MOLECULAR GENETICS – (E Content)

R. KOKILA

Assistant professor

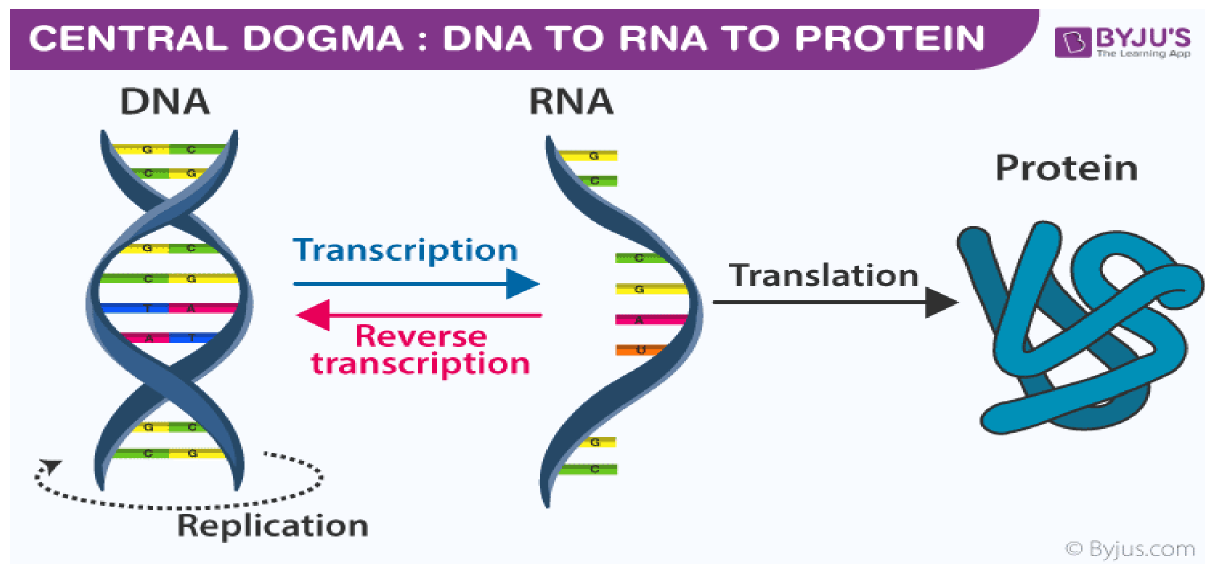
Department of Biotechnology

MOLECULAR GENETICS

UNIT 1

Central Dogma Definition

"Central dogma is the process in which the genetic information flows from DNA to RNA, to make a functional product protein."

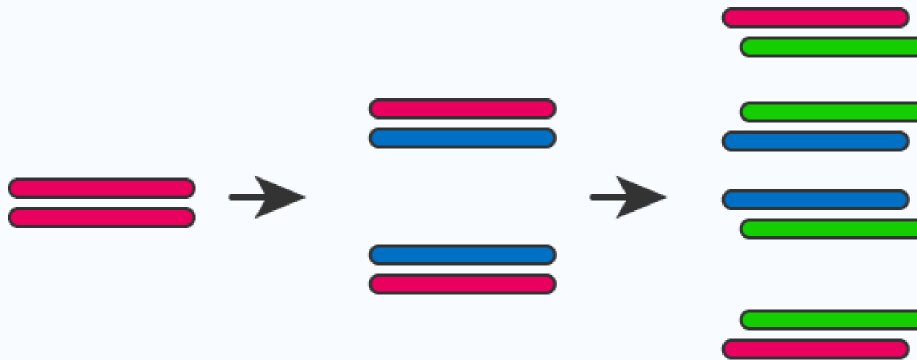


What is Central Dogma?

The central dogma illustrates the flow of genetic information in cells, the DNA replication, and coding for the RNA through the transcription process and further RNA codes for the proteins by translation.

The concept of a sequence of interaction can be understood through the framework. The most common includes biopolymers. The major category of biopolymers include Proteins, RNA and DNA that are further divided into general transfers, unknown transfers, and special transfers.

Special transfers occur in an exceptional case in the laboratory. General transfer occurs in almost all cells. It describes the regular flow of information through transcription and translation. Unknown transfers are said never to occur.



The new DNA strands are formed, with one strand of the parent DNA and the other is newly synthesized, this process is called semiconservative DNA replication.

Central Dogma Steps

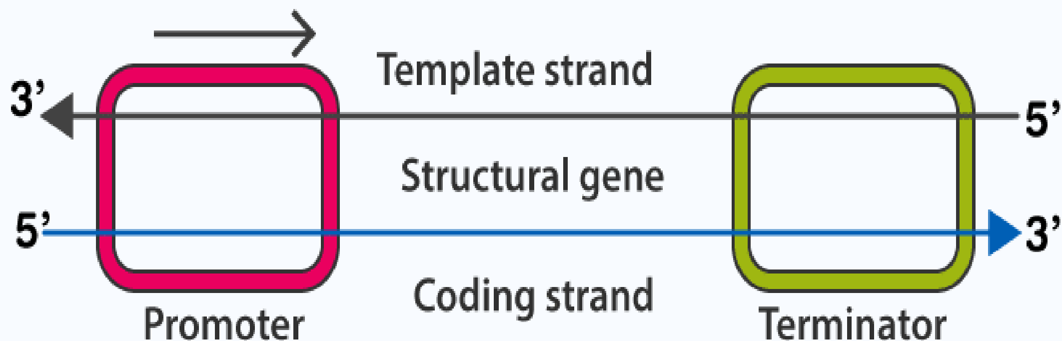
The central dogma takes place in two different steps:

Transcription

Transcription is the process by which the information is transferred from one strand of the DNA to RNA by the enzyme RNA Polymerase. The DNA strand which undergoes this process consists of three parts namely promoter, structural gene, and a terminator.

The DNA strand that synthesizes the RNA is called the template strand and the other strand is called the coding strand. The DNA-dependent RNA polymerase binds to the promoter and catalyzes the polymerization in the 3′ to 5′ direction.

As it approaches the terminator sequence, it terminates and releases the newly synthesized RNA strand. The newly released RNA strand further undergoes post-transcriptional modifications.



Translation

Translation is the process by which the RNA codes for specific proteins. It is an active process which requires energy. This energy is provided by the charged tRNA molecules.

Ribosomes initiate the translation process. The ribosomes consist of a larger subunit and a smaller subunit. The larger subunit, in turn, consists of two tRNA molecules placed close enough so that peptide bond can be formed at the expense of enough energy.

The mRNA enters the smaller subunit which is then held by the tRNA molecules of the complementary codon present in the larger subunit. Thus, two codons are held by two tRNA molecules placed close to each other and a peptide bond is formed between them. As this process repeats, long polypeptide chains of amino acids are synthesized.

DNA: Structure, Function and Discovery

Nucleic acids are the organic materials present in all organisms in the form of DNA or RNA. These nucleic acids are formed by the combination of nitrogenous bases, sugar molecules and phosphate groups that are linked by different bonds in a series of sequences. The DNA structure defines the basic genetic makeup of our body. In fact, it defines the genetic makeup of nearly all life on earth.

What is DNA?

"DNA is a group of molecules that is responsible for carrying and transmitting the hereditary materials or the genetic instructions from parents to offsprings."
This is also true for viruses, as most of these entities have either RNA or **DNA as their genetic material**. For instance, some viruses may have RNA as their genetic material, while others have DNA as the genetic material. The Human Immunodeficiency Virus (HIV) contains RNA, which is then converted into DNA after attaching itself to the host cell.

Apart from being responsible for the inheritance of genetic information in all living beings, DNA also plays a crucial role in the production of proteins. Nuclear DNA is the DNA contained within the nucleus of every cell in a eukaryotic organism. It codes for the majority of the organism's genomes while the mitochondrial DNA and plastid DNA handles the rest.

The DNA present in the mitochondria of the cell is termed mitochondrial DNA. It is inherited from the mother to the child. In humans, there are approximately 16,000 base pairs of mitochondrial DNA. Similarly, plastids have their own DNA, and they play an essential role in photosynthesis.

DNA is known as Deoxyribonucleic Acid. It is an organic compound that has a unique molecular structure. It is found in all **prokaryotic cells and eukaryotic cells.**

DNA Types

There are three different DNA types:

- **A-DNA:** It is a right-handed double helix similar to the B-DNA form. Dehydrated DNA takes an A form that protects the DNA during extreme conditions such as desiccation. Protein binding also removes the solvent from DNA, and the DNA takes an A form.
- **B-DNA:** This is the most common DNA conformation and is a right-handed helix. The majority of DNA has a B type conformation under normal physiological conditions.
- **Z-DNA:** Z-DNA is a left-handed DNA where the double helix winds to the left in a zig-zag pattern. It was discovered by Andres Wang and Alexander Rich. It is found ahead of the start site of a gene and hence, is believed to play some role in gene regulation.

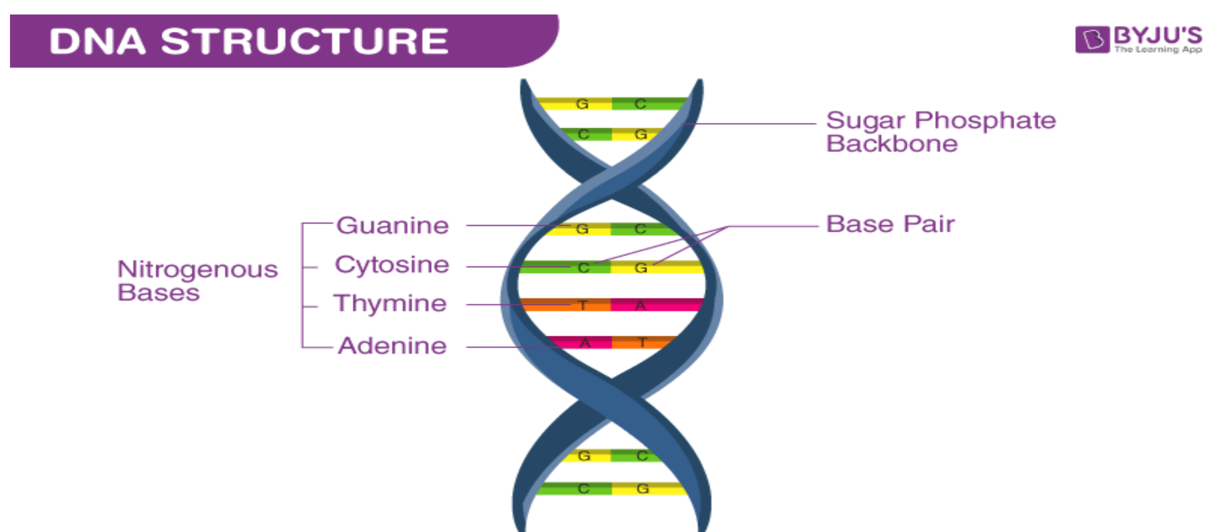
Who Discovered DNA?

DNA was first recognized and identified by the Swiss biologist **Johannes Friedrich Miescher** in 1869 during his research on white blood cells.

The double helix structure of a DNA molecule was later discovered through the experimental data by James Watson and Francis Crick. Finally, it was proved that DNA is responsible for storing genetic information in living organisms.

DNA Diagram

The following diagram explains the DNA structure representing the different parts of the DNA. DNA comprises a sugar-phosphate backbone and the nucleotide bases (guanine, cytosine, adenine and thymine).



DNA Diagram representing the DNA Structure

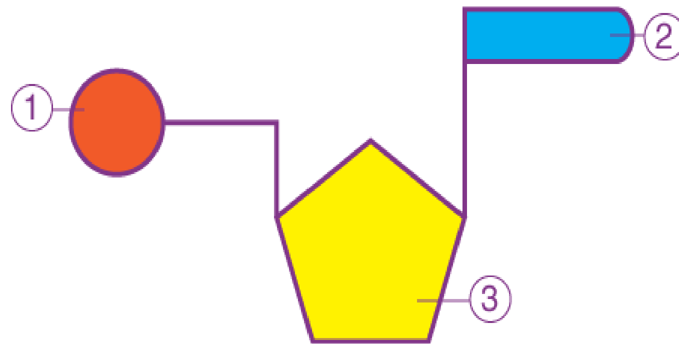
DNA Structure

The DNA structure can be thought of as a twisted ladder. This structure is described as a double-helix, as illustrated in the figure above. It is a nucleic acid, and all nucleic acids are made up of nucleotides. The DNA molecule is composed of units called nucleotides, and each nucleotide is composed of three different components such as sugar, phosphate groups and nitrogen bases.

The basic building blocks of DNA are nucleotides, which are composed of a sugar group, a phosphate group, and a nitrogen base. The sugar and phosphate groups link the nucleotides together to form each strand of DNA. Adenine (A), Thymine (T), Guanine (G) and Cytosine (C) are four types of nitrogen bases.

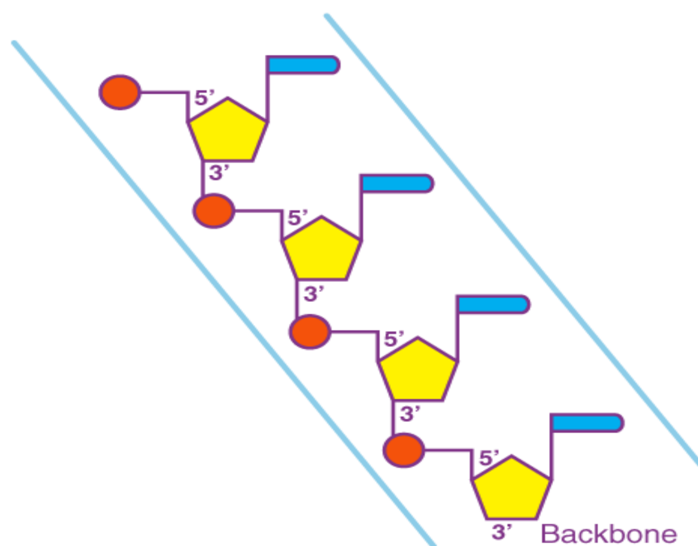
These 4 Nitrogenous bases pair together in the following way: **A** with **T,** and **C** with **G**. These base pairs are essential for the DNA's double helix structure, which resembles a twisted ladder.

The order of the nitrogenous bases determines the genetic code or the DNA's instructions.



Components of DNA Structure

Among the three components of DNA structure, sugar is the one which forms the backbone of the DNA molecule. It is also called deoxyribose. The nitrogenous bases of the opposite strands form hydrogen bonds, forming a ladder-like structure.



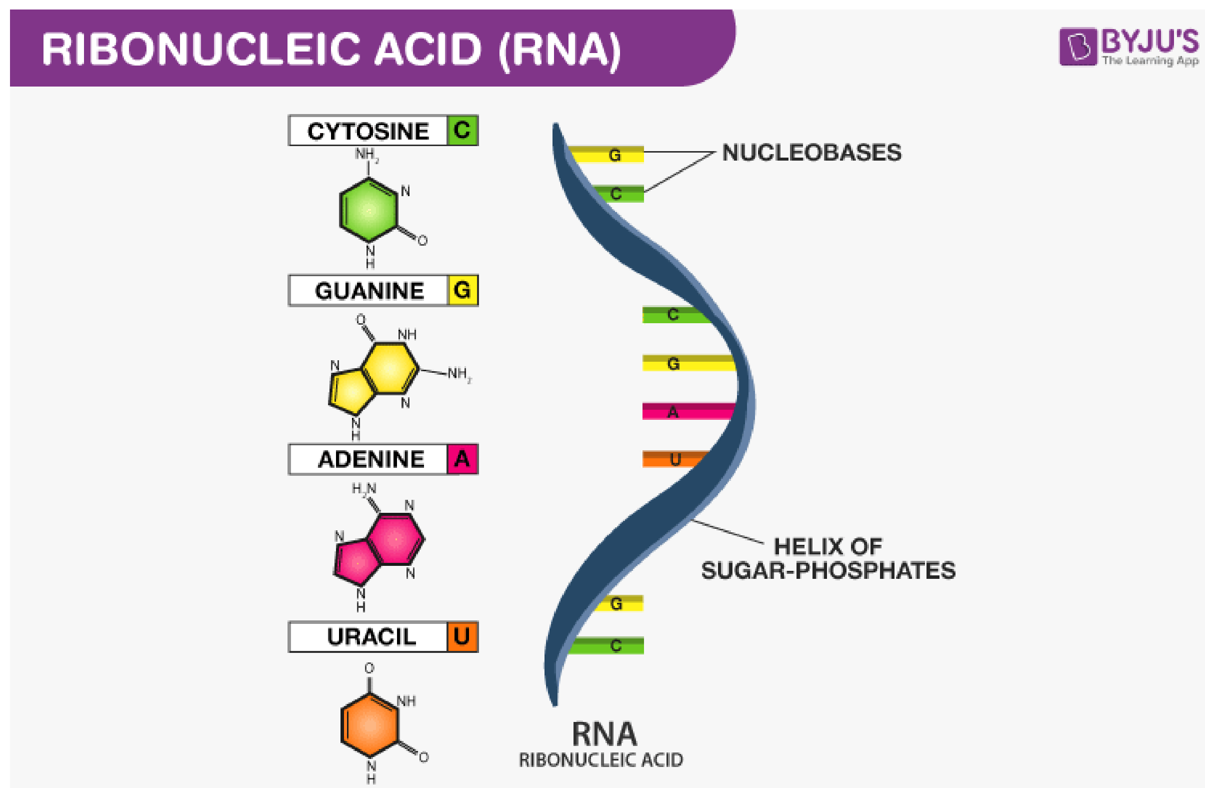
DNA Structure Backbone

The DNA molecule consists of 4 nitrogen bases, namely adenine (A), thymine (T), cytosine (C) and Guanine (G), which ultimately form the structure of a nucleotide. The A and G are purines, and the C and T are pyrimidines.

The two strands of DNA run in opposite directions. These strands are held together by the hydrogen bond that is present between the two complementary bases. The strands are helically twisted, where each strand forms a right-handed coil, and ten nucleotides make up a single turn.

Structure of RNA

RNA is a ribonucleic acid that helps in the synthesis of proteins in our body. This nucleic acid is responsible for the production of new cells in the human body. It is usually obtained from the DNA molecule. RNA resembles the same as that of DNA, the only difference being that it has a single strand unlike the DNA which has two strands and it consists of an only single ribose sugar molecule in it. Hence is the name Ribonucleic acid. RNA is also referred to as an enzyme as it helps in the process of chemical reactions in the body.



Basic Structure of RNA

The basic structure of RNA is shown in the figure below-

The ribonucleic acid has all the components same to that of the DNA with only 2 main differences within it. RNA has the same nitrogen bases called the adenine, Guanine, Cytosine as that of the DNA except for the Thymine which is replaced by the uracil. Adenine and uracil are considered as the major building blocks of RNA and both form base-pair with the help of 2 hydrogen bonds.

RNA resembles a hairpin structure and like the nucleotides in DNA, nucleotides are formed in this ribonucleic material (RNA). Nucleosides are nothing but the phosphate groups which sometimes also helps in the production of nucleotides in the DNA.

Functions of RNA

The ribonucleic acid – RNA, which are mainly composed of nucleic acids, are involved in a variety of functions within the cell and are found in all living organisms including bacteria, viruses, plants, and animals. These nucleic acid functions as a structural molecule in cell organelles and are also involved in the catalysis of biochemical reactions. The different types of RNA are involved in various cellular process. The primary functions of RNA:

- Facilitate the translation of DNA into proteins
- Functions as an adapter molecule in protein synthesis
- Serves as a messenger between the DNA and the ribosomes.
- They are the carrier of genetic information in all living cells
- Promotes the ribosomes to choose the right amino acid which is required in building up of new proteins in the body.

RNA Types

There are various types of RNA, out which most well-known and most commonly studied in the human body are :

tRNA – Transfer RNA

The transfer RNA is held responsible for choosing the correct protein or the amino acids required by the body in-turn helping the ribosomes. It is located at the endpoints of each amino acid. This is also called as soluble RNA and it forms a link between the messenger RNA and the amino acid.

rRNA-Ribosomal RNA

The rRNA is the component of the ribosome and are located within the in the cytoplasm of a cell, where ribosomes are found. In all living cells, the ribosomal RNA plays a fundamental role in the synthesis and translation of mRNA into proteins. The rRNA is mainly composed of cellular RNA and are the most predominant RNA within the cells of all living beings.

mRNA – Messenger RNA.

This type of RNA functions by transferring the genetic material into the ribosomes and pass the instructions about the type of proteins, required by the body cells. Based on the functions, these types of RNA is called the messenger RNA. Therefore, the mRNA plays a vital role in the process of transcription or during the protein synthesis process.

Griffith Experiment and Search of Genetic Material

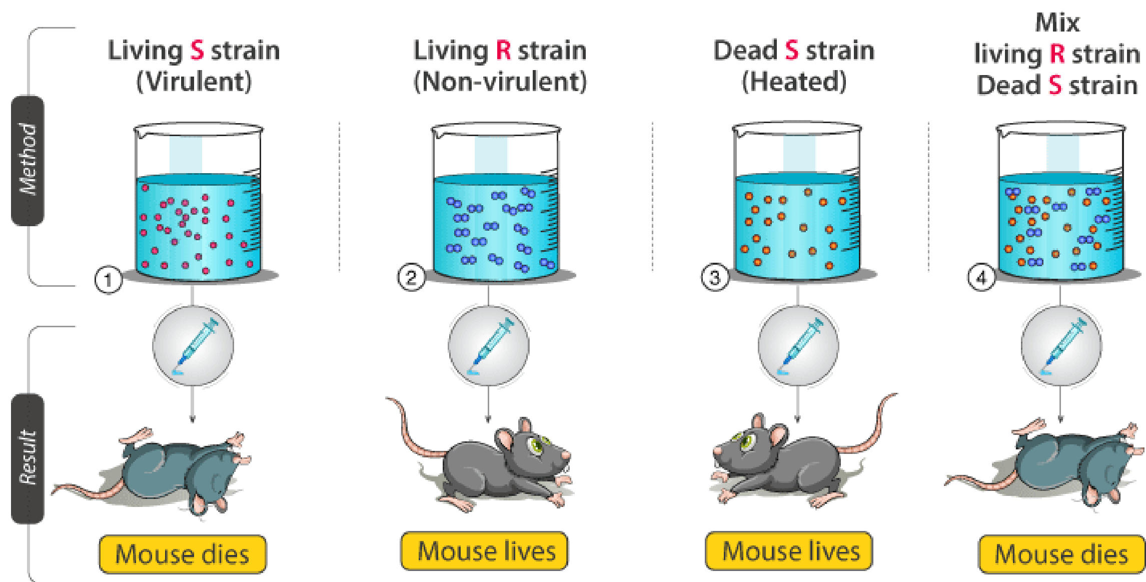
Griffith Experiment & Transforming Principle

Griffith experiment was a stepping stone for the discovery of genetic material. Frederick Griffith experiments were conducted with Streptococcus pneumoniae.

During the experiment, Griffith cultured Streptococcus pneumoniae bacteria which showed two patterns of growth. One culture plate consisted of smooth shiny colonies (S) while other consisted of rough colonies (R). The difference was due to the presence of mucous coat in S strain bacteria, whereas the R strain bacteria lacked them.

Experiment: Griffith injected both S and R strains to mice. The one which was infected with the S strain developed pneumonia and died while that infected with the R strain stayed alive.

In the second stage, Griffith heat-killed the S strain bacteria and injected into mice, but the mice stayed alive. Then, he mixed the heat-killed S and live R strains. This mixture was injected into mice and they died. In addition, he found living S strain bacteria in dead mice.



Conclusion: Based on the observation, Griffith concluded that R strain bacteria had been transformed by S strain bacteria. The R strain inherited some 'transforming principle' from the heat-killed S strain bacteria which made them virulent. And he assumed this transforming principle as genetic material.

DNA as Genetic Material

Griffith experiment was a turning point towards the discovery of hereditary material. However, it failed to explain the biochemistry of genetic material. Hence, a group of scientists, Oswald Avery, Colin MacLeod and Maclyn McCarty continued the Griffith experiment in search of biochemical nature of the hereditary material. Their discovery revised the concept of protein as genetic material to DNA as genetic material.

Avery and his team extracted and purified proteins, DNA, RNA and other biomolecules from the heat-killed S strain bacteria. They discovered that DNA is the genetic material and it is alone responsible for the transformation of the R strain bacteria. They observed that protein-digesting enzymes (proteases) and RNA-digesting enzymes (RNases) didn't inhibit transformation but DNase did. Although it was not accepted by all, they concluded DNA as genetic material.

DNA As Genetic Material - Hershey And Chase Experiment

Experiments of Hershey and Chase

We know about Griffith's experiment and experiments that followed to discover the hereditary material in organisms. Based on Griffith's experiment, Avery and his team isolated DNA and proved DNA to be the genetic material. But it was not accepted by all until Hershey and Chase published their experimental results.

In 1952, Alfred Hershey and Martha Chase took an effort to find the genetic material in organisms. Their experiments led to an unequivocal proof to DNA as genetic material. Bacteriophages (viruses that affect bacteria) were the key element for Hershey and Chase experiment.

The virus doesn't have their own mechanism of reproduction but they depend on a host for the same. Once they attach to the host cell, their genetic material is transferred to the host. Here in case of bacteriophages, bacteria are their host. The infected bacteria are manipulated by the bacteriophages such that bacterial cells start to replicate the viral genetic material. Hershey and Chase conducted an experiment to discover whether it was protein or DNA that acted as the genetic material that entered the bacteria.

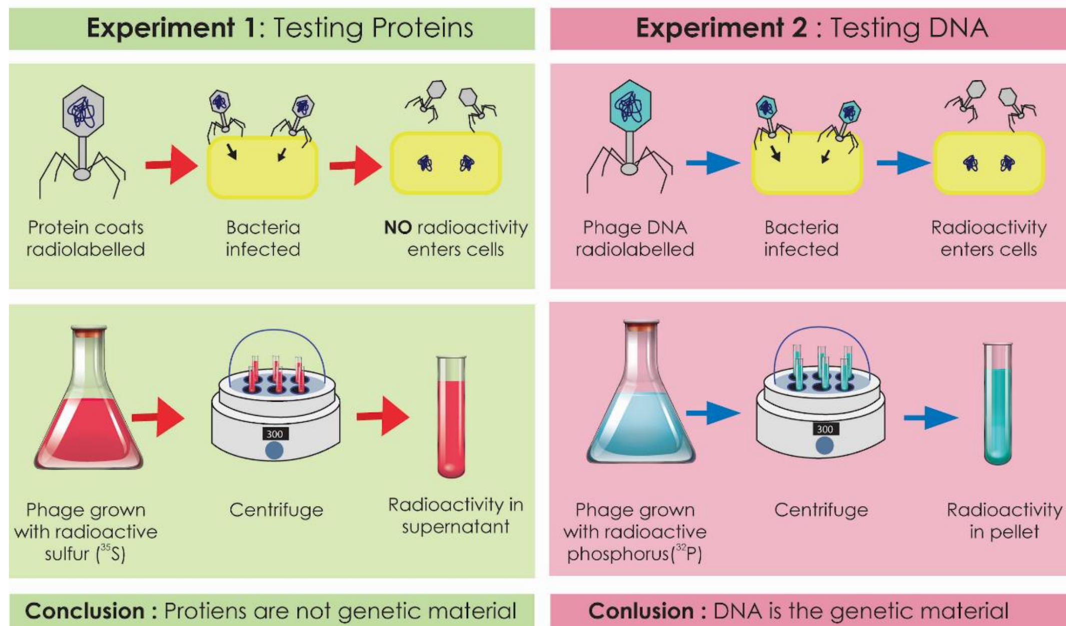
DNA as Genetic Material

Experiment: The experiment began with the culturing of viruses in two types of medium. One set of viruses (A) was cultured in a medium of radioactive phosphorus whereas another set (B) was cultured in a medium of radioactive sulfur. They observed that the first set of viruses (A) consisted of radioactive DNA but not radioactive proteins. This is because DNA is a phosphorus-based compound while protein is not. The latter set of viruses (B) consisted of radioactive protein but not radioactive DNA.

The host for infection was E.coli bacteria. The viruses were allowed to infect bacteria by removing the viral coats through a number of blending and centrifugation.

Observation: E.coli bacteria which were infected by radioactive DNA viruses (A) were radioactive but the ones that were infected by radioactive protein viruses (B) were non-radioactive.

Conclusion: Resultant radioactive and non-radioactive bacteria infer that the viruses that had radioactive DNA transferred their DNA to the bacteria but viruses that had radioactive protein didn't get transferred to the bacteria. Hence, DNA is the genetic material and not the protein.



Avery, McCarty, and MacLeod Experiment

During World War II, in 1943, Oswald Avery, Maclyn McCarty, and Colin MacLeod working at Rockefeller University in New York, dedicated themselves to continuing the work of Griffith in order to determine the biochemical nature of Griffith's transforming principle in an in vitro system. They used the phenotype of S. pneumoniae cells expressed on blood agar in order to figure out whether transformation had taken place or not, rather than working with mice. The transforming principle was partially purified from the cell extract (i.e., cell-free extract of heat-killed type III S cells) to determine which macromolecule of S cell transformed type II R-strain into the type III S-strain. They demonstrated DNA to be that particular transforming principle.

- Initially, type III S cells were heat-killed, and lipids and carbohydrates were removed from the solution.
- Secondly, they treated heat-killed S cells with digestive enzymes such as RNases and proteases to degrade RNA and proteins. Subsequently, they also treated it with DNases to digest DNA, each added separately in different tubes.

- Eventually, they introduced living type IIR cells mixed with heat-killed IIIS cells onto the culture medium containing antibodies for IIR cells. Antibodies for IIR cells were used to inactivate some IIR cells such that their number doesn't exceed the count of IIIS cells. that help to provide the distinct phenotypic differences in culture media that contained transformed S strain bacteria.

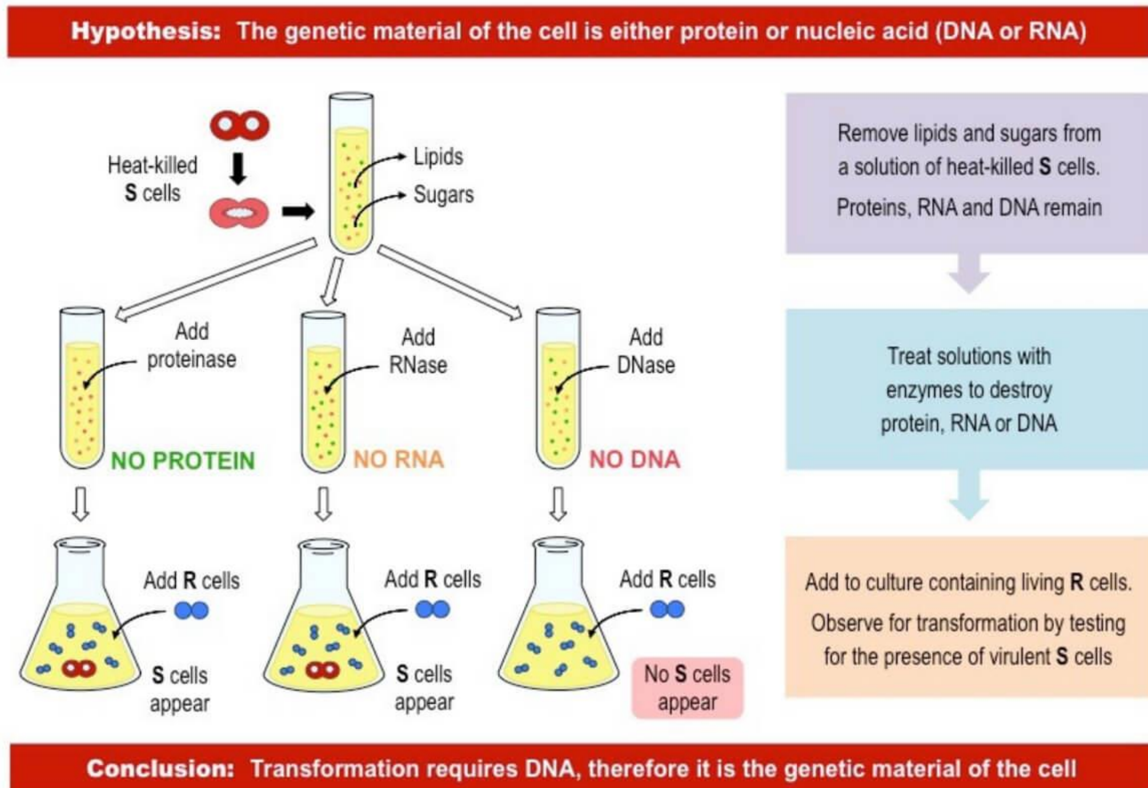


Figure: Avery, McCarty, and MacLeod Experiment

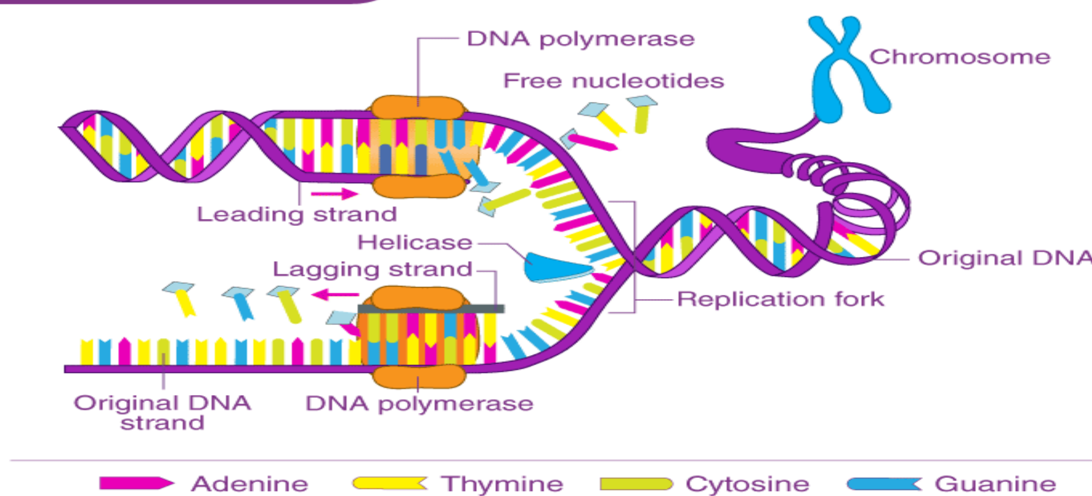
Observation of Avery, McCarty, and MacLeod Experiment
The culture treated with DNase did not yield transformed type III S strain bacteria which indicated that DNA was the hereditary material responsible for transformation.

Conclusion of Avery, McCarty, and MacLeod Experiment
DNA was found to be the genetic material that was being transferred between cells, not proteins.

UNIT 2

DNA Replication

In the process of DNA replication, the DNA makes multiple copies of itself. It is a biological polymerisation, which proceeds in the sequence of initiation, elongation, and termination. It is an enzyme-catalysed reaction. DNA Polymerase is the main enzyme in the replication process.



DNA Replication Process

DNA Replication Steps

Following are the important steps involved in DNA replication:

Initiation

DNA replication demands a high degree of accuracy because even a minute mistake would result in mutations. Thus, replication cannot initiate randomly at any point in DNA.

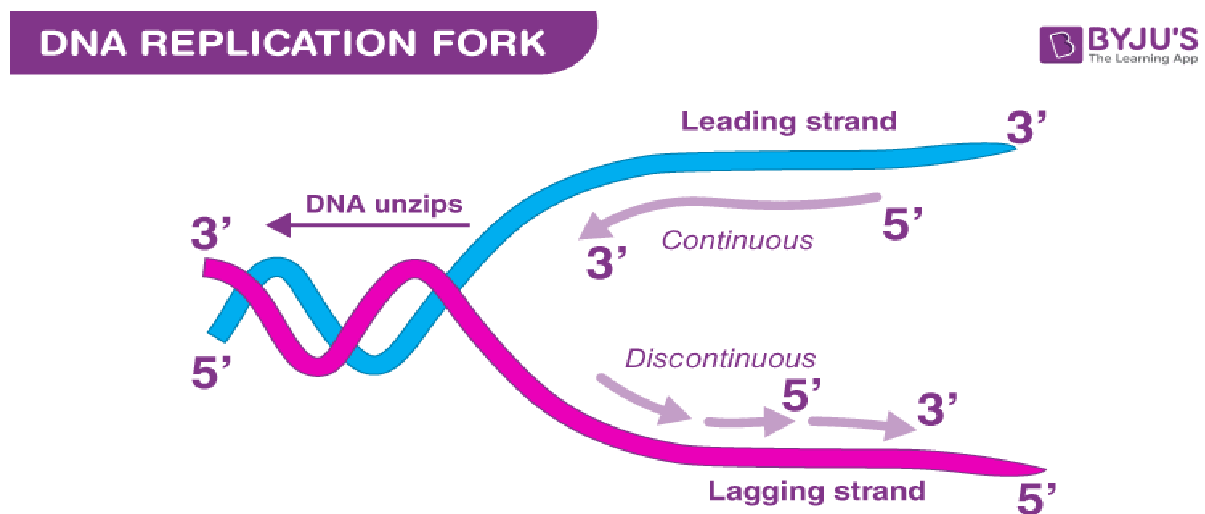
For the replication to begin there is a particular region called the origin of replication. This is the point where the replication originates. Replication begins with the spotting of this origin followed by the unwinding of the two DNA strands.

Unzipping of DNA strands in their entire length is not feasible due to high energy input. Hence, first, a replication fork is created catalysed by the helicase enzyme, which unzips the DNA strand.

Elongation

As the strands are separated, the polymerase enzymes start synthesising the complementary sequence in each of the strands. The parental strands will act as a template for newly synthesising daughter strands.

It is to be noted that elongation is unidirectional i.e. DNA is always polymerised only in the 5′ to 3′ direction. Therefore, in one strand (the template 3 ‘→5 ‘) it is continuous, hence called continuous replication while on the other strand (the template 5 ‘→3 ‘) it is discontinuous replication. They occur as fragments called Okazaki fragments. The enzyme called DNA ligase joins them later.



DNA Replication Fork

Termination

Termination of replication occurs in different ways in different organisms. In E.coli like organisms, chromosomes are circular. And this happens when the two replication forks between the two terminals meet each other.

Role of Enzymes in DNA Replication

DNA replication is a highly enzyme-dependent process. There are many enzymes involved in DNA replication, which includes the enzymes, DNA-dependent DNA polymerase, helicase, ligase, etc. Among them, DNA-dependent DNA polymerase is the main enzyme.

DNA-dependent DNA polymerase

It helps in the polymerisation, catalyses and regularises the whole process of DNA replication with the support of other enzymes. Deoxyribonucleoside triphosphates are the substrate as well as the energy provider for the replication process. DNA polymerase is of three types:

DNA Polymerase I

It is a DNA repair enzyme. It is involved in three activities:

- 5′-3′ polymerase activity
- 5′-3′ exonuclease activity
- 3′-5′ exonuclease activity

DNA Polymerase II

It is responsible for primer extension and proofreading.

DNA Polymerase III

It is responsible for in vivo DNA replication.

Helicase

Helicase is the enzyme, which unzips the DNA strands by breaking the hydrogen bonds between them. Thus, it helps in the formation of the replication fork.

Ligase

Ligase is the enzyme which joins together the Okazaki fragments of the discontinuous DNA strands.

Primase

This enzyme helps in the synthesis of RNA primer complementary to the DNA template strand.

Endonucleases

These produce a single-stranded or a double-stranded cut in a DNA molecule.

Single-stranded Binding Proteins

It binds to single-stranded DNA and protects it from forming secondary structures.

What are Okazaki Fragments?

Okazaki fragments are short sections of DNA formed at the time of discontinuous synthesis of the lagging strand during replication of DNA. It is essential as it allows for the synthesis of both the daughter strands required for cell division.

The role of Okazaki fragments is to permit the DNA polymerase to synthesise the lagging strands in the segments, as it is not correctly oriented for continuous synthesis.

DNA REPAIR

DNA repair can be divided into a set of mechanisms that identify and correct damage in DNA molecules. There are two general classes of DNA repair; the direct reversal of the chemical process generating the damage and the replacement of damaged nucleotide bases.

DNA encodes the cell genome and is therefore a permanent copy of a structure necessary for the correct functioning of a cell. Changes to the structure of DNA can cause mutations and genomic instability, leading to cancer. Damage to DNA is caused by the incorporation of incorrect nucleotide bases during DNA replication and the chemical changes caused by spontaneous mutation or exposure to environmental factors such as radiation.

The direct reversal DNA repair mechanism

Direct reversal of DNA damage is a mechanism of repair that does not require a template and is applied to two main types of damage. UV light induces the formation of pyrimidine dimers which can distort the DNA chain structure, blocking transcription beyond the area of damage.

Direct reversal through photoreactivation can inverse this dimerization reaction by utilizing light energy for the destruction of the abnormal covalent bond between adjacent pyrimidine bases. This type of photoreactivation does not occur in humans.

The damage caused by alkylating agents reacting with DNA can also be repaired through direct reversal. Methylation of guanine bases produces a change in the structure of DNA by forming a product that is complimentary to thymine rather than cytosine. The protein methyl guanine methyl transferase (MGMT) can restore the original guanine by transferring the methylation product to its active site.

DNA repair by excision

Excision is the general mechanism by which repairs are made when one of the double helix strands is damaged. The non-defective strand is used as a template with the damaged DNA on the other strand removed and replaced by the synthesis of new nucleotides. There are three types of excision repair:

1. **Base-excision repair.**
2. **Nucleotide excision repair.**
3. **Mismatch repair.**

Base-excision repair involves the recognition and removal of a single damaged base. The mechanism requires a family of enzymes called glycosylases. The enzymes remove the damaged base forming an AP site which is repaired by AP endonuclease before the nucleotide gap in the DNA strand is filled by DNA polymerase.

Nucleotide excision repair is a widespread mechanism for repairing damage to DNA and recognizes multiple damaged bases. This mechanism is used to repair the formation of pyrimidine dimers from UV light within humans. The process involves the recognition of damage which is then cleaved on both sides by endonucleases before resynthesis by DNA polymerase.

The third excision mechanism is called **mismatch repair** and occurs when mismatched bases are incorporated into the DNA strand during replication and are not removed by proofreading DNA polymerase. In mismatch repair, the missed errors are later corrected by enzymes which recognize and excise the mismatched base to restore the original sequence.

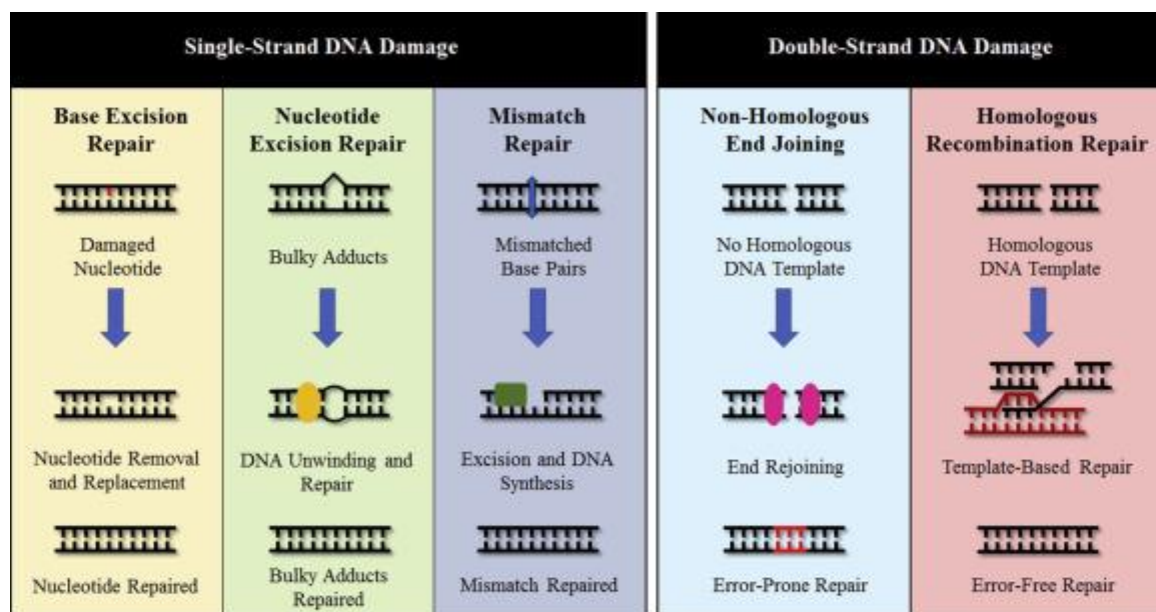
DNA double strand break repair

The repair of damage to both DNA strands is particularly important in maintaining genomic integrity. There are two main mechanisms for repairing double strand breaks: homologous recombination and classical nonhomologous end joining.

Homologous recombination involves the exchange of nucleotide sequences to repair damaged bases on both strands of DNA through the utilization of a sister chromatid. Classical nonhomologous end joining connects the break ends without a homologous template through the use of short DNA sequences called microhomologies. The mechanism is prone to error but protects genome integrity from possible chromosomal translocations that can occur through homologous recombination.

Studies have also found that double strand breaks can be repaired through alternative mechanisms such as single-stranded annealing and alternative joining during certain conditions. These mechanisms are mutagenic and can lead to a loss in genetic information.

Single-stranded annealing provides end joining between interspersed nucleotide repeats within the genome leading to one copy of the repeat and the intervening sequence being deleted in the process. Alternative joining has an undefined mechanism for repairing double strand breaks but is known to risk genomic integrity by joining end breaks on different chromosomes.



Recombination

- **Cleavage and rejoining of DNA molecules to generate new combinations of genes**
- **There are two types of recombination mechanisms occur in living system**

Genetic recombination
Site specific recombination
Genetic Recombination

- The production of gene combinations not found in the parents by the ***crossing over between homologous chromosomes during meiosis*** is called genetic recombination.
- Genetic recombination occurs in the germ cells (during meiosis I in oocytes and spermatocytes)
- Also called as homologous recombination as it takes place between homologous chromosomes
- It occurs during crossing over, in which homologous regions of chromosomes are exchanged and genes are shuffled into new combinations.
- It is an extremely important genetic process because it increases genetic variation.
- It is an ubiquitous process occurs in every bacteria and also in eukaryotes

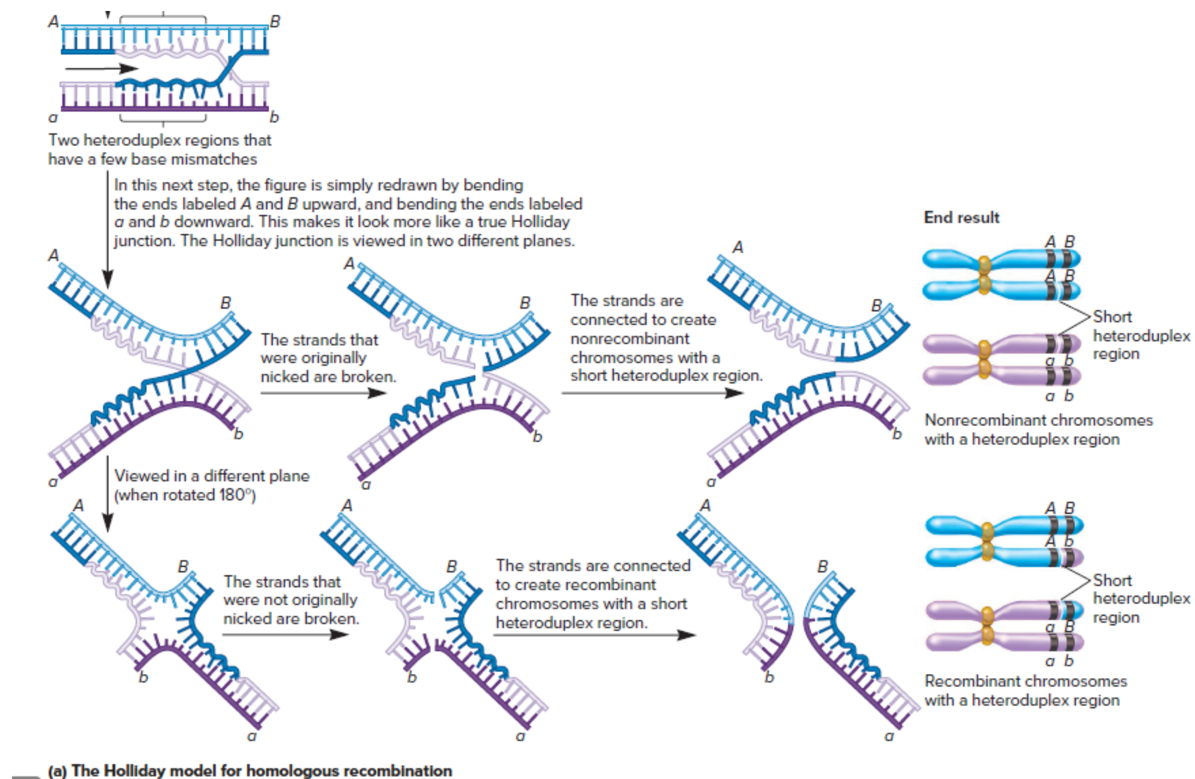
Models of Homologous Recombination

- **Holliday model (Robin Holliday 1964)**
- **Double strand repair model**

Holliday Model:

Holliday junction, cross-shaped structure that forms during the process of genetic recombination, when two double-stranded DNA molecules become separated into four strands in order to exchange segments of genetic information. This structure is named after British geneticist Robin Holliday, who proposed the original model for homologous (or general) recombination in 1964.

Homologous recombination occurs during meiosis and is characterized by the exchange of genes between a maternal chromatid and a paternal chromatid of a homologous chromosome pair. The two parent DNA molecules, which have long stretches of similar base sequences, are separated into single strands, resulting in base pairing that leads to a four-stranded DNA structure. The Holliday junction travels along the DNA duplex by “unzipping” one strand and reforming the hydrogen bonds on the second strand.



(a) The Holliday model for homologous recombination

UNIT 3

Mutation Definition

"Mutation is the change in our DNA base pair sequence due to various environmental factors such as UV light, or mistakes during DNA replication."

What Are Mutations?

The DNA sequence is specific to each organism. It can sometimes undergo changes in its base-pairs sequence. It is termed as a mutation. A mutation may lead to changes in proteins translated by the DNA. Usually, the cells can recognize any damage caused by mutation and repair it before it becomes permanent.

A mutation is a sudden, heritable modification in an organism's traits. The term "mutant" refers to a person who exhibits these heritable alterations. Mutations usually produce recessive genes.

Classification & Types of Mutations

Mutation Classifications	Types	Description	Examples of Human Disease(s)
Point mutation	Substitution	During replication, one base is inserted incorrectly, replacing the pair at the appropriate location on the complementary strand.	Sickle-cell anemia
	Insertion	In replicating DNA, one or more additional nucleotides are added, frequently causing a frameshift.	One form of beta-thalassemia
	Deletion	During replication, one or more nucleotides may be "skipped" or removed, which usually causes a frameshift.	Cystic fibrosis
Chromosomal mutation	Inversion	The flipping and reinserting of a single chromosomal region.	Opitz-Kaveggia syndrome
	Deletion	When a chromosome segment is lost, all the genes in that segment are also gone.	Cri du chat syndrome
	Duplication	A chromosomal segment is repeated, increasing the concentration of the genes in that area.	Some cancers
	Translocation	A section of one chromosome is inappropriately joined to another chromosome.	One form of leukemia
Copy number variation	Gene amplification	An increase is made in the tandem copies of a locus.	Some breast cancers
	Expanding trinucleotide	There are more repeating trinucleotide	Fragile X syndrome,

	repeat	sequences than usual.	Huntington’s disease

Causes of Mutations

The mutation leads to genetic variations among species. Positive mutations are transferred to successive generations.

E.g. Mutation in the gene coding for haemoglobin causes sickle cell anaemia. The R.B.Cs become sickle in shape. However, in the African population, this mutation provides protection against malaria.

A mutation in the gene controlling the cell division leads to cancer.

Let us have an overview of the causes and impacts of mutation.

Also Read: Mutagens

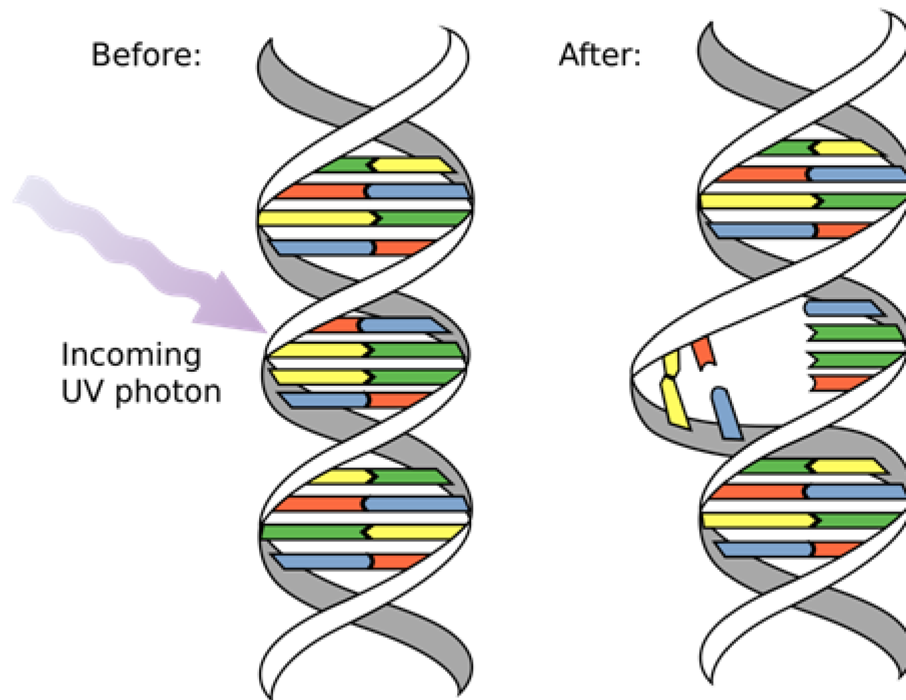
The mutation is caused due to the following reasons:

Internal Causes

Most of the mutations occur when the DNA fails to copy accurately. All these mutations lead to evolution. During cell division, the DNA makes a copy of its own. Sometimes, the copy of the DNA is not perfect and this slight difference from the original DNA is called a mutation.

External Causes

When the DNA is exposed to certain chemicals or radiations, it causes the DNA to break down. The ultraviolet radiations cause the thymine dimers to break resulting in a mutated DNA.



DNA Mutation

Effects of Mutation

There are several mutations that cannot be passed on to the offsprings. Such mutations occur in the somatic cells and are known as somatic mutations.

The germline mutations can be passed on to successive generations and occur in the reproductive cells.

Let us have a look at some of the effects of mutation:

Beneficial Effects of Mutation

1.
    1. Few mutations result in new versions of proteins and help the organisms to adapt to changes in the environment. Such mutations lead to evolution.
    2. Mutations in many bacteria result in antibiotic-resistant strains of bacteria that can survive in the presence of antibiotics.
    3. A unique mutation found in the population of Italy protects them from atherosclerosis, where fatty materials build up in the blood vessels.

Effects of Mutations

1. Genetic disorders can be caused by the mutation of one or more genes. Cystic fibrosis is one such genetic disorder caused by the mutation in one or more genes.

2. Cancer is another disease caused by the mutation in genes that regulate the cell cycle.

Spontaneous Mutation

Spontaneous mutations occur naturally in the genome. They generally occur due to error during replication, mitosis, meiosis, etc. Mutations may also occur due to mobile genetic elements or transposons. The main causes of spontaneous mutations are:

- Replication errors
- Slipped strand mispairing
- Wobble base pairing
- Depurination or deamination
- Tautomerism
- Unequal crossing over

Induced Mutation

Induced mutations do not occur spontaneously. They are induced through various chemical and physical agents known as mutagens. Mutagens greatly enhance the frequency of mutation.

Some of the mutagens are:

- Alkylating agents (Ethyl methanesulfonate or EMS, N-ethyl-N-nitrosourea or ENU)
- Base analogue (5-Bromouracil, Bromodeoxyuridine)
- Hydroxylamine modifies bases
- Deamination by nitrous acid
- DNA intercalating agents (ethidium bromide, proflavine)
- Oxidative damage (Reactive oxygen species, e.g. superoxide radical, hydrogen peroxide)
- Ionising and non-ionising radiations (gamma radiations, ultraviolet radiations, X-rays, etc.)

Difference between Spontaneous and Induced Mutation

The table below shows the main differences between Spontaneous and Induced Mutation.

Spontaneous Mutation	**Induced Mutation**
Spontaneous mutations occur naturally and mainly due to error in replication	Induced mutations occur due to physical or chemical agents
Occurs due to slippage in natural processes	Induced by mutagens

Caused due to replication error, tautomeric shift, transposable genetic elements, unequal cross overs, etc.	Caused due to base modification, base analogues, intercalating agents, base mispairing, radiations, etc.
E.g. sickle cell anaemia	E.g. skin cancer due to prolonged exposure to radiations

UNIT 4

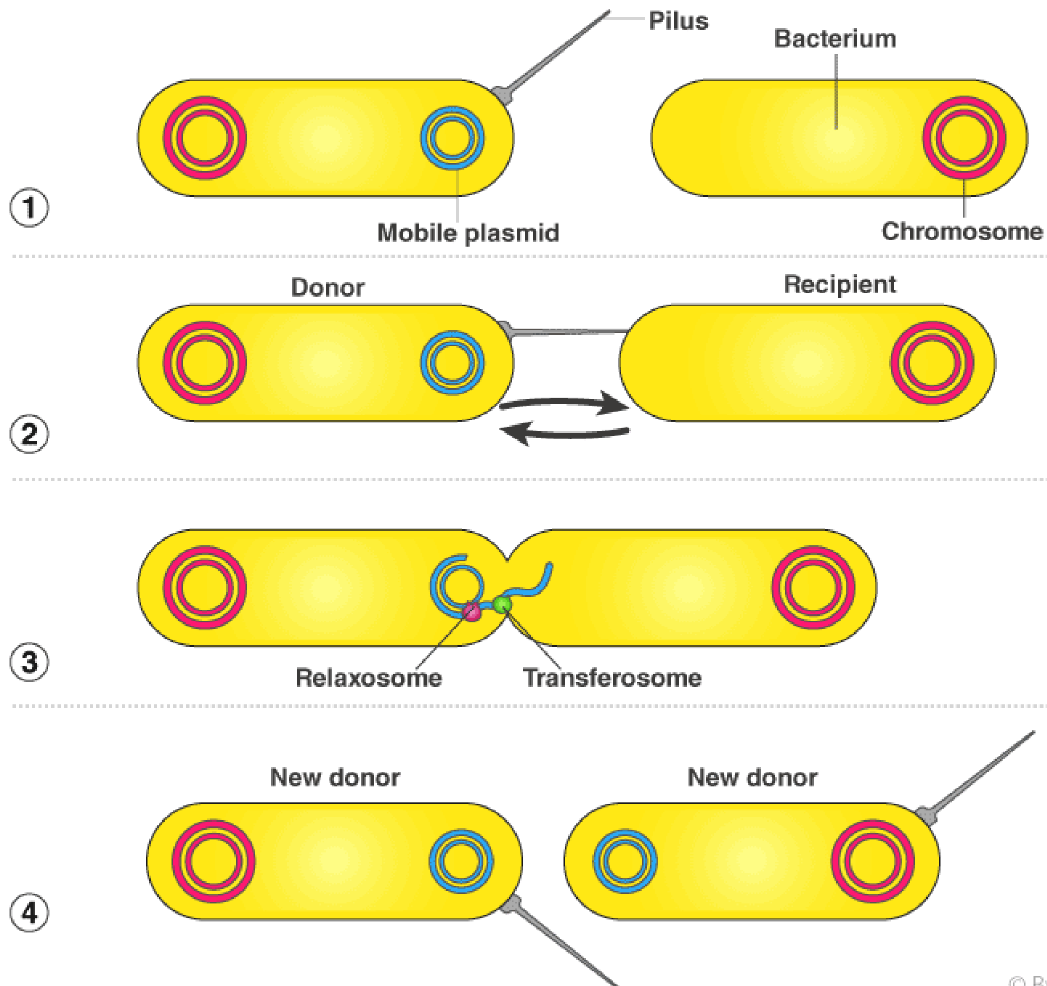
There are numerous bacteria found on planet earth. They divide quickly by binary fission producing identical daughter cells. Thus, the genetic information is transferred from the mother to the offspring and is known as vertical transmission.

The mutations are transferred from one bacteria to another through horizontal transmission. There are three different types of horizontal transmission for the transfer of genetic information.

- Conjugation
- Transduction
- Transformation

Bacterial Conjugation

Conjugation is the method of transfer of genetic material from one bacteria to another placed in contact. This method was proposed by Lederberg and Tatum. They discovered that the F-factor can move between *E.coli* cells and proposed the concept of conjugation.



There are various conjugal plasmids carried by various bacterial species. Conjugation is carried out in several steps:

- Mating pair formation
- Conjugal DNA synthesis
- DNA transfer
- Maturation

Also Read: R-Factor

Mechanism of Bacterial Conjugation

Bacterial conjugation involves the following steps:

Pilus Formation

The donor cells (F+ cells) form a sex pilus and begin contact with an F- recipient cell.

Physical Contact between Donor and Recipient Cell

The pilus forms a conjugation tube and enables direct contact between the donor and the recipient cells.

Transfer of F-Plasmid

The F-factor opens at the origin of replication. One strand is cut at the origin of replication, and the 5' end enters the recipient cell.

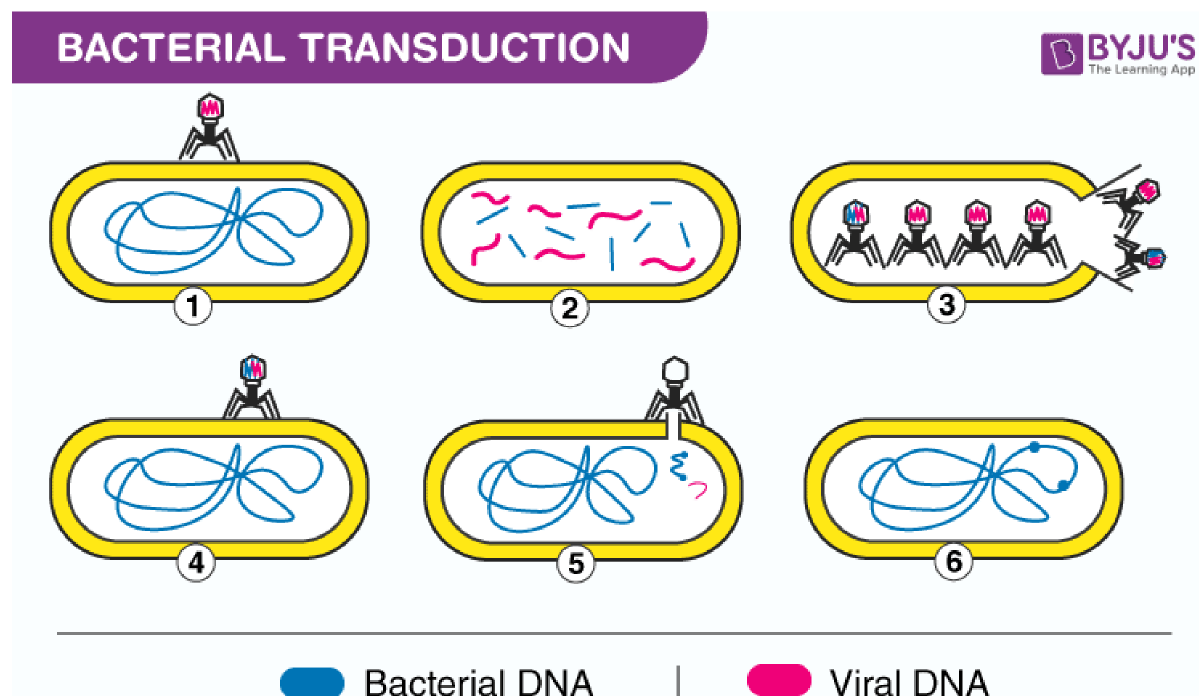
Synthesis of Complementary Strand

The donor and the recipient strand both contain a single strand of the F-plasmid. Thus, a complementary strand is synthesized in both the recipient and the donor. The recipient cell now contains a copy of F plasmid and becomes a donor cell.

Also Read: Difference between Virus and Bacteria

Bacterial Transduction

Transduction is the process of transfer of genes from the recipient to the donor through bacteriophage.



Transduction is of two types:

- Generalized Transduction
- Specialized Transduction

Generalized Transduction

In this type, the bacteriophage first infects the donor cells and begins the lytic cycle. The virus then develops its components using the host cell machinery. The host cell DNA is hydrolyzed into small fragments by the viral enzymes.

Small pieces of bacteria DNA is now integrated into viral genome. When the virus infects another bacteria the DNA is transferred into it.

Specialized Transduction

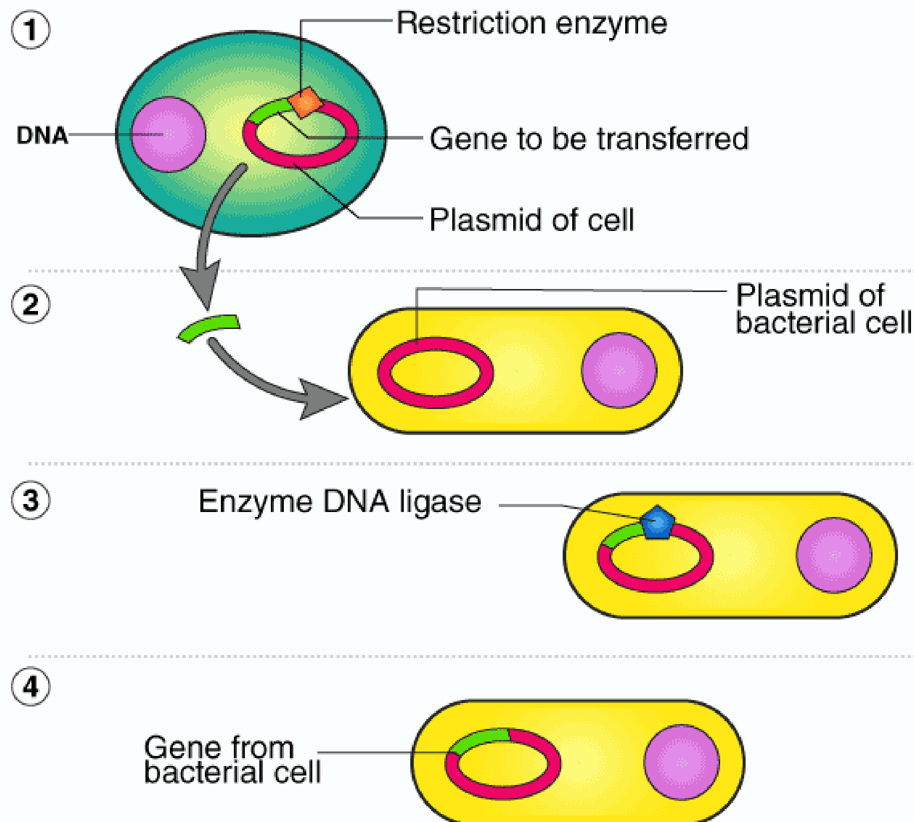
In this, only a few restricted bacteria are transferred from donor to recipient bacteria. This is carried out by temperate bacteriophage which undergoes the lysogenic cycle.

The virus enters the bacteria and integrates its genome within the host cell DNA. It remains dormant and passes on from generation to generation. When the lysogenic cell is exposed to some external stimulus, the lytic cycle begins.

The viral genome is induced in the host cell genome. Due to this, the phage genome sometimes carries the bacterial genome with it and integrates it into the genome of the recipient cell. Here, only the restricted genome has the possibility of entering the recipient cells.

Bacterial Transformation

Transformation is the process of DNA uptake by the bacteria from the surrounding environment. The cells that have the ability to uptake DNA are known as competent cells. This process was first reported in *Streptococcus pneumonia* by Griffith.



Bacterial Competence

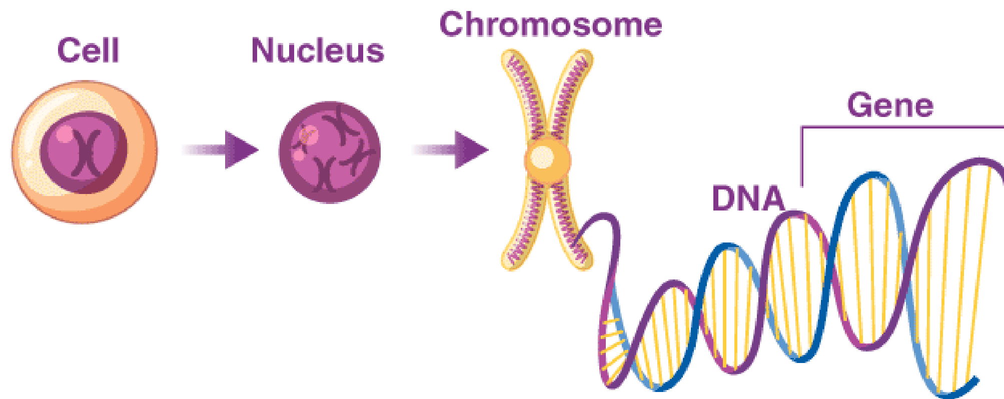
Not all bacteria are capable of taking up DNA from the surrounding environment. Such bacteria are made artificially competent. This is achieved by using chemicals and electrical pulses.

- **Chemicals-** The cells are chilled and made permeable in the presence of calcium phosphate. They are then incubated with the DNA and provided with a heat shock treatment that causes the DNA to enter the cells.
- **Electroporation-** The bacterial cells are subjected to electrical pulses to make them permeable and cause the DNA to enter into cells.

How to Identify Transformed Cells?

The bacteria are grown on an agar medium with antibiotics to check for transformed cells. Only the bacteria containing the antibiotic resistance gene will grow in the presence of antibiotics. The cells that survive and grow are transformed cells. The others are non-transformed.

UNIT 5



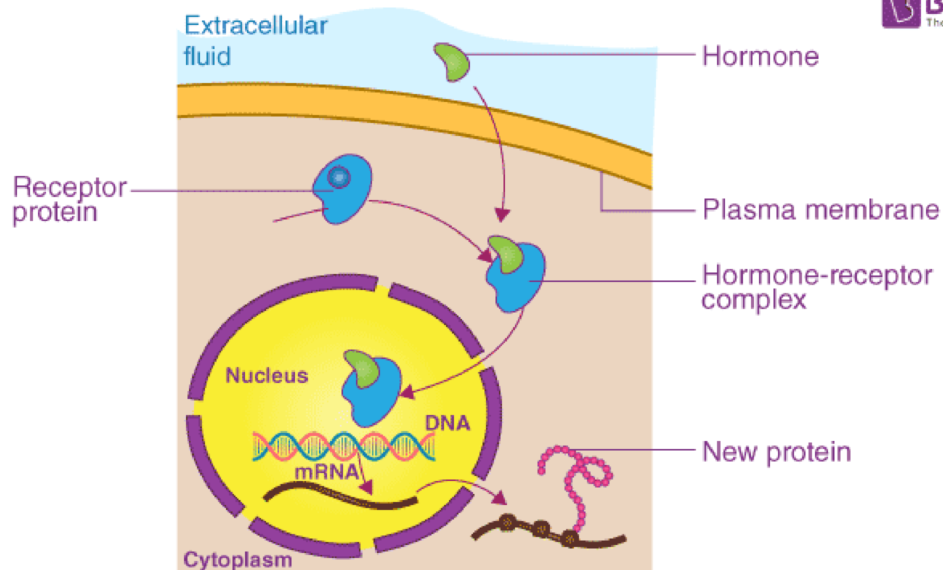
Genes

A gene is a part of DNA that codes for a particular protein. DNA is the information database of the cell and exists within the cell nucleus. It carries all the important genetic instructions that produce proteins required by our cells.

Each gene carries a particular set of instructions, which is usually in a coded format, used for an accurate function or for a distinct protein.

The said genes are first transcribed into mRNA and then get converted into a polypeptide chain. A polypeptide is then converted to a protein. All the hidden code inside our genes emerged as our physical traits, which are known as gene expression.

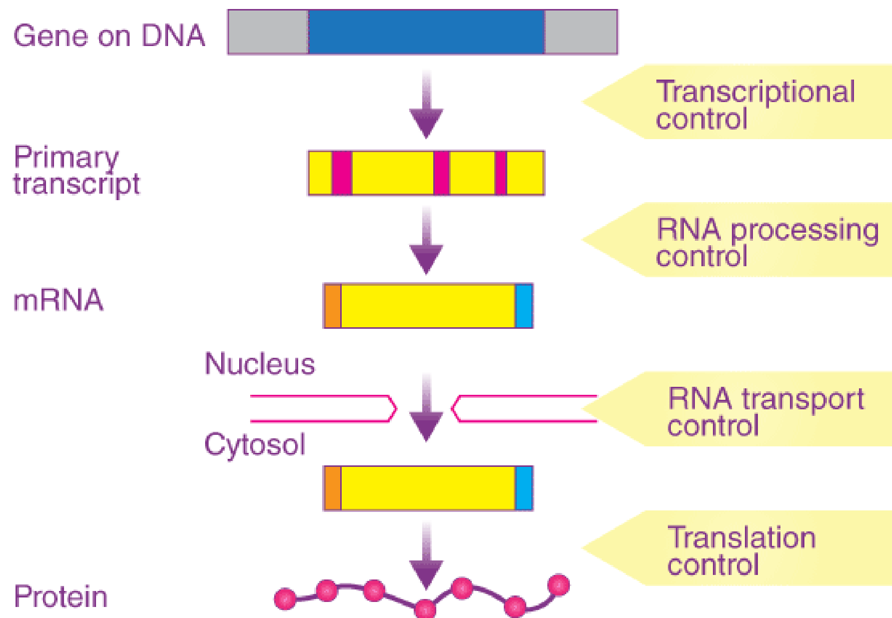
Gene Expression



This is a process where the gene's genetic codes are used in managing the protein synthesis that is required for our body to produce the cell structures. Genes that carry information required for the sequences of amino acids are termed structural genes. This process has two main steps:

1. Transcription- In this step, with the help of RNA polymerase enzymes, the messenger RNA is produced, resulting in the processing of mRNA molecules.
2. Translation- The main function of mRNA is to direct the synthesis of a protein resulting in the succeeding post-translational processing of the protein molecules.

Regulation of Gene Expression



Gene expression is the process by which the instructions present in our DNA are converted into a functional product, such as a protein. This process is a tightly coordinated process which allows a cell to respond to its changing environment.

During gene expression, genetic codes from the DNA code are converted into a protein with the help of translation and transcription. The genetic expression shows the process of the genetic makeup of an organism as its physical traits. In this process, the information flows from genes to proteins.

To understand this topic better, let us take the example of the Keratin genes. Keratin is a protein that helps in the formation of our hairs, nails, and skin. In most cases, these things grow at a continuous speed as our hairs, nails, and skin get worn down over a period of time. The production of excessive keratin could form many hairs on the skin, dry and hard skin, and thick and long nails. To avoid this, it is necessary to regulate the expression of the keratin gene.

Regulation of gene expression includes different mechanisms through which our cells manage the amount of produced protein by our genes.

Prokaryotic and Eukaryotic Transcription

Regulation of genes occurs differently, depending on the type of organisms- prokaryotic or eukaryotic. Eukaryotes refer to both multicellular and unicellular organisms like- animals,

fungi, plants, and protists possessing cells with nuclei and other organelles present within the cell. Prokaryotes are single-celled organisms like bacteria which do not have a well-defined nucleus. Regulation of prokaryotic and eukaryotic transcription is completely different as eukaryotes have a well-defined nucleus and prokaryotes do not.

Prokaryotic Transcription	Eukaryotic Transcription
Occurs in the cytoplasm.	Occurs within the nucleus.
The transcriptional unit has one or more genes	The transcriptional unit has just one gene
Transcription and translation are coupled	Transcription occurs in the nucleus, whereas translation in the cytoplasm
RNAs are released and processed in the cytoplasm	RNAs are processed in the nucleus and then released in the cytoplasm

Gene expression is basically the synthesis of the polypeptide chain encoded by a particular gene. Therefore, we can say that the expression of the gene can be quantified in terms of the amount of protein synthesised by the genes.

Let’s learn about the regulation of gene expression in prokaryotes with the example of Lac Operon.

Regulation of Gene Expression

We can infer that the gene regulation can take place at various steps of gene expression which includes the following:

- **Replication level** – Any error in copying the DNA may result in an altered expression.
- **Transcriptional level** – During transcription, any error in the polymerization may again lead to a change in expression of the gene.
- **Post-transcriptional level** – During the post-transcriptional modification i.e., RNA splicing, there may be some changes.
- **Translational level** – During translation, if there is an error in the attachment of mRNA to the tRNA molecules, there may arise some changes.

The regulation of the expression of the gene can be explained with the example of an enzyme, say peroxidase in bacteria. This enzyme catalyzes the hydrolysis of hydrogen peroxide to water and oxygen. Now, if the bacteria are moved to another environment where they fail to manufacture hydrogen peroxide, the enzyme no longer plays its role. In such a situation, the bacteria will not synthesize the enzyme anymore. Thus, we can say that the environmental, metabolic and physiological condition regulates the expression of genes.

Also Read: Protein Synthesis

Gene Regulation in Prokaryotes

Gene regulation in prokaryotes is most extensively observed at the initiation of transcription. Thus, the gene expression during transcription initiation is affected by regulation. The regulation usually takes place in the expression of the RNA polymerase at the promoter site. This affects the accessory proteins which bind to the recognition sites. These accessory proteins can regulate the promoter site in two ways:

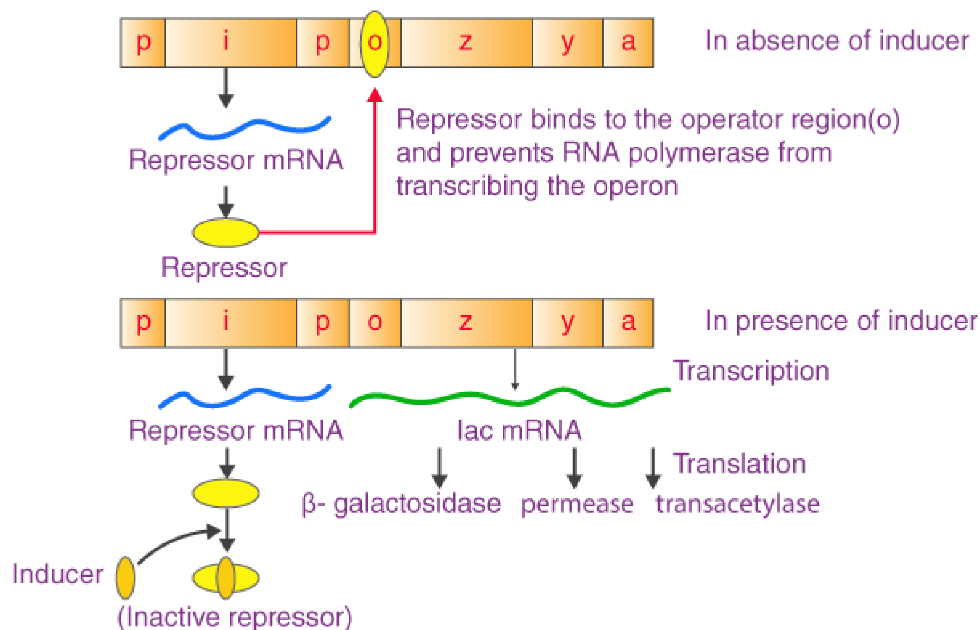
- Positive regulation by activators
- Negative regulation by repressors

In Operons, the operator is situated right next to the promoter where the regulator binds to control its entire functioning.

Lac Operon

"Lac operon is an operon or a group of genes with a single promoter that encode genes for the transport and metabolism of lactose in E.coli and other bacteria."

Lac Operon Concept



Gene regulation in prokaryotes can be explained with the help of the Lac Operon model. Here the alteration in physiological and environmental conditions can be observed leading to an alteration in expression in prokaryotes. It was observed by Jacob and Monod. The lac operon consists of:

- Regulatory gene *i* – It codes for the repressor protein.
- *z* gene – It codes for beta-galactosidase which catalyzes the hydrolysis of lactose into glucose and galactose.
- *y* gene – It codes for permease which regulates the lactose permeability in the cell.
- *a* gene – It codes for transacetylase which assists the enzyme beta-galactosidase.

Hence, all these genes help in lactose metabolism. In lac operon, lactose acts as an inducer. If lactose is provided in the medium for the bacteria, the regulatory gene is activated. The inducer will bind to the repressor protein and render it inactive which allows transcription of the operon. Thus, the lac operon is negatively regulated in this case.

Gene Regulation in Eukaryotes

Gene regulation in eukaryotes is regulated by transcriptional activators and repressors. The repressors bind to specific DNA sequences and inhibit transcription. In eukaryotes, transcription involves several steps. It occurs in both, nucleus (transcription) and cytoplasm (translation).

Lac operon Notes

- Lac operon contains genes involved in metabolism.
- The genes are expressed only when lactose is present and glucose is absent.
- The operon is turned on and off in response to the glucose and lactose levels: catabolite activator protein and lac repressor.
- The lac repressor blocks the transcription of the operon. In the presence of lactose, it stops acting as a repressor.
- catabolite activator protein activates the transcription of the operon, only when glucose levels are low.

TRP Operon

Tryptophan operon is found within the genome of *E.coli*, which carries a set of genes constructing an essential amino acid, **tryptophan**. Sometimes, it is also termed as trp operon. Unlike lactose or **lac operon**, trp operon is a **repressible system** discussed in this article.
A scientist named **Charles Yanofsky** and his co-workers explicitly studied the role of regulatory and structural genes of the trp operon. Trp operon aids biosynthesis of the amino acid tryptophan from a precursor molecule (**chorismic acid**). Tryptophan functions as an **effector molecule** that is required for building up the polypeptide chain.
The tryptophan is the **end product** of the biosynthetic pathway, whose combination or dissociation with the repressor protein can **turn on** or **turn off** the trp operon system. This post mainly describes the definition, structure and regulatory system of the tryptophan operon.

Content: Tryptophan Operon

1. Definition
2. Structure
3. Regulation
    - Repression
    - Derepression
    - Attenuation
4. Conclusion

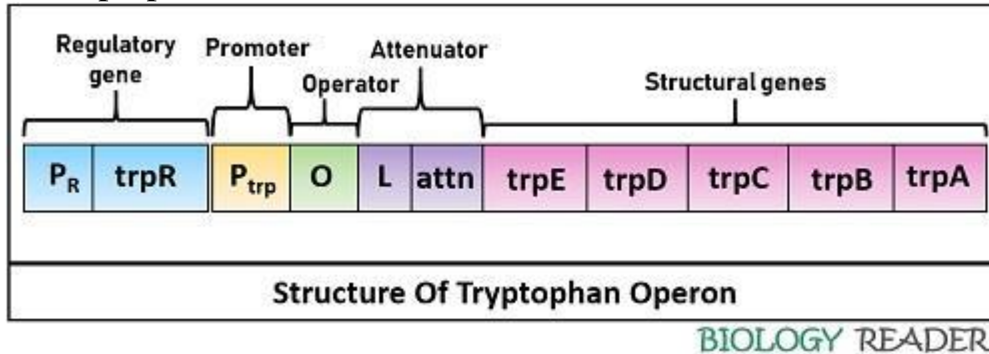
Definition

Trp operon is a repressible system that regulates gene expression for tryptophan biosynthesis according to the binding or uncoupling of a repressor with the operator region. The association or dissociation of the repressor protein strongly depends upon the **tryptophan level** in the surrounding.
A **high level** of tryptophan or effector molecule escalates the binding affinity of a repressor protein with the operator sequence, which in turn terminates the gene transcription. In contrast, a **low level** of tryptophan results in

detachment of the repressor from the operator region and allows gene expression.

Structure of Tryptophan Operon

The trp operon consists of:



- Promoter (P) region
- Operator (O) region
- Regulatory region
- Attenuator (A) region
- Structural genes (Trp A-E)

Promoter Gene

It refers to the region on a bacterial chromosome that comprises a specific nucleotide sequence, where an RNA polymerase can specifically bind to initiate transcription. The binding of a repressor protein with the operator inhibits the binding of **RNA polymerase** with the promoter sequence, which in turn terminates the transcription.

Operator Gene

It is the specific nucleotide sequence on chromosomal DNA of *E.coli*, where a **repressor** protein can bind by the association of an effector molecule or **tryptophan**. Here, the tryptophan molecule works as a corepressor that aids activation of aporepressor protein.

Regulatory Gene

Tryptophan operon is a repressor system, in which a regulatory gene of chromosome encodes **trp repressor** protein that recognizes the operator sequence. The repressor protein switches on the operon system at a low trp level in the surrounding while switches off the system at a high trp concentration in the environment.
Therefore, a repressor protein is associated with the synthesis of **five gene products**, which depends on the level of tryptophan in the surrounding. A

repressor becomes active or inactive in the presence or absence of a **corepressor** or effector molecule (tryptophan).

Attenuator Region

It is found in the middle of the operator region and structural genes. The attenuator region comprises leader sequences (**160 bp** in size) that regulate the transcription via attenuation.

Its mechanism is to attenuate transcription efficiency at sufficient tryptophan inside the bacterial cell by forming dimers. Oppositely, the RNAp can traverse through the attenuator region and transcribe genes necessary to build up trp at a low tryptophan level.

Structural Genes

Trp A, B, C, D and E are structural genes of the trp operon. The structural genes encoding for the enzymes and their subunits are necessary for tryptophan biosynthesis from chorismic acid.

- **trpE** encodes the enzyme Anthranilate synthase I.
- **trpD** encodes the enzyme Anthranilate synthase II.
- **trpC** encodes the enzyme N-5'-Phosphoribosyl anthranilate isomerase and Indole-3-glycerolphosphate synthase.
- **trpB** encodes the enzyme tryptophan synthase-B subunit.
- **trpA** encodes the enzyme tryptophan synthase-A subunit.

Regulation of Tryptophan Operon

Two mechanisms regulate the trp operon.

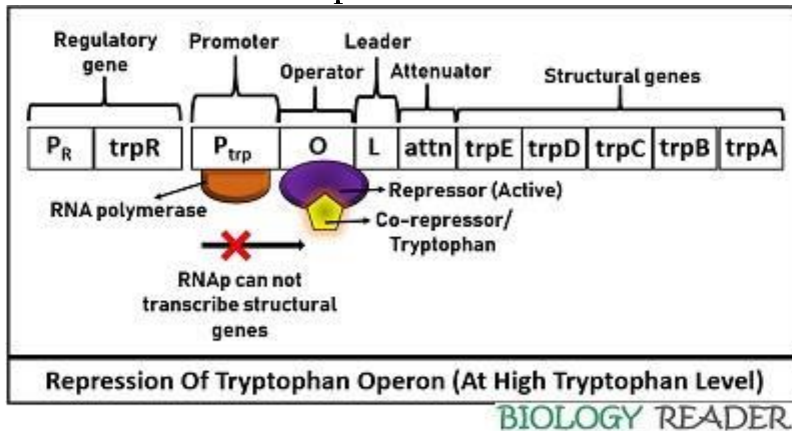
1. Repressor or derepression mechanism
2. Attenuation mechanism

Repression of Tryptophan Operon

It occurs when the trp level is **high** in the surrounding medium. In this case, the TrpR gene of the tryptophan operon releases **apo-repressor** (inactive) protein that alone cannot attach to the **operator** region. In the presence of a **corepressor** or tryptophan, the apo-repressor protein **activates** and blocks the RNA polymerase.

Thus, RNAP can not initiate structural genes transcription or inhibits enzymes necessary for the trp construction. Therefore, the respective RNA polymerase can neither bind with the operator gene nor transcribe structural genes at a high tryptophan concentration. Expression of trp operon during

availability of tryptophan indicates that the operon system will **switch off** to terminate the transcription.

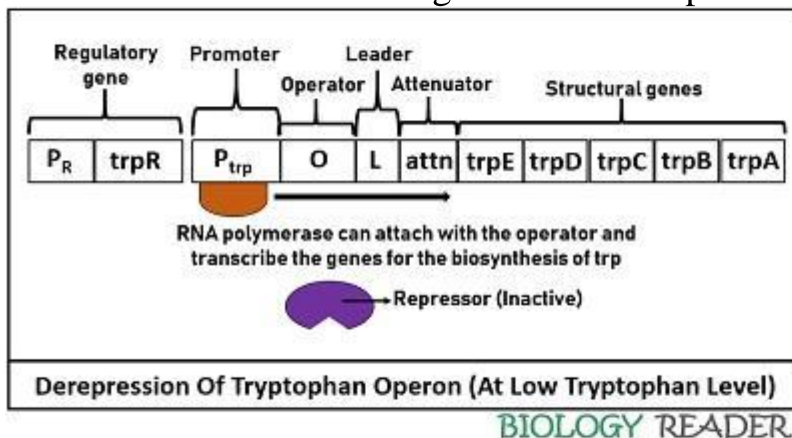


Repression Of Tryptophan Operon (At High Tryptophan Level)

Hence, the **repression** of trp operon is mediated via **complex** formed by the association of an allosteric repressor and an effector molecule. **Tryptophan** functions as a corepressor or an effector molecule that transforms the inactive apo-repressor to an active repressor, which ultimately adheres to the operator sequence to block gene expression.

Derepression of Tryptophan Operon

It occurs in the case of **low tryptophan** levels in the environment. Due to the lack of sufficient effector molecules, the active repressor protein attached to the operator region will detach. After its dissociation, the repressor remains **inactive** and functionless. As a result, RNA polymerase becomes free to transcribe structural genes further to produce tryptophan.

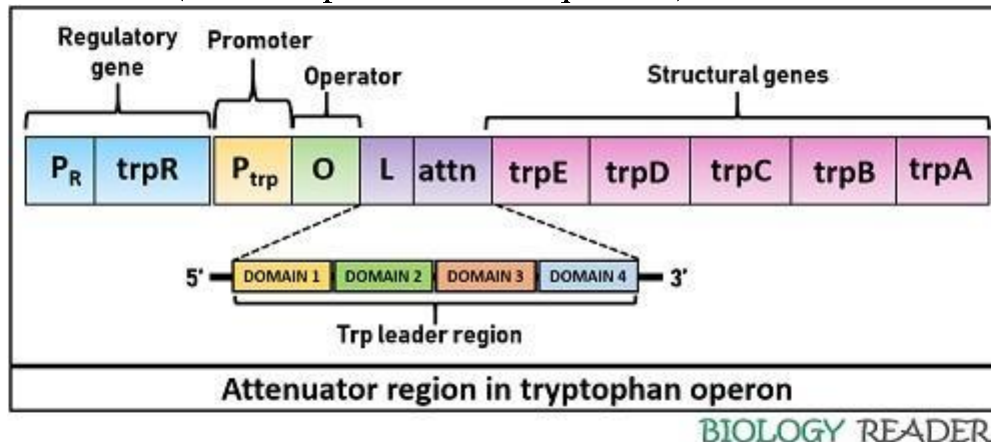


Derepression Of Tryptophan Operon (At Low Tryptophan Level)

The expression of trp operon during unavailability of tryptophan means that the operon system will **switch on** to conduct transcription of structural genes by the RNA polymerase. Hence, the **derepression** is achieved by the dissociation of repressor protein due to lack of trp or effector molecules, which together can form an active complex.

Attenuation

It is the second regulatory region of the trp operon controlled by the **trpL gene** or attenuator. A leader sequence controls the gene expression via an **attenuation** mechanism. It comprises a polypeptide sequence plus an attenuator (contains palindromic sequences).



Attenuator region in tryptophan operon

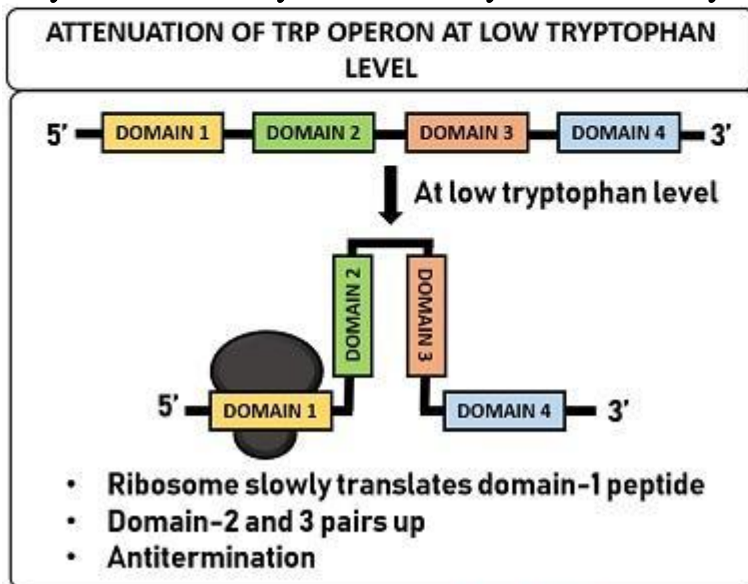
Once the bacterial DNA is transcribed into mRNA, the attenuator sequence can form **dimers** due to the pairing of palindromic sequences. There are **four domains** in the leader sequence, in which domain-3 can pair with either domain-2 or domain-4, and domain-1 can pair with domain-2. Besides, it comprises two trp residues.

The pairing of **domain-2 and 3** results in antitermination. Conversely, the pairing of **domain-3 and 4** causes a termination of trp biosynthesis. The presence of domain-4 (also called **attenuator**) is important to terminate the transcription because it only can facilitate **stem-loop** formation. The attenuation mechanism depends upon the pairing of the ribosome and the level of tryptophan inside a bacterial cell.

At low tryptophan level

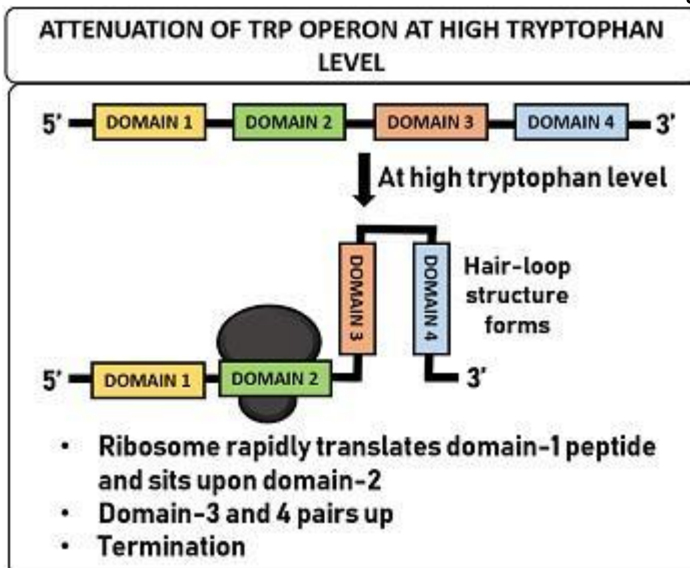
The ribosome sits at domain-1 of the mRNA transcript. Then, it translates the mRNA very slowly due to the low tryptophan level. As a result, domain-3 interacts with domain-2 due to the halt of the ribosome at domain-1. In such a case, the stem and loop structure will not form. As a result, the transcription

may continue to synthesize enzymes necessary for trp production.



At high tryptophan level

The ribosome rapidly translates the domain-1 and sits at the domain-2 when the concentration of tryptophan is high inside the cell. As a result, domain-3 associates with domain-4 and aids in forming a hair-loop structure.



The **dimerization** of domain-3 and 4 cause the RNAp to fall off and prevent mRNA from transcribing genes encoding enzymes for the trp biosynthesis. Therefore, an attenuator functions as a **barrier** at high trp concentration due to the pairing of self-complementary sequences.

Conclusion

Therefore, we can conclude that the tryptophan repressor and attenuation system decides when to switch on or switch off the expression of genes synthesizing trp, according to the **availability of tryptophan**.

Transposable elements definition

Transposable elements (TE) or transposons can be defined as small, mobile DNA sequences that move around chromosomes with no regard for homology, and insertion of these elements may produce deletions, inversions, chromosomal fusions, and even more complicated rearrangements.

- Transposons are mobile genetic elements that often carry an antimicrobial resistance gene.
- These elements can insert randomly, move from plasmids to the chromosome, and vice versa, and can be moved from one bacterium to another by conjugation, transformation, or transduction.
- Transposable elements make up a significant fraction of the genome and are responsible for much of the mass of DNA in a eukaryotic cell.
- Transposable elements were discovered by Barbara McClintock (1965) through an analysis of genetic instability in maize (corn).

Characteristics of transposable elements

Some salient features of transposable elements are:

1. These are the DNA sequences that code for enzymes which result in self-duplication and insertion into a new DNA site.
2. Transposons are involved in transposition events which include both recombination and replication, which usually generates two copies of the original transposable elements. One of the copies remains at the parent site, whereas the other one reaches the target site on the host chromosome.
3. The integrity of the target genes of these elements is invariably disrupted by the presence of those elements.
4. Because transposons carry the genes for initiation of RNA synthesis, some previously dormant genes might be activated.
5. A transposable element doesn't have a site for the origin of replication. As a result, it cannot replicate without the host chromosome as plasmids or phages.
6. There is no homology between the transposon and its target site for insertion. These elements can insert at almost any position in the host chromosome or a plasmid. Some transposons might seem likely to enter at some specific positions (hot spots), they barely insert at base-specific target sites.

Types of Transposable elements

Transposable elements are of the following two types:

1. Insertion Sequences (IS) or Simple Transposons

- The insertion sequences (IS) or simple transposons are shorter sequences (800 to 1500 bp) and do not code for proteins.
- These sequences carry the genetic information necessary for their transposition (*e.*, the gene for the enzyme transposase)
- Insertion sequences have been identified in bacteriophages, in F factor plasmid and many bacteria.

2. Transposons (Tn) or Complex Transposons

- Transposons (Tn) or complex transposons are several thousand base pair long, and have genes that code for one or more proteins which might include resistance factors in bacteria acting against antibiotics.
- A distinguishing character of a transposon is the presence of identical, inverted terminal repeat (IR) sequences ranging from 8 to 38 base pairs (b.p.)
- These inverted repeats are unique to different types of transposons.
- A short (less than ten b.p.) sequence is present on either side of a transposon.
- The insertion of transposon causes duplication of a singular target sequence, which then appear as direct repeats flanking the inserted transposable element.
- The flanking direct repeats are not considered part of the transposon. These repeat sequences themselves act like IS or IS-like segments.

Examples of Transposable elements

---

1. Tn 3 transposon of *E. coli*

- Tn 3 transposon has 4957 bp and contains three genes such as *tnp A*, *tnp R* and
- *Tnp A* codes for transposase having 1015 amino acids and required for transposition.
- *Tnp R* codes for a repressor (also called resolvase), containing 185 amino acids, which regulates the transposase.
- *Bla* codes for a β-lactamase enzyme which confers resistance to the antibiotic ampicillin.

2. Bacteriophage *Mu*

- The bacteriophage *Mu* (*Mu* = mutator) is a temperate bacteriophage having usual phage properties and could be regarded as a large transposon.
- It inserts itself into the *coli* chromosome at random locations resulting in a mutation.

3. Yeast *Ty* elements

- These elements are found in yeast *Saccharomyces cerevisiae* which are about 5900 base pairs long
- *Ty* elements are surrounded by five base-pair direct repeats which are created by the duplication of DNA present at the site of the *Ty*

Applications of transposable elements

- Transposable elements can be used as a genetic tool for the analysis of gene expression and protein functioning.
- These are used in genetic engineering to insert or remove specific genetic sequences, and also to cause frameshift mutation.
- The Tc1/mariner-class of TEs Sleeping Beauty transposon system is being studied for use in human gene therapy.

REFERENCE:

- https://www.britannica.com/science/colloid
- https://biotechnologymcq.com/homologous-recombination-holliday-model-and-dna-double-strand-repair-model-and-site-specific-recombination/#google_vignette
- https://www.bartleby.com/questions-and-answers/figure-20.2-the-holliday-model-for-homologous-recombination.-the-holliday-model-is-adapted-from-r.-h/c4ab2a97-bb06-48a8-abcf-2bdd43f5fe8b
- https://byjus.com/biology/mutation-genetic-change/
- https://byjus.com/neet/difference-between-spontaneous-and-induced-mutation/
- https://byjus.com/biology/bacterial-genetics/