DSE4 (BIOPHYSICS)

Dr. Debolina Das
Haldia Government College

[Sincere acknowledgments to numerous online articles, wikipedia, youtube videos and definitely Khan academy e-resources]

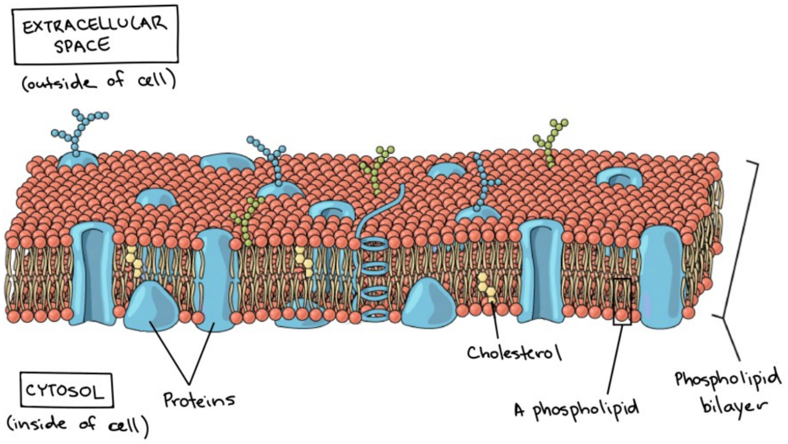
Biophysics is a study of the Physics and Chemistry of life. In short, it is a study of the Amino Acids. An amino acid is an organic molecule that is made up of a basic amino group ($-NH_2$), an acidic carboxyl group ($-COOH$), and an organic R group (or side chain) that is unique to each amino acid. These amino acids are essentially the basic building blocks of life, and the same 20 standard amino acids have been used in proteins throughout the history of life on Earth. Just like the 26 alphabets can be arranged in infinite ways to write a story, the 20 amino acids can be combined in various style to make different **proteins** that write the story of life. To start with...

The Cell Membrane:

The membrane is a delicate, two-layered structure of lipids and proteins, and it controls what can enter and exit the cell. Similarly, the cytoplasm of a eukaryotic cell consists not only of cytosol—a gel-like substance made up of water, ions, and macromolecules—but also of organelles and the structural proteins that make up the cytoskeleton, or "skeleton of the cell."

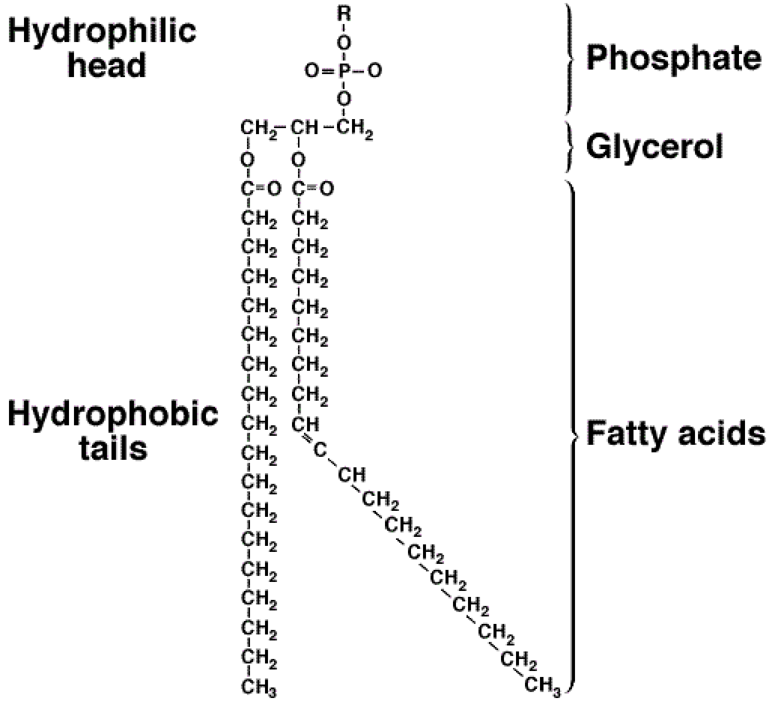
Both prokaryotic and eukaryotic cells have a **plasma membrane**, a double layer of lipids that separates the cell interior from the outside environment. This double layer consists largely of specialized lipids called phospholipids. A **phospholipid** is made up of a hydrophilic, water-loving, phosphate head, along with two hydrophobic, water-fearing, fatty acid tails. Phospholipids spontaneously arrange themselves in a double-layered structure with their hydrophobic tails pointing inward and their hydrophilic heads facing outward. This energetically favorable two-layer structure, called a **phospholipid bilayer**, is found in many biological membranes.

As shown below, proteins are also an important component of the plasma membrane. Some of them pass all the way through the membrane, serving as channels or signal receptors, while others are just attached at the edge. Different types of lipids, such as cholesterol, may also be found in the cell membrane and affect its fluidity.



The plasma membrane is the border between the interior and exterior of a cell. As such, it controls passage of various molecules—including sugars, amino acids, ions, and water—into and out of the cell. How easily these molecules can cross the membrane depends on their size and polarity. Some small, nonpolar molecules, such as oxygen, can pass directly through the phospholipid portion of the membrane.

Larger and more polar, hydrophilic, molecules, such as amino acids, must instead cross the membrane by way of protein channels, a process that is often regulated by the cell. The surface area of the plasma membrane limits the exchange of materials between a cell and its environment. Some cells are specialized in the exchange of wastes or nutrients and have modifications to increase the area of the plasma membrane. For instance, the membranes of some nutrient-absorbing cells are folded into fingerlike projections called **microvilli**, singular, **microvillus**. Cells with microvilli cover the inside surface of the small intestine, the organ that absorbs nutrients from digested food. The microvilli help intestinal cells maximize their absorption of nutrients from food by increasing plasma membrane surface area.



Structure of Phospholipid

What is Metabolism?

Metabolism is the term for a set of chemical reactions that occur in the cells of living organisms to sustain life. The metabolic processes lead to growth and reproduction and allow living organisms to maintain their structures and respond to the surrounding environment. All chemical reactions that occur in living organisms, from digestion to the transport of substances from cell to cell, can be part of metabolism. Intermediary or intermediate metabolism is the term for the transport of substances into and between different cells.

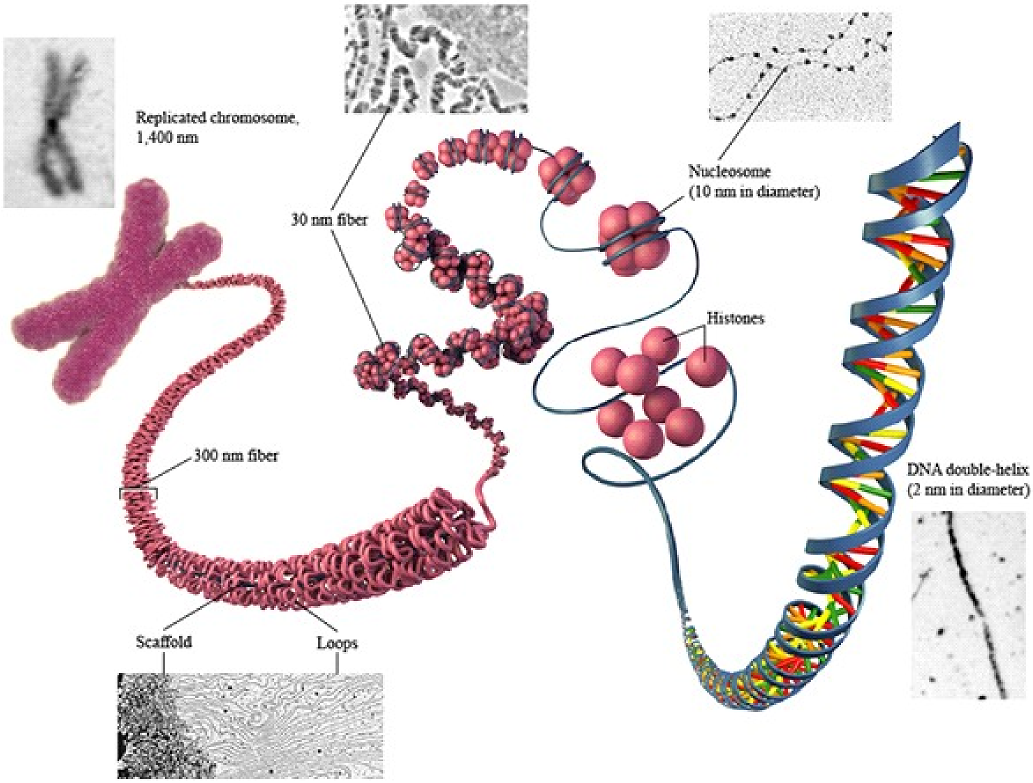
There are two categories of metabolism: **catabolism and anabolism**. Catabolism is the breakdown of organic matter, and anabolism uses energy to construct components of cells, such as proteins and nucleic acids.

The chemical reactions in the metabolic process are organized into metabolic pathways whereby one chemical is transformed through a series of steps into another chemical. Enzymes assist in this process by facilitating reactions and serving as catalysts for the reactions to occur. The reactions would not occur without enzymes, which respond to signals between cells and regulate the metabolic pathways. The speed of metabolism is called the metabolic rate.

The metabolism of a living organism allows it to determine which substances are nutritious and useful, and which are poisonous.

Some of the other chemicals and parts of an organism involved in the metabolic process are amino acids, proteins, lipids, carbohydrates, nucleotides, coenzymes and minerals and cofactors.

What's there in a cell?



Explanation:

- Let us start the story at the level of a cell. There is a nucleus, inside which you see chromosomes in pairs. These are called homologous pairs: they are in duplicate.
- Each chromosome contains a linear DNA molecule, closely associated with histone proteins. DNA is genetic material, made of only four different nucleotides arranged in a double helix. It contains chemical codes which guide life processes, written by using those four letters.
- Along the length of DNA molecule, only some stretches are 'meaningful' and are called genes. Hence, several genes are located on same chromosome. Homologous chromosomes carry same genes, generally in same order.

Cell Reproduction

Cell reproduction is the process by which cells divide to form new cells. Each time a cell divides, it makes a copy of all of its *chromosomes*, which are tightly coiled strands of DNA, the genetic material that holds the instructions for all life, and sends an identical copy to the new cell that is created. This is a process called Mitosis. There are two different types of cell reproduction--mitosis and meiosis. These processes are responsible for creating two different types of cells. Mitosis is a process that creates a nearly exact copy of the original cell. Somatic cells, which include nearly all human cells, are created by this process. Meiosis is a different form of reproduction that leads to the production of germ cells, or *sex cells*. All cells fall into one of these two categories. Some organisms, such as bacteria and single-celled organisms, use only mitosis for cell reproduction.

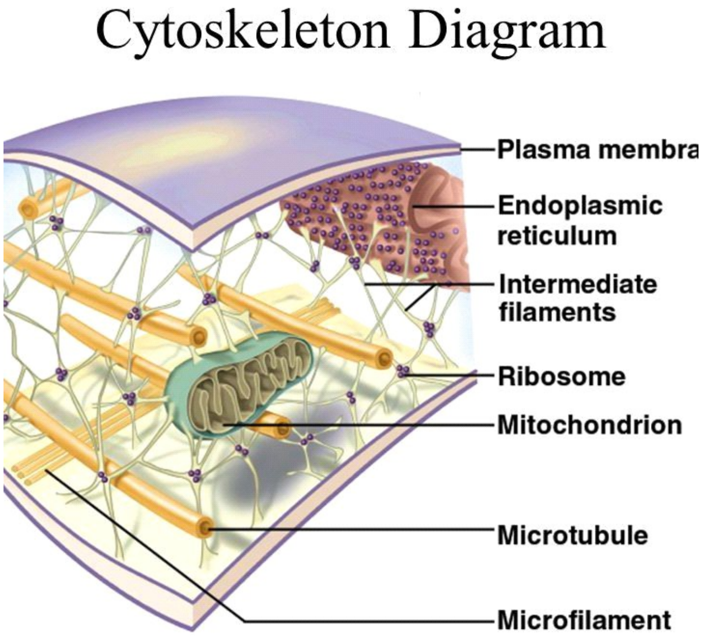
The difference between mitosis and meiosis can also be thought of as the difference between sexual and asexual reproduction. Humans obviously reproduce sexually, but this is not true for all organisms. Many lower-order cells create entirely new organisms with each round of **mitosis**: asexual reproduction. In humans and other organisms that reproduce sexually, **meiosis** is needed to take into account the genetic contribution of the two parent organisms.

Cell origin and Evolution

Chronology of cell evolution:

Formation of earth (5 billion years ago) => First cells (4 billion years) => First eukaryotes, Oxidative metabolism, Photosynthesis (3 billion years ago) => multicellular organism (2 billion years ago).

There are two types of cell: Prokaryotic (Cells lacking a nuclear envelope, cytoplasmic organelles, and a cytoskeleton (primarily bacteria).) and Eukaryotik (Cells having all the above). Cytoskeleton is the skeleton of a cell:



Fucntions of cytoskeleton:

1. Structural scaffold creating and supporting cell shape
2. Network of molecular roads for intracellular transport of materials
3. framework for whole cell movement
4. framework for cell division

Eukaryotik cells have a nucleus in which the genetic material is separated from the cytoplasm. Prokaryotic cells are generally smaller and simpler than Eukaryotik cells; in addition to the absence of a nucleus, their genomes are less complex. In spite of these differences, the same basic molecular mechanisms govern the lives of both prokaryotes and eukaryotes, indicating that all present-day cells are descended from a single primordial ancestor.

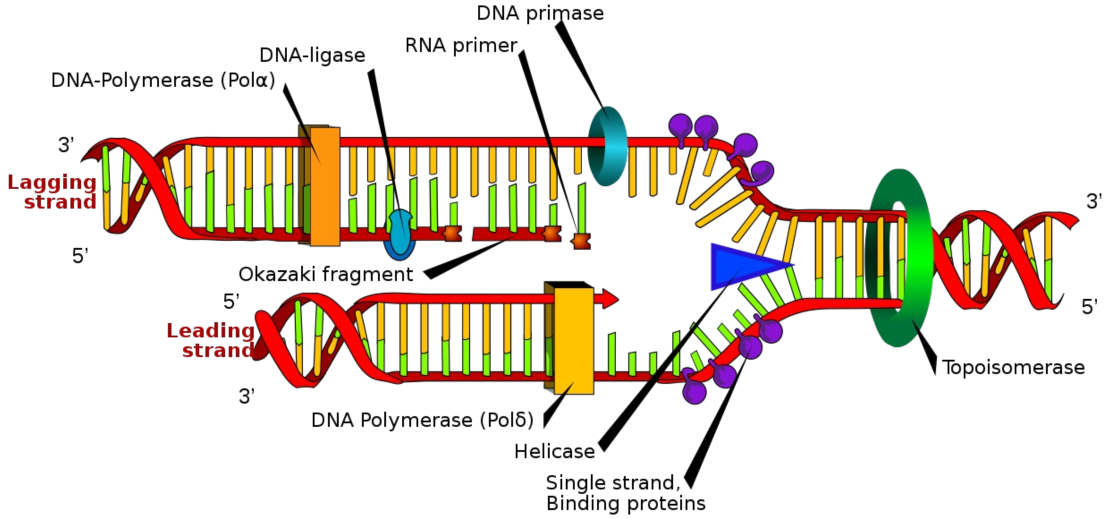
Steps of evolution:

1. **Formation of simple organic molecules**: At the time life arose, the atmosphere of Earth is thought to have contained little or no free oxygen, instead consisting principally of CO_2 and N_2 in addition to smaller amounts of gases such as H_2, H_2S, and CO. Such an atmosphere provides reducing conditions in which organic molecules, given a source of energy such as sunlight or electrical discharge, can form spontaneously.

2. **Spontaneous polymerization and formation of macromolecules**: Heating dry mixtures of amino acids, for example, results in their polymerization to form polypeptides.

3. **Self-replication**: But the critical characteristic of the macromolecule from which life evolved must have been the ability to replicate itself. Only a macromolecule capable of directing the synthesis of new copies of itself would have been capable of reproduction and further evolution. Of the two major classes of informational macromolecules in present-day cells, (nucleic acids and protein), only the nucleic acids are capable of directing their own self-replication. Nucleic acids can serve as templates for their own synthesis as a result of specific base pairing between complementary nucleotides. A critical step in understanding molecular evolution was reached in 1980 when it was discovered that RNA s capable of catalyzing a number of chemical reactions, including the polymerization of nucleotides. RNA is thus uniquely able both to serve as a template for and to catalyze its own replication. Consequently, RNA is generally believed to have been the initial genetic system, and an early stage of chemical evolution is thought to have been based on self-replicating RNA molecules—a period of evolution known as the **RNA world**. Ordered interactions between RNA and amino acids then evolved into the present-day genetic code, and **DNA eventually replaced RNA as the genetic material.**

Complementary pairing between nucleotides (adenine [A] with uracil [U] and guanine [G] with cytosine [C]) allows one strand of RNA to serve as a template for the synthesis of a new strand with the complementary sequence.



Step 1: Replication Fork Formation

Before DNA can be replicated, the double stranded molecule must be “unzipped” into two single strands. DNA has four bases called adenine (A), thymine (T), cytosine (C) and guanine (G) that form pairs between the two strands. Adenine only pairs with thymine and cytosine only binds with guanine. In order to unwind DNA, these interactions between base pairs must be broken. This is performed by an enzyme known as DNA helicase. DNA helicase disrupts the hydrogen bonding between base pairs to separate the strands into a Y shape known as the replication fork. This area will be the template for replication to begin.

DNA is directional in both strands, signified by a 5' and 3' end. This notation signifies which side group is attached the DNA backbone. The 5' end has a phosphate (P) group attached, while the 3' end has a hydroxyl (OH) group attached. This directionality is important for replication as it only progresses in the 5' to 3' direction. However, the replication fork is bi-directional; one strand is oriented in the 3' to 5' direction (leading strand) while the other is oriented 5' to 3' (lagging strand). The two sides are therefore replicated with two different processes to accommodate the directional difference.

Replication Begins

Step 2: Primer Binding

The leading strand is the simplest to replicate. Once the DNA strands have been separated, a short piece of RNA called a primer binds to the 3' end of the strand. The primer always binds as the starting point for replication. Primers are generated by the enzyme DNA primase.

Step 3: Elongation

Enzymes known as DNA polymerases are responsible creating the new strand by a process called elongation. There are five different known types of DNA polymerases in bacteria and human cells. In bacteria such as E. coli, polymerase III is the main replication enzyme, while polymerase I, II, IV and V are responsible for error checking and repair. DNA polymerase III binds to the strand at the site of the primer and begins adding new base pairs complementary to the strand during replication. In eukaryotic cells, polymerases alpha, delta, and epsilon are the primary polymerases involved in DNA replication. Because replication proceeds in the 5' to 3' direction on the leading strand, the newly formed strand is continuous.

The lagging strand begins replication by binding with multiple primers. Each primer is only several bases apart. DNA polymerase then adds pieces of DNA, called Okazaki fragments, to the strand between primers. This process of replication is discontinuous as the newly created fragments are disjointed.

Step 4: Termination

Once both the continuous and discontinuous strands are formed, an enzyme called exonuclease removes all RNA primers from the original strands. These primers are then replaced with appropriate bases. Another exonuclease “proofreads” the newly formed DNA to check, remove and replace any errors. Another enzyme called DNA ligase joins Okazaki fragments together forming a single unified strand. The ends of the linear DNA present a problem as DNA polymerase can only add nucleotides in the 5′ to 3′

direction. The ends of the parent strands consist of repeated DNA sequences called telomeres. Telomeres act as protective caps at the end of chromosomes to prevent nearby chromosomes from fusing. A special type of DNA polymerase enzyme called telomerase catalyzes the synthesis of telomere sequences at the ends of the DNA. Once completed, the parent strand and its complementary DNA strand coils into the familiar double helix shape. In the end, replication produces two DNA molecules, each with one strand from the parent molecule and one new strand.

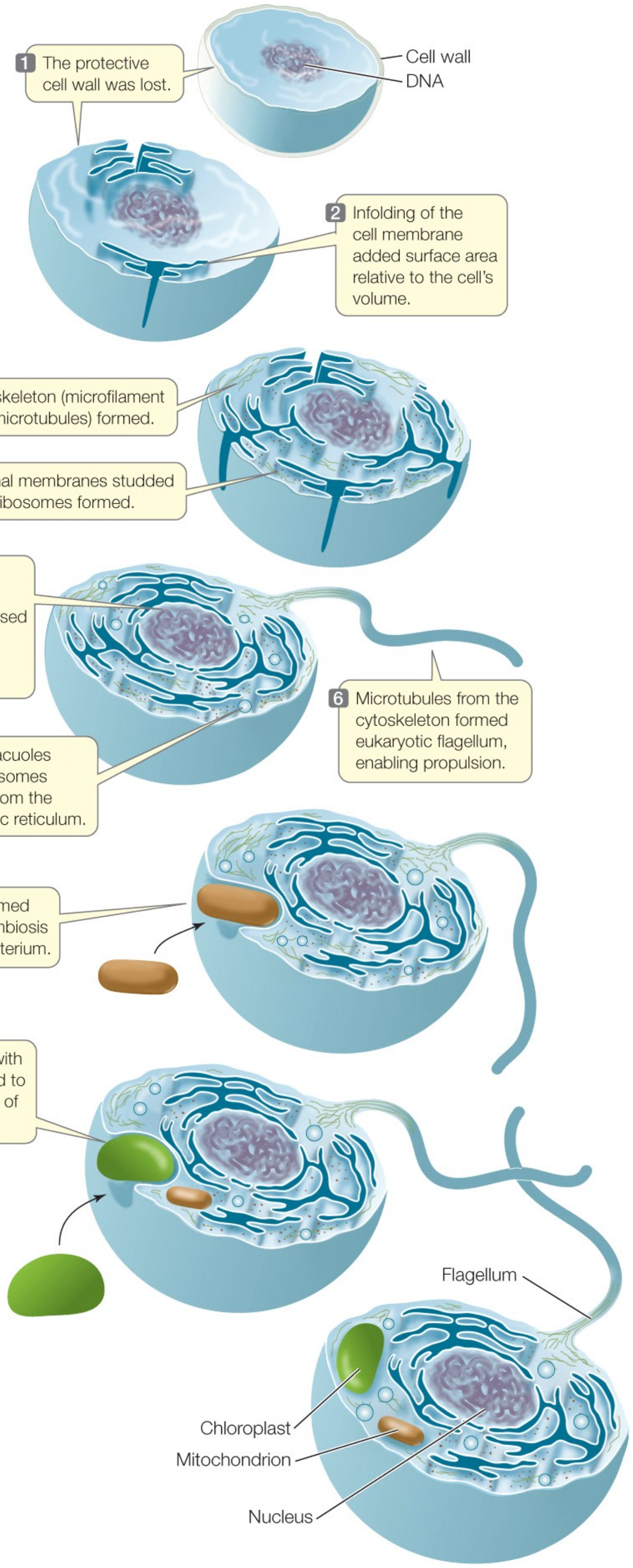
Replication Enzymes

DNA replication would not occur without enzymes that catalyze various steps in the process. Enzymes that participate in the eukaryotic DNA replication process include:

DNA helicase - unwinds and separates double stranded DNA as it moves along the DNA. It forms the replication fork by breaking hydrogen bonds between nucleotide pairs in DNA.
DNA primase - a type of RNA polymerase that generates RNA primers. Primers are short RNA molecules that act as templates for the starting point of DNA replication.
DNA polymerases - synthesize new DNA molecules by adding nucleotides to leading and lagging DNA strands.
Topoisomerase or DNA Gyrase - unwinds and rewinds DNA strands to prevent the DNA from becoming tangled or supercoiled.
Exonucleases - group of enzymes that remove nucleotide bases from the end of a DNA chain.
DNA ligase - joins DNA fragments together by forming phosphodiester bonds between nucleotides.

DNA replication steps and processes.pdf

4. **Further development of complex structure:**



Multicellularity:

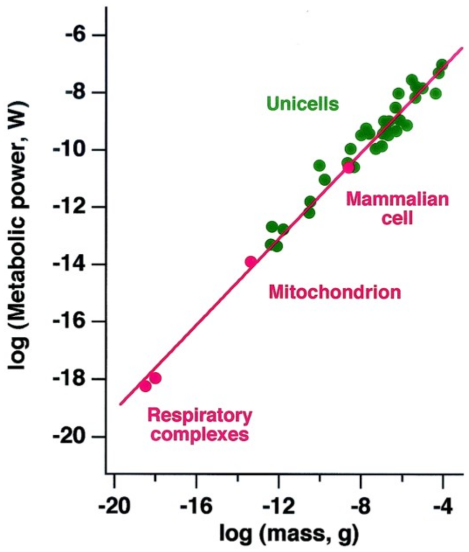
A **unicellular organism** is an **organism** that consists of a **single cell**. This means all life processes, such as reproduction, feeding, digestion, and excretion, occur in one cell. Amoebas, bacteria, and plankton are just some types of **unicellular organisms**. One of the greatest achievements in the evolution of complex life forms was the transition from unicellular organisms to multicellular organisms with different cell types.

Because the step from unicellular to multicellular life was taken early and frequently, the selective advantage of multicellularity seems to be quite large. In multicellular organisms like humans, a large number of cells form a cooperating cell community with specialized cell types and a division of labor among the various cells. Finding out how unicellular organisms can develop into multicellular organisms over the course of evolution is a central issue in biological research. Multicellular organisms arise in various ways, for example by cell division or by aggregation of many single cells.

BMR and Kleiber’s Law:

Basal Metabolic Rate is the rate of energy expenditure per unit time by endothermic animals at rest. An endotherm is an organism that maintains its body at a metabolically favorable temperature, largely by the use of heat set free by its internal bodily functions instead of relying almost purely on ambient heat. Proper measurement requires a strict set of criteria be met. These criteria include being in a physically and psychologically undisturbed state, in a thermally neutral environment, while in the post-absorptive stae (i.e., not actively digesting food). Basal metabolic rate is the amount of energy per unit time that a person needs to keep the body functioning at rest. Some of those processes are breathing, blood circulation, controlling body temperature, cell growth, brain and nerve function, and contraction of muscles. Basal metabolic rate (BMR) affects the rate that a person burns calories and ultimately whether that individual maintains, gains, or loses weight.

Kleiber's law, named after Max Kleiber for his biology work in the early 1930s, is the observation that, for the vast majority of animals, an animal's metabolic rate scales to the ¾ power of the animal's mass. This is called The classic allometric scaling relationship relating metabolic rate (B) to body mass (M): $B=B_0M^{3/4}$. The fact that metabolic rate scales as the three-quarter power of body mass (M) in unicellular, as well as multicellular, organisms suggests that the same principles of biological design operate at multiple levels of organization.



Metabolites

Metabolites are the products and intermediates of cellular metabolism. Metabolites can have a multitude of functions, including energy conversion, signaling, epigenetic influence, and cofactor activity. [DNA modifications that do not change the DNA sequence can affect gene activity. Chemical compounds that are added to single genes can regulate their activity; these modifications are known as epigenetic changes.] [A **cofactor** is a non-protein chemical compound or metallic ion that is required for an enzyme's **activity** as a catalyst, a substance that increases the rate of a chemical reaction. **Cofactors** can be considered "helper molecules" that assist in biochemical transformations.]

Size and type of metabolites, protein and nucleic acids and their role in structure and process

Primary and secondary metabolites are often used in industrial microbiology for the production of food, amino acids, and antibiotics. Primary metabolites are considered essential to microorganisms for proper growth. Secondary metabolites do not play a role in growth, development, and reproduction, and are formed during the end or near the stationary phase of growth. These metabolites can be used in industrial microbiology to obtain amino acids, develop vaccines and antibiotics, and isolate chemicals necessary for organic synthesis. Metabolites, the intermediates and products of metabolism, are typically characterized by small molecules with various functions. Metabolites can be categorized into both primary and secondary metabolites.

Primary metabolites

Primary metabolites are involved in growth, development, and reproduction of the organism. The primary metabolite is typically a key component in maintaining normal physiological processes; thus, it is often referred to as a central metabolite. Primary metabolites are typically formed during the growth phase as a result of energy metabolism, and are deemed essential for proper growth. Examples of primary metabolites include alcohols such as ethanol, lactic acid, and certain amino acids.

Secondary metabolites

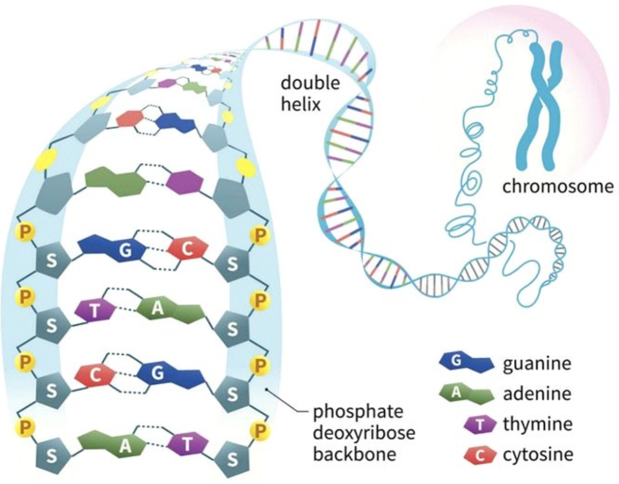
Secondary metabolites are typically organic compounds produced through the modification of primary metabolite synthases. Secondary metabolites do not play a role in growth, development, and reproduction like primary metabolites do, and are typically formed during the end or near the stationary phase of growth. Many of the identified secondary metabolites have a role in ecological function, including defense mechanism(s), by serving as antibiotics and by producing pigments. Examples of secondary metabolites with importance in industrial microbiology include atropine and antibiotics such as erythromycin and bacitracin. Atropine, derived from various plants, is a secondary metabolite with important use in the clinic. Atropine is a competitive antagonist for acetycholine receptors, specifically those of the muscarinic type, which can be used in the treatment of bradycardia. Antibiotics such as erythromcyin and bacitracin are also considered to be secondary metabolites. Erythromycin, derived from *Saccharopolyspora erythraea*, is a commonly used antibiotic with a wide antimicrobial spectrum. It is mass produced and commonly administered orally. Lastly, another example of an antibiotic which is classified as a secondary metabolite is bacitracin. Bacitracin, derived from organisms classified under *Bacillus subtilis*, is an antibiotic commonly used as topical drug.

There are five main classes of secondary metabolites such as terpenoids and steroids (Terpenoids have different variety of unrelated structures, while steroids have a common tetracyclic carbon skeleton), fatty acid-derived substances and polyketides, alkaloids (basic structure amine group), nonribosomal polypeptides (a compound consisting of two or more amino acids linked in a chain), and enzyme cofactors.

Proteins are large, complex macromolecules that play many critical roles in the body. They do most of the work in cells and are required for the structure, function, and regulation of the body's tissues and organs. Proteins are made up of hundreds or thousands of smaller units called amino acids, which are attached to one another in long chains. There are 20 different types of amino acids that can be combined to make a protein. The sequence of amino acids determines each protein's unique 3-dimensional structure and its specific function. Proteins can be described according to their large range of functions in the body, listed in alphabetical order:

Function	**Description**	**Example**
Antibody	Antibodies bind to specific foreign particles, such as viruses and bacteria, to help protect the body.	Immunoglobulin G
Enzyme	Enzymes carry out almost all of the thousands of chemical reactions that take place in cells. They also assist with the formation of new molecules by reading the genetic information stored in DNA.	Phenylalanine Hydroxylase
Messenger	Messenger proteins, such as some types of hormones, transmit signals to coordinate biological processes between different cells, tissues, and organs.	Growth hormone
Structural component	These proteins provide structure and support for cells. On a larger scale, they also allow the body to move.	Actin
Transport/storage	These proteins bind and carry atoms and small molecules within cells and throughout the body.	Ferritin

Nucleic acids are macromolecules that store genetic information, allow organisms to transfer genetic information from one generation to the next and enable protein production.



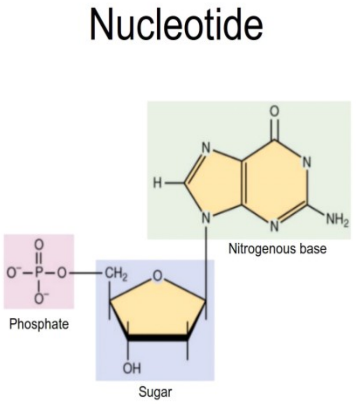
Nucleic acids are composed of **nucleotide** monomers linked together. Nucleotides have three parts:

A Nitrogenous Base

A Five-Carbon (Pentose) Sugar

A Phosphate Group

Nitrogenous bases include purine molecules (adenine and guanine) and pyrimidine molecules (cytosine, thymine, and uracil.) In DNA, the five-carbon sugar is deoxyribose, while ribose is the pentose sugar in RNA. Nucleotides are linked together to form polynucleotide chains. They are joined to one another by covalent bonds between the phosphate of one and the sugar of another. These linkages are called phosphodiester linkages. Phosphodiester linkages form the sugar-phosphate backbone of both DNA and RNA. Interestingly, some nucleotides perform important cellular functions as "individual" molecules, the most common example being adenosine triphosphate or ATP, which provides energy for many cell functions.



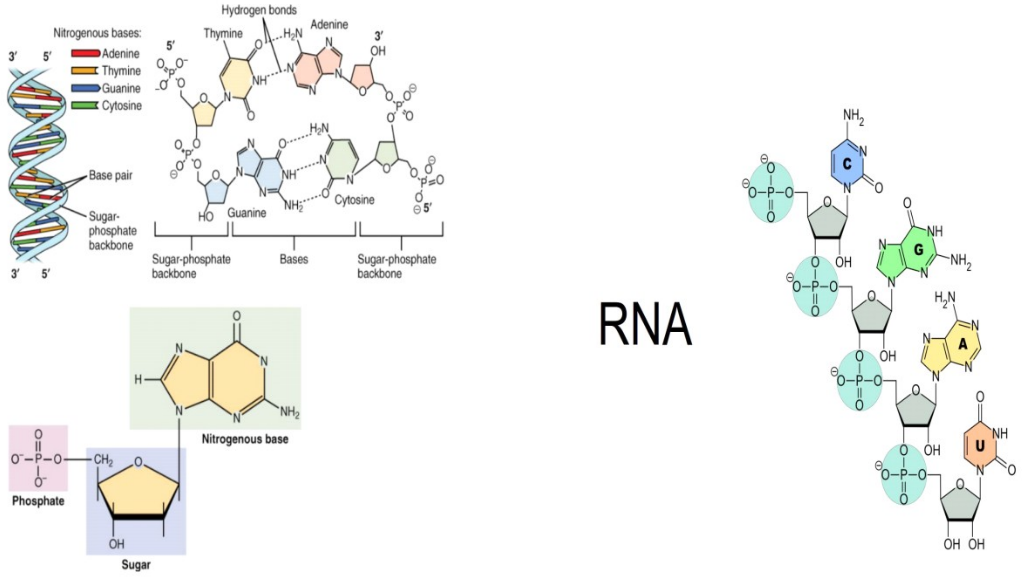
DNA:

DNA is the cellular molecule that contains instructions for the performance of all cell functions. When a cell divides, its DNA is copied and passed from one cell generation to the next. DNA is organized into chromosomes and found within the nucleus of our cells. It contains the "programmatic instructions" for cellular activities. When organisms produce offspring, these instructions are passed down through DNA. DNA commonly exists as a double-stranded molecule with a twisted double-helix shape. DNA is composed of a phosphate-deoxyribose sugar backbone and the four nitrogenous bases: adenine guanine cytosine and thymine. In double-stranded DNA, adenine pairs with thymine (A-T) and guanine pairs with cytosine (G-C).

RNA:

Ribonucleic acid, or **RNA** is one of the three major biological macromolecules that are essential for all known forms of life (along with DNA and proteins). RNA is essential for the synthesis of proteins. A central tenet of molecular biology states that the flow of genetic information in a cell is from DNA through **RNA** to proteins: “DNA makes **RNA** makes protein”. There are several types of RNA:

Messenger RNA (mRNA) is the RNA transcript or RNA copy of the DNA message produced during DNA transcription. Messenger RNA is translated to form proteins. **Transfer RNA (tRNA)** has a three-dimensional shape and is necessary for the translation of mRNA in protein synthesis. **Ribosomal RNA (rRNA**) is a component of ribosomes and is also involved in protein synthesis. **MicroRNAs (miRNAs**) are small RNAs that help to regulate gene expression.



__DNA structure and RNA structure__

RNA most commonly exists as a single-stranded molecule composed of a phosphate-ribose sugar backbone and the nitrogenous bases adenine, guanine, cytosine and uracil (U). When DNA is transcribed into an RNA transcript during DNA transcription, guanine pairs with cytosine (G-C) and adenine pairs with uracil (A-U).

The nucleic acids DNA and RNA differ in composition and structure. The differences are listed as follows:

DNA:

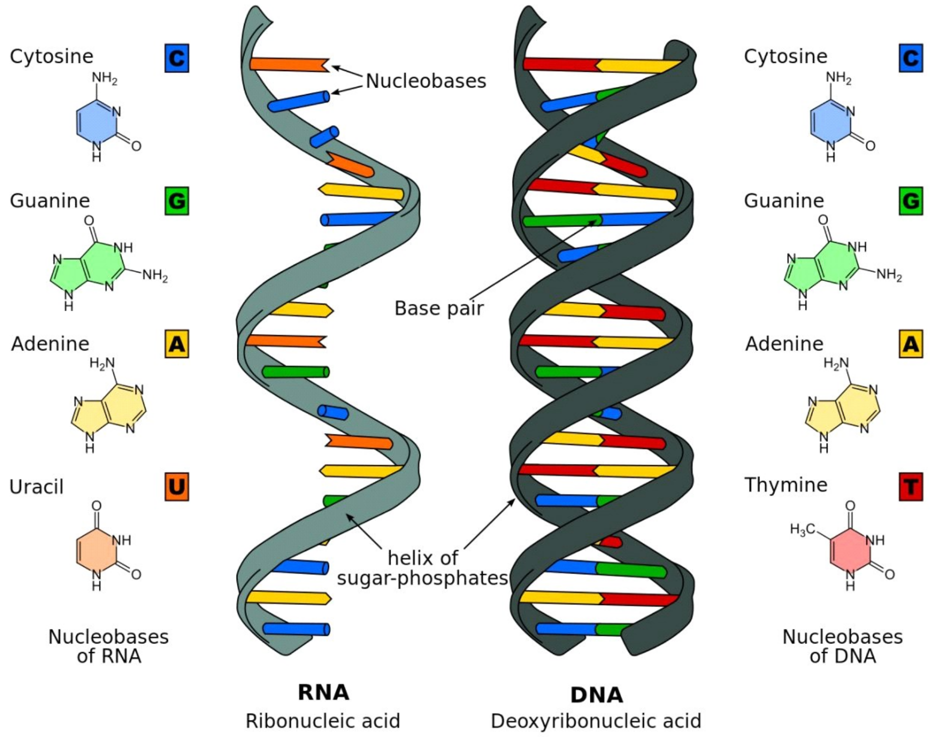
Nitrogenous Bases: Adenine, Guanine, Cytosine, and Thymine
Five-Carbon Sugar: Deoxyribose
Structure: Double-stranded

DNA is commonly found in its three-dimensional, double-helix shape. This twisted structure makes it possible for DNA to unwind for DNA replication and protein synthesis.

RNA:

Nitrogenous Bases: Adenine, Guanine, Cytosine, and Uracil
Five-Carbon Sugar: Ribose
Structure: Single-stranded

While RNA does not take on a double-helix shape like DNA, this molecule is able to form complex three-dimensional shapes. This is possible because RNA bases form complementary pairs with other bases on the same RNA strand. The base pairing causes RNA to fold, forming various shapes.



Transport Processes (in and out of cell)

Water and Solute Movement

Cell membranes act as barriers to most, but not all, molecules. Development of a cell membrane that could allow some materials to pass while constraining the movement of other molecules was a major step in the evolution of the cell. Cell membranes are differentially (or semi-) permeable barriers separating the inner cellular environment from the outer cellular (or external) environment.

Water potential is the tendency of water to move from an area of higher concentration to one of lower concentration. Energy exists in two forms: potential and kinetic. Water molecules move according to differences in potential energy between where they are and where they are going. Gravity and pressure are two enabling forces for this movement.

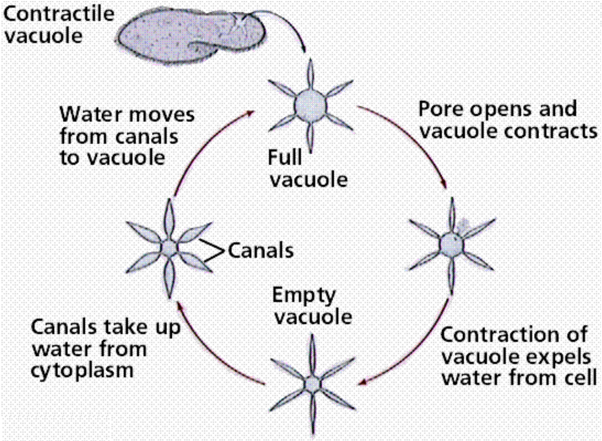
Diffusion is the net movement of a substance (liquid or gas) from an area of higher concentration to one of lower concentration. Since the molecules of any substance (solid, liquid, or gas) are in motion when that substance is above absolute zero (0 degrees Kelvin or -273 degrees C), energy is available for movement of the molecules from a higher potential state to a lower potential state, just as in the case of the water discussed above. The majority of the molecules move from higher to lower concentration, although there will be some that move from low to high. The overall (or net) movement is thus from high to low concentration. Eventually, if no energy is input into the system the molecules will reach a state of equilibrium where they will be distributed equally throughout the system.

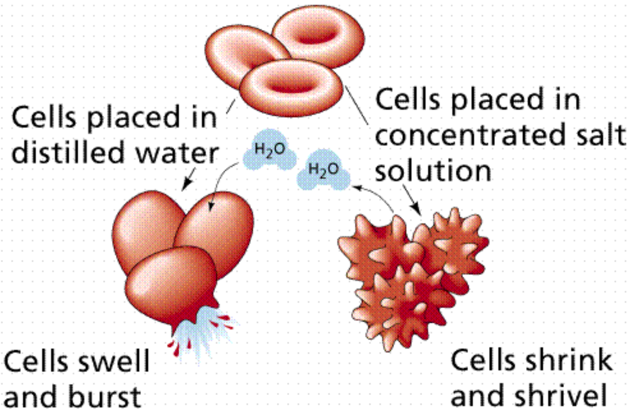
Cells and Diffusion

Water, carbon dioxide, and oxygen are among the few simple molecules that can cross the cell membrane by diffusion (or a type of diffusion known as osmosis). Osmosis is a process by which molecules of a solvent tend to pass through a semipermeable membrane from a less concentrated solution into a more concentrated one.

Diffusion is one principle method of movement of substances within cells, as well as the method for essential small molecules to cross the cell membrane. Gas exchange in gills and lungs operates by this process. Carbon dioxide is produced by all cells as a result of cellular metabolic processes. Since the source is inside the cell, the concentration gradient is constantly being replenished/re-elevated, thus the net flow of CO2 is out of the cell. Metabolic processes in animals and plants usually require oxygen, which is in lower concentration inside the cell, thus the net flow of oxygen is into the cell.

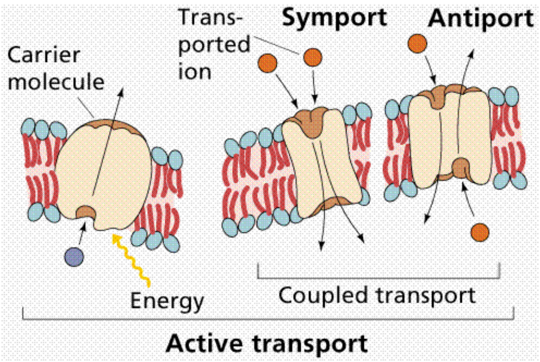
Osmosis is the diffusion of water across a semi-permeable (or differentially permeable or selectively permeable) membrane. The cell membrane, along with such things as dialysis tubing and cellulose acetate sausage casing, is such a membrane. The presence of a solute decreases the water potential of a substance. Thus there is more water per unit of volume in a glass of fresh-water than there is in an equivalent volume of sea-water. In a cell, which has so many organelles and other large molecules, the water flow is generally into the cell. Hypertonic solutions are those in which more solute (and hence lower water potential) is present. Hypotonic solutions are those with less solute (again read as higher water potential). Isotonic solutions have equal (iso-) concentrations of substances. Water potentials are thus equal, although there will still be equal amounts of water movement in and out of the cell, the net flow is zero. One of the major functions of blood in animals is the maintain an isotonic internal environment. This eliminates the problems associated with water loss or excess water gain in or out of cells. Again we return to homeostasis: The ability to maintain a relatively constant internal environment. Paramecium and other single-celled freshwater organisms have difficulty since they are usually hypertonic relative to their outside environment. Thus water will tend to flow across the cell membrane, swelling the cell and eventually bursting it. Not good for any cell! The contractile vacuole is the Paramecium's response to this problem. The pumping of water out of the cell by this method requires energy since the water is moving against the concentration gradient. Since ciliates (and many freshwater protozoans) are hypotonic, removal of water crossing the cell membrane by osmosis is a significant problem. One commonly employed mechanism is a contractile vacuole. Water is collected into the central ring of the vacuole and actively transported from the cell.

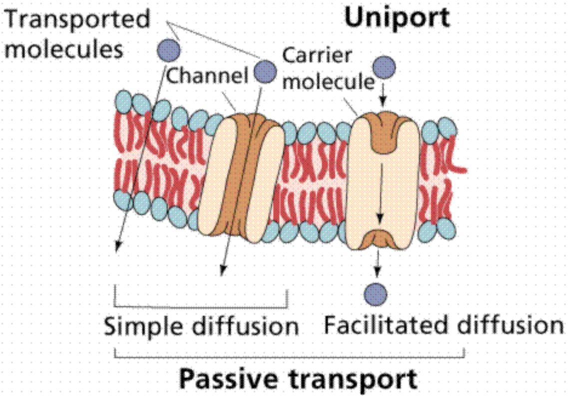


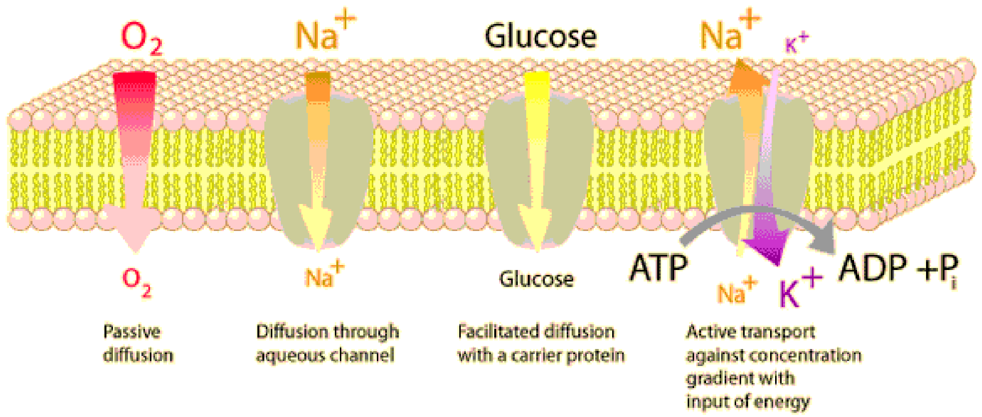


Active and Passive Transport

Passive transport requires no energy from the cell. Examples include the diffusion of oxygen and carbon dioxide, osmosis of water, and facilitated diffusion. Active Transport requires the cell to spend energy, usually in the form of ATP. Examples include transport of large molecules (non-lipid soluble) and the sodium-potassium pump. In active transport, molecules move across a membrane from a region of low concentration to an region of high concentration, and this requires a source of energy to "pump" the molecules uphill in concentration.







Carrier-Assisted Transport

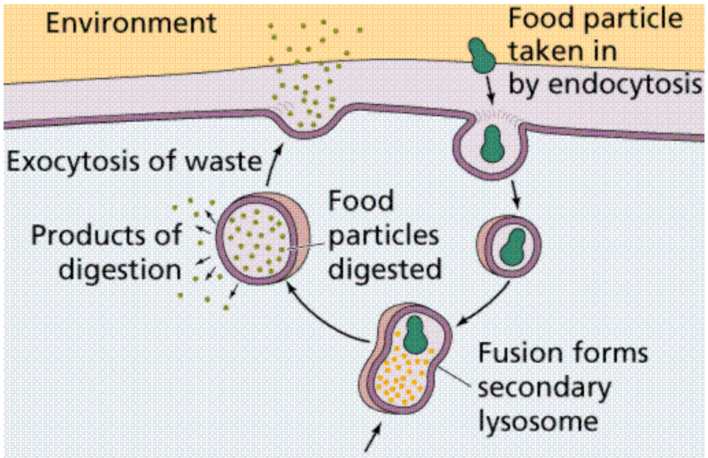
The transport proteins integrated into the cell membrane are often highly selective about the chemicals they allow to cross. Some of these proteins can move materials across the membrane only when assisted by the concentration gradient, a type of carrier-assisted transport known as facilitated diffusion. Both diffusion and facilitated diffusion are driven by the potential energy differences of a concentration gradient. Glucose enters most cells by facilitated diffusion. There seem to be a limiting number of glucose-transporting proteins. The rapid breakdown of glucose in the cell (a process known as glycolysis) maintains the concentration gradient. When the external concentration of glucose increases, however, the glucose transport does not exceed a certain rate, suggesting the limitation on transport. In the case of active transport, the proteins are having to move against the concentration gradient. For example the sodium-potassium pump in nerve cells. Na+ is maintained at low concentrations inside the cell and K+ is at higher concentrations. The reverse is the case on the outside of the cell. When a nerve message is propagated, the ions pass across the membrane, thus sending the message. After the message has passed, the ions must be actively transported back to their "starting positions" across the membrane. This is analogous to setting up 100 dominoes and then tipping over the first one. To reset them you must pick each one up, again at an energy cost. Up to one-third of the ATP used by a resting animal is used to reset the Na-K pump. [sodium-potassium pump The mechanism that uses ATP energy to reset the sodium and potassium ions after transmission of a nerve impulse.]

Types of Transport Molecules

Uniport transports one solute at a time. Symport transports the solute and a cotransported solute at the same time in the same direction. Antiport transports the solute in (or out) and the co-transported solute the opposite direction. One goes in the other goes out or vice-versa.

Vesicle-mediated Transport

Vesicles and vacuoles that fuse with the cell membrane may be utilized to release or transport chemicals out of the cell or to allow them to enter a cell. Exocytosis is the term applied when transport is out of the cell. Endocytosis is the case when a molecule causes the cell membrane to bulge inward, forming a vesicle. Phagocytosis is the type of endocytosis where an entire cell is engulfed. Pinocytosis is when the external fluid is engulfed. Receptor-mediated endocytosis occurs when the material to be transported binds to certain specific molecules in the membrane. Examples include the transport of insulin and cholesterol into animal cells.



Energy Storage

Living organisms use two major types of energy storage. Energy-rich molecules such as glycogen and triglycerides store energy in the form of covalent chemical bonds. Cells synthesize such molecules and store them for later release of the energy. The second major form of biological energy storage is electrochemical and takes the form of gradients of charged ions across cell membranes. Examples of energy storage molecules include following:

1. Glucose is a major energy storage molecule used to transport energy between different types of cells in the human body.
2. Starch
3. Fats: Fat itself has a high energy or calorfic value and can be directly burned in a fire. In the human body and presumably other animals, it serves a number of roles as there are different kinds of fats, but for the purpose of the discussion here, fats are frequently found associated with each of the organs in the body.

For example there is a deposit of fat on the heart and it was only relatively recently that it was realized that this acts as a temporary storage of buffer for energy. If one thinks of the blood system flowing around a heart, the glucose levels in it can fluctuate depending on a number of factors, such as whether one has recently digested a meal or engaged in strenous activity. As with most machines, steady inputs cause less strain on the system than large fluctuations, and the aforementioned cardiac fat assists in smoothing out during those lean periods and allows the heart to continue operating without additional strain.

4. Adenosine triphosphate or ATP is one of the key molecular energy carriers in the biological world.

Chamiosmosis

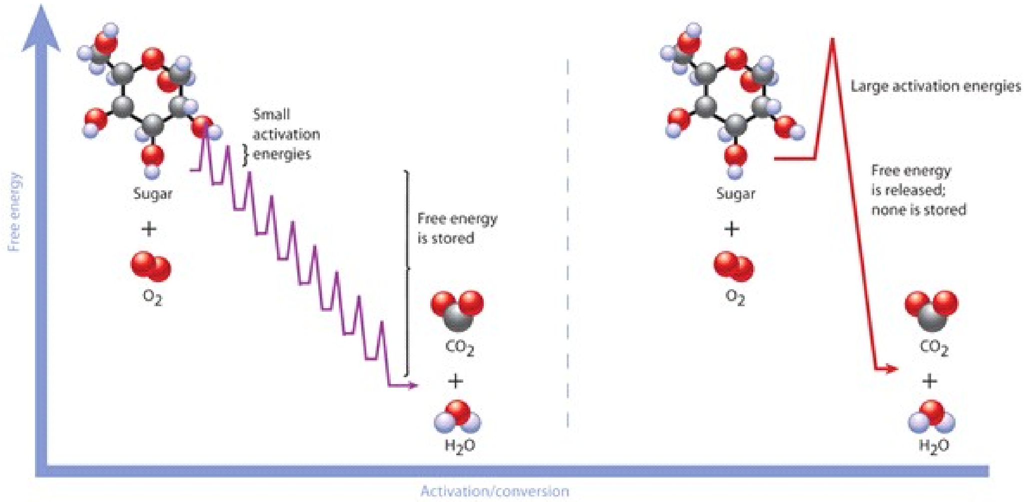
The term chemiosmosis refers to the inter-conversion of chemical energy (energy in the form of chemical bonds) and energy in the from of a transmembrane electrochemical gradient. The idea of "chemiosmotic coupling" arose largely from the work of Peter D. Mitchell and revolutionized the way biologists think about energy storage in biological systems.

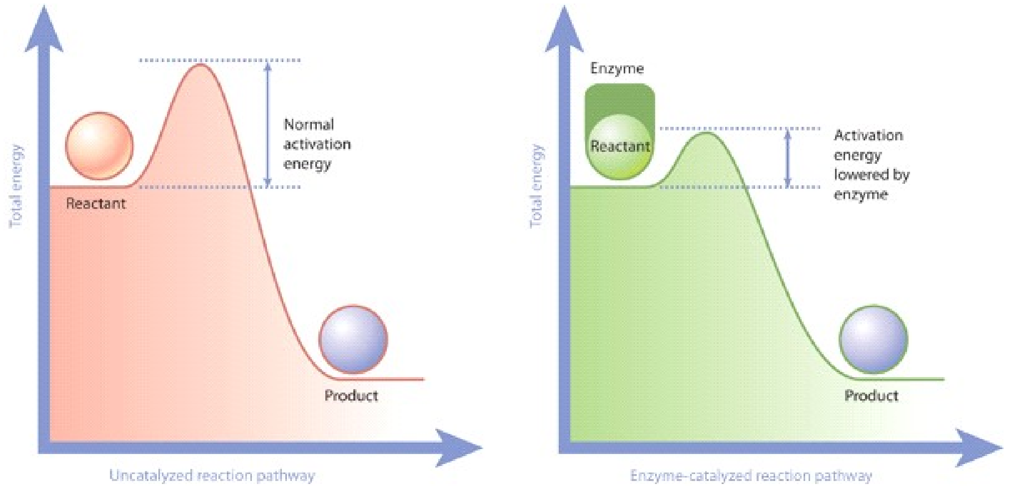
ATP

Adenosine Triphosphate (ATP) is synthesized by the F0-F1 ATP Synthase Complex. ATP Synthase is situated with one side facing into the inner membrane space, using a proton gradient (chemiosmotic gradient) to initiate changes in conformation of the subunits of the complex by allowing protons to move through the complex and into the inner membrane space. The changes in conformation (From Loose to Tight to Open) facilitate the binding of ADP and a phosphate (to the Loose site), covalently bonding of the ADP and Phosphate to form ATP (Tight site) and release of the completed ATP molecule (Open site). Four protons must move through the complex to initiate change in conformation necessary for change between Loose, Tight and Open; That is, twelve protons moving through the complex makes one complete turn, and 3 ATP.

How Do Cells Turn Nutrients into Usable Energy?

Complex organic food molecules such as sugars, fats, and proteins are rich sources of energy for cells because much of the energy used to form these molecules is literally stored within the chemical bonds that hold them together. Scientists can measure the amount of energy stored in foods using a device called a bomb calorimeter. With this technique, food is placed inside the calorimeter and heated until it burns. The excess heat released by the reaction is directly proportional to the amount of energy contained in the food. In reality, of course, cells don't work quite like calorimeters. Rather than burning all their energy in one large reaction, cells release the energy stored in their food molecules through a series of oxidation reactions. Oxidation describes a type of chemical reaction in which electrons are transferred from one molecule to another, changing the composition and energy content of both the donor and acceptor molecules. Food molecules act as electron donors. During each oxidation reaction involved in food breakdown, the product of the reaction has a lower energy content than the donor molecule that preceded it in the pathway. At the same time, electron acceptor molecules capture some of the energy lost from the food molecule during each oxidation reaction and store it for later use. Eventually, when the carbon atoms from a complex organic food molecule are fully oxidized at the end of the reaction chain, they are released as waste in the form of carbon dioxide. Cells do not use the energy from oxidation reactions as soon as it is released. Instead, they convert it into small, energy-rich molecules such as ATP and nicotinamide adenine dinucleotide (NADH), which can be used throughout the cell to power metabolism and construct new cellular components. In addition, workhorse proteins called enzymes use this chemical energy to catalyze, or accelerate, chemical reactions within the cell that would otherwise proceed very slowly. Enzymes do not force a reaction to proceed if it wouldn't do so without the catalyst; rather, they simply lower the energy barrier required for the reaction to begin.





What Specific Pathways Do Cells Use?

The particular energy pathway that a cell employs depends in large part on whether that cell is a eukaryote or a prokaryote. Eukaryotic cells use three major processes to transform the energy held in the chemical bonds of food molecules into more readily usable forms — often energy-rich carrier molecules. Adenosine 5'-triphosphate, or ATP, is the most abundant energy carrier molecule in cells. This molecule is made of a nitrogen base (adenine), a ribose sugar, and three phosphate groups. The word adenosine refers to the adenine plus the ribose sugar. The bond between the second and third phosphates is a high-energy bond (Figure 5).

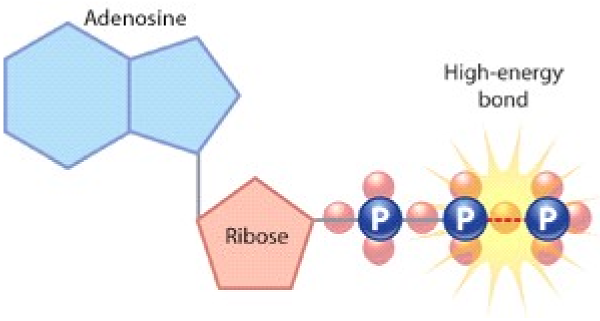
The first process in the eukaryotic energy pathway is glycolysis, which literally means "sugar splitting." During glycolysis, single molecules of glucose are split and ultimately converted into two molecules of a substance called pyruvate; because each glucose contains six carbon atoms, each resulting pyruvate contains just three carbons. Glycolysis is actually a series of ten chemical reactions that requires the input of two ATP molecules. This input is used to generate four new ATP molecules, which means that glycolysis results in a net gain of two ATPs. Two NADH molecules are also produced; these molecules serve as electron carriers for other biochemical reactions in the cell.

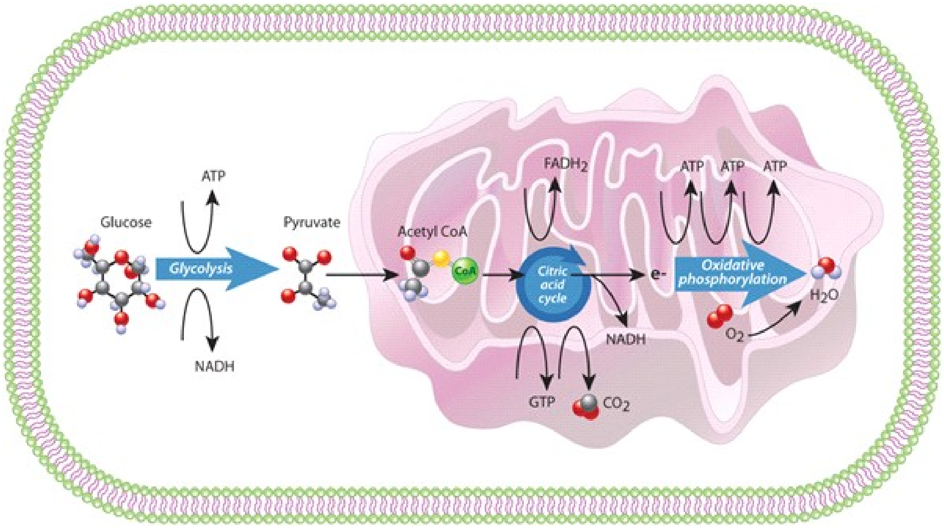
Glycolysis is an ancient, major ATP-producing pathway that occurs in almost all cells, eukaryotes and prokaryotes alike. This process, which is also known as fermentation, takes place in the cytoplasm and does not require oxygen. However, the fate of the pyruvate produced during glycolysis depends upon whether oxygen is present. In the absence of oxygen, the pyruvate cannot be completely oxidized to carbon dioxide, so various intermediate products result. For example, when oxygen levels are low, skeletal muscle cells rely on glycolysis to meet their intense energy requirements. This reliance on glycolysis results in the buildup of an intermediate known as lactic acid, which can cause a person's muscles to feel as if they are "on fire." Similarly, yeast, which is a single-celled eukaryote, produces alcohol (instead of carbon dioxide) in oxygen-deficient settings.

In contrast, when oxygen is available, the pyruvates produced by glycolysis become the input for the next portion of the eukaryotic energy pathway. During this stage, each pyruvate molecule in the cytoplasm enters the mitochondrion, where it is converted into acetyl CoA, a two-carbon energy carrier, and its third carbon combines with oxygen and is released as carbon dioxide. At the same time, an NADH carrier is also generated. Acetyl CoA then enters a pathway called the citric acid cycle, which is the second major energy process used by cells. The eight-step citric acid cycle generates three more NADH molecules and two other carrier molecules: FADH2 and GTP (Figure 6, middle).

The third major process in the eukaryotic energy pathway involves an electron transport chain, catalyzed by several protein complexes located in the mitochondrional inner membrane. This process, called oxidative phosphorylation, transfers electrons from NADH and FADH2 through the membrane protein complexes, and ultimately to oxygen, where they combine to form water. As electrons travel through the protein complexes in the chain, a gradient of hydrogen ions, or protons, forms across the mitochondrial membrane. Cells harness the energy of this proton gradient to create three additional ATP molecules for every electron that travels along the chain. Overall, the combination of the citric acid cycle and oxidative phosphorylation yields much more energy than fermentation - 15 times as much energy per glucose molecule! Together, these processes that occur inside the mitochondion, the citric acid cycle and oxidative phosphorylation, are referred to as respiration, a term used for processes that couple the uptake of oxygen and the production of carbon dioxide (Figure 6).

The electron transport chain in the mitochondrial membrane is not the only one that generates energy in living cells. In plant and other photosynthetic cells, chloroplasts also have an electron transport chain that harvests solar energy. Even though they do not contain mithcondria or chloroplatss, prokaryotes have other kinds of energy-yielding electron transport chains within their plasma membranes that also generate energy.

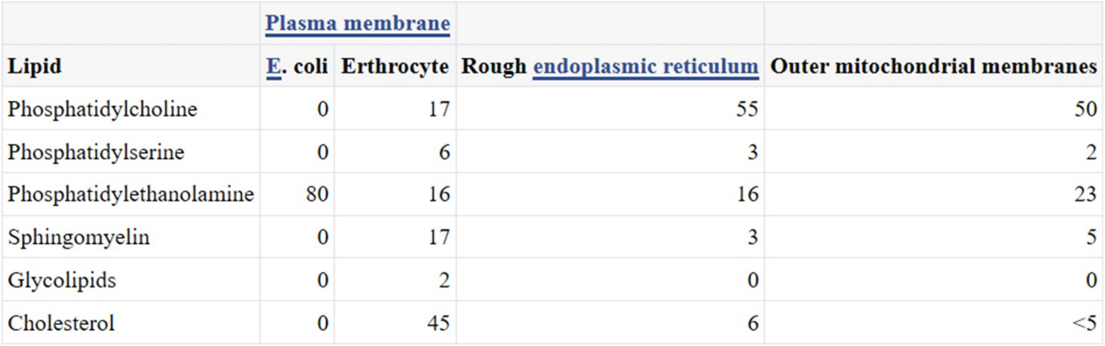




Membrane Formation

The formation of biological membranes is based on the properties of lipids, and all cell membranes share a common structural organization: bilayers of phospholipids with associated proteins. These membrane proteins are responsible for many specialized functions; some act as receptors that allow the cell to respond to external signals, some are responsible for the selective transport of molecules across the membrane, and others participate in electron transport and oxidative phosphorylation. In addition, membrane proteins control the interactions between cells of multicellular organisms. Lipids constitute approximately 50% of the mass of most cell membranes, although this proportion varies depending on the type of membrane. Plasma membranes, for example, are approximately 50% lipid and 50% protein. The inner membrane of mitochondria, on the other hand, contains an unusually high fraction (about 75%) of protein, reflecting the abundance of protein complexes involved in electron transport and oxidative phosphorylation. The lipid composition of different cell membranes also varies (Table 2.3). The plasma membrane of E. coli consists predominantly of phosphatidylethanolamine, which constitutes 80% of total lipid. Mammalian plasma membranes are more complex, containing four major phospholipids—phosphatidylcholine, phosphatidylserine, phosphatidylethanolamine, and sphingomyelin—which together constitute 50 to 60% of total membrane lipid. In addition to the phospholipids, the plasma membranes of animal cells contain glycolipids and cholesterol, which generally correspond to about 40% of the total lipid molecules.

Lipid composition of a cell membrane



	Plasma membrane			
Lipid	**E. coli**	**Erythrocyte**	**Rough endoplasmic reticulum**	**Outer mitochondrial membranes**
Phosphatidylcholine	0	17	55	50
Phosphatidylserine	0	6	3	2
Phosphatidylethanolamine	80	16	16	23
Sphingomyelin	0	17	3	5
Glycolipids	0	2	0	0
Cholesterol	0	45	6	<5

*Source:*Data from P. L. Yeagle, 1993. *The Membranes of Cells*, 2nd ed. San Diego, CA: Academic Press.

An important property of lipid bilayers is that they behave as two-dimensional fluids in which individual molecules (both lipids and proteins) are free to rotate and move in lateral directions (Figure 2.46). Such fluidity is a critical property of membranes and is determined by both temperature and lipid composition. For example, the interactions between shorter fatty acid chains are weaker than those between longer chains, so membranes containing shorter fatty acid chains are less rigid and remain fluid at lower temperatures. Lipids containing unsaturated fatty acids similarly increase membrane fluidity because the presence of double bonds introduces kinks in the fatty acid chains, making them more difficult to pack together. Because of its hydrocarbon ring structure (see Figure 2.9), cholesterol plays a distinct role in determining membrane fluidity. Cholesterol molecules insert into the bilayer with their polar hydroxyl groups close to the hydrophilic head groups of the phospholipids (Figure 2.47). The rigid hydrocarbon rings of cholesterol therefore interact with the regions of the fatty acid chains that are adjacent to the phospholipid head groups. This interaction decreases the mobility of the outer portions of the fatty acid chains, making this part of the membrane more rigid. On the other hand, insertion of cholesterol interferes with interactions between fatty acid chains, thereby maintaining membrane fluidity at lower temperatures.

Proteins are the other major constituent of cell membranes, constituting 25 to 75% of the mass of the various membranes of the cell. The current model of membrane structure, proposed by Jonathan Singer and Garth Nicolson in 1972, views membranes as a fluid mosaic in which proteins are inserted into a lipid bilayer. While phospholipids provide the basic structural organization of membranes, membrane proteins carry out the specific functions of the different membranes of the cell. These proteins are divided into two general classes, based on the nature of their association with the membrane. Integral membrane proteins are embedded directly within the lipid bilayer. Peripheral membrane proteins are not inserted into the lipid bilayer but are associated with the membrane indirectly, generally by interactions with integral membrane proteins. Many integral membrane proteins (called transmembrane proteins) span the lipid bilayer, with portions exposed on both sides of the membrane. The membrane-spanning portions of these proteins are usually α-helical regions of 20 to 25 nonpolar amino acids. The hydrophobic side chains of these amino acids interact with the fatty acid chains of membrane lipids, and the formation of an α helix neutralizes the polar character of the peptide bonds, as discussed earlier in this chapter with respect to protein folding. Like the phospholipids, transmembrane proteins are amphipathic molecules, with their hydrophilic portions exposed to the aqueous environment on both sides of the membrane. Some transmembrane proteins span the membrane only once; others have multiple membrane-spanning regions. Most transmembrane proteins of eukaryotic plasma membranes have been modified by the addition of carbohydrates, which are exposed on the surface of the cell and may participate in cell-cell interactions. Proteins can also be anchored in membranes by lipids that are covalently attached to the polypeptide chain. Distinct lipid modifications anchor proteins to the cytosolic and extracellular faces of the plasma membrane. Proteins can be anchored to the cytosolic face of the membrane either by the addition of a 14-carbon fatty acid (myristic acid) to their amino terminus or by the addition of either a 16-carbon fatty acid (palmitic acid) or 15- or 20-carbon prenyl groups to the side chains of cysteine residues. Alternatively, proteins are anchored to the extracellular face of the plasma membrane by the addition of glycolipids to their carboxy terminus.

The selective permeability of biological membranes to small molecules allows the cell to control and maintain its internal composition. Only small uncharged molecules can diffuse freely through phospholipid bilayers. Small nonpolar molecules, such as O_2 and CO_2, are soluble in the lipid bilayer and therefore can readily cross cell membranes. Small uncharged polar molecules, such as H_2O, also can diffuse through membranes, but larger uncharged polar molecules, such as glucose, cannot. Charged molecules, such as ions, are unable to diffuse through a phospholipid bilayer regardless of size; even H+ ions cannot cross a lipid bilayer by free diffusion. Although ions and most polar molecules cannot diffuse across a lipid bilayer, many such molecules (such as glucose) are able to cross cell membranes. These molecules pass across membranes via the action of specific transmembrane proteins, which act as transporters. Such transport proteins determine the selective permeability of cell membranes and thus play a critical role in membrane function. They contain multiple membrane-spanning regions that form a passage through the lipid bilayer, allowing polar or charged molecules to cross the membrane through a protein pore without interacting with the hydrophobic fatty acid chains of the membrane phospholipids. There are two general classes of membrane transport proteins. In contrast to channel proteins, carrier proteins selectively bind and transport specific small molecules, such as glucose. Rather than forming open channels, carrier proteins act like enzymes to facilitate the passage of specific molecules across membranes. In particular, carrier proteins bind specific molecules and then undergo conformational changes that open channels through which the molecule to be transported can pass across the membrane and be released on the other side.

As described so far, molecules transported by either channel or carrier proteins cross membranes in the energetically favorable direction, as determined by concentration and electrochemical gradients—a process known as passive transport. However, carrier proteins also provide a mechanism through which the energy changes associated with transporting molecules across a membrane can be coupled to the use or production of other forms of metabolic energy, just as enzymatic reactions can be coupled to the hydrolysis or synthesis of ATP. For example, molecules can be transported in an energetically unfavorable direction across a membrane (e.g., against a concentration gradient) if their transport in that direction is coupled to ATP hydrolysis as a source of energy—a process called active transport. The free energy stored as ATP can thus be used to control the internal composition of the cell, as well as to drive the biosynthesis of cell constituents.

Enzyme as Biological Catalyst

A fundamental task of proteins is to act as enzymes—catalysts that increase the rate of virtually all the chemical reactions within cells. Although RNAs are capable of catalyzing some reactions, most biological reactions are catalyzed by proteins. In the absence of enzymatic catalysis, most biochemical reactions are so slow that they would not occur under the mild conditions of temperature and pressure that are compatible with life. Enzymes accelerate the rates of such reactions by well over a million-fold, so reactions that would take years in the absence of catalysis can occur in fractions of seconds if catalyzed by the appropriate enzyme. Cells contain thousands of different enzymes, and their activities determine which of the many possible chemical reactions actually take place within the cell.

Like all other catalysts, enzymes are characterized by two fundamental properties. First, they increase the rate of chemical reactions without themselves being consumed or permanently altered by the reaction. Second, they increase reaction rates without altering the chemical equilibrium between reactants and products.

These principles of enzymatic catalysis are illustrated in the following example, in which a molecule acted upon by an enzyme (referred to as a substrate [S]) is converted to a product (P) as the result of the reaction. In the absence of the enzyme, the reaction can be written as follows:

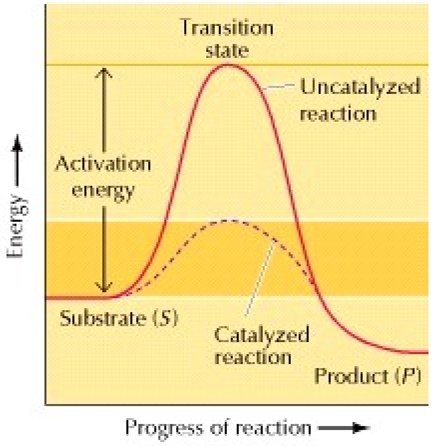
$$S \rightleftharpoons P$$

The chemical equilibrium between S and P is determined by the laws of thermodynamics (as discussed further in the next section of this chapter) and is represented by the ratio of the forward and reverse reaction rates ($S \rightarrow P$ and $P \rightarrow S$, respectively). In the presence of the appropriate enzyme, the conversion of S to P is accelerated, but the equilibrium between S and P is unaltered. Therefore, the enzyme must accelerate both the forward and reverse reactions equally. The reaction can be written as follows:

$$S \overset{E}{\rightleftharpoons} P$$

Note that the enzyme (E) is not altered by the reaction, so the chemical equilibrium remains unchanged, determined solely by the thermodynamic properties of S and P.

The effect of the enzyme on such a reaction is best illustrated by the energy changes that must occur during the conversion of S to P. The equilibrium of the reaction is determined by the final energy states of S and P, which are unaffected by enzymatic catalysis. In order for the reaction to proceed, however, the substrate must first be converted to a higher energy state, called the transition state. The energy required to reach the transition state (the activation energy) constitutes a barrier to the progress of the reaction, limiting the rate of the reaction. Enzymes (and other catalysts) act by reducing the activation energy, thereby increasing the rate of reaction. The increased rate is the same in both the forward and reverse directions, since both must pass through the same transition state.

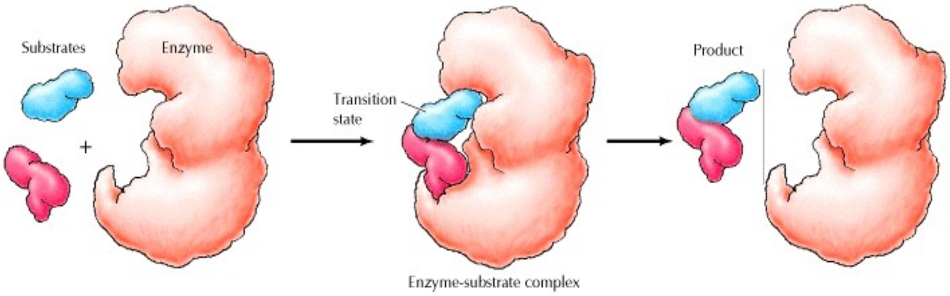


The catalytic activity of enzymes involves the binding of their substrates to form an enzyme-substrate complex (ES). The substrate binds to a specific region of the enzyme, called the active site. While bound to the active site, the substrate is converted into the product of the reaction, which is then released from the enzyme. The enzyme-catalyzed reaction can thus be written as follows:

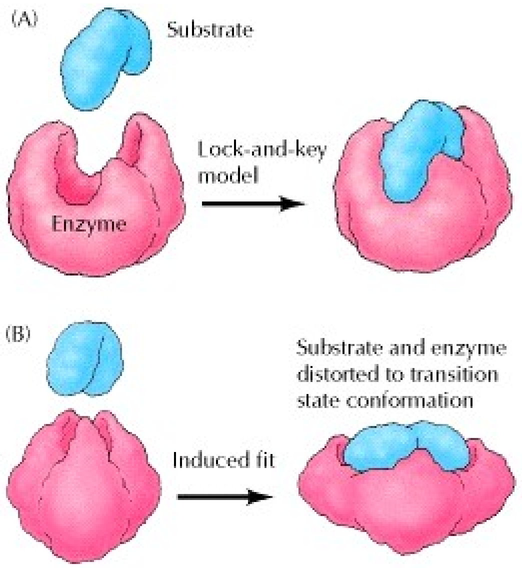
$$S + E \rightleftharpoons ES \rightleftharpoons E + P$$

Note that E appears unaltered on both sides of the equation, so the equilibrium is unaffected. However, the enzyme provides a surface upon which the reactions converting S to P can occur more readily. This is a result of interactions between the enzyme and substrate that lower the energy of activation and favor formation of the transition state.

Mechanism of enzymatic catalysis



Enzymes accelerate reactions also by altering the conformation of their substrates to approach that of the transition state. The simplest model of enzyme-substrate interaction is the lock-and-key model, in which the substrate fits precisely into the active site. In many cases, however, the configurations of both the enzyme and substrate are modified by substrate binding—a process called induced fit. In such cases the conformation of the substrate is altered so that it more closely resembles that of the transition state. The stress produced by such distortion of the substrate can further facilitate its conversion to the transition state by weakening critical bonds. Moreover, the transition state is stabilized by its tight binding to the enzyme, thereby lowering the required energy of activation.



In addition to bringing multiple substrates together and distorting the conformation of substrates to approach the transition state, many enzymes participate directly in the catalytic process. In such cases, specific amino acid side chains in the active site may react with the substrate and form bonds with reaction intermediates. The acidic and basic amino acids are often involved in these catalytic mechanisms, as illustrated in the following discussion of chymotrypsin as an example of enzymatic catalysis.

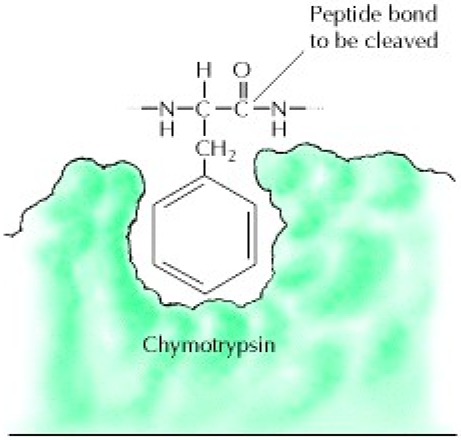
Chymotrypsin is a member of a family of enzymes (serine proteases) that digest proteins by catalyzing the hydrolysis of peptide bonds. The reaction can be written as follows:

$$\text{Protein} + H_2O \rightarrow \text{Peptide}_1 + \text{Peptide}_2$$

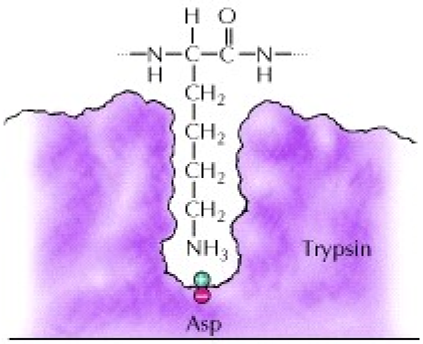
The different members of the serine protease family (including chymotrypsin, trypsin, elastase, and thrombin) have distinct substrate specificities; they preferentially cleave peptide bonds adjacent to different amino acids. For example, whereas chymotrypsin digests bonds adjacent to hydrophobic amino acids, such as tryptophan and phenylalanine, trypsin digests bonds next to basic amino acids, such as lysine and arginine. All the serine proteases, however, are similar in structure and use the same mechanism of catalysis. The active sites of these enzymes contain three critical amino acids—serine, histidine, and aspartate—that drive hydrolysis of the peptide bond. Indeed, these enzymes are called serine proteases because of the central role of the serine residue.

Substrates bind to the serine proteases by insertion of the amino acid adjacent to the cleavage site into a pocket at the active site of the enzyme. The nature of this pocket determines the substrate specificity of the different members of the serine protease family. For example, the binding pocket of chymotrypsin contains hydrophobic amino acids that interact with the hydrophobic side chains of its preferred substrates. In contrast, the binding pocket of trypsin contains a negatively charged acidic amino acid (aspartate), which is able to form an ionic bond with the lysine or arginine residues of its substrates.

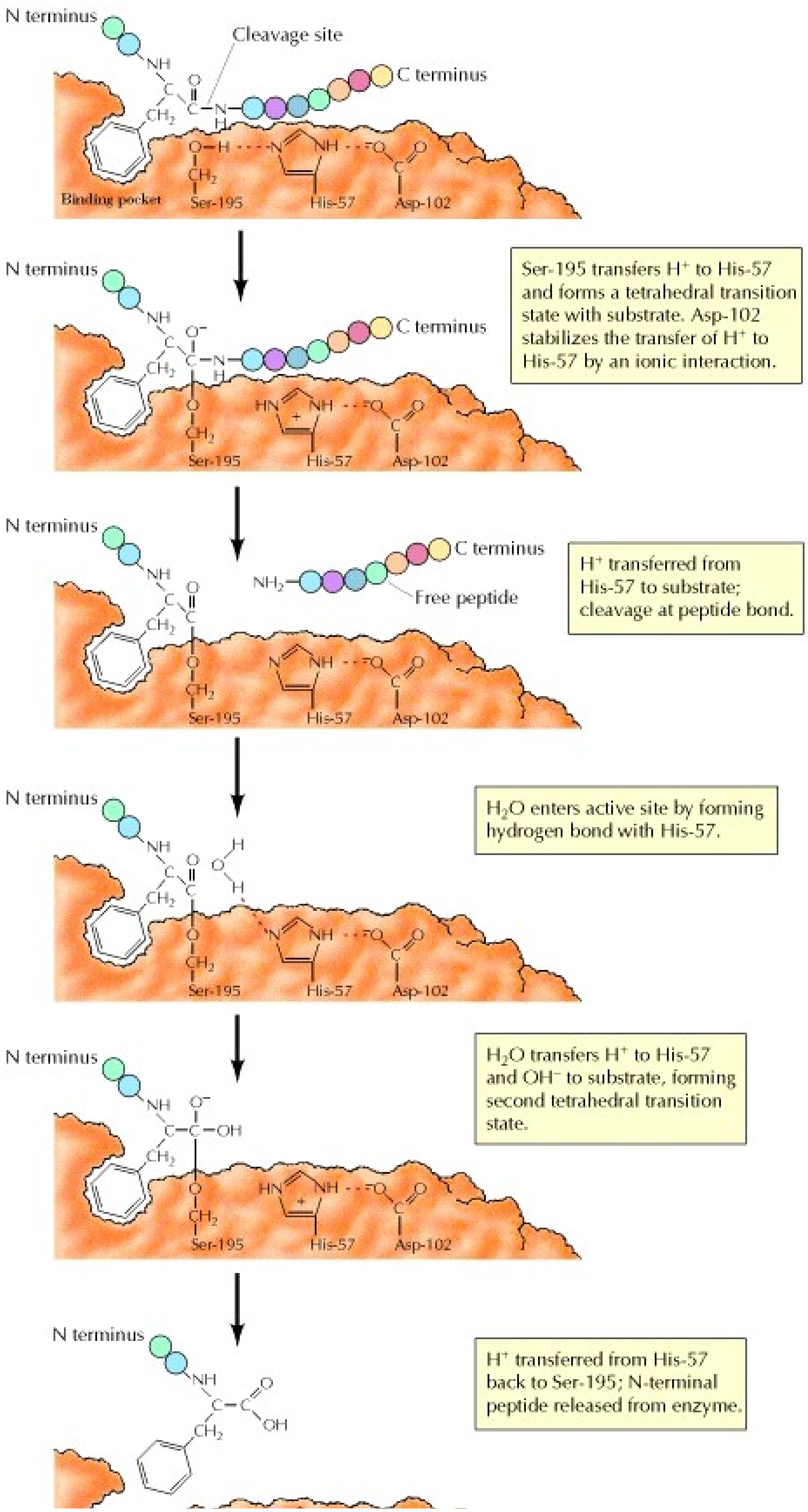
Substrate binding positions the peptide bond to be cleaved adjacent to the active site serine. The proton of this serine is then transferred to the active site histidine. The conformation of the active site favors this proton transfer because the histidine interacts with the negatively charged aspartate residue. The serine reacts with the substrate, forming a tetrahedral transition state. The peptide bond is then cleaved, and the C-terminal portion of the substrate is released from the enzyme. However, the N-terminal peptide remains bound to serine. This situation is resolved when a water molecule (the second substrate) enters the active site and reverses the preceding reactions. The proton of the water molecule is transferred to histidine, and its hydroxyl group is transferred to the peptide, forming a second tetrahedral transition state. The proton is then transferred from histidine back to serine, and the peptide is released from the enzyme, completing the reaction.



Hydrophobic interaction



Ionic interaction



Ser-195 transfers H^+ to His-57 and forms a tetrahedral transition state with substrate. Asp-102 stabilizes the transfer of H^+ to His-57 by an ionic interaction.

H^+ transferred from His-57 to substrate; cleavage at peptide bond.

H_2O enters active site by forming hydrogen bond with His-57.

H_2O transfers H^+ to His-57 and OH^- to substrate, forming second tetrahedral transition state.

H^+ transferred from His-57 back to Ser-195; N-terminal peptide released from enzyme.

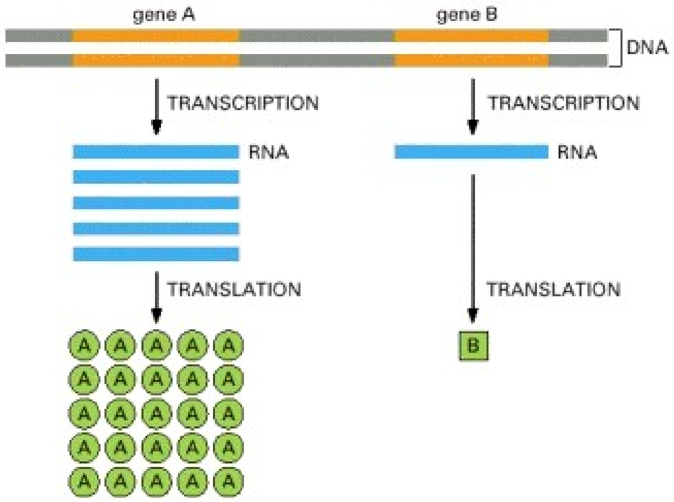
This example illustrates several features of enzymatic catalysis; the specificity of enzyme-substrate interactions, the positioning of different substrate molecules in the active site, and the involvement of active-site residues in the formation and stabilization of the transition state. Although the thousands of enzymes in cells catalyze many different types of chemical reactions, the same basic principles apply to their operation.

DNA Transcription

Transcription is the process by which the information in a strand of DNA is copied into a new molecule of messenger RNA (mRNA). DNA safely and stably stores genetic material in the nuclei of cells as a reference, or template. Meanwhile, mRNA is comparable to a copy from a reference book because it carries the same information as DNA but is not used for long-term storage and can freely exit the nucleus. Although the mRNA contains the same information, it is not an identical copy of the DNA segment, because its sequence is complementary to the DNA template.

Transcription is carried out by an enzyme called RNA polymerase and a number of accessory proteins called transcription factors. Transcription factors can bind to specific DNA sequences called enhancer and promoter sequences in order to recruit RNA polymerase to an appropriate transcription site. Together, the transcription factors and RNA polymerase form a complex called the transcription initiation complex. This complex initiates transcription, and the RNA polymerase begins mRNA synthesis by matching complementary bases to the original DNA strand. The mRNA molecule is elongated and, once the strand is completely synthesized, transcription is terminated. The newly formed mRNA copies of the gene then serve as blueprints for protein synthesis during the process of translation.

Because many identical RNA copies can be made from the same gene, and each RNA molecule can direct the synthesis of many identical protein molecules, cells can synthesize a large amount of protein rapidly when necessary. But each gene can also be transcribed and translated with a different efficiency, allowing the cell to make vast quantities of some proteins and tiny quantities of others.



Genes can be expressed with different efficiencies. Gene A is transcribed and translated much more efficiently than gene B. This allows the amount of protein A in the cell to be much greater than that of protein B.

Gene control regions

Start site. A start site for transcription.

Promoter. A region a few hundred nucleotides 'upstream' of the gene (toward the 5' end). It is not transcribed into mRNA, but plays a role in controlling the transcription of the gene. Transcription factors bind to specific nucleotide sequences in the promoter region and assist in the binding of RNA polymerases.

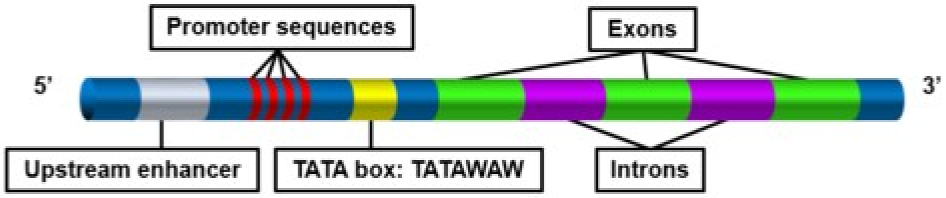
Enhancers. Some transcription factors (called activators) bind to regions called 'enhancers' that increase the rate of transcription. These sites may be thousands of nucleotides from the coding sequences or within an intron. Some enhancers are conditional and only work in the presence of other factors as well as transcription factors.

Silencers. Some transcription factors (called repressors) bind to regions called 'silencers' that depress the rate of transcription.

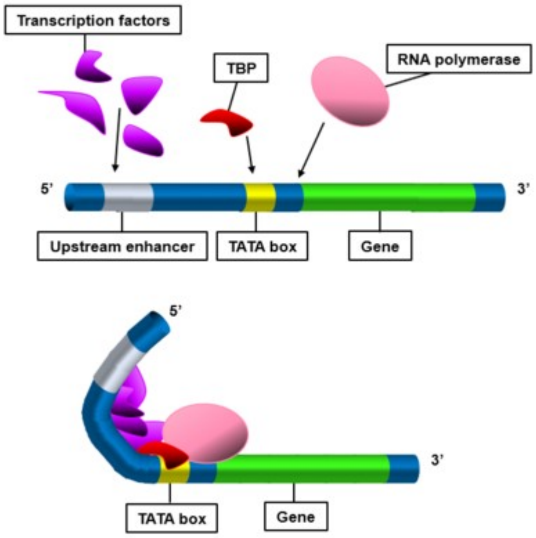
There are three major steps in DNA transcription: **Initiation, elongation and termination.**

Initiation

In molecular biology, the TATA box (also called the Goldberg-Hogness box) is a sequence of DNA found in the core promoter region of genes in archaea and eukaryotes. The prokaryotic homolog of the TATA box is called the Pribnow box which has a shorter consensus sequence. The TATA box is considered a non-coding DNA sequence (also known as a cis-regulatory element). It was termed the "TATA box" as it contains a consensus sequence characterized by repeating T and A base pairs.



The TATA-box is the site of preinitiation complex formation, which is the first step in transcription initiation in eukaryotes. Formation of the preinitiation complex begins when the multi-subunit transcription factor II D (TFIID) binds to the TATA box at its TATA-binding protein (TBP) subunit. The conformational changes induced by TBP binding to the TATA box allows for additional transcription factors and RNA polymerase II to bind to the promoter region. TFIID first binds to the TATA box, facilitated by TFIIA binding to the upstream part of the TFIID complex.[22][23] TFIIB then binds to the TFIID-TFIIA-DNA complex through interactions both upstream and downstream of the TATA box. RNA polymerase II is then recruited to this multi-protein complex with the help of TFIIF. Additional transcription factors then bind, first TFIIE and then TFIIH. This completes the assembly of the preinitiation complex for eukaryotic transcription.



Elongation

RNA polymerase moves along the template strand, synthesising an mRNA molecule. In prokaryotes RNA polymerase is a holoenzyme consisting of a number of subunits, including a sigma factor (transcription factor) that recognises the promoter. In eukaryotes there are three RNA polymerases: I, II and III. The process includes a proofreading mechanism.

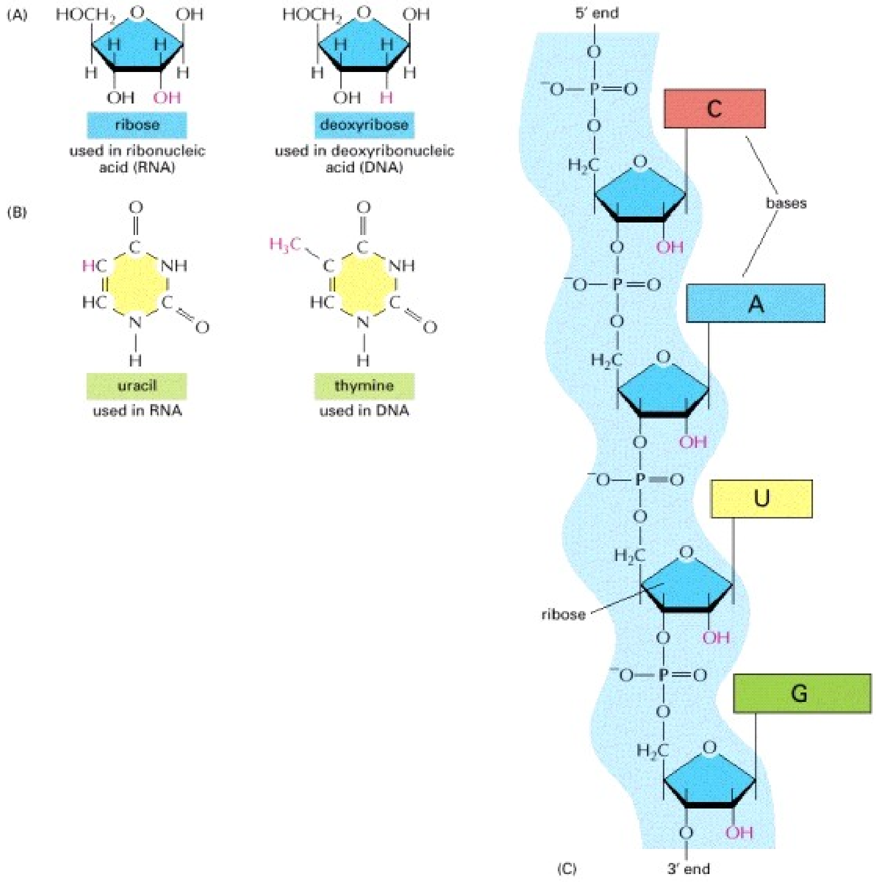
Termination

In prokaryotes there are two ways in which transcription is terminated. In Rho-dependent termination, a protein factor called "Rho" is responsible for disrupting the complex involving the template strand, RNA polymerase and RNA molecule. In Rho-independent termination, a loop forms at the end of the RNA molecule, causing it to detach itself. Termination in eukaryotes is more complicated, involving the addition of additional adenine nucleotides at the 3' of the RNA transcript (a process referred to as polyadenylation).

Processing

After transcription the RNA molecule is processed in a number of ways: introns are removed and the exons are spliced together to form a mature mRNA molecule consisting of a single protein-coding sequence. RNA synthesis involves the normal base pairing rules, but the base thymine is replaced with the base uracil.

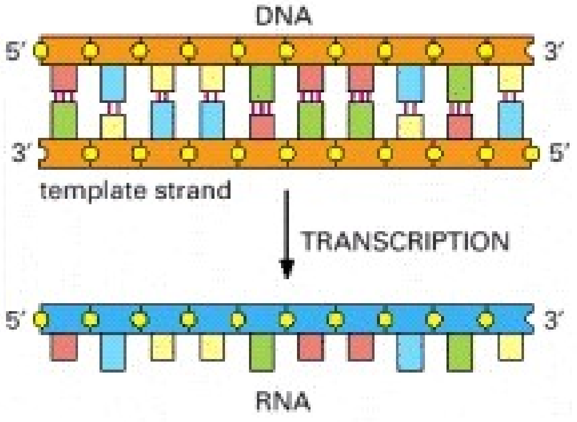
The first step a cell takes in reading out a needed part of its genetic instructions is to copy a particular portion of its DNA nucleotide sequence—a gene—into an RNA nucleotide sequence. The information in RNA, although copied into another chemical form, is still written in essentially the same language as it is in DNA—the language of a nucleotide sequence. Hence the name transcription. Like DNA, RNA is a linear polymer made of four different types of nucleotide subunits linked together by phosphodiester bonds (Figure 6-4). It differs from DNA chemically in two respects: (1) the nucleotides in RNA are ribonucleotides—that is, they contain the sugar ribose (hence the name ribonucleic acid) rather than deoxyribose; (2) although, like DNA, RNA contains the bases adenine (A), guanine (G), and cytosine (C), it contains the base uracil (U) instead of the thymine (T) in DNA. Since U, like T, can base-pair by hydrogen-bonding with A (Figure 6-5), the complementary base-pairing properties described for DNA in Chapters 4 and 5 apply also to RNA (in RNA, G pairs with C, and A pairs with U). It is not uncommon, however, to find other types of base pairs in RNA: for example, G pairing with U occasionally.



Despite these small chemical differences, DNA and RNA differ quite dramatically in overall structure. Whereas DNA always occurs in cells as a double-stranded helix, RNA is single-stranded. RNA chains therefore fold up into a variety of shapes, just as a polypeptide chain folds up to form the final shape of a protein (Figure 6-6). As we see later in this chapter, the ability to fold into complex three-dimensional shapes allows some RNA molecules to have structural and catalytic functions.

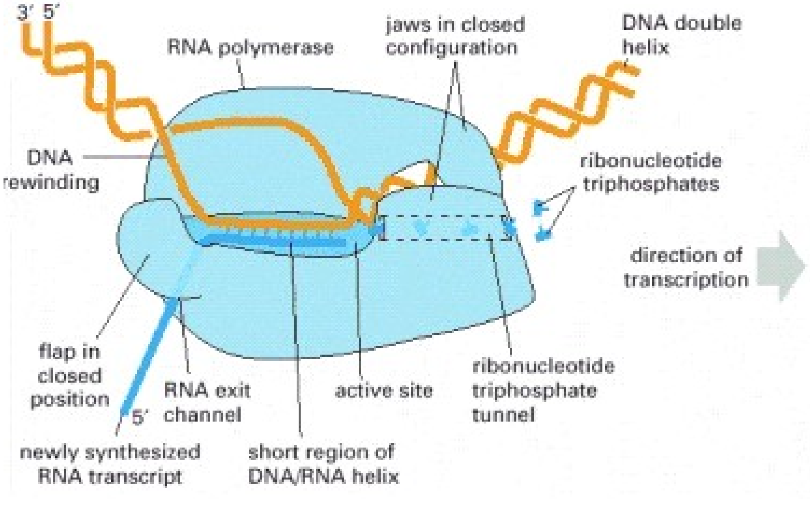
Transcription Produces RNA Complementary to One Strand of DNA

All of the RNA in a cell is made by DNA transcription, a process that has certain similarities to the process of DNA replication discussed in Chapter 5. Transcription begins with the opening and unwinding of a small portion of the DNA double helix to expose the bases on each DNA strand. One of the two strands of the DNA double helix then acts as a template for the synthesis of an RNA molecule. As in DNA replication, the nucleotide sequence of the RNA chain is determined by the complementary base-pairing between incoming nucleotides and the DNA template. When a good match is made, the incoming ribonucleotide is covalently linked to the growing RNA chain in an enzymatically catalyzed reaction. The RNA chain produced by transcription—the *transcript*—is therefore elongated one nucleotide at a time, and it has a nucleotide sequence that is exactly complementary to the strand of DNA used as the template.



Transcription, however, differs from DNA replication in several crucial ways. Unlike a newly formed DNA strand, the RNA strand does not remain hydrogen-bonded to the DNA template strand. Instead, just behind the region where the ribonucleotides are being added, the RNA chain is displaced and the DNA helix re-forms. Thus, the RNA molecules produced by transcription are released from the DNA template as single strands. In addition, because they are copied from only a limited region of the DNA, RNA molecules are much shorter than DNA molecules. A DNA molecule in a human chromosome can be up to 250 million nucleotide-pairs long; in contrast, most RNAs are no more than a few thousand nucleotides long, and many are considerably shorter.

The enzymes that perform transcription are called RNA polymerases. Like the DNA polymerase that catalyzes DNA replication (discussed in Chapter 5), RNA polymerases catalyze the formation of the phosphodiester bonds that link the nucleotides together to form a linear chain. The RNA polymerase moves stepwise along the DNA, unwinding the DNA helix just ahead of the active site for polymerization to expose a new region of the template strand for complementary base-pairing. In this way, the growing RNA chain is extended by one nucleotide at a time in the 5′-to-3′ direction (Figure 6-8). The substrates are nucleoside triphosphates (ATP, CTP, UTP, and GTP); as for DNA replication, a hydrolysis of high-energy bonds provides the energy needed to drive the reaction forward.



DNA is transcribed by the enzyme RNA polymerase

The RNA polymerase (pale blue) moves stepwise along the DNA, unwinding the DNA helix at its active site. As it progresses, the polymerase adds nucleotides (here, small “T” shapes) one by one to the RNA chain at the polymerization site using an exposed DNA strand as a template. The RNA transcript is thus a single-stranded complementary copy of one of the two DNA strands. The polymerase has a rudder (see Figure 6-11) that displaces the newly formed RNA, allowing the two strands of DNA behind the polymerase to rewind. A short region of DNA/RNA helix (approximately nine nucleotides in length) is therefore formed only transiently, and a “window” of DNA/RNA helix therefore moves along the DNA with the polymerase. The incoming nucleotides are in the form of ribonucleoside triphosphates (ATP, UTP, CTP, and GTP), and the energy stored in their phosphate-phosphate bonds provides the driving force for the polymerization reaction

The almost immediate release of the RNA strand from the DNA as it is synthesized means that many RNA copies can be made from the same gene in a relatively short time, the synthesis of additional RNA molecules being started before the first RNA is completed (Figure 6-9). When RNA polymerase molecules follow hard on each other's heels in this way, each moving at about 20 nucleotides per second (the speed in eucaryotes), over a thousand transcripts can be synthesized in an hour from a single gene.

Although RNA polymerase catalyzes essentially the same chemical reaction as DNA polymerase, there are some important differences between the two enzymes. First, and most obvious, RNA polymerase catalyzes the linkage of ribonucleotides, not deoxyribonucleotides. Second, unlike the DNA polymerases involved in DNA replication, RNA polymerases can start an RNA chain without a primer. This difference may exist because transcription need not be as accurate as DNA replication. Unlike DNA, RNA does not permanently store genetic information in cells. RNA polymerases make about one mistake for every 104 nucleotides copied into RNA (compared with an error rate for direct copying by DNA polymerase of about one in 107 nucleotides), and the consequences of an error in RNA transcription are much less significant than that in DNA replication.

Although RNA polymerases are not nearly as accurate as the DNA polymerases that replicate DNA, they nonetheless have a modest proofreading mechanism. If the incorrect ribonucleotide is added to the growing RNA chain, the polymerase can back up, and the active site of the enzyme can perform an excision reaction that mimics the reverse of the polymerization reaction, except that water instead of pyrophosphate is used RNA polymerase hovers around a misincorporated ribonucleotide longer than it does for a correct addition, causing excision to be favored for incorrect nucleotides. However, RNA polymerase also excises many correct bases as part of the cost for improved accuracy.

Cells Produce Several Types of RNA

The majority of genes carried in a cell's DNA specify the amino acid sequence of proteins; the RNA molecules that are copied from these genes (which ultimately direct the synthesis of proteins) are called messenger RNA (mRNA) molecules. The final product of a minority of genes, however, is the RNA itself. Careful analysis of the complete DNA sequence of the genome of the yeast S. cerevisiae has uncovered well over 750 genes (somewhat more than 10% of the total number of yeast genes) that produce RNA as their final product, although this number includes multiple copies of some highly repeated genes. These RNAs, like proteins, serve as enzymatic and structural components for a wide variety of processes in the cell. One of those RNAs, the template carried by the enzyme telomerase. Although not all of their functions are known, we see in this chapter that some small nuclear RNA (snRNA) molecules direct the splicing of pre-mRNA to form mRNA, that ribosomal RNA (rRNA) molecules form the core of ribosomes, and that transfer RNA (tRNA) molecules form the adaptors that select amino acids and hold them in place on a ribosome for incorporation into protein.

Principal types of RNA produced in cell

1. mRNA: messenger RNAs, code for proteins
2. rRNA: ribosomal RNAs, form the basic structure of the ribosome and catalyze protein synthesis
3. tRNA: transfer RNAs, central to protein synthesis as adaptors between mRNA and amino acids
4. snRNA: small nuclear RNAs, function in a variety of nuclear processes, including the splicing of pre-mRNA
5. snoRNA: small nucleolar RNAs, used to process and chemically modify rRNAs
6. Other non-coding RNA: function in diverse cellular processes, including telomere synthesis, X-chromosome inactivation, and the transport of proteins into the ER

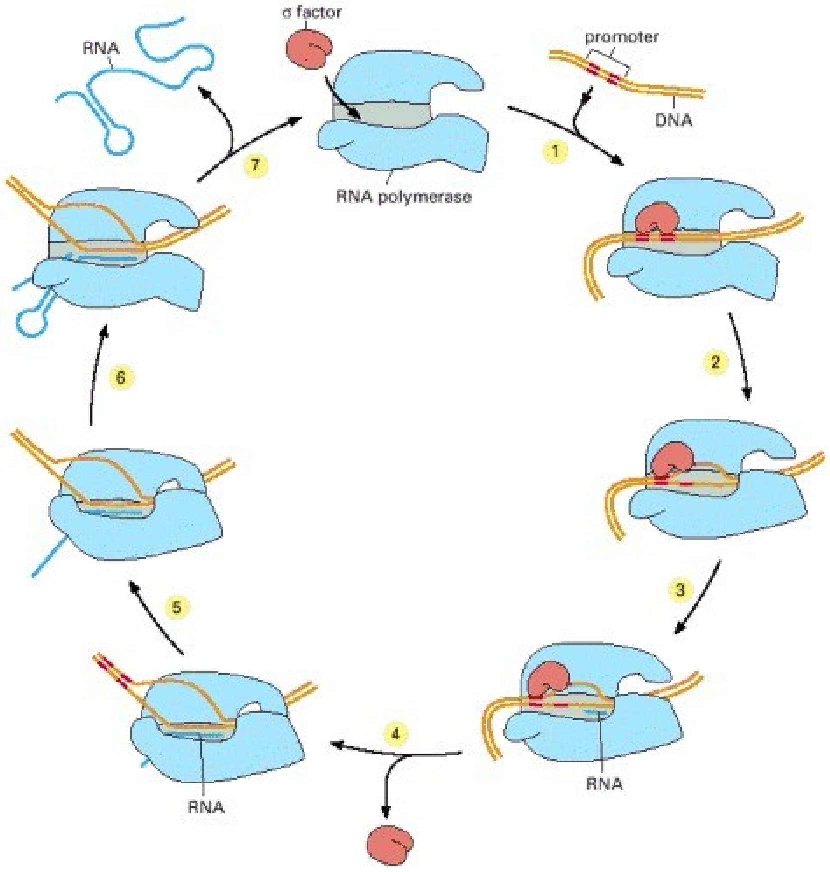
Each transcribed segment of DNA is called a *transcription unit*. In eucaryotes, a transcription unit typically carries the information of just one gene, and therefore codes for either a single RNA molecule or a single protein (or group of related proteins if the initial RNA transcript is spliced in more than one way to produce different mRNAs). In bacteria, a set of adjacent genes is often trans-cribed as a unit; the resulting mRNA molecule therefore carries the information for several distinct proteins. Overall, RNA makes up a few percent of a cell's dry weight. Most of the RNA in cells is rRNA; mRNA comprises only 3–5% of the total RNA in a typical mammalian cell. The mRNA population is made up of tens of thousands of different species, and there are on average only 10–15 molecules of each species of mRNA present in each cell.

Signals Encoded in DNA Tell RNA Polymerase Where to Start and Stop

To transcribe a gene accurately, RNA polymerase must recognize where on the genome to start and where to finish. The way in which RNA polymerases perform these tasks differs somewhat between bacteria and eucaryotes. Because the process in bacteria is simpler, we look there first.

The initiation of transcription is an especially important step in gene expression because it is the main point at which the cell regulates which proteins are to be produced and at what rate. Bacterial RNA polymerase is a multisubunit complex. A detachable subunit, called sigma (σ) factor, is largely responsible for its ability to read the signals in the DNA that tell it where to begin transcribing (Figure 6-10). RNA polymerase molecules adhere only weakly to the bacterial DNA when they collide with it, and a polymerase molecule typically slides rapidly along the long DNA molecule until it dissociates again. However, when the polymerase slides into a region on the DNA double helix called a promoter, a special sequence of nucleotides indicating the starting point for RNA synthesis, it binds tightly to it. The polymerase, using its σ factor, recognizes this DNA sequence by making specific contacts with the portions of the bases that are exposed on the outside of the helix (step-1).

The transcription cycle of bacteria RNA polymerase: In step 1, the RNA polymerase holoenzyme (core polymerase plus σ factor) forms and then locates a promoter. The polymerase unwinds the DNA at the position at which transcription is to begin (step 2) and begins transcribing (step 3). This initial RNA synthesis (sometimes called “abortive initiation”) is relatively inefficient. However, once RNA polymerase has managed to synthesize about 10 nucleotides of RNA, σ relaxes its grip, and the polymerase undergoes a series of conformational changes (which probably includes a tightening of its jaws and the placement of RNA in the exit channel). The polymerase now shifts to the elongation mode of RNA synthesis (step 4), moving rightwards along the DNA in this diagram. During the elongation mode (step 5) transcription is highly processive, with the polymerase leaving the DNA template and releasing the newly transcribed RNA only when it encounters a termination signal (step 6). Termination signals are encoded in DNA and many function by forming an RNA structure that destabilizes the polymerase's hold on the RNA, as shown here. In bacteria, all RNA molecules are synthesized by a single type of RNA polymerase and the cycle depicted in the figure therefore applies to the production of mRNAs as well as structural and catalytic RNAs.



After the RNA polymerase binds tightly to the promoter DNA in this way, it opens up the double helix to expose a short stretch of nucleotides on each strand *(Step 2*). Unlike a DNA helicase reaction, this limited opening of the helix does not require the energy of ATP hydrolysis. Instead, the polymerase and DNA both undergo reversible structural changes that result in a more energetically favorable state. With the DNA unwound, one of the two exposed DNA strands acts as a template for complementary base-pairing with incoming ribonucleotides, two of which are joined together by the polymerase to begin an RNA chain. After the first ten or so nucleotides of RNA have been synthesized (a relatively inefficient process during which polymerase synthesizes and discards short nucleotide oligomers), the σ factor relaxes its tight hold on the polymerase and evenutally dissociates from it. During this process, the polymerase undergoes additional structural changes that enable it to move forward rapidly, transcribing without the σ factor *(Step 4*). Chain elongation continues (at a speed of approximately 50 nucleotides/sec for bacterial RNA polymerases) until the enzyme encounters a second signal in the DNA, the terminator (described below), where the polymerase halts and releases both the DNA template and the newly made RNA chain

(Step 7). After the polymerase has been released at a terminator, it reassociates with a free σ factor and searches for a new promoter, where it can begin the process of transcription again.

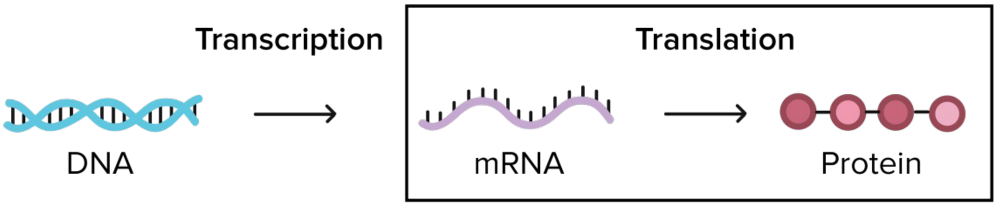
Several structural features of bacterial RNA polymerase make it particularly adept at performing the transcription cycle just described. Once the σ factor positions the polymerase on the promoter and the template DNA has been unwound and pushed to the active site, a pair of moveable jaws is thought to clamp onto the DNA. When the first 10 nucleotides have been transcribed, the dissociation of σ allows a flap at the back of the polymerase to close to form an exit tunnel through which the newly made RNA leaves the enzyme. With the polymerase now functioning in its elongation mode, a rudder-like structure in the enzyme continuously pries apart the DNA-RNA hybrid formed. We can view the series of conformational changes that takes place during transcription initiation as a successive tightening of the enzyme around the DNA and RNA to ensure that it does not dissociate before it has finished transcribing a gene. If an RNA polymerase does dissociate prematurely, it cannot resume synthesis but must start over again at the promoter. How do the signals in the DNA (termination signals) stop the elongating polymerase? For most bacterial genes a termination signal consists of a string of A-T nucleotide pairs preceded by a two-fold symmetric DNA sequence, which, when transcribed into RNA, folds into a “hairpin” structure through Watson-Crick base-pairing. As the polymerase transcribes across a terminator, the hairpin may help to wedge open the movable flap on the RNA polymerase and release the RNA transcript from the exit tunnel. At the same time, the DNA-RNA hybrid in the active site, which is held together predominantly by U-A base pairs (which are less stable than G-C base pairs because they form two rather than three hydrogen bonds per base pair), is not sufficiently strong enough to hold the RNA in place, and it dissociates causing the release of the polymerase from the DNA, perhaps by forcing open its jaws. Thus, in some respects, transcription termination seems to involve a reversal of the structural transitions that happen during initiation. The process of termination also is an example of a common theme in this chapter: the ability of RNA to fold into specific structures figures prominently in many aspects of decoding the genome.

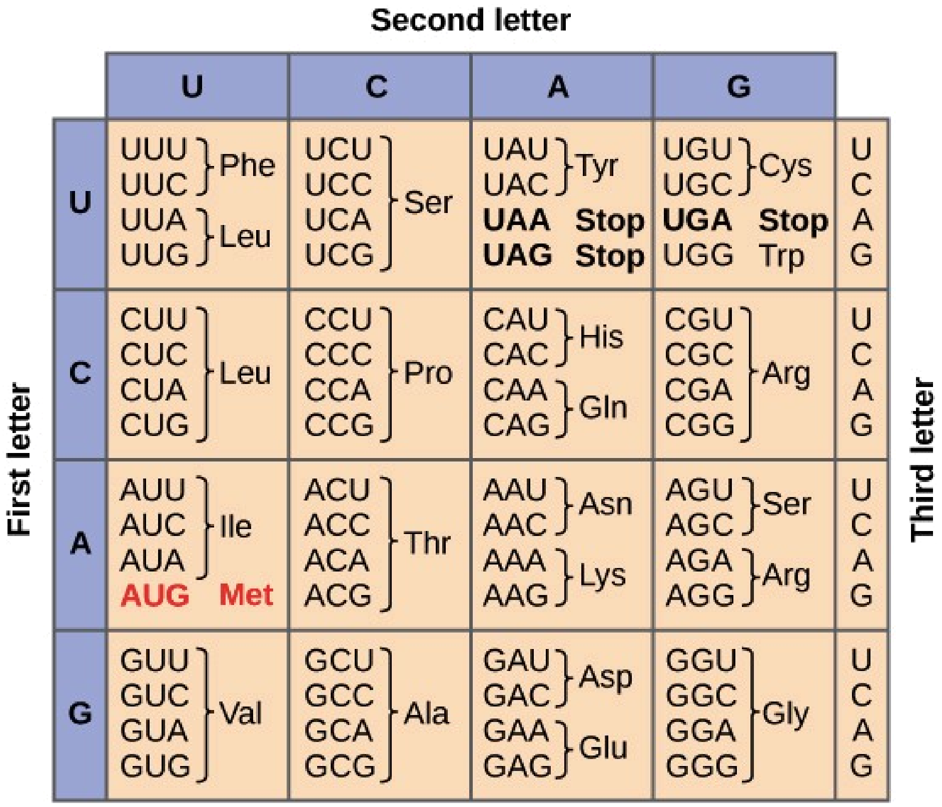
Translation in cell biology

Most genes in a cell produce mRNA molecules that serve as intermediaries on the pathway to proteins. In this section we examine how the cell converts the information carried in an mRNA molecule into a protein molecule. This feat of translation first attracted the attention of biologists in the late 1950s, when it was posed as the “coding problem”: how is the information in a linear sequence of nucleotides in RNA translated into the linear sequence of a chemically quite different set of subunits—the amino acids in proteins? This fascinating question stimulated great excitement among scientists at the time. Here was a cryptogram set up by nature that, after more than 3 billion years of evolution, could finally be solved by one of the products of evolution—human beings. And indeed, not only has the code been cracked step by step, but in the year 2000 the elaborate machinery by which cells read this code—the ribosome—was finally revealed in atomic detail.

Basically, a gene is used to build a protein in a two-step process:

- Step 1: transcription! Here, the DNA sequence of a gene is "rewritten" in the form of RNA. In eukaryotes like you and me, the RNA is processed (and often has a few bits snipped out of it) to make the final product, called a messenger RNA or mRNA.
- Step 2: translation! In this stage, the mRNA is "decoded" to build a protein (or a chunk/subunit of a protein) that contains a specific series of amino acids.





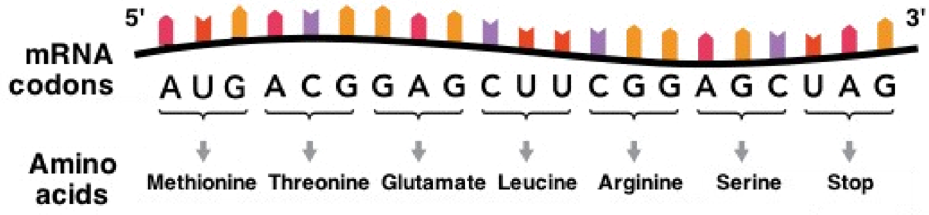
	Second letter				
First letter	U	C	A	G	Third letter
U	UUU UUC } Phe UUA UUG } Leu	UCU UCC UCA UCG } Ser	UAU UAC } Tyr **UAA Stop** **UAG Stop**	UGU UGC } Cys **UGA Stop** UGG Trp	U C A G
C	CUU CUC CUA CUG } Leu	CCU CCC CCA CCG } Pro	CAU CAC } His CAA CAG } Gln	CGU CGC CGA CGG } Arg	U C A G
A	AUU AUC AUA } Ile **AUG Met**	ACU ACC ACA ACG } Thr	AAU AAC } Asn AAA AAG } Lys	AGU AGC } Ser AGA AGG } Arg	U C A G
G	GUU GUC GUA GUG } Val	GCU GCC GCA GCG } Ala	GAU GAC } Asp GAA GAG } Glu	GGU GGC GGA GGG } Gly	U C A G

The genetic code

During translation, a cell “reads” the information in a messenger RNA (mRNA) and uses it to build a protein. Actually, to be a little more techical, an mRNA doesn’t always encode—provide instructions for—a whole protein. Instead, what we can confidently say is that it always encodes a polypeptide, or chain of amino acids. In an mRNA, the instructions for building a polypeptide are RNA nucleotides (As, Us, Cs, and Gs) read in groups of three. These groups of three are called codons.

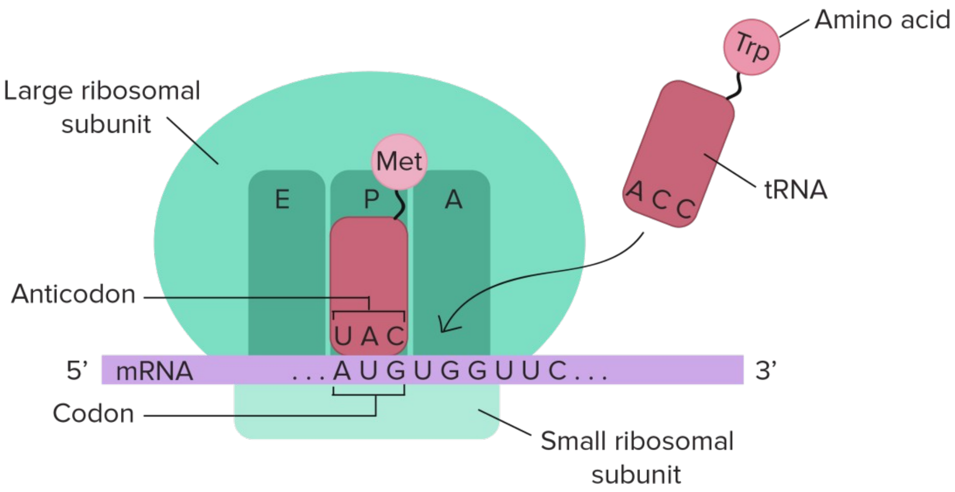
There are 61 codons for amino acids, and each of them is "read" to specify a certain amino acid out of the 20 commonly found in proteins. One codon, AUG, specifies the amino acid methionine and also acts as a start codon to signal the start of protein construction.

There are three more codons that do not specify amino acids. These stop codons, UAA, UAG, and UGA, tell the cell when a polypeptide is complete. All together, this collection of codon-amino acid relationships is called the genetic code, because it lets cells “decode” an mRNA into a chain of amino acids.



How is an mRNA "read" to make a polypeptide? Two types of molecules with key roles in translation are tRNAs and ribosomes.

Transfer RNAs, or tRNAs, are molecular "bridges" that connect mRNA codons to the amino acids they encode. One end of each tRNA has a sequence of three nucleotides called an anticodon, which can bind to specific mRNA codons. The other end of the tRNA carries the amino acid specified by the codons. There are many different types of tRNAs. Each type reads one or a few codons and brings the right amino acid matching those codons.



Ribosomes are the structures where polypeptides (proteins) are built. They are made up of protein and RNA (ribosomal RNA, or rRNA). Each ribosome has two subunits, a large one and a small one, which come together around an mRNA—kind of like the two halves of a hamburger bun coming together around the patty. The ribosome provides a set of handy slots where tRNAs can find their matching codons on the mRNA template and deliver their amino acids. These slots are called the A, P, and E sites. Not only that, but the ribosome also acts as an enzyme, catalyzing the chemical reaction that links amino acids together to make a chain.

Steps of translation

Your cells are making new proteins every second of the day. And each of those proteins must contain the right set of amino acids, linked together in just the right order. That may sound like a challenging task, but luckily, your cells (along with those of other animals, plants, and bacteria) are up to the job.

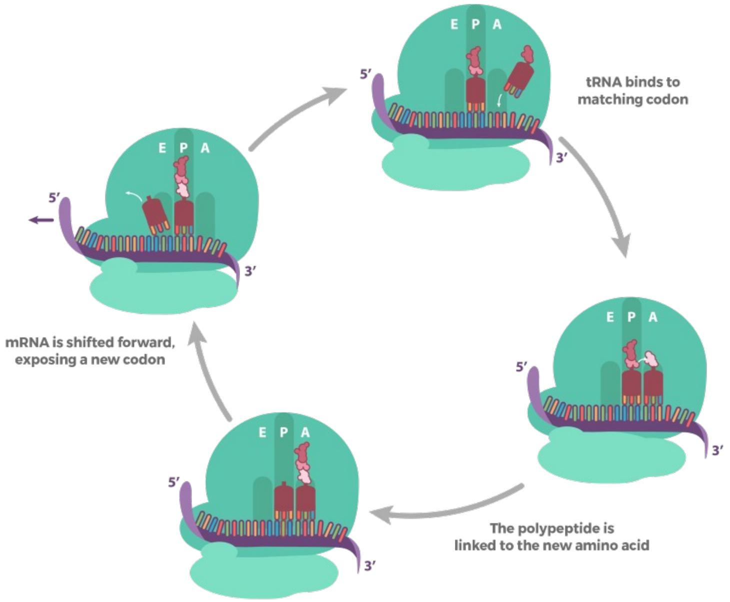
To see how cells make proteins, let's divide translation into three stages: initiation (starting off), elongation (adding on to the protein chain), and termination (finishing up).

In initiation, the ribosome assembles around the mRNA to be read and the first tRNA (carrying the amino acid methionine, which matches the start codon, AUG). This setup, called the initiation complex, is needed in order for translation to get started.

Elongation is the stage where the amino acid chain gets longer. In elongation, the mRNA is read one codon at a time, and the amino acid matching each codon is added to a growing protein chain.

Each time a new codon is exposed:

- A matching tRNA binds to the codon
- The existing amino acid chain (polypeptide) is linked onto the amino acid of the tRNA via a chemical reaction
- The mRNA is shifted one codon over in the ribosome, exposing a new codon for reading.

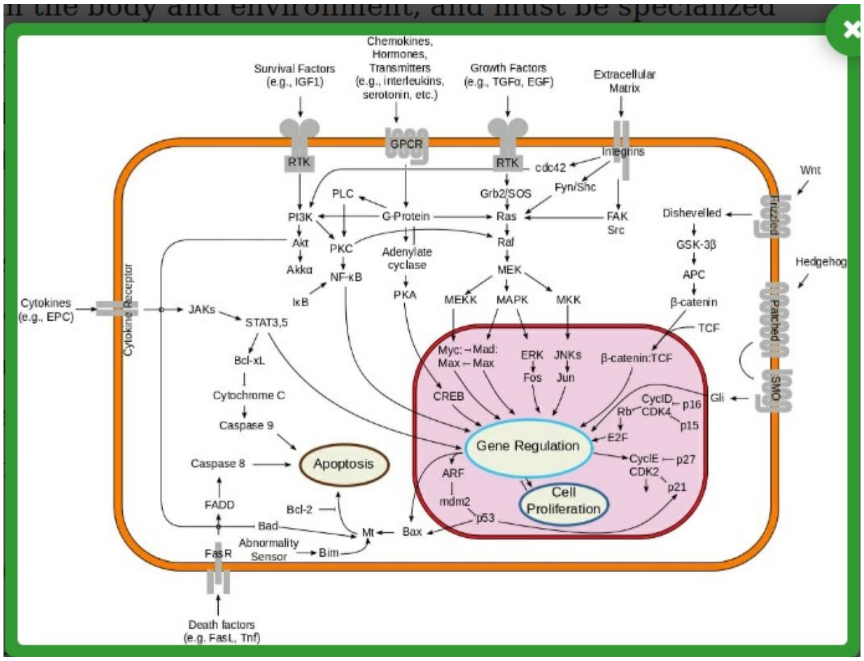


During elongation, tRNAs move through the A, P, and E sites of the ribosome, as shown above. This process repeats many times as new codons are read and new amino acids are added to the chain.

Termination is the stage in which the finished polypeptide chain is released. It begins when a stop codon (UAG, UAA, or UGA) enters the ribosome, triggering a series of events that separate the chain from its tRNA and allow it to drift out of the ribosome.

After termination, the polypeptide may still need to fold into the right 3D shape, undergo processing (such as the removal of amino acids), get shipped to the right place in the cell, or combine with other polypeptides before it can do its job as a functional protein.

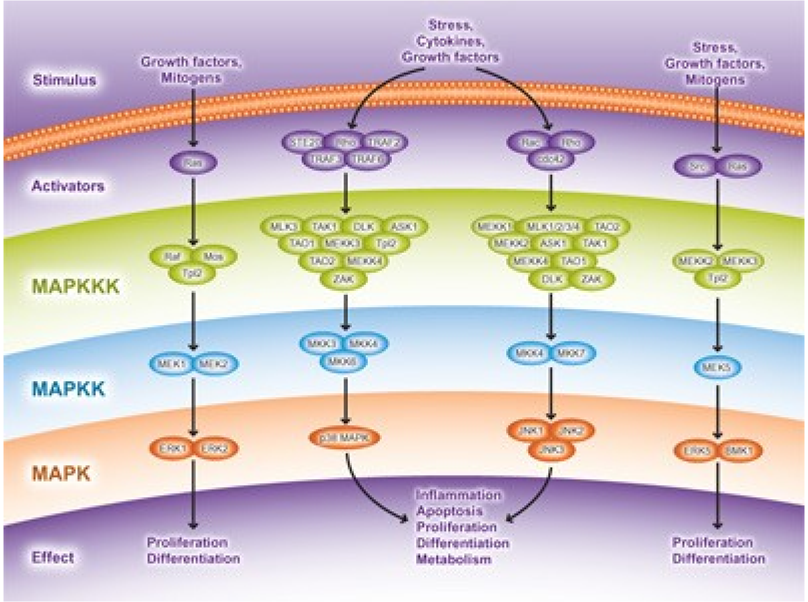
Cell Signaling Technology



Do not be intimidated by the above diagram of signaling pathways: its just to show you how complex and innovative cell signaling techniques can be. However, the basic mechanism is very simple: it consists of four steps- 1. Signal Sending (ligands) 2. Signal Reception (binding with protein acting as receptors) 3. Signal Transduction (conformational change and phosphorylation cascade) followed by 4. enzyme activation and starting of physiological process.

The following examples will clarify your concept of intracellular signaling system:

1. MAPK pathway:



Mitogen-activated protein kinases (MAPK) are a family serine/threonine kinases which function in tiered protein kinase cascades. They regulate proliferation, apoptosis, differentiation and survival - over and under-activation of this family of proteins have been linked to malignancy, neurodegeneration and cardiovascular disease. Each MAPK can be activated by associated MAPK kinases (MAPKK) through dual phosphorylation which is subsequently activated by MAPKK kinases (MAPKKK).

Through this mechanism MAPK signaling can amplify initial stimuli and regulate a diverse range of responses. Specificity and complexity of the MAPK signaling pathway is regulated by a number of mechanisms:

Activation and substrate specificity

Every level of the MAPK signaling cascade is subject to target specificity. This allows amplification of precise downstream signaling and transcription factors.

MAPKKK - Different intracellular (DNA damage, oxident stress etc.) and extracellular (cytokines, Growth Factor (GF), stress etc.) signals activate specific MAPKKKs. Once activated, many MAPKKKs have broad specificity towards target MAPKK proteins, particularly between JNK and p38 signaling (Stress-activated protein kinases (SAPK)/Jun amino-terminal kinases (JNK) are members of the MAPK family and are activated by a variety of environmental stresses, inflam- matory cytokines, growth factors etc).

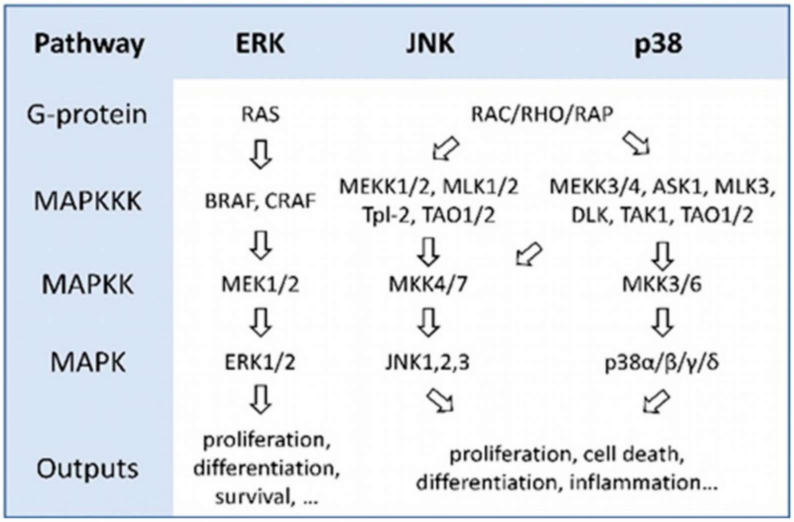
MAPKK - MAPKKs show increased specificity towards targets dependent on cell type and stimuli.

MAPK - Once phosphorylated each MAPK can activate a wide range of nuclear and cytoplasmic targets - over 100 have been identified for ERK (extracellular signal-regulated kinases, they are classical MAP kinases) signaling alone.

Tissue specific expression - All levels of MAPK signaling is subject to specific tissue distribution. These tissue distributions are a major form of signaling regulation ensuring the correct pathways are activated in specific areas.

Subcellular localization - Scaffold proteins regulate protein availability and subcellular localization which are important regulators of signal transduction and will heavily influence the final cellular outcome of MAPK activation.

Alternative activation - Activation of MAPK pathways does not always follow the traditional pattern. In some cases MAPK can be activated directly without MAPKKs and MAPKKKs.



What is "what"?

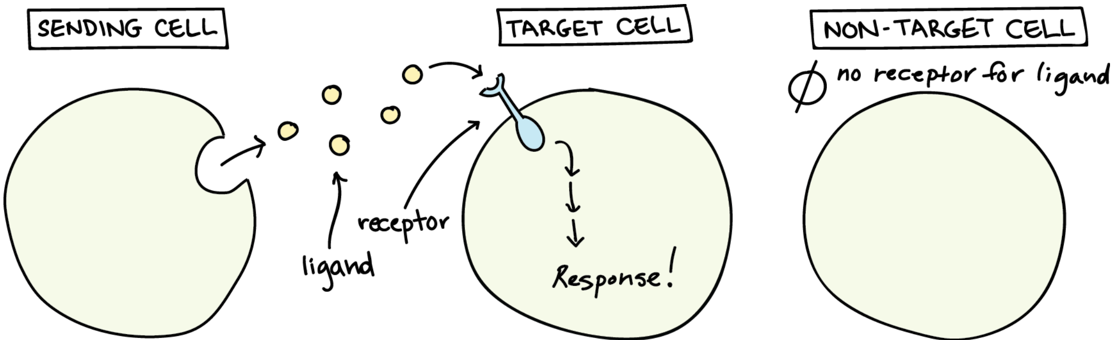
RAS/RAC/RHO/RAP: Ras is a family of related proteins which is expressed in all animal cell lineages and organs. All Ras protein family members belong to a class of protein called small GTPase, and are involved in transmitting signals within cells (cellular signal transduction). Small GTPases, also known as small G-proteins, are a family of hydrolase enzymes that can bind and hydrolyze guanosine triphosphate (GTP). They are a type of G-protein found in the cytosol. Because these signals result in cell growth and division, overactive Ras signaling can ultimately lead to cancer.

BRAF, CRAF: BRAF is a human gene that encodes a protein called B-Raf. RAF, and ERK (also known as MAPK) are both serine/threonine-selective protein kinases.

MEKK, MEK: They are proteins which participate in protein kinase phosphorylation cascade.

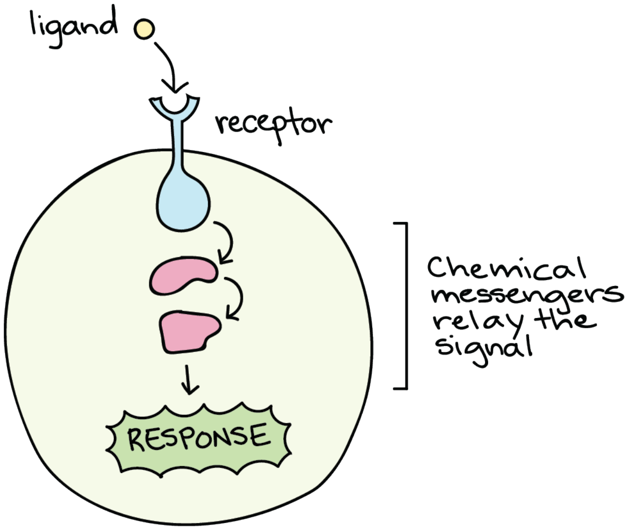
Overview of cell signaling

Cells typically communicate using chemical signals. These chemical signals, which are proteins or other molecules produced by a sending cell, are often secreted from the cell and released into the extracellular space. There, they can float – like messages in a bottle – over to neighboring cells.



Not all cells can “hear” a particular chemical message. In order to detect a signal (that is, to be a target cell), a neighbor cell must have the right receptor for that signal. When a signaling molecule binds to its receptor, it alters the shape or activity of the receptor, triggering a change inside of the cell. Signaling molecules are often called ligands, a general term for molecules that bind specifically to other molecules (such as receptors).

The message carried by a ligand is often relayed through a chain of chemical messengers inside the cell. Ultimately, it leads to a change in the cell, such as alteration in the activity of a gene or even the induction of a whole process, such as cell division. Thus, the original intercellular (between-cells) signal is converted into an intracellular (within-cell) signal that triggers a response.



Forms of signaling

Cell-cell signaling involves the transmission of a signal from a sending cell to a receiving cell. However, not all sending and receiving cells are next-door neighbors, nor do all cell pairs exchange signals in the same way.

There are four basic categories of chemical signaling found in multicellular organisms: paracrine signaling, autocrine signaling, endocrine signaling, and signaling by direct contact. The main difference between the different categories of signaling is the distance that the signal travels through the organism to reach the target cell.

Paracrine signaling

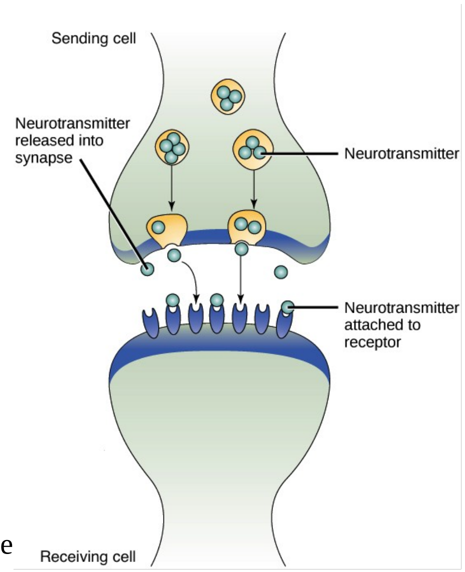
Often, cells that are near one another communicate through the release of chemical messengers (ligands that can diffuse through the space between the cells). This type of signaling, in which cells communicate over relatively short distances, is known as paracrine signaling.

Paracrine signaling allows cells to locally coordinate activities with their neighbors. Although they're used in many different tissues and contexts, paracrine signals are especially important during development, when they allow one group of cells to tell a neighboring group of cells what cellular identity to take on. [Example: spinal cord development]

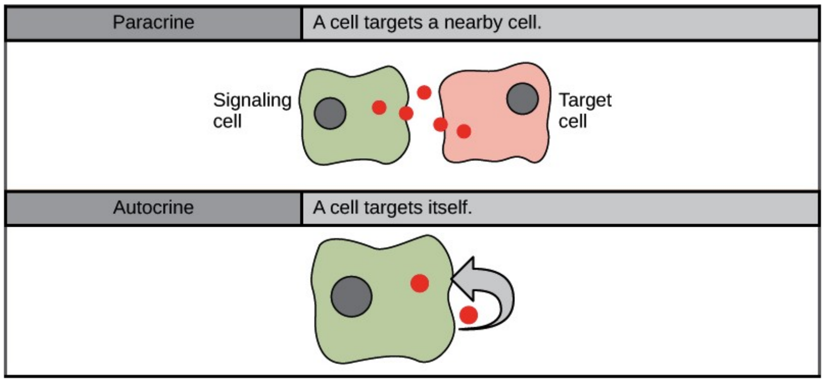
Synaptic signaling

One unique example of paracrine signaling is synaptic signaling, in which nerve cells transmit signals. This process is named for the synapse, the junction between two nerve cells where signal transmission occurs.

When the sending neuron fires, an electrical impulse moves rapidly through the cell, traveling down a long, fiber-like extension called an axon. When the impulse reaches the synapse, it triggers the release of ligands called neurotransmitters, which quickly cross the small gap between the nerve cells. When the neurotransmitters arrive at the receiving cell, they bind to receptors and cause a chemical change inside of the cell (often, opening ion channels and changing the electrical potential across the membrane).



The neurotransmitters that are released into the chemical synapse are quickly degraded or taken back up by the sending cell. This "resets" the system so they synapse is prepared to respond quickly to the next signal.



Autocrine signaling

In autocrine signaling, a cell signals to itself, releasing a ligand that binds to receptors on its own surface (or, depending on the type of signal, to receptors inside of the cell). This may seem like an odd thing for a cell to do, but autocrine signaling plays an important role in many processes.

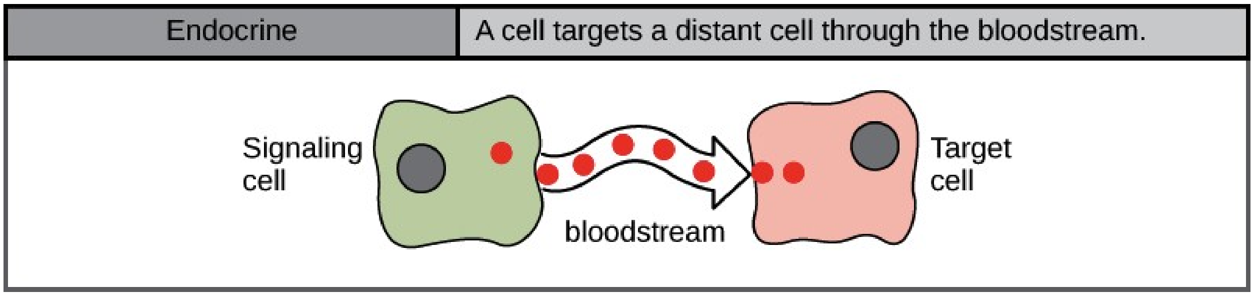
For instance, autocrine signaling is important during development, helping cells take on and reinforce their correct identities. From a medical standpoint, autocrine signaling is important in cancer and is thought to play a key role in metastasis (the spread of cancer from its original site to other parts of the body). In many cases, a signal may have both autocrine and paracrine effects, binding to the sending cell as well as other similar cells in the area.

Endocrine signaling

When cells need to transmit signals over long distances, they often use the circulatory system as a distribution network for the messages they send. In long-distance endocrine signaling, signals are produced by specialized cells and released into the bloodstream, which carries them to target cells in distant parts of the body. Signals that are produced in one part of the body and travel through the circulation to reach far-away targets are known as hormones.

In humans, endocrine glands that release hormones include the thyroid, the hypothalamus, and the pituitary, as well as the gonads (testes and ovaries) and the pancreas. Each endocrine gland releases one or more types of hormones, many of which are master regulators of development and physiology.

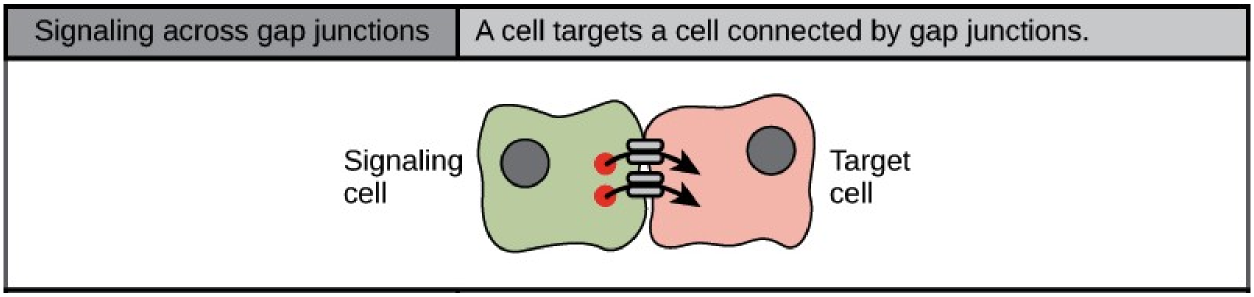
For example, the pituitary releases growth hormone (GH), which promotes growth, particularly of the skeleton and cartilage. Like most hormones, GH affects many different types of cells throughout the body. However, cartilage cells provide one example of how GH functions: it binds to receptors on the surface of these cells and encourages them to divide.



Direct Signaling

Gap junctions in animals and plasmodesmata in plants are tiny channels that directly connect neighboring cells. These water-filled channels allow small signaling molecules, called intracellular mediators, to diffuse between the two cells. Small molecules, such as calcium ions (Ca^{2+}), are able to move between cells, but large molecules like proteins and DNA cannot fit through the channels without special assistance.

The transfer of signaling molecules transmits the current state of one cell to its neighbor. This allows a group of cells to coordinate their response to a signal that only one of them may have received. In plants, there are plasmodesmata between almost all cells, making the entire plant into one giant network.



Signal Transmission Technique

Ligand and Receptor

Receptors and ligands come in many forms, but they all have one thing in common: they come in closely matched pairs, with a receptor recognizing just one (or a few) specific ligands, and a ligand binding to just one (or a few) target receptors. Binding of a ligand to a receptor changes its shape or activity, allowing it to transmit a signal or directly produce a change inside of the cell.

Types of receptors

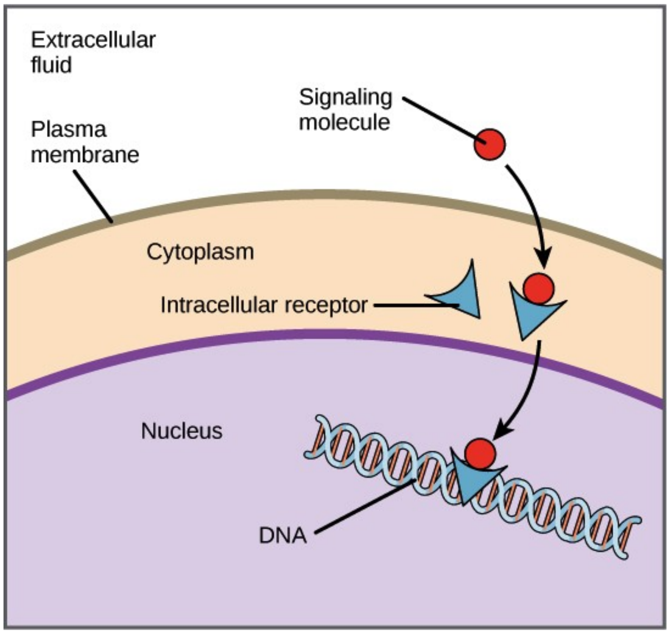
Receptors come in many types, but they can be divided into two categories: intracellular receptors, which are found inside of the cell (in the cytoplasm or nucleus), and cell surface receptors, which are found in the plasma membrane.

Intracellular receptors

Intracellular receptors are receptor proteins found on the inside of the cell, typically in the cytoplasm or nucleus. In most cases, the ligands of intracellular receptors are small, hydrophobic (water-hating) molecules, since they must be able to cross the plasma membrane in order to reach their receptors. For example, the primary receptors for hydrophobic steroid hormones, such as the sex hormones estradiol (an estrogen) and testosterone, are intracellular.

When a hormone enters a cell and binds to its receptor, it causes the receptor to change shape, allowing the receptor-hormone complex to enter the nucleus (if it wasn't there already) and regulate gene activity.

Hormone binding exposes regions of the receptor that have DNA-binding activity, meaning they can attach to specific sequences of DNA. These sequences are found next to certain genes in the DNA of the cell, and when the receptor binds next to these genes, it alters their level of transcription.



Many signaling pathways, involving both intracellular and cell surface receptors, cause changes in the transcription of genes. However, intracellular receptors are unique because they cause these changes very directly, binding to the DNA and altering transcription themselves.

Cell-surface receptors

Cell-surface receptors are membrane-anchored proteins that bind to ligands on the outside surface of the cell. In this type of signaling, the ligand does not need to cross the plasma membrane. So, many different kinds of molecules (including large, hydrophilic or "water-loving" ones) may act as ligands.

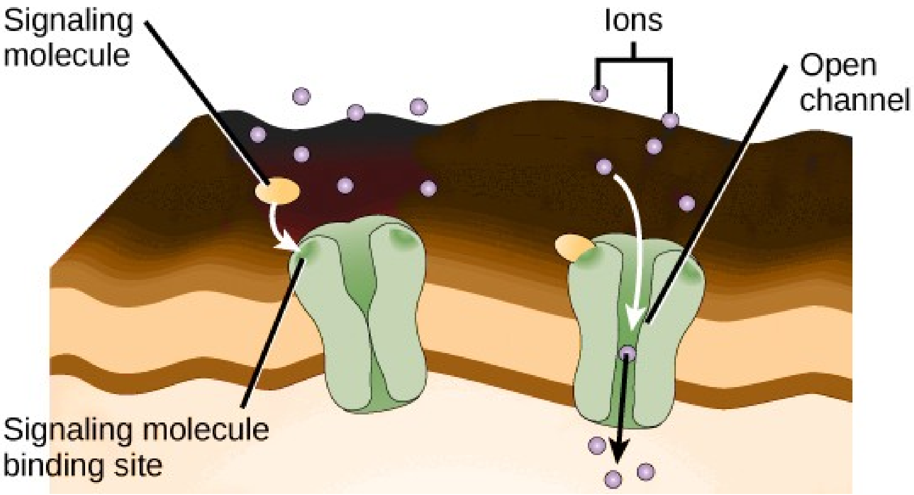
A typical cell-surface receptor has three different domains, or protein regions: a extracellular ("outside of cell") ligand-binding domain, a hydrophobic domain extending through the membrane, and an intracellular ("inside of cell") domain, which often transmits a signal. The size and structure of these regions can vary a lot depending on the type of receptor, and the hydrophobic region may consist of multiple stretches of amino acids that criss-cross the membrane.

There are many kinds of cell-surface receptors, but here we’ll look at three common types: ligand-gated ion channels, G protein-coupled receptors (GPCR), and receptor tyrosine kinases (RTK).

Ligand-gated ion channels

Ligand-gated ion channels are ion channels that can open in response to the binding of a ligand. To form a channel, this type of cell-surface receptor has a membrane-spanning region with a hydrophilic (water-loving) channel through the middle of it. The channel lets ions to cross the membrane without having to touch the hydrophobic core of the phospholipid bilayer.

When a ligand binds to the extracellular region of the channel, the protein's structure changes in such a way that ions of a particular type, such as Ca^{2+} or Cl^{-}, can pass through. In some cases, the reverse is actually true: the channel is usually open, and ligand binding causes it to close. Changes in ion levels inside the cell can change the activity of other molecules, such as ion-binding enzymes and voltage-sensitive channels, to produce a response. Neurons, or nerve cells, have ligand-gated channels that are bound by neurotransmitters.



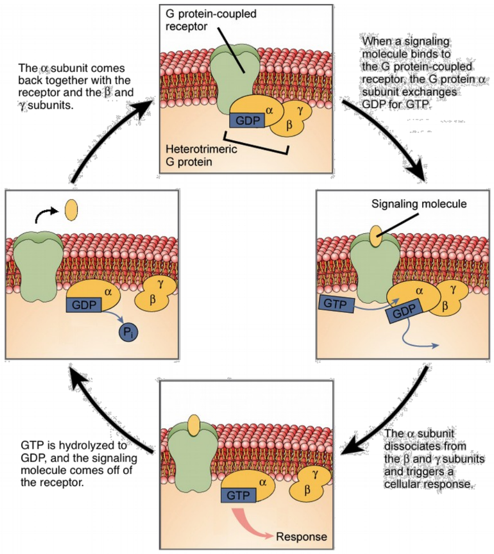
G protein-coupled receptors

G protein-coupled receptors (GPCRs) are a large family of cell surface receptors that share a common structure and method of signaling. The members of the GPCR family all have seven different protein segments that cross the membrane, and they transmit signals inside the cell through a type of protein called a G protein.

GPCRs are diverse and bind many different types of ligands. One particularly interesting class of GPCRs is the odorant (scent) receptors. There are about 800 of them in humans, and each binds its own "scent molecule" – such as a particular chemical in perfume, or a certain compound released by rotting fish – and causes a signal to be sent to the brain, making us smell a smell!

When its ligand is not present, a G protein-coupled receptor waits at the plasma membrane in an inactive state. For at least some types of GPCRs, the inactive receptor is already docked to its signaling target, a **G protein**.

G proteins come in different types, but they all bind the nucleotide guanosine triphosphate (GTP), which they can break down (hydrolyze) to form GDP. A G protein attached to GTP is active, or "on," while a G protein that's bound to GDP is inactive, or "off." The G proteins that associate with GPCRs are a type made up of three subunits, known as heterotrimeric G proteins. When they're attached to an inactive receptor, they're in the "off" form (bound to GDP).



Ligand binding, however, changes the picture: the GPCR is activated and causes the G protein to exchange GDP for GTP. The now-active G protein separates into two pieces (one called the α subunit, the other consisting of the β and γ subunits), which are freed from the GPCR. The subunits can interact with other proteins, triggering a signaling pathway that leads to a response.

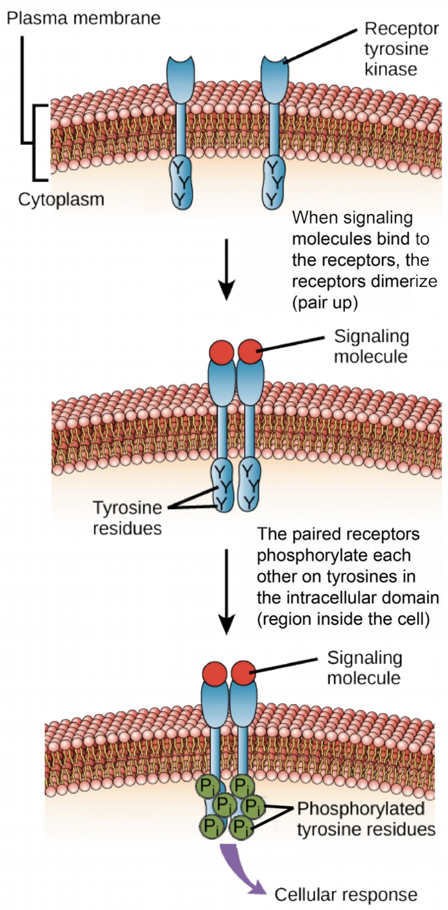
Eventually, the α subunit will hydrolyze GTP back to GDP, at which point the G protein becomes inactive. The inactive G protein reassembles as a three-piece unit associated with a GPCR. Cell signaling using G protein-coupled receptors is a cycle, one that can repeat over and over in response to ligand binding. G protein-coupled receptors play many different roles in the human body, and disruption of GPCR signaling can cause diseases like Cholera.

Receptor tyrosine kinases (RTK)

Enzyme-linked receptors are cell-surface receptors with intracellular domains that are associated with an enzyme. In some cases, the intracellular domain of the receptor actually is an enzyme that can catalyze a reaction. Other enzyme-linked receptors have an intracellular domain that interacts with an enzyme.

Receptor tyrosine kinases (RTKs) are a class of enzyme-linked receptors found in humans and many other species. A kinase is just a name for an enzyme that transfers phosphate groups to a protein or other target, and a receptor tyrosine kinase transfers phosphate groups specifically to the amino acid tyrosine.

How does RTK signaling work? In a typical example, signaling molecules first bind to the extracellular domains of two nearby receptor tyrosine kinases. The two neighboring receptors then come together, or dimerize. The receptors then attach phosphates to tyrosines in each others' intracellular domains. The phosphorylated tyrosine can transmit the signal to other molecules in the cell.



In many cases, the phosphorylated receptors serve as a docking platform for other proteins that contain special types of binding domains. A variety of proteins contain these domains, and when one of these proteins binds, it can initiate a downstream signaling cascade that leads to a cellular response.

Receptor tyrosine kinases are crucial to many signaling processes in humans. For instance, they bind to growth factors, signaling molecules that promote cell division and survival. Growth factors include platelet-derived growth factor (PDGF), which participates in wound healing, and nerve growth factor (NGF), which must be continually supplied to certain types of neurons to keep them alive. Because of their role in growth factor signaling, receptor tyrosine kinases are essential in the body, but their activity must be kept in balance: overactive growth factor receptors are associated with some types of cancers.

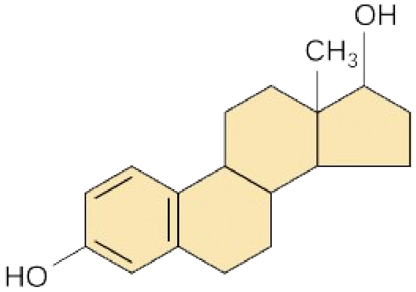
Types of ligands

Ligands, which are produced by signaling cells and interact with receptors in or on target cells, come in many different varieties. Some are proteins, others are hydrophobic molecules like steroids, and others yet are gases like nitric oxide. Here, we’ll look at some examples of different types of ligands.

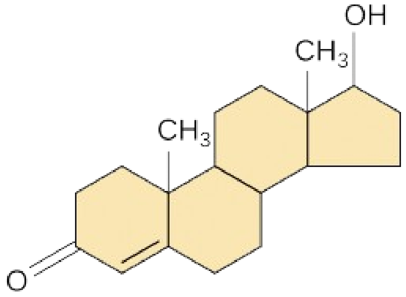
Ligands that can enter the cell

Small, hydrophobic ligands can pass through the plasma membrane and bind to intracellular receptors in the nucleus or cytoplasm. In the human body, some of the most important ligands of this type are the steroid hormones.

Familiar steroid hormones include the female sex hormone estradiol, which is a type of estrogen, and the male sex hormone testosterone. Vitamin D, a molecule synthesized in the skin using energy from light, is another example of a steroid hormone. Because they are hydrophobic, these hormones don't have trouble crossing the plasma membrane, but they must bind to carrier proteins in order to travel through the (watery) bloodstream.



Estradiol



Testosterone

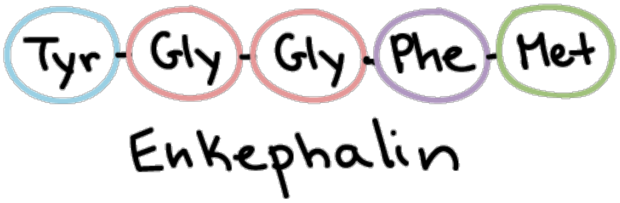
Nitric oxide (NO) is a gas that acts as a ligand. Like steroid hormones, it can diffuse directly across the plasma membrane thanks to its small size. One of its key roles is to activate a signaling pathway in the smooth muscle surrounding blood vessels, one that makes the muscle relax and allows the blood vessels to expand (dilate). In fact, the drug nitroglycerin treats heart disease by triggering the release of NO, dilating vessels to restore blood flow to the heart.

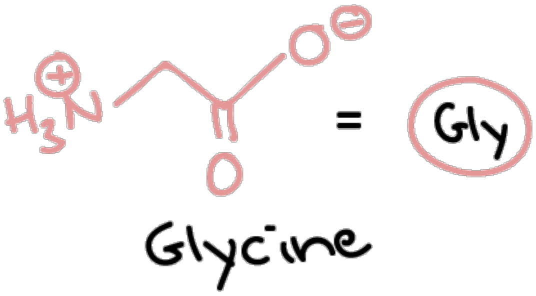
NO has become better-known recently because the pathway that it affects is targeted by prescription medications for erectile dysfunction, such as Viagra.

Ligands that bind on the outside of the cell

Water-soluble ligands are polar or charged and cannot readily cross the plasma membrane. So, most water-soluble ligands bind to the extracellular domains of cell-surface receptors, staying on the outer surface of the cell.

Peptide (protein) ligands make up the largest and most diverse class of water-soluble ligands. For instance, growth factors, hormones such as insulin, and certain neurotransmitters fall into this category. Peptide ligands can range from just a few amino acids long, as in the pain-suppressing enkephalins, to a hundred or more amino acids in length. As mentioned above, some neurotransmitters are proteins. Many other neurotransmitters, however, are small, hydrophilic (water-loving) organic molecules. Some neurotransmitters are standard amino acids, such as glutamate and glycine, and others are modified or non-standard amino acids.



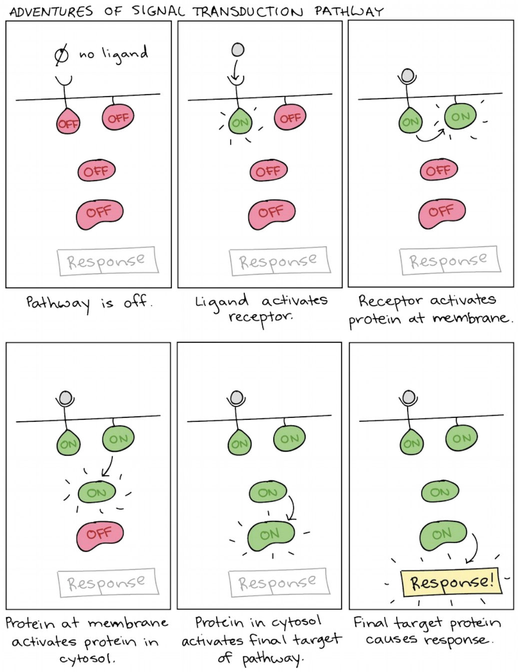


The chains of molecules that relay signals inside a cell are known as **intracellular signal transduction pathways.**

Binding initiates a signaling pathway

When a ligand binds to a cell-surface receptor, the receptor’s intracellular domain (part inside the cell) changes in some way. Generally, it takes on a new shape, which may make it active as an enzyme or let it bind other molecules.

The change in the receptor sets off a series of signaling events. For instance, the receptor may turn on another signaling molecule inside of the cell, which in turn activates its own target. This chain reaction can eventually lead to a change in the cell's behavior or characteristics, as shown in the cartoon below.

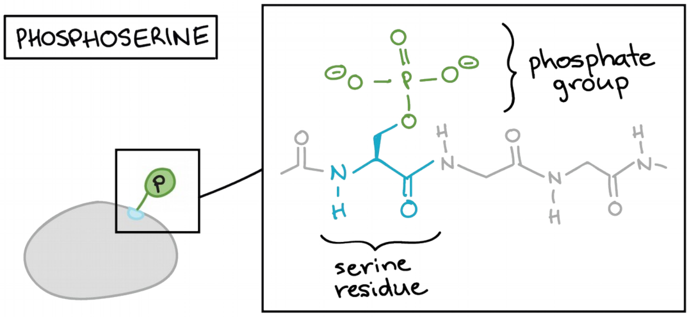


Because of the directional flow of information, the term upstream is often used to describe molecules and events that come earlier in the relay chain, while downstream may be used to describe those that come later (relative to a particular molecule of interest). For instance, in the diagram, the receptor is downstream of the ligand but upstream of the the proteins in the cytosol. Many signal transduction pathways amplify the initial signal, so that one molecule of ligand can lead to the activation of many molecules of a downstream target.

The molecules that relay a signal are often proteins. However, non-protein molecules like ions and phospholipids can also play important roles.

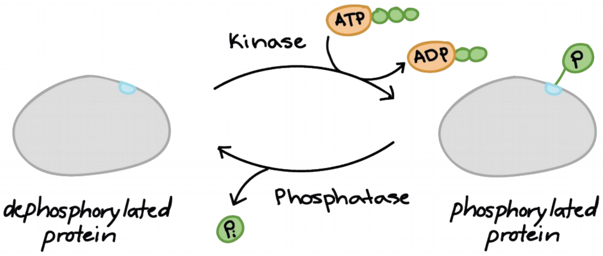
Phosphorylation

The cartoon above features a bunch of blobs (signaling molecules) labeled as “on” or “off.” What does it actually mean for a blob to be on or off? Proteins can be activated or inactivated in a variety of ways. However, one of the most common tricks for altering protein activity is the addition of a phosphate group to one or more sites on the protein, a process called phosphorylation.



Phosphate groups can’t be attached to just any part of a protein. Instead, they are typically linked to one of the three amino acids that have hydroxyl (-OH) groups in their side chains: tyrosine, threonine, and serine. The transfer of the phosphate group is catalyzed by an enzyme called a kinase, and cells contain many different kinases that phosphorylate different targets. Phosphorylation often acts as a switch, but its effects vary among proteins. Sometimes, phosphorylation will make a protein more active (for instance, increasing catalysis or letting it bind to a partner). In other cases, phosphorylation may inactivate the protein or cause it to be broken down.

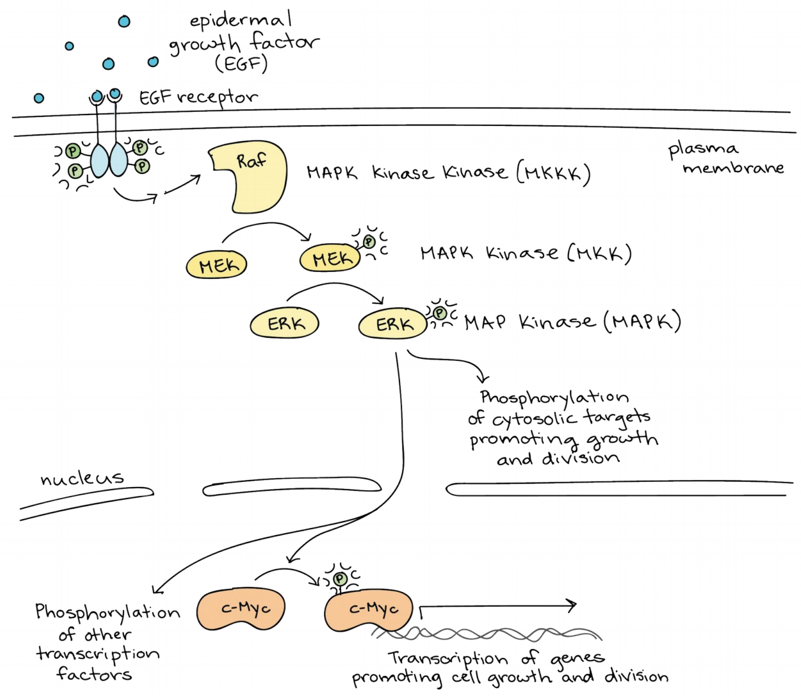
In general, phosphorylation isn’t permanent. To flip proteins back into their non-phosphorylated state, cells have enzymes called phosphatases, which remove a phosphate group from their targets.



Phosphorylation example: MAPK signaling cascade

To get a better sense of how phosphorylation works, let’s examine a real-life example of a signaling pathway that uses this technique: growth factor signaling. Specifically, we'll look at part of the epidermal growth factor (EGF) pathway that acts through a series of kinases to produce a cellular response.

This diagram shows part of the epidermal growth factor signaling pathway:



Phosphorylation (marked as a P) is important at many stages of this pathway:

1. When growth factor ligands bind to their receptors, the receptors pair up and act as kinases, attaching phosphate groups to one another’s intracellular tails.

2. The activated receptors trigger a series of events (skipped here because they don't involve phosphorylation). These events activate the kinase Raf.

3. Active Raf phosphorylates and activates MEK, which phosphorylates and activates the ERKs.

4. The ERKs phosphorylate and activate a variety of target molecules. These include transcription factors, like c-Myc, as well as cytoplasmic targets. The activated targets promote cell growth and division.

Together, Raf, MEK, and the ERKs make up a three-tiered kinase signaling pathway called a mitogen-activated protein kinase (MAPK) cascade. (A mitogen is a signal that causes cells to undergo mitosis, or divide.) Because they play a central role in promoting cell division, the genes encoding the growth factor receptor, Raf, and c-Myc are all proto-oncogenes, meaning that overactive forms of these proteins are associated with cancer.

MAP kinase signaling pathways are widespread in biology: they are found in a wide range of organisms, from humans to yeast to plants. The similarity of MAPK cascades in diverse organisms suggests that this pathway emerged early in the evolutionary history of life and was already present in a common ancestor of modern-day animals, plants, and fungi.

Second messengers

Although proteins are important in signal transduction pathways, other types of molecules can participate as well. Many pathways involve second messengers, small, non-protein molecules that pass along a signal initiated by the binding of a ligand (the "first messenger") to its receptor.

Second messengers include Ca^{2+}; cyclic AMP (cAMP), a derivative of ATP; and inositol phosphates, which are made from phospholipids.

Calcium ions

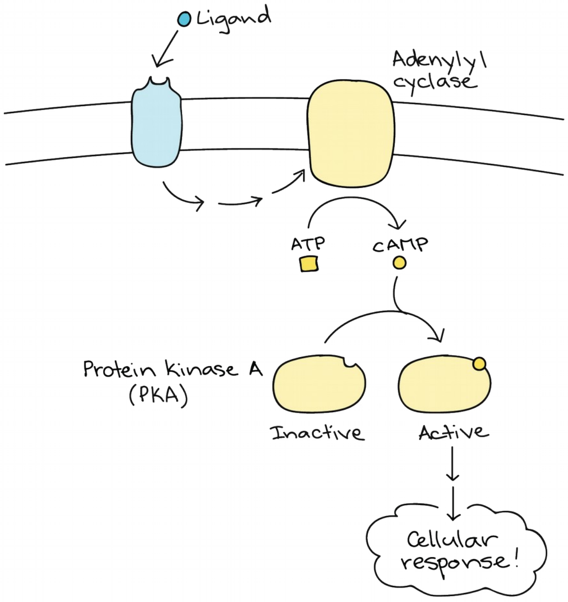
Calcium ions are a widely used type of second messenger. In most cells, the concentration of calcium ions (Ca^{2+}) in the cytosol is very low, as ion pumps in the plasma membrane continually work to remove it. For signaling purposes, Ca^{2+} may be stored in compartments such as the endoplasmic reticulum.

In pathways that use calcium ions as a second messenger, upstream signaling events release a ligand that binds to and opens ligand-gated calcium ion channels. These channels open and allow the higher levels of Ca^{2+} that are present outside the cell (or in intracellular storage compartments) to flow into the cytoplasm, raising the concentration of cytoplasmic Ca^{2+}.

How does the released Ca^{2+} help pass along the signal? Some proteins in the cell have binding sites for Ca^{2+} ions, and the released ions attach to these proteins and change their shape (and thus, their activity). The proteins present and the response produced are different in different types of cells. For instance, Ca^{2+} signaling in the β-cells of the pancreas leads to the release of insulin, while Ca^{2+} signaling in muscle cells leads to muscle contraction.

Cyclic AMP (cAMP)

Another second messenger used in many different cell types is cyclic adenosine monophosphate (cyclic AMP or cAMP), a small molecule made from ATP. In response to signals, an enzyme called adenylyl cyclase converts ATP into cAMP, removing two phosphates and linking the remaining phosphate to the sugar in a ring shape. Once generated, cAMP can activate an enzyme called protein kinase A (PKA), enabling it to phosphorylate its targets and pass along the signal. Protein kinase A is found in a variety of types of cells, and it has different target proteins in each. This allows the same cAMP second messenger to produce different responses in different contexts.



cAMP signaling is turned off by enzymes called phosphodiesterases, which break the ring of cAMP and turn it into adenosine monophosphate (AMP).

Inositol phosphates

Although we usually think of plasma membrane phospholipids as structural components of the cell, they can also be important participants in signaling. Phospholipids called phosphatidylinositols can be phosphorylated and snipped in half, releasing two fragments that both act as second messengers.

One lipid in this group that's particularly important in signaling is called PIP_2. In response to a signal, an enzyme called phospholipase C cleaves (chops) PIP_2 into two fragments, DAG and IP_3. These fragments made can both act as second messengers.

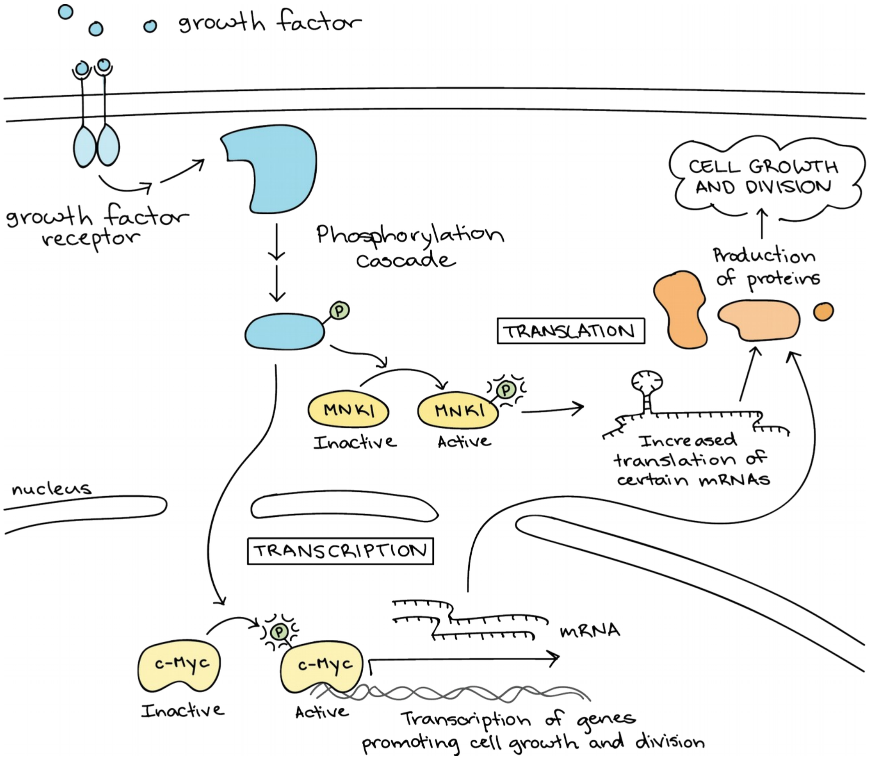
DAG stays in the plasma membrane and can activate a target called protein kinase C (PKC), allowing it to phosphorylate its own targets. IP_3 diffuses into the cytoplasm and can bind to ligand-gated calcium channels in the endoplasmic reticulum, releasing Ca^{2+} that continues the signal cascade.

Cellular Response

1. Gene Expression (Growth Factor Signaling)

An example to see how signaling pathways alter transcription and translation:

This growth factor pathway has many targets, which it activates through a signaling cascade that involves phosphorylation (addition of phosphate groups to molecules). Some of the pathway's targets are transcription factors, proteins that increase or decrease transcription of certain genes. In the case of growth factor signaling, the genes have effects that lead to cell growth and division. One transcription factor targeted by the pathway is c-Myc, a protein that can lead to cancer when it is too active ("too good" at promoting cell division).



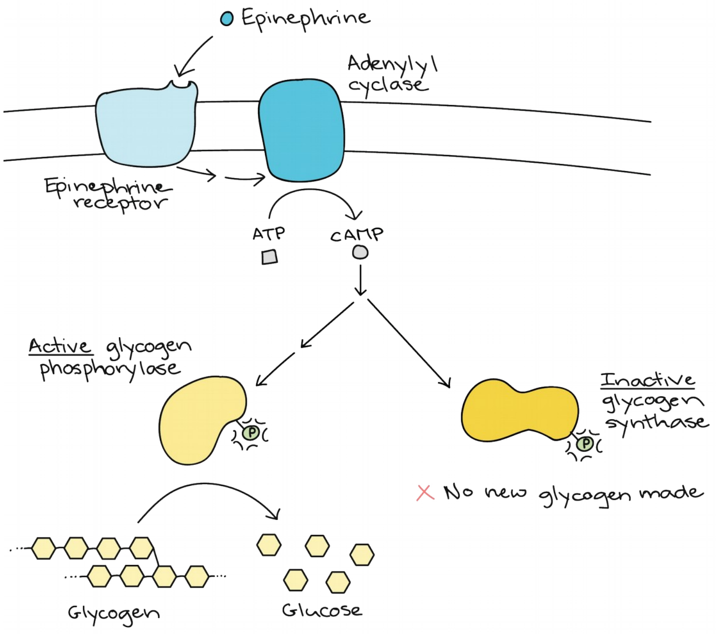
The growth factor pathway also affects gene expression at the level of translation. For instance, one of its targets is a translational regulator called MNK1. Active MNK1 increases the rate of mRNA translation, especially for certain mRNAs that fold back on themselves to make hairpin structures (which would normally block translation). Many key genes regulating cell division and survival have mRNAs that form hairpin structures, and MNK1 allows these genes to be expressed at high levels, driving growth and division.

Notably, neither c-Myc nor MNK1 is a "final responder" in the growth factor pathway. Instead, these regulatory factors, and others like them, promote or repress the production of other proteins (the orange blobs in the illustration above) that are more directly involved in carrying out cell growth and division.

2. Cellular metabolism

Some signaling pathways produce a metabolic response, in which metabolic enzymes in the cell become more or less active. We can see how this works by considering adrenaline signaling in muscle cells. Adrenaline, also known as epinephrine, is a hormone (produced by the adrenal gland) that readies the body for short-term emergencies. If you’re nervous before a test or competition, your adrenal gland is likely to be pumping out epinephrine.

When epinephrine binds to its receptor on a muscle cell (a type of G protein-coupled receptor), it triggers a signal transduction cascade involving production of the second messenger molecule cyclic AMP (cAMP). This cascade leads to phosphorylation of two metabolic enzymes— that is, addition of a phosphate group, causing a change in the enzymes' behavior.



The first enzyme is glycogen phosphorylase (GP). The job of this enzyme is to break down glycogen into glucose. Glycogen is a storage form of glucose, and when energy is needed, glycogen must be broken down. Phosphorylation activates glycogen phosphorylase, causing lots of glucose to be released.

The second enzyme that gets phosphorylated is glycogen synthase (GS). This enzyme is involving in building up glycogen, and phosphorylation inhibits its activity. This ensures that no new glycogen molecules are built when the current need is for glycogen to be broken down.

Through regulation of these enzymes, a muscle cell rapidly gets a large, ready pool of glucose molecules. The glucose is available for use by the muscle cell in response to a sudden surge of adrenaline—the “fight or flight” response.

The Complexity of Life: (At the level of a cell)

- ✓ There are some 42 million protein molecules in a simple cell.
- ✓ **What are the concentrations of free metabolites in cells?**

The cell’s canonical components of proteins, nucleic acids, lipids and sugars are complemented by a host of small metabolites that serve a number of key roles. These metabolites are broadly defined as members of the many families of molecules within cells having a molecular weight of less than 1000 Daltons. Recent measurements have made it possible to take a census of these metabolites in bacteria as shown in Table 1. Perhaps the most familiar role for these metabolites is as the building blocks for the polymerization reactions leading to the assembly of the key macromolecules of the cell. However, their biochemical reach is much larger than this restricted set of reactions. These metabolites also serve as energy sources, key activity regulators, signal transducers, electron donors and buffers of both pH and osmotic pressure. An inventory of which metabolites are present and at what concentrations is of great interest since it provides a picture of the stocks available to the cell as reserves for building its macromolecules. In addition, this inventory tells us which compounds are most ubiquitous and how we should think about the various chemical reactions (both specific and nonspecific) that they are part of. The molecular census of metabolites in E. coli reveals some overwhelmingly dominant molecular players. The amino acid glutamate wins out in Table 1 at about 100mM, which is higher than all other amino acids combined as depicted in Figure 1. Our intuition and memory is much better with absolute numbers than with concentrations so we recall our rule of thumb that a concentration of 1 nM corresponds to roughly one copy of the molecule of interest per E. coli cell. Hence, 100 mM means that there are roughly 10^8 copies of glutamate in each bacterium.

✓ How many protein equivalents is this? these 108 glutamates are equivalent to roughly 3 x 10^5 proteins, roughly 10% of the ≈3 x 10^6 proteins making up the entire protein census of the cell.

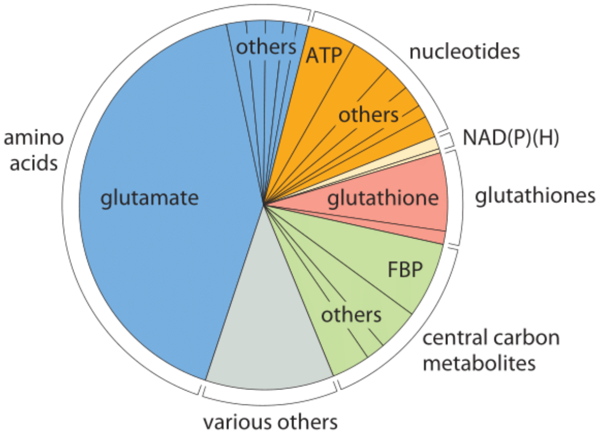
Effective metabolite concentrations:

•1-10 mM(very abundant) =1-10 fmol/ cell

•1-100 µM(typical) = 1-100 amol/cell

•1-100 nmol(rare) = 1-100 zmol/cell

✓ 109 types of metabolites in a single E. coli cell (table 1)



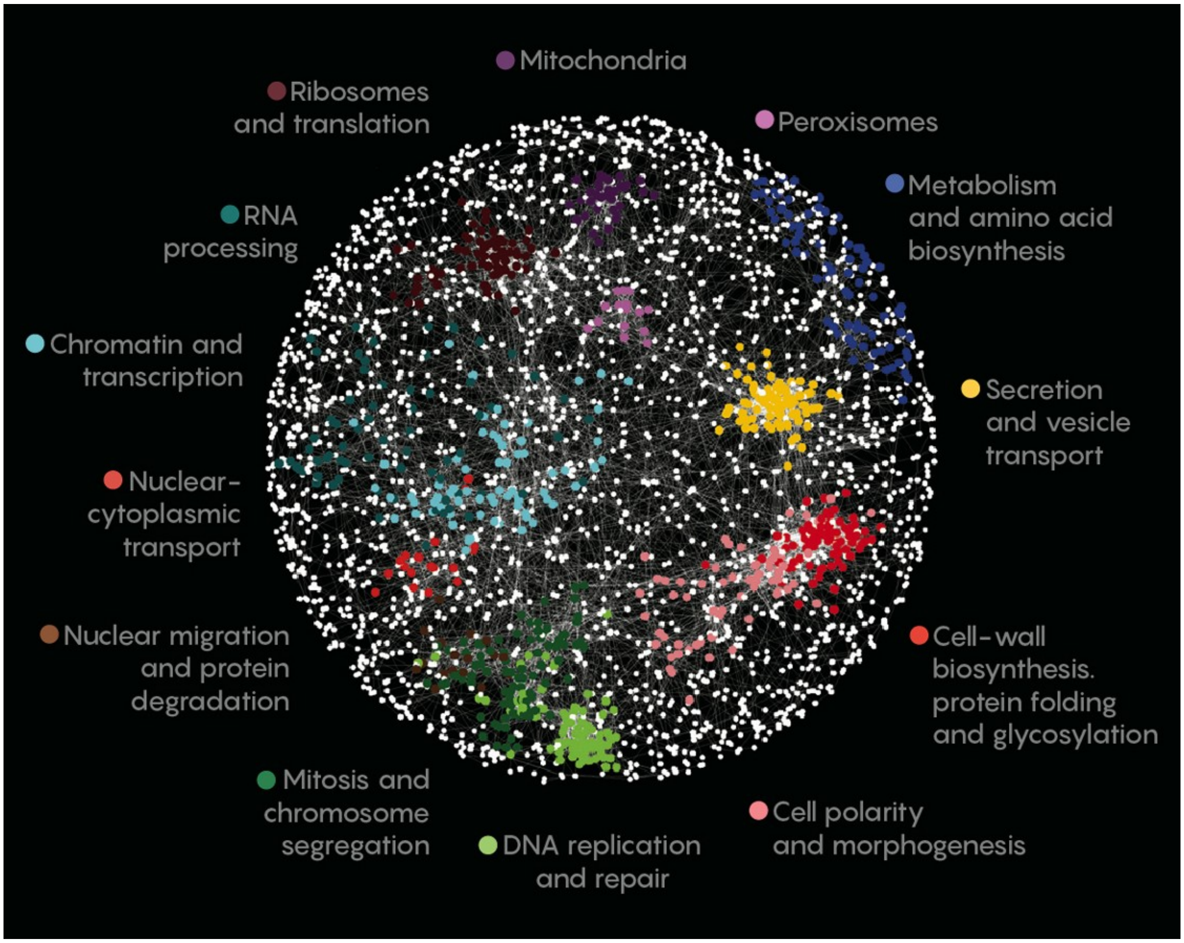
sum of concentrations ≈ 200 mM

metabolite	mM
glutamate	96
glutathione	17
fructose-1,6-bisphosphate	15
ATP	9.6
UDP-N-acetyl-glucosamine	9.2
hexose-P	8.8
UTP	8.3
GTP	4.9
dTTP	4.6
aspartate	4.2
valine	4.0
glutamine	3.8
6-phospho-D-gluconate	3.8
CTP	2.7
NAD	2.6
alanine	2.5
UDP-glucose	2.5
glutathionedisulfide	2.4
uridine	2.1
citrate	2.0
UDP	1.8
malate	1.7
3-phosphoglycerate	1.5
glycerate	1.4
coenzyme-A	1.4
citrulline	1.4
pentose-P	1.3
glucosamine-6_phosphate	1.2
acetylphosphate	1.1
gluconolactone	1.0
GDP	0.68
acetyl-CoA	0.61
carbamyl-aspartate	0.59
succinate	0.57
arginine	0.57
UDP-glucaronate	0.57
ADP	0.55
asparagine	0.51
2-ketoglutarate	0.44
lysine	0.40
proline	0.38
dTDP	0.38
dihydroxyacetone-phosphate	0.37
homocysteine	0.37
CMP	0.36
isoleucine+leuicine	0.30
deoxyribose-5-P	0.30
AMP	0.28
inosine-monophosphate	0.27
PRPP	0.26
succinyl-CoA	0.23
inosine-triphosphate	0.20
guanine	0.19
S-adenosyl-L-methionine	0.18
phosphoenolpyruvate	0.18
threonine	0.18
FAD	0.17
methionine	0.14
2,3-dihydroxybenzoicacid	0.14
NADPH	0.12
fumarate	0.11
phenylpyruvate	0.090
NADH	0.083
N-acetyl-glucosamine-1P	0.082
serine	0.068
histidine	0.068
flavinmononucleotide	0.054
4-hydroxybenzoate	0.052
dGMP	0.051
glycerolphosphate	0.049
N-acetyl-ornithine	0.043
gluconate	0.042
malonyl-CoA	0.035
cyclic-AMP	0.035
dCTP	0.034
tyrosine	0.029
inosine-diphosphate	0.024
GMP	0.024
acetoacetyl-CoA	0.022
riboflavin	0.019
phenylalanine	0.018
aconitate	0.016
dATP	0.016
cytosine	0.014
shikimate	0.014
histidinol	0.013
tryptophan	0.012
dihydroorotate	0.012
quinolinate	0.012
ornithine	0.010
dAMP	0.0088
adenosine-phosphosulfate	0.0066
myo-inositol	0.0057
propionyl-CoA	0.0053
ADP-glucose	0.0043
anthranilate	0.0035
deoxyadenosine	0.0028
cytidine	0.0026
NADP+	0.0021
guanosine	0.0016
adenine	0.0015
deoxyguanosine	0.00052
adenosine	0.00013
Sum	231

How many genes in a cell?

By knocking out genes three at a time, scientists have painstakingly deduced the web of genetic interactions that keeps a cell alive. Researchers long ago identified essential genes that yeast cells can't live without, but new work, which appears today in Science, shows that looking only at those gives a skewed picture of what makes cells tick: Many genes that are inessential on their own become crucial as others disappear. The result implies that the true minimum number of genes that yeast — and perhaps, by extension, other complex organisms — need to survive and thrive may be surprisingly large.

About 20 years ago, Charles Boone and Brenda Andrews decided to do something slightly nuts. The yeast biologists, both professors at the University of Toronto, set out to systematically destroy or impair the genes in yeast, two by two, to get a sense of how the genes functionally connected to one another. Only about 1,000 of the 6,000 genes in the yeast genome, or roughly 17 percent, are considered essential for life: If a single one of them is missing, the organism dies. But it seemed that many other genes whose individual absence was not enough to spell the end might, if destroyed in tandem, sicken or kill the yeast. Those genes were likely to do the same kind of job in the cell, the biologists reasoned, or to be involved in the same process; losing both meant the yeast could no longer compensate.



This figure maps the interactions among various genes (represented as dots) in the yeast genome. Genes with linked effects are connected by lines. A recount of human genes ups the number to at least 46,831.

Metabolic Network

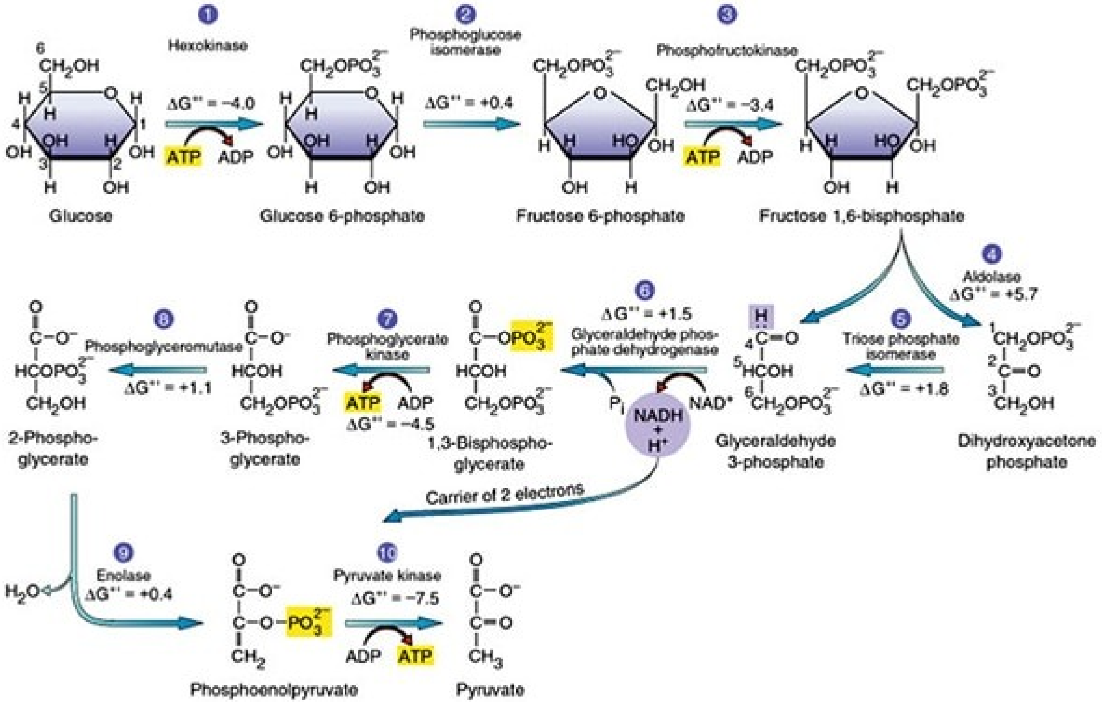
Metabolism is sum of all the chemical reactions that take place in every cell of a living organism, providing energy for the processes of life and synthesizing new cellular material.

Metabolism forms a network of interconnected metabolites and reactions. e.g. **Saccharomyces cerevisiae** has 646 metabolites and 1149 reactions. [Saccharomyces cerevisiae (commonly known as baker's yeast) is a single-celled eukaryote that is frequently used in scientific research. S. cerevisiae is an attractive model organism due to the fact that its genome has been sequenced, its genetics are easily manipulated, and it is very easy to maintain in the lab.] Metabolic network is a set of metabolites connected by reactions, and it consists of pathways. systems of successive chemical reactions, “set of oriented reactions interacting under given physiological conditions via simple or apparently simple intermediates”.

<u>Stoichiometry Matrix</u>

Stoichiometry is determination of the proportions (by weight or number of molecules) in which elements or compounds react with one another. Charge balance is considered in the determination of stoichiometric coefficients.

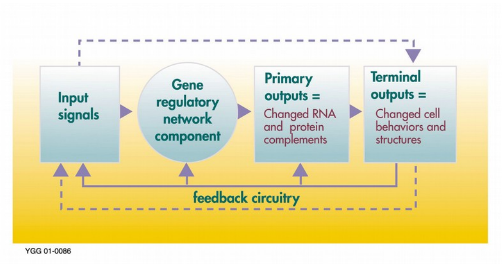
Example of Metabolic Network: Glycolysis

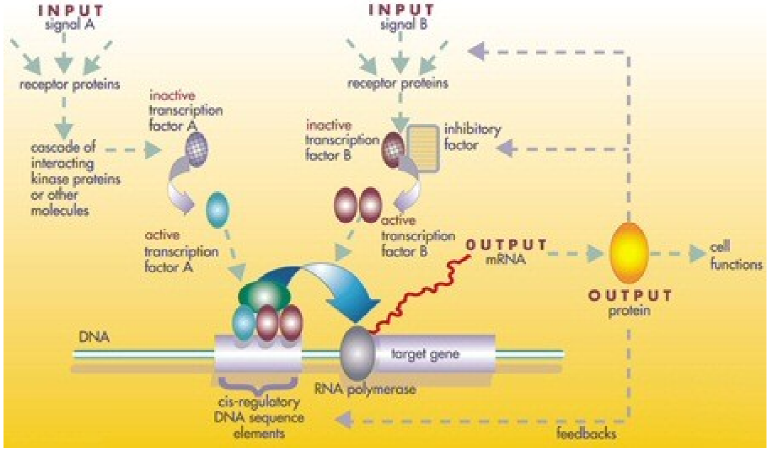


Regulatory Network

A gene (or genetic) regulatory network (GRN) is a collection of molecular regulators that interact with each other and with other substances in the cell to govern the gene expression levels of mRNA and proteins. These play a central role in morphogenesis, the creation of body structures, which in turn is central to evolutionary developmental biology (evo-devo). Morphogenesis (from the Greek morphê shape and genesis creation, literally, "beginning of the shape") is the biological process that causes an organism to develop its shape. It is one of three fundamental aspects of developmental biology along with the control of cell growth and cellular differentiation, unified in evolutionary developmental biology (evo-devo).

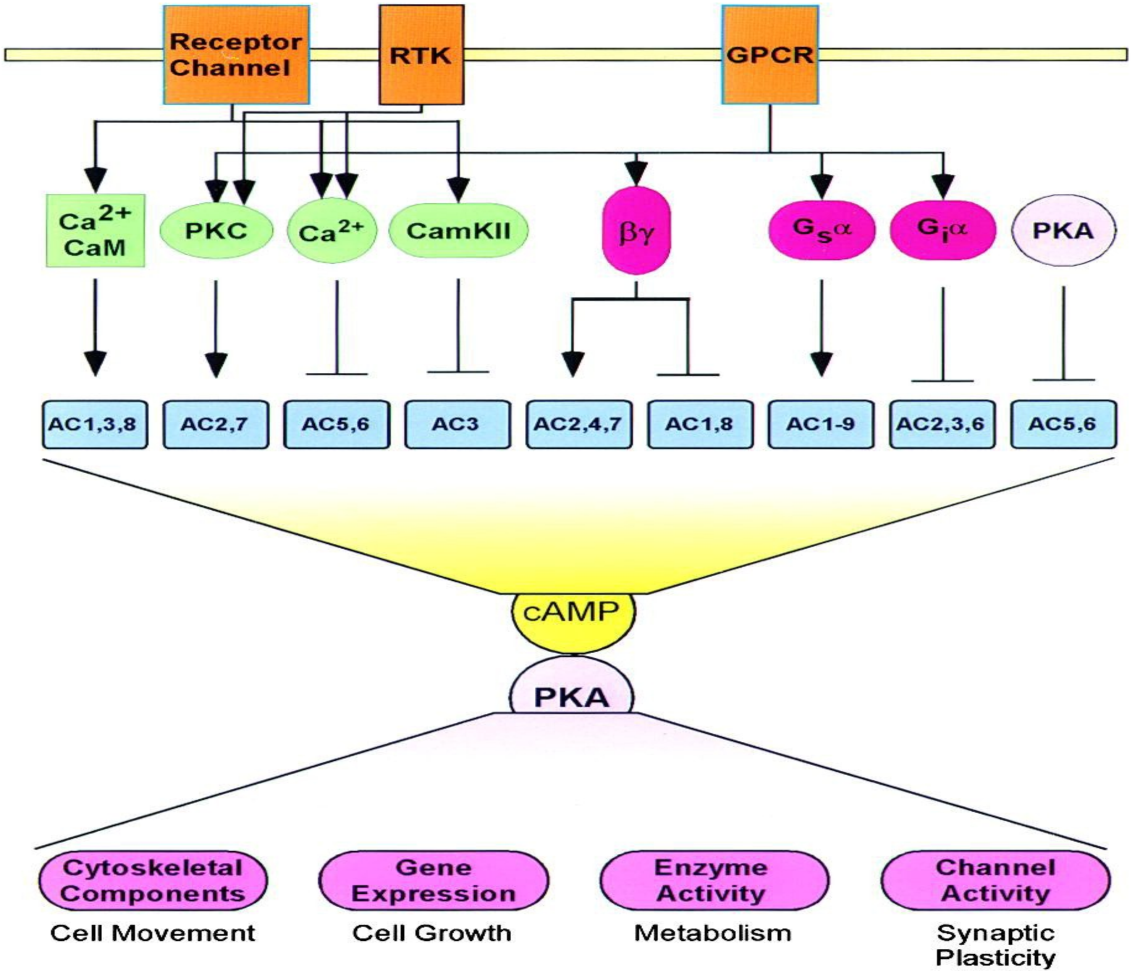
The regulator can be DNA, RNA, protein and complexes of these. The interaction can be direct or indirect (through transcribed RNA or translated protein). In general, each mRNA molecule goes on to make a specific protein (or set of proteins). In some cases this protein will be structural, and will accumulate at the cell membrane or within the cell to give it particular structural properties. In other cases the protein will be an enzyme, i.e., a micro-machine that catalyses a certain reaction, such as the breakdown of a food source or toxin. Some proteins though serve only to activate other genes, and these are the transcription factors that are the main players in regulatory networks or cascades. By binding to the promoter region at the start of other genes they turn them on, initiating the production of another protein, and so on. Some transcription factors are inhibitory. In single-celled organisms, regulatory networks respond to the external environment, optimising the cell at a given time for survival in this environment.





Signaling Network

Signaling Networks result from the interconnections between signaling pathways. Such interconnections occur because the same signaling component is capable of receiving signals from multiple inputs. There are two general classes of interconnections: junctions, which are signal integrators and nodes which split the signal and route them to multiple outputs.



Adenylyle Cyclase as example of a junction: The signal receiving capabilities of the various adenylyl cyclase isoforms and the capability of the cAMP-dependent protein kinase (PKA) to regulate various physiological functions are shown. Receptor channel, ligand gated channel (e.g., NMDA receptor); RTK, receptor tyrosine kinase; GPCR, G protein–coupled receptor. Stimulatory signals are shown as arrows and inhibitory signals as plungers. The various cellular components or processes regulated by PKA are shown in the red ovals and the resultant physiological functions are given below.

Systems Biology

What is Systems Biology?

Systems biology is a comprehensive quantitative analysisof the manner in which all the components of a biological system interact functionally over time.

Why is it difficult to define?

Because there always appears to be a delicate balance between opposing aspects:

Scale: genome wide vs small scale networks

Discipline: biological vs physical

Method: computational vs experimental

Analysis: deterministic vs probablistic

Comparing with a chemical plant or electronic circuit:

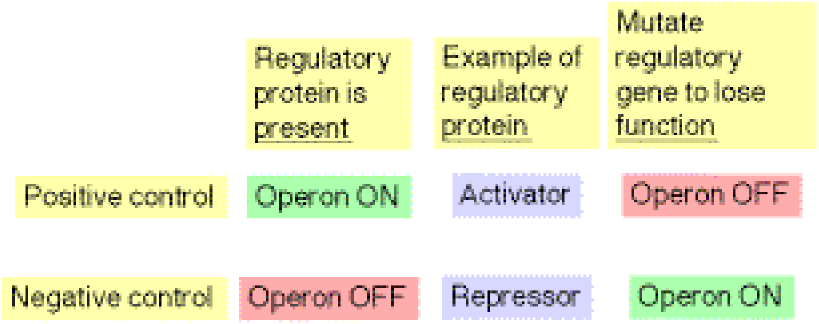
1. There are no pipes lines inside the cell
2. Reactants and products are not restricted with ‘reactors’
3. Characteristic times of cellular events vary over a wide range (10 9s to 103 s)
4. Petroleum plants cannot self replicate
5. Signal molecules are not restricted to electrical channels
6. Genetics circuits are not restricted by “circuit boards”
7. Cells generate their own energy
8. Cells are “ALIVE”

Systems Biology is a combination of Electrical (signal, gene regulation), Mechanical (membrane transport) and Chemical (cellular processes) interactions.

Cell Mechanism: Positive control and negative control

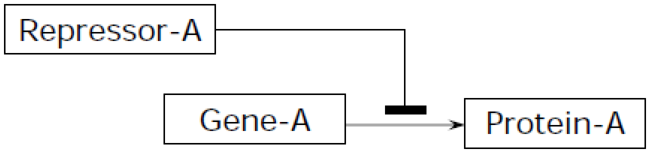
Operons

An **operon** is a cluster of coordinately regulated genes. It includes **structural genes** (generally encoding enzymes), **regulatory genes** (encoding, e.g. activators or repressors) and **regulatory sites** (such as promoters and operators). The type of control is defined by the response of the operon when no regulatory protein is present. In the case of <u>negative control</u>, the genes in the operon are expressed unless they are switched off by a repressor protein. Thus the operon will be turned on constitutively (the genes will be expressed) when the repressor in inactivated. In the case of <u>positive control</u>, the genes are expressed only when an active regulator protein, e.g. an activator, is present. Thus the operon will be turned off when the positive regulatory protein is absent or inactivated.

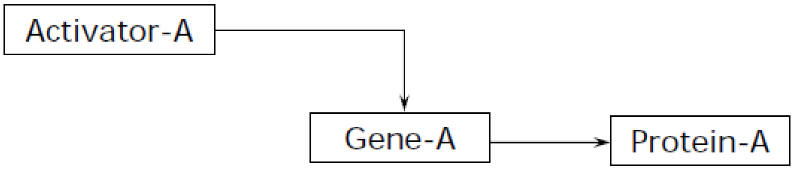


	Regulatory protein is present	Example of regulatory protein	Mutate regulatory gene to lose function
Positive control	Operon ON	Activator	Operon OFF
Negative control	Operon OFF	Repressor	Operon ON

Depiction of negative and positive control



Negative control



Positive control

Example:
Lac Operon. *What is Lac Operon?*

The lactose operon of E. coli contains genes involved in **lactose metabolism**. It's expressed only when lactose is present and glucose is absent.
Two regulators turn the operon "on" and "off" in response to lactose and glucose levels: the lac repressor and catabolite activator protein (CAP).
The lac repressor (negative control) acts as a lactose sensor. It normally blocks transcription of the operon, but stops acting as a repressor when lactose is present. The lac repressor senses lactose indirectly, through its isomer **allolactose**.

Lac repressor keeps the Operon turned off.

Catabolite activator protein (CAP) (positive control) acts as a glucose sensor. It activates transcription of the operon, but only when glucose levels are low. CAP senses glucose indirectly, through the "hunger signal" molecule cAMP.

CAP activator switches on the Operon.

Low glucose=> CAP on
High Lactose=> Lac repressor off
== Lactose operon ON.

<u>Genetic Circuits</u>

Synthetic biological circuits are an application of synthetic biology (Synthetic biology (SynBio) is a multidisciplinary area of research that seeks to create new biological parts, devices, and systems, or to redesign systems that are already found in nature) where biological parts inside a cell are designed to perform logical functions mimicking those observed in electronic circuits. The applications range from simply inducing production to adding a measurable element, like GFP, to an existing natural biological circuit, to implementing completely new systems of many parts. A ribosome is a biological machine.

The goal of synthetic biology is to generate an array of tunable and characterized parts, or modules, with which any desirable synthetic biological circuit can be easily designed and implemented. These circuits can serve as a method to modify cellular functions, create cellular responses to environmental conditions, or influence cellular development. By implementing rational, controllable logic elements in cellular systems, researchers can use living systems as engineered "biological machines" to perform a vast range of useful functions.

What is GFP?

Green fluorescent protein (GFP) has been widely used to tag proteins and to study the dynamic changes of cellular processes in living cells.

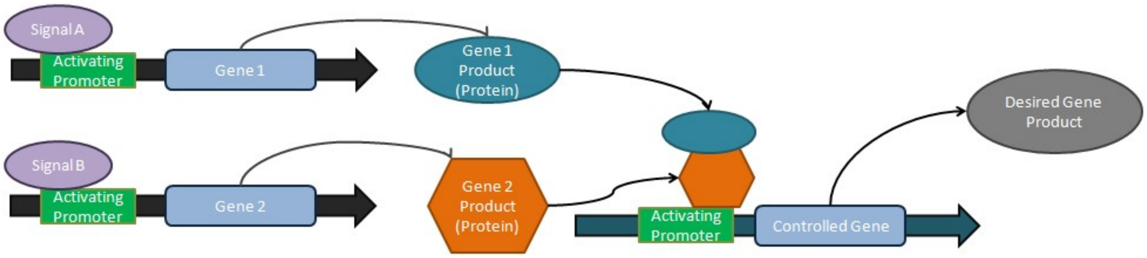
List of currently synthesized circuits:

- Repressilator
- Toggle-switch
- Mammalian tunable synthetic oscillator
- Bacterial tunable synthetic oscillator
- Coupled bacterial oscillator
- Globally coupled bacterial oscillator

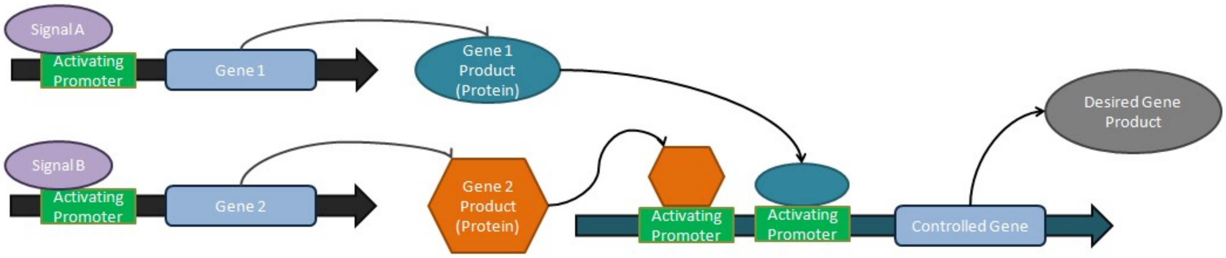
Repressilator:

The repressilator is a genetic regulatory network consisting of at least one feedback loop with at least three genes, each expressing a protein that represses the next gene in the loop.[1] In biological research, repressilators have been used to build cellular models and understand cell function. There are both artificial and naturally-occurring repressilators. Recently, the naturally-occurring repressilator clock gene circuit in Arabidopsis thaliana (A. thaliana) and mammalian systems have been studied.

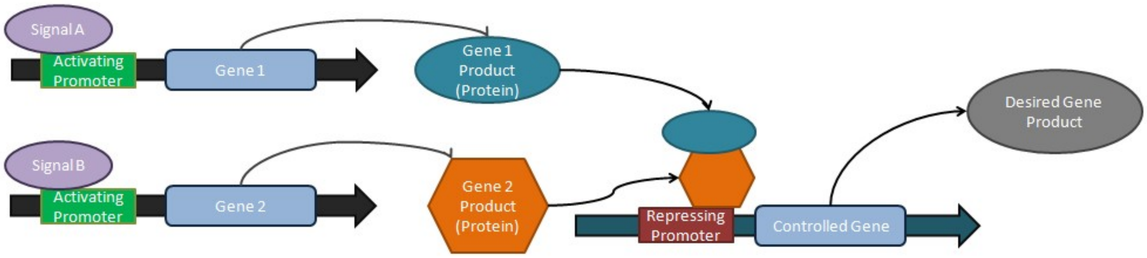
Examples of genetic logic circuits:



The logical AND gate. If Signal A AND Signal B are present, then the desired gene product will result. All promoters shown are inducible, activated by the displayed gene product. Each signal activates expression of a separate gene (shown in light blue). The expressed proteins then can either form a complete complex in cytosol, that is capable of activating expression of the output (shown), or can act separately to induce expression, such as separately removing an inhibiting protein and inducing activation of the uninhibited promoter.



The logical OR gate. If Signal A OR Signal B are present, then the desired gene product will result. All promoters shown are inducible. Either signal is capable of activating the expression of the output gene product, and only the action of a single promoter is required for gene expression. Post-transcriptional regulation mechanisms can prevent the presence of both inputs producing a compounded high output, such as implementing a low binding affinity ribosome binding site.



The logical Negated AND gate. If Signal A AND Signal B are present, then the desired gene product will NOT result. All promoters shown are inducible. The activating promoter for the output gene is constitutive, and thus not shown. The constitutive promoter for the output gene keeps it "on" and is only deactivated when (similar to the AND gate) a complex as a result of two input signal gene products blocks the expression of the output gene.

Positive and Negative Feedback in Systems Biology

Positive Feedback Loop

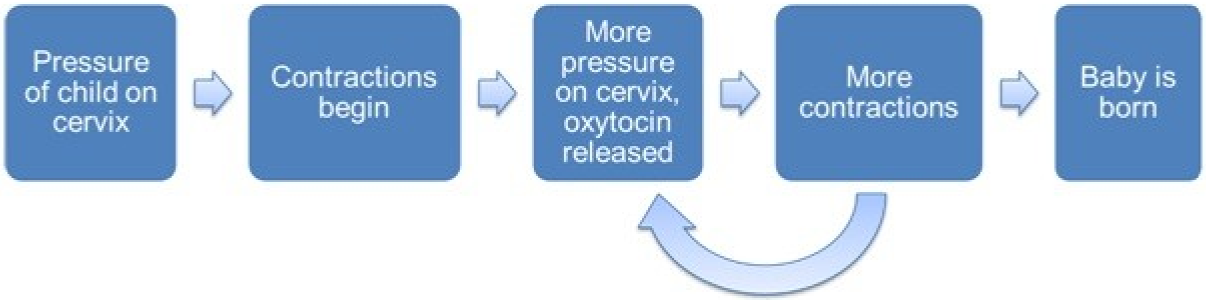
A positive feedback loop occurs in nature when the product of a reaction leads to an increase in that reaction. If we look at a system in homeostasis, a positive feedback loop moves a system further away from the target of equilibrium. It does this by amplifying the effects of a product or event and occurs when something needs to happen quickly.

Example 1: Fruit Ripening

There is a surprising effect in nature where a tree or bush will suddenly ripen all of its fruit or vegetables, without any visible signal. This is our first example of a positive biological feedback loop. If we look at an apple tree, with many apples, seemingly overnight they all go from unripe to ripe to overripe. This will begin with the first apple to ripen. Once ripe, it gives off a gas known as ethylene (C2H4) through its skin. When exposed to this gas, the apples near to it also ripen. Once ripe, they too produce ethylene, which continues to ripen the rest of the tree in an effect much like a wave. This feedback loop is often used in fruit production, with apples being exposed to manufactured ethylene gas to make them ripen faster.

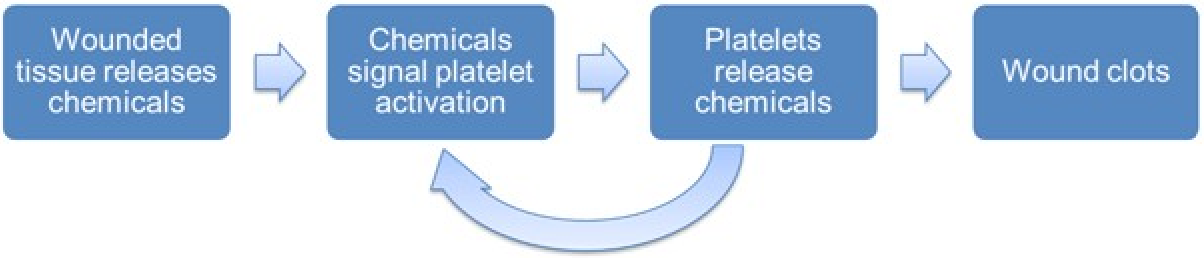
Example 2: Childbirth

When labor begins, the baby’s head is pushed downwards and results in increased pressure on the cervix. This stimulates receptor cells to send a chemical signal to the brain, allowing the release of oxytocin. This oxytocin diffuses to the cervix via the blood, where it stimulated further contractions. These contractions stimulate further oxytocin release until the baby is born.



Example 3: Blood Clotting

When tissue is torn or injured, a chemical is released. This chemical causes platelets in the blood to activate. Once these platelets have activated, they release a chemical which signals more platelets to activate, until the wound is clotted.

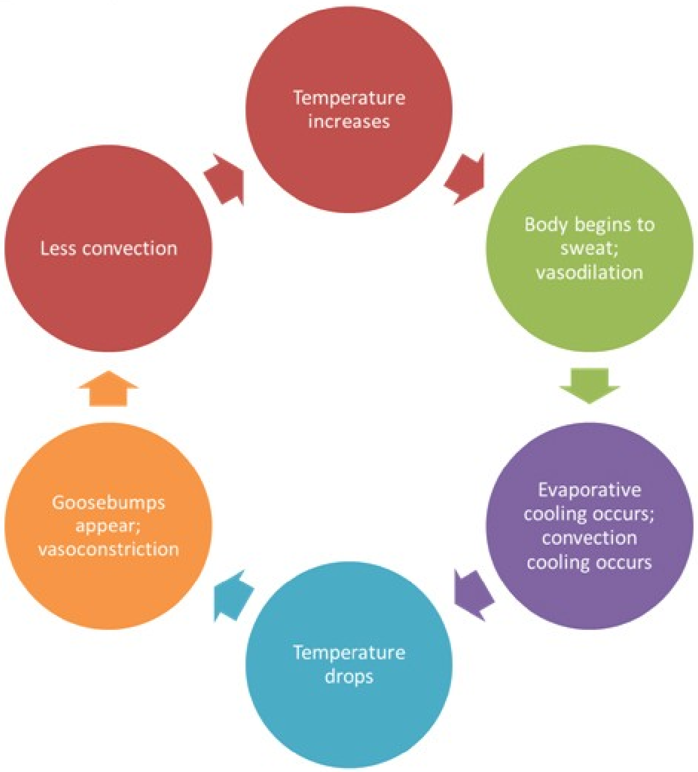


Negative Feedback Loop

A negative feedback loop occurs in biology when the product of a reaction leads to a decrease in that reaction. In this way, a negative feedback loop brings a system closer to a target of stability or homeostasis. Negative feedback loops are responsible for the stabilization of a system, and ensure the maintenance of a steady, stable state. The response of the regulating mechanism is opposite to the output of the event.

Example 1: Temperature Regulation

Temperature regulation in humans occurs constantly. Normal human body temperature is approximately 98.6°F. When body temperature rises above this, two mechanisms kick in the body begins to sweat, and vasodilation occurs to allow more of the blood surface area to be exposed to the cooler external environment. As the sweat cools, it causes evaporative cooling, while the blood vessels cause convective cooling. Normal temperature is regained. Should these cooling mechanisms continue, the body will become cold. The mechanisms which then kick in are the formation of goose bumps, and vasoconstriction. Goosebumps in other mammals raise the hair or fur, allowing more heat to be retained. In humans, they tighten the surrounding skin, reducing (slightly) the surface area from which to lose heat. Vasoconstriction ensures that only a small surface area of the veins is exposed to the cooler outside temperature, retaining heat. Normal temperature is regained.



Example 2: Blood Pressure Regulation (Baroreflex)

Blood pressure needs to remain high enough to pump blood to all parts of the body, but not so high as to cause damage while doing so. While the heart is pumping, baroreceptors detect the pressure of the blood going through the arteries. If the pressure is too high or too low, a chemical signal is sent to the brain via the glossopharyngeal nerve. The brain then sends a chemical signal to the heart to adjust the rate of pumping: if blood pressure is low, heart rate increases, while if blood pressure is high, heart rate decreases.

Example 3: Osmoregulation

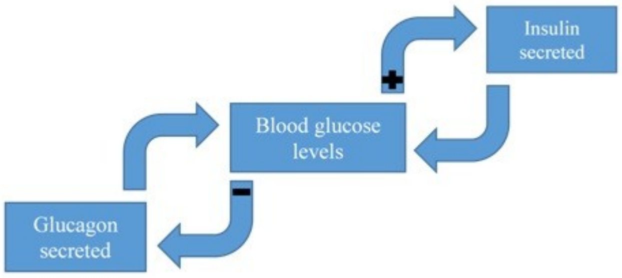
Osmoregulation refers to the control of the concentration of various liquids within the body, to maintain homeostasis. We will again look at an example of a fish, living in the ocean. The concentration of salt in the water surrounding the fish is much higher than that of the liquid in the fish. This water enters the fish diffusion through the gills, through food consumption, and through drinking. Also, because the concentration of salt is higher outside than inside the fish, there is passive diffusion of salt into the fish and water out of the fish. The salt concentration is then too high in the fish, and salt ions must be released through excretion. This occurs via the skin, and in very concentrated urine. In addition, high salt levels in the blood are removed via active transport by the chloride secretory cells in the gills. The correct salt concentration is thus maintained.

Positive vs. Negative Feedback

The key difference between positive and negative feedback is their response to change: positive feedback amplifies change while negative feedback reduces change. This means that positive feedback will result in more of a product: more apples, more contractions, or more clotting platelets. Negative feedback will result in less of a product: less heat, less pressure, or less salt. Positive feedback moves away from a target point while negative feedback moves towards a target.

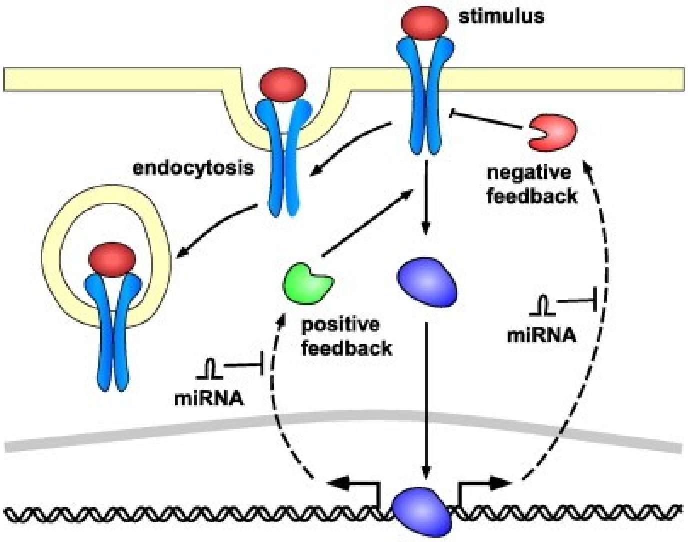
Why is Feedback Important?

Without feedback, homeostasis cannot occur. This means that an organism loses the ability to self-regulate its body. Negative feedback mechanisms are more common in homeostasis, but positive feedback loops are also important. Changes in feedback loops can lead to various issues, including diabetes mellitus.



In a normal glucose cycle, increases in blood glucose levels detected by the pancreas will result in the beta cells of the pancreas secreting insulin until normal blood glucose levels are reached. Whereas if low blood glucose levels are detected, the alpha cells of the pancreas will release glucagon to raise blood glucose levels to be normal.

Feedback in Signal Transduction

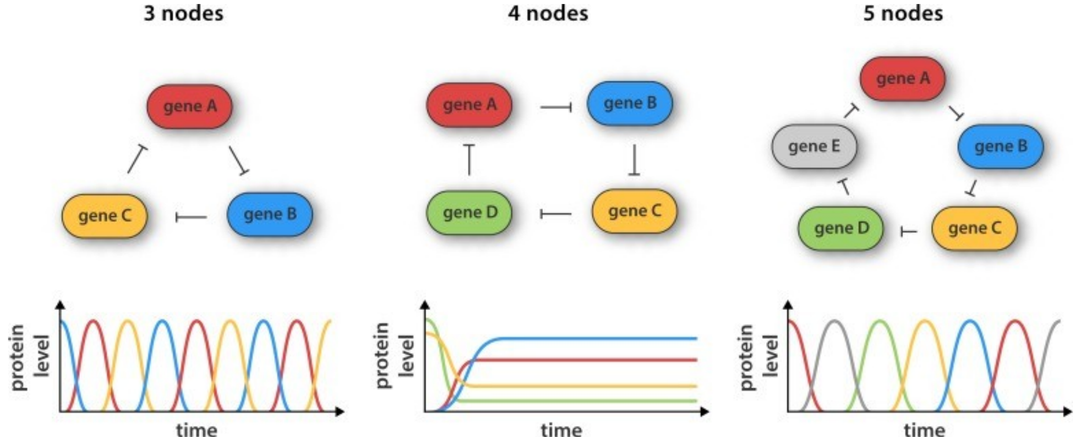


Extracellular ligands bind to and activate cell surface receptors. This induces intracellular signaling events, often resulting in changes in gene expression. Positive or negative feedback regulators are expressed, modulating the signaling dynamics. miRNA have been reported to affect the translation of these feedback proteins, adding another layer of control. In addition, signal attenuation can be induced by endocytosis, leading to a depletion of receptor or ligand.

Negative Feedback and System Oscillation

The rise of synthetic biology has made it possible to design and construct synthetic networks in living cells that perform a specific role. In an early example of this, researchers at Princeton reported that they had constructed an oscillatory gene network in E. coli based on a cyclic network of three genes called the repressilator (Elowitz and Leibler, 2000). Theory predicts that the repressilator and other ring oscillators that have an odd number of genes (nodes) should be capable of producing sustained oscillations. However, since designing, building and testing new gene networks in living cells is a lengthy process, ring oscillators with more than three nodes have not been reported.

Now Maerkl and co-workers – who are based at the École Polytechnique Fédérale de Lausanne and the California Institute of Technology – have made ring architectures containing three, four and five genes.



Synthetic gene networks containing three, four and five genes.

The genes in each circuit (top) are translated into protein products, with each protein product repressing the activity of another gene in the network (as indicated by the arrows). Theory predicts that cyclic networks of genes display oscillatory behavior when the number of nodes in the network is odd. Niederholtmeyer et al. found that a circuit consisting of three genes gave rise to well-defined oscillations with a period of up to 8 hr, and that a circuit containing five genes oscillated with a period of 19 hours. In contrast, and in line with theoretical predictions, a network consisting of four nodes did not oscillate: instead it reached a steady state where the activity of all the genes was constant over time.

Modeling Cell Dynamics

Cell migration is essential for embryogenesis, wound healing, immune surveillance, andprogression of diseases, such as cancer metastasis. During embryogenesis, coordinatedcollective migration to particular sites is essential to the development of an organism.Tissue repair and immune response rely on directed movement of cells following externalcues, which are produced after the tissue layer is damaged and infected with pathogens.Likewise, cancer cells migrate away from the tumor into surrounding tissue and distantorgans to form metastases, which is the leading cause of death among cancer patients.

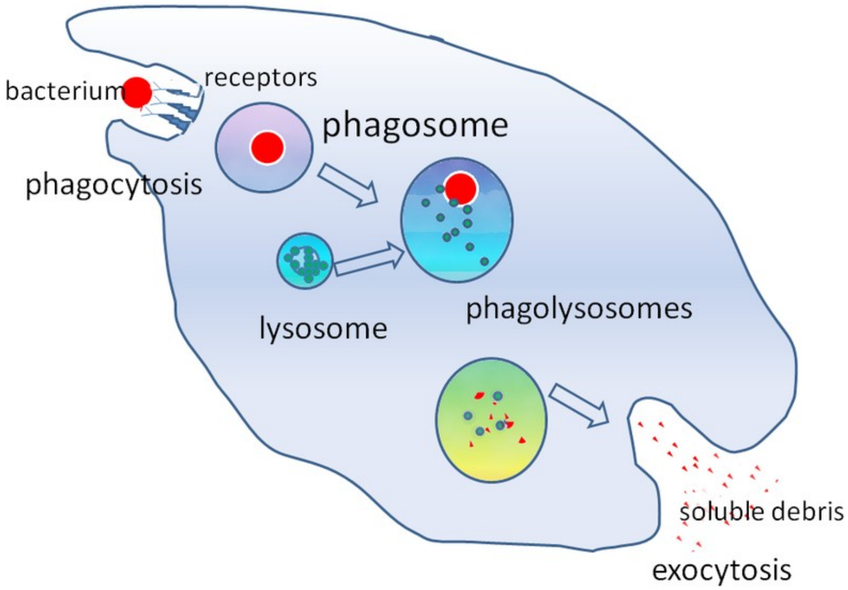
Cell motility is a cyclical process, involving morphological changes of the cell body andadhesive contacts with the underlying substrate. The cycle can be divided into three steps: protrusion of the leading edge, assembly of adhesions to the substrate at the front and disassembly in the rear, and contraction of the cell body, thereby producing locomotion. This type of crawling movement requires transmission of contractile forces, generatedwithin the cell cytoskeleton, through substrate adhesions. Such mechanical interactionsbetween various cellular structures is attained by integrating numerous signaling molecules,the most prominent of which are Rho GTPases. There are other modes of motility relying on, for example, flagellar activity (e.g. spermatozoa or E.coli) or rolling inthe bloodstream (e.g. leukocytes), apart from this type of crawling type of movement.

Examples of cell motion and signaling:

1. Neutrophil: Neutrophils are white blood cells that play some very important roles in our innate immune system. They circulate around our body in the bloodstream, and when they sense signals that an infection is present, they are the first cells to migrate to the site of the infection to begin killing the invading microbes.

2. Macrophage: The macrophage is a large white blood cell that is an integral part of our immune system. Its job is to locate microscopic foreign bodies and 'eat' them. Macrophages use the process of phagocytosis to engulf particles and then digest them. Phagocytosis is the process by which a cell uses its

plasma membrane to engulf a large particle, giving rise to an internal compartment called the phagosome. It is one type of endocytosis. In a multicellular organism's immune system, phagocytosis is a major mechanism used to remove pathogens and cell debris.

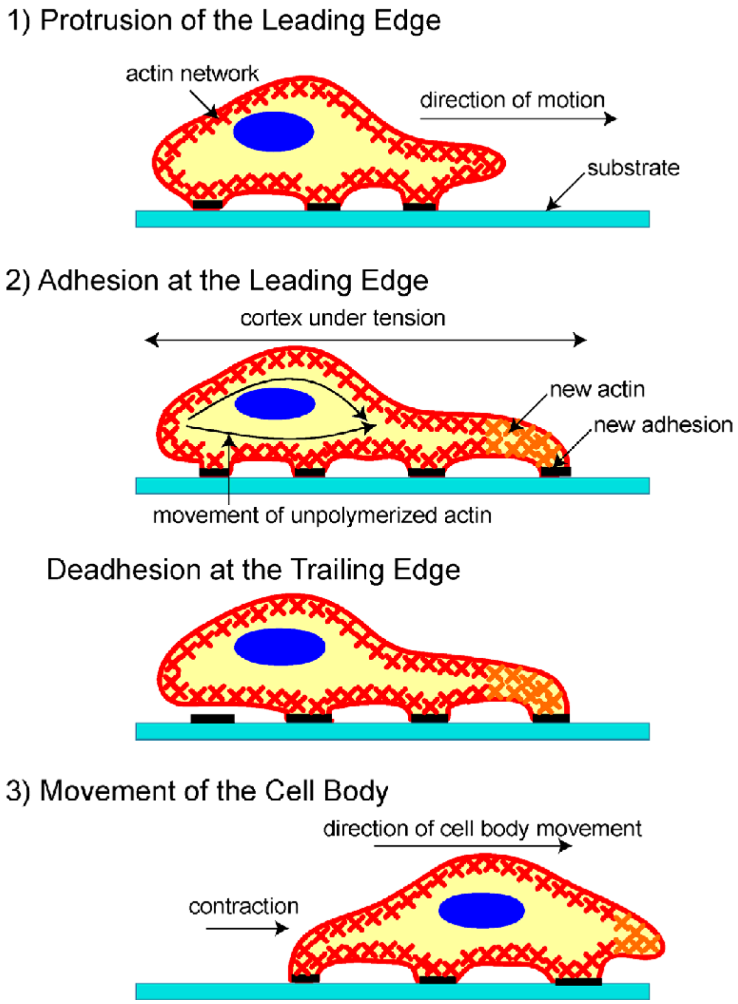


3. Neuron: Neurons (also known as neurones, nerve cells and nerve fibers) are electrically excitable cells in the nervous system that function to process and transmit information. In vertebrate animals, neurons are the core components of the brain, spinal cord and peripheral nerves.

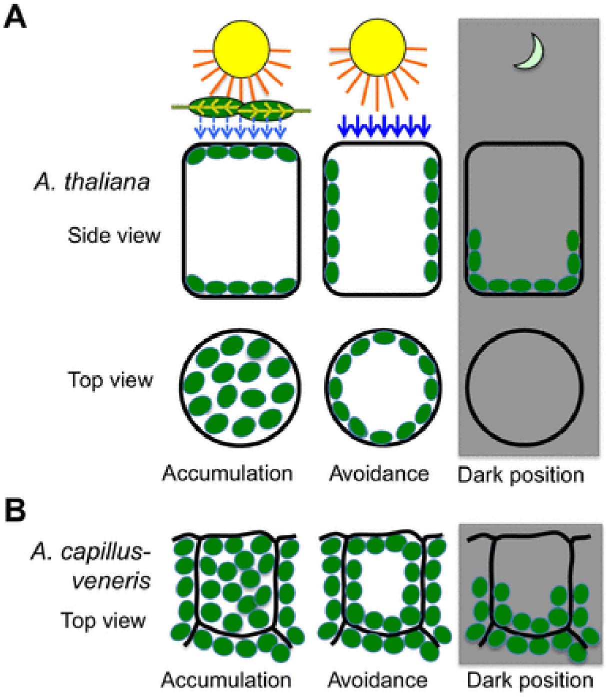
4. T cells: T cells are lymphocyte immune cells that protect the body from pathogens and cancer cells. T cells originate from bone marrow and mature in the thymus. They are important for cell mediated immunity and the activation of immune cells to fight infection.

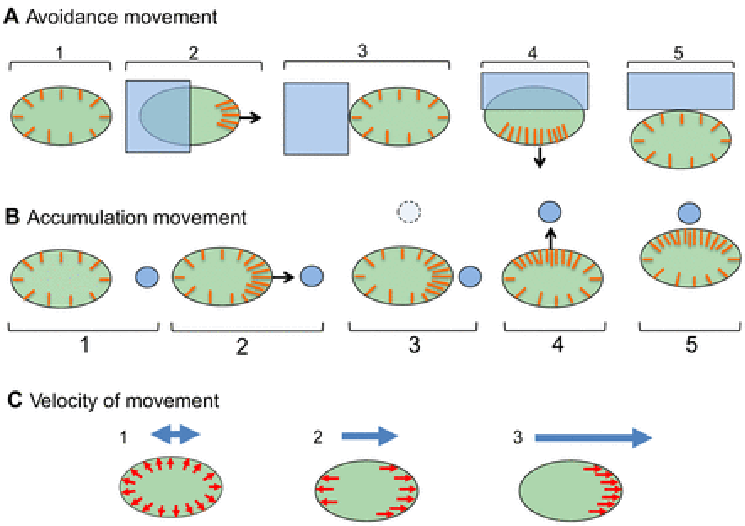
Examples of cell mechanics:

1. Actin motility: Motility is the ability of an organism to move independently, using metabolic energy. This is in contrast to mobility, which describes the ability of an object to be moved. Motility is genetically determined, but may be affected by environmental factors. Spatially controlled polymerization of actin is at the origin of cell motility and is responsible for the formation of cellular protrusions like lamellipodia.



2. Actin mediated movement: Plants are sessile and require diverse strategies to adapt to fluctuations in the surrounding light conditions. Consequently, the photorelocation movement of chloroplasts is essential to prevent damages that are induced by intense light (avoidance response) and to ensure efficient photosynthetic activities under weak light conditions (accumulation response). The mechanisms that underlie chloroplast movements have been revealed through analysis of the behavior of individual chloroplasts and it has been found that these organelles can move in any direction without turning. This implies that any part of the chloroplast periphery can function as the leading or trailing edge during movement. This ability is mediated by a special structure, which consists of short actin filaments that are polymerized at the leading edge of moving chloroplasts and are specifically localized in the space between the chloroplast and the plasma membrane, and is called chloroplast-actin. In addition, several of the genes that encode proteins that are involved in chloroplast-actin polymerization or maintenance have been identified.





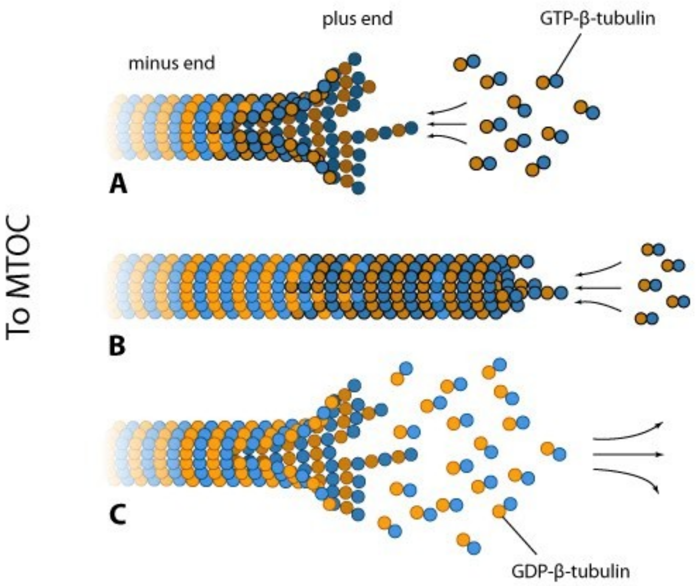
Cp-actin filaments in the directional and velocity control of chloroplast movements. (A) Avoidance movement induced by strong light. When half of a stable chloroplast (1) is illuminated with strong blue light [blue rectangle, (2)], cp-actin filaments (orange bars) at the irradiated part disappear and new cp-actin filaments polymerize at the opposite side (i.e. leading edge) (2). After the chloroplast has moved away from the irradiated area, cp-actin filaments become evenly distributed (3). If a second beam of light is applied to a different half of the chloroplast (see diagram), the asymmetric distribution of cp-actin filaments is re-established and the avoidance response re-occurs (4,5). (B) Accumulation movement induced by weak light. When part of a periclinal wall is illuminated with a microbeam (1) of either high or low intensity blue light (or red light in the case of ferns), chloroplasts that are positioned far from the spot migrate toward the irradiated area with an increased abundance of cp-actin filaments at the leading edge (2,3). The cp-actin filaments at the trailing edge are not depolymerized. If a second microbeam is applied (3, dotted circle), the moving chloroplast changes directions, with an increased abundance of cp-actin filaments at the new leading edge (4,5). (C) Velocity control through the balance of cp-actin filament amounts at the leading and trailing edges of a chloroplast. Red arrows indicate the vectors of centrifugal forces caused by cp-actin filament polymerization around the chloroplast periphery. (1) Stable chloroplast, for which the cp-actin filaments are evenly distributed around the chloroplast periphery so that all vector signals cancel each other. (2,3) In chloroplasts moving toward the right, their velocities are based on the sum of the centrifugal forces from the center to the chloroplast periphery. Blue arrows indicate the forces driving chloroplast movement.

3. Microtubule dynamics:

Microtubules are dynamic polymers of αβ-tubulin that are essential for intracellular organization, organelle trafficking and chromosome segregation. Microtubule growth and shrinkage occur via addition and loss of αβ-tubulin subunits, which are biochemical processes. Microtubules are highly dynamic and switch stochastically between growing and shrinking phases both in vivo and in vitro. This non-equilibrium behavior, known as dynamic instability, is based on the binding and hydrolysis of GTP at the nucleotide exchangeable site (E-site) in β-tubulin.

In most cell types, thirteen protofilaments associate laterally to form a microtubule. In a few cases microtubules contain more or fewer protofilaments. Numerous interactions between the subunits give microtubules their stiffness and resistance to bending forces. The lateral interactions between protofilaments are however comparatively weak and as a result the plus end of microtubules frequently have a ‘frayed’ appearance. Microtubules are highly dynamic and will frequently grow and shrink at a rapid yet constant rate. During this phenomenon, known as ‘dynamic instability’, tubulin subunits will both associate and dissociate from the plus end of the protofilament. A number of factors regulate the dynamics of microtubule formation however the primary determinant of whether microtubules grow or shrink is the rate of GTP hydrolysis, a factor that is both intrinsic and essential to filament assembly. In their stable state microtubles are hollow, cylindrical structures predominately composed of GDP-bound β-tubulin protofilaments. As GTP-bound protofilaments are straight, with multiple lateral contacts, assembly into the final cylindrical conformation is therefore dependent upon GTP hydrolysis. Without this, assembly would result in a flat sheet-like lattice of tubulin protofilaments.

Whilst GTP hyodrolysis of β-tubulin is therefore essential in filament production, it must be noted that the rate of assembly will often outpace the rate of hydrolysis. When this occurs a blunt end or GTP-cap is produced, which effectively constrains the curvature of the protofilaments.When hydrolysis does occur, the constraint is removed and the protofilaments become highly unstable as the stored energy in the lattice is released. This results in rapid shrinking of the microtubule. A typical microtubule will fluctuate every few minutes between growing and shrinking.



(A) The intracellular concentration of tubulin found in most cells (10-20µM) favors microtubule assembly at the plus end. This end may appear frayed in slow growing filaments due to fewer lateral interactions between the protofilaments and because GDP-tubulin dimers have an inherent curve.

(B) At high free GTP-tubulin dimer concentrations, hydrolysis is outpaced by rapid assembly at the plus end, thereby forming a rigid GTP-cap. (C) Concerted GTP hydrolysis at the plus end weakens the tubulin dimer interactions and the protofilaments rapidly disassemble.

4. Adhesion: Cell adhesion to the extracellular matrix (ECM) plays an essential role in regulation of fundamental cellular processes, such as cell survival, proliferation and migration.

Models of dynamics:

1. **Continuum model of cell migration (e.g. Keller & Segel -1971)**: The continuum modeling of biological tissue growth explicitly addresses cell division, using a homogenized representation of cells and their extracellular matrix (ECM). The model relies on the description of the cell as containing a solution of water and osmolytes, and having a porous solid matrix. The division of a cell into two nearly identical daughter cells is modeled as the doubling of the cell solid matrix and osmolyte content, producing an increase in water uptake via osmotic effects. This framework is also generalized to account for the growth of ECM-bound molecular species that impart a fixed charge density (FCD) to the tissue, such as proteoglycans. This FCD similarly induces osmotic effects, resulting in extracellular water uptake and osmotic pressurization of the ECM interstitial fluid, with concomitant swelling of its solid matrix.

	Advantages	Disadvantages
Discrete/Stochastic	Detailed structure and explicit implementation	High computational cost
	Direct connection to experimental data	Multiple simulations required
	Incorporation of randomness	Relatively inaccessible to mathematical analysis
Continuum/Deterministic	Fast to simulate	Lack of fine detailed structure
	Amenable to mathematical analysis	Difficult to link experimental data
	Suitable for systems of large numbers of cells	Ignore the effects of randomness

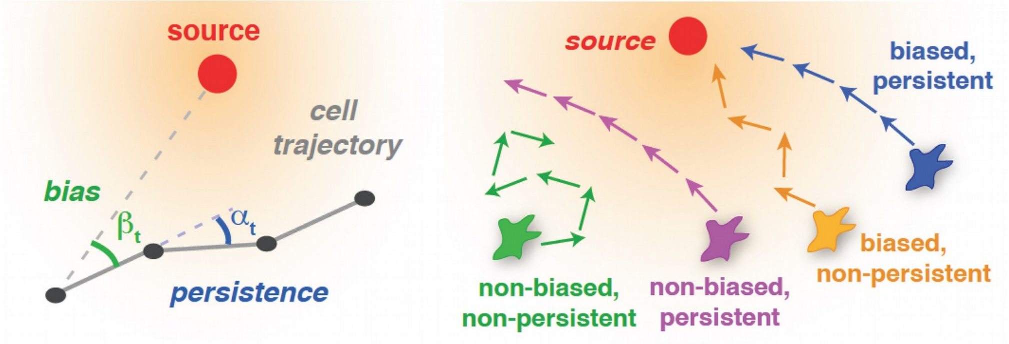
Collective cell migration plays a central role in tissue development, morphogenesis, wound repair and cancer progression. With the growing realization that physical forces mediate cell motility in development and physiology, a key biological question is how cells integrate molecular activities for force generation on multicellular scales.

2. Random walk with bias (e.g. Alt – 1980): There are different types of random walks that are commonly described in Biology. We can classify them into random walks that describe the step length distribution and random walks that describe the angular distributions. The definition of random walks via step length distribution is somewhat more frequently used. However, to investigate if a cell or a molecule is targeted in its movement, it is easier to look at angular distributions. A random walk can be compared to a drunkard trying to go home after a long pub night.Random walks are not always completely ‘random’. In this example, our walker has a clear target: home. So he is biased towards a specific direction. However, the more drunk he is the less able he is to follow his target. In other terms, he performs a biased random walk, with a bias strength defined by his drunkenness.

The most prominent random walk is Brownian motion. The angular distribution is isotropic, meaning that at each step a cell or molecule has equal probability to move in any direction. If we measure the angles between a motion vector (cell step) and a reference direction, we will find that the resulting angular distribution is flat (uniformly distributed). If, on the contrary, a cell has a specific target direction, then the cell has higher probability to move towards that target direction compared to all remaining directions. In this case we speak about a biased random walk. The expected angular distribution will have a peak at the angle which points towards the target direction. The remaining characteristics of the angular distribution of such biased random walk depend on the details of the exhibited walk, which are usually unknown. However, a commonly used description of the angular distribution is a wrapped normal distribution (a normal distribution wrapped around a circle to describe circular variables such as angles). The mean of the wrapped normal distribution indicates the bias direction and the variance indicates the strength of the bias. The lower the variance, the narrower is the distribution and the stronger is the exhibited bias.

A further type of random walk frequently used to describe animal movement and cell migration is a persistent random walk. A cell exhibiting this type of walk has higher probability of moving in the same direction as in the previous step compared to changing its direction. If we measure the angles between consecutive motion vectors (consecutive cell steps) we will observe a peak at 0, i.e. no change of direction. As for the biased random walk, the persistent random walk can also be described using a wrapped normal distribution with 0 mean and a variance which indicates the strength of the persistence (the lower the variance the stronger the persistence).

All three types of walks have been described for migration of immune and other cells, migration of animals, and movement of molecules inside the cell. Often a mix of these three types is observed.



Types of random walks and their analysis. Cell bias and persistence are computed from the angles βt (green, B) between the motion vector and the direction towards the source (red dot, B) and the angles αt (blue, B) between the current and proceeding motion vector, using an inference-based approach (D).

3. Stochastic or discreet model (e.g. Tranquillo – 1988): Scratch assays are routinely used to study the collective spreading of cell populations. In general, the rate at which a population of cells spreads is driven by the combined effects of cell migration and proliferation. To examine the effects of cell migration separately from the effects of cell proliferation, scratch assays are often performed after treating the cells with a drug that inhibits proliferation. In order to describe a scratch assay that incorporates the effects of cell-to-cell crowding, cell-to-cell adhesion, and dynamic changes in cell size, we present stochastic models are used that incorporates these mechanisms. The agent-based stochastic model takes the form of a system of Langevin equations that is the system of stochastic differential equations governing the evolution.

4. Hyperbolic continuum model (Hillen & Stevens – 2000): In the mathematical literature, there are various models that consider the effect of non-local social interactions on the collective movement of cells and animals. A large number of models for the collective movement of animals consider the interplay between all three social interactions: repulsion, attraction and alignment. It is very important to develop non-local models that consider cell polarization and describe the way that all three social forces affect the velocity and the turning behavior of cells. In this model, nonlinear non-local first-order hyperbolic equations are used which describe the dynamics of polarised early- and late-stage cancer cell populations. In addition to cell movement and cell turning behaviours (which depend on repulsive, attractive and polarising forces), mutation and proliferation have also been considered in this model. The patterns generated by this hyperbolic model are investigated by focusing on the effect of the following parameters: (i) magnitude of repulsive/attractive/polarisation (alignment) interactions; (ii) turning rates; (iii) proliferation rates; and (iv) baseline speed.

All these are top-down models.

1. Molecular dynamics models (~ $7\mathrm{x}10^{15}$ atoms per cell): Molecular dynamics (MD) is a computer simulation method for analyzing the physical movements of atoms and molecules. The atoms and molecules are allowed to interact for a fixed period of time, giving a view of the dynamic "evolution" of the system. It is an atomic approach, that is bottom-up.

2. Membrane models

3. Cytoskeleton models

4. Adhesion modulation models: In response to varying external biochemical and biophysical stimuli, cells regulate their adhesive state by modulating the number and binding capacity of their receptors to ligands. This modulation of adhesion can be modeled to explain cell mechanics, because cell adhesion regulates the motion of the cells.

These are all bottom-up models.

Knowns and Unknowns in solving the origin of life: Probability Estimate

Our uncertainty about the origin of life can be summed up succinctly as our ignorance in calculating the probability, P_{life}, for matter to transition from the non-living to living state. This is an important parameter to know, not just for understanding life on Earth, but also for estimating the distribution of life in the universe. Of the major historical unknowns in constraining the origin of life on our planet, the where, when and how life first emerged,- we have the least certainty in how and where. Only the timescale for life's emergence seems reasonably constrained. Fossil evidence for cellular life exists early in the geological record, dating back at least 3.5 billion years. Assuming the simplest case, that life did indeed emerge on Earth, we can reasonably assume that the early evidence for complex cellular life suggests that the origin of life was a rapid event (in a geological sense), occurring nearly as soon as conditions were favorable. A corollary to this argument is that since life appeared so rapidly on Earth, the origin of life must be a relatively common event on Earth-like worlds. One might therefore conclude that P_{life} -> 1, at least for Earth-like planets in the habitable zone of their parent star.

In the absence of discovery of life forms in other planets, with current empirical data it is impossible to determine with any certainty whether life is common or rare. The challenge arises because we have only one sample of life on which to inform estimates of P_{life}. All known life on Earth shares a common ancestry. Evidence for this common ancestry derives from phylogenetic reconstruction of the history of life on Earth, as inferred from modern organisms. Another line of evidence corroborating phylogenetic reconstruction is the existence of universal features of biochemistry shared by all organisms on Earth (such as DNA and RNA which allow phylogenetic histories to be generated in the first place). Because we are constrained by a single biochemical sample of life, despite our best attempts at logical arguments to the contrary, we cannot say whether the origin of life is easy or hard. That is, the probability P_{life} is unconstrained.

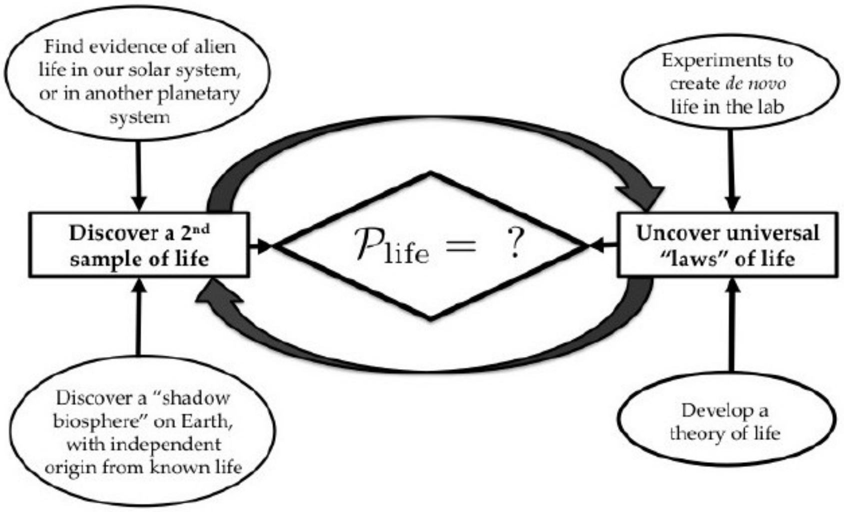


Figure 1. Possible constraints on $\mathcal{P}_{\text{life}}$, the probability that life emerges from non-living matter.

According to Carter's theory, based on Bayesian theorem of conditional probability, if d: observed data, t: theory, and t_r: life is rare and t_c denotes the event that life is common, then:

$$\frac{P(t_c|d)}{P(t_r|d)} = \frac{P(d|t_c)P(t_c)}{P(d|t_r)P(t_r)}$$

Carter's point is this: the effects
of experimental bias and observational selection must be taken into account when computing the values for the likelihood of the observed data (P(d 'given' t_c) and P(d 'given' t_r)) on the right-hand side of the above equation. These represent probabilities based on current knowledge, and are often incomplete - that is, they are not ab initio probabilities derived from a fundamental theory and do not represent absolute states of knowledge.

Multicellularity

Human body is made up of 100 trillion cells.

Cell types: Bone cell, nerve cell, ovum cell, sperm cell, blood cell etc.

Stem Cells

Stem cells are unique in that they originate as unspecialized cells and have the ability to develop into specialized cells that can be used to build specific organs or tissues. Stem cells can divide and replicate many times in order to replenish and repair tissue. In the field of stem cell research, scientists take advantage of the renewal properties of these structures by utilizing them to generate cells for tissue repair, organ transplantation, and for the treatment of disease.

Bone Cells

Bones are a type of mineralized connective tissue that comprise a major component of the skeletal system. Bones are made up of a matrix of collagen and calcium phosphate minerals. There are three primary types of bone cells in the body: osteoclasts, osteoblasts, and osteocytes. Osteoclasts are large cells that decompose bone for resorption and assimilation while they heal. Osteoblasts regulate bone mineralization and produce osteoid, an organic substance of the bone matrix, which mineralizes to form bone. Osteoblasts mature to form osteocytes. Osteocytes aid in the formation of bone and help maintain calcium balance.

Blood Cells

From transporting oxygen throughout the body to fighting infection, blood cell activity is vital to life. Blood cells are produced by bone marrow. The three major types of cells in the blood are red blood cells, white blood cells, and platelets. Red blood cells determine blood type and are responsible for transporting oxygen. White blood cells are immune system cells that destroy pathogens and provide immunity. Platelets help clot blood to prevent excessive blood loss due to broken or damaged blood vessels.

Muscle Cells

Muscle cells form muscle tissue, which enables all bodily movement. The three types of muscle cells are skeletal, cardiac, and smooth. Skeletal muscle tissue attaches to bones to facilitates voluntary movement. These muscle cells are covered by connective tissue, which protects and supports muscle fiber bundles.

Cardiac muscle cells form involuntary muscle, or muscle that doesn't require conscious effort to operate, found in the heart. These cells aid in heart contraction and are joined to one another by intercalated discs that allow for heartbeat synchronization. Smooth muscle tissue is not striated like cardiac and skeletal muscle. Smooth muscle is involuntary muscle that lines body cavities and forms the walls of many organs such as kidneys, intestines, blood vessels, and lung airways.

Fat Cells

Fat cells, also called adipocytes, are a major cell component of adipose tissue. Adipocytes contain droplets of stored fat (triglycerides) that can be used for energy. When fat is stored, its cells become round and swollen. When fat is used, its cells shrink. Adipose cells also have a critical endocrine function: they produce hormones that influence sex hormone metabolism, blood pressure regulation, insulin sensitivity, fat storage and use, blood clotting, and cell signaling.

Skin Cells

The skin is composed of a layer of epithelial tissue (epidermis) that is supported by a layer of connective tissue (dermis) and an underlying subcutaneous layer. The outermost layer of the skin is composed of flat, squamous epithelial cells that are closely packed together. The skin covers a wide range of roles. It protects internal structures of the body from damage, prevents dehydration, acts as a barrier against germs, stores fat, and produces vitamins and hormones.

Nerve Cells

Nerve cells or neurons are the most basic unit of the nervous system. Nerves send signals between the brain, spinal cord, and other body organs via nerve impulses. Structurally, a neuron consists of a cell body and nerve processes. The central cell body contains the neuron's nucleus, associated cytoplasm, and organelles. Nerve processes are "finger-like" projections (axons and dendrites) that extend from the cell body and transmit signals.

Endothelial Cells

Endothelial cells form the inner lining of the cardiovascular system and lymphatic system structures. They make up the inner layer of blood vessels, lymphatic vessels, and organs including the brain, lungs, skin, and heart. Endothelial cells are responsible for angiogenesis or the creation of new blood vessels. They also regulate the movement of macromolecules, gases, and fluid between the blood and surrounding tissues as well as help manage blood pressure.

Sex Cells

Sex cells or gametes are reproductive cells created in male and female gonads that bring new life into existence. Male sex cells or sperm are motile and have long, tail-like projections called flagella. Female sex cells or ova are non-motile and relatively large in comparison to male gametes. In sexual reproduction, sex cells unite during fertilization to form a new individual. While other body cells replicate by mitosis, gametes reproduce by meiosis.

Pancreatic Cells

The pancreas functions as both an exocrine and endocrine organ, meaning that it discharges hormones both through ducts and directly into other organs. Pancreatic cells are important for regulating blood glucose concentration levels as well as for the digestion of proteins, carbohydrates, and fats.

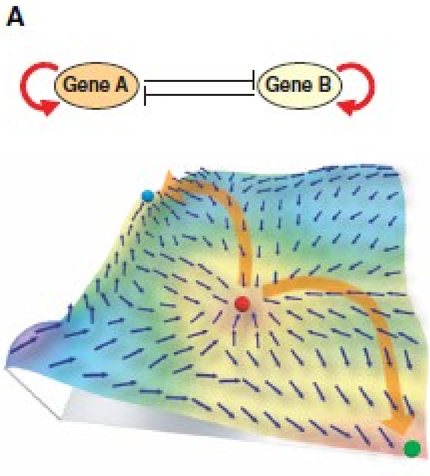
Exocrine acinar cells, which are produced by the pancreas, secrete digestive enzymes that are transported by ducts to the small intestine. A very small percentage of pancreatic cells have an endocrine function or secrete hormones into cells and tissues. Pancreatic endocrine cells are found in small clusters called islets of Langerhans. Hormones produced by these cells include insulin, glucagon, and gastrin.

Cancer Cells

Unlike all of the other cells listed, cancer cells work to destroy the body. Cancer results from the development of abnormal cell properties that cause cells to divide uncontrollably and spread to other locations. Cancer cell development can originate from mutations stemming from exposure to chemicals, radiation, and ultraviolet light. Cancer can also have genetic origins such as chromosome replication errors and cancer-causing viruses of the DNA. Cancer cells are allowed to spread rapidly because they develop decreased sensitivity to anti-growth signals and proliferate quickly in the absence of stop commands. They also lose the ability to undergo apoptosis or programmed cell death, making them even more formidable.

Cell types as Attractors in Cell Dynamics

Cells contain many components, including genes, proteins, and metabolites. The cellular state at a particular time can be represented as a point in multidimensional state space in which each axis represents the abundance of a component. Gene (or protein) expressions are a major part of such components (for simplicity, we write “gene expression level” to describe the abundance). Interplay among genes, such as the activation and repression of gene expressions, causes the cellular state to shift, a phenomenon that can be depicted as a trajectory in the state space. Temporal changes in the expressions restrict the cellular state to a certain region, which is defined as an “attractor” in dynamical-systems theory. After a slight perturbation (change in gene expression levels), a state returns to its original attractor with the aforementioned temporal change. The attractor can be in a fixed state over time (i.e., fixed-point attractor) where the synthesis and degradation of each product are balanced, or a set of dynamically changing states with temporally oscillating gene expressions (e.g., orange trajectory in the figure). A system can have multiple attractors of different composition. Each attractor then can be regarded as a distinct cell type corresponding to the different valleys into which a ball can fall in Waddington’s landscape. For example, when two genes, A and B, mutually suppress each other’s expression (known as a toggle switch), two fixed-point attractors form: one with activated A and suppressed B, and the other with activated B and suppressed A. These two attractors can be regarded as two distinct cell types.



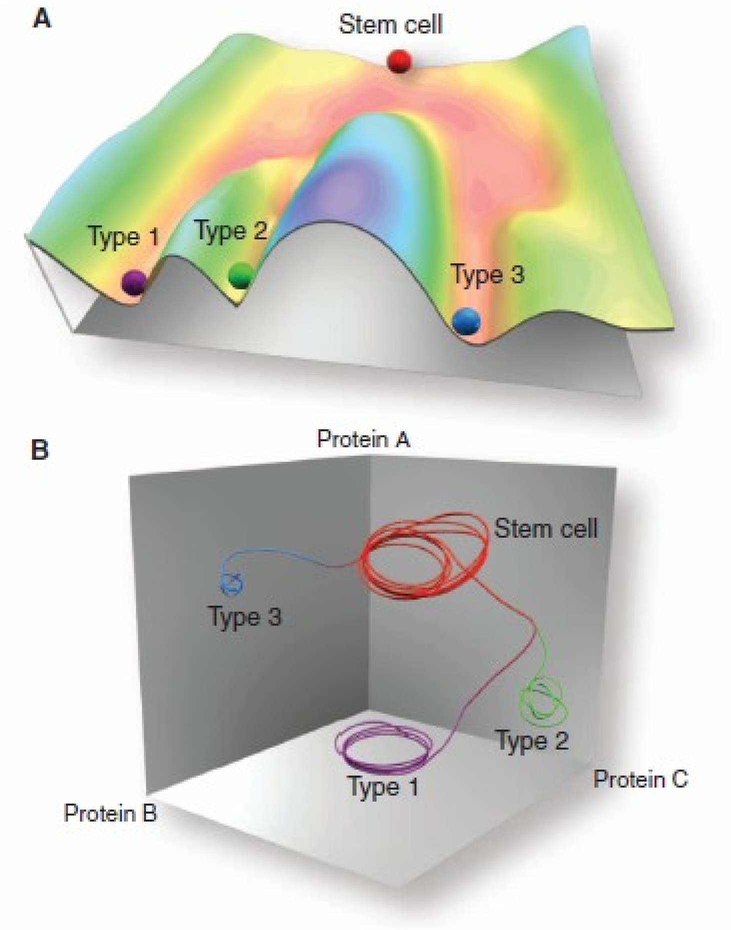


Fig. 1. (**A**) Waddington's epigenetic landscape. The development of a cellular state is represented by a ball rolling down a landscape of bifurcating valleys, each representing different cell types. (**B**) Dynamical-systems representation of cellular states. Each axis represents the expression of a protein whose time development is depicted as a trajectory in space. Final states are attractors and correspond to distinct cell types.

Stem Cell and Cellular Differentiation

Stem cell is a primitive cell which can either self renew (reproduce itself) or give rise to more specialised cell types. The stem cell is the ancestor at the top of the family tree of related cell types. One blood stem cell gives rise to red cells, white cells and platelets. Stem cells are the foundation for every organ and tissue in your body. There are many different types of stem cells that come from different places in the body or are formed at different times in our lives. These include embryonic stem cells that exist only at the earliest stages of development and various types of tissue-specific (or adult) stem cells that appear during fetal development and remain in our bodies throughout life.

All stem cells can self-renew (make copies of themselves) and differentiate (develop into more specialized cells). Beyond these two critical abilities, though, stem cells vary widely in what they can and cannot do and in the circumstances under which they can and cannot do certain things. This is one of the reasons researchers use all types of stem cells in their investigations.

Types of Stem Cells:

Embryonic stem cells

Tissue-specific stem cells

Mesenchymal stem cells

Induced pluripotent stem cells

Embryonic stem cells

Embryonic stem cells are obtained from the inner cell mass of the blastocyst, a mainly hollow ball of cells that, in the human, forms three to five days after an egg cell is fertilized by a sperm. A human blastocyst is about the size of the dot above this "i."

In normal development, the cells inside the inner cell mass will give rise to the more specialized cells that give rise to the entire body—all of our tissues and organs. However, when scientists extract the inner cell mass and grow these cells in special laboratory conditions, they retain the properties of embryonic stem cells. Embryonic stem cells are pluripotent, meaning they can give rise to every cell type in the fully formed body, but not the placenta and umbilical cord. These cells are incredibly valuable because they provide a renewable resource for studying normal development and disease, and for testing drugs and other therapies. Human embryonic stem cells have been derived primarily from blastocysts created by in vitro fertilization (IVF) for assisted reproduction that were no longer needed.

Tissue-specific stem cells

Tissue-specific stem cells (also referred to as somatic or adult stem cells) are more specialized than embryonic stem cells. Typically, these stem cells can generate different cell types for the specific tissue or organ in which they live.

For example, blood-forming (or hematopoietic) stem cells in the bone marrow can give rise to red blood cells, white blood cells and platelets. However, blood-forming stem cells don't generate liver or lung or brain cells, and stem cells in other tissues and organs don't generate red or white blood cells or platelets.

Some tissues and organs within your body contain small caches of tissue-specific stem cells whose job it is to replace cells from that tissue that are lost in normal day-to-day living or in injury, such as those in your skin, blood, and the lining of your gut.

Tissue-specific stem cells can be difficult to find in the human body, and they don't seem to self-renew in culture as easily as embryonic stem cells do. However, study of these cells has increased our general knowledge about normal development, what changes in aging, and what happens with injury and disease.

Mesenchymal stem cells:

You may hear the term "mesenchymal stem cell" or MSC to refer to cells isolated from stroma, the connective tissue that surrounds other tissues and organs. Cells by this name are more accurately called “stromal cells” by many scientists. The first MSCs were discovered in the bone marrow and were shown to be capable of making bone, cartilage and fat cells. Since then, they have been grown from other tissues, such as fat and cord blood. Various MSCs are thought to have stem cell, and even immunomodulatory, properties and are being tested as treatments for a great many disorders, but there is little evidence to date that they are beneficial. Scientists do not fully understand whether these cells are actually stem cells or what types of cells they are capable of generating. They do agree that not all MSCs are the same, and that their characteristics depend on where in the body they come from and how they are isolated and grown.

Induced pluripotent stem cells

Induced pluripotent stem (iPS) cells are cells that have been engineered in the lab by converting tissue-specific cells, such as skin cells, into cells that behave like embryonic stem cells. IPS cells are critical tools to help scientists learn more about normal development and disease onset and progression, and they are also useful for developing and testing new drugs and therapies.

While iPS cells share many of the same characteristics of embryonic stem cells, including the ability to give rise to all the cell types in the body, they aren’t exactly the same. Scientists are exploring what these differences are and what they mean. For one thing, the first iPS cells were produced by using viruses to insert extra copies of genes into tissue-specific cells. Researchers are experimenting with many alternative ways to create iPS cells so that they can ultimately be used as a source of cells or tissues for medical treatments.

Differentiation

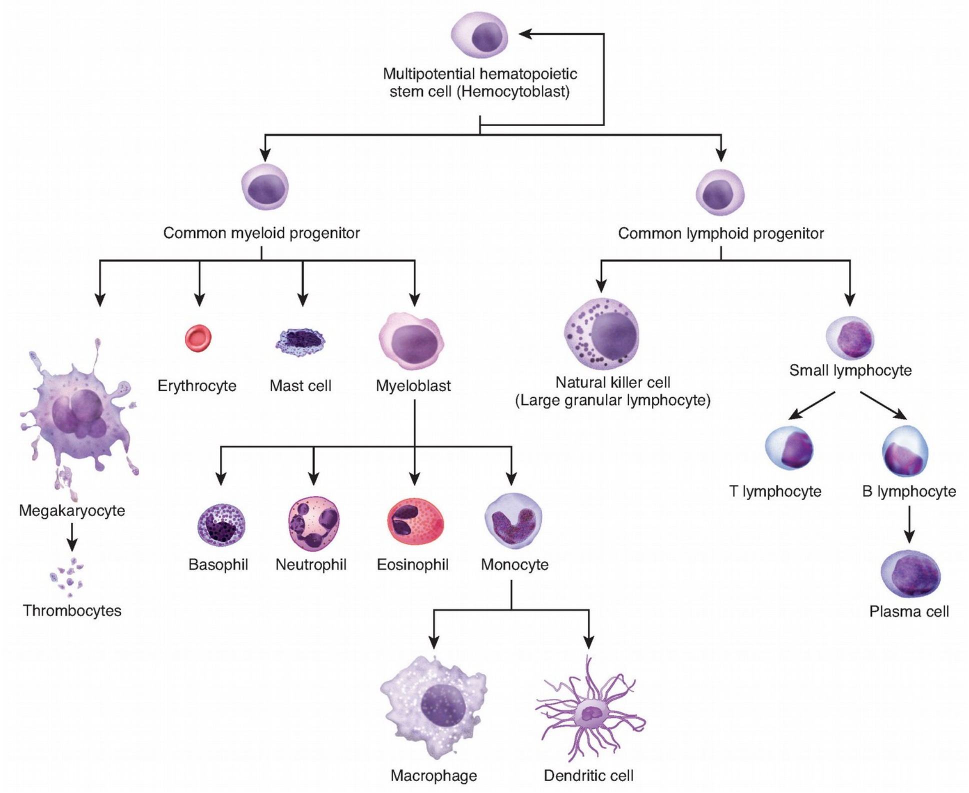
How does a complex organism such as a human develop from a single cell—a fertilized egg—into the vast array of cell types such as nerve cells, muscle cells, and epithelial cells that characterize the adult? Throughout development and adulthood, the process of cellular differentiation leads cells to assume their final morphology and physiology. Differentiation is the process by which unspecialized cells become specialized to carry out distinct functions.

A stem cell is an unspecialized cell that can divide without limit as needed and can, under specific conditions, differentiate into specialized cells. Stem cells are divided into several categories according to their potential to differentiate.

The first embryonic cells that arise from the division of the zygote are the ultimate stem cells; these stems cells are described as totipotent because they have the potential to differentiate into any of the cells needed to enable an organism to grow and develop. The embryonic cells that develop from totipotent stem cells and are precursors to the fundamental tissue layers of the embryo are classified as pluripotent. A pluripotent stem cell is one that has the potential to differentiate into any type of human tissue but cannot support the full development of an organism. These cells then become slightly more specialized, and are referred to as multipotent cells. A multipotent stem cell has the potential to differentiate into different types of cells within a given cell lineage or small number of lineages, such as a red blood cell or white blood cell.

Finally, multipotent cells can become further specialized oligopotent cells. An oligopotent stem cell is limited to becoming one of a few different cell types. In contrast, a unipotent cell is fully specialized and can only reproduce to generate more of its own specific cell type.

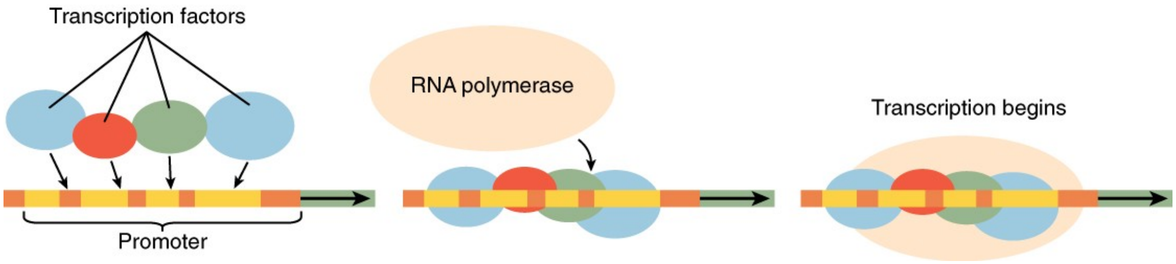
Stem cells are unique in that they can also continually divide and regenerate new stem cells instead of further specializing. There are different stem cells present at different stages of a human’s life. They include the embryonic stem cells of the embryo, fetal stem cells of the fetus, and adult stem cells in the adult. One type of adult stem cell is the epithelial stem cell, which gives rise to the keratinocytes in the multiple layers of epithelial cells in the epidermis of skin. Adult bone marrow has three distinct types of stem cells: hematopoietic stem cells, which give rise to red blood cells, white blood cells, and platelets (Figure 1); endothelial stem cells, which give rise to the endothelial cell types that line blood and lymph vessels; and mesenchymal stem cells, which give rise to the different types of muscle cells.



Hematopoiesis. The process of hematopoiesis involves the differentiation of multipotent cells into blood and immune cells. The multipotent hematopoietic stem cells give rise to many different cell types, including the cells of the immune system and red blood cells.

When a cell differentiates (becomes more specialized), it may undertake major changes in its size, shape, metabolic activity, and overall function. Because all cells in the body, beginning with the fertilized egg, contain the same DNA, how do the different cell types come to be so different? The answer is analogous to a movie script. The different actors in a movie all read from the same script, however, they are each only reading their own part of the script. Similarly, all cells contain the same full complement of DNA, but each type of cell only "reads" the portions of DNA that are relevant to its own function. In biology, this is referred to as the unique genetic expression of each cell.

In order for a cell to differentiate into its specialized form and function, it need only manipulate those genes (and thus those proteins) that will be expressed, and not those that will remain silent. The primary mechanism by which genes are turned "on" or "off" is through transcription factors. A transcription factor is one of a class of proteins that bind to specific genes on the DNA molecule and either promote or inhibit their transcription.



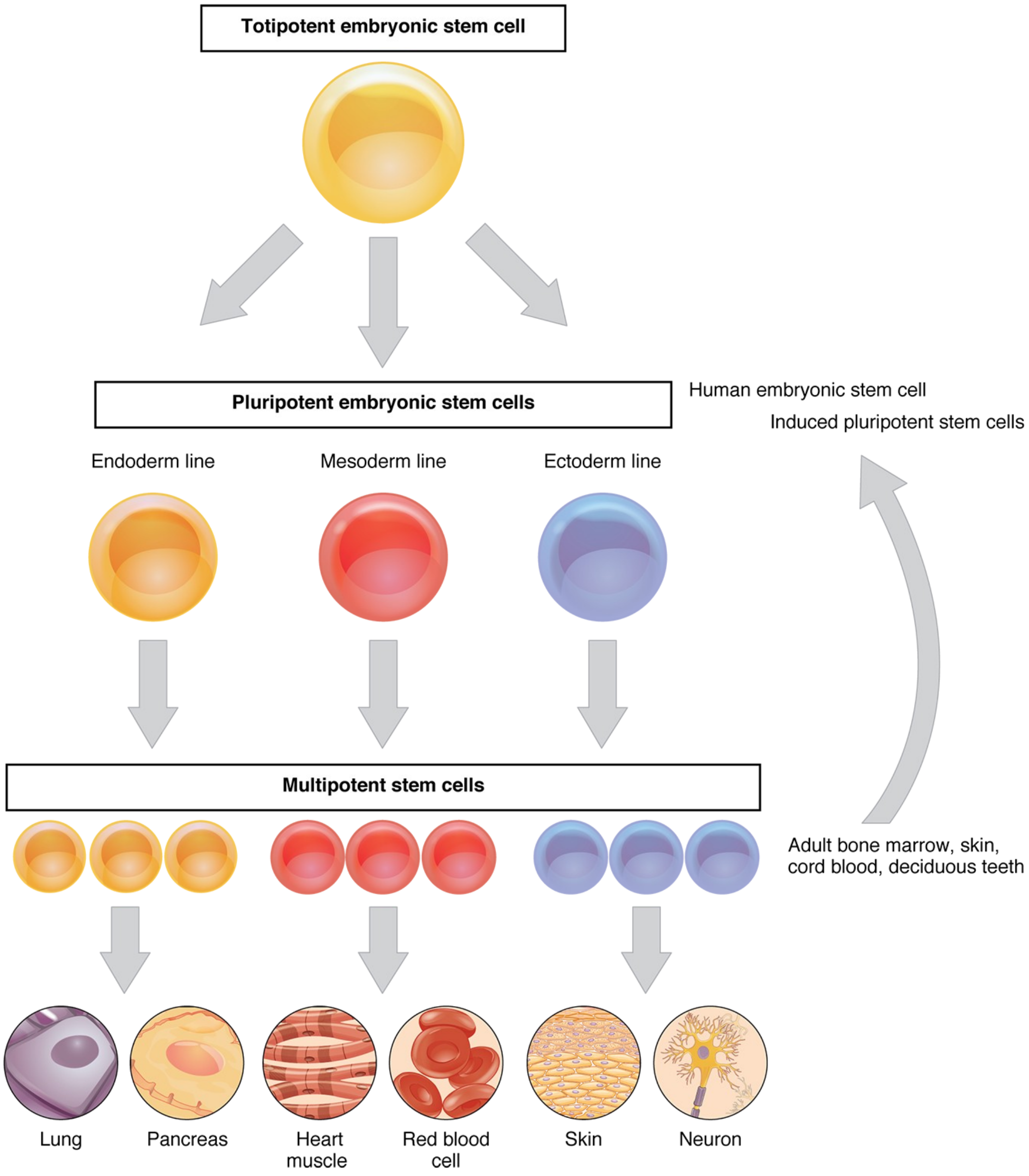
Transcription Factors Regulate Gene Expression. While each body cell contains the organism's entire genome, different cells regulate gene expression with the use of various transcription factors. Transcription factors are proteins that affect the binding of RNA polymerase to a particular gene on the DNA molecule.

Stem Cell Research

Stem cell research aims to find ways to use stem cells to regenerate and repair cellular damage. Over time, most adult cells undergo the wear and tear of aging and lose their ability to divide and repair themselves. Stem cells do not display a particular morphology or function. Adult stem cells, which exist as a small subset of cells in most tissues, keep dividing and can differentiate into a number of specialized cells generally formed by that tissue. These cells enable the body to renew and repair body tissues.

The mechanisms that induce a non-differentiated cell to become a specialized cell are poorly understood. In a laboratory setting, it is possible to induce stem cells to differentiate into specialized cells by changing the physical and chemical conditions of growth. Several sources of stem cells are used experimentally and are classified according to their origin and potential for differentiation. Human embryonic stem cells (hESCs) are extracted from embryos and are pluripotent. The adult stem cells that are present in many organs and differentiated tissues, such as bone marrow and skin, are multipotent, being limited in differentiation to the types of cells found in those tissues. The stem cells isolated from umbilical cord blood are also multipotent, as are cells from deciduous teeth (baby teeth). Researchers have recently developed induced pluripotent stem cells (iPSCs) from mouse and human adult stem cells. These cells are genetically reprogrammed multipotent adult cells that function like embryonic stem cells; they are capable of generating cells characteristic of all three germ layers.

Because of their capacity to divide and differentiate into specialized cells, stem cells offer a potential treatment for diseases such as diabetes and heart disease. Cell-based therapy refers to treatment in which stem cells induced to differentiate in a growth dish are injected into a patient to repair damaged or destroyed cells or tissues. Many obstacles must be overcome for the application of cell-based therapy. Although embryonic stem cells have a nearly unlimited range of differentiation potential, they are seen as foreign by the patient's immune system and may trigger rejection. Also, the destruction of embryos to isolate embryonic stem cells raises considerable ethical and legal questions.



Stem Cells. The capacity of stem cells to differentiate into specialized cells make them potentially valuable in therapeutic applications designed to replace damaged cells of different body tissues.

In contrast, adult stem cells isolated from a patient are not seen as foreign by the body, but they have a limited range of differentiation. Some individuals bank the cord blood or deciduous teeth of their child, storing away those sources of stem cells for future use, should their child need it. Induced pluripotent stem cells are considered a promising advance in the field because using them avoids the legal, ethical, and immunological pitfalls of embryonic stem cells.

Review Questions

1. Arrange the following terms in order of increasing specialization: oligopotency, pleuripotency, unipotency, multipotency.

multipotency, pleuripotency, oligopotency, unipotency

pleuripotency, oligopotency, multipotency unipotency

oligopotency, pleuripotency, unipotency, multipotency

pleuripotency, multipotency, oligopotency, unipotency

2. Which type of stem cell gives rise to red and white blood cells?

endothelial

epithelial

hematopoietic

mesenchymal

3. What multipotent stem cells from children sometimes banked by parents?

fetal stem cells

embryonic stem cells

cells from the umbilical cord and from baby teeth

hematopoietic stem cells from red and white blood cells

Pattern Formation and Development

Evolution of metazoan organisms has produced, over hundreds of millions of years, both phenotypic complexity and the developmental mechanisms by which such complexity is generated. During development a single cell becomes an organism composed of multiple cell types arranged in spatial distributions that can be both architecturally complex and functionally coherent. How this distribution of cellular phenotypes (‘cell states’) is attained through spatiotemporal regulation of gene interactions and cell behaviors is one of the main questions of developmental biology. To this end, considerable knowledge has been acquired during the last few decades about the genetic composition of multicellular organisms, how various genes and gene products interact, where are they expressed, and in which developmental processes are they involved.

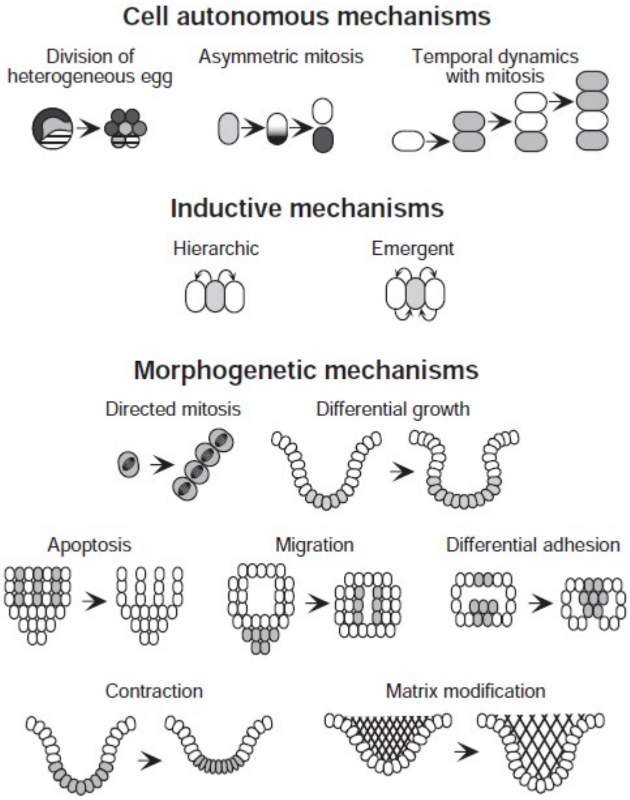
Organismal development is enabled by developmental mechanisms. A developmental mechanism can be explained as gene product interactions and changes in cellular behaviors (such as mitosis, apoptosis, secretion of molecular signals, cellular adhesion, differentiation, and so forth) that are required for and cause the formation of a particular arrangement of cell states in three-dimensional space (i.e., a ‘pattern’). In formal terms the development of an organism can be described as transformation from one set of patterns to another set of patterns.

Causal explanations of pattern formation in an embryonic primordium require knowledge of all the genes, epigenetic determinants (that is, surrounding cell arrangements and other microenvironmental conditions in the embryo), and their interactions necessary for generating such a pattern from a previous pattern. In practice, causality can be inferred by testing how well a developmental mechanism predicts the ‘variational’ properties – the range of potential morphological outcomes. It is common in theoretical discussions of development to distinguish two components of pattern formation. First, pattern formation through cell-cell signaling mechanisms (inductive mechanisms) establishes cells with different states and different spatial relationships by signaling in two and three dimensions in developing planar and solid tissues, respectively. Second, mechanisms that use cell behaviors other than signaling (morphogenetic mechanisms) act on the previously established pattern to cause the formation of three dimensional tissues and organs. Morphogenetic mechanisms change the spatial distribution of cells without changing cell states.

Morphogenetic and inductive mechanism act at all stages of development. Inductive mechanisms are generally implicated in developmental changes that produce new patterns. Induction is a prerequisite for development to proceed. However, the relative timing, including possible coincidence, of inductive and morphogenetic mechanisms can have major consequences for developmental dynamics and the range of potential morphological outcomes, and is therefore of central importance for the understanding of both development and morphological evolution.

Basic Developmental Mechanism leading to various patterning:

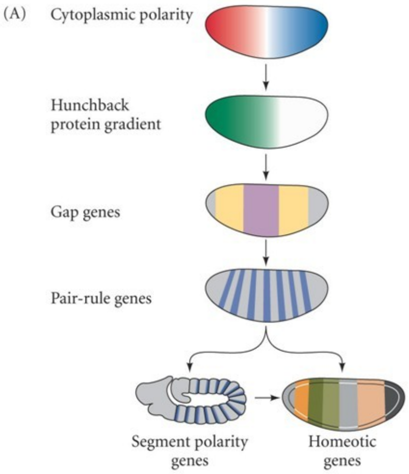
1. Cell autonomous mechanism Cell autonomous developmental mechanisms all involve one cellular behavior: mitosis. Thus, cells do not interact mechanically or by signaling.



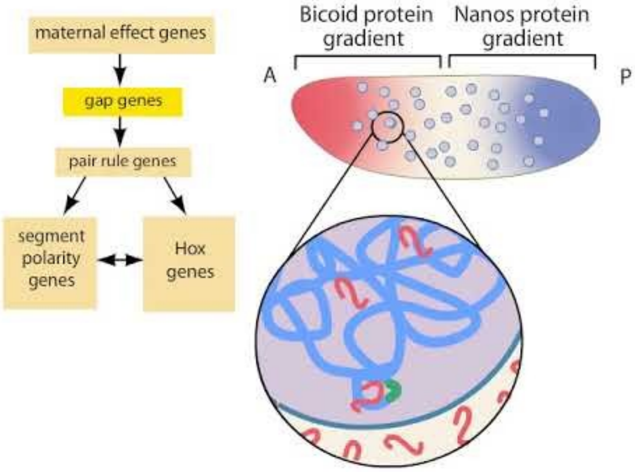
Schematic examples of the basic developmental mechanisms. Division of an heterogeneous egg: different parts of the egg bind different molecules (indicated by different shading) resulting in different blastomere cells. Asymmetric mitosis: molecules are differentially transported into different parts of a cell resulting in different daughter cells. Internal temporal dynamics coupled to mitosis: cells that have oscillating levels of molecules before their division can produce spatial patterns. Hierarchic induction: inducing cell (gray) affects neighboring cells but the induced cells (white) do not affect the production of the inducing signal. Emergent induction: inducing cell affects neighboring cells, which in turn signal back affecting the production of the inducing signal. Directed mitosis: consistently oriented mitotic spindles may direct tissue growth. Differential growth: cells dividing at a higher rate (gray) can alter tissue shape. Apoptosis: transformation of an established pattern into another can result from apoptosis affecting specific cells (gray). Migration: cells can migrate to a new location. Adhesion: a change in pattern can result if a set of cells have differential adhesion properties (strong adhesion among gray cells). Contraction: differential contraction of cells can cause buckling of a tissue. Matrix swelling, deposition, and loss: matrix swelling can cause budding.

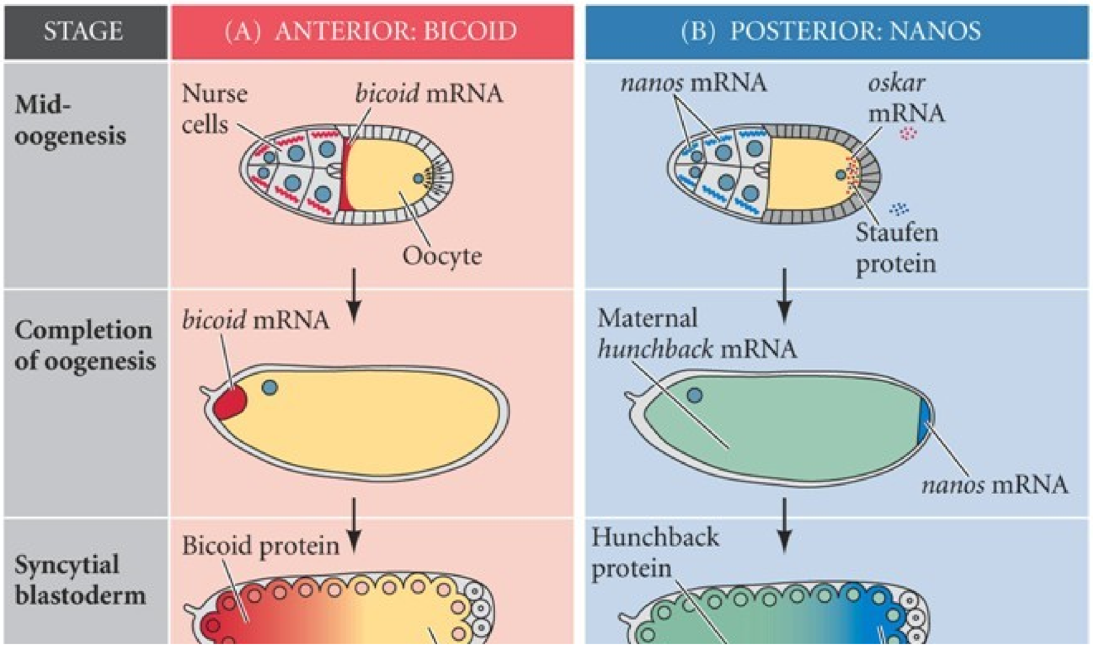
2. Inductive Mechanism Cells can affect each other by secreting diffusible molecules, by means of membrane-bound molecules or by chemical coupling through gap junctions. A large number of mechanisms which use only these developmental functions are capable of pattern formation. In inductive mechanisms tissue pattern changes as a direct consequence of changes in cell state. This, in turn, is due to

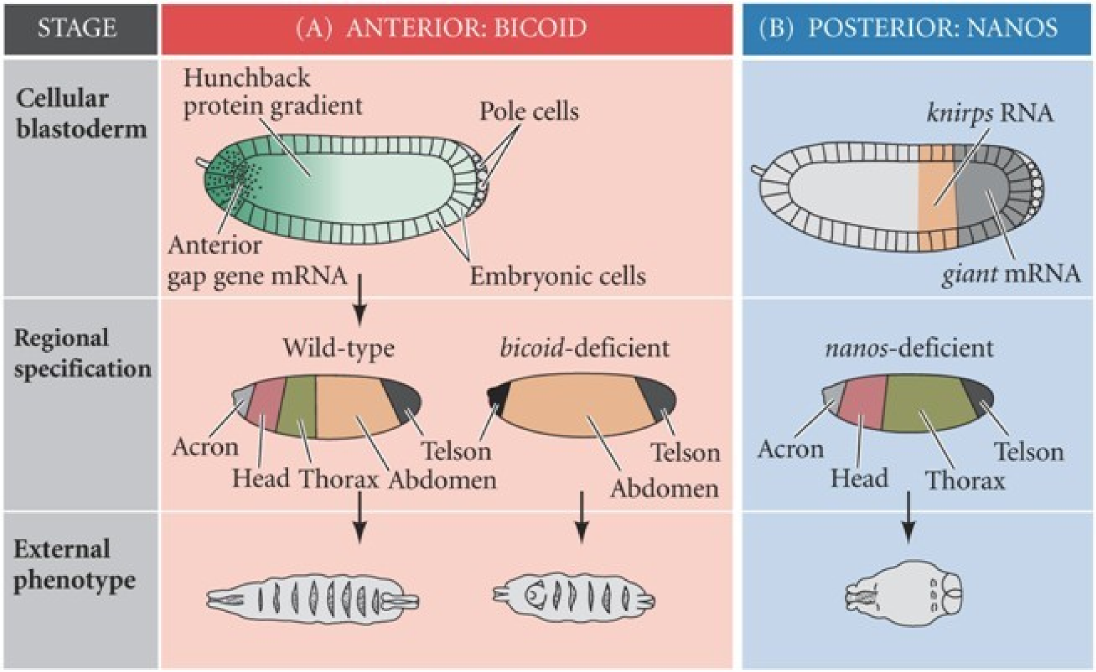
the processing or interpretation of signals sent by other cells. In certain cases, inductive pattern formation assumes a simple form, that is, one cell or tissue type will change the state of another cell or tissue type from what it would have been without the interaction, with no morphological consequence following directly from this. In other cases a morphological consequence accompanies, or follows closely upon, the change in state of the induced target cells. Examples include generation of the gradient patterns of gap gene products in the Drosophila syncytial blastula induced by the patterns of maternal gene products, and the subsequent induction of striped patterns of pair-rule gene products, based on these gap patterns.



DEVELOPMENTAL BIOLOGY, Eighth Edition, Figure 9.17 (Part 1) © 2006 Sinauer Associates, Inc.







DEVELOPMENTAL BIOLOGY, Eighth Edition, Figure 9.19 (Part 2) © 2006 Sinauer Associates, Inc.

<u>3. Morphogenetic mechanism</u>

A number of patterning mechanisms use cellular behaviors other than signaling (although signaling may have been active at a prior stage). These mechanisms alter pattern by affecting form. This can be defined as a mechanism that changes the relative arrangement of cells over space without affecting their states.

a) Directed Mitosis: Intracellular or extracellular signals can affect the direction of the mitotic spindle. Once the mitotic spindle assumes a set direction, new cells are forced to be positioned at specific places. The central nervous system of Drosophila, for example, forms by the dorsally directed budding of presumptive neuroblasts from the ectoderm.

b) Differential Growth: A change in a pattern can be produced if, in a previously existing pattern, cells with different states divide at different rates. The new pattern depends on the previous pattern, the relative rates and directions of mitosis and on other epigenetic factors such as the adhesion between cells and the influences of surrounding matrices. One such example is the establishment, maintenance, and waning of the growth plate during the formation of long bones in vertebrates.

c) Adoptosis: A pattern can be transformed into another if some of the cells undergo apoptosis. Apoptosis can be strictly dependent on a cell’s lineage, or triggered by interaction, or abrogation of interaction, with surrounding cells. (Adoptosis is the death of cells which occurs as a normal and controlled part of an organism's growth or development.)

d) Migration: Cells can rearrange their relative positions without changing their states simply by migrating. Migration can be directionally random, random but speeded up by an ambient chemical signal (‘chemokinesis’), or have a preferred direction in relation to a chemical gradient (‘chemotaxis’) or an insoluble substrate gradient (‘haptotaxis’). The migration of premuscle cells into the developing vertebrate limb is regulated by both chemokinetic and chemotactic responses to hepatocyte growth factor.

e) Differential adhesion: Cell adhesion is the defining property of multicellular organisms. It is an indispensable requirement for cell shape, differentiation and migration. A large, but limited number of pattern changes can be produced in tissues by constituent cells expressing different adhesion molecules or the same molecules at different levels.

f) Contraction: Individual cell contraction mediated by actin-myosin complexes can have morphogenetic effects on neighboring cells and the tissue as a whole.

g) Matrix swelling, deposition and loss: The cells of mesenchymal and connective tissues are surrounded and separated by semi-solid or solid extracellular matrices. Changes in pattern may be accomplished by increased hydration or swelling of a preexisting matrix, increase in the amount of matrix separating the cells, or matrix degradation. During development of the avian eye, the primary corneal stroma swells in anticipation of its invasion by mesenchymal cells from the periphery.

Brain Structure: neurons and neural networks

Neurons and neural networks.pdf

Brain as an information processing system

Brain as information processing system.pdf

Associative memory models

Associative memory models.pdf

Neural network dynamics and memory

Attractor networks have been proposed as models of learning and memory in the cerebral cortex (Hopkins). In these models, synaptic connectivity in a recurrent neural network is set up in such a way that the network dynamics have multiple attractor states, each of which represents a particular item that is stored in memory. Each attractor state is a specific pattern of activity of the network that is correlated with the state of the network when the particular item is presented through external inputs. The attractor property means that the network converges to the stored pattern even when the external inputs are correlated, but not identical, to the pattern, a necessary requirement for an associative memory model. In many of these models, the appropriate synaptic connectivity is assumed to be generated through a Hebbian learning process according to which synaptic efficacies are modified by the activity of pre- and post-synaptic neurons (Hebb, 1949).

What is an Ecosystem?

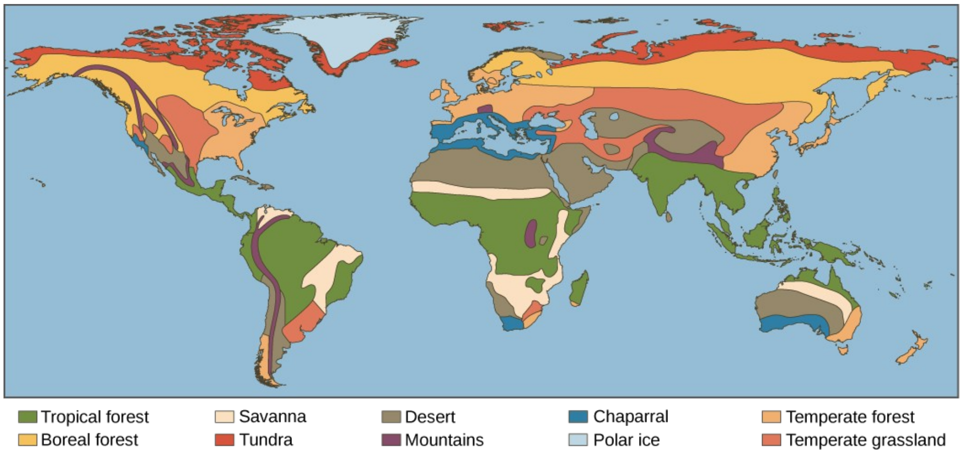
An ecosystem consists of a community of organisms together with their physical environment. Ecosystems can be of different sizes and can be marine, aquatic, or terrestrial. Broad categories of terrestrial ecosystems are called biomes. In ecosystems, both matter and energy are conserved. Energy flows through the system—usually from light to heat—while matter is recycled. Ecosystems with higher biodiversity tend to be more stable with greater resistance and resilience in the face of disturbances, disruptive events.

As a reminder, a community consists of all the populations of all the species that live together in a particular area. The concepts of ecosystem and community are closely related—the difference is that an ecosystem includes the physical environment, while a community does not. In other words, a community is the biotic, or living, component of an ecosystem. In addition to this biotic component, the ecosystem also includes an abiotic component—the physical environment. Ecosystems can be small, such as the tide pools found near the rocky shores of many oceans, or very large, such as the Amazon Rainforest in South America. It's basically up to the ecologist studying the ecosystem to define its boundaries in a way that makes sense for their questions of interest.

What are ecosystems like?

The short answer: incredibly diverse! Not only can ecosystems vary in size, but they can also differ in just about every imaginable biotic or abiotic feature.

Some ecosystems are marine, others freshwater, and others yet terrestrial—land based. Ocean ecosystems are most common on Earth, as oceans and the living organisms they contain cover 75% of the Earth's surface. Freshwater ecosystems are the rarest, covering only 1.8% of the Earth's surface. Terrestrial, land, ecosystems cover the remainder of Earth. Terrestrial ecosystems can be further grouped into broad categories called biomes, based largely on climate. Examples of terrestrial biomes include tropical rain forests, savannas, deserts, coniferous forests, deciduous forests, and tundra. The map below shows the broad distribution of biomes on Earth.



Energy and matter in ecosystems

Ecosystem ecologists are often most interested in tracing the movement of energy and matter through ecosystems.

We'll take a closer look at the movement of energy and matter when we consider food webs, networks of organisms that feed on one another, and biogeochemical cycles, the pathways taken by chemical elements as they move through the biosphere. The organisms found in an ecosystem tend to have adaptations, beneficial features arising by natural selection, that help them get energy and matter in the context of that particular ecosystem. Before we get into details, though, let's look at the key features of how energy and matter travel through ecosystems. Both energy and matter are conserved, neither created nor destroyed, but take different routes through ecosystems:

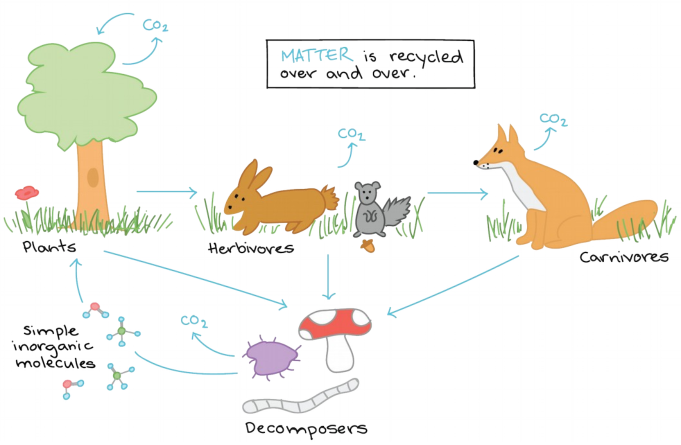
a) Matter is recycled; the same atoms are reused over and over.

b) Energy flows through the ecosystem, usually entering as light and exiting as heat.

Matter is recycled

Matter is recycled through Earth's ecosystems—though it may move from one ecosystem to another as it does when nutrients are washed away into a river1^11start superscript, 1, end superscript. The same atoms are used over and over again, assembled into different chemical forms and incorporated into the bodies of different organisms.

As an example, let's see how chemical nutrients move through a terrestrial ecosystem. A land plant takes in carbon dioxide from the atmosphere and other nutrients, such as nitrogen and phosphorous, from the soil to build the molecules that make up its cells. When an animal eats the plant, it uses the plant's molecules for energy and as building material for its own cells, often rearranging atoms and molecules into new forms. When plants and animals carry out cellular respiration—break down molecules as fuel—carbon dioxide is released into the atmosphere. Similarly, when they excrete waste or die, their chemical compounds are used for energy and building material by bacteria and fungi. These decomposers release simple molecules back into the soil and atmosphere, where they can be taken up anew in the next round of the cycle.

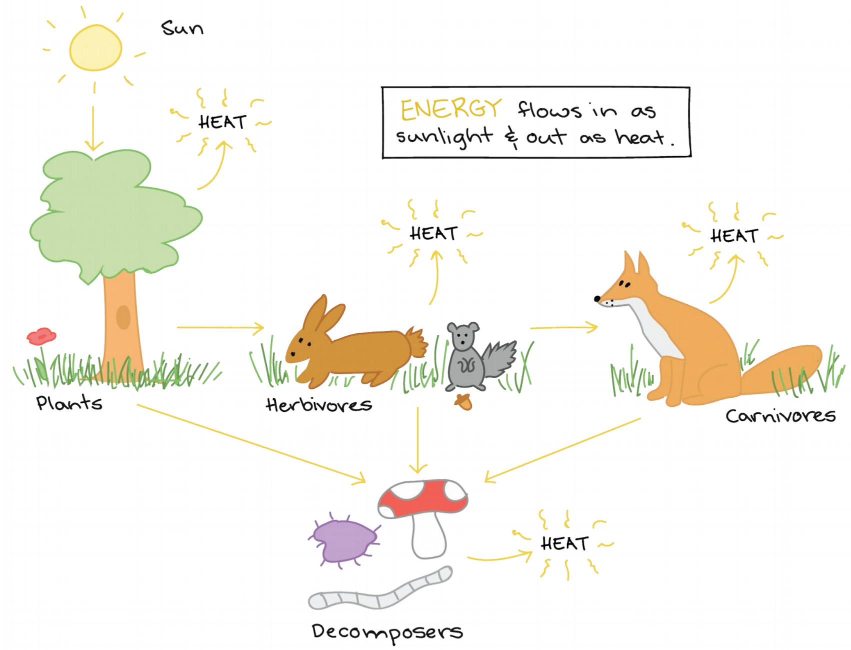


Energy flow is unidirectional, or one-way

Energy, unlike matter, cannot be recycled in ecosystems. Instead, energy flow through an ecosystem is a one-way street—generally, from light to heat.

Energy usually enters ecosystems as sunlight and is captured in chemical form by photosynthesizers like plants and algae. The energy is then passed through the ecosystem, changing forms as organisms metabolize, produce waste, eat one another, and eventually, die and decompose.

Each time energy changes forms, some of it is converted to heat. Heat still counts as energy—and thus no energy has been destroyed—but it generally can't be used as an energy source by living organisms. Ultimately, energy that entered the ecosystem as sunlight is dissipated as heat and radiated back into space.



This one-way flow of energy through ecosystems means that every ecosystem needs a constant supply of energy, usually from the sun, in order to function. Energy can be passed between organisms, but it cannot be recycled because some of it is lost as heat in each transfer.

Food chain and Food web

1. Producers, or autotrophs, make their own organic molecules. Consumers, or heterotrophs, get organic molecules by eating other organisms.

2. A food chain is a linear sequence of organisms through which nutrients and energy pass as one organism eats another.

3. In a food chain, each organism occupies a different trophic level, defined by how many energy transfers separate it from the basic input of the chain.

4. Food webs consist of many interconnected food chains and are more realistic representation of consumption relationships in ecosystems.

5. Energy transfer between trophic levels is inefficient—with a typical efficiency around 10%. This inefficiency limits the length of food chains.

Organisms of different species can interact in many ways. They can compete, or they can be symbionts—longterm partners with a close association. Or, of course, they can do what we so often see in nature programs: one of them can eat the other—chomp! That is, they can form one of the links in a food chain.

In ecology, a food chain is a series of organisms that eat one another so that energy and nutrients flow from one to the next. For example, if you had a hamburger for lunch, you might be part of a food chain that looks like this: grass → goat → human. But what if you had lettuce on your hamburger? In that case, you're also part of a food chain that looks like this: lettuce → human. As this example illustrates, we can't always fully describe what an organism—such as a human—eats with one linear pathway. For situations like the one above, we may want to use a food web that consists of many intersecting food chains and represents the different things an organism can eat and be eaten by.

Autotrophs vs. heterotrophs

What basic strategies do organisms use to get food? Some organisms, called autotrophs, also known as self-feeders, can make their own food—that is, their own organic compounds—out of simple molecules like carbon dioxide. There are two basic types of autotrophs:

1. Photoautotrophs, such as plants, use energy from sunlight to make organic compounds—sugars—out of carbon dioxide in photosynthesis. Other examples of photoautotrophs include algae and cyanobacteria.

2. Chemoautotrophs use energy from chemicals to build organic compounds out of carbon dioxide or similar molecules. This is called chemosynthesis. For instance, there are hydrogen sulfide-oxidizing chemoautotrophic bacteria found in undersea vent communities where no light can reach.

Autotrophs are the foundation of every ecosystem on the planet. That may sound dramatic, but it's no exaggeration! Autotrophs form the base of food chains and food webs, and the energy they capture from light or chemicals sustains all the other organisms in the community. When we're talking about their role in food chains, we can call autotrophs producers.

Heterotrophs, also known as other-feeders, can't capture light or chemical energy to make their own food out of carbon dioxide. Humans are heterotrophs. Instead, heterotrophs get organic molecules by eating other organisms or their byproducts. Animals, fungi, and many bacteria are heterotrophs. When we talk about heterotrophs' role in food chains, we can call them consumers. As we'll see shortly, there are many different kinds of consumers with different ecological roles, from plant-eating insects to meat-eating animals to fungi that feed on debris and wastes.

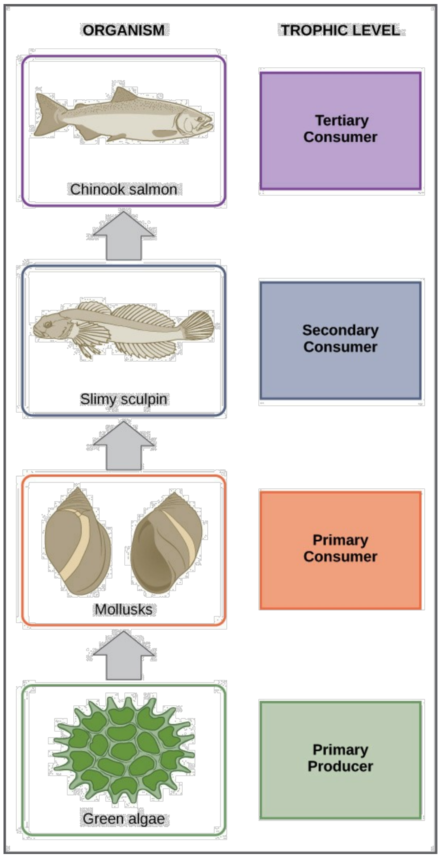
Food chains

Now, we can take a look at how energy and nutrients move through a ecological community. Let's start by considering just a few who-eats-who relationships by looking at a food chain.

A food chain is a linear sequence of organisms through which nutrients and energy pass as one organism eats another. Let's look at the parts of a typical food chain, starting from the bottom—the producers—and moving upward. At the base of the food chain lie the primary producers. The primary producers are autotrophs and are most often photosynthetic organisms such as plants, algae, or cyanobacteria. The organisms that eat the primary producers are called primary consumers. Primary consumers are usually herbivores, plant-eaters, though they may be algae eaters or bacteria eaters.

The organisms that eat the primary consumers are called secondary consumers. Secondary consumers are generally meat-eaters—carnivores. The organisms that eat the secondary consumers are called tertiary consumers. These are carnivore-eating carnivores, like eagles or big fish. Some food chains have additional levels, such as quaternary consumers—carnivores that eat tertiary consumers. Organisms at the very top of a food chain are called apex consumers.

We can see examples of these levels in the diagram below. The green algae are primary producers that get eaten by mollusks—the primary consumers. The mollusks then become lunch for the slimy sculpin fish, a secondary consumer, which is itself eaten by a larger fish, the Chinook salmon—a tertiary consumer.



Each of the categories above is called a trophic level, and it reflects how many transfers of energy and nutrients—how many consumption steps—separate an organism from the food chain's original energy source, such as light. As we'll explore further below, assigning organisms to trophic levels isn't always clear-cut. For instance, humans are omnivores that can eat both plants and animals.

Decomposers

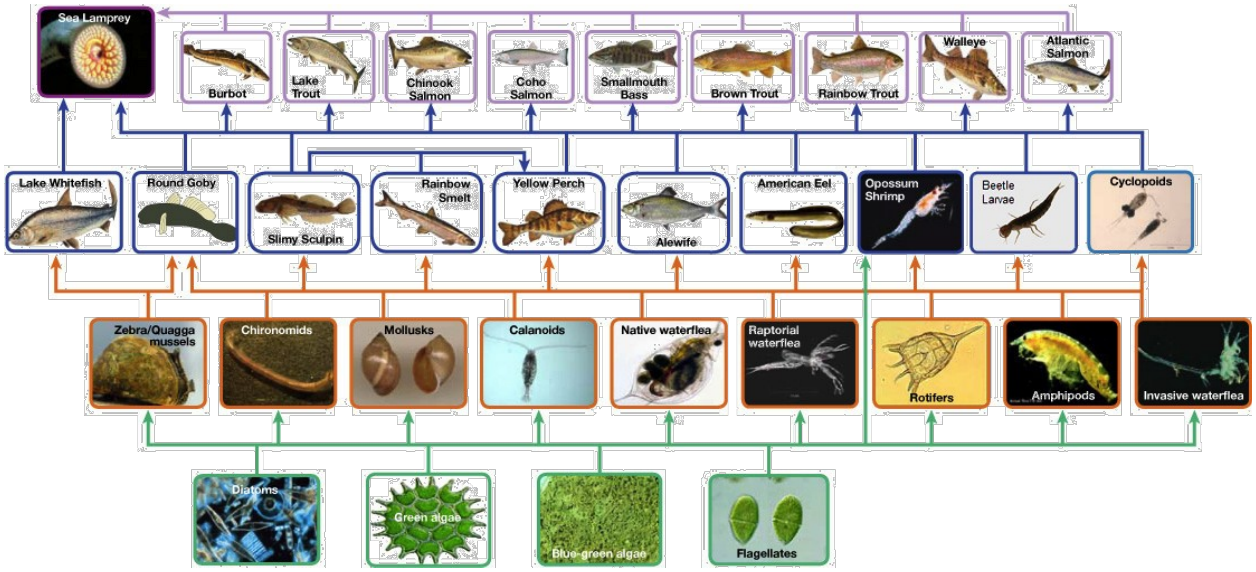
One other group of consumers deserves mention, although it does not always appear in drawings of food chains. This group consists of decomposers, organisms that break down dead organic material and wastes.

Decomposers are sometimes considered their own trophic level. As a group, they eat dead matter and waste products that come from organisms at various other trophic levels; for instance, they would happily consume decaying plant matter, the body of a half-eaten squirrel, or the remains of a deceased eagle. In a sense, the decomposer level runs parallel to the standard hierarchy of primary, secondary, and tertiary consumers.

Fungi and bacteria are the key decomposers in many ecosystems; they use the chemical energy in dead matter and wastes to fuel their metabolic processes. Other decomposers are detritivores—detritus eaters or debris eaters. These are usually multicellular animals such as earthworms, crabs, slugs, or vultures. They not only feed on dead organic matter but often fragment it as well, making it more available for bacterial or fungal decomposers.

Decomposers as a group play a critical role in keeping ecosystems healthy. When they break down dead material and wastes, they release nutrients that can be recycled and used as building blocks by primary producers.

Food webs



Food chains give us a clear-cut picture of who eats whom. However, some problems come up when we try and use them to describe whole ecological communities. For instance, an organism can sometimes eat multiple types of prey or be eaten by multiple predators, including ones at different trophic levels. This is what happens when you eat a hamburger patty! The cow is a primary consumer, and the lettuce leaf on the patty is a primary producer. To represent these relationships more accurately, we can use a food web, a graph that shows all the trophic—eating-related—interactions between various species in an ecosystem. The diagram below shows an example of a food web from Lake Ontario. Primary producers are marked in green, primary consumers in orange, secondary consumers in blue, and tertiary consumers in purple.

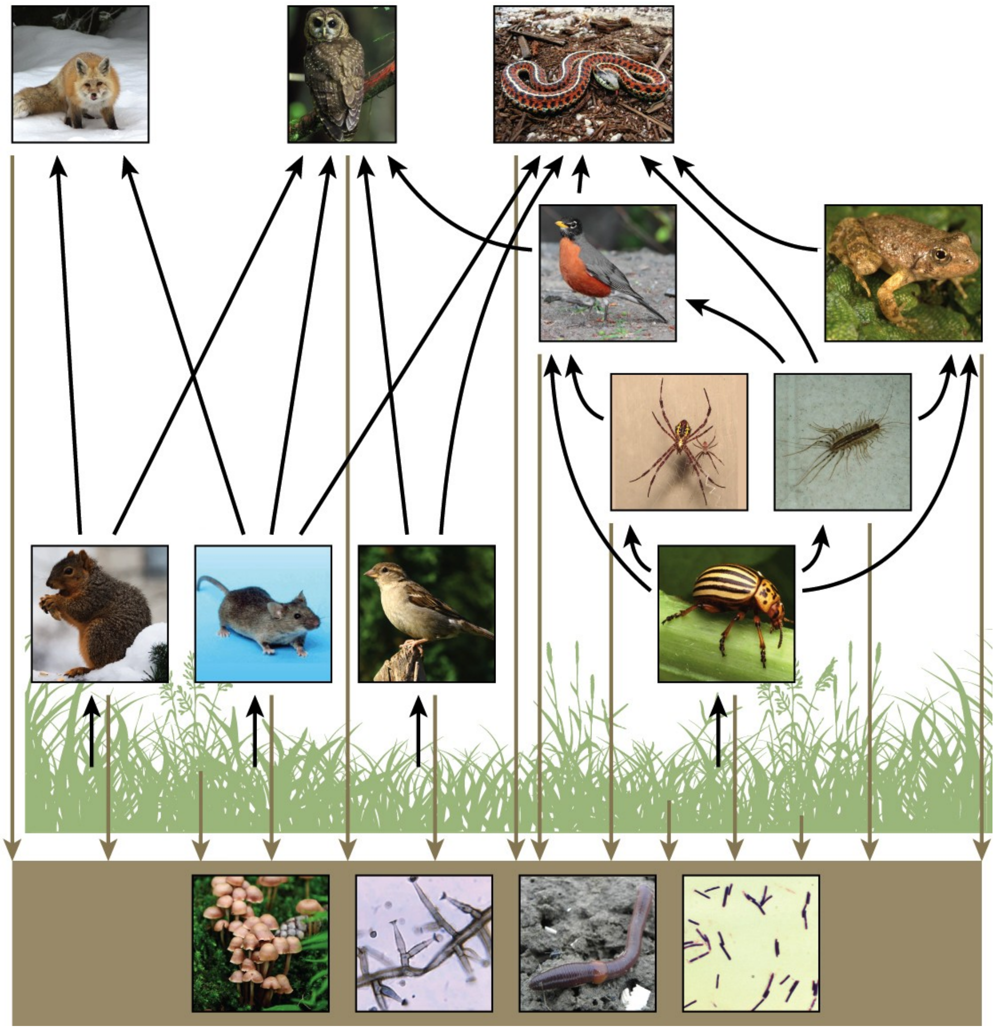
In food webs, arrows point from an organism that is eaten to the organism that eats it. As the food web above shows, some species can eat organisms from more than one trophic level. For example, opossum shrimp eat both primary producers and primary consumers.

Bonus question: This food web contains the food chain we saw earlier in the article—green algae → mollusks → slimy sculpin → salmon. Can you find it?

Grazing vs. detrital food webs

Food webs don't usually show decomposers—you might have noticed that the Lake Ontario food web above does not. Yet, all ecosystems need ways to recycle dead material and wastes. That means decomposers are indeed present, even if they don't get much air time.

For example, in the meadow ecosystem shown below, there is a grazing food web of plants and animals that provides inputs for a detrital food web of bacteria, fungi, and detritovores. The detrital web is shown in simplified form in the brown band across the bottom of the diagram. In reality, it would consist of various species linked by specific feeding interactions—that is, connected by arrows, as in the grazing food web aboveground. Detrital food webs can contribute energy to grazing food webs, as when a robin eats an earthworm.



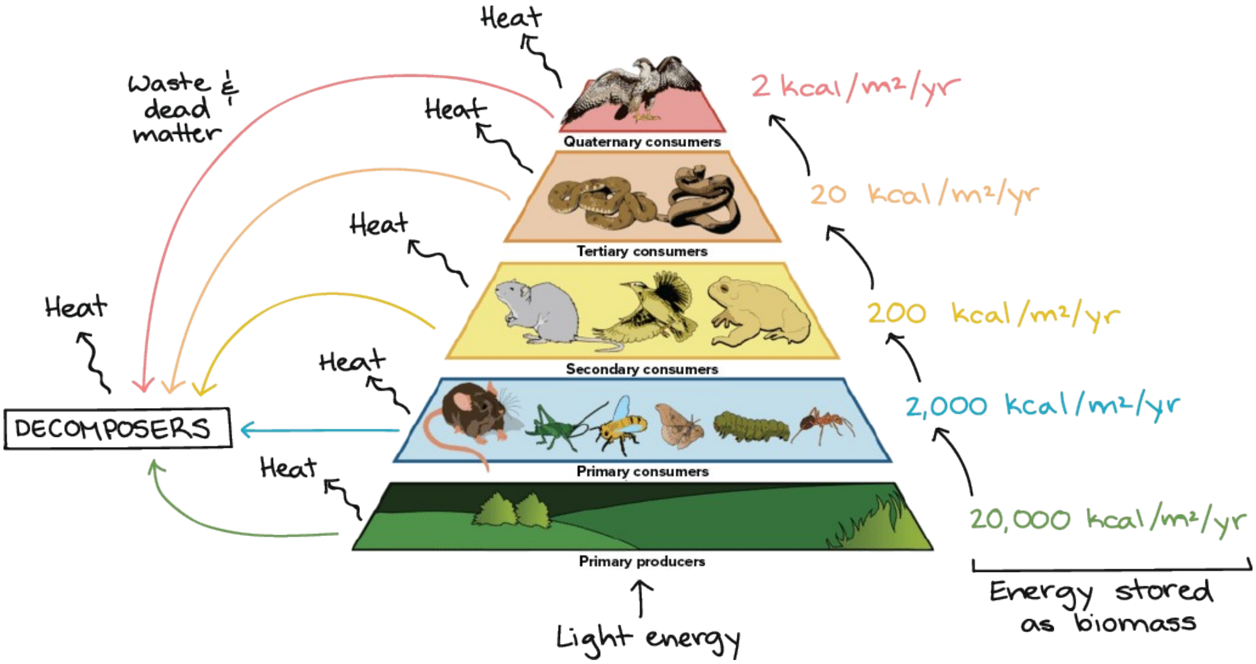
Energy transfer efficiency limits food chain lengths

Energy is transferred between trophic levels when one organism eats another and gets the energy-rich molecules from its prey's body. However, these transfers are inefficient, and this inefficiency limits the length of food chains.

When energy enters a trophic level, some of it is stored as biomass, as part of organisms' bodies. This is the energy that's available to the next trophic level since only energy storied as biomass can get eaten. As a rule of thumb, only about 10% of the energy that's stored as biomass in one trophic level—per unit time—ends up stored as biomass in the next trophic level—per the same unit time. This 10% rule of energy transfer is a good thing to commit to memory.

As an example, let's suppose the primary producers of an ecosystem store 20,000 kcal/m^2/year of energy as biomass. This is also the amount of energy per year that's made available to the primary consumers, which eat the primary producers. The 10% rule would predict that the primary consumers store only 2,000 kcal/m^2/year of energy in their own bodies, making energy available to their predators—secondary consumers—at a lower rate.

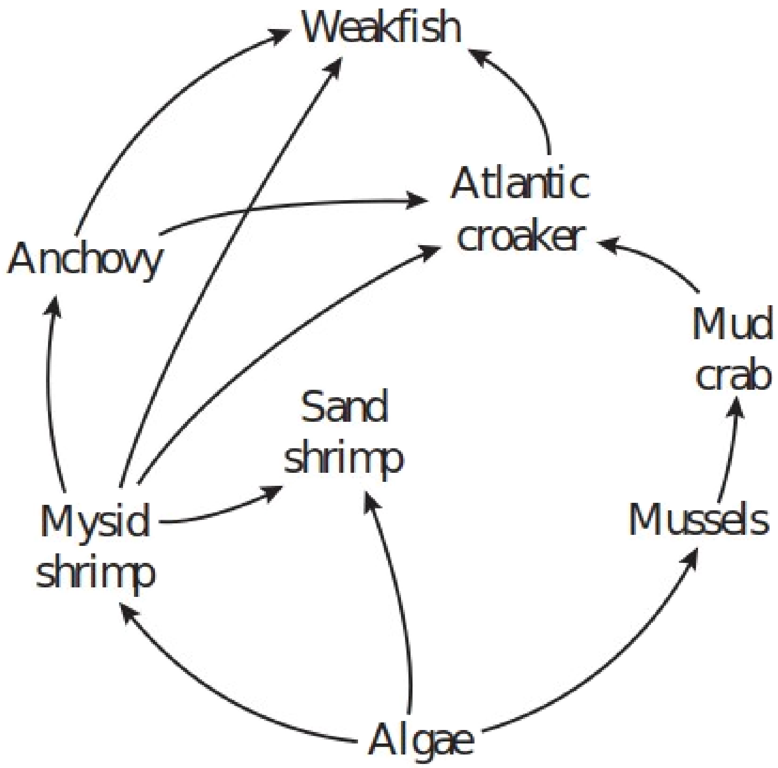
This pattern of fractional transfer limits the length of food chains; after a certain number of trophic levels—generally three to six, there is too little energy flow to support a population at a higher level.



Why does so much energy exit the food web between one trophic level and the next? Here are a few of the main reasons for inefficient energy transfer:

In each trophic level, a significant amount of energy is dissipated as heat as organisms carry out cellular respiration and go about their daily lives. Some of the organic molecules an organism eats cannot be digested and leave the body as feces, poop, rather than being used. Not all of the individual organisms in a trophic level will get eaten by organisms in the next level up. Some instead die without being eaten.

The feces and uneaten, dead organisms become food for decomposers, who metabolize them and convert their energy to heat through cellular respiration. So, none of the energy actually disappears—it all winds up as heat in the end.



Which of the organisms in this food web gets energy from both producers and consumers?

Choose 1 answer:

A. Algae

B. Mysid shrimp

C. Sand shrimp

D. Anchovy

Ecological Feedback Loop

Ecosystem feedback is the effect that change in one part of an ecosystem has on another and how this effect then feeds back to effect the source of the change inducing more or less of it. These feedback loops form the basic dynamics for regulating the state of the ecosystem. A negative feedback loop is where the state of one element affects the other in the opposite direction, with the net result of this being a stable system where different forces are counterbalancing each other out creating some equilibrium. As such negative feedback can be identified as providing stability. All ecosystems are composed of many negative feedback loops that keep every part of the system within the boundaries necessary for the whole system to continue functioning. Population regulation is a classical example of negative feedback. Because the resources that sustain populations are limited, no population can exceed the carrying capacity of the ecosystem for long. Negative feedback loops between predators and prey work to keep plant and animal populations within the limits of the carrying capacity of their environment and thus maintain some form of stability.

Positive Feedback Loops

Positive feedback stimulates change and it is responsible for the sudden appearance of rapid changes within ecosystems. Positive feedback is a circular link of effects that are self-reinforcing. When part of the system increases, another part of the system also changes in a way that makes the first part increase even more. Positive feedback is a source of instability and strong force of change as it can drive the system outside of its normal operating parameters. As an example we could cite exponential population growth, when there is a surplus of resources, or lack of predators, this allows a plant or animal population to grow without limit. More population leads to more births, and more births lead to an increasing population creating a compounding effect over time.

Positive & Negative Feedback

Ecosystems and complex systems, in general, have a tension between forces that resist change, the negative feedback, and forces that promote change, the positive feedback. Negative feedback may dominate at some times and positive feedback may dominate at other times, depending on the situation. As a result, ecosystems may stay more or less the same for long periods, but they can also change very suddenly. This change can be like a rapid switch from one state to another, this flipping is known to be a characteristic of nonlinear systems and complex systems in general.

Examples

As an example of these two counteracting forces we can look at the succession of an ecosystem from grass to shrub community, beginning with an ecosystem in which the ground is covered with grasses. Shrubs may be present, but they are young and scattered. The ecosystem may stay this way for five to ten years, or possibly longer, because shrub seedlings grow very slowly. They grow slowly because grass roots are located in the topsoil, while most of the shrub roots are lower down. Grasses intercept most of the rainwater before it reaches the roots of the shrubs. Because the grasses limit the supply of water to the shrub seedlings, they maintain the integrity of the ecosystem as a grass ecosystem. At this stage, negative feedback is acting to keep the biological community the same. However, after a number of years, some of the trees and shrubs, which have been growing slowly, are finally tall enough to shade the grasses below them. The grasses then have less sunlight for photosynthesis, and their growth is restricted. This results in more water for the shrubs, which grow faster and shade the grasses even more. This process of positive feedback allows the shrubs to take over in a relatively short period of time. They now dominate the available sunlight and water, and the grasses decrease dramatically.

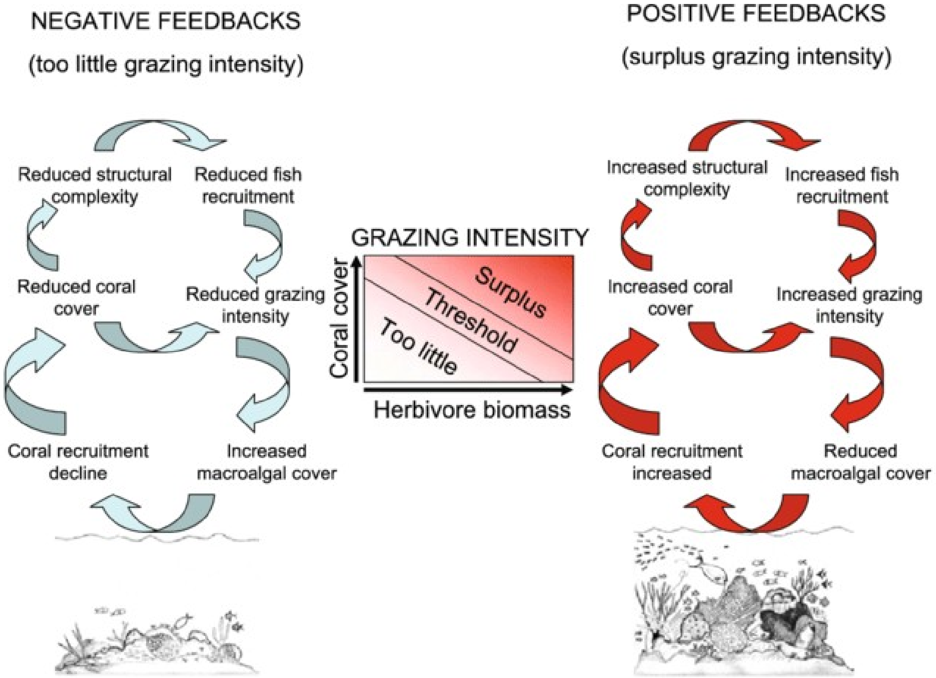
Vicious Cycles

The term vicious circle refers to a complex chain of events which reinforce themselves through a feedback loop. If the outcome is a negative result this would be termed a vicious cycle. The melting of the polar ice caps is an example of a vicious cycle, as the reflective ice caps melt they reflect less sunlight and heat back to the atmosphere, with more of this heat being trapped by the dark ocean which is now exposed by the loss of ice cap. This retained heat then increases the temperature feeding back to induce the melting of more ice caps creating what we would call a vicious circle.

As another example, we might cite the pollution of the lagoons that surround small South Pacific islands. Many South Pacific communities now consume imported packaged and canned foods, disposing of the empty cans and other waste in dumps. Rainwater runoff from the dumps pollutes the lagoons, reducing the quantity of fish and other seafood. With less seafood, people are forced to buy more and more cheap canned food, the pollution becomes worse and the lagoon has fewer fish. This positive feedback loop changes the lagoon ecosystem while also degrading the people’s diet again creating a vicious circle.

Virtuous Cycles

The term virtuous circle refers to the opposite phenomena, a chain of events which reinforce themselves through a positive feedback loop creating some favorable outcome. As an example of a virtuous ecological cycle we might cite the Philippine’s fishery after World War II. With the introduction of destructive fishing methods such as dynamite, cyanide, and small-mesh fishing nets a number of interlocking and mutually reinforcing vicious cycles were set in motion to significantly degrade the state of the marine ecosystem surrounding Apo island up to this point. The positive tipping point for Apo Island was the creation of a marine sanctuary, setting in motion a cascade of changes that reversed the vicious cycle with additional virtuous cycles arising in association with the marine sanctuary. The sanctuary served as a nursery, contributing directly to the recovery of fish stocks in the island’s fishing grounds. Success with the sanctuary stimulated the fishermen to set up sustainable management for the fishing grounds. A virtuous cycle of increasing fish stocks, accompanied by growing management experience, pride, and commitment to the sanctuary, was set in motion. As fishing improved around the island, fishermen were no longer compelled to travel far away for their work. Fishing right at home, where they had to live with the consequences of their fishing practices, reinforced their motivation for sustainable fishing, this compounding of positive outcomes is a virtuous cycle.

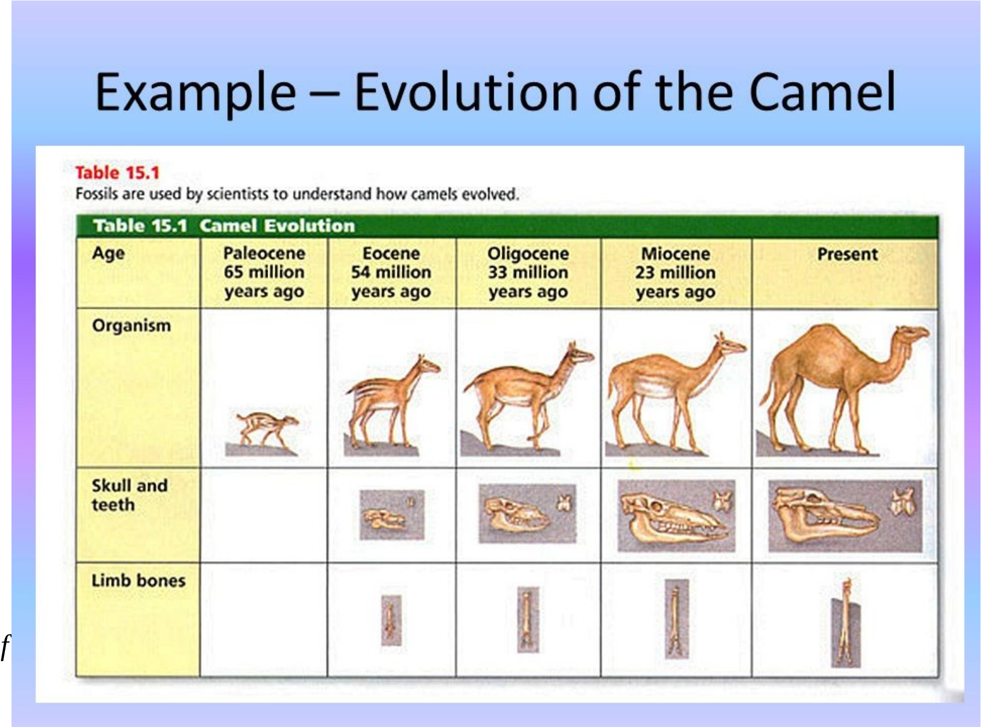


Principles of ecosystem sustainability.pdf

Evolution: A short story

What is Evolution?

Evolution is change in inherited characteristics of a species over time. By the word “species”, we mean a group of organisms that share similar characteristics and can reproduce among themselves.



Example – Evolution of the Camel

Table 15.1
Fossils are used by scientists to understand how camels evolved.

Table 15.1 Camel Evolution

Age	Paleocene 65 million years ago	Eocene 54 million years ago	Oligocene 33 million years ago	Miocene 23 million years ago	Present
Organism					
Skull and teeth					
Limb bones					

Early models of Evolution

1. Theory of Acquired Characteristics (Lamark's Model)

Jean Baptiste de Lamark proposed hypothesis suggesting that traits acquired by the parents during their lifetime were inherited by their offspring.

Ex.- if parents were muscular, the children would be too.

Evidence does not support this theory.

Biologists define an **acquired characteristic** as one that has developed in the course of the life of an individual in the somatic or body cells, usually as a direct response to some external change in the environment or through the use or disuse of a part. The inheritance of such a characteristic means its reappearance in one or more individuals in the next or in succeeding generations. The blacksmith's arm (or any other set of muscles) enlarges when used continually against an external resistance, such as the weight of the hammer. If the effect were inherited, the smith's children at birth would have unusually large arms—if not at birth, then when they became adults, even though they had not used their arms excessively. There is no evidence supporting this case. A more subtle illustration is found in the supposed inheritance of an increased dexterity of the hands of a musician through practice. The skill acquired, although causing no visible increase in the size of the fingers, might be imagined to be passed along to the musician's children, and they might then be expected to play skillfully with minimal practice. Just how the intricate interplay of cerebral sequences that has given the dexterity to the musician's fingers could ever be transferred to the musician's sex cells (spermatozoa or ova), and through them to any potential children, has never been brought within the range of biological possibilities.

2. Darwin's Model of Evolution (1859)

In 1831 Charles Darwin set sail on the HMS Beagle to explore the South American coast.

1. He recorded observations of the different plant and animal species he saw.

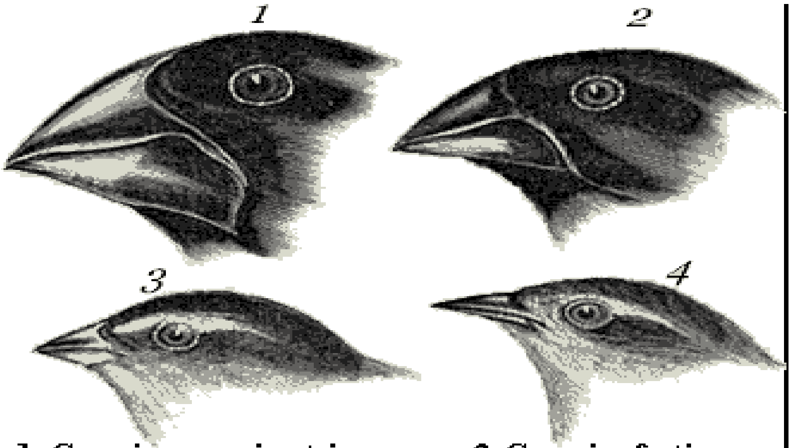
2. amazed by the variety of life he saw on the Galapagos Islands he hypothesized they must have come from Central and South America. He also saw forms of life never seen anywhere else.

3. He also observed 13 different species of finches.

a. they were similar except for body size, beak shape, and eating habits

b. they looked a lot like finches he had seen in South America

c. he hypothesized they must have all evolved from that S American species. Those that had the traits necessary to survive did and those that didn’t died out.



1. Geospiza magnirostris. **2. Geospiza fortis.**
3. Geospiza parvula. **4. Certhidea olivacea.**

FINCHES FROM GALAPAGOS ARCHIPELAGO.

Darwin’s Theory of Natural Selection

Darwinism is a theory of biological evolution developed by the English naturalist Charles Darwin (1809–1882) and others, stating that all species of organisms arise and develop through the natural selection of small, inherited variations that increase the individual's ability to compete, survive, and reproduce.

These are the basic tenets of evolution by natural selection as defined by Darwin:

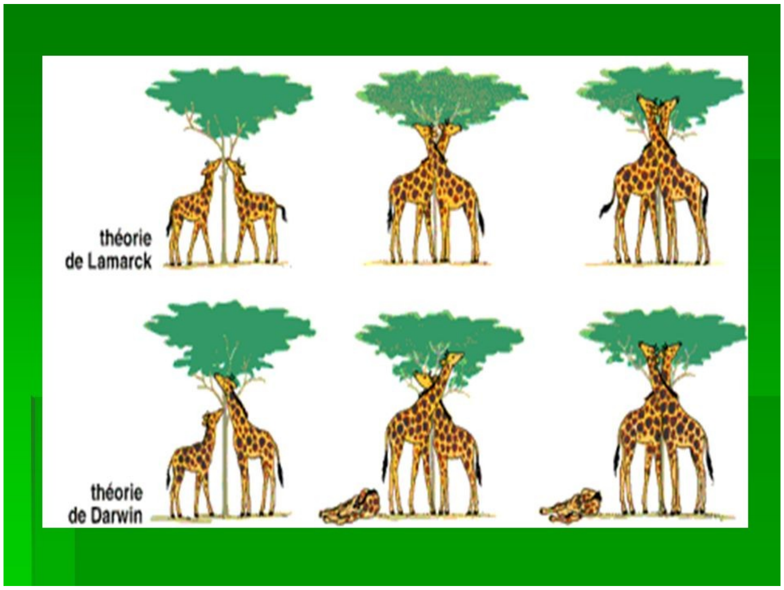
1. More individuals are produced each generation than can survive.
2. Phenotypic variation exists among individuals and the variation is heritable.

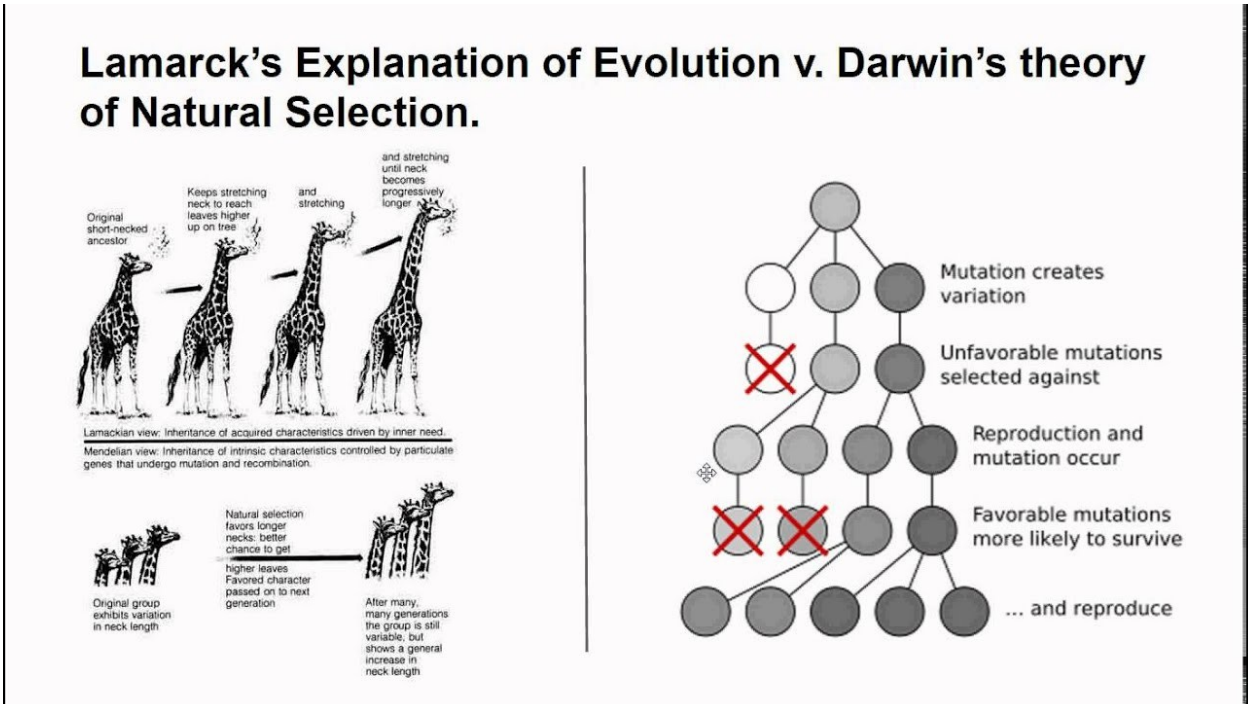
3. The individuals compete for resources (Struggle for Existence). Those individuals with heritable traits better suited to the environment will survive (Survival of the fittest).

4. When reproductive isolation occurs new species will form.

5. The bottomline is: Species alive today are descended with modifications from common ancestors.

Lamark vs Darwin





<u>3. Modern Synthesis Model (1942):</u>

The Modern Synthetic theory of Evolution explains the evolution of life in terms of genetic changes occurring in the population that leads to the formation of new species. It also explains the genetic population or Mendelian population, gene pool and the gene frequency. The concepts coming under this synthetic theory of evolution include the genetic variations, reproductive and geographical isolation and the natural selection.

The Modern Synthetic populations of Evolution describes the merging of the Darwinian evolution with the Mendelian genetics, resulting in a unified theory of the evolution. This theory is also referred to as the Neo-Darwinian theory. Synthetic theory of Evolution was introduced to us by few legendary evolutionary biologists naming T. Dobzhansky, J.B.S. Haldane, R.A. Fisher, Sewall Wright, G.L. Stebbins, Ernst Mayr in the years 1930 and 1940. The Modern Synthetic theory of Evolution showed a number of changes as to how the evolution and the process of evolution are conceived. The theory gave a new definition about the evolution as “the changes occurring in the allele frequencies within the populations, ” which emphasizes on the genetics of evolution. The modern synthetic theory includes scientific evidence from genetics. It explains the concepts which occur when the allele frequency of the population changes. According to this theory, when the changes are great enough, there is a formation of new species. A species is a group of individuals who are capable of interbreeding and producing a fertile offspring.

Factors of Modern Synthetic Theory of Evolution

There are some factors describing the modern theory of synthetic evolution which are as explained below- (In addition to these reactions, the other factors affecting the working of the process are the migration of the individuals from one form of the population to other, hybridization between the races of species increases the genetic variability of the population.

a. Recombination or Variation

Recombination of the new genotypes from the existing genes. The gene combinations having same individuals with two kinds of alleles, mixing of the chromosomes during sexual reproduction of two parents produce new individuals, an exchange of the chromosomal pairs of alleles during the meiosis which is called crossing overproduce the new form of gene combinations. Chromosomal mutations like deletion, inversion, duplication, translocation, polyploidy result in the recombination.

b. Mutation

The changes that occur in the gene due to phenotypic effect differential as the mutation. This produces a variety of changes that may be harmful. Many of the mutant forms of genes are recessive to the normal genes in a homozygous condition. These mutations cause variations in offsprings.

c. Heredity

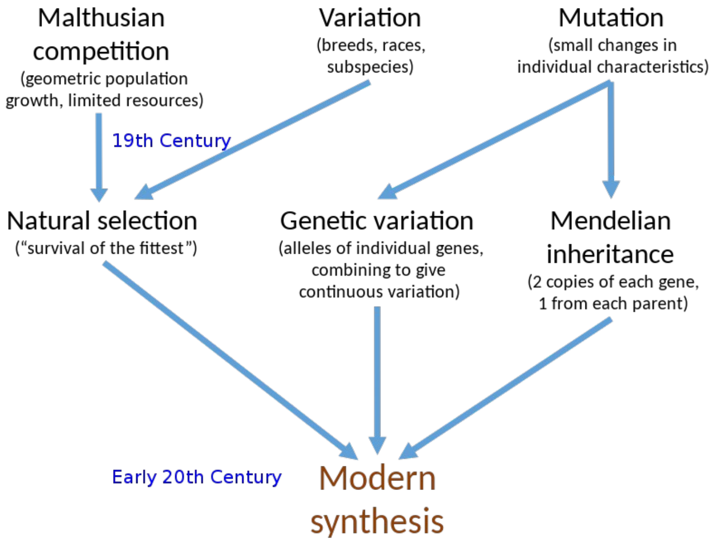
The transmission occurring in the variations from the parents to their offsprings is a primary mechanism in the evolution. The organisms which possess hereditary properties are favoured in the struggle for the existence. By this, the offsprings benefit from the characteristics of parents.

d. Natural selection

Natural selection produces a change in the frequency of the genes from one generation to the other favoring the differential form of the reproduction. The natural selection process creates an adaptive relation between the environment and the population through various combinations of genes.

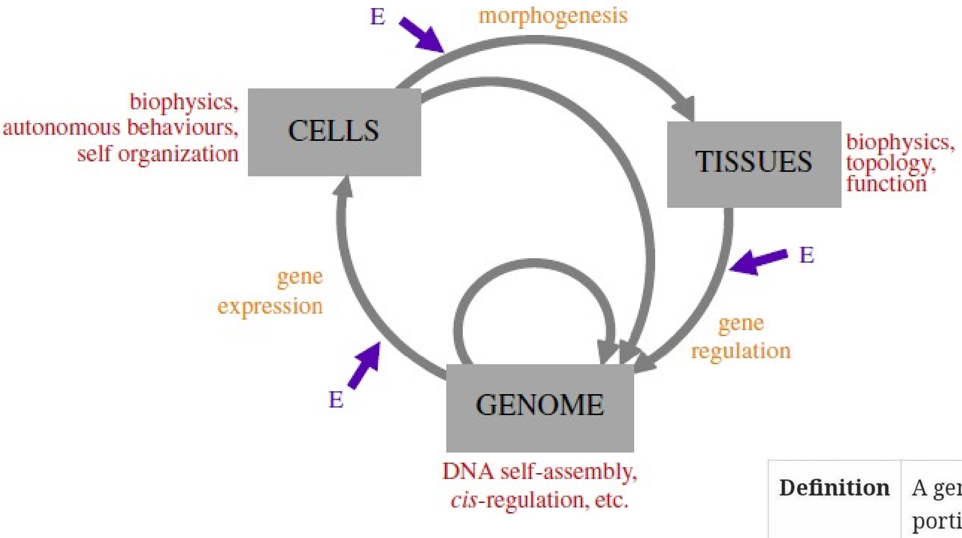
e. Isolation

It is one among the significant factors responsible for the synthetic theory of evolution. The isolation helps in preventing the interbreeding of related organisms which is a reproductive form of isolation.



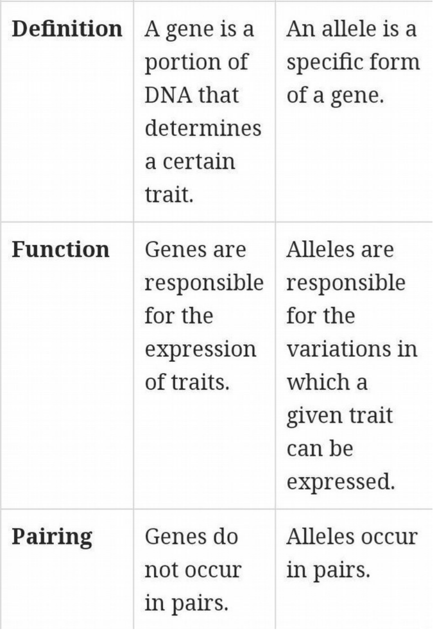
4. Integrated Synthesis Model (2012)

Evolutionary developmental biology (evo–devo) that forms the heart of Integrated Synthesis Model is that part of biology concerned with how changes in embryonic development during single generations relate to the evolutionary changes that occur between generations. Charles Darwin argued for the importance of development (embryology) in understanding evolution. After the discovery in 1900 of Mendel’s research on genetics, however, any relationship between development and evolution was either regarded as unimportant for understanding the process(es) of evolution or as a black box into which it was hard to see. A suite of new concepts emerged from evo-devo or Evolutionary Developmental Biology, a field of research that arose in the early 1980s from a discontent with the exclusion of developmental biology from evolutionary theory. The subsequent rise of new molecular methodologies for a comparative analysis of gene regulation resulted in a huge increase of our understanding of how the processes of development evolve. In its theoretical domain, the evo-devo approach starts from the premise that the genotype–phenotype relation is not merely a statistical correlation, but that the rules of developmental processes govern phenotypic outcomes while relying on additional inputs not coming from the genome. It is abundantly clear that development is not a linear reading out of a code or program but a systemic process of feedback interactions between genetic and non-genetic templates, cells and tissues that mobilizes physical and autonomous properties at different scales and depends on local as well as global environments.



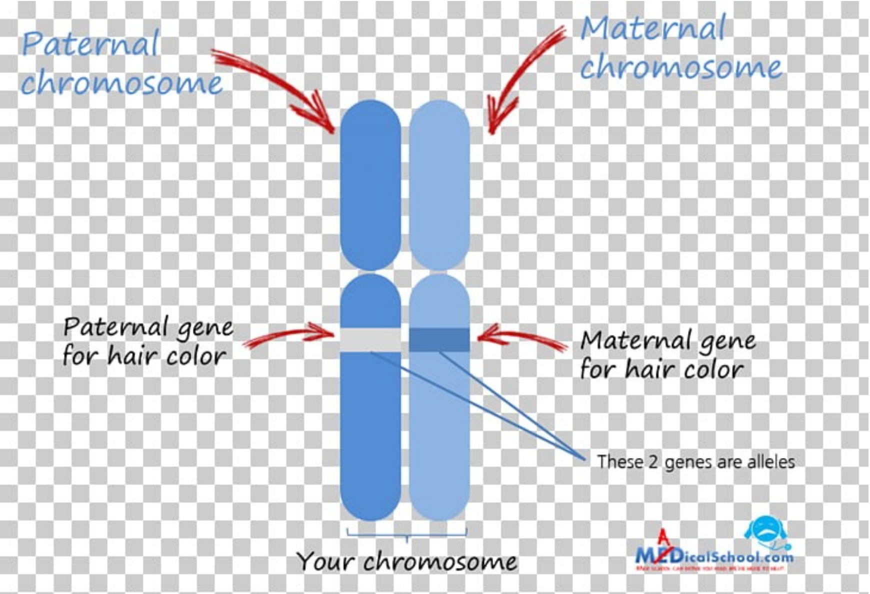
[At this moment, let us take a quick look at the genotype-phenotype map:

What is the definition of a genotype?



Definition	A gene is a portion of DNA that determines a certain trait.	An allele is a specific form of a gene.
Function	Genes are responsible for the expression of traits.	Alleles are responsible for the variations in which a given trait can be expressed.
Pairing	Genes do not occur in pairs.	Alleles occur in pairs.

In biology, a gene is a section of DNA that encodes a trait (like eye color). The precise arrangement of nucleotides (each composed of a phosphate group, sugar and a base) in a gene can differ between copies of the same gene. Therefore, a gene can exist in different forms across organisms. These different forms are known as alleles (like: what is the eye color? Black or blue?). The exact fixed position on the chromosome that contains a particular gene is known as a locus. An allele is therefore one of the possible forms of a gene. Most genes have two alleles, a dominant allele and a recessive allele. If an organism is heterozygous for that trait, or possesses one of each allele, then the dominant trait is expressed. A recessive allele is only expressed if an organism is homozygous for that trait, or possesses two recessive alleles. Alleles were first defined by Gregor Mendel in the law of segregation.



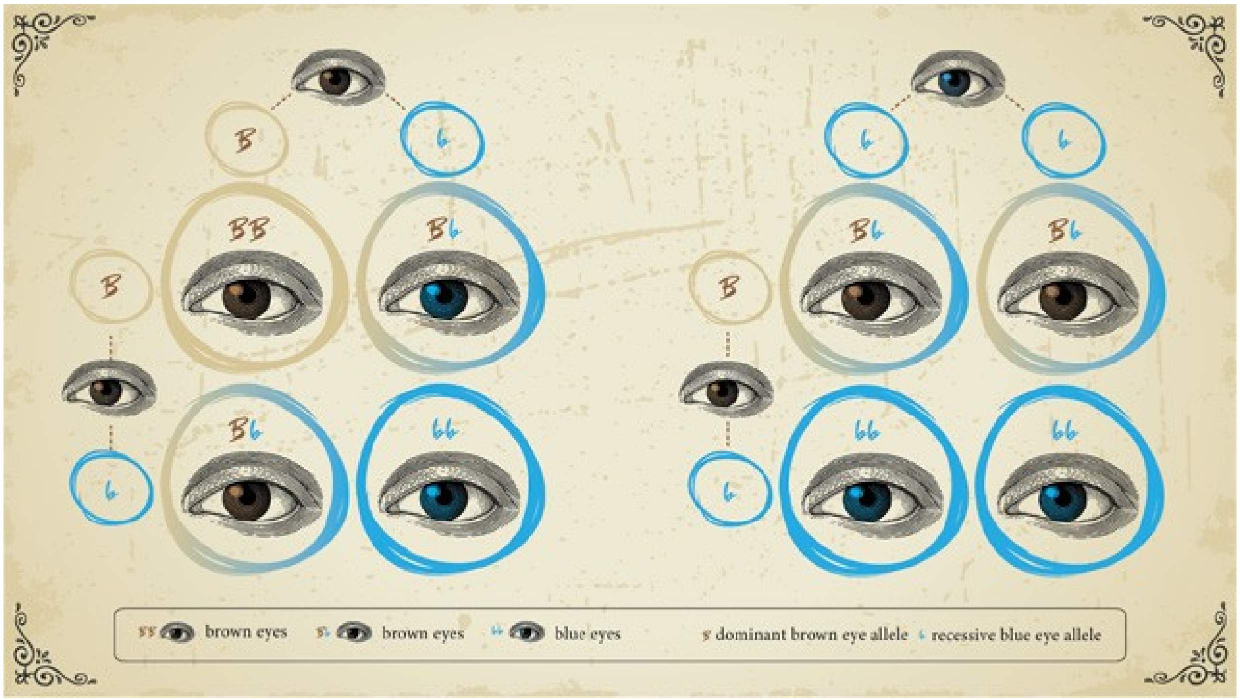
A diploid organism either inherits two copies of the same allele or one copy of two different alleles from their parents. If an individual inherits two identical alleles, their genotype is said to be homozygous at that locus. However, if they possess two different alleles, their genotype is classed as heterozygous for that locus. Alleles of the same gene are either autosomal dominant or recessive. An autosomal dominant allele will always be preferentially expressed over a recessive allele. The subsequent combination of alleles that an individual possesses for a specific gene is their genotype.

Genotype examples

Let’s look at a classic example – eye color.

A gene encodes eye color.

In this example, the allele is either brown, or blue, with one inherited from the mother, and the other inherited from the father. The brown allele is dominant (B), and the blue allele is recessive (b). If the child inherits two different alleles (heterozygous) then they will have brown eyes. For the child to have blue eyes, they must be homozygous for the blue eye allele.



Other examples of genotype include:

Hair color

Height

Shoe size

What is the definition of a phenotype?

The sum of an organism’s observable characteristics is their phenotype. A key difference between phenotype and genotype is that, whilst genotype is inherited from an organism’s parents, the phenotype is not. Whilst a phenotype is influenced the genotype, genotype does not equal phenotype. The phenotype is influenced by the genotype and factors including:

Epigenetic modifications

Environmental and lifestyle factors

Phenotype examples

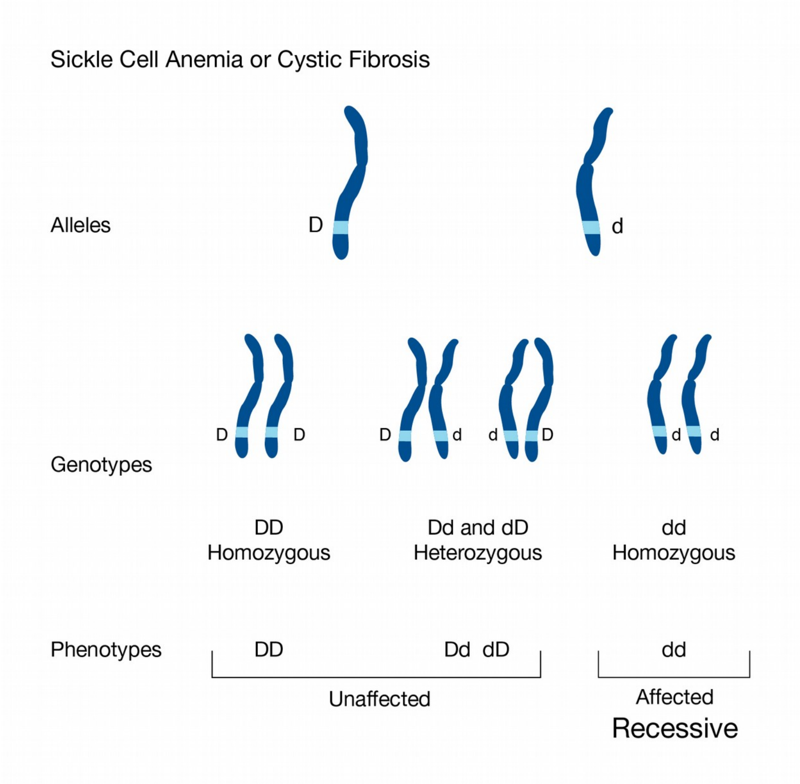
Environmental factors that may influence the phenotype include nutrition, temperature, humidity and stress. Flamingos are a classic example of how the environment influences the phenotype. Whilst renowned for being vibrantly pink, their natural color is white – the pink color is caused by pigments in the organisms in their diet.

A second example is an individual's skin color. Our genes control the amount and type of melanin that we produce, however, exposure to UV light in sunny climates causes the darkening of existing melanin and encourages increased melanogenesis and thus darker skin.

Genotype vs phenotype: observing

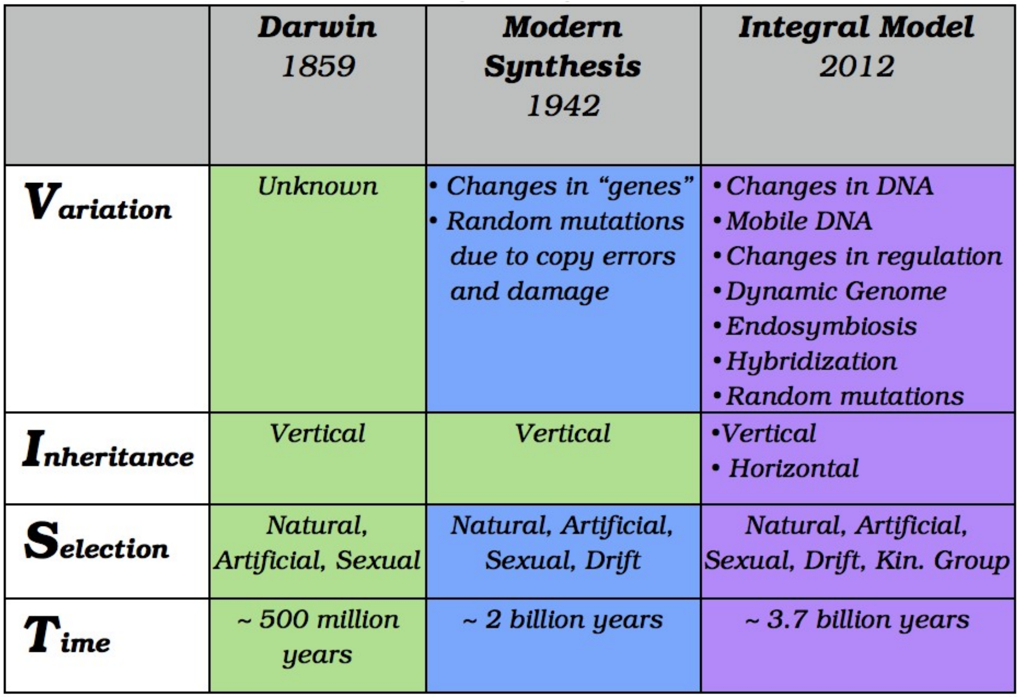
Observing the phenotype is simple – we take a look at an organism’s outward features and characteristics, and form conclusions about them. Observing the genotype, however, is a little more complex.

Genotyping is the process by which differences in the genotype of an individual are analyzed using biological assays. The data obtained can then be compared against either a second individual’s sequence, or a database of sequences.]

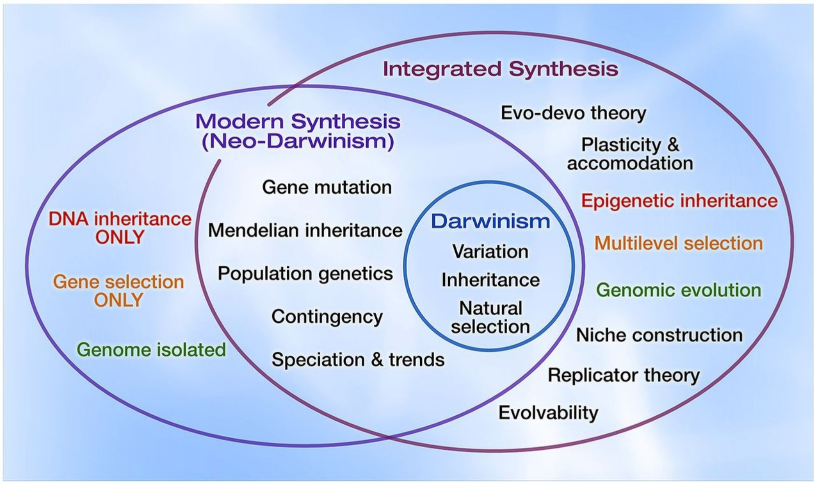


Cystic fibrosis is a hereditary disease that affects the lungs and digestive system. The body produces thick and sticky mucus that can clog the lungs and obstruct the pancreas. Cystic fibrosis (CF) can be life-threatening, and people with the condition tend to have a shorter-than-normal life span. In this figure, cystic fibrosis allele is carried by “d”.

Thus, different models of evolution compare with each other as follows:



	Darwin *1859*	***Modern Synthesis*** *1942*	***Integral Model*** *2012*
Variation	*Unknown*	*• Changes in “genes”* *• Random mutations due to copy errors and damage*	*• Changes in DNA* *• Mobile DNA* *• Changes in regulation* *• Dynamic Genome* *• Endosymbiosis* *• Hybridization* *• Random mutations*
Inheritance	*Vertical*	*Vertical*	*• Vertical* *• Horizontal*
Selection	*Natural, Artificial, Sexual*	*Natural, Artificial, Sexual, Drift*	*Natural, Artificial, Sexual, Drift, Kin. Group*
Time	*~ 500 million years*	*~ 2 billion years*	*~ 3.7 billion years*



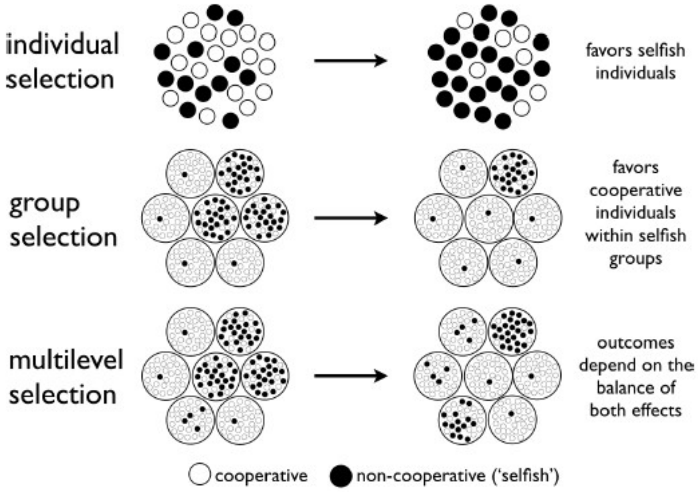
Basic tenets of Integrated Synthesis Model:

1. Evo-devo theory: discussed earlier

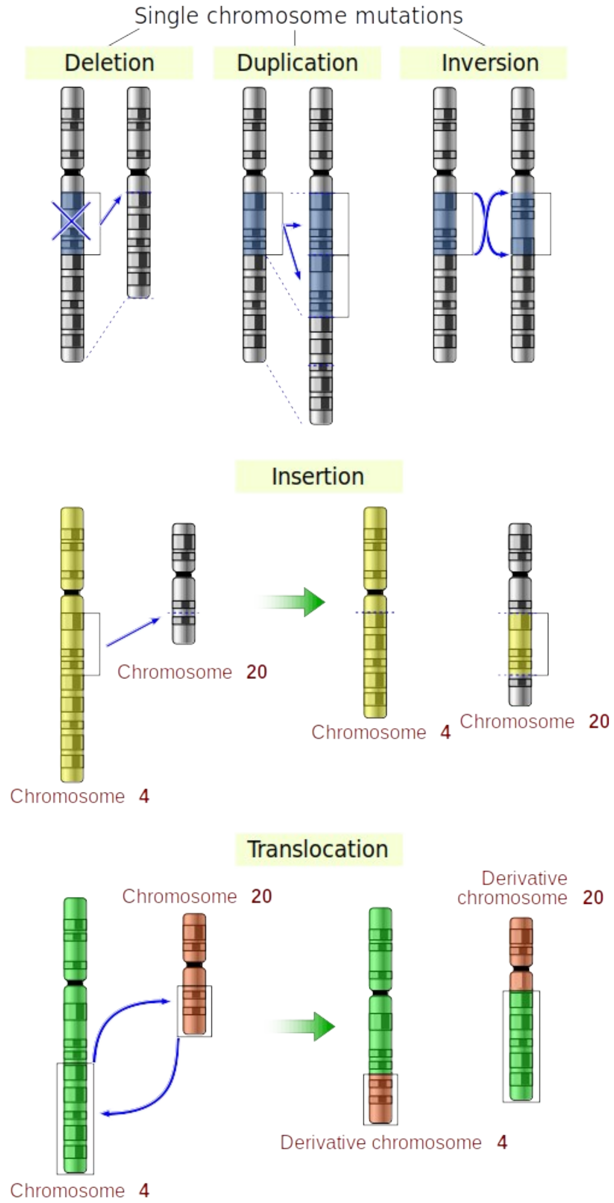
2. Epigenetic Inheritance: Alterations in gene expression passed on to the next generation, keeping DNA unchaged.

3. Multilevel selection: Multilevel selection (MLS) theory in biology clarifies evolutionary processes when populations are structured in groups. Within groups, selection is driven by differences in fitness between members, or "relative fitness." Likewise, if groups compete within a supergroup, selection of groups will depend on the relative fitness of groups. MLS states that processes at both levels matter.

Fig. 1. A simplified graphical model of multilevel selection. In social dilemmas, outcomes depend on the level of organization on which selection operates most strongly. To determine the dominant level of selection, the direction and magnitude of selection at the relevant levels should be estimated and compared.

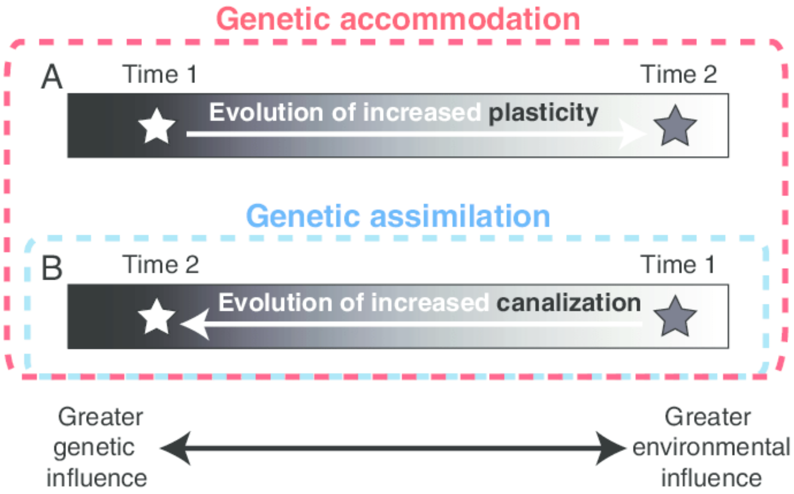


4. Genomic Evolution: Genome evolution is the process by which a genome changes in structure (sequence) or size over time, via the route of mutations, duplications etc. The study of genome evolution involves multiple fields such as structural analysis of the genome, the study of genomic parasites, gene and ancient genome duplications, polyploidy (Polyploidy is the state of a cell or organism having more than two paired (homologous) sets of chromosomes) and comparative genomics.



Chromosomal Mutations: Chromosomal mutations over time can accumulate and promote diversity and evolution if a produced trait is favorable.

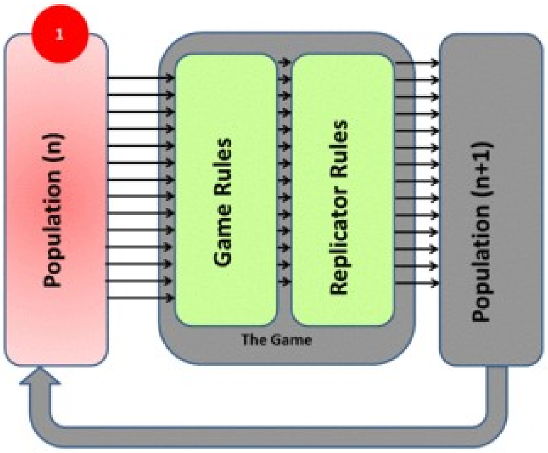
5. Plasticity and accommodation: The following diagram illustrating the distinction between genetic accommodation and genetic assimilation. Genetic accommodation is any adaptive genetic change in the environmental regulation of a phenotype. For example, a trait may evolve either (A) increased or (B) decreased environmental sensitivity (i.e. phenotypic plasticity). The complete loss of phenotypic plasticity (i.e. increased canalization) is an extreme form of genetic accommodation known as genetic assimilation.



6. Niche construction: Niche construction is the process by which an organism alters its own local environment. These alterations can be a physical change to the organism's environment or encompass when an organism actively moves from one habitat to another to experience a different environment.

7. Replicator theory: Evolutionary game theory analyses Darwinian mechanisms with a system model with three main components – Population, Game, and Replicator Dynamics. The model (as evolution itself) deals with a Population (Pn). The population will exhibit Variation among Competing individuals (Game). The mix of strategies in the population affects the results, i.e. the Fitness (Production rate of offspring). Based on this resulting fitness each member of the population then undergoes replication or culling determined by the exact mathematics of the Replicator Dynamics Process. This overall process then produces a New Generation P(n+1). Each surviving individual now has a new fitness level determined by the result of the competition. The new generation then takes the place of the previous one

and the cycle repeats. The population mix may converge to an Evolutionarily Stable State that cannot be invaded by any mutant strategy.



8. Evolvability: It defines a population with respect to mutation. Now what is mutation? Definition of mutation depends on the level of organization. In Genotype mutation, gene sequence is changed. In Phenotype mutation, the external structure is altered. If in a genotype mutation, phenotype does not change, then it is called neutral mutation. If phenotype changes, it is called non-neutral. Now, ability of a system to withstand mutations is called its robustness. On the other hand, ability of mutations to produce heritable phenotypic variations is called Evolvability.

All these highly complex fields discussed above, together, form the latest theory of evolution called Integrated Model.

Variations and Adaptation

1. Variation- appearance of inherited trait that makes an individual different from other members of same species.

a. Can be small (# of petals on a flower) or large (albino animal)

b. New species can evolve from variations if enough of the variation spreads throughout a population.

2. adaptation- helpful variation that makes an organism better suited to their environment

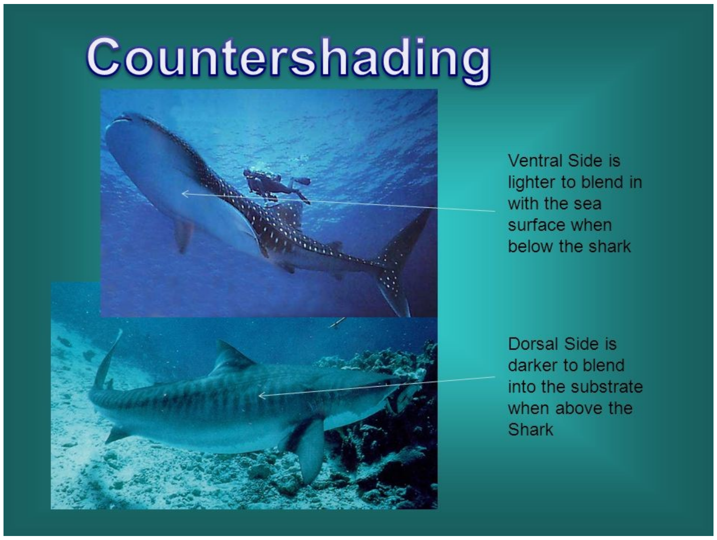
3. mutation- source of variation caused by a change in the DNA

4. The isolation of some individuals from others by geography or climate can also result in evolutionary change.

Concealing coloration (arctic white animals), **disruptive coloration** (tiger, zebra), **disguises** (the insect called the walking stick), **counter shading** (When the coloration of the upper parts of an animal is darker than its undersides. This causes the effect of sunlight to be counteracted. Most whales are counter shaded.

If you look up at a whale you would see a light color, just as if the whale was not there- blends in with the sky. If you look down on a whale you see darkness like the ocean floor.), **mimicry** (hairstreak butterflies: they have false antennae and spots on the back of their wings. Birds peck at the wrong end trying to grab the head, and the butterfly can get away.) fall among various techniques of Variation and Adaptation.

The walking stick



Countershading

Ventral Side is lighter to blend in with the sea surface when below the shark

Dorsal Side is darker to blend into the substrate when above the Shark

Viceroy butterfly
(The mimic)

Monarch butterfly
(The model)

In summary, Darwin’s Theory of Evolution was a groundbreaking advancement, explaining how natural selection results in the inherited biological change within a population. This is evolution. Biological fitness is central to this theory, and although many people understand that the fittest survive, not all understand what this truly means. Biological fitness is measured by the ability of an organism to reproduce and successfully pass on its genes to future generations. Misconceptions arise when individuals perceive the largest, strongest organisms within a population to be the most biologically fit. To demonstrate fitness in the context of evolution, one need only look at butterflies. They come in all shapes, sizes and colors, sometimes adopting another species’ physical characteristics in a process known as mimicry. Mimicry comes in several varieties, including Batesian mimicry, which is when a palatable organism mimics a species that is unpalatable to predators. Consequently, they are avoided by predators, increasing their fitness.

A vivid example of Batesian mimicry is depicted by Viceroy and Monarch Butterflies. Monarch butterflies are unpalatable due to toxic milkweeds they consume as larvae, which results in low levels of predation in their natural environment. Viceroy butterflies have wings emblazoned with similar shape and color schemes, ostensibly reducing the predation rate. Colors must be matched very closely as avian predators have some of the most developed eyes in the animal kingdom.

Mechanism of Evolution

Key points:

1. Microevolution is a change in the frequency of gene variants, alleles, in a population, typically occurring over a relatively short time period.

2. Population genetics is the field of biology that studies allele frequencies in populations and how they change over time.

3. Allele frequency refers to how common an allele is in a population. It is determined by counting how many times the allele appears in the population then dividing by the total number of copies of the gene.

Frequency of allele A=Number of copies of allele A in population/ Total number of copies of gene in population

4. The gene pool of a population consists of all the copies of all the genes in that population.

5. When a population is in Hardy-Weinberg equilibrium for a gene, it is not evolving, and allele frequencies will stay the same across generations.

6. There are five basic Hardy-Weinberg assumptions: no mutation, random mating, no gene flow, infinite population size, and no selection.

7. If the assumptions are not met for a gene, the population may evolve for that gene (the gene's allele frequencies may change).

8. Mechanisms of evolution correspond to violations of different Hardy-Weinberg assumptions. They are: mutation, non-random mating, gene flow, finite population size (genetic drift), and natural selection.

Introduction

In nature, populations are usually evolving. The grass in an open meadow, the wolves in a forest, and even the bacteria in a person's body are all natural populations. And all of these populations are likely to be evolving for at least some of their genes. Evolution is happening right here, right now!

To be clear, that doesn't mean these populations are marching towards some final state of perfection. All evolution means is that a population is changing in its genetic makeup over generations. And the changes may be subtle—for instance, in a wolf population, there might be a shift in the frequency of a gene variant for black rather than gray fur. Sometimes, this type of change is due to natural selection. Other times, it comes from migration of new organisms into the population, or from random events—the evolutionary "luck of the draw."

Darwin meets Mendel—not literally

When Darwin came up with his theories of evolution and natural selection, he knew that the processes he was describing depended on heritable variation in populations. That is, they relied on differences in the features of the organisms in a population and on the ability of these different features to be passed on to offspring.

Darwin did not, however, know how traits were inherited. Like other scientists of his time, he thought that traits were passed on via blending inheritance. In this model, parents' traits are supposed to permanently blend in their offspring. The blending model was disproven by Austrian monk Gregor Mendel, who found that traits are specified by non-blending heritable units called genes.

Although Mendel published his work on genetics just a few years after Darwin published his ideas on evolution, Darwin probably never read Mendel’s work. Today, we can combine Darwin’s and Mendel’s ideas to arrive at a clearer understanding of what evolution is and how it takes place.

Microevolution and population genetics

Microevolution, or evolution on a small scale, is defined as a change in the frequency of gene variants, alleles, in a population over generations. The field of biology that studies allele frequencies in populations and how they change over time is called population genetics.

Microevolution is sometimes contrasted with macroevolution, evolution that involves large changes, such as formation of new groups or species, and happens over long time periods. However, most biologists view microevolution and macroevolution as the same process happening on different timescales. Microevolution adds up gradually, over long periods of time to produce macroevolutionary changes.

Let's look at three concepts that are core to the definition of microevolution: populations, alleles, and allele frequency.

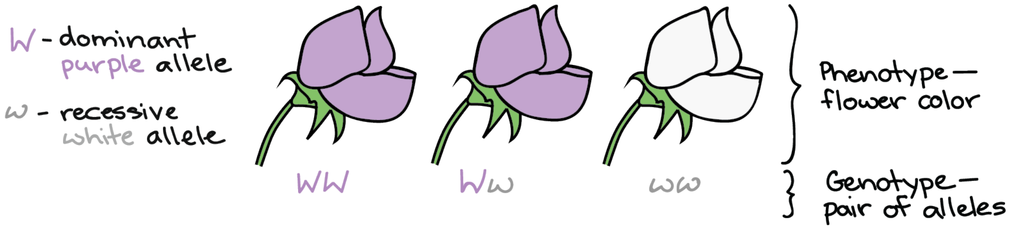
Populations

A population is a group of organisms of the same species that are found in the same area and can interbreed. A population is the smallest unit that can evolve—in other words, an individual can’t evolve.

Alleles

An allele is a version of a gene, a heritable unit that controls a particular feature of an organism.

For instance, Mendel studied a gene that controls flower color in pea plants. This gene comes in a white allele, w, and a purple allele, W. Each pea plant has two gene copies, which may be the same or different alleles. When the alleles are different, one—the dominant allele, W—may hide the other—the recessive allele, w. A plant's set of alleles, called its genotype, determines its phenotype, or observable features, in this case flower color.



Allele frequency

Allele frequency refers to how frequently a particular allele appears in a population. For instance, if all the alleles in a population of pea plants were purple alleles, W, the allele frequency of W would be 100% or 1.0. However, if half the alleles were W and half were w, each allele would have an allele frequency of 50% or 0.5.

In general, we can define allele frequency as:

Frequency of allele A= Number of copies of allele A in population / Total no. of A/a gene copies in population

Sometimes there are more than two alleles in a population (e.g., there might be A, a, and A_i alleles of a gene). In that case, you would want to add up all of the different alleles to get your denominator.

It's also possible to calculate genotype frequencies—the fraction of individuals with a given genotype—and phenotype frequencies—the fraction of individuals with a given phenotype. Keep in mind, though, that these are different concepts from allele frequency. We'll see an example of this difference next.

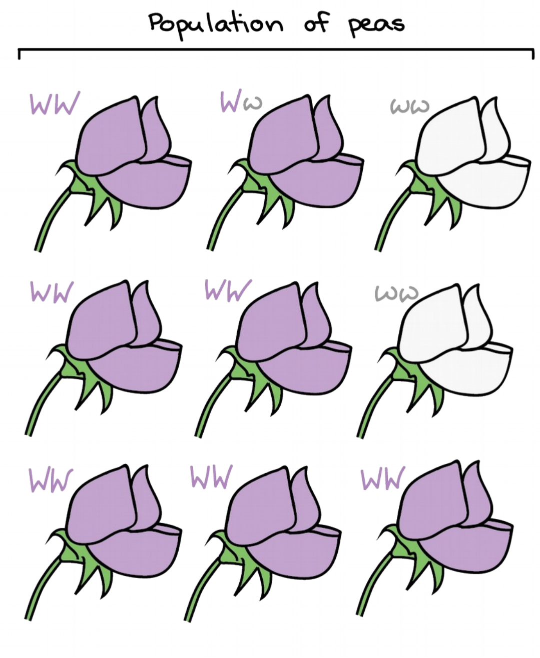
Example: Finding allele frequency

Let’s look at an example. Consider the very small population of nine pea plants shown below. Each pea plant has two copies of the flower color gene. If we look at the two gene copies in each plant and count up how many W copies are present, we find there are 13. If we count up how many w copies are present, we find that there are five. The total number of gene copies in the whole population is 13+5=18.

We can divide the number of copies of each allele by the total number of copies to get the allele frequency. By convention, when there are just two alleles for a gene in a population, their frequencies are given the symbols p and q:

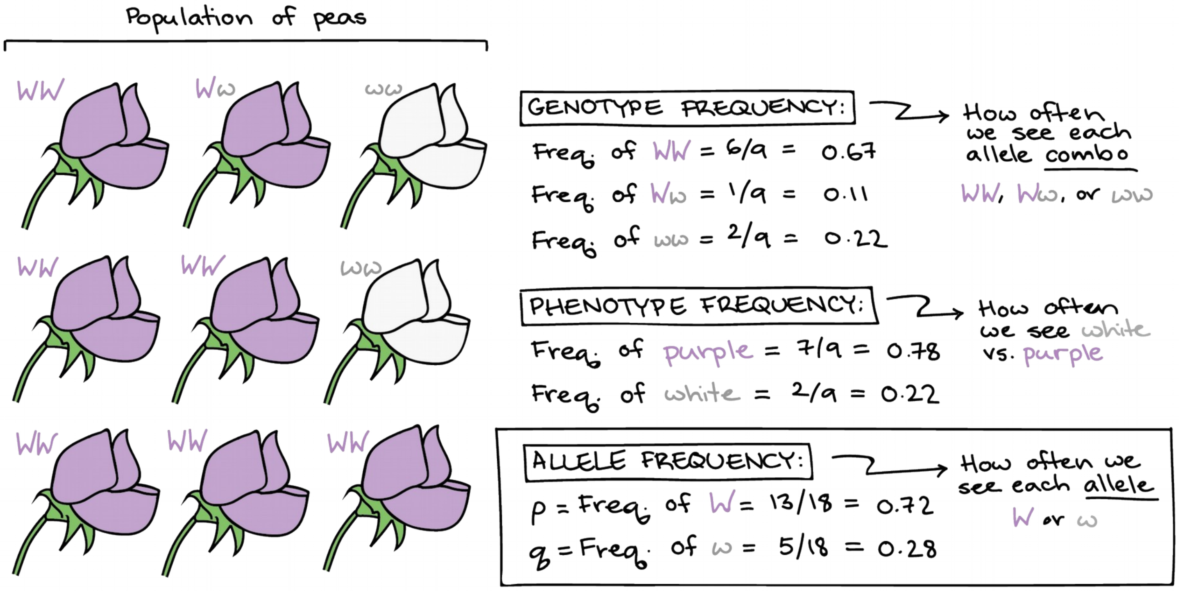
p= frequency of W=13/18=0.72 or 72%

q=frequency of w=5/18=0.28 or 28%

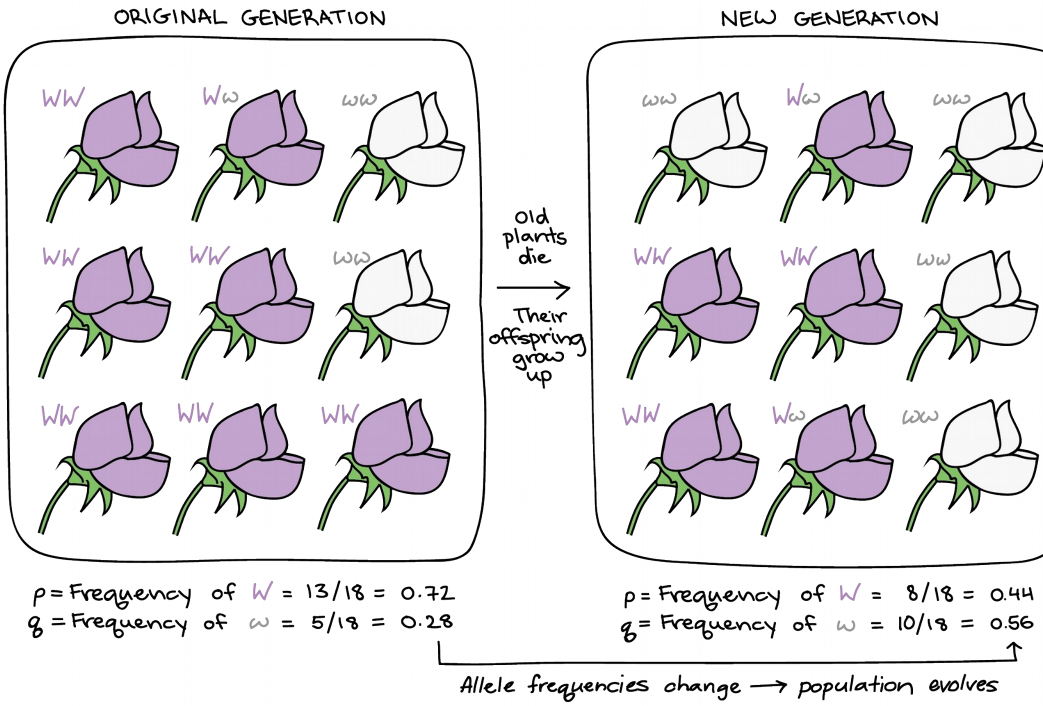


The frequencies of all the alleles of a gene must add up to one, or 100%.

Allele frequency is different from genotype frequency or phenotype frequency. Genotype and phenotype frequencies can also be calculated and are important for understanding how populations evolve, but they are not the same thing as allele frequency. The diagram below shows the difference:



Now, let's suppose we come back a generation later and check the genotypes of the new pea plants that now make up the population. To find the allele frequencies, we again look at each individual's genotype, count the number of copies of each allele, and divide by the total number of gene copies. Now, we find the frequency of W has dropped to 8/18=0.44 and the frequency of w has risen to 10/18 or 0.56.



There has been a change in allele frequencies in the population over generations, so—by the definition of microevolution—we can say that the population has evolved. If we were actually doing research, we might want to use a statistical test to confirm that these proportions were really different.

We’ll examine the factors that cause a population to evolve, including natural selection, genetic drift—random change—and others factors.

The gene pool

The total set of gene copies for all genes in a population is referred to as its gene pool. The gene pool gets its name from the idea that we are essentially taking all the gene copies—for all genes—in the individuals of a population and dumping them into one large, common pool.

What would this look like? In the example above, we went through all nine individuals in the population and looked at their copies of the flower color gene. There were 18 individual gene copies, each of which was a W or a w allele. Now, imagine that we went through this same process for every single gene in the pea plant, including genes that control height, seed color, seed shape, metabolism, etc. There would be 18 copies of each gene pulled out and dumped into the common pool. At the end of this process, the common pool of gene copies will be the gene pool of our population.

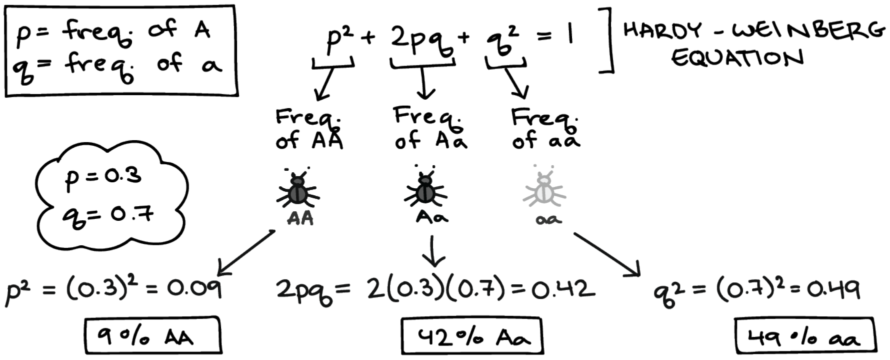
By looking at all the copies of all the genes in a population, we can see globally how much genetic variation there is in the population. The more variation a population has, the better its ability to adapt to changes in its environment through natural selection. If there is more variation, the odds are better that there will be some alleles already present that allow organisms to survive and reproduce effectively under the new conditions.

Hardy-Weinberg equilibrium

First, let's see what it looks like when a population is not evolving. If a population is in a state called Hardy-Weinberg equilibrium, the frequencies of alleles, or gene versions, and genotypes, or sets of alleles, in that population will stay the same over generations (and will also satisfy the Hardy-Weinberg equation). Formally, evolution is a change in allele frequencies in a population over time, so a population in Hardy-Weinberg equilibrium is not evolving.

That's a little bit abstract, so let's break it down using an example. Imagine we have a large population of beetles. In fact, just for the heck of it, let's say this population is infinitely large. The beetles of our infinitely large population come in two colors, dark gray and light gray, and their color is determined by the A gene. AA and Aa beetles are dark gray, and aa beetles are light gray.

In our population, let's say that the A allele has a frequency of 0.3, while the a allele has a frequency of 0.7. If a population is in Hardy-Weinberg equilibrium, allele frequencies will be related to genotype frequencies by a specific mathematical relationship, the Hardy-Weinberg equation. So, we can predict the genotype frequencies we'd expect to see (if the population is in Hardy-Weinberg equilibrium) by plugging in allele frequencies as shown below:



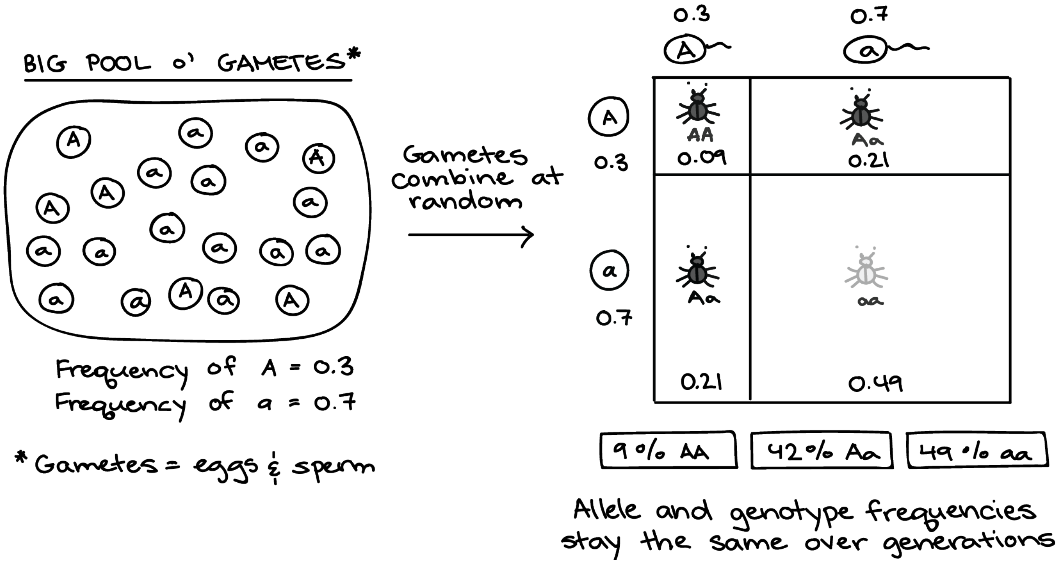
Let's imagine that these are, in fact, the genotype frequencies we see in our beetle population (9% AA, 42% Aa, 49% aa). Excellent—our beetles appear to be in Hardy-Weinberg equilibrium! Now, let's imagine that the beetles reproduce to make a next generation. What will the allele and genotype frequencies will be in that generation?

To predict this, we need to make a few assumptions:

First, let's assume that none of the genotypes is any better than the others at surviving or getting mates. If this is the case, the frequency of A and a alleles in the pool of gametes (sperm and eggs) that meet to make the next generation will be the same as the overall frequency of each allele in the present generation.

Second, let's assume that the beetles mate randomly (as opposed to, say, black beetles preferring other black beetles). If this is the case, we can think of reproduction as the result of two random events: selection of a sperm from the population's gene pool and selection of an egg from the same gene pool. The probability of getting any offspring genotype is just the probability of getting the egg and sperm combo(s) that produce that genotype.

We can use a modified Punnett square to represent the likelihood of getting different offspring genotypes. Here, we multiply the frequencies of the gametes on the axes to get the probability of the fertilization events in the squares:



As shown above, we'd predict an offspring generation with the exact same genotype frequencies as the parent generation: 9% AA, 42% Aa, and 49% aa. If genotype frequencies have not changed, we also must have the same allele frequencies as in the parent generation: 0.3 for A and 0.7 for a.

What we've just seen is the essence of Hardy-Weinberg equilibrium. If alleles in the gamete pool exactly mirror those in the parent generation, and if they meet up randomly (in an infinitely large number of events), there is no reason—in fact, no way—for allele and genotype frequencies to change from one generation to the next.

In the absence of other factors, you can imagine this process repeating over and over, generation after generation, keeping allele and genotype frequencies the same. Since evolution is a change in allele frequencies in a population over generations, a population in Hardy-Weinberg equilibrium is, by definition, not evolving.

But is that realistic?

As we mentioned at the beginning of the article, populations are usually not in Hardy-Weinberg equilibrium (at least, not for all of the genes in their genome). Instead, populations tend to evolve: the allele frequencies of at least some of their genes change from one generation to the next.

In fact, population geneticists often check to see if a population is in Hardy-Weinberg equilibrium because they suspect other forces may be at work. If the population’s allele and genotype frequencies are changing over generations (or if the allele and genotype frequencies don't match the predictions of the Hardy-Weinberg equation), the race is on to find out why.

Hardy-Weinberg assumptions and evolution

What causes populations to evolve? In order for a population to be in Hardy-Weinberg equilibrium, or a non-evolving state, it must meet five major assumptions:

1. No mutation. No new alleles are generated by mutation, nor are genes duplicated or deleted.

2. Random mating. Organisms mate randomly with each other, with no preference for particular genotypes.

3. No gene flow. Neither individuals nor their gametes (e.g., windborne pollen) enter or exit the population.

4. Very large population size. The population should be effectively infinite in size.

5. No natural selection. All alleles confer equal fitness (make organisms equally likely to survive and reproduce).

If any one of these assumptions is not met, the population will not be in Hardy-Weinberg equilibrium. Instead, it may evolve: allele frequencies may change from one generation to the next. Allele and genotype frequencies within a single generation may also fail to satisfy the Hardy-Weinberg equation.

Some genes may satisfy Hardy-Weinberg, while others do not

Note that we can think about Hardy-Weinberg equilibrium in two ways: for just one gene, or for all the genes in the genome.

1. If we look at just one gene, we check whether the above criteria are true for that one gene. For example, we would ask if there were mutations in that gene, or if organisms mated randomly with regards to their genotype for that gene.

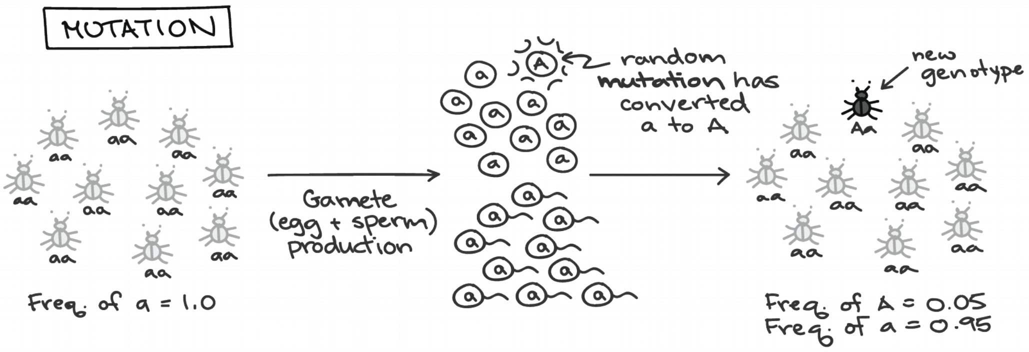
2. If we look at all the genes in the genome, the conditions have to be met for every single gene.

While it’s possible that the conditions will be more or less met for a single gene under certain circumstances, it’s very unlikely that they would be met for all the genes in the genome. So, while a population may be in Hardy-Weinberg equilibrium for some genes (not evolving for those genes), it’s unlikely to be in Hardy-Weinberg equilibrium for all of its genes (not evolving at all).

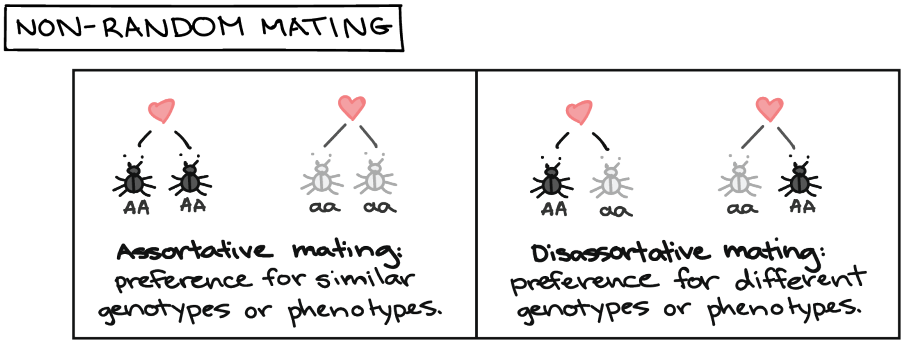
Mechanisms of evolution

Different Hardy-Weinberg assumptions, when violated, correspond to different mechanisms of evolution.

1. Mutation: Although mutation is the original source of all genetic variation, mutation rate for most organisms is pretty low. So, the impact of brand-new mutations on allele frequencies from one generation to the next is usually not large. (However, natural selection acting on the results of a mutation can be a powerful mechanism of evolution!)

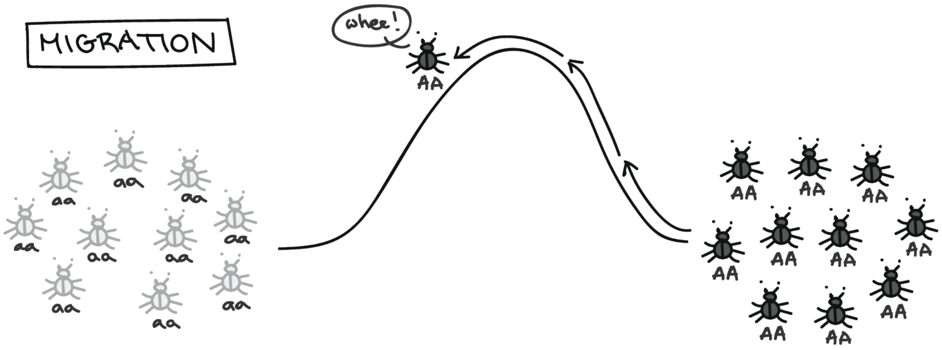


2. Non-random mating: In non-random mating, organisms may prefer to mate with others of the same genotype or of different genotypes. Non-random mating won't make allele frequencies in the population change by itself, though it can alter genotype frequencies. This keeps the population from being in Hardy-Weinberg equilibrium, but it’s debatable whether it counts as evolution, since the allele frequencies are staying the same.

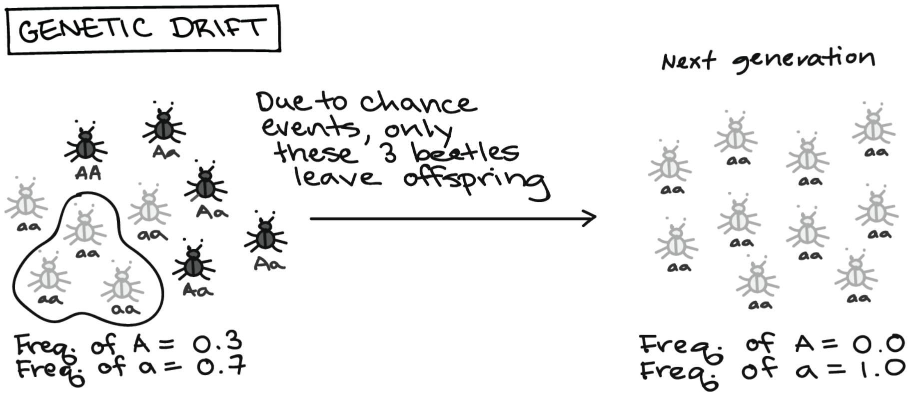


3. Gene

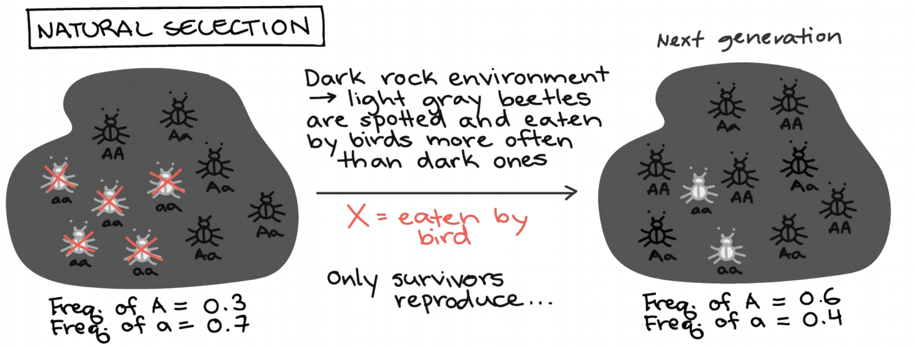
flow: Gene flow involves the movement of genes into or out of a population, due to either the movement of individual organisms or their gametes (eggs and sperm, e.g., through pollen dispersal by a plant). Organisms and gametes that enter a population may have new alleles, or may bring in existing alleles but in different proportions than those already in the population. Gene flow can be a strong agent of evolution.



4. Non-infinite population size (genetic drift): Genetic drift involves changes in allele frequency due to chance events – literally, "sampling error" in selecting alleles for the next generation. Drift can occur in any population of non-infinite size, but it has a stronger effect on small populations. We will look in detail at genetic drift and the effects of population size.



5. Natural selection: Finally, the most famous mechanism of evolution! Natural selection occurs when one allele (or combination of alleles of different genes) makes an organism more or less fit, that is, able to survive and reproduce in a given environment. If an allele reduces fitness, its frequency will tend to drop from one generation to the next. We will look in detail at different forms of natural selection that occur in populations.



All five of the above mechanisms of evolution may act to some extent in any natural population. In fact, the evolutionary trajectory of a given gene (that is, how its alleles change in frequency in the population across generations) may result from several evolutionary mechanisms acting at once. For instance, one gene’s allele frequencies might be modified by both gene flow and genetic drift. For another gene, mutation may produce a new allele, which is then favored (or disfavored) by natural selection.

Natural selection in populations

Key points:

1. Natural selection can cause microevolution (change in allele frequencies), with fitness-increasing alleles becoming more common in the population.

2. Fitness is a measure of reproductive success (how many offspring an organism leaves in the next generation, relative to others in the group).

3. Natural selection can act on traits determined by alternative alleles of a single gene, or on polygenic traits (traits determined by many genes).

4. Natural selection on traits determined by multiple genes may take the form of stabilizing selection, directional selection, or disruptive selection.

Introduction

We've already met a few different mechanisms of evolution. Genetic drift, migration, mutation...the list goes on. All of these mechanisms can make a population evolve, or change in its genetic makeup over generations.

But there's one mechanism of evolution that's a bit more famous than the others, and that's natural selection. What makes natural selection so special? Out of all the mechanisms of evolution, it's the only one that can consistently make populations adapted, or better-suited for their environment, over time.

We will examine natural selection at the level of population genetics, in terms of allele, genotype, and phenotype frequencies.

Quick review of natural selection

Here is a quick reminder of how a population evolves by natural selection:

1. Organisms with heritable (genetically determined) features that help them survive and reproduce in a particular environment tend to leave more offspring than their peers.

2. If this continues over generations, the heritable features that aid survival and reproduction will become more and more common in the population.

3. The population will not only evolve (change in its genetic makeup and inherited traits), but will evolve in such a way that it becomes adapted, or better-suited, to its environment.

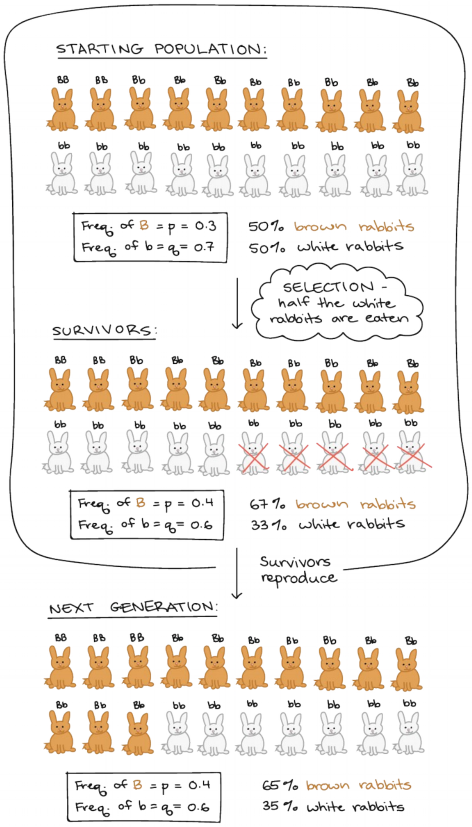
Natural selection can cause microevolution

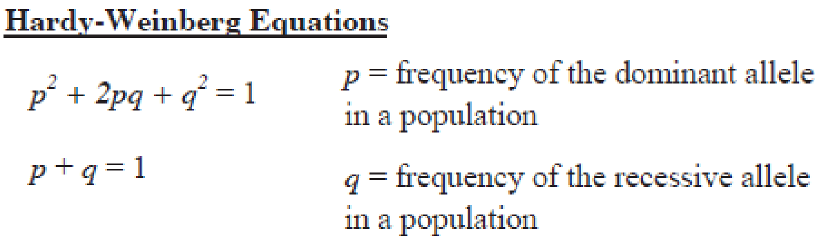
Natural selection acts on an organism's phenotype, or observable features. Phenotype is often largely a product of genotype (the alleles, or gene versions, the organism carries). When a phenotype produced by certain alleles helps organisms survive and reproduce better than their peers, natural selection can increase the frequency of the helpful alleles from one generation to the next – that is, it can cause microevolution.

Example: Rabbit coat color

As an example, let's imagine a population of brown and white rabbits, whose coat color is determined by dominant brown (B) and recessive white (b) alleles of a single gene. If a predator such as a hawk can see white rabbits (genotype bb) more easily than brown rabbits (BB and Bb) against the backdrop of a grassy field, brown rabbits are more likely than white rabbits to survive hawk predation. Because more brown than white rabbits will survive to reproduce, the next generation will probably contain a higher frequency of B alleles.

We can demonstrate this to ourselves by working through an example. Let's start with a set of allele and phenotype frequencies, shown in the diagram below, and see how they change in a generation if half of the white rabbits (but none of the brown rabbits) are eaten by hawks:





Hardy-Weinberg Equations

$p^2 + 2pq + q^2 = 1$

$p + q = 1$

p = frequency of the dominant allele in a population

q = frequency of the recessive allele in a population

An Example Calculation:

For that, we can use the Hardy-Weinberg equation.

You may think, that shouldn't work, because there's selection! You're right that the population is not in Hardy-Weinberg equilibrium, because its allele frequencies are changing from generation to generation. And if you check, you'll find that the allele frequencies and genotype frequencies of the survivors group do not match Hardy-Weinberg expectations.

However, once we know the allele frequencies in the survivors group, we can assume that the survivors will mate randomly with one another and use the Hardy-Weinberg equilibrium to predict the likely genotype and phenotype proportions of their offspring.

Allele frequencies:

frequency of B=p=0.4

frequency of b=q=0.6

Calculation of genotypes based on allele frequencies using Hardy-Weinberg:

Frequency of BB=p^2=0.16

Frequency of Bb=2pq=0.48

Frequency of bb=q^2=0.36

Calculation of number of individuals of each genotype (population size =20)

Number of BB rabbits= 0.16*20

Number of Bb rabbits= 0.48*20

Number of bb rabbits=0.36*20

We would predict about 3 BB, 10 Bb, and 7 bb rabbits. Since BB and Bb rabbits are brown, we would expect 3+10=13 brown rabbits. Only bb rabbits are white, so we would expect 7 white rabbits.

In this example, the frequency of the survival-promoting B allele rose from 0.3 to 0.4 in a single generation. The percent of the population with the survival-promoting brown phenotype also rose from 50% to 65%. (We can predict the next generation by assuming that the survivors mate randomly and leave equal numbers of offspring on average.) This is a made-up example, but it gives us a concrete sense of how natural selection can shift allele and phenotype frequencies to make a population better-suited to its environment. Will the recessive b alleles disappear from the population due to selection? Maybe someday, but not right away. That's because they can “hide” from predators in the heterozygous (Bb) brown rabbits. This is a good reminder that natural selection acts on phenotypes, not genotypes. A hawk can tell a brown rabbit from a white rabbit, but it can't tell an BB rabbit from an Bb rabbit.

Fitness = reproductive success

The phenotypes and genotypes favored by natural selection aren't necessarily just the ones that survive best. Instead, they're the ones with the highest overall fitness. Fitness is a measure of how well organisms survive and reproduce, with emphasis on "reproduce." Officially, fitness is defined as the number of offspring that organisms with a particular genotype or phenotype leave behind, on average, as compared to others in the population.

Survival is one important component of fitness. In order to leave any offspring at all in the next generation, an organism has to reach reproductive age. For instance, in the example above, brown rabbits had higher fitness than white rabbits, because a larger fraction of brown rabbits than white rabbits survived to reproduce. Living for a longer period of time may also allow an organisms to reproduce more separate times (e.g., with more mates or in multiple years).

However, survival is not the only part of the fitness equation. Fitness also depends on the ability to attract a mate and the number of offspring produced per mating. An organism that survived for many years, but never successfully attracted a mate or had offspring, would have very (zero) low fitness.

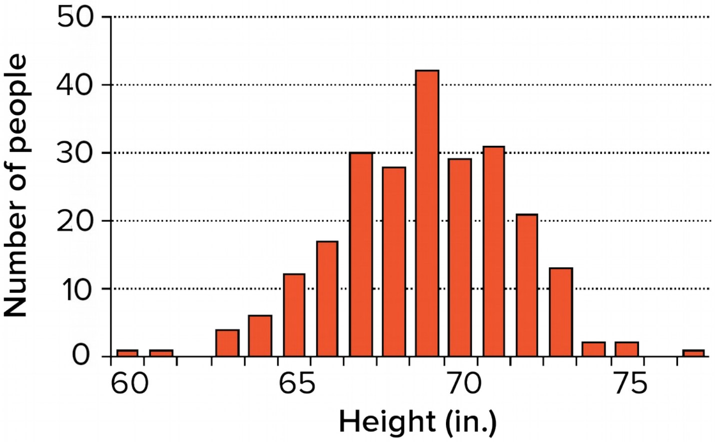
Fitness depends on the environment

Which traits are favored by natural selection (that is, which features make an organism more fit) depends on the environment. For example, a brown rabbit might be more fit than a white rabbit in a brownish, grassy landscape with sharp-eyed predators. However, in a light-colored landscape (such as sand dunes), white rabbits might be better than brown rabbits at avoiding predators. And if there weren't any predators, the two coat colors might be equally fit!

In many cases, a trait also involves tradeoffs. That is, it may have some positive and some negative effects on fitness. For instance, a particular coat color might make a rabbit less visible to predators, but also less attractive to potential mates. Since fitness is a function of both survival and reproduction, whether the coat color is a net "win" will depend on the relative strengths of the predation and the mate preference.

Natural selection can act on traits controlled by many genes

In some cases, different phenotypes in a population are determined by just one gene. For example, this was the case with our hypothetical rabbits. It's also true in some real cases of natural selection for coat color (e.g., in mice). However, in many cases, phenotypes are controlled by multiple genes that each make a small contribution overall result. Such phenotypes are



often called polygenic traits, and they typically form a spectrum, taking many slightly different forms. Plotting the frequency of the different forms in a population often results in a graph with a bell curve shape. Height (see graph below) and many other traits in humans are polygenic.

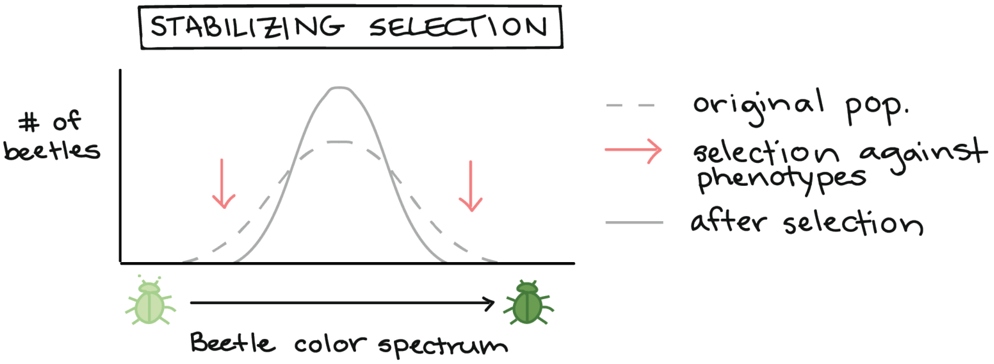
We can see if natural selection is acting on a polygenic trait by watching how the distribution of phenotypes in the population changes over time. Certain characteristic shifts tell us selection is occurring, even if we don't know exactly which genes control the trait.

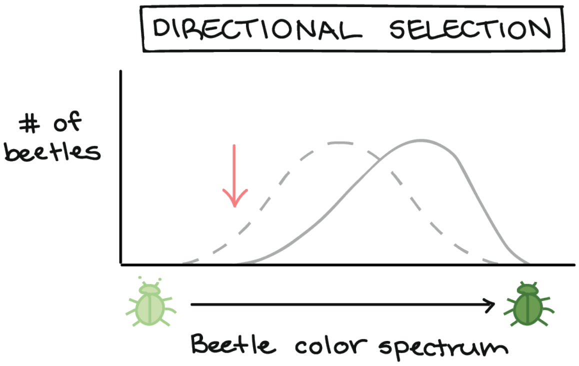
How natural selection can shift phenotype distributions

There are three basic ways that natural selection can influence distribution of phenotypes for polygenic traits in a population. To illustrate these forms of selection, let's use an imaginary beetle population, in which beetle color is controlled by many genes and varies in a spectrum from light to dark green.

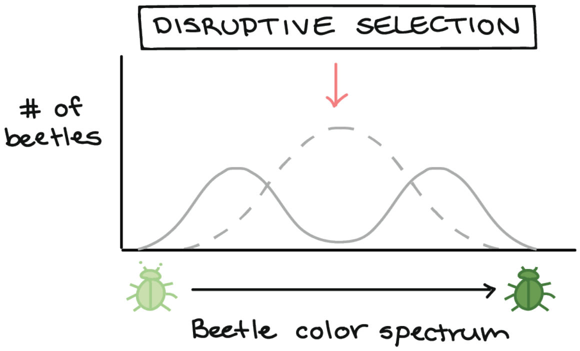
1. Stabilizing selection In stabilizing selection, intermediate phenotypes are more fit than extreme ones. For example, medium-green beetles might be the best camouflaged, and thus survive best, on a forest floor covered by medium-green plants. Stabilizing selection tends to narrow the curve.

2. Directional selection One extreme phenotype is more fit than all the other phenotypes. For example, if the beetle population moves into a new environment with dark soil and vegetation, the dark green beetles might be better hidden and survive better than medium or light beetles. Directional selection shifts the curve towards the favorable phenotype.





3. Disruptive selection Both extreme phenotypes are more fit than those in the middle. For example, if the beetles move into a new environment with patches of light-green moss and dark-green shrubs, both light and dark beetles might be better hidden (and survive better) than medium-green beetles. Diversifying selection makes multiple peaks in the curve.



Summary

Natural selection can cause microevolution, or a change in allele frequencies over time, with fitness-increasing alleles becoming more common in the population over generations. Fitness is a measure of relative reproductive success. It refers to how many offspring organisms of a particular genotype or phenotype leave in the next generation, relative to others in the group.

Natural selection can act on traits determined by different alleles of a single gene, or on polygenic traits (traits determined by many genes). Polygenic traits in a population often form a bell curve distribution. Natural selection on polygenic traits can take the form of:

1. **Stabilizing selection**: Intermediate phenotypes have the highest fitness, and the bell curve tends to narrow.

2. **Directional selection**: One of the extreme phenotypes has the highest fitness. The bell curve shifts towards the more fit phenotype.

3. **Disruptive selection**: Both extreme phenotypes have a higher fitness than intermediate phenotypes. The bell curve develops two peaks.

Mathematical Models of Evolution

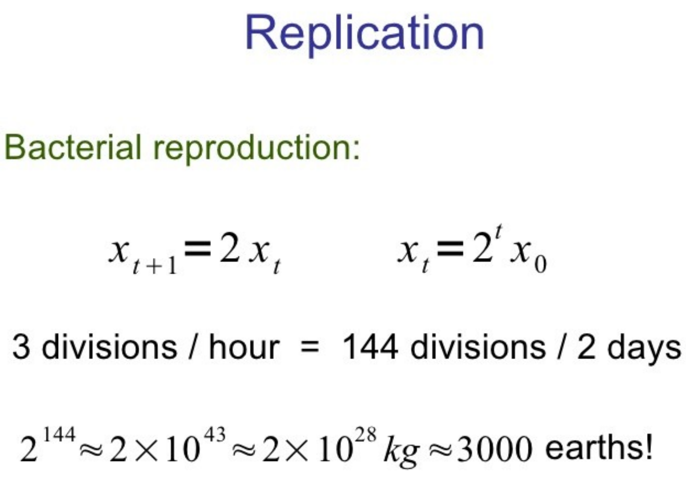
Now we can say after all our study, that Evolution is standing on three basic building blocks:

Replication

Selection

and mutation

Mathematical treatment of Replication:



Replication

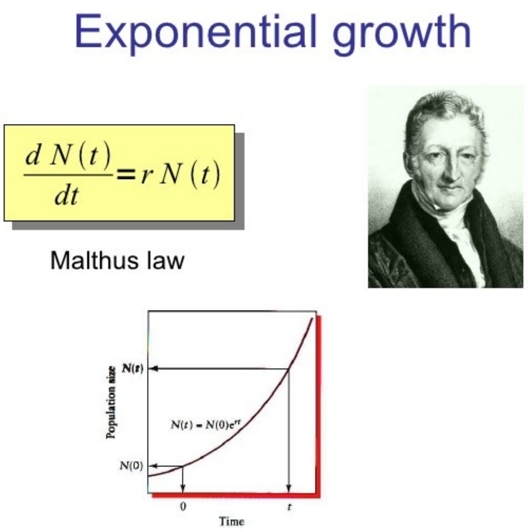
Bacterial reproduction:

$$x_{t+1}=2\,x_t \qquad x_t=2^t\,x_0$$

3 divisions / hour = 144 divisions / 2 days

$$2^{144}\approx 2\times 10^{43}\approx 2\times 10^{28}\,kg\approx 3000 \text{ earths!}$$

This relation hints at a equation where rate of increase of population size is proportional to the size of the population at a particular moment. This leads to the famous Malthus Law of exponential growth of organism.

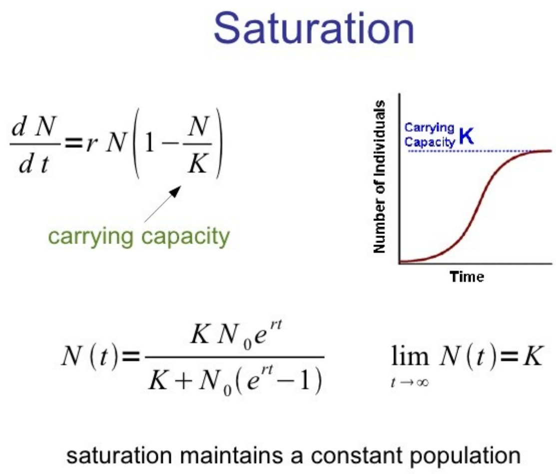


Exponential growth

$$\frac{d\,N(t)}{dt} = r\,N(t)$$

Malthus law

But this kind of growth is unphysical, it does not saturate the population and does not take into account the death of organism. Hence this equation is bit modified by introducing "carrying capacity".



Saturation

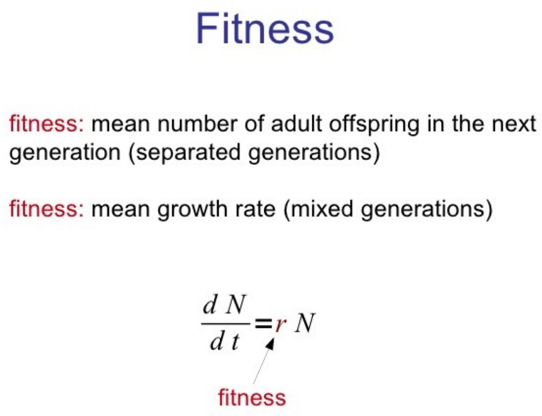
$$\frac{d\,N}{d\,t} = r\,N\left(1 - \frac{N}{K}\right)$$

carrying capacity

$$N(t) = \frac{K\,N_0 e^{rt}}{K + N_0(e^{rt} - 1)} \qquad \lim_{t \to \infty} N(t) = K$$

saturation maintains a constant population

Next we will switch to "Selection". We have encountered so many times the word "fitness". The entire theory of evolution seems to be pinned to this one word: Fitness. So, what, infact, fitness is? Mathematically?



Fitness

fitness: mean number of adult offspring in the next generation (separated generations)

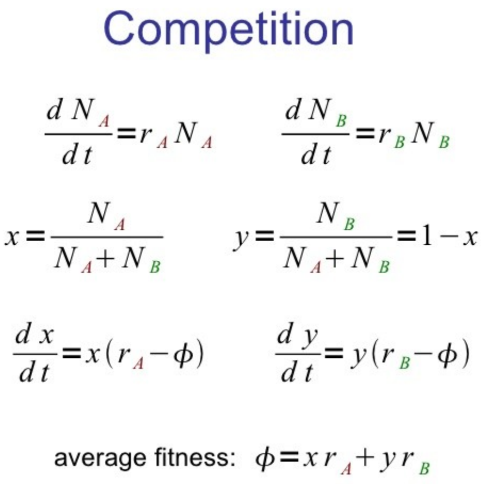
fitness: mean growth rate (mixed generations)

$$\frac{d\,N}{d\,t} = r\,N$$

fitness

How will you model mathematically "Competition"?

Combining both x and y into a single equation, we will get that master equation defining Survival of the Fittest:



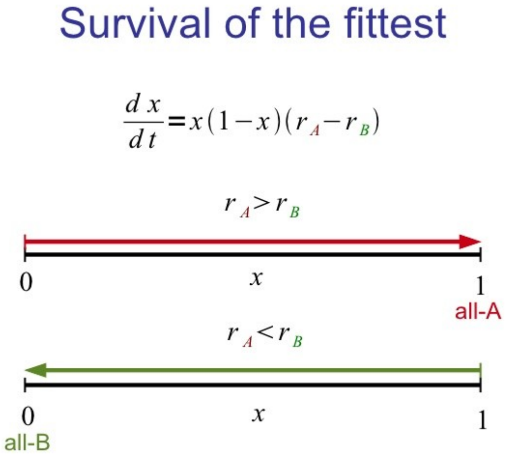
Competition

$$\frac{dN_A}{dt} = r_A N_A \qquad \frac{dN_B}{dt} = r_B N_B$$

$$x = \frac{N_A}{N_A + N_B} \qquad y = \frac{N_B}{N_A + N_B} = 1 - x$$

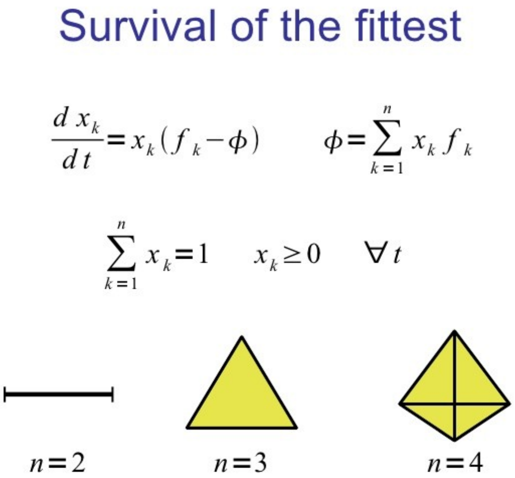
$$\frac{dx}{dt} = x(r_A - \phi) \qquad \frac{dy}{dt} = y(r_B - \phi)$$

average fitness: $\phi = x r_A + y r_B$



Survival of the fittest

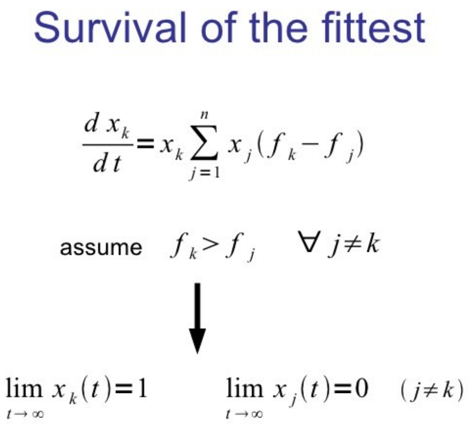
$$\frac{dx}{dt} = x(1-x)(r_A - r_B)$$



Survival of the fittest

$$\frac{dx_k}{dt} = x_k(f_k - \phi) \qquad \phi = \sum_{k=1}^{n} x_k f_k$$

$$\sum_{k=1}^{n} x_k = 1 \qquad x_k \geq 0 \qquad \forall t$$



Survival of the fittest

$$\frac{dx_k}{dt}=x_k\sum_{j=1}^{n}x_j(f_k-f_j)$$

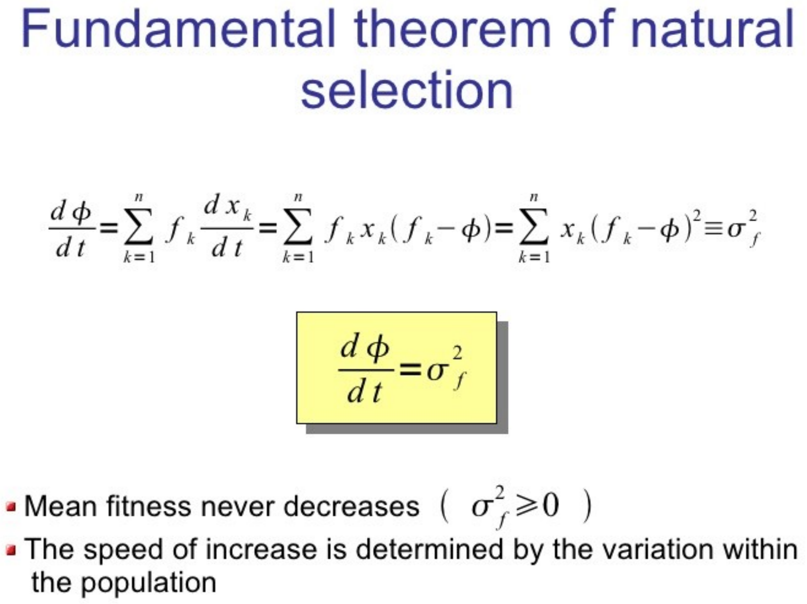
assume $f_k>f_j \quad \forall\, j\neq k$

$$\Big\downarrow$$

$$\lim_{t\to\infty}x_k(t)=1 \qquad \lim_{t\to\infty}x_j(t)=0 \quad (j\neq k)$$

Next, we will discuss the concept of symbiosis and competition between two species in case of composition-dependent fitness:

and thus, we will get the fundamental theorem of Natural Selection:

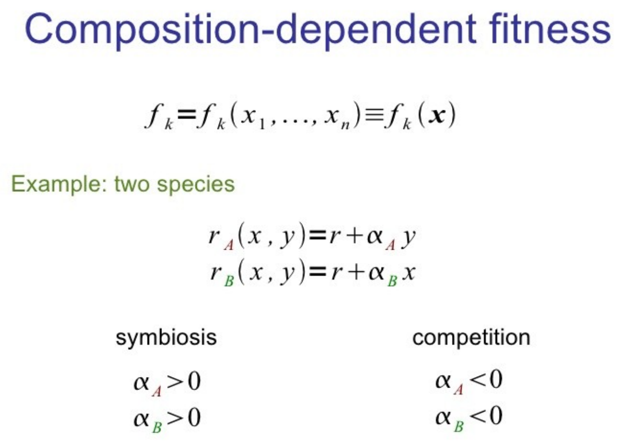


Fundamental theorem of natural selection

$$\frac{d\phi}{dt}=\sum_{k=1}^{n}f_k\frac{dx_k}{dt}=\sum_{k=1}^{n}f_k x_k(f_k-\phi)=\sum_{k=1}^{n}x_k(f_k-\phi)^2\equiv\sigma_f^2$$

$$\boxed{\frac{d\phi}{dt}=\sigma_f^2}$$

- Mean fitness never decreases ($\sigma_f^2\geqslant 0$)
- The speed of increase is determined by the variation within the population



Composition-dependent fitness

$$f_k = f_k(x_1, \ldots, x_n) \equiv f_k(\boldsymbol{x})$$

Example: two species

$$r_A(x, y) = r + \alpha_A y$$
$$r_B(x, y) = r + \alpha_B x$$

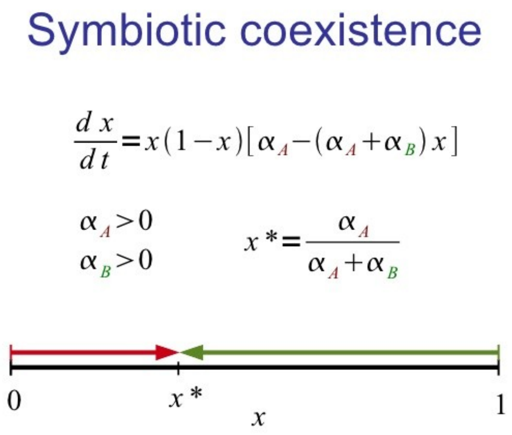
symbiosis

$$\alpha_A > 0$$
$$\alpha_B > 0$$

competition

$$\alpha_A < 0$$
$$\alpha_B < 0$$

In case of symbiosis, there is an equilibrium value of population size where component population coexist peacefully:

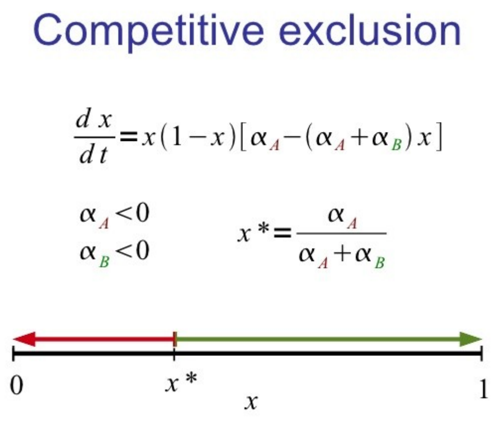


Symbiotic coexistence

$$\frac{dx}{dt} = x(1-x)[\alpha_A - (\alpha_A + \alpha_B)x]$$

$$\alpha_A > 0 \qquad x^* = \frac{\alpha_A}{\alpha_A + \alpha_B}$$
$$\alpha_B > 0$$

But in case of competition, this same condition would imply Competitive Exclusion:



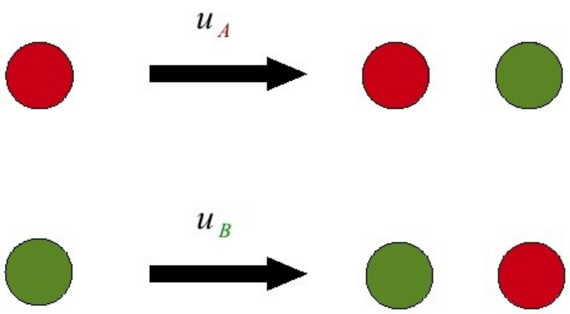
Competitive exclusion

$$\frac{dx}{dt} = x(1-x)[\alpha_A - (\alpha_A + \alpha_B)x]$$

$$\alpha_A < 0 \qquad x^* = \frac{\alpha_A}{\alpha_A + \alpha_B}$$
$$\alpha_B < 0$$

Next, we will deal with mathematical model of Mutation: here u_A and u_B represent the error rate. Therefore the rate equation will be a modified version of rate equation for Competition:

Replication with error



Replication with error

$$\frac{dx}{dt} = x r_A (1-u_A) + y r_B u_B - \phi x$$

$$\frac{dy}{dt} = x r_A u_A + y r_B (1-u_B) - \phi y$$

$$\phi = x r_A + y r_B$$

$$\frac{dx}{dt} = x(1-x)(r_A - r_B) + (1-x) r_B u_B - x r_A u_A$$

$x=0$ $x=1$ are not equilibria

y has been replaced by the mean fitness parameter here.

Replication with error

$$x^* \approx 1 - \frac{r_A u_A}{r_A - r_B} \qquad r_A > r_B$$

$$x^* \approx \frac{r_B u_B}{r_B - r_A} \qquad r_A < r_B$$

$$x^* = \frac{u_A}{u_A + u_B} \qquad r_A = r_B$$

mutation is the source of variability

Evolution with mutation

mutation matrix $Q=(q_{ij})$ $\qquad \boldsymbol{u}=(1,\ldots,1)$

$$\sum_{j=1}^{n} q_{ij}=1 \qquad \Leftrightarrow \qquad Q\,\boldsymbol{u}^T=\boldsymbol{u}^T$$

$$R=\begin{pmatrix} r_1 & & 0 \\ & \ddots & \\ 0 & & r_n \end{pmatrix} \qquad W\equiv RQ$$

mutation-selection matrix

$$\boxed{\frac{d\boldsymbol{x}}{dt}=\boldsymbol{x}\,W-\phi\,\boldsymbol{x} \qquad \phi=\boldsymbol{x}\,W\,\boldsymbol{u}^T=\sum_{j=1}^{n} x_j r_j}$$

quasispecies equation

Evolution with mutation

fundamental theorem

may be negative!

$$\frac{d\phi}{dt}=\frac{d\boldsymbol{x}}{dt}W\,\boldsymbol{u}^T=\boldsymbol{x}\,W^2\boldsymbol{u}^T-\phi^2=\boldsymbol{x}\,(W-\phi\,I)^2\,\boldsymbol{u}^T$$

equilibria: $\quad \boldsymbol{x}^*W=\phi^*\boldsymbol{x}^*$

if Q is irreducible, ϕ^* is the largest eigenvalue of W

as $\boldsymbol{x}^*$ normally corresponds to a mixed population,
ϕ^* need not be the absolute maximum