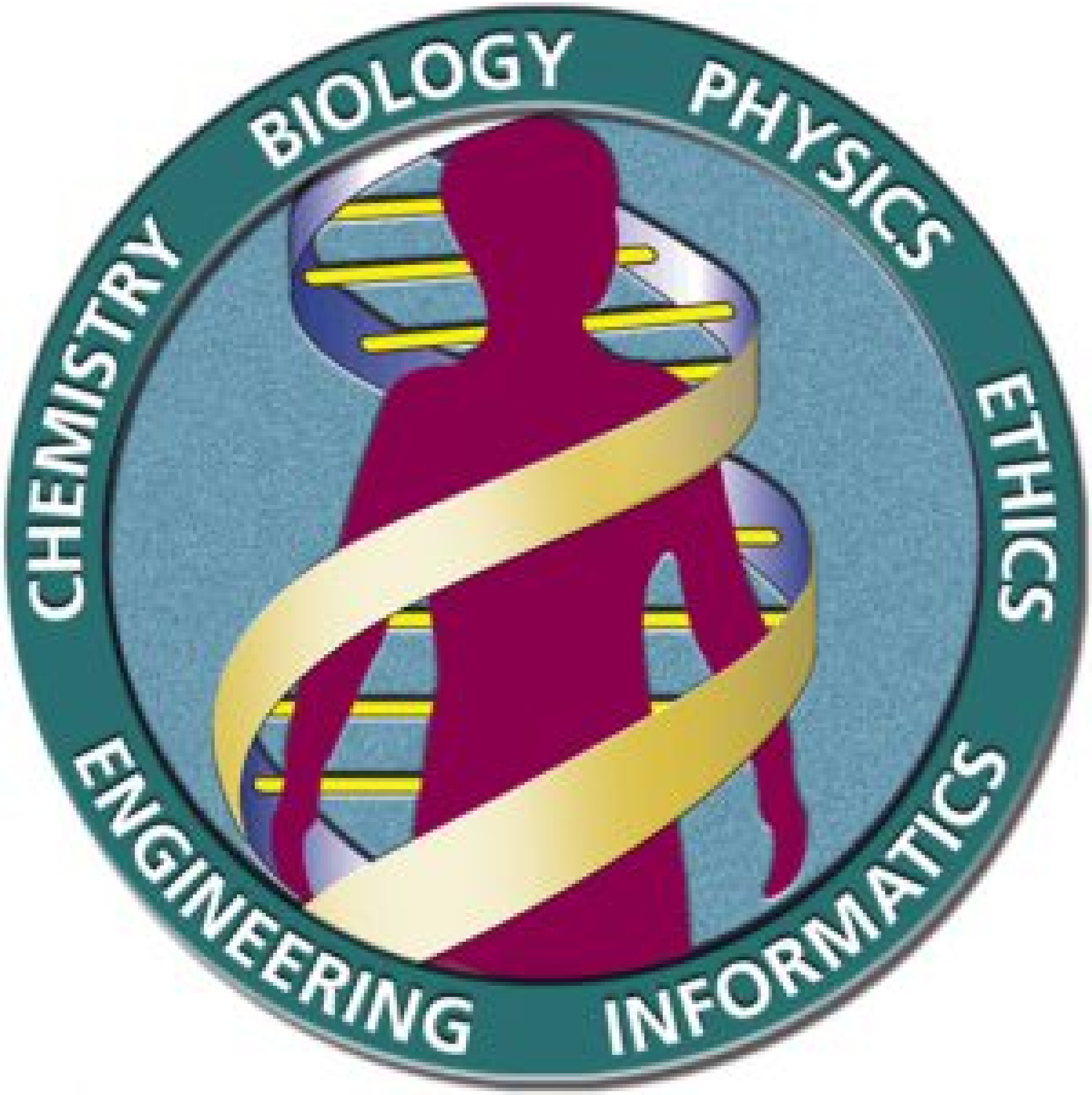


Figure 5. *The Human Genome Project was a 13-year collaborative effort among researchers working in several different fields of science. The project was completed in 2003. (credit: the U.S. Department of Energy Genome Programs)*

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.

Reporting Scientific Work

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.

Section Summary

Biology is the science that studies living organisms and their interactions with one another and their environments. Science attempts to describe and understand the nature of the universe in whole or in part. Science has many fields; those fields related to the physical world and its phenomena are considered natural sciences.

A hypothesis is a tentative explanation for an observation. A scientific theory is a well-tested and consistently verified explanation for a set of observations or phenomena. A scientific law is a description, often in the form of a mathematical formula, of the behavior of an aspect of nature under certain circumstances. Two types of logical reasoning are used in science. Inductive reasoning uses results to produce general scientific principles. Deductive reasoning is a form of logical thinking that predicts results by applying general principles. The common thread throughout scientific research is the use of the scientific method. Scientists present their results in peer-reviewed scientific papers published in scientific journals.

Science can be basic or applied. The main goal of basic science is to expand knowledge without any expectation of short-term practical application of that knowledge. The primary goal of applied research, however, is to solve practical problems.

Glossary

applied science

a form of science that solves real-world problems

basic science
: science that seeks to expand knowledge regardless of the short-term application of that knowledge

control
: a part of an experiment that does not change during the experiment

deductive reasoning
: a form of logical thinking that uses a general statement to forecast specific results

descriptive science
: a form of science that aims to observe, explore, and find things out

falsifiable
: able to be disproven by experimental results

hypothesis
: a suggested explanation for an event, which can be tested

hypothesis-based science
: a form of science that begins with a specific explanation that is then tested

inductive reasoning
: a form of logical thinking that uses related observations to arrive at a general conclusion

life science
: a field of science, such as biology, that studies living things

natural science
: a field of science that studies the physical world, its phenomena, and processes

peer-reviewed article
: a scientific report that is reviewed by a scientist’s colleagues before publication

physical science
: a field of science, such as astronomy, physics, and chemistry, that studies nonliving matter

science
: knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method

scientific law
: a description, often in the form of a mathematical formula, for the behavior of some aspect of nature under certain specific conditions

scientific method
: a method of research with defined steps that include experiments and careful observation

scientific theory
: a thoroughly tested and confirmed explanation for observations or phenomena

variable
: a part of an experiment that can vary or change

46.

Essential Mathematics

Exponential Arithmetic

Exponential notation is used to express very large and very small numbers as a product of two numbers. The first number of the product, the *digit term*, is usually a number not less than 1 and not greater than 10. The second number of the product, the *exponential term*, is written as 10 with an exponent. Some examples of exponential notation are:

$$1000 = 1 \times 10^3$$
$$100 = 1 \times 10^2$$
$$10 = 1 \times 10^1$$
$$1 = 1 \times 10^0$$
$$0.1 = 1 \times 10^{-1}$$
$$0.001 = 1 \times 10^{-3}$$
$$2386 = 2.386 \times 1000 = 2.386 \times 10^3$$
$$0.123 = 1.23 \times 0.1 = 1.23 \times 10^{-1}$$

The power (exponent) of 10 is equal to the number of places the decimal is shifted to give the digit number. The exponential method is particularly useful notation for every large and very small numbers. For example, 1,230,000,000 = 1.23 x 10^9, and 0.00000000036 = 3.6 x 10^{-10}.

Addition of Exponentials

Convert all numbers to the same power of 10, add the digit terms of the numbers, and if appropriate, convert the digit term back to a number between 1 and 10 by adjusting the exponential term.

Adding Exponentials

Add 5.00 x 10^{-5} and 3.00 x 10^{-3}.

Solution

$$3.00 \times 10^{-3} = 300 \times 10^{-5}$$
$$(5.00 \times 10^{-5}) + (300 \times 10^{-5}) = 305 \times 10^{-5} = 3.05 \times 10^{-3}$$

Subtraction of Exponentials

Convert all numbers to the same power of 10, take the difference of the digit terms, and if appropriate, convert the digit term back to a number between 1 and 10 by adjusting the exponential term.

Subtracting Exponentials

Subtract 4.0×10^{-7} from 5.0×10^{-6}.

Solution

$$4.0 \times 10^{-7} = 0.40 \times 10^{-6}$$
$$(5.0 \times 10^{-6}) - (0.40 \times 10^{-6}) = 4.6 \times 10^{-6}$$

Multiplication of Exponentials

Multiply the digit terms in the usual way and add the exponents of the exponential terms.

Multiplying Exponentials

Multiply 4.2×10^{-8} by 2.0×10^{3}.

Solution

$$(4.2 \times 10^{-8}) \times (2.0 \times 10^{3}) = (4.2 \times 2.0) \times 10^{(-8)+(+3)} = 8.4 \times 10^{-5}$$

Division of Exponentials

Divide the digit term of the numerator by the digit term of the denominator and subtract the exponents of the exponential terms.

Dividing Exponentials

Divide 3.6 × 10^5 by 6.0 × 10^{-4}.

Solution

$$\frac{3.6 \times 10^{-5}}{6.0 \times 10^{-4}} = (\frac{3.6}{6.0}) \times 10^{(-5)-(-4)} = 0.60 \times 10^{-1} = 6.0 \times 10^{-2}$$

Squaring of Exponentials

Square the digit term in the usual way and multiply the exponent of the exponential term by 2.

Squaring Exponentials

Square the number 4.0 × 10^{-6}.

Solution

$$(4.0 \times 10^{-6})^2 = 4 \times 4 \times 10^{2\times(-6)} = 16 \times 10^{-12} = 1.6 \times 10^{-11}$$

Cubing of Exponentials

Cube the digit term in the usual way and multiply the exponent of the exponential term by 3.

Cubing Exponentials

Cube the number 2 × 10^4.

Solution

$$(2 \times 10^4)^3 = 2 \times 2 \times 2 \times 10^{3\times4} = 8 \times 10^{12}$$

Taking Square Roots of Exponentials

If necessary, decrease or increase the exponential term so that the power of 10 is evenly divisible by 2. Extract the square root of the digit term and divide the exponential term by 2.

Finding the Square Root of Exponentials

Find the square root of 1.6×10^{-7}.

Solution

$$1.6 \times 10^{-7} = 16 \times 10^{-8}$$

$$\sqrt{16 \times 10^{-8}} = \sqrt{16} \times \sqrt{10^{-8}} = \sqrt{16} \times 10^{-\frac{8}{2}} = 4.0 \times 10^{-4}$$

Significant Figures

A beekeeper reports that he has 525,341 bees. The last three figures of the number are obviously inaccurate, for during the time the keeper was counting the bees, some of them died and others hatched; this makes it quite difficult to determine the exact number of bees. It would have been more accurate if the beekeeper had reported the number 525,000. In other words, the last three figures are not significant, except to set the position of the decimal point. Their exact values have no meaning useful in this situation. In reporting any information as numbers, use only as many significant figures as the accuracy of the measurement warrants.

The importance of significant figures lies in their application to fundamental computation. In addition and subtraction, the sum or difference should contain as many digits to the right of the decimal as that in the least certain of the numbers used in the computation (indicated by underscoring in the following example).

Addition and Subtraction with Significant Figures

Add 4.383 g and 0.0023 g.

Solution

$$4.38\underline{3}\text{g}$$
$$0.002\underline{3}\text{g}$$

$$\overline{}$$

$$4.38\underline{5}\text{g}$$

In multiplication and division, the product or quotient should contain no more digits than that in the factor containing the least number of significant figures.

Multiplication and Division with Significant Figures

Multiply 0.6238 by 6.6.

Solution

$$0.623\underline{8} \times 6.\underline{6} = 4.\underline{1}$$

When rounding numbers, increase the retained digit by 1 if it is followed by a number larger than 5 ("round up"). Do not change the retained digit if the digits that follow are less than 5 ("round down"). If the retained digit is followed by 5, round up if the retained digit is odd, or round down if it is even (after rounding, the retained digit will thus always be even).

The Use of Logarithms and Exponential Numbers

The common logarithm of a number (log) is the power to which 10 must be raised to equal that number. For example, the common logarithm of 100 is 2, because 10 must be raised to the second power to equal 100. Additional examples are shown in ***Table 1***.

Table 1. Logarithms and exponential numbers

Number	Number Expressed Exponentially	Common Logarithm
1000	10^3	3
10	10^1	1
1	10^0	0
0.1	10^{-1}	-1
0.001	10^{-3}	-3

What is the common logarithm of 60? Because 60 lies between 10 and 100, which have logarithms of 1 and 2, respectively, the logarithm of 60 is 1.7782; that is,

$$60 = 10^{1.7782}$$

The common logarithm of a number less than 1 has a negative value. The logarithm of 0.03918 is −1.4069, or

$$0.03918 = 10^{-1.4069} = \frac{1}{10^{1.4069}}$$

To obtain the common logarithm of a number, use the log button on your calculator. To calculate a number from its logarithm, take the inverse log of the logarithm, or calculate 10^x (where *x* is the logarithm of the number).

The natural logarithm of a number (ln) is the power to which *e* must be raised to equal the number; *e* is the constant 2.7182818. For example, the natural logarithm of 10 is 2.303; that is,

$$10 = e^{2.303} = 2.7182818^{2.303}$$

To obtain the natural logarithm of a number, use the *ln* button on your calculator. To calculate a number from its natural logarithm, enter the natural logarithm and take the inverse *ln* of the natural logarithm, or calculate e^x (where *x* is the natural logarithm of the number).

Logarithms are exponents; thus, operations involving logarithms follow the same rules as operations involving exponents.

Rule #1:

The logarithm of a product of two numbers is the sum of the logarithms of the two numbers.

$$\log xy = \log x + \log y, \text{ and } \ln xy = \ln x + \ln y$$

Rule #2:

The logarithm of the number resulting from the division of two numbers is the difference between the logarithms of the two numbers.

$$\log \frac{x}{y} = \log x - \log y, \text{ and } \ln \frac{x}{y} = \ln x - \ln y$$

Rule #3:

The logarithm of a number raised to an exponent is the product of the exponent and the logarithm of the number.

$$\log x^n = n\log x \text{ and } \ln x^n = n\ln x$$

The Solution of Quadratic Equations

Mathematical functions of this form are known as second-order polynomials or, more commonly, quadratic functions.

$$ax^2 + bx + c = 0$$

The solution or roots for any quadratic equation can be calculated using the following formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Solving Quadratic Equations

Solve the quadratic equation $3x^2 + 13x - 10 = 0$.

Solution

Substituting the values $a = 3$, $b = 13$, $c = -10$ in the formula, we obtain

$$x = \frac{-13 \pm \sqrt{(13)^2 - 4 \times 3 \times (-10)}}{2 \times 3}$$

$$x = \frac{-13 \pm \sqrt{169 + 120}}{6} = \frac{-13 \pm \sqrt{289}}{6} = \frac{-13 \pm 17}{6}$$

The two roots are therefore

$$x = \frac{-13 + 17}{6} = \frac{2}{3} \text{ and } x = \frac{-13 - 17}{6} = -5$$

Quadratic equations constructed on physical data always have real roots, and of these real roots, often only those having positive values are of any significance.

Two-Dimensional (x-y) Graphing

The relationship between any two properties of a system can be represented graphically by a two-dimensional data plot. Such a graph has two axes: a horizontal one corresponding to the independent variable, or the variable whose value is being controlled (x), and a vertical axis corresponding to the dependent variable, or the variable whose value is being observed or measured (y).

When the value of y is changing as a function of x (that is, different values of x correspond to different values of y), a graph of this change can be plotted or sketched (***Figure 1***). The graph can be produced by using specific values for (x,y) data pairs (***Table 2***).

Graphing the Dependence of y on x

As shown, (x,y) points can be plotted on a graph and connected to produce a graphical representation of the dependence of y on x.

Table 2. This table contains the following points: (1,5), (2,10), (3,7), and (4,14).

x	***y***
1	5
2	10
3	7
4	14

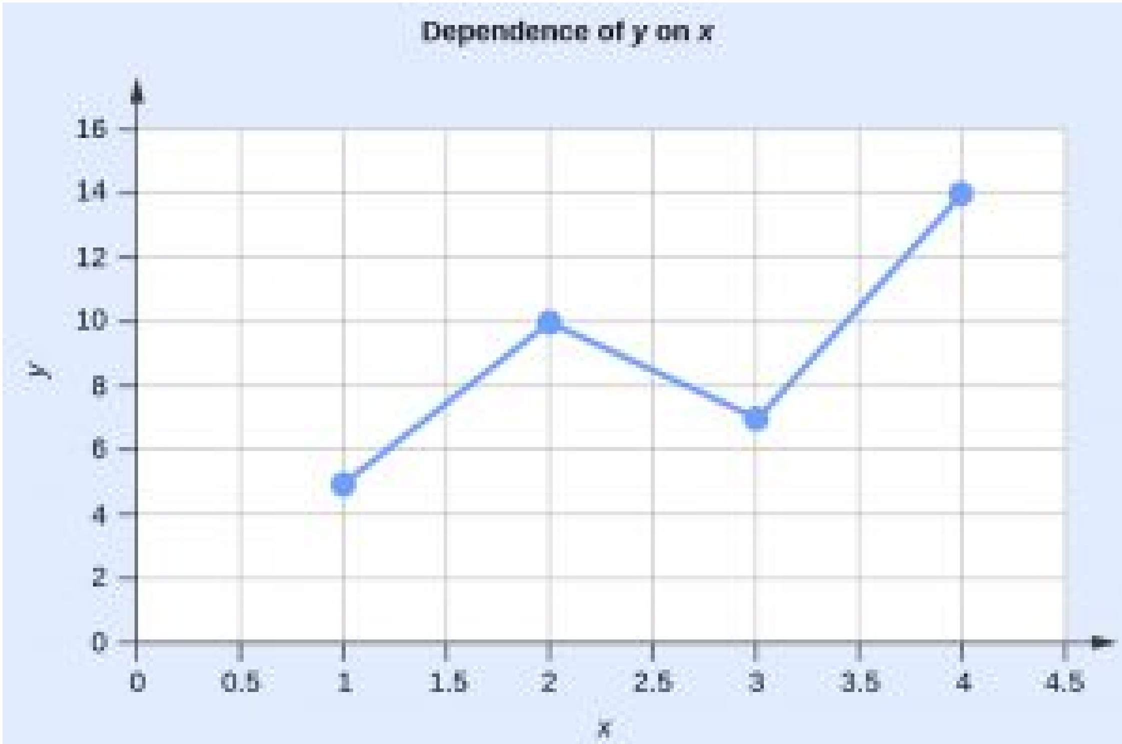


Figure 1. *A graphical representation of the dependence of y on x, using given points.*

If the function that describes the dependence of *y* on *x* is known, it may be used to compute *x*,*y* data pairs (***Table 3***) that may subsequently be plotted (***Figure 2***).

Plotting Data Pairs

If we know that $y = x^2 + 2$, we can produce a table of a few (*x*,*y*) values and then plot the line based on the data shown here.

Table 3. Data used to plot the line $y = x^2 + 2$

x	$y = x^2 + 2$
1	3
2	6
3	11
4	18

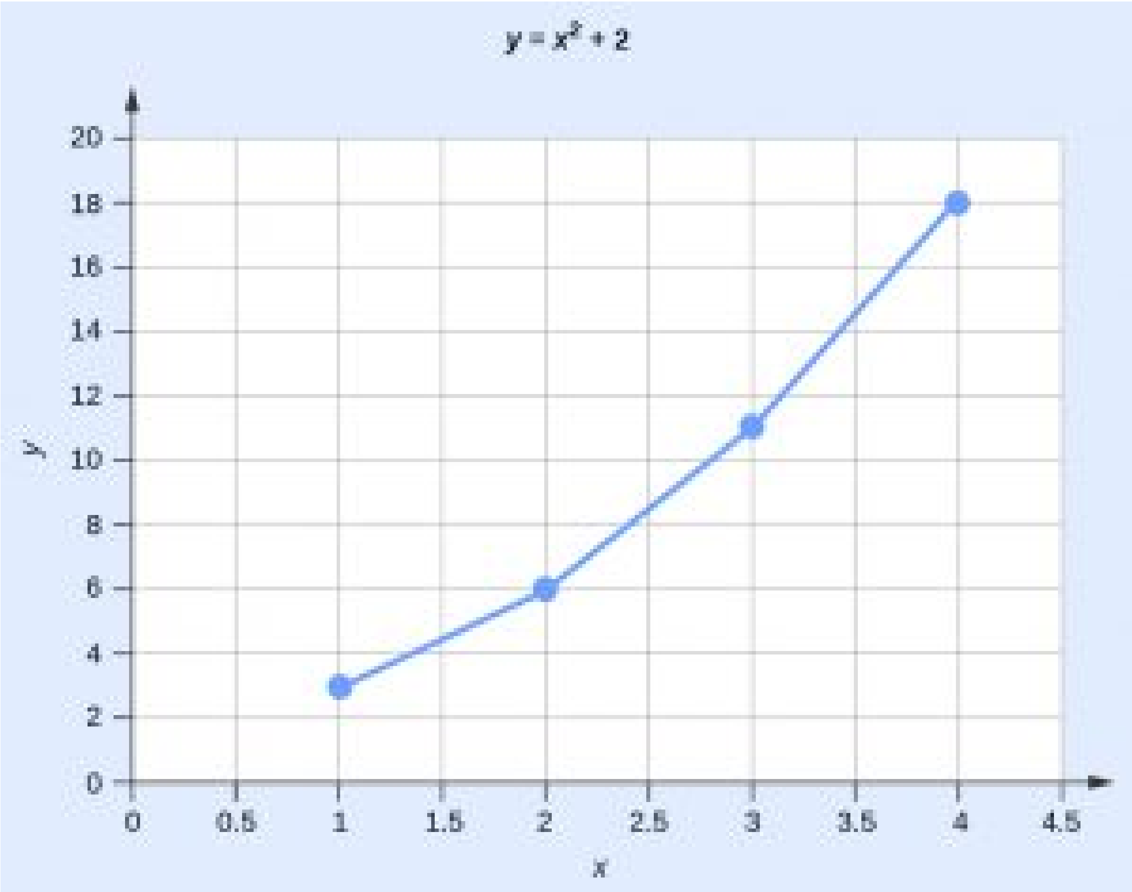


Figure 2. *If $y = x^2 + 2$, we can plot a line based on a few different (x,y) values as shown here.*