**BRIEF CONTENT OF LECTURE MATERIALS – MOLECULAR BIOLOGY**

 **Introduction to molecular biology. Part I**

History and subject of molecular biology. Nucleotides and nucleic acids

1. The Chargaff, Griffith, Avery-MacLeod-McCarty, Hershey-Chase experiments and explain their significance.

2. Informational properties of macromolecules.

3.The central dogma of molecular biology.

4. The role of molecular biology in medicine.

5. The components of nucleosides and nucleotides.

6. The chains of nucleic acids in DNA and RNA.

Many people believe that American biologist James Watson and English physicist Francis Crick discovered DNA in the 1950s. In reality, this is not the case. Rather, DNA was first identified in the late 1860s by Swiss chemist Friedrich Miescher. Then, in the decades following Miescher's discovery, other scientists--notably, Phoebus Levene and Erwin Chargaff--carried out a series of research efforts that revealed additional details about the DNA molecule, including its primary chemical components and the ways in which they joined with one another. Without the scientific foundation provided by these pioneers, Watson and Crick may never have reached their groundbreaking conclusion of 1953: that the DNA molecule exists in the form of a three-dimensional double helix. Many people believe that American biologist James Watson and English physicist Francis Crick discovered DNA in the 1950s. In reality, this is not the case. Rather, DNA was first identified in the late 1860s by Swiss chemist Friedrich Miescher. Then, in the decades following Miescher's discovery, other scientists--notably, Phoebus Levene and Erwin Chargaff--carried out a series of research efforts that revealed additional details about the DNA molecule, including its primary chemical components and the ways in which they joined with one another. Without the scientific foundation provided by these pioneers, Watson and Crick may never have reached their groundbreaking conclusion of 1953: that the DNA molecule exists in the form of a three-dimensional double helix.

Although few people realize it, 1869 was a landmark year in genetic research, because it was the year in which Swiss physiological chemist Friedrich Miescher first identified what he called "nuclein" inside the nuclei of human white blood cells. More than 50 years passed before the significance of Miescher's discovery of nucleic acids was widely appreciated by the scientific community. Meanwhile, even as Miescher's name fell into obscurity by the twentieth century, other scientists continued to investigate the chemical nature of the molecule formerly known as nuclein. One of these other scientists was Russian biochemist Phoebus Levene. Levene is credited with many firsts. For instance, he was the first to discover the order of the three major components of a single nucleotide (phosphate-sugar-base); the first to discover the carbohydrate component of RNA (ribose); the first to discover the carbohydrate component of DNA (deoxyribose); and the first to correctly identify the way RNA and DNA molecules are put together.

During the early years of Levene's career, neither Levene nor any other scientist of the time knew how the individual nucleotide components of DNA were arranged in space; discovery of the sugar-phosphate backbone of the DNA molecule was still years away. The large number of molecular groups made available for binding by each nucleotide component meant that there were numerous alternate ways that the components could combine. Several scientists put forth suggestions for how this might occur, but it was Levene's "polynucleotide" model that proved to be the correct one.

Erwin Chargaff was one of a handful of scientists who expanded on Levene's work by uncovering additional details of the structure of DNA, thus further paving the way for Watson and Crick. Chargaff, an Austrian biochemist, had read the [famous 1944 paper by Oswald Avery and his colleague](https://www.nature.com/scitable/topicpage/Isolating-Hereditary-Material-Frederick-Griffith-Oswald-Avery-336%22%20%5Co%20%22famous%201944%20paper%20by%20Oswald%20Avery%20and%20his%20colleague)[s](https://www.nature.com/scitable/topicpage/Isolating-Hereditary-Material-Frederick-Griffith-Oswald-Avery-336%22%20%5Co%20%22s) at Rockefeller University, which demonstrated that hereditary units, or genes, are composed of DNA. This paper had a profound impact on Chargaff, inspiring him to launch a research program that revolved around the chemistry of nucleic acids. Of Avery's work, Chargaff (1971) wrote the following:

*"This discovery, almost abruptly, appeared to foreshadow a chemistry of heredity and, moreover, made probable the nucleic acid character of the*gene*... Avery gave us the first text of a new language, or rather he showed us where to look for it. I resolved to search for this text."*

As his first step in this search, Chargaff set out to see whether there were any differences in DNA among different species. After developing a new paper chromatography method for separating and identifying small amounts of organic material, Chargaff reached two major conclusions (Chargaff, 1950). First, he noted that the nucleotide composition of DNA varies among species. In other words, the same nucleotides do not repeat in the same order, as proposed by Levene. Second, Chargaff concluded that almost all DNA--no matter what organism or tissue type it comes from--maintains certain properties, even as its composition varies. In particular, the amount of adenine (A) is usually similar to the amount of thymine (T), and the amount of guanine (G) usually approximates the amount of cytosine (C). In other words, the total amount of purines (A + G) and the total amount of pyrimidines (C + T) are usually nearly equal. (This second major conclusion is now known as "Chargaff's rule.") Chargaff's research was vital to the later work of Watson and Crick, but Chargaff himself could not imagine the explanation of these relationships--specifically, that A bound to T and C bound to G within the molecular structure of DNA

Chargaff's realization that A = T and C = G, combined with some crucially important X-ray crystallography work by English researchers Rosalind Franklin and Maurice Wilkins, contributed to Watson and Crick's derivation of the three-dimensional, double-helical model for the structure of DNA. Watson and Crick's discovery was also made possible by recent advances in model building, or the assembly of possible three-dimensional structures based upon known molecular distances and bond angles, a technique advanced by American biochemist Linus Pauling. In fact, Watson and Crick were worried that they would be "scooped" by Pauling, who proposed a different model for the three-dimensional structure of DNA just months before they did. In the end, however, Pauling's prediction was incorrect.

Using cardboard cutouts representing the individual chemical components of the four bases and other nucleotide subunits, Watson and Crick shifted molecules around on their desktops, as though putting together a puzzle. They were misled for a while by an erroneous understanding of how the different elements in thymine and guanine (specifically, the carbon, nitrogen, hydrogen, and oxygen rings) were configured. Only upon the suggestion of American scientist Jerry Donohue did Watson decide to make new cardboard cutouts of the two bases, to see if perhaps a different atomic configuration would make a difference. It did. Not only did the complementary bases now fit together perfectly (i.e., A with T and C with G), with each pair held together by hydrogen bonds, but the structure also reflected Chargaff's rule

Although scientists have made some minor changes to the Watson and Crick model, or have elaborated upon it, since its inception in 1953, the model's four major features remain the same yet today. These features are as follows:

* DNA is a double-stranded helix, with the two strands connected by hydrogen bonds. A bases are always paired with Ts, and Cs are always paired with Gs, which is consistent with and accounts for Chargaff's rule.
* Most DNA double helices are right-handed; that is, if you were to hold your right hand out, with your thumb pointed up and your fingers curled around your thumb, your thumb would represent the axis of the helix and your fingers would represent the sugar-phosphate backbone. Only one type of DNA, called Z-DNA, is left-handed.
* The DNA double helix is anti-parallel, which means that the 5' end of one strand is paired with the 3' end of its complementary strand (and vice versa). As shown in Figure 4, nucleotides are linked to each other by their phosphate groups, which bind the 3' end of one sugar to the 5' end of the next sugar.
* Not only are the DNA base pairs connected via hydrogen bonding, but the outer edges of the nitrogen-containing bases are exposed and available for potential hydrogen bonding as well. These hydrogen bonds provide easy access to the DNA for other molecules, including the proteins that play vital roles in the replication and expression of DNA

**Nucleic acid**, naturally occurring [chemical compound](https://www.britannica.com/science/chemical-compound) that is capable of being broken down to yield [phosphoric acid](https://www.britannica.com/science/phosphoric-acid), sugars, and a mixture of organic bases (purines and pyrimidines). Nucleic acids are the main information-carrying molecules of the [cell](https://www.britannica.com/science/cell-biology), and, by directing the process of [protein synthesis](https://www.britannica.com/science/translation-genetics), they determine the inherited characteristics of every living thing. The two main classes of nucleic acids are deoxyribonucleic acid ([DNA](https://www.britannica.com/science/DNA)) and ribonucleic acid ([RNA](https://www.britannica.com/science/RNA)). DNA is the master blueprint for life and [constitutes](https://www.merriam-webster.com/dictionary/constitutes) the genetic material in all free-living organisms and most viruses. RNA is the genetic material of certain viruses, but it is also found in all living cells, where it plays an important role in certain processes such as the making of proteins.

Nucleic acids are polynucleotides—that is, long chainlike molecules composed of a series of nearly identical building blocks called [nucleotides](https://www.britannica.com/science/nucleotide). Each [nucleotide](https://www.britannica.com/science/nucleotide) consists of a nitrogen-containing aromatic base attached to a pentose (five-carbon) [sugar](https://www.britannica.com/science/sugar-chemical-compound), which is in turn attached to a [phosphate](https://www.britannica.com/science/phosphate) group. Each nucleic acid contains four of five possible nitrogen-containing [base](https://www.britannica.com/science/base-nucleic-acid)s: [adenine](https://www.britannica.com/science/adenine) (A), [guanine](https://www.britannica.com/science/guanine) (G), [cytosine](https://www.britannica.com/science/cytosine) (C), [thymine](https://www.britannica.com/science/thymine) (T), and [uracil](https://www.britannica.com/science/uracil) (U). A and G are categorized as [purines](https://www.britannica.com/science/purine), and [C](https://www.britannica.com/science/carbon-chemical-element), T, and U are collectively called [pyrimidines](https://www.britannica.com/science/pyrimidine). All nucleic acids contain the bases A, C, and G; T, however, is found only in DNA, while U is found in RNA. The pentose sugar in DNA ([2′-deoxyribose](https://www.britannica.com/science/deoxyribose)) differs from the sugar in RNA (ribose) by the absence of a hydroxyl group (―OH) on the 2′ carbon of the sugar ring. Without an attached phosphate group, the sugar attached to one of the bases is known as a [nucleoside](https://www.britannica.com/science/nucleoside). The phosphate group connects successive sugar residues by bridging the 5′-hydroxyl group on one sugar to the 3′-hydroxyl group of the next sugar in the chain. These nucleoside linkages are called phosphodiester bonds and are the same in RNA and DNA.

DNA is a [polymer](https://www.britannica.com/science/polymer) of the four nucleotides A, [C](https://www.britannica.com/science/carbon-chemical-element), G, and T, which are joined through a backbone of alternating [phosphate](https://www.britannica.com/science/phosphate) and [deoxyribose](https://www.britannica.com/science/deoxyribose) [sugar](https://www.britannica.com/science/sugar-chemical-compound) residues. These nitrogen-containing bases occur in complementary pairs as determined by their ability to form hydrogen bonds between them. A always pairs with T through two hydrogen bonds, and G always pairs with C through three hydrogen bonds. The spans of A:T and G:C hydrogen-bonded pairs are nearly identical, allowing them to bridge the sugar-phosphate chains uniformly. This structure, along with the [molecule’s](https://www.britannica.com/science/molecule) chemical stability, makes DNA the ideal genetic material. The bonding between complementary bases also provides a mechanism for the replication of DNA and the transmission of genetic information. The double helical structure of normal DNA takes a right-handed form called the B-helix. The helix makes one complete turn approximately every 10 base pairs. B-DNA has two principal grooves, a wide major groove and a narrow minor groove. Many proteins interact in the space of the major groove, where they make sequence-specific contacts with the bases. In addition, a few proteins are known to make contacts via the minor groove.

RNA is a single-stranded nucleic acid [polymer](https://www.britannica.com/science/polymer) of the four nucleotides A, C, G, and U joined through a backbone of alternating [phosphate](https://www.britannica.com/science/phosphate) and [ribose](https://www.britannica.com/science/ribose) [sugar](https://www.britannica.com/science/sugar-chemical-compound) residues. It is the first intermediate in converting the information from DNA into proteins essential for the working of a cell. Some RNAs also serve direct roles in cellular metabolism. RNA is made by copying the base sequence of a section of double-stranded DNA, called a [gene](https://www.britannica.com/science/gene), into a piece of single-stranded nucleic acid. This process, called transcription

**The Concept of Information Flow**

The DNA macromolecule is an elegant structure that accommodates the need to store and transmit genetic information. A chromosomal DNA molecule is, in fact, two single chains of nucleotide subunits that are complementary to each other (Figure [**1**](https://onlinelibrary.wiley.com/doi/full/10.1038/npg.els.0000812#a0000812-fig-0001)a). Thus there is present both the primary information, and the template necessary to make a new copy of that information within this chemically inert macromolecule. The redundancy helps to ensure that essential information is not lost, because if one copy is damaged, the other can serve as a template for repair. The dangers inherent in transmitting genetic information are also reduced. If only a single chain were present, essential information might be lost either when the necessary template was assembled, or subsequently when the copies were produced.

 

 Information transfer between DNA, RNA and protein macromolecules. (a) Complementary base pairing in a short segment of DNA. The four types of nucleotides are represented by the bars (solid or open, long or short), and each nucleotide can pair with only one complement. (b) Information transfer between DNA (drawn as in part a) and RNA (green). RNA can also be used to specify DNA by an analogous mechanism. (c) Information transfer from RNA to protein. The 20 types of amino acids (the four oval shapes represent four of these) are linked according to the sequence of nucleotides in the messenger RNA, which in turn was copied from the gene in part b. The next amino acid which will be added is shown with its adaptor transfer RNA (tRNA) attached. (d) The central dogma of molecular biology specifies the forbidden information transfers.

The central dogma of molecular biology predicts that a particular sequence of amino acids (a protein) cannot be used to specify or even alter a particular sequence of nucleotides (a gene). Instead, information flows from nucleic acids to proteins, in that an elaborate machinery exists to ‘translate’ the nucleic acid ‘alphabet’ to the amino acid ‘alphabet’ according to the rules of the genetic code. Cells exhibit no trace of a ‘back-translation’ machinery, and organisms can transmit only their genes to their offspring. Even though the genetic material is not entirely constant, advantageous mutations do not arise in a directed manner. The predictions of the central dogma have withstood every challenge, and are likely to remain as the central organizing principles of molecular biology.

Modern molecular medicine encompasses the utilization of many molecular biological techniques in the analysis of disease, disease genes and disease gene function. The study of disease genes and their function in an unaffected individual has been possible by the development of recombinant DNA and cloning techniques. The basis of the term recombinant DNA refers to the recombining of different segments of DNA. Cloning refers to the process of preparing multiple copies of an individual type of recombinant DNA molecule. The classical mechanisms for producing recombinant molecules involves the insertion of exogenous fragments of DNA into either bacterially derived plasmid (circular double stranded autonomously replicating DNAs found in bacteria) vectors or bacteriophage (viruses that infect bacteria) based vectors.

**Introduction to molecular biology. Part II**

Replication of DNA. Hypothetical DNA replication mechanisms: conservative, semi-conservative, dispersive. Enzymology of replication. Molecular bases of DNA biosynthesis.

*Maximal point: 7*

1. Describe the three hypotheses of DNA replication.

2. Describe the Meselson-Stahl experiment and explain its significance.

3. Explain the molecular mechanism of semiconservative DNA replication.

4. Explain the role of main enzymes implicated in the replication process.

5. Explain proofreading mechanisms and error correction during DNA replication.

in 1953, after the double helix structure of DNA has just been discovered. One big question concerned DNA replication. The structure of the DNA double helix provided a tantalizing hint about how copying might take place^{1,2}1,2start superscript, 1, comma, 2, end superscript. It seemed likely that the two complementary strands of the helix might separate during replication, each serving as a template for the construction of a new, matching strand. There were three basic models for DNA replication that had been proposed by the scientific community after the discovery of DNA's structure

* **Semi-conservative replication.** In this model, the two strands of DNA unwind from each other, and each acts as a template for synthesis of a new, complementary strand. This results in two DNA molecules with one original strand and one new strand.
* **Conservative replication.** In this model, DNA replication results in one molecule that consists of both original DNA strands (identical to the original DNA molecule) and another molecule that consists of two new strands (with exactly the same sequences as the original molecule).
* **Dispersive replication.** In the dispersive model, DNA replication results in two DNA molecules that are mixtures, or “hybrids,” of parental and daughter DNA. In this model, each individual strand is a patchwork of original and new DNA.

The evidence that DNA replication was semi-conservative came from an elegant experiment completed by Matthew Meselson and Franklin Stahl. They labelled the parental DNA with a heavy isotope of nitrogen (15N) by growing bacteria in a growth medium that contained 15NH4Cl. They then grew the bacteria, in a medium that contained 14NH4Cl, in conditions such that any newly synthesised DNA would contain 14N. Since DNA replication is semi-conservative, after one round of DNA replication, each cell would have a DNA molecule that contains one ‘old’ parental strand labelled with 15N and one ‘new’ daughter strand labelled with 14N. This was shown by analysing the density of the DNA using density-gradient centrifugation. As predicted, they observed that the new daughter DNA molecule had a density consistent with the fact that it contained both 15N and 14N and that this daughter DNA contained one strand with 15N and another strand with 14N.

DNA replication is **semiconservative**, meaning that each strand in the DNA double helix acts as a template for the synthesis of a new, complementary strand.

 

DNA replication can be thought of in three stages; **Initiation, Elongation, Termination**

### Initiation

DNA synthesis is initiated at particular points within the DNA strand known as ‘**origins**’, which are specific coding regions. These origins are targeted by initiator proteins, which go on to recruit more proteins that help aid the replication process, forming a replication complex around the DNA origin. There are multiple origin sites, and when replication of DNA begins, these sites are referred to as **replication forks**.

Within the replication complex is the enzyme **DNA Helicase**, which unwinds the double helix and exposes each of the two strands, so that they can be used as a template for replication. It does this by hydrolysing the ATP used to form the bonds between the nucleobases, therefore breaking the bond holding the two strands together.

DNA Primase is another enzyme that is important in DNA replication. It synthesises a small **RNA primer**, which acts as a ‘kick-starter’ for **DNA Polymerase.** DNA Polymerase is the enzyme that is ultimately responsible for the creation and expansion of the new strands of DNA.

### Elongation

Once the DNA Polymerase has attached to the original, unzipped two strands of DNA (i.e. the **template** strands), it is able to start synthesising the new DNA to match the templates. It is essential to note that DNA polymerase is only able to extend the primer by adding free nucleotides to the **3’ end.**

One of the templates is read in a 3’ to 5’ direction, which means that the new strand will be formed in a 5’ to 3’ direction. This newly formed strand is referred to as the **Leading Strand.** Along this strand, DNA Primase only needs to synthesise an **RNA primer** once, at the beginning, to initiate DNA Polymerase. This is because DNA Polymerase is able to extend the new DNA strand by reading the template 3′ to 5′, synthesising in a 5′ to 3′ direction as noted above.

However, the other template strand (the **lagging strand**) is antiparallel, and is therefore read in a **5’ to 3’** direction. Continuous DNA synthesis, as in the **leading strand**, would need to be in the 3′ to 5′ direction, which is impossible as we cannot add bases to the 5′ end. Instead, as the helix unwinds, RNA primers are added to the newly exposed bases on the **lagging strand** and DNA synthesis occurs **in fragments,**but still in the 5′ to 3′ direction as before. These fragments are known as **Okazaki fragments.**

### Termination

The process of expanding the new DNA strands continues until there is either no more DNA template left to replicate (i.e. at the end of the chromosome), or two replication forks meet and subsequently **terminate.** The meeting of two replication forks is not regulated and happens randomly along the course of the chromosome.

Once DNA synthesis has finished, it is important that the newly synthesised strands are bound and stabilized.  With regards to the lagging strand, two enzymes are needed to achieve this; **RNAase H** removes the RNA primer that is at the beginning of each Okazaki fragment, and **DNA Ligase** joins fragments together to create one complete strand.

 **Transcription of genetic information**

Structure of gene: promoter, exons, introns, terminator. Enzymology of transcription. Mechanism of gene transcription: initiation, elongation, termination. Post-transcriptional maturation of mRNA: 3 'polyadenylation, 5 'capping, cutting out introns.

1. Define the terms: transcription, promoter, enhancer, terminator.

2. Describe prokaryotic and eukaryotic RNA-polymerases' structure and functions.

3. Describe phases of transcription, explain the processes happening at each phase and their importance.

4. Explain the process, importance and difference of Rho-independent and Rho-dependent termination of transcription.

5. Explain the mechanism of polyadenylation, its importance.

6. Describe the structure of the cap fragment, its synthesis and functions.

7. Describe the splicing mechanism and its meaning.

8. Explain the effect of splicing on gene expression.

**Transcription**, the synthesis of [RNA](https://www.britannica.com/science/RNA) from [DNA](https://www.britannica.com/science/DNA). Genetic information flows from DNA into [protein](https://www.britannica.com/science/protein), the substance that gives an organism its form. This flow of information occurs through the sequential processes of transcription (DNA to RNA) and [translation](https://www.britannica.com/science/translation-genetics) (RNA to protein). Transcription occurs when there is a need for a particular gene product at a specific time or in a specific tissue.



During transcription, only one strand of DNA is usually copied. This is called the template strand, and the RNA molecules produced are single-stranded [messenger RNAs](https://www.britannica.com/science/messenger-RNA) (mRNAs). The DNA strand that would correspond to the [mRNA](https://www.britannica.com/science/messenger-RNA) is called the coding or sense strand. In [eukaryotes](https://www.britannica.com/science/eukaryote) (organisms that possess a [nucleus](https://www.britannica.com/science/nucleus-biology)) the initial product of transcription is called a pre-mRNA. Pre-mRNA is extensively edited through splicing before the mature mRNA is produced and ready for translation by the [ribosome](https://www.britannica.com/science/ribosome), the cellular [organelle](https://www.britannica.com/science/organelle) that serves as the site of protein synthesis. Transcription of any one [gene](https://www.britannica.com/science/gene) takes place at the chromosomal location of that gene, which is a relatively short segment of the [chromosome](https://www.britannica.com/science/chromosome). The active transcription of a gene depends on the need for the activity of that particular gene in a specific [cell](https://www.britannica.com/science/cell-biology) or [tissue](https://www.britannica.com/science/tissue) or at a given time.

Small segments of DNA are transcribed into RNA by the [enzyme](https://www.britannica.com/science/enzyme) [RNA polymerase](https://www.britannica.com/science/RNA-polymerase), which achieves this copying in a strictly controlled process. The first step is to recognize a specific sequence on DNA called a promoter that signifies the start of the gene. The two strands of DNA become separated at this point, and RNA polymerase begins copying from a specific point on one strand of the DNA using a special type of [sugar](https://www.britannica.com/science/sugar-chemical-compound)-containing [nucleoside](https://www.britannica.com/science/nucleoside) called ribonucleoside 5’-triphosphate to begin the growing chain. Additional ribonucleoside triphosphates are used as the substrate, and, by cleavage of their high-energy [phosphate](https://www.britannica.com/science/phosphate) bond, ribonucleoside monophosphates are incorporated into the growing RNA chain. Each successive ribonucleotide is directed by the complementary base pairing rules of DNA. For example, a C ([cytosine](https://www.britannica.com/science/cytosine)) in DNA directs the incorporation of a G ([guanine](https://www.britannica.com/science/guanine)) into RNA. Likewise, a G in DNA is copied into a C in RNA, a T ([thymine](https://www.britannica.com/science/thymine)) into an A ([adenine](https://www.britannica.com/science/adenine)), and an A into a U ([uracil](https://www.britannica.com/science/uracil); RNA contains U in place of the T of DNA). Synthesis continues until a termination signal is reached, at which point the RNA polymerase drops off the DNA, and the RNA molecule is released.

Ahead of many genes in [prokaryotes](https://www.britannica.com/science/prokaryote) (organisms that lack a nucleus), there are signals called “operators” (see [operons](https://www.britannica.com/science/operon)) where specialized proteins called repressors bind to the DNA just upstream of the start point of transcription and prevent access to the DNA by RNA polymerase. These repressor proteins thus prevent transcription of the gene by physically blocking the action of the RNA polymerase. Typically, repressors are released from their blocking action when they receive signals from other molecules in the cell indicating that the gene needs to be expressed. Ahead of some prokaryotic genes are signals to which activator proteins bind to stimulate transcription.



Transcription in eukaryotes is more complicated than in prokaryotes. First, the RNA polymerase of higher organisms is a more complicated enzyme than the relatively simple five-subunit enzyme of prokaryotes. In addition, there are many more accessory factors that help to control the [efficiency](https://www.merriam-webster.com/dictionary/efficiency) of the individual promoters. These accessory proteins are called [transcription factors](https://www.britannica.com/science/transcription-factor) and typically respond to signals from within the cell that indicate whether transcription is required. In many human genes, several transcription factors may be needed before transcription can proceed efficiently. A [transcription factor](https://www.britannica.com/science/transcription-factor) can cause either repression or activation of gene expression in eukaryotes.

**Translation of genetic information**

Structure of ribosome: rRNA and ribosomal proteins. Genetic code: properties and key experiments. tRNA, aminoacyl - tRNA synthetase. Mechanism of translation: initiation, elongation, termination.

1. Explain the ribosome cycle and fidelity of translation.

2. Define the genetic code, tRNA, mRNA, codon, anticodon.

3. Describe the structure of tRNA and the mechanism of its charging.

4. Explain the scanning model of translation.

5. Explain the mechanism of translation and its phases.

6. Describe the structure of ribosomes and polysomes.

Translation, the synthesis of [protein](https://www.britannica.com/science/protein) from [RNA](https://www.britannica.com/science/RNA). Hereditary information is contained in the [nucleotide](https://www.britannica.com/science/nucleotide) sequence of [DNA](https://www.britannica.com/science/DNA) in a code. The coded information from DNA is copied faithfully during [transcription](https://www.britannica.com/science/transcription-genetics) into a form of RNA known as [messenger RNA](https://www.britannica.com/science/messenger-RNA) (mRNA), which is then translated into chains of [amino acids](https://www.britannica.com/science/amino-acid). Amino acid chains are folded into helices, zigzags, and other shapes to form proteins and are sometimes associated with other [amino acid](https://www.britannica.com/science/amino-acid) chains.

The specific amounts of amino acids in a protein and their sequence determine the protein’s unique properties; for example, [muscle](https://www.britannica.com/science/muscle) protein and [hair](https://www.britannica.com/science/hair-anatomy) protein contain the same 20 amino acids, but the sequences of these amino acids in the two proteins are quite different. If the nucleotide sequence of mRNA is thought of as a written message, it can be said that this message is read by the translation apparatus in “words” of three nucleotides, starting at one end of the mRNA and proceeding along the length of the molecule. These three-letter words are called [codons](https://www.britannica.com/science/codon). Each codon stands for a specific amino acid, so if the message in mRNA is 900 nucleotides long, which corresponds to 300 codons, it will be translated into a chain of 300 amino acids.

Translation takes place on [ribosomes](https://www.britannica.com/science/ribosome)—complex particles in the [cell](https://www.britannica.com/science/cell-biology) that contain RNA and protein. In [prokaryotes](https://www.britannica.com/science/prokaryote) (organisms that lack a [nucleus](https://www.britannica.com/science/nucleus-biology)) the ribosomes are loaded onto the mRNA while transcription is still ongoing. The mRNA sequence is read three bases at a time from its 5’ end toward its 3’ end, and one amino acid is added to the growing chain from its respective [transfer RNA](https://www.britannica.com/science/transfer-RNA) (tRNA), until the complete protein chain is assembled. Translation stops when the [ribosome](https://www.britannica.com/science/ribosome) encounters a termination codon, normally UAG, UAA, or UGA (where U, A, and G represent the RNA bases [uracil](https://www.britannica.com/science/uracil), [adenine](https://www.britannica.com/science/adenine), and [guanine](https://www.britannica.com/science/guanine), respectively). Special release factors associate with the ribosome in response to these codons, and the newly synthesized protein, tRNAs, and mRNA all dissociate. The ribosome then becomes available to interact with another mRNA molecule.

Any one mRNA is translated by several ribosomes along its length, each one at a different stage of translation. In [eukaryotes](https://www.britannica.com/science/eukaryote) (organisms that possess a nucleus) ribosomes that produce proteins to be used in the same cell are not associated with membranes. However, proteins that must be exported to another location in the organism are synthesized on ribosomes located on the outside of flattened membranous chambers called the [endoplasmic reticulum](https://www.britannica.com/science/endoplasmic-reticulum) (ER). A completed amino acid chain is extruded into the inner cavity of the ER. Subsequently, the ER transports the proteins via small vesicles to another cytoplasmic [organelle](https://www.britannica.com/science/organelle) called the [Golgi apparatus](https://www.britannica.com/science/Golgi-apparatus), which in turn buds off more vesicles that eventually fuse with the [cell membrane](https://www.britannica.com/science/cell-biology/Intercellular-communication#ref37365). The protein is then released from the cell.

Translation has pretty much the same three parts, but they have fancier names: initiation, elongation, and termination.

Initiation ("beginning"): in this stage, the ribosome gets together with the mRNA and the first tRNA so translation can begin.

Elongation ("middle"): in this stage, amino acids are brought to the ribosome by tRNAs and linked together to form a chain.

Termination ("end"): in the last stage, the finished polypeptide is released to go and do its job in the cell.

The key components required for translation are mRNA, ribosomes, and transfer RNA (tRNA).

During translation, mRNA nucleotide bases are read as codons of three bases. Each ‘codon’ codes for a particular amino acid. Every tRNA molecule possesses an anticodon that is complementary to the mRNA codon, and at the opposite end lies the attached amino acid. tRNA molecules are therefore responsible for bringing amino acids to the ribosome in the correct order ready for polypeptide assembly

It is important to know that a single amino acid may be coded for by more than one codon. There are also specific codons that signal the start and the end of translation.

Aminoacyl-tRNA synthetases are enzymes that link amino acids to their corresponding tRNA molecules. The resulting complex is charged and is referred to as an aminoacyl-tRNA.

Initiation

For translation to begin, the start codon 5’AUG must be recognised. This is a codon specific to the amino acid methionine, which is nearly always the first amino acid in a polypeptide chain.

At the 5’ cap of mRNA, the small 40s subunit of the ribosome binds. Subsequently, the larger 60s subunit binds to complete the initiation complex. The next step (elongation) can now commence.

**Elongation**

The ribosome has two tRNA binding sites; the P site which holds the peptide chain and the A site which accepts the tRNA.

While Methionine-tRNA occupies the P site, the aminoacyl-tRNA that is complementary to the next codon binds to the A site, using energy yielded from the hydrolysis of GTP.

Methionine moves from the P site to the A site to bond to new amino acid there, and so the growth of the peptide has begun. The tRNA molecule in the P site no longer has an attached amino acid, and so leaves the ribosome.

The ribosome then translocates along the mRNA molecule to the next codon again using energy yielded from the hydrolysis of GTP. Now, the growing peptide lies at the P site and the A site is open for the binding of the next aminoacyl-tRNA, and the cycle continues. The polypeptide chain is built up in the direction from the N terminal (methionine) to the C terminal (the final amino acid).

**Termination**

One of the three stop codons enters the A site. No tRNA molecules bind to these codons so the peptide and tRNA in the P site become hydrolysed releasing the polypeptide into the cytoplasm.

The small and large subunits of the ribosome dissociate ready for the next round of translation.

**Post-translational protein modifications and folding.**

Posttranslational modification of protein. Folding of protein: chaperones

1. Draw a functional connection between primary structure and higher-order spatial organization of polypeptides.

2. Explain the auxiliary role of chaperones in protein folding.

3. Give detailed examples of human disorders linked with protein misfolding.

Within the last few decades, scientists have discovered that the human proteome is vastly more complex than the human genome. While it is estimated that the human genome comprises between 20,000 and 25,000 genes, the total number of proteins in the human proteome is estimated at over 1 million. These estimations demonstrate that single genes encode multiple proteins. Genomic recombination, transcription initiation at alternative promoters, differential transcription termination, and alternative splicing of the transcript are mechanisms that generate different mRNA transcripts from a single gene.

The increase in complexity from the level of the genome to the proteome is further facilitated by protein post-translational modifications (PTMs). PTMs are chemical modifications that play a key role in functional proteomic because they regulate activity, localization, and interaction with other cellular molecules such as proteins, nucleic acids, lipids and cofactors.

**Types of post-translational modification**

There are many typs of protein modification, which are  mostly catalyzed by enzymes that recognize specific target sequences in proteins. These modifications regulate protein folding by targeting specific subcellular compartments, interacting with ligands or other proteins, or by bringing about a change in their functional state including catalytic activity or signaling. The most common PTMs  are:

***Based on the addition of chemical groups***

* Phosphorylation
* [Acetylation](https://www.news-medical.net/health/What-is-Acetylation.aspx)
* Hydroxylation
* Methylation

***Based on the addition of complex groups***

* Glycosylation
* AMPylation
* Lipidation

***Based on the addition of polypeptides***

* Ubiquitination

***Based on the cleavage of proteins***

* Proteolysis

***Based on the amino acid modification***

* Deamidation

**Chemical groups**

***Phosphorylation***

Reversible phosphorylation of proteins involves addition of a phosphate group on serine, threonine, or tyrosine residues and is one of the important and extensively studied PTM in both prokaryotes and eukaryotes.

Several enzymes or signaling proteins are switched ‘on’ or ‘off’ by phosphorylation or dephosphorylation. Phosphorylation is performed by enzymes called ‘kinases’, while dephosphorylation is performed by ‘phosphatases’.

Addition of a phosphate group can convert a previously uncharged pocket of protein into a negatively charged and hydrophilic protein thereby inducing conformational changes in the protein.

Phosphorylation has implications in several cellular processes, including cell cycle, growth, apoptosis and signal transduction pathways. One example is the activation of p53, a tumor suppressor protein. p53 is used in cancer therapeutics and is activated by phosphorylation of its N-terminal by several kinases.

**Acetylation**

Acetylation refers to addition of acetyl group in a protein. It is involved in several biological functions, including protein stability, location, synthesis; apoptosis; cancer; DNA stability. Acetylation and deacetylation of histone form a critical part of gene regulation.

Acetylation of histones reduces the positive charge on histone, reducing its interaction with the negatively charged phosphate groups of DNA, making it less tightly wound to DNA and accessible to gene transcription. Acetylation of p53, a tumor suppressor gene, is crucial for its growth suppressing properties.

**Hydroxylation**

This process adds a hydroxyl group (-OH) to the proteins. It is catalyzed by enzymes termed as ‘hydroxylases’ and aids in converting hydrophobic or lipophilic compounds into hydrophilic compounds.

**Methylation**

Methylation refers to addition of a methyl group to lysine or arginine residue of a protein. Arginine can be methylated once or twice, while lysine can be methylated once, twice, or thrice. Methylation is achieved by enzymes called methyltransferases. Methylation has been widely studied in histones wherein histone methylation can lead to gene activation or repression based on the residue that is methylated.

## Complex groups

**Glycosylation**

Glycosylation involves addition of an oligosaccharide termed ‘glycan’ to either a nitrogen atom (N-linked glycosylation) or an oxygen atom (O-linked glycosylation). N-linked glycosylation occurs in the amide nitrogen of asparagine, while the O-linked glycosylation occurs on the oxygen atom of serine or threonine.

Carbohydrates present in the form of N-linked or O-linked oligosaccharides are present on the surface of cells and secrete proteins. They have critical roles in protein sorting, immune recognition, receptor binding, inflammation, and pathogenicity. For example, N-linked glycans on an immune cell can dictate how it migrates to specific sites. Similarly, it can also determine how a cell recognizes ‘self’ and ‘non-self’.

**AMPylation**

AMPylation refers to reversible addition of AMP to a protein. It involves formation of a phosphodiester bond between the hydroxyl group of the protein and the phosphate group of AMP.

**Lipidation**

The covalent binding of a lipid group to a protein is called lipidation. Lipidation can be further subdivided into prenylation, N-myristoylation, palmitoylation, and glycosylphosphatidylinositol (GPI)-anchor addition.

Prenylation involves the addition of isoprenoid moiety to a cysteine residue of a substrate protein. It is critical in controlling the localization and activity of several proteins that have crucial functions in biological regulation.

Myristoylation involves the addition of myristoyl group to a [glycine](https://www.news-medical.net/health/What-is-Glycine.aspx) residue by an amide bond. It has functions in membrane association and apoptosis. In palmitoylation, a palmitoyl group is added to a cysteine residue of a protein.

In GPI-anchor addition, the carboxyl-terminal signal peptide of the protein is split and replaced by a GPI anchor. Recent research in human genetics has revealed that GPI anchors are important for human health. Any defects in the assembling, attachment or remodeling of GPI anchors lead to genetic diseases known as inherited GPI deficiency.

## Polypeptides

**Ubiquitination**

Ubiquitination involves addition of a protein found ubiquitously, termed ‘ubiquitin’, to the lysine residue of a substrate. Either a single ubiquitin molecule (monoubiquitination) or a chain of several ubiquitin molecules may be attached (polyubiquitination).

Polyubiquitinated proteins are recognized by the 26S proteasome and are subsequently targeted for proteolysis or degradation. Monoubiquitinated proteins may influence cell tracking and endocytosis.

## Protein cleavage

**Proteolysis**

Proteolysis refers to breakdown of proteins into smaller polypeptides or amino acids. For example, removal of N-terminal methionine, a signal peptide, after translation leads to conversion of an inactive or non-functional protein to an active one.

## Amino acid modification

**Deamidation**

Deamidation is the removal or conversion of asparagine or glutamine residue to another functional group. Asparagine is converted to [aspartic acid](https://www.news-medical.net/health/What-is-Aspartic-Acid.aspx) or isoaspartic acid, while glutamine is converted to glutamic acid or pyroglutamic acid. This modification can change the protein structure, stability, and function.

Molecular chaperones are present in all organisms and are essential for cell survival (**Figure 1**). One of the major functions of molecular chaperones is to facilitate [protein folding](https://www.sciencedirect.com/topics/medicine-and-dentistry/protein-folding%22%20%5Co%20%22Learn%20more%20about%20Protein%20Folding%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages). Although the [amino acid sequence](https://www.sciencedirect.com/topics/medicine-and-dentistry/peptide-sequence%22%20%5Co%20%22Learn%20more%20about%20Peptide%20Sequence%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages) of a protein contains the information required to adopt the native conformation, not all proteins can fold spontaneously. Unfolded [polypeptides](https://www.sciencedirect.com/topics/medicine-and-dentistry/polypeptide%22%20%5Co%20%22Learn%20more%20about%20Polypeptide%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages) are generated during normal growth as the product of [protein synthesis](https://www.sciencedirect.com/topics/medicine-and-dentistry/protein-synthesis%22%20%5Co%20%22Learn%20more%20about%20Protein%20Synthesis%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages), but misfolded proteins arise as a consequence of cellular stresses, such as heat shock, [oxidative stress](https://www.sciencedirect.com/topics/medicine-and-dentistry/oxidative-stress%22%20%5Co%20%22Learn%20more%20about%20Oxidative%20Stress%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages), as well as pathological conditions. Molecular chaperones, including Hsp60s, [Hsp70s](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/hsp70%22%20%5Co%20%22Learn%20more%20about%20Hsp70%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages), Hsp90s and sHsps, assist in the folding of unfolded and misfolded polypeptides by stabilization of folding intermediates and prevention of protein misfolding and aggregation. Several chaperones also function to reactivate aggregated proteins. For example both Clps and Hsp90s work with Hsp70s to salvage aggregated proteins. Additionally, some chaperones, such as Clps, interact with specific [proteases](https://www.sciencedirect.com/topics/medicine-and-dentistry/proteinase%22%20%5Co%20%22Learn%20more%20about%20Proteinase%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages) and deliver unfolded and non-native proteins to compartmentalized proteases for degradation. There are also dedicated chaperones that act on one or a small number of proteins. The combined action of molecular chaperones therefore increases the cellular pool of native proteins while minimizing inactive proteins and potentially harmful protein aggregates. Since molecular chaperones have important roles in protein folding and remodeling, modulation of chaperone activity is linked to numerous diseases including cancer and amyloid disorders.

**Regulation of gene expression in prokaryotes and eukaryotes**

Gene structure in prokaryotes. Bacterial operons: lac, ara, trp, gal. Gene structure in eukaryotes. Regulation of transcription: transcription factors.

1. Define the terms: operon, cistron, promoter.

2. Explain the functioning and regulation of the following operons: lac, ara, trp, gal.

3. Explain positive and negative controls of operons.

4. Differentiate between constitutive and inducible promoters.

5. Explain the mechanism of transcription regulation in eukaryotes.

6. Describe the structure of the promoter: TATA-box, GC-box.

7. Explain the functions of enhancers and silencers.

8. Describe the role of transcription factors and activators in the regulation of transcription

9. Describe the structure and significance of DNA-binding domains and transcription activation domains.

10. Compare translation regulation in pro- and eukaryotes

Genes determine everything about us, from the outward physical traits we can see to the behind the scenes structures inside our cells that allow them to carry out all of our body functions. Genes are transcribed into copies of themselves called RNA, which is then translated to protein. **Structural genes** are genes that code for proteins in the body needed for structure or function. Unlike structural genes, **regulatory genes** code for protein products that control other genes, instead of making structures of their own. Regulatory genes code for proteins that act like switches, turning other genes on or off. These genes are essential to controlling cell function, and without them, cells can grow out of control, causing diseases, like cancer, in the body.

Gene expression is the phenotypic manifestation of genes by the processes of transcription and translation. Gene expression via transcription and translation is a fundamental principle of molecular biology that is often referred to as the central dogma of molecular biology. Gene expression is the process by which the information encoded in a gene is used to direct the assembly of a protein molecule. The cell reads the sequence of the gene in groups of three bases. Each group of three bases (codon) corresponds to one of 20 different amino acids used to build the protein.

Gene expression in humans is complex and highly regulated. Regulation occurs at many points during the transcription and translation processes and involves epigenomic compounds, which are chemical compounds and proteins that can attach to DNA and influence gene expression.

The number of genes in an organism’s genome (the entire set of chromosomes) varies significantly between species. For example, whereas the genome of the bacterium Escherichia coli O157:H7 houses precisely 5,416 genes. Arabidopsis thaliana—the first [plant](https://www.britannica.com/plant/plant) for which a complete genomic sequence was recovered—has roughly 25,500 genes; its genome is one of the smallest known to plants. Among [extant](https://www.merriam-webster.com/dictionary/extant) independently replicating organisms, the bacterium Mycoplasma genitalium has the fewest number of genes, just 517. In the human genome, there are a little less than 20,000 genes. In some cells, many genes are active--say, 10,000--and the other 10,000 would be inactive. In other kinds of cells, maybe the other 10,000 would be active and the first 10,000 would be inactive.

And so, gene regulation is the process by which the cell determines which genes will be active and which genes will not be active. And gene regulation is at the bottom of what makes a cell decide to become a red blood cell, or a neuron, or a hepatocyte in the liver, or a muscle cell. So different gene regulation will give you a different program of genes and different genes expressed.

There are several different kinds of gene regulation. Some genes, called housekeeping genes, are expressed in almost every cell. And these require a regulatory network or machinery that keeps them on in almost every cell, so these are the enzymes that help make DNA, and perform glycolysis, and burn sugar, and things like that. There are other genes that are called tissue-specific genes. These are genes that, say, would only be expressed in a red blood cell or a neuron. Very often, these genes have transcription factors, which are proteins that bind to DNA, near these genes. And those transcription factors actually help the RNA machinery get there and transcribe that gene in those cells, and those tissues, transcription factors, rather, are expressed specifically in those tissues. There are also factors expressed in those tissues that will be suppressors that can turn a gene off. And then there are genes that are regulated during development. Sometimes they're expressed in fetal life and then turned off in adults, and sometimes it's vice versa. So there are very complex different ways that genes are regulated.

Gene regulation is the process of turning genes on or off. Gene regulation can occur at any point of the transcription-translation process but most often occurs at the transcription level.

Proteins that can be activated by other cells and signals from the environment are called transcription factors. Transcription factors bind to regulatory regions of the gene and increase or decrease the level of transcription. Other mechanisms of gene regulation include regulating the processing of RNA, the stability of mRNA and the rate of translation.

Turning the correct genes on and off is an essential component to maintaining a cell’s functionality.

**Gene regulation** is how a cell controls which genes, out of the many genes in its genome, are "turned on" (expressed). Thanks to gene regulation, each cell type in your body has a different set of active genes – despite the fact that almost all the cells of your body contain the exact same DNA. These different patterns of gene expression cause your various cell types to have different sets of proteins, making each cell type uniquely specialized to do its job.

Differences in gene regulation makes the different cell types in a multicellular organism (such as yourself) unique in structure and function. Gene regulation can also help us explain some of the differences in form and function between different species with relatively similar gene sequences.

For instance, humans and chimpanzees have genomes that are about 98%, percent identical at the DNA level. The protein-coding sequences of some genes are different between humans and chimpanzees, contributing to the differences between the species. However, researchers also think that changes in gene regulation play a major role in making humans and chimps different from one another. For instance, some DNA regions that are present in the chimpanzee genome but missing in the human genome contain known gene-regulatory sequences that control when, where, or how strongly a gene is expressed.

Experiments have shown that many of the genes within the cells of organisms are inactive much or even all of the time. Thus, at any time, in both eukaryotes and prokaryotes, it seems that a gene can be switched on or off. The regulation of genes between eukaryotes and prokaryotes differs in important ways.

Gene expression can be regulated at any step: from [transcriptional initiation](https://en.wikipedia.org/wiki/Transcriptional_regulation), to [RNA processing](https://en.wikipedia.org/wiki/RNA_processing), to [post-translational modification](https://en.wikipedia.org/wiki/Post-translational_modification) of the protein. The regulation of [lactose](https://en.wikipedia.org/wiki/Lactose) metabolism genes in [*E. coli*](https://en.wikipedia.org/wiki/E._coli) ([*lac* operon](https://en.wikipedia.org/wiki/Lac_operon)) was the first such mechanism to be described in 1961.

Regulation of gene expression in prokaryotes

We tend to think of bacteria as simple. But even the simplest bacterium has a complex task when it comes to gene regulation! The bacteria in your gut or between your teeth have genomes that contain thousands of different genes. Most of these genes encode proteins, each with its own role in a process such as fuel metabolism, maintenance of cell structure, and defense against viruses.

Some of these proteins are needed routinely, while others are needed only under certain circumstances. Thus, cells don't express all the genes in their genome all the time. You can think of the genome as being like a cookbook with many different recipes in it. The cell will only use the recipes (express the genes) that fit its current needs.

There are various forms of **gene regulation**, that is, mechanisms for controlling which genes get expressed and at what levels. However, a lot of gene regulation occurs at the level of transcription.

Bacteria have specific regulatory molecules that control whether a particular gene will be transcribed into mRNA. Often, these molecules act by binding to DNA near the gene and helping or blocking the transcription enzyme, RNA polymerase. Let's take a closer look at how genes are regulated in bacteria.

In bacteria, related genes are often found in a cluster on the chromosome, where they are transcribed from one **promoter** (RNA polymerase binding site) as a single unit. Such a cluster of genes under control of a single promoter is known as an **operon**. Operons are common in bacteria, but they are rare in eukaryotes such as humans.

A bacterial cell with has a circular bacterial chromosome A small segment of the chromosome is an operon. The DNA of the operon contains three genes, Gene 1, Gene 2, and Gene 3, which are found in a row in the DNA. They are under control of a single promoter (site where RNA polymerase binds) and they are transcribed together to make a single mRNA that has contains sequences coding for all three genes. When the mRNA is translated, the three different coding sequences of the mRNA are read separately, making three different proteins (Protein 1, Protein 2, and Protein 3). The operon does not consist of just the three genes. Instead, it also includes the promoter and other regulatory sequences that regulate expression of the genes.

In general, an operon will contain genes that function in the same process. For instance, a well-studied operon called the [*lac* operon](https://www.khanacademy.org/science/ap-biology/gene-expression-and-regulation/regulation-of-gene-expression-and-cell-specialization/a/overview-gene-regulation-in-bacteria) contains genes that encode proteins involved in uptake and metabolism of a particular sugar, lactose. Operons allow the cell to efficiently express sets of genes whose products are needed at the same time.

**Anatomy of an operon**

Operons aren't just made up of the coding sequences of genes. Instead, they also contain **regulatory DNA sequences** that control transcription of the operon. Typically, these sequences are binding sites for **regulatory proteins**, which control how much the operon is transcribed. The promoter, or site where RNA polymerase binds, is one example of a regulatory DNA sequence.

 The promoter is found in the DNA of the operon, upstream of (before) the genes. When the RNA polymerase binds to the promoter, it transcribes the operon and makes some mRNAs.

Most operons have other regulatory DNA sequences in addition to the promoter. These sequences are binding sites for regulatory proteins that turn expression of the operon "up" or "down."

* Some regulatory proteins are **repressors** that bind to pieces of DNA called **operators**. When bound to its operator, a repressor reduces transcription (e.g., by blocking RNA polymerase from moving forward on the DNA).

 A repressor protein binds to a site called on the operator. In this case (and many other cases), the operator is a region of DNA that overlaps with or lies just downstream of the RNA polymerase binding site (promoter). That is, it is in between the promoter and the genes of the operon. When the repressor binds to the operator, it prevents RNA polymerase from binding to the promoter and/or transcribing the operon. When the repressor is bound to the operator, no transcription occurs and no mRNA is made.

* Some regulatory proteins are **activators**. When an activator is bound to its DNA binding site, it increases transcription of the operon (e.g., by helping RNA polymerase bind to the promoter).

The activator protein binds to a specific sequence of DNA, in this case immediately upstream of (before) the promoter where RNA polymerase binds. When the activator binds, it helps the polymerase attach to the promoter (makes promoter binding more energetically favorable). This causes the RNA polymerase to bind firmly to the promoter and transcribe the genes of the operon much more frequently, leading to the production of many molecules of mRNA.

Like any other protein produced in an organism, they are encoded by genes in the bacterium's genome. The genes that encode regulatory proteins are sometimes called **regulatory genes**.

Many regulatory proteins can themselves be turned "on" or "off" by specific small molecules. The small molecule binds to the protein, changing its shape and altering its ability to bind DNA. For instance, an activator may only become active (able to bind DNA) when it's attached to a certain small molecule.

When the small molecule is absent, the activator is "off" - it takes on a shape that makes it unable to bind DNA. When the small molecule that activates the activator is added, it binds to the activator and changes its shape. This shape change makes the activator able to bind its target DNA sequence and activate transcription.

**Operons may be inducible or repressible**

Some operons are usually "off," but can be turned "on" by a small molecule. The molecule is called an **inducer**, and the operon is said to be **inducible**.

* For example, the [*lac* operon](https://www.khanacademy.org/science/biology/gene-regulation/gene-regulation-in-bacteria/a/the-lac-operon) is an inducible operon that encodes enzymes for metabolism of the sugar lactose. It turns on only when the sugar lactose is present (and other, preferred sugars are absent). The inducer in this case is allolactose, a modified form of lactose.

Other operons are usually "on," but can be turned "off" by a small molecule. The molecule is called a **corepressor**, and the operon is said to be **repressible**.

* For example, the *[trp](https://www.khanacademy.org/science/biology/gene-regulation/gene-regulation-in-bacteria/a/the-trp-operon%22%20%5Ct%20%22_blank)*[operon](https://www.khanacademy.org/science/biology/gene-regulation/gene-regulation-in-bacteria/a/the-trp-operon%22%20%5Ct%20%22_blank) is a repressible operon that encodes enzymes for synthesis of the amino acid tryptophan. This operon is expressed by default, but can be repressed when high levels of the amino acid tryptophan are present. The corepressor in this case is tryptophan.

These examples illustrate an important point: that gene regulation allows bacteria to respond to changes in their environment by altering gene expression (and thus, changing the set of proteins present in the cell).

**Some genes and operons are expressed all the time**

Many genes play specialized roles and are expressed only under certain conditions, as described above. However, there are also genes whose products are constantly needed by the cell to maintain essential functions. These **housekeeping genes** are constantly expressed under normal growth conditions ("constitutively active"). Housekeeping genes have promoters and other regulatory DNA sequences that ensure constant expression.

*E. coli* bacteria can break down lactose, but it's not their favorite fuel. If glucose is around, they would much rather use that. Glucose requires fewer steps and less energy to break down than lactose. However, if lactose is the only sugar available, the *E. coli* will go right ahead and use it as an energy source.

To use lactose, the bacteria must express the *lac* operon genes, which encode key enzymes for lactose uptake and metabolism. To be as efficient as possible, *E. coli* should express the *lac* operon only when two conditions are met:

* Lactose is available, **and**
* Glucose is not available

How are levels of lactose and glucose detected, and how do changes in levels affect *lac* operon transcription? Two regulatory proteins are involved:

* One, the *lac* repressor, acts as a lactose sensor.
* The other, catabolite activator protein (CAP), acts as a glucose sensor.

These proteins bind to the DNA of the *lac* operon and regulate its transcription based on lactose and glucose levels. Let's take a look at how this works.

**Structure of the *lac* operon**

The *lac* operon contains three genes: *lacZ*, *lacY*, and *lacA*. These genes are transcribed as a single mRNA, under control of one promoter.

Genes in the *lac* operon specify proteins that help the cell utilize lactose. *lacZ* encodes an enzyme that splits lactose into monosaccharides (single-unit sugars) that can be fed into glycolysis. Similarly, *lacY* encodes a membrane-embedded transporter that helps bring lactose into the cell.

In addition to the three genes, the *lac* operon also contains a number of regulatory DNA sequences. These are regions of DNA to which particular regulatory proteins can bind, controlling transcription of the operon.

 The DNA of the *lac* operon contains (in order from left to right): CAP binding site, promoter (RNA polymerase binding site), operator (which overlaps with promoter), *lacZ* gene, *lacY* gene, and *lacA* gene. The activator protein CAP, when bound to a molecule called cAMP , binds to the CAP binding site and promotes RNA polymerase binding to the promoter. The *lac* repressor protein binds to the operator and blocks RNA polymerase from binding to the promoter and transcribing the operon.

\_ The **promoter** is the binding site for RNA polymerase, the enzyme that performs transcription.

* The **operator** is a negative regulatory site bound by the *lac* repressor protein. The operator overlaps with the promoter, and when the *lac* repressor is bound, RNA polymerase cannot bind to the promoter and start transcription.
* The **CAP binding site** is a positive regulatory site that is bound by catabolite activator protein (CAP). When CAP is bound to this site, it promotes transcription by helping RNA polymerase bind to the promoter.

Let's take a closer look at the *lac* repressor and CAP and their roles in regulation of the *lac* operon.

**The *lac* repressor**

The *lac* repressor is a protein that represses (inhibits) transcription of the *lac* operon. It does this by binding to the operator, which partially overlaps with the promoter. When bound, the *lac* repressor gets in RNA polymerase's way and keeps it from transcribing the operon.

When lactose is not available, the *lac* repressor binds tightly to the operator, preventing transcription by RNA polymerase. However, when lactose is present, the *lac* repressor loses its ability to bind DNA. It floats off the operator, clearing the way for RNA polymerase to transcribe the operon.

Upper panel: No lactose. When lactose is absent, the *lac* repressor binds tightly to the operator. It gets in RNA polymerase's way, preventing transcription.

Lower panel: With lactose. Allolactose (rearranged lactose) binds to the *lac* repressor and makes it let go of the operator. RNA polymerase can now transcribe the operon.

This change in the *lac* repressor is caused by the small molecule **allolactose**, an isomer (rearranged version) of lactose. When lactose is available, some molecules will be converted to allolactose inside the cell. Allolactose binds to the *lac* repressor and makes it change shape so it can no longer bind DNA.

Allolactose is an example of an **inducer**, a small molecule that triggers expression of a gene or operon. The *lac* operon is considered an **inducible operon** because it is usually turned off (repressed), but can be turned on in the presence of the inducer allolactose.

**Catabolite activator protein (CAP)**

When lactose is present, the *lac* repressor loses its DNA-binding ability. This clears the way for RNA polymerase to bind to the promoter and transcribe the *lac* operon. That sounds like the end of the story, right?

Well...not quite. As it turns out, RNA polymerase alone does not bind very well to the *lac* operon promoter. It might make a few transcripts, but it won't do much more unless it gets extra help from **catabolite activator protein** (**CAP**). CAP binds to a region of DNA just before the *lac* operon promoter and helps RNA polymerase attach to the promoter, driving high levels of transcription.

When glucose levels are low, cAMP is produced. The cAMP attaches to CAP, allowing it to bind DNA. CAP helps RNA polymerase bind to the promoter, resulting in high levels of

When glucose levels are high, no cAMP is made. CAP cannot bind DNA without cAMP, so transcription occurs only at a low level.

CAP isn't always active (able to bind DNA). Instead, it's regulated by a small molecule called **cyclic AMP** (**cAMP**). cAMP is a "hunger signal" made by *E. coli* when glucose levels are low. cAMP binds to CAP, changing its shape and making it able to bind DNA and promote transcription. Without cAMP, CAP cannot bind DNA and is inactive.

CAP is only active when glucose levels are low (cAMP levels are high). Thus, the *lac* operon can only be transcribed at high levels when glucose is absent. This strategy ensures that bacteria only turn on the *lac* operon and start using lactose after they have used up all of the preferred energy source (glucose).

**So, when does the *lac* operon really turn on?**

The *lac* operon will be expressed at high levels if two conditions are met:

* *Glucose must be unavailable:* When glucose is unavailable, cAMP binds to CAP, making CAP able to bind DNA. Bound CAP helps RNA polymerase attach to the *lac* operon promoter.
* *Lactose must be available*: If lactose is available, the *lac* repressor will be released from the operator (by binding of allolactose). This allows RNA polymerase to move forward on the DNA and transcribe the operon.

These two events in combination – the binding of the activator and the release of the repressor – allow RNA polymerase to bind strongly to the promoter and give it a clear path for transcription. They lead to strong transcription of the *lac* operon and production of enzymes needed for lactose utilization.

**Mutations**

Mutations: gene, chromosomal, genomic. Hereditary diseases. The value of mutations for the evolution of living nature.

1. Explain what a mutation is and its importance for evolution of life.

2. Classify and characterize the main types of mutations.

3. Define the terms: deletion, insertion, inversion, duplication, translocation, and explain what type of mutation each term belongs to and why.

4. Give specific examples of hereditary diseases

All organisms differ from each other in a varying degree. These small differences constitute variation, which may be the result of genetic changes taking place during the formation of the gametes, or of the influence of the environment, or a combination of both. In some cases it is dif- ficult to determine what contribution is made by heredity and what is due to the environment, es- pecially if the differences are very small. In humans, factors such as colour of the skin, hair colour, weight, shape of head and facial features all show variation and we know that many of these are inherited characteristics. Some of these factors, for example weight, can be affected by the level of nutrition or exercise, which are both environmental influences.

From viewpoint of genetics we recognize two forms of variation: **inheritable and non-inheritable.**

During the process of replication, DNA is normally copied exactly so that the-genetic material remains the same from generation to generation. However, very occasionally, changes can occur so that an organism may inherit altered genetic material. Such inherited changes are known as **mutations**. Mutations give rise to all variation, but their survival in the genome is influenced by many factors including effects on reproductive fitness, human population history, chromosomal location and recombination rates.

Mutations are the *source* of [variation](https://www.ncbi.nlm.nih.gov/books/n/iga/A4529/def-item/A5507/), but the *process* of [mutation](https://www.ncbi.nlm.nih.gov/books/n/iga/A4529/def-item/A5117/) does not itself drive evolution. The rate of change in [gene frequency](https://www.ncbi.nlm.nih.gov/books/n/iga/A4529/def-item/A4874/) from the mutation process is very low because [spontaneous mutation](https://www.ncbi.nlm.nih.gov/books/n/iga/A4529/def-item/A5396/) rates are low. The [mutation rate](https://www.ncbi.nlm.nih.gov/books/n/iga/A4529/def-item/A5121/) is defined as the probability that a copy of an [allele](https://www.ncbi.nlm.nih.gov/books/n/iga/A4529/def-item/A4550/) changes to some other allelic form in one generation. Suppose that a population were completely homozygous [*A*](https://www.ncbi.nlm.nih.gov/books/n/iga/A4529/def-item/A4530/) and mutations to *a* occurred at the rate of 1/100,000 Then, in the next generation, the frequency of *a* 0.00001 and the frequency of = 1/100,000 × alleles would be only 1.0 *A* alleles would be 0.99999. After yet another generation of mutation, the frequency of *a* 0.00009 to a new frequency of 0.000019, whereas the original allele would be reduced in frequency to 0.999981. It is obvious that the rate of increase of the new allele is extremely slow and that = 1/100,000 × would be increased by 0.99999 *it gets slower every generation* because there are fewer copies of the old allele still left to mutate.

Genetic variation is generated continuously by the mutational process, but its persistence in the genome is determined by different historical and genomic factors. Some of these factors leave an imprint on sequence variation across the whole genome, others only influence local patterns of variation.

Most new mutations that affect gene function have deleterious effects on reproductive fitness. But because genes represent only a small fraction of the human genome, most mutations are thought to have no effect on reproductive fitness and are effectively invisible to natural selection – a category referred to as ‘selectively neutral’. Most deoxyribonucleic acid (DNA) variants in the human genome are thought to be selectively neutral for three main reasons. First, the main portion of the genome, estimated as about 97%, neither codes for a functional product, such as protein or ribonucleic acid (RNA), nor indirectly affects gene function, by regulating expression or replication. Second, if a new variant does occur in the 1.5% of the genome that encodes a functional product (coding regions), it may not result in a change of amino acid (i.e. it may be a ‘synonymous’ substitution). Third, variants that do affect regulatory regions or coding regions and do change an amino acid (nonsynonymous substitutions) may have no effect on reproductive fitness.

Mutations may be subdivided according to:

A. Cause of mutation:

1. spontaneous
2. induced by exogenous and endogenous agents

B. Type of change brought about by a mutation:

1. genome mutations (numerical chromosomal aberrations)
2. chromosome mutations (structural chromosomal aberrations)
3. gene or point mutation (alteration in the DNA at the molecular level)

C. Place (cells) where a mutation occurs:

1. somatic mutations (occurring in body cells)
2. germ cell mutations (occurring in germ cells—gametes).

D. Phenotypic properties: morphological (shape, size, quantity, coloration), biochemical, lethal, be- havioral, silent.

E. Regulatory: increased or decreased expression, altered message processing, stability, or rate of translation.

The mutant genes can be divided into *dominant, recessive, autosomal, sex-linked*. It has been suggested that the polypeptide products of the genes involved in dominant conditions make up struc- tural proteins, whereas those concerned in recessive conditions make up enzymes. The sex-linked conditions are determined by mutant genes located on the X- or Y-chromosomes.

Mutagens

Mutations can be induced by *mutagens*, substances that cause a much higher rate of mutation.A mutagen is a natural or human-made agent which can greatly increase the mutation rate.

Mutagens can be subdivided into three groups: *physical, chemical* and *biological*.

Point, chromosomal and genomic mutations

**Point mutations** involve only one base pair of DNA and include both ***substitutions*** (*transitions* and *transversions*), and a **insertions** or a **deletion** of a single base pair..

*Transitions* occur when a purine is converted to a purine (A to G or G to A) or a pyrimide is con- verted to a pyrimidine (T to C or C to T). A *transversion* results when a purine is converted to a pyrimidine or a pyrimidine is converted to a purine. Substitution of one nucleotide for another is a result of tautomeric shift, a rare process by which the hydrogen atoms of a deoxyribonucleotide base move in a way that changes the properties of its hydrogen bonding.

Other common forms of genetic variation include insertions, deletions and inversions of one or more bases.

**Insertion**

An [insertion](https://medlineplus.gov/images/PX00009G_PRESENTATION.jpeg) changes the number of DNA bases in a gene by adding a piece of DNA. As a result, the protein made by the gene may not function properly.

**Deletion**

A [deletion](https://medlineplus.gov/images/PX000080_PRESENTATION.jpeg) changes the number of DNA bases by removing a piece of DNA. Small deletions may remove one or a few base pairs within a gene, while larger deletions can remove an entire gene or several neighboring genes. The deleted DNA may alter the function of the resulting protein(s).

**Duplication**

A [duplication](https://medlineplus.gov/images/PX00008K_PRESENTATION.jpeg) consists of a piece of DNA that is abnormally copied one or more times. This type of mutation may alter the function of the resulting protein.

**Repeat expansion**

Nucleotide repeats are short DNA sequences that are repeated a number of times in a row. For example, a trinucleotide repeat is made up of 3-base-pair sequences, and a tetranucleotide repeat is made up of 4-base-pair sequences. A [repeat expansion](https://medlineplus.gov/images/PX00004S_PRESENTATION.jpeg) is a mutation that increases the number of times that the short DNA sequence is repeated. This type of mutation can cause the resulting protein to function improperly.

Addition or deletion mutations are actually of nucleotide pairs; nevertheless, the convention is to call them base-pair additions or deletions. The simplest of these mutations are single-base-pair additions or single-base-pair deletions. There are examples in which mutations arise through simultaneous addition or deletion of multiple base pairs at once. Mechanisms that selectively produce certain kinds of multiple-base-pair additions or deletions are the cause of certain human genetic diseases.

A point mutation can result in **missense** (amino acid substitution), **nonsense** (insertion of a stop codon), or **frameshift** as result of insertions and deletions.

 Base substitutions can have different consequences at the protein level. Some base substitutions are “silent,” meaning that they result in a new codon that codes for the same [amino acid](https://www.britannica.com/science/amino-acid) as the wild type codon at that position or a codon that codes for a different amino acid that happens to have the same properties as those in the wild type. Substitutions that result in a functionally different amino acid are called “[missense](https://www.britannica.com/science/missense-mutation)” mutations; these can lead to alteration or loss of protein function. A more severe type of [base substitution](https://www.britannica.com/science/point-mutation), called a “[nonsense](https://www.britannica.com/science/nonsense-mutation)” mutation, results in a stop codon in a position where there was not one before, which causes the premature termination of [protein synthesis](https://www.britannica.com/science/methionine) and, more than likely, a complete loss of function in the finished protein.

There are the **silent mutations** that don't alter the phenotype. Silent because either:

* 1. Mutation occurs in non-coding or non-regulatory region
	2. Mutation occurs in an intron
	3. Mutation changes a codon such that it codes for the same amino acid.

A [mutation](https://www.ncbi.nlm.nih.gov/books/n/mcb/A7315/def-item/A7669/) involving deletion of a few base pairs generally affects the function of a single [gene](https://www.ncbi.nlm.nih.gov/books/n/mcb/A7315/def-item/A7519/). So **Gene mutations** are defined as those that occur entirely within one gene (and its upstream regulatory sequences) and may be either point mutations or other small disruptions of normal chro- mosomal structure that occur entirely within one gene

 **Chromosomal mutations** are defined as those that involve deletion, inversion, duplication, or other changes of a chromosomal region that is large enough so the change can be detected cytologically. Many types of chromosomal abnormalities exist, but they can be categorized as either numerical or structural. Numerical abnormalities are whole chromosomes either missing from or extra to the normal pair. Structural abnormalities are when part of an individual chromosome is missing, extra, switched to another chromosome, or turned upside down.

Chromosomal abnormalities can occur as an accident when the egg or the sperm is formed or during the early developmental stages of the fetus. The age of the mother and certain environmental factors may play a role in the occurrence of genetic errors. Prenatal screening and testing can be performed to examine the chromosomes of the fetus and detect some, but not all, types of chromosomal abnormalities.

**Genomic mutations** are defined as those that involve loss or gain of whole chromosomes, translocation from one chromosome to another or other gross chromosomal rearrangements. Note that both chromosomal and genomic mutations can cause **aneuploidy**.

**DNA repair.**

Sources of DNA damage in the cell. Enzymology of DNA repair. Repair of single-strand damage: excisional repair of nucleotides, excisional repair of bases, repair of mismatched bases. Repair of double-stranded damage: homologous recombination, non-homologous end joining.

1. List and describe the sources of DNA damage in the cell.

2. Explain the significance of DNA repair.

3. Explain the mechanisms of base excision, nucleotide excision, homologous recombination, non-homologous end joining modes of repair.

Damage to cellular DNA is involved in mutagenesis and the development of cancer. The DNA in a human cell undergoes several thousand to a million damaging events per day, generated by both external (exogenous) and internal metabolic (endogenous) processes. Changes to the cellular genome can generate errors in the transcription of DNA and ensuing translation into proteins necessary for signaling and cellular function. Genomic mutations can also be carried over into daughter generations of cells if the mutation is not repaired prior to mitosis.

Damage to DNA may occur as a result of normal cellular metabolism (endogenous damage) or under influences of origin external to the cell (exogenous damage). Therefore, DNA damaging agents and the associated damage mechanisms may be broadly classified as agents/mechanisms of endogenous or of exogenous origin.

DNA damage can also result from endogenous metabolic and biochemical reactions, some of which are not well understood

Hydrolysis reactions can partially or completely cleave the nucleotide base from the DNA strand. The chemical bond connecting a purine base (adenine or guanine) to the deoxyribosyl phosphate chain can spontaneously break in the process known as depurination. An estimated 10,000 depurination events occur per day in a mammalian cell.**[7](https://www.sigmaaldrich.com/technical-documents/articles/biofiles/dna-damage-and-repair.html%22%20%5Cl%20%22ref)** Depyrimidination (loss of pyrimidine base from thymine or cytosine) also occurs, but at a rate 20 to 100-fold lower than depurination.

Deamination occurs within the cell with the loss of amine groups from adenine, guanine, and cytosine rings, resulting in hypoxanthine, xanthine, and uracil, respectively. DNA repair enzymes are able to recognize and correct these unnatural bases. However, an uncorrected uracil base may be misread as a thymine during subsequent DNA replication and generate a C→T point mutation.

DNA methylation, a specific form of alkylation, occurs within the cell due to a reaction with S-adenosyl methionine (SAM). SAM is an intracellular metabolic intermediate that contains a highly reactive methyl group. In mammalian cells, methylation occurs at the 5-position of the cytosine ring of a cytidine base (C) that is 5’ to a guanosine base (G), i.e., sequence CpG. A significant source of mutation error is the spontaneous deamination of the 5-methylcytosine product of methyl-ation. Loss of the amine group results in a thymine base, which is not detected by DNA repair enzymes as an unnatural base. The resulting substitution is retained in DNA replication, creating a C→T point mutation

Normal metabolic processes generate reactive oxygen species (ROS), which modify bases by oxidation. Both purine and pyrimidine bases are subject to oxidation. The most common mutation is guanine oxidized to 8-oxo-7,8-dihydroguanine, resulting in the nucleotide 8-oxo-deoxy guano sine (8-oxo-dG). The 8-oxo-dG is capable of base pairing with deoxyadenosine, instead of pairing with deoxycytotidine as expected. If this error is not detected and corrected by mismatch repair enzymes, the DNA subsequently replicated will contain a C→A point mutation. ROS may also cause depurination, depyrimidination, and single-strand or double strand breaks in the DNA.

Other genomic mutations may be introduced during DNA replication in the S phase of the cell cycle. Polymerases that duplicate template DNA have a small but significant error rate, and may incorporate an incorrect nucleotide based on Watson-Crick pairing versus the template DNA. Chemically altered nucleotide precursors may be incorporated into the generated DNA by the polymerase, instead of normal bases. In addition, polymerases are prone to “stuttering” when copying sections of DNA that contain a large number of repeating nucleotides or repeating sequences (microsatellite regions). This enzymatic “stuttering” is due to a strand slippage, when the template and replicated strands of DNA slip out of proper alignment. As a result, the polymerase fails to insert the correct number of nucleotides indicated by the template DNA, resulting in too few or too many nucleotides in the daughter strand.

Single strand and double strand cleavage of the DNA may occur. Single strand breaks may result from damage to the deoxyribose moiety of the DNA deoxyribosylphosphate chain. Breaks also result as an intermediate step of the base excision repair pathway after the removal of deoxyribose phosphate by AP-endonuclease 1.**[8](https://www.sigmaaldrich.com/technical-documents/articles/biofiles/dna-damage-and-repair.html%22%20%5Cl%20%22ref)** When a single strand break occurs, both the nucleotide base and the deoxyribose backbone are lost from the DNA structure. Double strand cleavage most often occurs when the cell is passing through S-phase, as the DNA may be more susceptible to breakage while it is unraveling for use as a template for replication.

Because DNA uniquely serves as a permanent copy of the cell genome, however, changes in its structure are of much greater consequence than are alterations in other cell components, such as RNAs or [proteins](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3297/). Such damage to DNA can block replication or [transcription](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3391/), and can result in a high frequency of mutations—consequences that are unacceptable from the standpoint of cell reproduction. Because DNA is the repository of genetic information in each living cell, its integrity and stability are essential to life. To maintain the integrity of their genomes, cells have therefore had to evolve mechanisms to repair damaged DNA. DNA, is not inert; rather, it is a chemical entity subject to assault from the environment, and any resulting damage, if not repaired, will lead to mutation and possibly disease

DNA repair processes exist in both prokaryotic and eukaryotic organisms, and many of the proteins involved have been highly conserved throughout evolution. In fact, cells have evolved a number of mechanisms to detect and repair the various types of damage that can occur to DNA, no matter whether this damage is caused by the environment or by errors in replication. Because DNA is a molecule that plays an active and critical role in cell division, control of DNA repair is closely tied to regulation of the cell cycle.

Defects in DNA repair underlie a number of human genetic diseases that affect a wide variety of body systems but share a constellation of common traits, most notably a predisposition to cancer (Table 2). These disorders include ataxia-telangiectasia (AT), a degenerative motor condition caused by failure to repair oxidative damage in the cerebellum, and xeroderma pigmentosum (XP), a condition characterized by sensitivity to sunlight and linked to a defect in an important ultraviolet (UV) damage repair pathway

Most damage to [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/) is repaired by removal of the damaged bases followed by resynthesis of the excised region. Some lesions in DNA, however, can be repaired by direct reversal of the damage, which may be a more efficient way of dealing with specific types of DNA damage that occur frequently. Only a few types of DNA damage are repaired in this way, particularly pyrimidine dimers resulting from exposure to ultraviolet (UV) light and alkylated [guanine](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3105/) residues that have been modified by the addition of methyl or ethyl groups at the O6 position of the purine ring.

UV light is one of the major sources of damage to [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/) and is also the most thoroughly studied form of DNA damage in terms of repair mechanisms. Its importance is illustrated by the fact that exposure to solar UV irradiation is the cause of almost all skin [cancer](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A2944/) in humans. The major type of damage induced by UV light is the formation of [pyrimidine dimers](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/#A3304), in which adjacent pyrimidines on the same strand of DNA are joined by the formation of a cyclobutane ring resulting from saturation of the double bonds between carbons 5 and 6.

The formation of such dimers distorts the structure of the DNA chain and blocks [transcription](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3391/) or replication past the site of damage, so their repair is closely correlated with the ability of cells to survive UV irradiation. One mechanism of repairing UV-induced pyrimidine dimers is direct reversal of the dimerization reaction. The process is called [photoreactivation](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/%22%20%5Cl%20%22A3252) because energy derived from visible light is utilized to break the cyclobutane ring structure. The original pyrimidine bases remain in DNA, now restored to their normal state. As might be expected from the fact that solar UV irradiation is a major source of DNA damage for diverse cell types, the repair of pyrimidine dimers by [photoreactivation](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3252/) is common to a variety of prokaryotic and [eukaryotic cells](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3051/), including *E*. *coli*, [yeasts](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3425/), and some species of plants and animals. Curiously, however, photoreactivation is not universal; many species (including humans) lack this mechanism of DNA repair.

Another form of direct repair deals with damage resulting from the reaction between alkylating agents and [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/). Alkylating agents are reactive compounds that can transfer methyl or ethyl groups to a DNA base, thereby chemically modifying the base. A particularly important type of damage is methylation of the O6 position of [guanine](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3105/), because the product, O6-methylguanine, forms complementary base pairs with [thymine](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3386/) instead of [cytosine](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3007/). This lesion can be repaired by an enzyme (called O6-methylguanine methyltransferase) that transfers the methyl group from O6-methylguanine to a cysteine residue in its [active site](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A2893/) . The potentially mutagenic chemical modification is thus removed, and the original guanine is restored. Enzymes that catalyze this direct repair reaction are widespread in both prokaryotes and eukaryotes, including humans.

## Excision Repair

Although direct repair is an efficient way of dealing with particular types of [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/) damage, excision repair is a more general means of repairing a wide variety of chemical alterations to DNA. Consequently, the various types of excision repair are the most important DNA repair mechanisms in both prokaryotic and [eukaryotic cells](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3051/). In excision repair, the damaged DNA is recognized and removed, either as free bases or as nucleotides. The resulting gap is then filled in by synthesis of a new DNA strand, using the undamaged complementary strand as a template. Three types of excision repair—[base-excision repair](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A2934/), [nucleotide](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3214/)-excision repair, and [mismatch repair](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3182/)—enable cells to cope with a variety of different kinds of DNA damage.

The repair of [uracil](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3416/)-containing [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/) is a good example of [base-excision repair](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/#A2934), in which single damaged bases are recognized and removed from the DNA molecule. Uracil can arise in DNA by two mechanisms: (1) Uracil (as dUTP [deoxyuridine triphosphate]) is occasionally incorporated in place of [thymine](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3386/) during DNA synthesis, and (2) uracil can be formed in DNA by the deamination of [cytosine](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3007/) . The second mechanism is of much greater biological significance because it alters the normal pattern of complementary base pairing and thus represents a mutagenic event. The excision of uracil in DNA is catalyzed by [DNA glycosylase](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/#A3017), an enzyme that cleaves the bond linking the base (uracil) to the deoxyribose of the DNA backbone. This reaction yields free uracil and an apyrimidinic site—a sugar with no base attached. DNA glycosylases also recognize and remove other abnormal bases, including hypoxanthine formed by the deamination of [adenine](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A2896/), pyrimidine dimers, alkylated purines other than O6-alkylguanine, and bases damaged by oxidation or ionizing radiation

Nucleotide-excision repair systems have also been studied extensively in eukaryotes, particularly in [yeasts](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3425/) and in humans. In yeasts, as in *E*. *coli*, several genes involved in [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/) repair (called *RAD* genes for *rad*iation sensitivity) have been identified by the isolation of mutants with increased sensitivity to UV light. In humans, DNA repair genes have been identified largely by studies of individuals suffering from inherited diseases resulting from deficiencies in the ability to repair DNA damage. The most extensively studied of these diseases is xeroderma pigmentosum (XP), a rare genetic disorder that affects approximately one in 250,000 people. Individuals with this disease are extremely sensitive to UV light and develop multiple skin cancers on the regions of their bodies that are exposed to sunlight. In 1968 James Cleaver made the key discovery that cultured cells from XP patients were deficient in the ability to carry out [nucleotide](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3214/)-excision repair. This observation not only provided the first link between DNA repair and [cancer](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A2944/), but also suggested the use of XP cells as an experimental system to identify human DNA repair genes. The identification of human DNA repair genes has been accomplished by studies not only of XP cells, but also of two other human diseases resulting from DNA repair defects (Cockayne's syndrome and trichothiodystrophy) and of UV-sensitive mutants of rodent cell lines. The availability of mammalian cells with defects in DNA repair has allowed the cloning of repair genes based on the ability of wild-type alleles to restore normal UV sensitivity to mutant cells in [gene transfer](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3082/) assays, thereby opening the door to experimental analysis of nucleotide-excision repair in mammalian cells.

The direct reversal and excision repair systems act to correct [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/) damage before replication, so that replicative DNA synthesis can proceed using an undamaged DNA strand as a template. Should these systems fail, however, the cell has alternative mechanisms for dealing with damaged DNA at the [replication fork](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3316/). Pyrimidine dimers and many other types of lesions cannot be copied by the normal action of DNA polymerases, so replication is blocked at the sites of such damage. Downstream of the damaged site, however, replication can be initiated again by the synthesis of an Okazaki fragment and can proceed along the damaged template strand ([Figure 5.27](https://www.ncbi.nlm.nih.gov/books/NBK9900/figure/A809/?report=objectonly" \t "object)). The result is a daughter strand that has a gap opposite the site of damage to the parental strand. One of two types of mechanisms may be used to repair such gaps in newly synthesized DNA: recombinational repair or error-prone repair.

**Recombinational repair** depends on the fact that one strand of the parental [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/) was undamaged and therefore was copied during replication to yield a normal daughter molecule (see [Figure 5.27](https://www.ncbi.nlm.nih.gov/books/NBK9900/figure/A809/?report=objectonly" \t "object)). The undamaged parental strand can be used to fill the gap opposite the site of damage in the other daughter molecule by [recombination](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3315/) between homologous DNA sequences (see the next section). Because the resulting gap in the previously intact parental strand is opposite an undamaged strand, it can be filled in by [DNA polymerase](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3019/). Although the other parent molecule still retains the original damage (e.g., a [pyrimidine dimer](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3304/)), the damage now lies opposite a normal strand and can be dealt with later by excision repair. By a similar mechanism, recombination with an intact DNA molecule can be used to repair double strand breaks, which are frequently introduced into DNA by radiation and other damaging agents.

In **error-prone repair**, a gap opposite a site of [DNA](https://www.ncbi.nlm.nih.gov/books/n/cooper/A2886/def-item/A3010/) damage is filled by newly synthesized DNA. Since the new DNA is synthesized from a damaged template strand, this form of DNA synthesis is very inaccurate and leads to frequent mutations. It is used only in bacteria that have been subjected to potentially lethal conditions, such as extensive UV irradiation. Such treatments induce the SOS response, which may be viewed as a mechanism for dealing with extreme environmental stress. The SOS response includes inhibition of cell division and induction of repair systems to cope with a high level of DNA damage. Under these conditions, error-prone repair mechanisms are used, presumably as a way of dealing with damage so extensive that cell death is the only alternative.

**Epigenetics.**

Significance of epigenetic regulation of gene expression. Mechanisms of epigenetic regulation: DNA methylation, RNA interference. Mechanisms of epigenetic regulation: histone modifications, histone variants.

1. Explain the importance of epigenetic regulation and its role in heritability of cellular traits.

2. Explain the role of DNA methylation in regulation of gene expression.

3. Explain the mechanism of RNAi.

4. Describe chromatin structure at the levels of organization: nucleosome, 30-nm fiber, chromosome.

5. Explain the effects of histones on transcription.

6. Explain how transcription is affected by: nucleosome positioning, histone acetylation and methylation, chromatin remodeling.

7. Describe the mechanisms and major players of above mentioned processes.

Since the early days of genetics research, scientists have noted certain heritable phenotypic differences that are not due to differences in the nucleotide sequence of DNA. Current evidence suggests that these “epigenetic” phenomena might be controlled by a number of mechanisms, including the modification of DNA cytosine bases with methyl groups, the addition of various chemical groups to histone proteins, and the recruitment of protein factors to specific DNA sites via interactions with non-protein-coding RNAs.

Epigenetics is the study of heritable changes in gene expression (active versus inactive genes) that do not involve changes to the underlying DNA sequence — a change in phenotype without a change in genotype — which in turn affects how cells read the genes. Epigenetic change is a regular and natural occurrence but can also be influenced by several factors including age, the environment/lifestyle, and disease state. Epigenetic modifications can manifest as commonly as the manner in which cells terminally differentiate to end up as skin cells, liver cells, brain cells, etc. Or, epigenetic change can have more damaging effects that can result in diseases like cancer.

At least three systems including DNA methylation, histone modification and non-coding RNA (ncRNA)-associated gene silencing are currently considered to initiate and sustain epigenetic change.[**1**](https://www.whatisepigenetics.com/fundamentals/#fn-12-1) New and ongoing research is continuously uncovering the role of epigenetics in a variety of human disorders and fatal diseases.

### DNA Methylation

DNA methylation works by adding a chemical group to DNA. Typically, this group is added to specific places on the DNA, where it blocks the proteins that attach to DNA to “read” the gene. This chemical group can be removed through a process called demethylation. Typically, methylation turns genes “off” and demethylation turns genes “on.”

### Histone modification

DNA wraps around proteins called histones. DNA wrapped tightly around histones cannot be accessed by proteins that “read” the gene. Some genes are wrapped around histones and are turned “off” while some genes are not wrapped around histones and are turned “on.” Chemical groups can be added or removed from histones and change whether a gene is unwrapped or wrapped (“on” or “off”).

### Non-coding RNA

Your DNA is used as instructions for making coding and non-coding RNA. Coding RNA is used to make proteins. Non-coding RNA helps control gene expression by attaching to coding RNA, along with certain proteins, to break down the coding RNA so that it cannot be used to make proteins. Non-coding RNA may also recruit proteins to modify histones to turn genes “on” or “off.”

While epigenetic changes are required for normal development and health, they can also be responsible for some disease states. Disrupting any of the three systems that contribute to epigenetic alterations can cause abnormal activation or silencing of genes. Such disruptions have been associated with cancer, syndromes involving chromosomal instabilities, and mental retardation .

### Epigenetics and Cancer

The first human disease to be linked to epigenetics was cancer, in 1983. Researchers found that diseased tissue from patients with colorectal cancer had less DNA methylation than normal tissue from the same patients (Feinberg & Vogelstein, 1983). Because methylated genes are typically turned off, loss of DNA methylation can cause abnormally high gene activation by altering the arrangement of chromatin. On the other hand, too much methylation can undo the work of protective tumor suppressor genes.

As previously mentioned, DNA methylation occurs at CpG sites, and a majority of CpG cytosines are methylated in mammals. However, there are stretches of DNA near promoter regions that have higher concentrations of CpG sites (known as CpG islands) that are free of methylation in normal cells. These CpG islands become excessively methylated in cancer cells, thereby causing genes that should not be silenced to turn off. This abnormality is the trademark epigenetic change that occurs in tumors and happens early in the development of cancer. Hypermethylation of CpG islands can cause tumors by shutting off tumor-suppressor genes. In fact, these types of changes may be more common in human cancer than DNA sequence mutations/

Furthermore, although epigenetic changes do not alter the sequence of DNA, they can cause mutations. About half of the genes that cause familial or inherited forms of cancer are turned off by methylation. Most of these genes normally suppress tumor formation and help repair DNA, including O6-methylguanine-DNA methyltransferase (*MGMT*), MLH1 cyclin-dependent kinase inhibitor 2B (*CDKN2B*), and *RASSF1A*. For example, hypermethylation of the promoter of *MGMT* causes the number of G-to-A mutations to increase

### Epigenetics and Mental Retardation

Fragile X syndrome is the most frequently inherited mental disability, particularly in males. Both sexes can be affected by this condition, but because males only have one X chromosome, one fragile X will impact them more severely. Indeed, fragile X syndrome occurs in approximately 1 in 4,000 males and 1 in 8,000 females. People with this syndrome have severe intellectual disabilities, delayed verbal development, and "autistic-like" behavior .

Fragile X syndrome gets its name from the way the part of the X chromosome that contains the gene abnormality looks under a microscope; it usually appears as if it is hanging by a thread and easily breakable (Figure 3). The syndrome is caused by an abnormality in the *FMR1* (fragile X mental retardation 1) gene. People who do not have fragile X syndrome have 6 to 50 repeats of the trinucleotide CGG in their *FMR1* gene. However, individuals with over 200 repeats have a full mutation, and they usually show symptoms of the syndrome. Too many CGGs cause the CpG islands at the promoter region of the *FMR1* gene to become methylated; normally, they are not. This methylation turns the gene off, stopping the *FMR1* gene from producing an important protein called fragile X mental retardation protein. Loss of this specific protein causes fragile X syndrome. Although a lot of attention has been given to the CGG expansion mutation as the cause of fragile X, the epigenetic change associated with *FMR1* methylation is the real syndrome culprit.

Fragile X syndrome is not the only disorder associated with mental retardation that involves epigenetic changes. Other such conditions include Rubenstein-Taybi, Coffin-Lowry[, Prader-Willi](https://www.nature.com/scitable/topicpage/Imprinting-and-Genetic-Disease-Angelman-Prader-Willi-923%22%20%5Co%20%22%2C%20Prader-Willi), [Angelman](https://www.nature.com/scitable/topicpage/Imprinting-and-Genetic-Disease-Angelman-Prader-Willi-923%22%20%5Co%20%22Angelman), [Beckwith-Wiedemann](https://www.nature.com/scitable/topicpage/Imprinting-and-Genetic-Disease-Angelman-Prader-Willi-923%22%20%5Co%20%22Beckwith-Wiedemann), ATR-X, and Rett syndromes .

## Combating Diseases with Epigenetic Therapy

Because so many diseases, such as cancer, involve epigenetic changes, it seems reasonable to try to counteract these modifications with epigenetic treatments. These changes seem an ideal target because they are by nature reversible, unlike DNA sequence mutations. The most popular of these treatments aim to alter either DNA methylation or histone acetylation.

Inhibitors of DNA methylation can reactivate genes that have been silenced. Two examples of these types of drugs are 5-azacytidine and 5-aza-2′-deoxycytidine (Egger *et al*., 2004). These medications work by acting like the nucleotide cytosine and incorporating themselves into DNA while it is replicating. After they are incorporated into DNA, the drugs block DNMT enzymes from acting, which inhibits DNA methylation.

Drugs aimed at histone modifications are called histone deacetylase (HDAC) inhibitors. HDACs are enzymes that remove the acetyl groups from DNA, which condenses chromatin and stops transcription. Blocking this process with HDAC inhibitors turns on gene expression. The most common HDAC inhibitors include phenylbutyric acid, SAHA, depsipeptide, and valproic acid (Egger *et al.*, 2004).

Caution in using epigenetic therapy is necessary because epigenetic processes and changes are so widespread. To be successful, epigenetic treatments must be selective to irregular cells; otherwise, activating gene transcription in normal cells could make them cancerous, so the treatments could cause the very disorders they are trying to counteract. Despite this possible drawback, researchers are finding ways to specifically target abnormal cells with minimal damage to normal cells, and epigenetic therapy is beginning to look increasingly promising.

.

**Intracellular signaling.**

Intracellular signaling pathways. Membrane receptors. Secondary intermediaries. Cytoplasmic and nuclear receptors.

1. Give the definition of intracellular signaling (signal transduction).

2. Classify and characterize membrane cell receptors, give specific examples.

3. Describe and provide examples of secondary intermediaries.

4. Predict the signaling pathways when the cell is exposed to insulin and thyroid hormone, steroid hormones.

5. Characterize the cytoplasmic and nuclear receptors.

6. Give examples of signaling pathways when a cell is exposed to steroid hormones

The environment in which higher organisms live is changing all the time, and the coordination of the functions of the organism requires a perfect mechanism for mutual recognition, response and interaction between cells. The mechanism is called cell communication. In this system, cells recognize cells in contact with them, or recognize various signals present in the surrounding environment (from surrounding or distant cells) and transform them into functional changes in various molecules within the cell, thereby changing certain metabolic processes within the cell, affecting the growth rate of the cells, and even inducing cell death. The ultimate goal of signal transduction is to permit the body to respond appropriately to the changes in the environment at an overall level

The essence is that a part of the cell in the body emits signals, and another part of the cells receive signals and transforms them into changes in cell function. Therefore, signal transduction will directly affect the regulation of cell proliferation, differentiation, metabolism and death.

Signal transduction (also known as cell signaling) is the transmission of molecular signals from a cell's exterior to its interior. Signals received by cells must be transmitted effectively into the cell to ensure an appropriate response.

Cell signal transduction refers to the binding of extracellular factors to a receptor (membrane receptor or nuclear receptor), triggering a series of biochemical reactions and protein interactions in the cell, until the genes which required for cellular physiological reactions begin to express and the process of forming biological effects. It is now known that there is a variety of signal transduction methods and pathways in cells, and there are multiple levels of cross-regulation between various methods and pathways, which are a very complicated network system.

Signals are most often chemicals that can be found in the extracellular fluid around cells. Some cells also respond to mechanical stimuli. For example, sensory cells in the skin respond to the pressure of touch, whereas similar cells in the ear react to the movement of sound waves. In addition, specialized cells in the human vascular system detect changes in blood pressure — information that the body uses to maintain a consistent cardiac load.

Summarizing our words we can say that there 2 processes occur during signal transduction, the original **intercellular** (between-cells) signal is converted into an **intracellular** (within-cell) signal that triggers a response

**Cell signaling process can be divided into 3 stages.**

**1. Reception**: A cell detects a signaling molecule from the outside of the cell. A signal is detected when the chemical signal (also known as a ligand) binds to a receptor protein on the surface of the cell or inside the cell.

**2. Transduction**: When the signaling molecule binds the receptor it changes the receptor protein in some way. This change initiates the process of transduction. Signal transduction is usually a pathway of several steps. Each relay molecule in the signal transduction pathway changes the next molecule in the pathway.

**3. Response**: Finally, the signal triggers a specific cellular response.

***Types 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).

Synaptic signaling. Neurotransmitter is released from vesicles at the end of the axon of the sending cell. It diffuses across the small gap between sending and target neurons and binds to receptors on the target neuron.

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.

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**.

Cells usually communicate with each other through extracellular messenger molecules. Extracellular messengers can travel a short distance and stimulate cells that are in close proximity to the origin of the message, or they can travel throughout the body, potentially stimulating cells that are far away from the source. Cell signaling is initiated with the release of a messenger molecule by a cell that is engaged in sending messages to other cells in the body.

The extracellular environments of cells contain hundreds of different informational molecules, ranging from small compounds (e.g., steroids and neurotransmitters ) to small, soluble protein hormones (e.g., glucagon and insulin) to huge glycoproteins bound to the surfaces of other cells.

Cells can only respond to a particular extracellular message if they express receptors that specifically recognize and bind that messenger molecule. The molecule that binds to the receptor is called a **ligand**. Different types of cells possess different complements of receptors, which allow them to respond to different extracellular messengers. Even cells that share a specific receptor may respond very differently to the same extracellular messenger

**Receptors** are protein molecules inside the target cell or on its surface that receive a chemical signal.

**Internal receptors**, also known as intracellular or cytoplasmic receptors, are found in the cytoplasm of the cell and respond to hydrophobic ligand molecules that are able to travel across the plasma membrane. Once inside the cell, many of these molecules bind to proteins that act as regulators of mRNA synthesis. Recall that mRNA carries genetic information from the DNA in a cell’s nucleus out to the ribosome, where the protein is assembled. When the ligand binds to the internal receptor, a change in shape is triggered that exposes a DNA-binding site on the receptor protein. The ligand-receptor complex moves into the nucleus, then binds to specific regions of the DNA and promotes the production of mRNA from specific genes Internal receptors can directly influence gene expression (how much of a specific protein is produced from a gene) without having to pass the signal on to other receptors or messengers.

**Cell-surface receptors**, also known as **transmembrane receptors**, are proteins that are found attached to the cell membrane. These receptors bind to external ligand molecules (ligands that do not travel across the cell membrane). This type of receptor spans the plasma membrane and performs **signal transduction**, in which an extracellular signal is converted into an intercellular signal. Ligands that interact with cell-surface receptors do not have to enter the cell that they affect. Cell-surface receptors are also called cell-specific proteins or markers because they are specific to individual cell types.

#### Structure of Cell Surface Receptors

Cell surface receptors are integral membrane proteins and, as such, have regions that contribute to three basic domains:

* **Extracellular domains:** Some of the residues exposed to the outside of the cell interact with and bind the hormone - another term for these regions is the *ligand-binding domain*.
* **Transmembrane domains:** Hydrophobic stretches of amino acids are "comfortable" in the lipid bilayer and serve to anchor the receptor in the membrane.
* **Cytoplasmic or intracellular domains:** Tails or loops of the receptor that are within the cytoplasm react to hormone binding by interacting in some way with other molecules, leading to generation of second messengers. Cytoplasmic residues of the receptor are thus the *effector region* of the molecule.
* Signals received by receptors at the cell surface or, in some cases, within the cell are often relayed throughout the cell via generation of small, rapidly diffusing molecules referred to as **second messengers.** These second messengers broadcast the initial signal (the “first message”) that occurs when a ligand binds to a specific cellular receptor; ligand binding alters the protein conformation of the receptor such that it stimulates nearby effector proteins that catalyze the production or, in the case of ions, release or influx of the second messenger. The second messenger then diffuses rapidly to protein targets elsewhere within the cell, altering the activities as a response to the new information received by the receptor.
* There are 3 second messenger pathways
* (1) activation of adenylyl cyclase by G-protein-coupled receptors (GPCRs) to generate the cyclic nucleotide second messenger 3′-5′-cyclic adenosine monophosphate (cAMP);
* (2) stimulation of phosphoinositide 3-kinase (PI3K) by growth factor receptors to generate the lipid second messenger phosphatidylinositol 3,4,5-trisphosphate (PIP3);
* and (3) activation of phospholipase C by GPCRs to generate the two second messengers membrane-bound messenger diacylglycerol (DAG) and soluble messenger inositol 1,4,5-trisphosphate (IP3), which binds to receptors on subcellular organelles to release calcium into the cytosol.

The activation of multiple effector pathways by a single plasma membrane receptor and the production of multiple second messengers by each effector can generate a high degree of amplification in signal transduction, and stimulate diverse, pleiotropic, responses depending on the cell type.

Second messengers fall into four major classes:

* cyclic nucleotides, such as cAMP and other soluble molecules that signal within the cytosol;
* lipid messengers that signal within cell membranes;
* ions that signal within and between cellular compartments;
* and gases and free radicals that can signal throughout the cell and even to neighboring cells.

Second messengers from each of these classes bind to specific protein targets, altering their activity to relay downstream signals. In many cases, these targets are enzymes whose catalytic activity is modified by direct binding of the second messengers. The activation of multiple target enzymes by a single second messenger molecule further amplifies the signal.Second messengers are not only produced in response to extracellular stimuli, but also in response to stimuli from within the cell. Moreover, their levels are exquisitely controlled by various homeostatic mechanisms to ensure precision in cell signaling.

**Cell differentiation and development of a multicellular organism.**

Differentiation. Morphogenesis. Ontogenesis. Stem cells. Totipotency. Pluripotency. Ageing mechanisms of the body.

*Maximal point: 7*

1. Give the definitions to the following terms: cell differentiation, morphogenesis, embryogenesis, ontogenesis, stem cells, totipotency, pluripotency.

2.Explain how the level of expression of various genes changes during cell differentiation and at different stages of development of a multicellular organism.

3. Describe the use of stem cells in medicine and cosmetology, analyze the advantages and disadvantages of these methods.

4. Analyze the various theories of aging in the body and the possible relationship of the aging process with stem cells and molecular biological processes.

Living organisms can be made of a single cell, such as bacteria and protists, or they can be [multicellular](https://www.texasgateway.org/glossary/multicellular), like plants, animals, and fungi. [Unicellular](https://www.texasgateway.org/glossary/unicellular)organisms, like bacteria, are able to perform all life functions within one single cell. They can transport molecules, metabolize nutrients, and reproduce within this one cell.

Multicellular organisms need many different types of cells to carry out the same life processes. Each of these special types of cells has a different structure that helps it perform a specific function. Humans have many different types of cells with different jobs, such as blood cells that carry oxygen and nerve cells that transmit signals to all parts of the body

 An animal or plant starts its life as a single cell—a fertilized [egg](https://www.ncbi.nlm.nih.gov/books/n/mboc4/A4754/def-item/A5114/). The entire set of transformations of an individual from the moment of inception to his death is called ontogeny. The development of a whole organism from a fertilized egg is a complex and multi-stage process called embryogenesis. In the process of embryonic development, many phenomena occur, which together can be characterized as an interconnected and mutually consistent process of reproduction, growth, differentiation and specialization of cells, the progenitor of which is 1 fertilized cell.

.  Differentiation is the process of the emergence and development of structural and functional differences between initially homogeneous embryonic cells, as a result of which specialized cells, tissues and organs of a multicellular organism are formed. Cell differentiation is an essential part of the formation of a multicellular organism.

During [development](https://www.ncbi.nlm.nih.gov/books/n/mboc4/A4754/def-item/A5071/), this cell divides repeatedly to produce many different cells in a final pattern of spectacular complexity and precision. In a developing embryo, all these processes are happening at once, in a kaleidoscopic variety of different ways in different parts of the organism.

(1) *cell proliferation*, producing many cells from one,

(2) *cell specialization,* creating cells with different characteristics at different positions,

 (3) *cell interactions,* coordinating the behavior of one cell with that of its neighbors,

 (4) *cell movement,* rearranging the cells to form structured tissues and organs

[Cell differentiation](https://www.texasgateway.org/glossary/celldifferentiation) is the process by which cells become specialized in order to perform different functions. The processes of organizing differentiated cells into tissues and organs are called morphogenesis.

Features of cell differentiation include: The potential of differentiation gradually appears with the development of the individual. During embryonic development, the cells gradually change from "all-around" to "multi-energy", and finally to the "single-energy", which is the general rule of cell differentiation. In the process of individual development, multicellular biological cells have both temporal differentiation and spatial differentiation; cell differentiation is compatible with the state and speed of cell division, and differentiation must be based on division, that is, differentiation is inevitable with division, but the dividing cells do not necessarily need differentiate. The higher the degree of differentiation, the worse the ability to divide; the cell differentiation is highly stable. Under normal physiological conditions, cells that have differentiated into a specific, stable type are generally impossible to reverse to undifferentiated state or become other types. In the general case, differentiation is irreversible, i.e. highly differentiated cells cannot transform into cells of another type. This phenomenon is called terminal differentiation and is inherent mainly in animal cells. Differences manifested in the process of differentiation are retained by the cells during reproduction, that is, they are hereditarily fixed (for example, liver cells, when multiplying, produce only liver cells, and muscle cells - only muscle cells, etc.).

Cell differentiation is plastic, and the differentiated cells re-enter the undifferentiated state or transdifferentiate into another type of cell under special conditions. Under certain conditions, the differentiated cells are also unstable, and their gene expression patterns can also undergo reversible changes and return to their undifferentiated state. This process is called dedifferentiation.

Differentiation mainly occurs in the embryonic period, as well as in the early stages of postembryonic development. In addition, differentiation occurs in some organs of the adult organism. Adult or 'somatic' stem cells are thought to be undifferentiated

Cell differentiation takes place in three situations:

* The *growth* of an immature organism into an adult.
* Normal *turnover* of cells such as blood cells in mature organisms.
* The *repair* of damaged tissues when specialized cells have to be replaced.

During the differentiation process, cells gradually become committed towards developing into a given cell type. Here, the state of commitment may be described as "specification" representing a reversible type of commitment or "determination" representing irreversible commitment.

Although the two represent differential gene activity, the properties of cells in this stage is not completely similar to that of fully differentiated cells. For instance, in the specification state, cells are not stable over a long period of time.

There are two mechanisms that bring about altered commitments in the different regions of the early embryo.

These include:

* Cytoplasmic localization
* Induction

**Cytoplasmic Localization** - This occurs during the earliest stage of embryo development. Here, the embryo divides without growth and undergoes cleavage divisions that produce blastomeres (separate cells). Each of these cells inherit a given region of the cytoplasm of the original cell that may contain cytoplasmic determinants (reuratory substances). Once the embryo becomes a morula (solid mass of blastomeres) it is composed of two or more differently committed cell populations. The cytoplasmic determinants may contain mRNA or protein a given state of activation that influence specific development.

 **Induction** - In induction, a substance secreted by one group of cells causes changes in the development of another group. During early development, induction tends to be instructive in that tissue assumes a given state of commitment in the presence of the signal.

In induction, inductive signals also evoke various responses at varying concentrations which results in the formation of a sequence of groups of cells, each being in a different state of specification.

 During the final phase of cell differentiation, there is formation of several types of differentiated cells from one population of stem cells of the precursor. Here, terminal differentiation occurs both in embryonic development as well as in tissues during postnatal life.

Stem cells have the ability to differentiate into specialized cells and can self-renew; dividing to give rise to new stem cells. Once the female egg has been fertilized, the cells formed after cell division contain [DNA](https://www.microscopemaster.com/dna-under-the-microscope.html) that is identical. Beginning with the totipotent and pluripotent stem cells which give rise to all of the specialized tissues in the body, the DNA sequence of the cells does not change. The [genome](https://www.ncbi.nlm.nih.gov/books/n/mboc4/A4754/def-item/A5222/) is normally identical in every cell; the cells differ not because they contain different genetic information, but because they express different sets of genes. Individual development from one cell to a multicellular mature organism is the result of sequential, selective inclusion in different gene regions of chromosomes in different cells. This selective [gene](https://www.ncbi.nlm.nih.gov/books/n/mboc4/A4754/def-item/A5215/) [expression](https://www.ncbi.nlm.nih.gov/books/n/mboc4/A4754/def-item/A5163/) controls essential processes by which the embryo is constructed.

 Cellular differentiation is often controlled by [cell signaling](https://en.wikipedia.org/wiki/Cell_signaling%22%20%5Co%20%22Cell%20signaling). Many of the signal molecules that convey information from cell to cell during the control of cellular differentiation are called [growth factors](https://en.wikipedia.org/wiki/Growth_factor%22%20%5Co%20%22Growth%20factor). Although the details of specific [signal transduction](https://en.wikipedia.org/wiki/Signal_transduction%22%20%5Co%20%22Signal%20transduction) pathways vary, these pathways often share the following general steps. A ligand produced by one cell binds to a receptor in the extracellular region of another cell, inducing a conformational change in the receptor. The shape of the cytoplasmic domain of the receptor changes, and the receptor acquires enzymatic activity. The receptor then catalyzes reactions that phosphorylate other proteins, activating them. A cascade of phosphorylation reactions eventually activates a dormant transcription factor or cytoskeletal protein, thus contributing to the differentiation process in the target cell.[[22]](https://en.wikipedia.org/wiki/Cellular_differentiation%22%20%5Cl%20%22cite_note-Gilbert-22) Cells and tissues can vary in competence, their ability to respond to external signals.[[23]](https://en.wikipedia.org/wiki/Cellular_differentiation%22%20%5Cl%20%22cite_note-Rudel-23)

Signal induction refers to cascades of [signaling](https://en.wikipedia.org/wiki/Cell_signaling%22%20%5Co%20%22Cell%20signaling) events, during which a cell or tissue signals to another cell or tissue to influence its developmental fate.

Although differentiation is not thought to occur by permanent loss of genetic material, DNA can be modified in a way that affects gene expression. For instance, [DNA and its associated histone proteins](https://www.nature.com/scitable/topicpage/DNA-Packaging-Nucleosomes-and-Chromatin-310%22%20%5Co%20%22DNA%20and%20its%20associated%20histone%20proteins) (together known as chromatin) can be chemically modified by a cell's own machinery. Chromatin modification can affect gene expression by changing the accessibility of genes to transcription factors, in either a positive or a negative manner.

Two major classes of such chemical modifications include [DNA methylation](https://www.nature.com/scitable/topicpage/The-Role-of-Methylation-in-Gene-Expression-1070%22%20%5Co%20%22DNA%20methylation) and histone modification (methylation and/or acetylation). These [changes are often described as epigenetic](https://www.nature.com/scitable/topicpage/Epigenetic-Influences-and-Disease-895%22%20%5Co%20%22changes%20are%20often%20described%20as%20epigenetic) because they do not act to alter the primary DNA sequence but instead act at a level just above the DNA sequence. Although DNA methylation and histone modification are not genetic, cells have mechanisms to copy this epigenetic information during their division so that their daughter cells contain the same regulatory data.

Changes in chromatin modification play an important role in regulating gene expression during developmental cell-type specification as well. For example, chromatin-modifying proteins play an essential role in muscle cell differentiation via interactions with key muscle-promoting transcription factors MyoD and MEF. That is, these factors are thought to help recruit chromatin modifying factors, such as histone acetyltransferases and deacetylases. In so doing, MyoD and MEF alter access to their target sites upstream of muscle differentiation genes. For instance, MyoD binds histone acetyltransferases p300 and PCAF, and this activity is essential for muscle cell differentiation. This example provides evidence for a link among chromatin modifications, transcription factors, and, ultimately, cell-fate-specific changes in gene expression.

Chromatin modification can be stable over the life of an organism, thereby effectively permanently influencing gene expression. However, that is not to say that chromatin modification is irreversible. For instance, chromatin can become mismodified in certain cancers, suggesting that, although important, the change is not permanent. Moreover, chromatin modifications are usually erased and reset during the production of gametes, such that the adult program of intrinsic cues is replaced with a program more suited to embryonic development

In fact, embryonic cell types are known to contain a unique set of chromatin modifications that are different from those found in adult cell types (Bernstein *et al*., 2006; Meshorer *et al*., 2006). This has led to the tantalizing proposal that chromatin modification helps lock in changes in gene expression that are required during development. The permanent silencing of the genes involved only in embryogenesis could then drive the development of cells toward more mature cell types. By blocking accessibility of transcription machinery, for example, chromatin modification could prevent the need for continued repression through active binding of a repressive transcription factor. Alternatively, the genes required for an adult cell type might contain chromatin modifications (especially histone acetylation) that cause the DNA to become open and, therefore, more accessible to the transcription machinery.

Interestingly, embryonic cell types have been found to contain a signature chromatin modification in the regions that regulate the expression of genes involved in early embryonic development (Bernstein *et al*., 2006). Such regions were found to contain chromatin modifications with both silencing and promoting characteristics. The finding of these bivalent (two-directional) markers in association with genes important for embryonic development has led to the belief that embryonic cells exist in a special epigenetic state, wherein they can choose to remain embryonic (as in an embryonic stem cell) or to differentiate (as in normal development), and bivalent domains provide a means by which to quickly choose between the two options.

Together, these lines of evidence have led to an emerging hypothesis that cell-cell signaling and epigenetic changes converge to guide cell differentiation decisions both during development and beyond

Basic cell differentiation occurs after a sperm cell fertilizes an egg and the resulting [zygote](https://sciencing.com/what-is-a-zygote-13714437.html) reaches a certain size. At that point the zygote starts developing different cell types and needs differentiated cells to take on the specialized functions. A cell capable of differentiating into any type of cell is known as "totipotent". For mammals, totipotent includes the zygote and products of the first few cell divisions. There are also certain types of cells that can differentiate into many types of cells. These cells are known as "pluripotent" or stem cells in animals (meristemic cells in higher plants).Embryonic stem cells (ES cells) have the potential to develop into different types of cells.

**Stem cell differentiation** involves the changing of a cell to a more specialized cell type, involving a switch from proliferation to specialization. This involves a succession of alterations in cell morphology, membrane potential, metabolic activity and responsiveness to certain signals. Differentiation leads to the commitment of a cell to developmental lineages and the acquisition of specific functions of committed cells depending upon the tissue in which they will finally reside. Stem cell differentiation is tightly regulated by signaling pathways and modifications in gene expression.

Stem cells can be categorized into groups depending on their ability to differentiate.

* Totipotent: can differentiate into all cell types;
* Pluripotent: can differentiate into almost all cell types;
* Multipotent: can differentiate into a related family of cell types;
* Oligopotent: can differentiate into a few different cells;
* Unipotent: can produce one cell type only.

Differentiation mainly occurs in the embryonic period, as well as in the early stages of postembryonic development. In addition, differentiation occurs in some organs of the adult organism. Adult or 'somatic' stem cells are thought to be undifferentiated. Their primary role is to self-renew and maintain or repair the tissue in which they reside. Examples *of*[*stem*](https://www.microscopemaster.com/stem-cells.html)*and*[*progenitor cells*](https://www.microscopemaster.com/progenitor-cells.html) *in mature organism  include:*

***H*ematopoietic Stem Cells** - These are from the bone marrow and are involved in the production of red and white blood cells as well as the platelets.

[Mesenchymal Stem Cells](https://www.microscopemaster.com/mesenchymal-stem-cells.html) - Also from the bone marrow, these cells are involved in the production of fat cells, [stromal cells](https://www.microscopemaster.com/stromal-cells.html) as well as a given type of bone cell.

[Epithelial](https://www.microscopemaster.com/epithelial-cells.html)**Stem Cells** - These are progenitor cells and are involved in the production of certain skin cells.

**Muscle Satellite Cells** - These are progenitor cells that contribute to differentiated muscle tissue.

The rapid development of stem cell biology has provided us with a strong support for further understanding of the precise molecular regulation mechanisms in the development of organisms, as well as new treatments for cancer, cardiovascular and cerebrovascular diseases, neurodegenerative diseases, diabetes, and other diseases. It brought hopes to neurological diseases. Therefore, before the therapeutic potential of stem cells is widely applied to the clinic, it is necessary to have a deeper understanding of the characteristics and regulatory mechanisms of stem cell proliferation and differentiation that determine the stem cell fate, to survive and proliferate through the endogenous cells.

Stem cells with the potential for the treatment of a wide range of degenerative disorders may be obtained from a variety of sources but, for practical reasons, some of them are more likely to find earlier clinical application than others. The main types that have been studied in the context of stem cell therapy are embryonic stem cells, fetal stem cells and adult stem cells.

Embryonic stem cells and, to a lesser extent, fetal stem cells have the potential to repair many types of tissue because they are totipotent  Embryonic stem cells can be greatly increased in number in culture as cell lines *in vitro* and may be immuno-priviledged. These attributes mean that they can be used to treat multiple patients. However, their use has been confounded by serious ethical issues  and the very real likelihood that, being immortal, they will form tumors after they have been transplanted into patients . Undoubtedly, however, these barriers to widespread application will be overcome in the future. Stem cells themselves do not serve any single purpose but are important for several reasons.

First, with the right stimulation, many stem cells can take on the role of any type of cell, and they can regenerate damaged tissue, under the right conditions.

Aging is an unavoidable physiological consequence of the living animals. Mammalian aging is mediated by the complex cellular and organismal processes, driven by diverse acquired and genetic factors[[1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5316899/%22%20%5Cl%20%22B1)]. Aging is among the greatest known risk factors for most human diseases[[2](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5316899/%22%20%5Cl%20%22B2)-[5](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5316899/#B5)], and of roughly 150000 people who die each day across the globe, about two thirds die from age-related causes

### There are two major theories of organismal aging: program and damage-based

### Program theories

Programmed aging theories, sometimes referred to as active or adaptive aging theories, suggest that there is a deliberate deterioration with age because a limited life span results in evolutionary benefits

For many years, programmed aging has been debated and some studies have substantiated this hypothesis. For example,  have suggested that there are mechanisms that preserve the integrity of spores of aging diploid yeast cells. Through these mechanisms, aging diploid cells that are induced to sporulate appear to lose all age-associated damage to a point that is no longer detectable, though the assumption that these findings can be extrapolated to higher organisms has been put into question Yet, though development and morphogenesis can be easily understood as programmed, as they are the end-result of a determined sequence of molecular and cellular events designed to produce a given phenotype, aging is mostly thought of as decay. If aging is indeed programmed, the purposes of such program remain unclear. Some have suggested that aging may constitute an altruistic plan, by eliminating post-reproductive age individuals, who would compete for resources, by avoiding overpopulation and by promoting adaptation through a succession of generations . The supporters of this view underscore that the similarities between the biochemical pathways that regulate aging in organisms such as yeasts, flies and mice, together with evidence consistent with programmed death in salmon and other organisms, hint at the possibility that programmed aging can occur in higher eukaryotes. Moreover, this plan could be the result of “aging genes”. Nonetheless, if this was the case, than certainly such mechanisms would be susceptible to inactivation, and, despite many gene mutations have been described as life-extending mutations none has been reported that abolishes the process of aging. It should be noted that, in some model organisms, genes have been demonstrated to play a pivotal role in aging. In fact, the first described mutation to yield a significant extension in the lifespan of Caenorhabditis elegans was in the age-I gene, which was shown to result in a 65% increase in mean lifespan and a 110% increase in maximum lifespan of this organism. Since then, many mutations that result in lifespan extension in C. elegans have been identified, most of which involving genes that are homologs of the of components of the insulin/IGF (insulin-like growth factor) pathway , namely, daf-2/daf-16  and sir2.1 , which, interestingly, have been shown to interact to extend lifespan in C. elegans .

Composed mostly of post-mitotic cells, C. elegans is one of the most widely studied model organisms. With a lifespan ranging from days to a few weeks, it has been noted that, under caloric restriction (CR) and/or crowded conditions, C. elegans can enter an alternative stasis-like developmental pathway, called dauer . This pathway consists of a developmental arrest, leading to an increased adult phase. This arrest suggests that, at least partially, aging and development are coupled in C. elegans, as well as in other invertebrates. However, in addition to the severity of the restriction (30–70% fewer calories than the control group), the degree of lifespan lengthening in C. elegans depends on numerous factors, namely, age at onset of restriction. Though providing some key insights into longevity, invertebrates are, nevertheless, distant animal models and are likely unrepresentative of human biology and physiology.

The endocrine system has also been viewed as involved in “telling the time”. Because the levels of hormones such as growth hormone (GH) and its corresponding downstream target insulinlike growth factor I (IGF-1) decline with age, the idea that such changes cause aging has been suggested a few decades ago, and, in rats, deficiency in growth hormone production (loss of function mutations at the Pit-I locus) has been linked to lifespan extension and delayed immune aging. Due to the fact that the brain regulates the endocrine system, the neuroendocrine theory of aging has emerged as the main hormone-based theory of aging, and, not surprisingly, many anti-aging products aim at restoring the levels of specific hormones in older people. Some studies have supported the idea that the insulin pathway is associated with human longevity, as individuals with mutated Prop-I gene – a pituitary transcription factor whose mutation causes dwarfism – may live longer and patients with GH and IGF-1 deficiencies have shown signs of early aging, despite actually living longer. Some have proposed that such mechanisms could be activated by decreasing cellular replication or that it may operate on the basis of antioxidant regulation. Whatever the mechanism, it is now clear that the early assumption that the aging process is driven by hormone changes that occur with age is unsubstantiated. If anything, the decrease in GH/IGF-1 signaling increases lifespan, not the contrary and, more broadly, hormonal changes may regulate aging as an indirect consequence of the developmental program. The imbalance on chemical processes caused by differential gene expression and hormonal changes may contribute to aging, but, so far, such assertions remain in the realm of speculation. Furthermore, the significant lifespan differences observed in numerous species, under identical conditions, seems to indicate that there is no pre-determined timeline for aging. Thus, under certain conditions, it may be possible to prolong or to curtail lifespan, leading to the hypothesis that aging is not predetermined, but rather the end-result of a “wear-and-tear” mechanism.

### (2) Damage theories

Evolutionary biologists may argue that aging occurs due to the absence of natural selection at the post-reproductive stage of life . Hence, aging is not programmed; instead, it is the absence of selection for maintenance Although such aging theories are subjectively appealing, as they convey a cure for aging, the accumulation of damage is a spontaneous entropy-driven process, and, as such, its kinetics can be genetically and environmentally modulated, resulting in the wide range of life-spans we observe

Among the damage theories, a prevailing idea is that of oxidative damage. Reactive oxygen species (ROS) – partially reduced intermediates of oxygen that can be radical or non-radical molecules – are generated during metabolism through a number of inter-related reactions  and are considered to lead to the cumulative DNA, protein and lipid damage observed over a lifetime. Approximately 2–3% of oxygen taken up is chemically reduced by the addition of single electrons. Incomplete reduction of oxygen can generate a variety of biologically relevant ROS such as, hydrogen peroxide, the anion radical superoxide and the hydroxyl radical. The electron transport chain in the mitochondria, the nicotinamide adenine dinucleotide phosphate oxidases (NADPH oxidase) and the 5-lipoxygenase as the three major sources of ROS in living cells . Multiple studies have highlighted the relatively haphazard molecular damage that ROS cause to lipids, proteins and nucleic acids and exposure to ROS have been demonstrated to trigger specific mechanisms aimed at neutralizing their effects

Adult stem cells, also known as somatic stem cells, are found throughout the body in every tissues and organs after development and function as self-renewing cell pools to replenish dying cells and regenerate damaged tissues throughout life. However, adult stem cells appear to age with the person. As stem cells age, their functional ability also deteriorates. Specifically, this regenerative power appears to decline with age, as injuries in older individuals heal more slowly than in childhood. For example, healing of a fractured bone takes much longer time in elderly than in young individuals. There is a substantial amount of evidence showing that deterioration of adult stem cells in the adult phase can become an important player in the initiation of several diseases in aging. The following is some of the examples of aging-associated effects on stem cells.

Neural stem cells (NSCs) are multipotent and self-renewing cells and located primarily in the neural tissues. In response to a complex combination of signaling pathways, NSCs differentiate into various specific cell types locally in the central nervous system (CNS), like neurons, astrocytes, and oligodendrocytes. NSCs in humans maintain brain homeostasis and it continuously replenishes new neurons, which are important for cognitive functions. However, there is now strong evidence for the aging-associated cognitive deficits, such as olfactory dysfunction, spatial memory deficits, and neurodegenerative disorders, which are caused by deterioration of NSC proliferation and differentiation and enhanced NSC senescence as a consequence of aging

Mesenchymal stem cells (MSCs) are multipotent stromal cells that can differentiate into cells of mesenchyme tissues, including osteoblasts (bone cells), chondrocytes (cartilage cells), myocytes (muscle cells) and adipocytes (fat cells). MSCs were first isolated from the bone marrow of guinea pigs in 1970’s and after that it was isolated from almost every organ in mice including fat, liver, spleen, pancreas, kidney, lung, muscle, and brain. Human MSCs have also been isolated from umbilical cord tissue and cord blood, placenta, bone and joints. However, the major sources of MSCs are the bone marrow-derived MSCs (BM-MSCs) and the adipose tissue-derived MSCs (A-MSCs); and they are currently the most studied MSCs. Aging also affects MSCs in humans and in animal models as indicated by the decrease in the bone marrow MSC pool and also shifts their lineage differentiation from one that usually favors osteoblastic differentiation to one that prefers adipogenic differentiation, which is largely responsible for the gradual and aging-associated shift of hematopoietic (red) marrows to fatty (yellow) marrows, and which also contributes significantly to the etiology of senile osteoporosis. It is also evident that with increasing donor age, MSCs from both bone marrow and adipose tissues have been shown to have reduced capacity to handle oxidative stress. During the aging process, oxidative stress leads to hyperactivity of pro-growth pathways, such as insulin/IGF-1 and mTOR pathways, and the subsequent accumulation of toxic aggregates and cellular debris ultimately lead to apoptosis, necrosis, or autophagy. In addition, in some non-skeletal tissues, particularly the hematopoietic system, MSCs is a key niche component for hematopoietic cells. Aging of MSCs has been shown to be detrimental with respect to this important function.

Adult skeletal muscle stem cells (satellite cells) have a remarkable capacity to regenerate. Similarly, their regeneration capacity declines with aging, although it is not clear whether this is due to extrinsic changes in the environment and/or to cell-intrinsic mechanisms associated to aging. This impaired regenerative capacity of skeletal muscle during aging is due to accumulation of the altered progeny, which leads to progressive deterioration of tissue structure and function, manifesting after injury or in response to the depletion of memory B cells and naive T cells in the hematopoietic system in the elderly.

Hematopoietic stem cells (HSCs) are the blood-forming stem cells through the process of hematopoiesis. They are located in the red bone marrow within marrow cavity of most bones. HSCs also produce immune cells of the body. Since blood cells are responsible for constant maintenance and immune protection of every cell type of the body, the constant production of billions of new blood cells each day by HSCs is very important for mammal life. HSC-derived monocytes can give rise to osteoclasts, macrophage and granulocyte. Osteoclasts are giant cells with numerous nuclei that work in synergy with osteoblasts through complicated bone coupling mechanisms to maintain bone homeostasis. All these activities of HSCs are carefully modulated by a complex interplay between cell-intrinsic mechanisms and cell-extrinsic factors produced by the microenvironment; and aging altered this fine-tuned regulatory network, leading to aberrant HSC cell cycle regulation, degraded HSC function, and hematological malignancy.

There are several potential mechanisms that are believed to contribute to the aging-associated stem cell dysfunction; and they probably are in part responsible for many aging-associated diseases. The functions of aged stem cells become impaired as the result of cell-intrinsic pathways and surrounding environmental changes. With the sharp rise in the aging-associated diseases, the need for effective regenerative medicine strategies for the aged is more important than ever.

**Molecular and genetic basis of immunity.**

Major histocompatibility complex. Humoral and cellular immunity. Antibodies. Cytokines, interferons and the complement system.

*Maximal point: 7*

1. Describe the main histocompatibility complex and its role in human immunity.

2. Explain what humoral and cellular immunity is.

3. Classify and characterize proteins involved in humoral and cellular immunity.

4. Describe congenital and acquired disorders of human immunity.

Every multicellular organism, including our own, constantly has to be on guard not to be gobbled up by others, as it constitutes a potential source of valuable organic molecules. The ability to resist being used as "food" automatically confers a selective advantage. Over the course of evolution, this has led to the development of highly sophisticated defense systems in multicellular organisms.

To maintain the integrity of our organism, it is essential to distinguish between biological structures that have to be fought off –ideally, everything that poses a danger to our organism—and structures that must not be attacked, e.g., the cells of our own body, or useful bacteria in our gut. This problem is not at all trivial, as dangerous attackers from the worlds of viruses, bacteria and parasites consist of largely the same molecules as the human body.

Immunity can be defined as a complex biological system endowed with the capacity to recognize and tolerate whatever belongs to the self, and to recognize and reject what is foreign (non-self).

**Antigen**, substance that is capable of stimulating an immune response, specifically activating lymphocytes, which are the body’s infection-fighting [white blood cells](https://www.britannica.com/science/white-blood-cell). In general, two main divisions of antigens are recognized: foreign antigens (or heteroantigens) and autoantigens (or [self-antigens](https://www.britannica.com/science/self-antigen)). Foreign antigens originate from outside the body. Examples include parts of or substances produced by [viruses](https://www.britannica.com/science/virus) or microorganisms (such as [bacteria](https://www.britannica.com/science/bacteria) and [protozoa](https://www.britannica.com/science/protozoan)), as well as substances in snake venom, certain [proteins](https://www.britannica.com/science/protein) in foods, and components of serum and [red blood cells](https://www.britannica.com/science/red-blood-cell) from other individuals. Autoantigens, on the other hand, originate within the body. Normally, the body is able to distinguish self from nonself, but in persons with [autoimmune disorders](https://www.britannica.com/science/autoimmunity), normal bodily substances provoke an immune response, leading to the generation of autoantibodies. An antigen that induces an [immune response](https://www.britannica.com/science/antigen-antibody-reaction)—i.e., stimulates the lymphocytes to produce [antibody](https://www.britannica.com/science/antibody) or to attack the antigen directly—is called an [immunogen](https://www.britannica.com/science/immunogen). On the surface of antigens are regions, called antigenic determinants, that fit and bind to [receptor](https://www.britannica.com/science/receptor-nerve-ending) molecules of complementary structure on the surface of the lymphocytes. The binding of the lymphocytes’ receptors to the antigens’ surface molecules stimulates the lymphocytes to multiply and to initiate an immune response—including the production of [antibody](https://www.britannica.com/science/antibody), the activation of cytotoxic cells, or both—against the antigen. The amount of antibody formed in response to stimulation depends on the kind and amount of antigen involved, the route of entry to the body, and individual characteristics of the host.

Molecular and cellular components make up the immune system. The function of these components is divided up into nonspecific mechanisms, those which are **innate** to an organism, and responsive responses, which are **adaptive** to specific pathogens. These two systems work closely together and take on different tasks.

The innate immune system is the phylogenically oldest component of the human immune system.

The innate immune system is the body's first line of defense against germs entering the body. It responds in the same way to all germs and foreign substances, which is why it is sometimes referred to as the "nonspecific" immune system. It acts very quickly. The innate immune system has only limited power to stop germs from spreading, though. The innate immune system is always general, or nonspecific, meaning anything that is identified as foreign or non-self is a target for the innate immune response.

The innate immune system consists of

* Protection offered by the skin and mucous membranes
* Protection offered by the immune system cells (defense cells) and proteins

Defects in innate immunity are associated with invasive, life-threatening infection. Inappropriate activation of the innate immune system can lead to autoinflammatory states. The innate immune system directs the subsequent development of [adaptive immune responses](https://www.sciencedirect.com/topics/immunology-and-microbiology/adaptive-immune-system)

The adaptive immune system, also called acquired immunity, uses specific antigens to strategically mount an immune response. Unlike the innate immune system, which attacks only based on the identification of general threats, the adaptive immunity is activated by exposure to pathogens, and uses an immunological memory to learn about the threat and enhance the immune response accordingly. The adaptive immune response is much slower to respond to threats and infections than the innate immune response, which is primed and ready to fight at all times.

 The adaptive immune system is made up of:

* T lymphocytes in the tissue between the body's cells
* B lymphocytes, also found in the tissue between the body's cells
* Antibodies in the blood and other bodily fluids

The major proteins of the immune system are predominantly signaling рroteins (often called cytokines), antibodies, and complement proteins.

The cells of the immune system are represented by a variety of cell populations of lymphocytes, macrophages, microphages and platelets. The total weight of immune cells in an adult is about 1 kg.

The cells of the immune system can be categorized as lymphocytes (T-cells, B-cells and NK cells), neutrophils, and monocytes/macrophages. These are all types of white blood cells.

**Characteristics of immune system**

There are 4 main properties of immunity: existence of

1- cell immunity

2-humoral immunity

3-passive immunity

4 –ability to individual “education” of immune system

## Immune system disorders

The immune system is a complex and highly developed system, yet its mission is simple: to seek and kill invaders

The [immune system](https://www.britannica.com/science/immune-system) protects against [infectious disease](https://www.britannica.com/science/infectious-disease), but it may also at times cause [disease](https://www.britannica.com/science/disease). Disorders of the immune system fall into two broad categories: (1) those that arise when some aspect of the host’s immune mechanism fails to prevent infection (immune deficiencies) . Immune deficiencies could be primary and acquired

(2) those that occur when the immune response is directed at an inappropriate [antigen](https://www.britannica.com/science/antigen), such as a noninfectious agent in an allergic reaction, the body’s own antigens in an autoimmune response, or the cells of a transplanted organ in graft rejection.

**The human genome. Part I, II**

History of the Human Genome Project. The structure of the human genome: protein-coding genes and non-coding DNA. Satellite DNA. Tandem repeats. Single nucleotide polymorphisms (SNP). Transposed elements of the genome: transposons, retrotransposons.

1. Describe the structure of the human genome: protein-coding genes, intergenic regions (spacers), satellites, tandem repeats, single nucleotide polymorphisms (SNPs).

2. Explain the role of non-coding DNA in the human genome.

3. Discuss the prospects for applying knowledge about the human genome in medicine and pharmaceuticals.

4. Describe DNA transposons, retrotransposons, retroviral integration

Methods for the study of nucleic acids and proteins. Proteomic methods of analysis. Bioinformation databases. DNA diagnostics: polymerase chain reaction, restriction analysis, FISH hybridization. Linked immunosorbent assay. Bioethics of genetic experiments with humans.

1. Give the definitions of genomics, proteomics and bioinformatics, describe their research methods.

2. Explain the Sanger, Maxam-Gilbert, NGS (New Generation Sequencing) and other methods of genome sequencing.

3. Characterize and analyze the main methods of protein research: two-dimensional gel electrophoresis, mass spectrometry, chromatography, X-ray structural analysis, nuclear magnetic resonance.

4. Describe EMBL-EBI, DDJB, NCBI, PIR, MIPS, NBRF, SwissProt, UniProt and other bioinformatical databases.

5. Give the definition of molecular diagnostics and describe its various methods.

**Human Genome Project (HGP)**, an international collaboration that successfully determined, stored, and rendered publicly available the sequences of almost all the genetic content of the [chromosomes](https://www.britannica.com/science/chromosome) of the human organism, otherwise known as the [human genome](https://www.britannica.com/science/human-genome).

The Human Genome Project (HGP) refers to the international 13-year effort, formally begun in October 1990 and completed in 2003, to discover all the estimated 20,000-25,000 human genes and make them accessible for further biological study.

Project [goals](https://web.ornl.gov/sci/techresources/Human_Genome/hg5yp/index.shtml)

* *identify* all the approximately 20,000-25,000 genes in human DNA,
* *determine* the sequences of the 3 billion chemical base pairs that make up human DNA,
* *store* this information in databases,
* *improve* tools for data analysis,
* *transfer* related technologies to the private sector, and
* *address* the ethical, legal, and social issues (ELSI) that may arise from the project.

 As part of the HGP, parallel studies were carried out on selected model organisms such as the bacterium E. coli and the mouse to help develop the technology and interpret human gene function.

In 1988, Congress appropriated funds to the Department of Energy (DOE) and the National Institutes of Health (NIH) to begin planning the Human Genome Project. Planners set a 15-year time frame, estimated that the price tag would be $3 billion, and laid out formal goals to get the job done.[**1**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106) On October 1, 1990, the Human Genome Project officially began.[**2**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106) According to early plans, the human race would witness its own blueprint in fine detail in the year 2005.

 In the fall of 1998, however, improvements in technology, success in achieving early mapping goals, emerging research opportunities, and a growing demand for the human DNA sequence prompted project leaders in the United States and abroad to promise the blueprint — the complete DNA sequence of the human genome — two years ahead of schedule, in 2003.  Indeed, only six months later, in March 1999, 15 percent of the sequence was in a finished or nearly finished state. The largest centers participating in the Human Genome Project received new grants to begin full-scale sequencing of the human genome, and the timetable was moved up yet again. Pilot sequencing projects had been so successful that the planners of the Human Genome Project now felt confident that at least 90 percent of the human sequence could be completed in “working draft” form by the spring of 2000, considerably earlier than expected.

 Until the entire human genome sequence has been completed, in 2003 or perhaps earlier, the working-draft sequence will be very useful, especially for finding genes, exons, and other genomic features. But because the working draft will contain gaps and will not be entirely accurate, it will not be as useful as the finished sequence for studying DNA features that span large regions or require a high degree of accuracy over long stretches of the sequence. The HGP was further intended to improve the technologies needed to interpret and analyze genomic sequences, to identify all the [genes](https://www.britannica.com/science/gene) encoded in human DNA, and to address the [ethical](https://www.merriam-webster.com/dictionary/ethical), legal, and social [implications](https://www.merriam-webster.com/dictionary/implications) that might arise from defining the entire human genomic sequence.

 The final product must have four characteristics — the four A's of the Human Genome Project.

 First, the sequence must be*accurate* — that is, the DNA spellings must have an accuracy of 99.99 percent or better. The HGP was further intended to improve the technologies needed to interpret and analyze genomic sequences, to identify all the [genes](https://www.britannica.com/science/gene) encoded in human DNA, and to address the [ethical](https://www.merriam-webster.com/dictionary/ethical), legal, and social [implications](https://www.merriam-webster.com/dictionary/implications) that might arise from defining the entire human genomic sequence. Second, large-scale sequencing requires that the shorter lengths of sequenced DNA be accurately *assembled* into longer, genomic-scale pieces that reflect the original genomic DNA. Third, the human DNA sequence must also be *affordable*, and technology development will aim to reduce the cost as much as possible. Finally, the high-quality, finished human DNA sequence should be *accessible* within 24 hours through public data bases.

**Implications for Understanding Genetic Illness**

Maps and other forms of genome technology provide the tools for a gene-isolation technique known as positional cloning.[**9**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106) This technique allows a researcher to confirm the genetic basis of a disease and identify the responsible gene, even when little is known about the gene's function. So far, over 100 disease-linked genes have been isolated with the use of the positional-cloning technique. Whereas gene discovery by this route once took years to decades, an investigator using these powerful tools can now sometimes map and isolate a gene in a matter of weeks. Increasingly, gene hunters are combining positional-cloning techniques with information in EST data bases to narrow their gene searches to rational candidates. This method, called positional candidate cloning,**[10](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** has been used to isolate many altered genes associated with human disease.

Gene isolation provides the best hope for understanding human disease at its most fundamental level (. Knowledge about genetic control of cellular functions will underpin future strategies to prevent or treat disease phenotypes. The recent isolation of genes for Parkinson's disease,**[11-14](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** for example, has greatly advanced molecular research on this baffling disease. In one study of families with early-onset Parkinson's disease, gene hunters mapped a suspect gene to a region of chromosome 4.[**11**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106) The region contained approximately 100 genes, among which was 1 known to encode the protein α-synuclein. Earlier research had shown that α-synuclein accumulates in brain cells of people with Alzheimer's disease,**[15](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** and people with Parkinson's disease have similar deposits in the substantia nigra. In just a few months, the researchers showed conclusively that a missense mutation in the α-synuclein gene caused Parkinson's disease in the study families. Further research has shown that a mutation in a gene encoding a protein critical to the breakdown of α-synuclein and other proteins also results in the Parkinson's disease phenotype in a different family.**[14](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** Although most cases of Parkinson's disease appear to have limited heritability, studying rare families of this sort provides crucial clues to the pathways involved — α-synuclein is found in the Lewy bodies in virtually all cases of Parkinson's disease. An understanding of the genetic control of the proteolytic processes of brain proteins may provide new targets for interventions in a number of related neurodegenerative disorders characterized by the accumulation of protein deposits, including Alzheimer's disease, Huntington's disease, and spinocerebellar ataxia.

Even before a gene's role in disease is fully understood, diagnostic applications can be useful in preventing or minimizing the development of health consequences. DNA tests that look for the presence of disease-linked mutations, for example, are proving to be the most immediate commercial application of gene discovery and the one now used most frequently by clinicians. These tests may help establish the diagnosis of a genetic disease, foreshadow the development of disease later in life, or identify healthy heterozygote carriers of recessive diseases. Genetic tests can be performed at any stage of the human life cycle, and the sampling procedures are becoming less invasive. Whereas genetic testing was once sought almost exclusively by couples with a family history of early-onset disease, for the purpose of family planning, information about genetic status is increasingly sought by persons who wish to learn about their own predisposition to adult-onset illness.

In a growing number of instances, strategies can be implemented to reduce or prevent illness when a genetic cause or predisposition is known. Successes in reducing disease through treatment have been achieved for the hereditary disorders hemochromatosis, phenylketonuria, and familial hypercholesterolemia, among others. Risk reduction through early detection and lifestyle changes may be possible in the case of disorders associated with predisposing mutations, such as some cancers. As therapies build on knowledge gained about the molecular basis of disease, increasing numbers of illnesses that are now refractory to treatment may yield to molecular medicine in the future.

The recent discovery of an altered gene (*HFE*) that leads to hereditary hemochromatosis,**[16](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** a common disorder of iron metabolism, provides an interesting example of the potential for using information about mutations to prevent an adult-onset disease phenotype. A recessive condition, hereditary hemochromatosis affects about 1 in 300 persons of northern European descent and is easily treatable if diagnosed early. Its major symptoms — liver cirrhosis, heart failure, diabetes, arthritis, and other organ damage — do not occur until midlife and are easily misdiagnosed. Untreated, the disease causes early death, but treatment by phlebotomy to remove excess iron allows people with hereditary hemochromatosis to live a normal life span. A single substitution of the amino acid tyrosine for cysteine at codon 282 accounts for the majority of cases.**[17](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)**

At first glance, hereditary hemochromatosis seems to be an ideal target for public health approaches to the prevention of hereditary disease: the disorder is common, the number of disease-linked mutations in the gene are few, and effective treatment can minimize or eliminate the effects of the disease. But closer examination reveals a number of complexities that have so far militated against the rapid introduction of this genetic test as a tool for disease prevention.**[16](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** Because the penetrance of the altered *HFE* gene is reduced, especially in females, clinical signs can range from none that are detectable to severe organ damage from iron overload. At the moment, simple detection of the mutation does not predict the most likely clinical course. Before population testing for *HFE* mutations is considered, further research is needed to explain the variations in phenotype among mutation carriers and to correlate the genotype more closely with health outcomes.

The rather straightforward mendelian rules that govern the inheritance of disease traits have been worked out for many rare disorders that result from highly penetrant changes in a single gene. But teasing out the genetic components of the so-called complex disorders — diabetes, heart disease, most common cancers, autoimmune disorders, and psychiatric disorders — that result from the interplay of environment, lifestyle, and the small effects of many genes remains a formidable task. Most of the successful efforts to identify genes associated with common diseases have focused on highly heritable subgroups, including the *BRCA1* **[18](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** and *BRCA2* **[19](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** genes in breast cancer, the gene for hepatocyte nuclear factor 4α (*HNF-4* α) in maturity-onset diabetes of the young (MODY) type 1,**[20](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** the gene for glucokinase (*GCK* ) in MODY type 2,**[21](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** the gene for hepatocyte nuclear factor 1α (*HNF-1*α) in MODY type 3,**[22](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** the gene for human Mut S homologue 2 (*hMSH2*)**[23,24](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** and the gene for human Mut L homologue 1 (*hMLH1* )**[25](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** in hereditary nonpolyposis colon cancer, and the gene for α-synuclein in Parkinson's disease.**[11](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** Linkage analysis and positional-cloning techniques are well suited to discovering genes with such strong influences. But these strategies are not as easily applied to the multiple, low-penetrance variants, which in the aggregate account for a larger percentage of illnesses. Identification of weakly penetrant alleles that contribute to common disorders requires new and more powerful approaches.

To assist in these efforts, the Human Genome Project is initiating new studies of genetic variation in the human population to provide a dense map of common DNA variants. DNA sequence variations include insertions and deletions of nucleotides, differences in the copy number of repeated sequences, and single-nucleotide polymorphisms, or SNPs (pronounced “snips”), which occur most frequently throughout the human genome. About 1 in every 300 to 500 bases in human DNA may be a SNP.

SNPs can be used as markers in whole-genome linkage analysis of families with affected members, as well as in association studies of individuals in a population. Association studies may directly test a variant with potential functional importance or may take advantage of the phenomenon of linkage disequilibrium — in which a marker and a gene are inherited together — to map gene variants associated with disease. Because the human species consists of relatively few generations, recombination events have not disrupted linkage disequilibrium over distances of 3000 to 100,000 bases in most populations. Consequently, association studies view large human populations as evolutionary families and do not rely on studies of nuclear families for gene mapping.[**26,27**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)

Some SNPs may contribute directly to a trait or disease phenotype by altering function. Though most SNPs are located outside protein-coding sequences, those within coding sequences, called cSNPs, are of particular interest because they are more likely to affect gene function. A large, well-characterized collection of SNPs will become increasingly important for the discovery of DNA sequence variations that affect biologic function. Work is already under way with NIH support to develop a catalogue of 60,000 or more SNPs. A recently formed pharmaceutical consortium will support the production of 300,000 more, with the work being done at the publicly funded genome centers and all the data deposited in the public domain. This is a wonderful example of a public–private partnership to develop a powerful set of research tools that all can use.

## New Forms of Technology for Genetic Analysis and Risk Assessment

The transition from genetics to genomics marks the evolution from an understanding of single genes and their individual functions to an understanding of the actions of multiple genes and their control of biologic systems. Whereas the tools of the Human Genome Project initially advanced research on single genes, they are now forming the basis for genomic-scale analysis of the human organism.

The so-called DNA chip currently provides one promising approach to genome-scale studies of genetic variation,**[28](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** detection of heterogeneous gene mutations,[**29**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106) and gene expression.[**30**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106) The result of an adaptation of dot blot hybridization techniques, DNA chips, also called microarrays, generally consist of a thin slice of glass or silicon about the size of a postage stamp on which threads of synthetic nucleic acids are arrayed. Sample probes are added to the chip, and matches are read by an electronic scanner. As with semiconductors, the capacity of DNA chips has doubled about every two years, so chips that held a few hundred arrays not so long ago now hold hundreds of thousands.

Microarray technology has been applied to the detection of DNA variations as well as expression of messenger RNA in individual cells and tissues. Microarrays are used clinically to detect human immunodeficiency virus sequence variations, p53 gene mutations in breast tissue, and expression of cytochrome P450 genes. In the laboratory, microarray technology has also been applied to genomic comparisons across species,**[31](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** genetic recombination,[**32**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106) and large-scale analysis of gene copy number and expression, as well as protein expression, in cancerous tissues.[**33**](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)

Use of microarrays and other new technologies to detect DNA variations holds promise, along with family histories and data from large population studies, for establishing a person's risk of contracting common, adult-onset disorders. A base-line genome scan could provide helpful information about a person's risk profile and point to the prevention strategies — if available — that should be used.

## Genetic Knowledge and Individualized Medicine

Identifying human genetic variations will eventually allow clinicians to subclassify diseases and adapt therapies to the individual patient. There may be large differences in the effectiveness of medicines from one person to the next. Toxic reactions can also occur and in many instances are likely to be a consequence of genetically encoded host factors. That basic observation has spawned the burgeoning new field of pharmacogenomics, which attempts to use information about genetic variation to predict responses to drug therapies.

For example, researchers discovered that patients with Alzheimer's who have the ε4 subtype of the gene for apolipoprotein E (*APOE* ε*4*), which affects cholinergic function in the brain, are less likely to benefit from the cholinomimetic drug tacrine than are patients without this subtype.**[34](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** This finding will help in the analysis of data from clinical trials of Alzheimer's therapies and will promote the development of new therapies specifically designed for *APOE* ε*4* carriers.

In another example, cholesteryl ester transfer protein (CETP) plays an important part in the metabolism of high-density lipoprotein, a lipoprotein associated with lowered susceptibility to atherosclerosis. A certain genetic variant of the *CETP* gene is correlated with higher plasma CETP levels and lower levels of plasma high-density lipoprotein. One study showed that in men who carried this genetic variant, treatment with pravastatin slowed the progression of coronary atherosclerosis.**[35](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** This finding may allow physicians to predict which patients with coronary artery disease will benefit from treatment with pravastatin.

In a third example, the formation of venous blood clots in the brain and legs is a rare but serious side effect of birth-control pills. One study has shown a dramatically increased risk of cerebral-vein thrombosis among women taking oral contraceptives who also carry the blood-clotting variant factor V Leiden.**[36](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** The risk of other venous thrombotic events is also increased in this group. Foreknowledge of the presence of this variant and consideration of alternative forms of birth control might be useful in minimizing the risk of thrombosis in these women.

Not only will genetic tests predict responsiveness to drugs on the market today, but also genetic approaches to disease prevention and treatment will include an expanding array of gene products for use in developing tomorrow's drug therapies. Since the Food and Drug Administration's approval of recombinant human insulin in 1982, over 50 additional gene-based drugs have become available for clinical use. These include drugs for the treatment of cancer, heart attack, stroke, and diabetes, as well as many vaccines.**[37](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)**

Not all therapeutic advances for gene discovery will be genes or gene products. In other instances, molecular insights into a disorder, derived from gene discovery, will suggest a new treatment. Sodium phenylbutyrate, for example, which is approved for the regulation of blood ammonia levels, is being tested in clinical trials for the treatment of cystic fibrosis.**[38](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** The main clinical phenotype in people with cystic fibrosis results from a mutation in the gene for cystic fibrosis transmembrane conductance regulator (CFTR) protein. The mutation prevents normal amounts of CFTR protein from crossing the cell membrane, diminishing the ability of chloride and water to enter and exit the cell. Sodium phenylbutyrate apparently stimulates expression of CFTR protein, allowing more of it to reach the correct location.

## Ethical, Legal, and Social Implications

## One of the most active areas of the ELSI program has been policy development related to the privacy and fair use of genetic information, particularly in health insurance, employment, and medical research. Debates in this area focus largely on the potential of genetic information to predict an increased likelihood of the eventual development of a disease phenotype in a currently healthy person.

Although many states have attempted to address “genetic discrimination” in health insurance and employment, federal legislation would provide the most comprehensive protection. Concern about the confidentiality of genetic information may make people reluctant to volunteer for studies involving disease-linked gene mutations or genetic therapy, for fear that the results could result in the loss of a job or the loss of insurance coverage.

Largely on the basis of recommendations formulated in workshops held by the Human Genome Project and the National Action Plan on Breast Cancer,**[39,40](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** the Clinton administration endorsed the need for congressional action to protect against genetic discrimination in health insurance and employment. In 1996, Congress enacted the Health Insurance Portability and Accountability Act, which represented a large step toward protecting access to health insurance in the group-insurance market but left several serious gaps in the individual-insurance market that must still be closed.

In the area of workplace discrimination, the Equal Employment Opportunity Commission has interpreted the Americans with Disabilities Act as covering on-the-job discrimination based on “genetic information relating to illness, disease or other disorders.”**[41](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** But no claims of genetic discrimination have been brought to the commission, and the guidance has yet to be tested in court, so the degree of protection actually provided by the act remains uncertain.

In the area of privacy, as part of the partnership between the National Action Plan on Breast Cancer and the Human Genome Project, medical researchers, policy makers, and representatives of law, government, the insurance industry, and public health have recently assessed current policies and practices designed to protect confidentiality in genetics research and have identified areas where new or modified policies or practices might enhance the protection of privacy and promote the conduct of research. The group is developing a set of principles for researchers, research institutions, state and federal agencies, and policy makers to consider in formulating measures to protect the privacy of research information.

Other important steps have been taken to ensure the responsible integration of genetic tests into clinical practice. For the most part, genetic testing in the United States has developed successfully, providing options for avoiding, preventing, and treating inherited disorders. But the rapid pace of test development combined with the rush to market new products may create an environment in which the tests are made available before they have been adequately validated. On the recommendation of the Task Force on Genetic Testing,**[42](https://www.nejm.org/doi/full/10.1056/nejm199907013410106)** assembled by the Human Genome Project's NIH–DOE ELSI Working Group, the Secretary of the Department of Health and Human Services has established an advisory panel to ensure the safe introduction of genetic tests into clinical practice.

Completion of the first human-genome sequence and the expansion of human genetic research to include studies of genetic variation among subpopulations have raised new questions about ethical, legal, and social issues. The 1998–2003 plan includes an examination of these issues as well as the integration of genetic technology and information into health care and public health activities; the use of knowledge about genomics and gene–environment interactions in nonclinical settings; examination of a variety of philosophical, theological, and ethical perspectives on new genetic knowledge; and consideration of the ways in which racial, ethnic, and socioeconomic factors affect the use, understanding, and interpretation of genetic information, the use of genetic services, and the development of policy.

 **Genome structure**

Humans have two genomes, nuclear and mitochondrial. Normal diploid cells contain two copies of the nuclear genome and a much larger but variable number of copies of the mitochondrial genome.

The human haploid genome consists of about 3 x 109 base pairs of DNA. Although the sequence of the human genome has been (almost) completely determined by DNA sequencing, it is not yet fully understood. Most (though probably not all) [genes](https://en.wikipedia.org/wiki/Gene) have been identified by a combination of high throughput experimental and [bioinformatics](https://en.wikipedia.org/wiki/Bioinformatics) approaches, yet much work still needs to be done to further elucidate the biological functions of their protein and [RNA](https://en.wikipedia.org/wiki/RNA) products. Recent results suggest that most of the vast quantities of noncoding DNA within the genome have associated biochemical activities, including [regulation of gene expression](https://en.wikipedia.org/wiki/Regulation_of_gene_expression), organization of [chromosome architecture](https://en.wikipedia.org/wiki/Chromatin), and signals controlling [epigenetic inheritance](https://en.wikipedia.org/wiki/Epigenetic_inheritance).Genomic DNA exists as single linear pieces of DNA that are associated with a protein called a nucleoprotein complex. The DNA-protein complex is the basis for the formation of chromosomes, virtually all of the genomic DNA is distributed among the 23 chromosomes that reside in the cellular nucleus.

Human [genomes](https://en.wikipedia.org/wiki/Genome) include both protein-coding DNA genes and [noncoding DNA](https://en.wikipedia.org/wiki/Noncoding_DNA). [Protein](https://en.wikipedia.org/wiki/Protein%22%20%5Co%20%22Fertilization)-coding sequences account for only a very small fraction of the genome (approximately 1.5%), and the rest is associated with [non-coding RNA](https://en.wikipedia.org/wiki/Non-coding_RNA%22%20%5Co%20%22Non-coding%20RNA) genes, [regulatory DNA sequences](https://en.wikipedia.org/wiki/Regulatory_sequences%22%20%5Co%20%22Regulatory%20sequences), [LINEs](https://en.wikipedia.org/wiki/LINEs%22%20%5Co%20%22LINEs), [SINEs](https://en.wikipedia.org/wiki/Short_interspersed_nuclear_element%22%20%5Co%20%22Short%20interspersed%20nuclear%20element), [introns](https://en.wikipedia.org/wiki/Intron%22%20%5Co%20%22Intron), and sequences for which as yet [no function](https://en.wikipedia.org/wiki/Noncoding_DNA%22%20%5Co%20%22Noncoding%20DNA) has been determined

The content of the human genome is commonly divided into coding and noncoding DNA sequences. [Coding DNA](https://en.wikipedia.org/wiki/Coding_region%22%20%5Co%20%22Coding%20region) is defined as those sequences that can be transcribed into [mRNA](https://en.wikipedia.org/wiki/MRNA%22%20%5Co%20%22MRNA) and [translated](https://en.wikipedia.org/wiki/Translation_%28biology%29%22%20%5Co%20%22Translation%20%28biology%29) into proteins during the human life cycle; these sequences occupy only a small fraction of the genome (<2%). [Noncoding DNA](https://en.wikipedia.org/wiki/Noncoding_DNA%22%20%5Co%20%22Noncoding%20DNA) is made up of all of those sequences (ca. 98% of the genome) that are not used to encode proteins.

Some noncoding DNA contains genes for RNA molecules with important biological functions ([noncoding RNA](https://en.wikipedia.org/wiki/Noncoding_RNA%22%20%5Co%20%22Noncoding%20RNA), for example [ribosomal RNA](https://en.wikipedia.org/wiki/Ribosomal_RNA%22%20%5Co%20%22Ribosomal%20RNA) and [transfer RNA](https://en.wikipedia.org/wiki/Transfer_RNA%22%20%5Co%20%22Transfer%20RNA)). The exploration of the function and evolutionary origin of noncoding DNA is an important goal of contemporary genome research, including the [ENCODE](https://en.wikipedia.org/wiki/ENCODE%22%20%5Co%20%22ENCODE) (Encyclopedia of DNA Elements) project, which aims to survey the entire human genome, using a variety of experimental tools whose results are indicative of molecular activity.

Because non-coding DNA greatly outnumbers coding DNA, the concept of the sequenced genome has become a more focused analytical concept than the classical concept of the DNA-coding gene.

One of the main open questions in the field of the [origin of life](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/biogenesis%22%20%5Co%20%22Learn%20more%20about%20Biogenesis%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages) is the biogenesis of proteins and [nucleic acids](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/nucleic-acids%22%20%5Co%20%22Learn%20more%20about%20Nucleic%20Acids%20from%20ScienceDirect%27s%20AI-generated%20Topic%20Pages) as ordered sequences of monomeric residues, possibly in many identical copies. The days of peering down the microscope to detect chromosome abnormalities are gone, replaced by chromosome analysis at the genomic level.

* Polymerase chain reaction (PCR)
* Sanger sequencing
* Southern blotting
* Multiplex ligation probe amplification (MLPA)
* Array comparative genomic hybridisation (array CGH)
* Karyotyping
* Fluorescent in situ hybridisation (FISH)
* Quantitative fluorescent PCR (QF-PCR)
* Single nucleotide polymorphism (SNP) genotyping and genome wide association studies (GWAS)
* The extraction and analysis of cell free fetal DNA, including non-invasive prenatal testing (NIPT).

**Molecular biomedicine.**

The use of genetic engineering in the production of vaccines and drugs. Genome Editing Technologies (CRISPR-Cas9). *Ex vivo* and *in vivo* gene therapy. Technologies for targeted delivery of drugs and gene therapy vectors: liposomes, dendrimers, aptamers, nanoparticles, genetically modified viruses, etc. Quantum dots. Prospects for the use of nanorobots in medicine. Pharmacogenomics. Pharmacogenetics.

1. Describe recombinant DNA technology.

2. Discuss about perspectives and dangers of creating the genetically modified organisms.

3. Describe the use of genetic engineering in the production of vaccines and drugs.

4. Explain the principles of CRISPR-Cas9 technology.

*5.* Explain what gene therapy is ex vivo and in vivo, analyze the problems and prospects of genomic technologies in medicine.

6. Give definitions of nanotechnology and bionanotechnology.

7. Describe and provide examples of various bionanotechnologies for targeted delivery of drugs and gene therapy vectors into the cells of the human body.

8. Analyze bionanotechnological methods for the diagnosis and treatment of cancer: quantum dots, magnetic and radioactive nanoparticles, etc.

9. Analyze the prospects for the use of nanorobots in biomedicine.

10. Give definitions and explain the difference between the terms "pharmacogenomics", "pharmacogenetics", "personalized medicine".

11. Explain how a hereditary predisposition can affect the individual reactions of the human body to drugs and dietary supplements, give specific examples.

Molecular Biomedicine is a branch of the Life Sciences that overlaps with Molecular Biology in many ways. It strives to elucidate molecular disease principles by studying the genome, transcriptome, proteome and metabolome, and their components, under healthy and disease conditions. It exploits our understanding of molecular biological and physiological mechanisms to devise new technologies for disease prediction, diagnosis and therapy, thereby improving healthcare and clinical practice and advancing the field of personalized medicine.

Genetic engineering, also called genetic modification, is the direct manipulation of an organism’s genome using biotechnology.Genetic engineering means the manipulation of organisms to make useful products and it has broad applications.

* Genetic engineering has applications in medicine, research, industry and agriculture and can be used on a wide range of plants, animals and microorganisms.
* In medicine, genetic engineering has been used to mass-produce insulin, human growth hormones, follistim (for treating infertility), human albumin, monoclonal antibodies, antihemophilic factors, vaccines, and many other drugs.
* In research, organisms are genetically engineered to discover the functions of certain genes.
* Industrial applications include transforming microorganisms such as bacteria or yeast, or insect mammalian cells with a gene coding for a useful protein. Mass quantities of the protein can be produced by growing the transformed organism in bioreactors using fermentation, then purifying the protein.
* Genetic engineering is also used in agriculture to create genetically-modified crops or genetically-modified organisms.

New DNA may be inserted in the host genome by first isolating and copying the genetic material of interest, using molecular-cloning methods to generate a DNA sequence; or by synthesizing the DNA, and then inserting this construct into the host organism. Genes may be removed, or “knocked out”, using a nuclease.

Gene targeting is a different technique that uses homologous recombination to change an endogenous gene, and can be used to delete a gene, remove exons, add a gene, or introduce point mutations. Genetic engineering has applications in medicine, research, industry and agriculture and can be used on a wide range of plants, animals and microorganisms.

Two fundamental kinds of cell are somatic cells and reproductive cells. Most of the cells in our bodies are somatic – cells that make up organs like skin, liver, heart, lungs, etc., and these cells vary from one another.  Changing the genetic material in these cells is not passed along to a person’s offspring.  Reproductive cells are sperm cells, egg cells, and cells from very early embryos.  Changes in the genetic make-up of reproductive cells would be passed along to the person’s offspring.  Those reproductive cell changes could result in different genetics in the offspring’s somatic cells than otherwise would have occurred because the genetic makeup of somatic cells is directly linked to that of the germ cells from which they are derived.

Two problems must be confronted when changing genes.  The first is what kind of change to make to the gene.  The second is how to incorporate that change in all the other cells that are must be changed to achieve a desired effect.

There are several options for what kind of change to make to the gene.  DNA in the gene could be replaced by other DNA from outside (called “homologous replacement).  Or the gene could be forced to mutate (change structure – “selective reverse mutation.”)  Or a gene could just be added.  Or one could use a chemical to simply turn off a gene and prevent it from acting.

There are also several options for how to spread the genetic change to all the cells that need to be changed.  If the altered cell is a reproductive cell, then a few such cells could be changed and the change would reach the other somatic cells as those somatic cells were created as the organism develops.  But if the change were made to a somatic cell, changing all the other relevant somatic cells individually like the first would be impractical due to the sheer number of such cells.  The cells of a major organ such as the heart or liver are too numerous to change one-by-one.  Instead, to reach such somatic cells a common approach is to use a carrier, or vector, which is a molecule or organism.  A virus, for example, could be used as a vector.  The virus would be an innocuous one or changed so as not to cause disease.  It would be injected with the genetic material and then as it reproduces and “infects” the target cells it would introduce the new genetic material.  It would need to be a very specific virus that would infect heart cells, for instance, without infecting and changing all the other cells of the body.  Fat particles and chemicals have also been used as vectors because they can penetrate the cell membrane and move into the cell nucleus with the new genetic material.

Genetic engineering has produced a variety of drugs and hormones for medical use. The first successful products of genetic engineering were protein drugs like insulin, which is used to treat diabetes, and growth hormone somatotropin. These proteins are made in large quantities by genetically engineered bacteria or yeast in large “bioreactors. ” Some drugs are also made in transgenic plants, such as tobacco. Other human proteins that are used as drugs require biological modifications that only the cells of mammals, such as cows, goats, and sheep, can provide. For these drugs, production in transgenic animals is a good option. Using farm animals for drug production has many advantages because they are reproducible, have flexible production, are easily maintained, and have a great delivery method

Interferon, which is used to eliminate certain viruses and kill cancer cells, also is a product of genetic engineering, as are tissue plasminogen activator and urokinase, which are used to dissolve blood clots.

Recombinant DNA technology not only allows therapeutic proteins to be produced on a large scale but using the same methodology protein molecules may be purposefully engineered. Genetic modifications introduced to a protein have many advantages over chemical modifications. Genetically engineered entities are biocompatible and biodegradable. The changes are introduced in 100% of the molecules with the exclusion of rare errors in gene transcription or translation. The preparations do not contain residual amounts of harsh chemicals used in the conjugation process. Bacterial expression systems, due to their simplicity, are often not able to produce a recombinant human protein identical to the naturally occurring wild type.

The evolving field of gene therapy involves manipulating human genes to treat or cure genetic diseases and disorders. Modified plasmids or viruses often are the messengers to deliver genetic material to the body’s cells, resulting in the production of substances that should correct the illness. Sometimes cells are genetically altered inside the body; other times scientists modify them in the laboratory and return them to the patient’s body.

Since the 1990s, gene therapy has been used in clinical trials to treat diseases and conditions such as AIDS, cystic fibrosis, cancer, and high cholesterol. Drawbacks of gene therapy are that sometimes the person’s immune system destroys the cells that have been genetically altered, and also that it is hard to get the genetic material into enough cells to have the desired effect.

Gene therapy and genetic engineering are two closely related technologies that involve altering the genetic material of organisms. The distinction between the two is based on purpose. Gene therapy seeks to alter genes to correct genetic defects and thus prevent or cure genetic diseases. Genetic engineering aims to modify the genes to enhance the capabilities of the organism beyond what is normal.

Gene therapy is an experimental technique that uses genes to treat or prevent disease. In the future, this technique may allow doctors to treat a disorder by inserting a gene into a patient’s cells instead of using drugs or surgery. Researchers are testing several approaches to gene therapy, including:

* Replacing a mutated gene that causes disease with a healthy copy of the gene.
* Inactivating, or “knocking out,” a mutated gene that is functioning improperly.
* Introducing a new gene into the body to help fight a disease.

Genome editing (also called gene editing) is a group of technologies that give scientists the ability to change an organism's DNA. These technologies allow genetic material to be added, removed, or altered at particular locations in the genome.

Several approaches to genome editing have been developed. A recent one is known as CRISPR-Cas9, which is short for clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9. The CRISPR-Cas9 system has generated a lot of excitement in the scientific community because it is faster, cheaper, more accurate, and more efficient than other existing genome editing methods.

CRISPR-Cas9 was adapted from a naturally occurring genome editing system in bacteria. The bacteria capture snippets of DNA from invading viruses and use them to create DNA segments known as CRISPR arrays. The CRISPR arrays allow the bacteria to "remember" the viruses (or closely related ones). If the viruses attack again, the bacteria produce RNA segments from the CRISPR arrays to target the viruses' DNA. The bacteria then use Cas9 or a similar enzyme to cut the DNA apart, which disables the virus.

The CRISPR-Cas9 system works similarly in the lab. Researchers create a small piece of RNA with a short "guide" sequence that attaches (binds) to a specific target sequence of DNA in a genome. The RNA also binds to the Cas9 enzyme. As in bacteria, the modified RNA is used to recognize the DNA sequence, and the Cas9 enzyme cuts the DNA at the targeted location. Although Cas9 is the enzyme that is used most often, other enzymes (for example Cpf1) can also be used. Once the DNA is cut, researchers use the cell's own DNA repair machinery to add or delete pieces of genetic material, or to make changes to the DNA by replacing an existing segment with a customized DNA sequence.

Genome editing is of great interest in the prevention and treatment of human diseases. Currently, most research on genome editing is done to understand diseases using cells and animal models. Scientists are still working to determine whether this approach is safe and effective for use in people. It is being explored in research on a wide variety of diseases, including single-gene disorders such as [cystic fibrosis](https://medlineplus.gov/genetics/condition/cystic-fibrosis/), [hemophilia](https://medlineplus.gov/genetics/condition/hemophilia/), and [sickle cell disease](https://medlineplus.gov/genetics/condition/sickle-cell-disease/). It also holds promise for the treatment and prevention of more [complex diseases](https://medlineplus.gov/genetics/understanding/mutationsanddisorders/complexdisorders/), such as cancer, heart disease, mental illness, and human immunodeficiency virus (HIV) infection.

Ethical concerns arise when genome editing, using technologies such as CRISPR-Cas9, is used to alter human genomes. Most of the changes introduced with genome editing are limited to somatic cells, which are cells other than egg and sperm cells. These changes affect only certain tissues and are not passed from one generation to the next. However, changes made to genes in egg or sperm cells (germline cells) or in the genes of an embryo could be passed to future generations. Germline cell and embryo genome editing bring up a number of ethical challenges, including whether it would be permissible to use this technology to enhance normal human traits (such as height or intelligence). Based on concerns about ethics and safety, germline cell and embryo genome editing are currently illegal in many countries.

**Nanotechnology**, the manipulation of matter at the atomic and molecular scale to create materials with remarkably varied and new properties, is a rapidly expanding area of research with huge potential in many sectors, ranging from healthcare to construction and electronics. In medicine, it promises to revolutionize drug delivery, gene therapy, diagnostics, and many areas of research, development and clinical application.

 The use of nanotechnology in medicine offers some exciting possibilities. The aim of nanomedicine may be broadly defined as the comprehensive monitoring, control, construction, repair, defence and improvement of all human biological systems, working from the molecular level using engineered devices and nanostructures, ultimately to achieve medical benefits. In this context, nanoscale should be taken to include active components or objects in the size range from one nanometre to hundreds of nanometres. These may be included in a micro-device (that have a macro-interface) or in a biological environment. The focus, however, is always on nano-interactions within the framework of a larger device or directly within a sub-cellular (or cellular) system.

 Nanotechnology in medicine involves applications of nanoparticles currently under development, as well as longer range research that involves the use of manufactured nano-robots to make repairs at  the cellular level.

Nanoparticulate drug delivery systems In the short and medium term, the main use of nanoparticle medicinal products (NMP) is vectorisation of active principles, corresponding to several products already marketed like Doxil™ or more recently Abraxane™

Generally three vector generations are considered:

 • First generation vectors: nanospheres and nanocapsules (the best known and most accessible); • Second generation vectors: nanoparticles coated with hydrophilic polymers such as polyethylene glycol (PEG), PEGylated nanoparticles;

 • Third generation vectors, still under development, combining a biodegradable core and a polymer envelope (PEG) with a membrane recognition ligand.

Today, most current research projects in nano delivery systems are focused on the third type. Conventional chemotherapy employs drugs that are known to kill cancer cells effectively. But these cytotoxic drugs kill healthy cells in addition to tumor cells, leading to adverse side effects such as nausea, neuropathy, hair-loss, fatigue, and compromised immune function. Nanoparticles can be used as drug carriers for chemotherapeutics to deliver medication directly to the tumor while sparing healthy tissue [9] (Fig. 7).

 Nanocarriers present several advantages over conventional chemotherapy. They can:

• Protect drugs from being degraded in the body before they reach their target;

 • Enhance drug absorption into tumors and the cancerous cells themselves; • Allow for better control over the timing and distribution of drugs to the tissue, making it easier for oncologists to assess how well they work;

• Prevent drugs from interacting with normal cells, thus avoiding side effects.

Passive targeting

There are now several nanocarrier-based drugs on the market, which rely on passive targeting through a process known as “enhanced permeability and retention”. Because of their size and surface properties, certain nanoparticles can escape through blood vessel walls into tissues. In addition, tumors tend to have leaky blood vessels and defective lymphatic drainage, causing nanoparticles to accumulate in them, thereby concentrating the attached cytotoxic drug where it’s needed, protecting healthy tissue and greatly reducing adverse side effects. Another strategy for passive targeting consists in using myeloid cells like macrophages which absorb nanoparticles and concentrate them in the site to be treated, like a Trojan horse.

 Active targeting On the horizon are nanoparticles that will actively target drugs to cancerous cells, based on the molecules that they express on their surface. Molecules that bind particular cellular receptors can be attached to a nanoparticle so that it specifically targets cells expressing this receptor. Active targeting can even be used to bring drugs into the cancerous cell, by inducing the cell to absorb the nanocarrier. Active targeting can be combined with passive targeting to further reduce interaction of carried drugs with healthy tissue. Nanotechnology-enabled active and passive targeting can also increase the efficiency of a chemotherapeutic, achieving more significant tumor reduction with lower drug doses.

. Destruction from within Moving away from conventional chemotherapeutic agents that activate normal molecular mechanisms to induce cell death, researchers are exploring ways to physically destroy cancerous cells from within. One such technology – nanoshells – is being used in the laboratory to thermally destroy tumors from the inside. Nanoshells can be designed to absorb light at different wavelengths, generating heat (hyperthermia). Once the cancer cells take up the nanoshells (via active targeting), scientists apply near-infrared light that is absorbed by the nanoshells, creating an intense heat inside the tumor that selectively kills tumor cells without disturbing neighbouring healthy cells. Similarly, new targeted magnetic nanoparticles are in development that will both be visible through Magnetic Resonance Imaging (MRI) and can also destroy cells by hyperthermia.

 Drug delivery (mechanical) devices Implanted drug delivery devices – DDD – can take benefit of nanotechnology. Conventional water-soluble drugs can create difficulties in treatment, such as failed absorption in the diseased areas. However, nanomedicine applications such as diagnostic nanomachines provide the ability to monitor the internal chemistry of the body’s organs, providing direct access to diseased areas. Moreover, technology such as nanobots can be equipped with wireless transmitters, and this offers doctors opportunities to change the treatment method if a patient’s medical condition gets worse. Nanobots in medicine could also be planted into a patient’s nervous system to monitor pulse and brainwave activities.

According to scientists, nanobots can completely replace pacemakers by treating the heart’s cell directly. Research regarding nanobots in medicine offer several opportunities such as artificial antibodies, artificial white blood cells (WBCs) and red blood cells (RBCs), and antiviral nanobots. The major advantage that nanobots provide is that they are extremely durable. Theoretically, they can operate for years without any damage owing to their miniature size, which reduces mechanical damage.