Chapter 02

Molecular Basis of Inheritance

CONTENT

INTRODUCTION

- At the time of Mendel, the nature of those 'factors' regulating the pattern of inheritance was not clear.
- Over the next hundred years, the nature of the putative genetic material was investigated culminating in the realisation that DNA (Deoxyribonucleic acid) is the genetic material, at least for the majority of organisms.
- Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are the two types of nucleic acids found in living systems.
- DNA acts as the genetic material in most of the organisms. RNA though it also acts as a genetic material in some viruses, mostly functions as a messenger. RNA has additional roles as well. It functions as adapter, structural and in some cases as a catalytic molecule.

NUCLEIC ACIDS

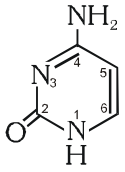
- **F. Meischer** discovered nucleic acid in nucleus of pus cell in 1869 and called it "**Nuclein**". The term nucleic acid was coined by "**Altman.**"
- Nucleic acids are polymer of **nucleotides**.
- Nucleotide = Nitrogen base + pentose sugar + phosphate
- Nucleoside = Nitrogen base + pentose sugar.

STRUCTURE OF POLYNUCLEOTIDE CHAIN: -

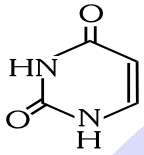
1. Nitrogen base:

- On the basis of structure, nitrogen bases are broadly of two types: –

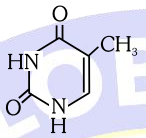
(a) **Pyrimidines**: It consists of one pyrimidine ring. Skeleton of ring composed of two nitrogen and four carbon atoms. e.g. Cytosine, Thymine and Uracil.	**(b)** **Purines:** It consists of two rings i.e. one pyrimidine ring (2N + 4C) and one imidazole ring (2N + 3C) e.g. Adenine and Guanine.



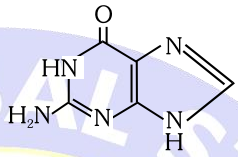
CYTOSINE



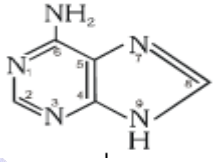
THYMINE



URACIL

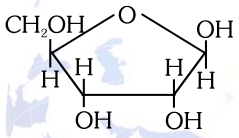


ADENINE

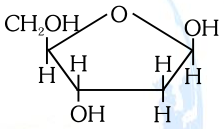


GUANINE

2. Pentose Sugar: –

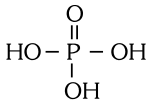


Ribose



Deoxyribose

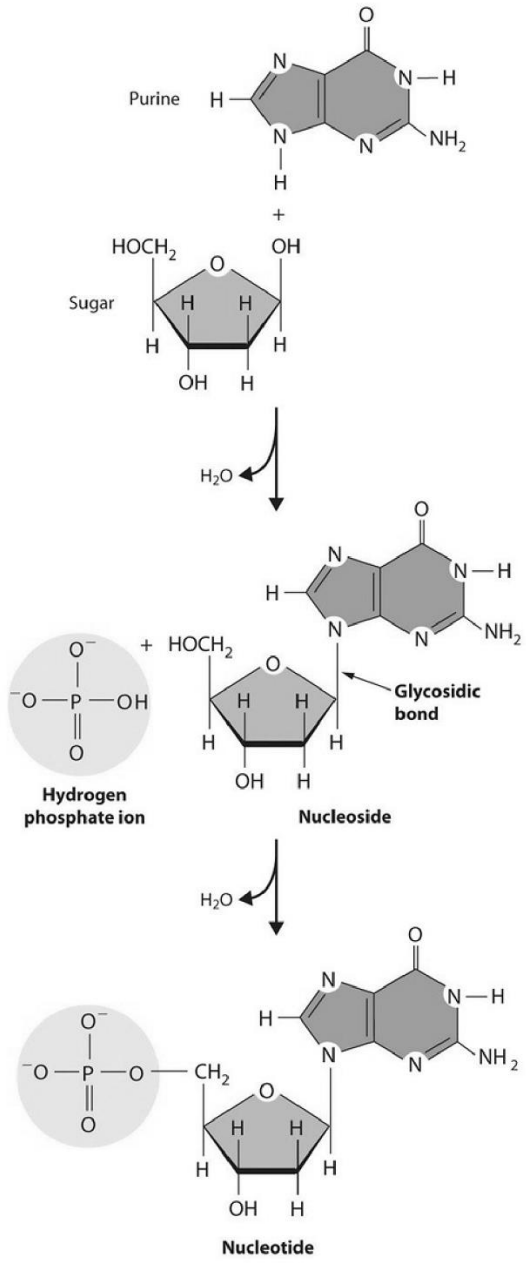
3. Phosphate → It is negatively charged and acidic in nature.



Phosphate

- Nitrogen base forms bond with OH of first carbon of pentose sugar through a **N-glycosidic linkage** to form a nucleoside. Nitrogen of **first place** (N_1) forms bond with sugar in case of pyrimidines while in purines nitrogen of **ninth place** (N_9) forms bond with sugar.
- Phosphate forms ester bond (covalent bond) with fifth Carbon of sugar to form a complete nucleotide.
- When a phosphate group is linked to OH of 5'C of a nucleoside through **phosphoester linkage** a corresponding nucleotide (depending on sugar) is formed.
- Two nucleotides are linked through 3'-5' phosphodiester linkage to form a dinucleotide and later on it forms a polynucleotide chain after adding more nucleotides in same manner.
- A polymer thus formed has at one end a free phosphate moiety at 5' -end of sugar, which is referred to as 5'-end of polynucleotide chain. Similarly, at the other end of the polymer the sugar has a free OH of 3'C group which is referred to as 3' -end of the polynucleotide chain. The backbone of a polynucleotide chain is formed due to sugar and phosphates. The nitrogenous bases linked to sugar moiety project from the backbone.

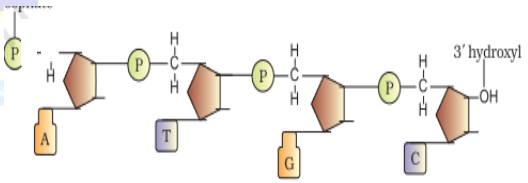
Formation of nucleotide



Difference between nucleoside and nucleotide

	Nucleoside	Nucleotide
(i)	Nucleoside is a compound formed by the union of a nitrogen base with a pentose sugar.	Nucleotide is a compound formed by the union of a nitrogen base, a pentose sugar and phosphate.
(ii)	It is a component of nucleotide	Nucleotide is formed through phosphorylation of nucleoside.
(iii)	It is slightly basic in nature	A nucleotide is acidic in nature

(c) Structure of one polynucleotide chain



A Polynucleotide chain

♦ TYPES OF NUCLEOSIDES AND NUCLEOTIDES:

(i) Nucleoside and Nucleotide of RNA

Base		Ribonucleoside		Ribonucleotide (5^o monophosphate)
Adenine (A)	+ Ribose →	Adenosine + Po_4	→	Adenylate (AMP)
Guanine (G)	+ Ribose →	Guanosine + Po_4	→	Guanylate (GMP)
Uracil (U)	+ Ribose →	Uridine + Po_4	→	Uridylate (UMP)
Cytosine (C)	+ Ribose →	Cytidine + Po_4	→	Cytidylate (CMP)

(ii) Nucleoside and Nucleotide of DNA

Base		Deoxyribonucleoside		Deoxyribonucleotide (5^o monophosphate)
Adenine (A)	+ Deoxyribose →	Deoxyadenosine + Po_4	→	Deoxyadenylate (dAMP)
Guanine (G)	+ Deoxyribose →	Deoxyguanosine + Po_4	→	Deoxyguanylate (dGMP)
Thymine (T)	+ Deoxyribose →	Deoxythymidine + Po_4	→	Deoxythymidylate (dTMP)
Cytosine (C)	+ Deoxyribose →	Deoxy cytidine + Po_4	→	Deoxycytidylate (dCMP)

DEOXYRIBONUCLEIC ACID (DNA)

- Most important biomolecule.
- DNA is a long polymer of deoxyribonucleotide (longest biomolecule).
- The length of DNA is usually defined as number of nucleotides (or a pair of nucleotide referred to as base pairs) present in it.
- For example, a bacteriophage known as $\phi \times 174$ has 5386 nucleotides, Bacteriophage lambda has 48502 base pairs (bp), *Escherichia* coli has 4.6×10^6 bp, and haploid content of human DNA is 3.3×10^9 bp.

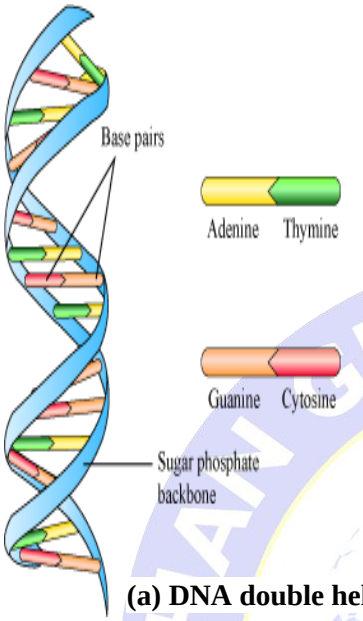
SPOT LIGHT

- DNA act as a genetic material in most of the organisms (Except RNA virus).
- It has some unique properties (like self replication, self repair, proof reading)
- In RNA, every nucleotide residue has an additional –OH group present at 2' -position in the ribose. Also, in RNA the uracil is found at the place of thymine (5-methyl uracil, another chemical name for thymine).

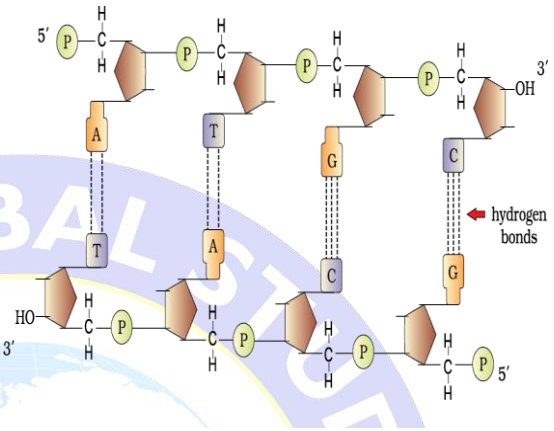
Structure of DNA:

- DNA is an acidic substance present in nucleus. It was first identified by **Friedrich Meischer** in 1869.
- He named it as '**Nuclein'**. However, due to technical limitation in isolating such a long polymer intact, the elucidation of structure of DNA remained elusive for a very long period of time.
- It was only in 1953 that **James Watson** and **Francis Crick**, based on the **X-ray diffraction data** produced by **Maurice Wilkins** and **Rosalind Franklin**, proposed a very simple but famous **Double Helix model** for the structure of DNA.
- **Watson** and **Crick** (1953) proposed a double helix model for DNA. For this model Watson, Crick and Wilkins were awarded by Noble Prize in 1962.
- One main hallmark (main point) of double helix model is complementary base pairing between Purine and Pyrimidine.
- According to this model, DNA is composed of two polynucleotide chains.
- Two strands of DNA are helically coiled like a revolving ladder. Back bone of this ladder (Reiling) is composed of phosphates and sugars while steps (bars) are composed of pairs of nitrogen bases.
- Both polynucleotide chains are complementary and antiparallel to each other.
- In both strand of DNA direction of phosphodiester bond is opposite. i.e. If direction of phosphodiester bond in one strand is 3'-5' then it is 5'-3' in another strand.

- Both strand of DNA held together by hydrogen bonds. These hydrogen bonds are present between nitrogen bases of both strand.
- Adenine binds to Thymine by two hydrogen bonds (A = T) and Cytosine binds to Guanine by three hydrogen bonds $(G \equiv C)$.



(a) DNA double helix



(b) Double stranded polynucleotide chain

- In a DNA molecule, one purine always pairs with a pyrimidine. Hence number of purines and pyrimidines are equal. This generates approximately uniform distance between the two strands of DNA.
- In DNA plane of one base pair stacks over the other in double helix. **This, in addition to H-bonds, confers stability of the helical structure of DNA.**

SPOT LIGHT

Chargaff's equivalency rule – According to Erwin Chargaff, in a double stranded DNA. amount of purine nucleotides is equal to amount of pyrimidine nucleotides.

Purine = Pyrimidine

$[A] + [G] = [T] + [C]$

$$\frac{[A]+[G]}{[T]+[C]} = 1$$

- Sugar deoxyribose and phosphate occurs in equimolar proportions.
- **Base ratio** = $\frac{[A]+[G]}{[T]+[C]} = 1$ = constant for a given species. i.e. species specific.
- Base ratio $\frac{A+T}{G+C}$ is used to identify the species. It is less than one in prokaryotes, e.g., E. coli = 0.92 and more than one in eukaryotes, e.g., Humans = =1.52.
- In a DNA A + T > G + C $\Rightarrow$ A – T type DNA. Base ratio of A – T type of DNA is more than one. eg. Eukaryotic DNA
- In a DNA G + C > A + T $\Rightarrow$ G – C type DNA. Base ratio of G – C type of DNA is less than one. eg. Prokaryotic DNA

SPOT LIGHT

Rosalind Franklin and Maurice Wilkins made X-ray diffraction study of B- DNA and draw some important conclusions. They are as follow

- B-DNA is a helically coiled molecule.
- The length of one complete turn is 34 Å (3.4 nm). This is called helix length.
- In one complete turn of DNA molecule, there are such 10 steps (10 pairs of nitrogen bases).
- Distance between two consecutive base pair is 3.4 Å (0.34 nm).
- Diameter of DNA molecule i.e. distance between phosphates of two strands is 20 Å (2 nm).

DETECTIVE MIND

Properties of DNA

1. **Denaturation and renaturation of DNA -** If a normal DNA molecule is placed at high temperature (80-90°C) then both strands of DNA will separate from each other due to breaking of hydrogen bonds. It is called DNA-denaturation.
- When denatured DNA molecule is placed at normal temperature then both strand of DNA gets attached and recoiled to each other. It is called renaturation of DNA.
- **Hyperchromicity -** When a double stranded DNA is denatured by heating then the denatured DNA molecule absorbs more amount of UV light, this phenomenon is called **hyperchromicity.**
- **Hypochromicity -** When denatured DNA molecule or strand cool down slowly then it becomes double stranded and it absorb less amount of UV light. This phenomenon is called **hypochromicity.**

2. **Temperature of melting (T_m)**
- Melting point of DNA depends on G – C contents. More G – C contents means more melting point.
- T_m of prokaryotic DNA > T_m of Eukaryotic DNA

3. **Palindromic DNA** – Sequence of nucleotides same from both ends.

→
C C G G T A C C G G
G G C C A T G G C C
←

4. **Sense and antisense strands of DNA**
- Out of two strand of DNA, only one strand participates in transcription, it is called **Antisense strand/ Non coding strand/Template strand.**
- Other strand of DNA which does not participate in transcription is called **Sense strand/Coding strand**.

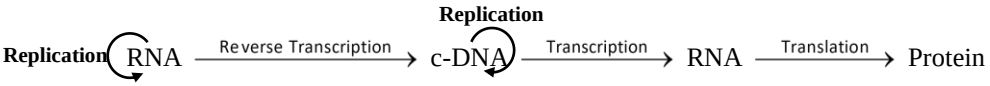
CENTRAL DOGMA

➢ Central dogma term was given by **Crick**. The formation (production) of mRNA from DNA and then synthesis of protein from it, is known as **Central Dogma**.

DNA $\xrightarrow{\text{Transcription}}$ RNA $\xrightarrow{\text{Translation}}$ Protein

MODIFIED CENTRAL DOGMA

➢ In some viruses the flow of information is in reverse direction from RNA to DNA



Reverse Transcription: –

➢ The formation of DNA from RNA is known as **Reverse - transcription**. It was discovered by **Temin and Baltimore** in Rous - sarcoma virus. So it is also called **Teminism.**

DETECTIVE MIND

- ss–RNA of Rous–Sarcoma virus (Retro virus) produces ds–DNA in host's cell with the help of enzyme reverse transcriptase (DNA–polymerase). This DNA is called c–DNA (Complimentary DNA). Sometimes, this DNA moves in host genome. Such mobile DNA is called "Retroposon" (Oncogene).

PACKAGING OF DNA HELIX

- Taken the distance between two consecutive base pairs as 0.34 nm (0.34×10^{-9} m), if the length of DNA double helix in a typical mammalian cell is calculated (simply by multiplying the total number of bp with distance between two consecutive bp, that is, **6.6×10^9 bp $\times$ 0.34 $\times$ 10^{-9} m/bp)**, it comes out to be approximately **2.2 metres**. A length that is far greater than the dimension of a typical nucleus (approximately 10^{-6} m).
- **In prokaryotes**, such as, *E. coli*, though they do not have a defined nucleus, the DNA is not scattered throughout the cell. DNA (being negatively charged) is held with some proteins called as **polyamines** (that have positive charges) in a region termed as **'nucleoid'**.
- The DNA in nucleoid is organised in large loops held by proteins.
- **In eukaryotes**, this organisation is much more complex. There is a set of positively charged, basic proteins called **histones**. A protein acquires charge depending upon the abundance of amino acids residues with charged side chains.
- Histones are rich in the basic amino acid residues **lysines** and **arginines**. Both the amino acid residues carry positive charges in their side chains.
- There are five types of histone proteins i.e. H_1, H_{2A}, H_{2B}, H_3 and H_4. Four of them occur in pairs to produce **histone octamer** (two copies of each H_{2A}, H_{2B}, H_3 and H_4).
- The negatively charged DNA is wrapped around the positively charged histone octamer to form a structure called **nucleosome**.
- DNA present between two adjacent nucleosomes is called **linker DNA**. A typical nucleosome contains **200 bp of DNA helix**.
- Nucleosomes constitute the repeating unit of a structure in nucleus called **chromatin**, thread-like stained (coloured) bodies seen in nucleus.
- The nucleosomes in chromatin are seen as **'beads-on-string'** structure when viewed under electron microscope (EM).
- The nucleosomes further coils to form **solenoids** (Diatmeter-30nm).
- Chromatin fibres are further coiled and condensed at metaphase stage of cell division to form chromosomes.
- The packaging of chromatin at higher level requires additional set of proteins that collectively are referred to as **Non-histone Chromosomal (NHC) proteins**.

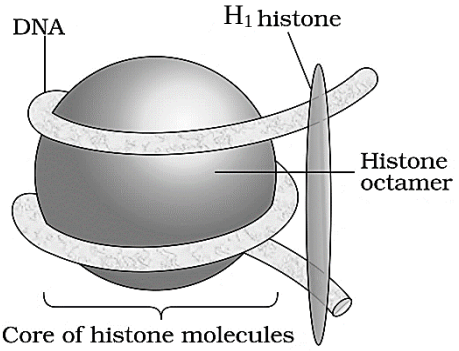
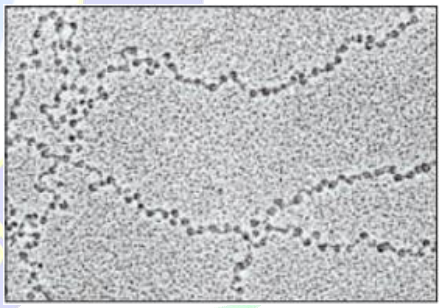


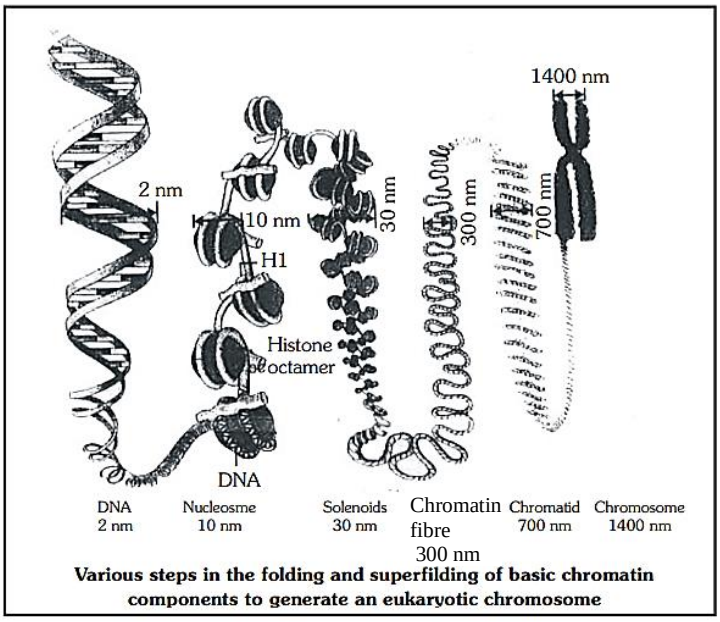
Figure : Nucleosome



EM picture – 'Beads-on-String'

SPOT LIGHT

In a typical nucleus, some region of chromatin are loosely packed (and stains light) and are referred to as **euchromatin**. The chromatin that is more densely packed and stains dark are called as **Heterochromatin**. Euchromatin is said to be transcriptionally active chromatin, whereas heterochromatin is said to be transcriptionally inactive chromatin.



Various steps in the folding and superfilding of basic chromatin components to generate an eukaryotic chromosome

GENETIC MATERIAL

♦ **The search for genetic material:**

- Even though the discovery of nuclein by Meischer and the proposition for principles of inheritance by Mendel were almost at the same time, but that the DNA acts as a genetic material took long to be discovered and proven. By 1926, the quest to determine the mechanism for genetic inheritance had reached the molecular level.
- Previous discoveries by Gregor Mendel, Walter Sutton, Thomas Hunt Morgan and numerous other scientists had narrowed the search to the chromosomes located in the nucleus of most cells. But the question of what molecule was actually the genetic material, had not been answered.

♦ **Evidence from bacterial transformation:**

- In 1928, **Frederick Griffith**, in a series of experiments with ***Streptococcus pneumoniae*** (bacterium responsible for pneumonia), witnessed a miraculous transformation in the bacteria.
- During the course of his experiment, a living organism (bacteria) had changed in physical form.
- When ***Streptococcus pneumoniae*** (pneumococcus) bacteria are grown on a culture plate, some produce **smooth shiny colonies (S)** while others produce **rough colonies (R)**.
- This is because the S strain bacteria have a mucous (polysaccharide) coat, while R strain does not. Mice infected with the S strain (virulent) die from pneumonia infection but mice infected with the R strain do not develop pneumonia.

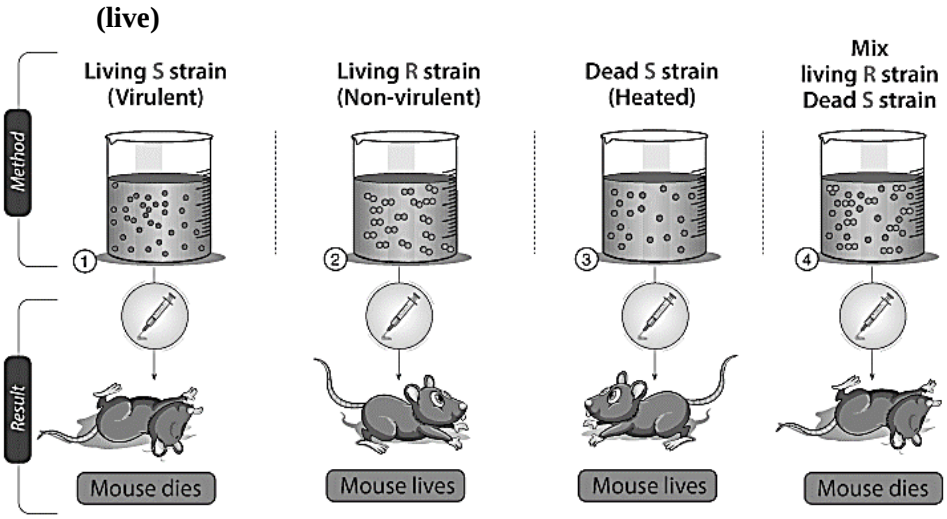
  S strain ⟶ Inject into mice ⟶ Mice die

  R strain ⟶ Inject into mice ⟶ Mice live

- Griffith was able to kill bacteria by heating them. He observed that heat-killed S strain bacteria injected into mice did not kill them. When he injected a mixture of heat-killed S and live R bacteria, the mice died.
- Moreover, he recovered living S bacteria from the dead mice.

S strain (heat-killed) ⟶ Inject into mice ⟶ Mice live

S strain (heat-killed) + R strain ⟶ Inject into mice ⟶ Mice die



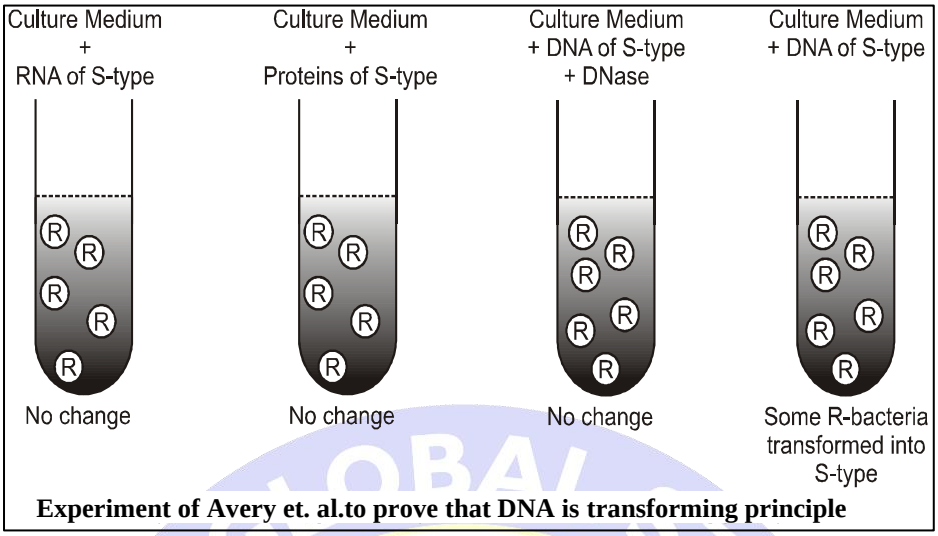
- He concluded that the R strain bacteria had somehow been transformed by the heat-killed S strain bacteria. Some 'transforming principle', transferred from the heat-killed S strain, had enabled the R strain to synthesise a smooth polysaccharide coat and become virulent.
- This must be due to the transfer of the genetic material. However, the biochemical nature of genetic material was not defined from his experiments.

DETECTIVE MIND

- There was no need of a mice in the above experiment this can also be done in testube by heating 'S' killed & mixing with 'R' type in test tube only.
- DNAs →DNAs are nucleic acids present in nucleus that carry the genetic information from parents to offsprings.
- DNase- It is a protein found in cytoplasm that catalyzes the hydrolytic cleavage of phosphodiester linkages in backbone of DNA.

BIOCHEMICAL CHARACTERISATION OF TRANSFORMING PRINCIPLE:

- **Oswald Avery, Colin MacLeod and Maclyn McCarty (1933 – 1944)**, determined the biochemical nature of 'transforming principle' in Griffith's experiment.
- They purified biochemicals (proteins, DNA, RNA, etc.) from the heat-killed S cells to see which ones could transform live R cells into S cells. They discovered that DNA alone from S bacteria caused R bacteria to become transformed.
- They also discovered that protein-digesting enzymes **(proteases)** and RNA-digesting enzymes **(RNases)** did not affect transformation, so the transforming substance was not a protein or RNA.
- Digestion with **DNase** did inhibit transformation, suggesting that the DNA caused the transformation. They concluded that DNA is the hereditary material, but not all biologists were convinced.



Experiment of Avery et. al.to prove that DNA is transforming principle

EVIDENCE FROM EXPERIMENTS WITH BACTERIOPHAGE:

- The unequivocal proof that DNA is the genetic material came from the experiments of **Alfred Hershey** and **Martha Chase (1952)**. They worked with viruses that infect bacteria called bacteriophages.
- The bacteriophage attaches to the bacteria and its genetic material then enters the bacterial cell. The bacterial cell treats the viral genetic material as if it was its own and subsequently manufactures more virus particles.
- **Hershey and Chase** worked to discover whether it was protein or DNA from the viruses that entered the bacteria.
- They grew some viruses on a medium that contained radioactive phosphorus (p^{32}) and some others on medium that contained radioactive **sulphur** (S^{35}).
- Viruses grown in the presence of radioactive phosphorus contained radioactive DNA but not radioactive protein because DNA contains phosphorus but protein does not.
- Similarly, viruses grown on radioactive sulphur contained radioactive protein but not radioactive DNA because DNA does not contain sulphur.
- **After labelling, three steps were followed i.e., infection**, **blending** and **centrifugation**.
  - **(i) Infection:** Both types of labelled phages were allowed to infect normally cultured bacteria in separate experiments.
  - **(ii) Blending:** These bacterial cells were agitated in a blender to break the contact between virus and bacteria.
  - **(iii) Centrifugation:** The virus particles were separated from the bacteria by spinning them in a centrifuge.
- After the centrifugation the bacterial cells showed the presence of radioactive DNA labelled with P^{32} while radioactive protein labelled with S^{35} appeared outside the bacterial cells i.e., in the medium. Labelled DNA was also found in the next generation of phage.
- Bacteria that were infected with viruses that had radioactive proteins were not radioactive. This indicated that proteins did not enter the bacteria from the viruses.

- **DNA is therefore the genetic material that is passed from virus to bacteria.**

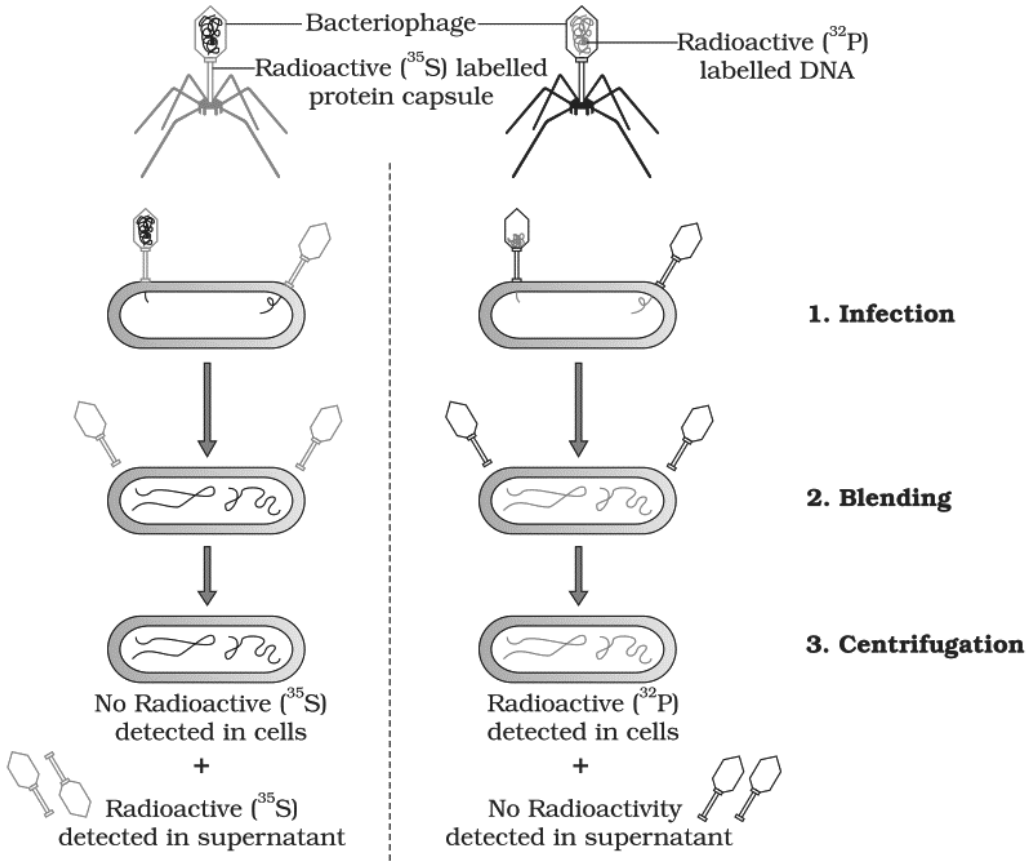


Figure: The Hershey-Chase experiment

PROPERTIES OF GENETIC MATERIAL (DNA VERSUS RNA) :

- From the foregoing discussion, it is clear that the debate between proteins versus DNA as the genetic material was unequivocally resolved from Hershey-Chase experiment.
- It became an established fact that it is DNA that acts as genetic material. However, it subsequently became clear that in some viruses, RNA is the genetic material (for example, Tobacco Mosaic viruses, QB bacteriophage, etc.).
- Answer to some of the questions such as, why DNA is the predominant genetic material, whereas RNA performs dynamic functions of messenger and adapter has to be found from the differences between chemical structures of the two nucleic acid molecules.
- A molecule that can act as a genetic material must fulfill the following criteria:
  (i) It should be able to generate its replica (Replication).
  (ii) It should chemically and structurally be stable.
  (iii) It should provide the scope for slow changes (mutation) that are required for evolution.
  (iv) It should be able to express itself in the form of 'Mendelian Characters'.
- If one examines each requirement one by one, because of rule of base pairing and complementarity, both the nucleic acids (DNA and RNA) have the ability to direct their duplications.
- The other molecules in the living system, such as proteins fail to fullfill first criteria itself.
- The genetic material should be stable enough not to change with different stages of life cycle, age or with change in physiology of the organism.
- Stability as one of the properties of genetic material was very evident in Griffith's 'transforming principle' itself that heat, which killed the bacteria, at least did not destroy some of the properties of genetic material.
- This now can easily be explained in light of the DNA that the two strands being complementary if separated by heating come together, when appropriate conditions are provided. Further, 2'-OH group present at every nucleotide in RNA is a reactive group and makes RNA labile and easily degradable. **RNA is also now known to be catalytic, hence reactive.**

- Therefore, **DNA chemically is less reactive and structurally more stable** when compared to RNA. Therefore, among the two nucleic acids, the DNA is a better genetic material.
- In fact, the **presence of thymine** at the place of uracil also confers **additional stability to DNA.**
- Both DNA and RNA are able to mutate. In fact, RNA being unstable, mutate at a faster rate. Consequently, viruses having RNA genome and having shorter life span mutate and evolve faster.
- RNA can directly code for the synthesis of proteins, hence can easily express the characters. DNA, however, is dependent on RNA for synthesis of proteins. The protein synthesising machinery has evolved around RNA.
- The above discussion indicate that both RNA and DNA can function as genetic material, but **DNA being more stable is preferred for storage of genetic information. For the transmission of genetic information, RNA is better.**

RNA WORLD:

- **RNA was the first genetic material**. There is now enough evidence to suggest that essential life processes (such as **metabolism, translation, splicing,** etc.), evolved around RNA.
- RNA used to act as a genetic material as well as a catalyst (there are some important biochemical reactions in living systems that are catalysed by RNA catalysts and not by protein enzymes).
- But, RNA being a catalyst was reactive and hence unstable. Therefore, DNA has evolved from RNA with chemical modifications that make it more stable.
- DNA being double stranded and having complementary strand further resists changes by evolving a process of repair.

DNA REPLICATION:

- DNA is the only molecule capable of self duplication so it is termed as a " Living molecule"
- DNA replication takes place in "S - Phase" of the cell cycle.
- Watson and Crick had immediately proposed a scheme for DNA replication while proposing the double helical structure of DNA. The scheme suggested that the two strands would separated and act as template for the synthesis of new complementary strands.
- After the completion of replication, each DNA molecule would have one parental and one newly synthesised strand. This scheme was termed as **semiconservative DNA replication**.

SPOT LIGHT

"It has not escaped our notice that the specific pairing we have postulated immediately suggest a possible copying mechanism for the genetic material" Watson and crick [1953]

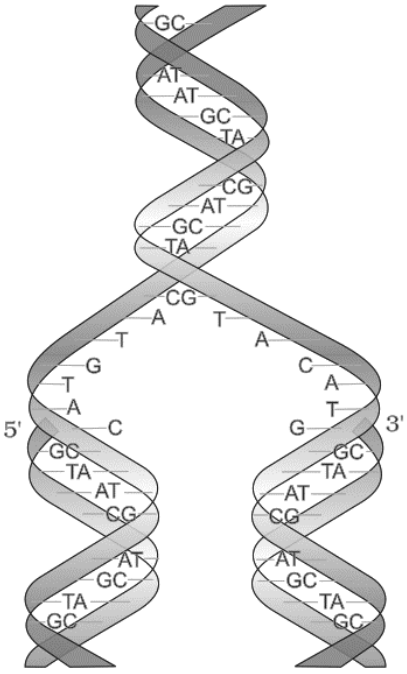


Fig. Watson-Crick model for semi conservative DNA replication

SEMI CONSERVATIVE MODE OF DNA REPLICATION

- Semi conservative mode of DNA replication was experimentally proved by **Matthew Meselson & Franklin Stahl** (1958) on ***E - Coli*** and ***Taylor*** on ***Vicia faba*** **(1958)**
- To prove this method, Taylor used Radiotracer Technique in which Radioisotopes (tritiated thymidine = $_1H^3$) were used. Meselson and Stahl used heavy isotope (N^{15}).

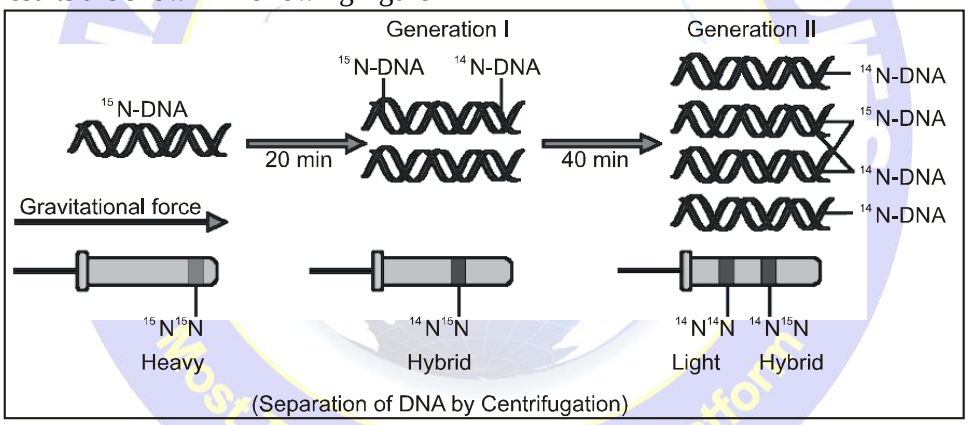
  Matthew Meselson and Franklin Stahl performed the following experiment in 1958:

  (i) They grew *E. coli* in a medium containing $^{15}NH_4Cl$ (^{15}N is the heavy isotope of nitrogen) as the only nitrogen source for many generations. The result was that ^{15}N was incorporated into newly synthesised DNA (as well as other nitrogen containing compounds). This heavy DNA molecule could be distinguished from the normal DNA by centrifugation in a cesium chloride (CsCl) density gradient (Please note that ^{15}N is not a radioactive isotope, and it can be separated from ^{14}N only based on densities).

  (ii) Then they transferred the cells into a medium with normal $^{14}NH_4Cl$ and took samples at various definite time intervals as the cells multiplied, and extracted the DNA that remained as double-stranded helices. The various samples were separated independently on CsCl gradients to measure the densities of DNA.

  (iii) Thus, the DNA that was extracted from the culture one generation after the transfer from ^{15}N to ^{14}N medium [that is after 20 minutes; *E. coli* divides in 20 minutes] had a hybrid or intermediate density. DNA extracted from the culture after another generation [that is after 40 minutes, II generation] was composed of equal amounts of this hybrid DNA and of 'light' DNA.

  The results are shown in following figure



- Very similar experiments involving use of radioactive thymidine to detect distribution of newly synthesised DNA in the chromosomes was performed on *Vicia faba* (faba beans) by Taylor and colleagues in 1958. The experiments proved that the DNA in chromosomes also replicate semi-conservatively.

DNA REPLICATION MECHANISM:

The following steps are included in DNA replication :-

(1) Activation of deoxyribonucleotides: -

- Energetically replication is a very expensive process. Deoxyribonucleoside triphosphate (activated deoxyribonucleotides) serve dual purposes in addition to acting as substrates they provide energy for polymerisation (Just like ATP, the two terminal phosphates in a deoxynucleoside triphosphates are high – energy phosphates).

(2) Initiation of DNA replication: -

- DNA replication always starts only at some specifics sites. This are known as initiation point/ origin of replication (Ori).

(3) Unzipping (Unwinding) :–

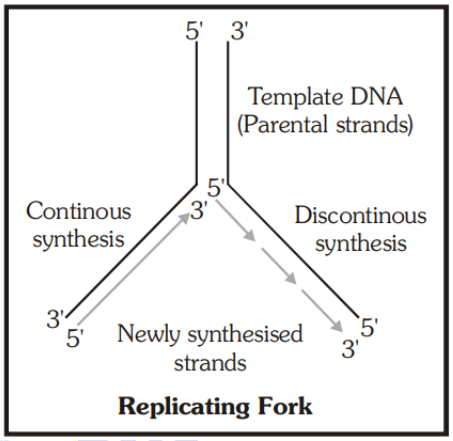
- The separation of 2 chains of DNA is termed as unzipping. And it takes place due to the breaking of H–bonds.

- The process of unzipping starts at a certain specific point which is termed as initiation point or origin of replication. In prokaryotes, there occurs only one origin of replication but in eukaryotes, there occurs many origin of replication i.e. unzipping starts at many points simultaneously.
- At the place of origin the enzyme responsible for unzipping (breaking the hydrogen bonds) is **Helicase**.
- For long DNA molecules, since the two strands of DNA cannot be separated in its entire length (due to very high energy requirement) the replication occur within a small opening of the DNA helix, referred to **replication fork.**
- **SSB (single stranded DNA binding protein)** prevents the reformation of H-bonds.
- **Topoisomerase** (in prokaryotes also called as **DNA gyrase**) release the tension arises due to supercoiling.

Note : The process of DNA replication takes a few minutes in prokaryotes and a few hours in Eukaryotes.

(4) Formation and elongation of new DNA strand: -

- To start the synthesis of new chain, special type of RNA is required which is termed as RNA Primer.
- The formation of RNA primer is catalysed by an enzyme - RNA Polymerase (primase). Synthesis of RNA–primer takes place in 5' → 3' direction.
- After the formation of new chain, this RNA is removed. For the formation of new chain Nucleotides are obtained from Nucleoplasm.
- In the nucleoplasm, Nucleotides are present in the form of triphosphates like dATP, dGTP, dCTP, dTTP etc.
- During replication, the 2 phosphate groups of all nucleotides are separated. In this process energy is yielded which is consumed in DNA replication.



Replicating Fork

SPOT LIGHT

In eukaryotes, the replication of DNA takes place at S-phase of the cell-cycle. The replication of DNA and cell division cycle should be highly coordinated. A failure in cell division after DNA replication results into **polyploidy**(a chromosomal anomaly).

- The formation of new chain always takes place in 5' - 3' direction. As a result of this, one chain of DNA is continuously formed and it is termed as **Leading strand**. The formation of second chain begins from the centre and not from the terminal points, so this chain is discontinuous and is made up of small segments called **Okazaki Fragments**.
- This discontinuous chain is termed as **Lagging strand**. Ultimately all these segments joined together and a complete new chain is formed.
- The Okazaki fragments are joined together by an enzyme **DNA Ligase**.
- The formation of new chains is catalysed by an enzyme **DNA- dependent DNA Polymerase.** In prokaryotes, it is of 3 types:

(1) DNA - Polymerase I: - This was discovered by KORNBERG (1957). So it is also called as 'Kornberg's enzyme'. Kornberg also synthesized DNA first of all, in the laboratory. This enzyme functions as exonuclease.

- It separates RNA - primer from DNA and also fills the gaps. It is also known as DNA-repairing enzyme.

(2) DNA - Polymerase II: - It is least reactive in replication process. It is also helpful in DNA-repairing in absence of DNA-polymerase-I and DNA polymerase-III

(3) DNA - Polymerase III: - This is the main enzyme in DNA - Replication. It is most important. The larger chains are formed by this enzyme. This is also known as Replicase.

DETECTIVE MIND

- All DNA polymerase I, II and III enzymes have 5' → 3' polymerisation activity and 3' → 5' exonuclease activity.
- In eukaryotes, DNA polymerases are of 5 types, these are DNA polymerase α, β, γ, δ and ε.
- Synthesis of leading or continuous strand is fast with the help of single primer; while the synthesis of lagging or discontinuous strand is slow and requires many primers.
- **Main polymerizing enzyme is DNA polymerase III.**
- Difference between DNAs and DNase is that DNAs means many DNA and DNase means DNA digestive enzymes.
- In the semi conservative mode of replication each daughter DNA molecule receives one chain of polynucleotides from the mother DNA - molecule and the second chain is synthesized.

RIBO NUCLEIC ACID (RNA)

RNA Molecules: -

- RNA or ribonucleic acid is a single chain poly-ribonucleotide. It carries the genetic code or hereditary information from DNA to cytoplasm for taking part in protein and enzyme synthesis.
- The back bone is formed of alternate residues of phosphate and ribose sugar. Phosphate combines with carbon 5' of its sugar and carbon 3' of next sugar similar to the arrangement found in DNA strand. Nitrogen bases are attached to sugars at carbon 1' of the latter.
- RNA has four types of nitrogen bases - adenine (A), guanine (G), both purines; cytosine (C) and uracil (U), both pyrimidines. The sequence of nitrogen bases is complementary to their sequence on DNA template.
- RNA molecules are produced by transcription of genes along the DNA. Some RNA species called messenger RNA, carry the genetic code for a protein.
- Others termed transfer RNA, serve to decode the sequence of ribonucleotides along the messenger RNA into the primary sequence of a protein. Still others play a structural role.
- Regardless of the function, all types of RNA have many structural features in common.

Types of RNA:

1. Genetic RNA:

- In the absence of DNA, sometime RNA works as genetic material and it transfers informations from one generation to next generation.
  e.g. Reo virus, TMV, QB bacteriophage.

2. Non-genetic RNA: 3 types:

(A) rRNA (B) tRNA (3) mRNA

- RNA functions as adapter, plays structural and in same cases as a catalytic role (Ribozyme)
- mRNA- Provides the template during translation.
  tRNA- Brings amino acids and reads the genetic code .
  rRNA- Plays structural and catalytic role in translation.

(A) Ribosomal RNA (rRNA):

- This RNA is 80% of the cell's total RNA. It is the most stable form of RNA. It is found in ribosomes and it is produced in nucleolus.
- There are **80S** type of ribosomes present in Eukaryotic cells. Their subunits are **60S** and **40S**. In **60S** sub unit of ribosome three types of rRNA are found – 5s, 5.8s, 28s while 40**s** sub unit of ribosome has only one type of rRNA i.e. 18**S**. So 80**S** ribosome has total 4 types of rRNA.
- Prokaryotic cells have 70**s** type of ribosomes and its subunits are 50S and 30**S**. 50**S** sub unit of ribosome contains **2 types** of rRNA i.e. 5**S** and 23**S.** 30**S** sub unit of ribosome has 16**S** type of rRNA. So 70**S** RNA has total 3 types of rRNA.

In eukaryotes → 80s → 60s → 5s, 5.8s, 28s; 40s

In Prokaryotes → 70s 70s → 50s → 5s, 23s; 30s → 16 s

FUNCTION:

- At the time of protein synthesis, rRNA provides attachment site to tRNA and mRNA and attaches them on the ribosome.

(B) Transfer – RNA (tRNA) :

- It is 10-15% of total RNA. It is synthesized in the nucleus by DNA. It is also known as soluble RNA **(sRNA)/Adapter** RNA. It is the **smallest RNA (4S)**.
- **Function** : At the time of protein synthesis, it acts as a carrier of amino-acids.
- **Structure :** The structure of tRNA is most complicated.
- A scientist named **Holley** presented **Clover leaf model** of its structure. In two dimensional structure the tRNA appears clover leaf like but in three dimensional structure (by Kim) it appears inverted L–shaped.
- The molecule of tRNA is of single stranded. There are three nucleotides present in a particular sequence at 3' end of tRNA and the sequence is CCA. All the 5' ends i.e. last ends are having G (guanine).
- 3' end is known as Acceptor end. tRNA accepts amino acids at acceptor points. Amino acid binds to 3' end by its – COOH group.
- The molecule of tRNA is folded and due to folding some complementary nitrogenous bases come across with each other and form hydrogen bonds. There are some places where hydrogen bonds are not formed, these places are known as loop.

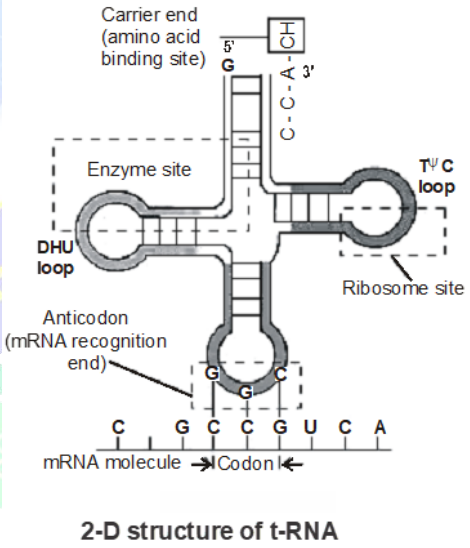
LOOPS OF tRNA:

(i) T Ψ C Loop or Attachment loop:

- This loop connects tRNA to the larger subunit of ribosome.

(ii) Recognition Loop (Anticodon loop): -

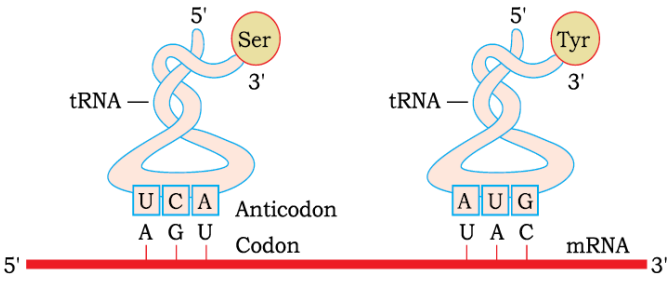
- This is the most specific loop of tRNA and different types of tRNA are different due to this loop. There is a specific sequence of three nucleotides called **Anticodon,** is present at the end of this loop.
- On the basis of Anticodon, there are total 61 types of tRNA or we can also say that there are 61 types of **Anticodon**. tRNA recognizes its place on mRNA with the help of Anticodon.
- The anticodon of tRNA recognises its complimentary sequence on mRNA. This complimentary sequence is known as codon.



2-D structure of t-RNA

(iii) DHU Loop:

- It is also known as **Amino - acyl synthetase recognition loop**. Amino - acyl synthetase is a specific type of enzyme. The function of this enzyme is to activate a specific type of amino acid. After activation this enzyme attaches the amino acid to the 3' end of tRNA.



tRNA - the adapter molecule

(C) Messenger RNA (mRNA):

- The mRNA is 1 - 5% of the cell's total RNA.
  The name mRNA was given by **Jacob** and **Monad.**
- The mRNA is produced by genetic DNA in the nucleus. This process is known as Transcription. It is least stable RNA.

DETECTIVE MIND

Types of mRNA – mRNA is of 2 types:

(1) Monocistronic :

- The mRNA in which genetic signal is present for the formation of only one polypeptide chain e.g. mostly in eukaryotes.

(2) Polycistronic :

- The mRNA, in which genetic signal is present for the formation of more than one polypeptide chains e.g. mostly in bacteria or prokaryotes.

TRANSCRIPTION

- Formation of RNA over DNA template is called transcription. Out of two strand of DNA only one strand participates in transcription and called ***''Antisense strand''***.

SPOT LIGHT

Question: **Why are both the strands not copied during transcription?**

Answer: If both strands act as a template during transcription they would code for RNA molecule with different sequence and If they code for proteins the sequence of aminoacid in these protein would be different Thus, one segment of the DNA code for two different proteins. And another reason that if the two RNA molecule produced they would be complementary to each other and form a ds RNA which prevent translation of RNA.

THE GENE:

- A gene is defined as the functional unit of inheritance. It is difficult to literally define a gene in terms of DNA sequence, because the DNA sequence coding for tRNA or rRNA molecule is also define a gene (But information of protein is present on the DNA segment which code mRNA. So generally it is reffered for it)
- The segment of DNA involved in transcription is ***''Cistron''***.

DETECTIVE MIND

- RNA polymerase enzyme is involved in transcription. In eukaryotes there are three types of RNA polymerases.
- **RNA polymerase–I** for 28S rRNA, 18S rRNA, 5.8S rRNA synthesis.
- **RNA polymerase–II** for m–RNA synthesis.
- **RNA polymerase–III** for t–RNA, 5S rRNA, SnRNA synthesis **(NEET 2023)** .

- Prokaryotes have only one type of RNA polymerase which synthesizes all types of RNAs.
- RNA polymerase of *E. Coli* has two parts -

(i) Sigma factor
- Small & Detachable part
- Help in recognition of initiation point

(i) Core enzyme
- Main enzyme
- forms RNA

TRANSCRIPTION UNIT:

- A transcription unit in DNA is defined primarily by the three regions in the DNA:
  (i) A Promoter
  (ii) The Structural gene
  (iii) A Terminator
- There is a convention in defining the two strands of the DNA in the structural gene of a transcription unit. Since the two strands have opposite polarity and the **DNA-dependent RNA polymerase** also catalyse the polymerisation in only one direction, that is, 5'→3', the strand that has the polarity 3'→5' acts as a template, and is also referred to as **template strand**.
- The other strand which has the polarity (5'→3') and the sequence same as RNA (except thymine at the place of uracil), is displaced during transcription.
- Strangely, this strand (which does not code for anything) is referred to as **coding strand**. All the reference point while defining a transcription unit is made with coding strand.
- To explain the point, a hypothetical sequence from a transcription unit is represented below:

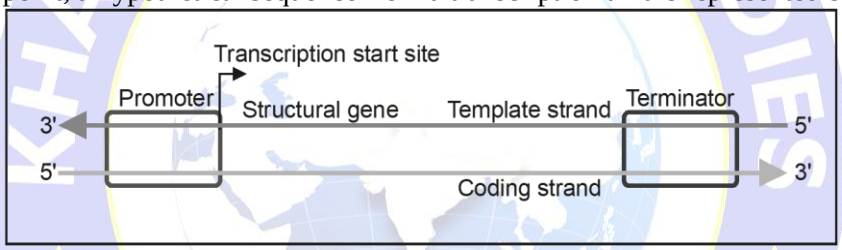


Fig. Schematic Structure of a Transcription Unit

- The **promoter** and **terminator** flank the **structural** gene in a transcription unit. The **promoter** is said to be located towards **5'-end (upstream)** of the structural gene (the reference is made with respect to the polarity of coding strand).
- It is a DNA sequence that provides binding site for RNA polymerase, and it is the presence of a promoter in a transcription unit that also defines the template and coding strands.
- By switching its position with terminator, the definition of coding and template strands could be reversed.
- The **terminator** is located towards **3'-end (downstream)** of the coding strand and it usually defines the end of the process of transcription.
- There are additional regulatory sequences that may be present further upstream or downstream to the promoter.

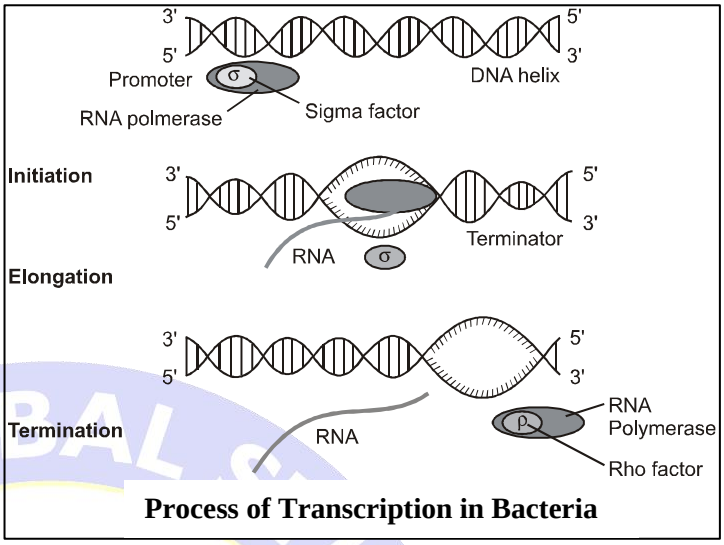
MECHANISM OF TRANSCRIPTION:

Following steps are present in transcription:

(1) Initiation:

- DNA has a ***''Promoter site or initiation site''*** where transcription begins and a ***''Terminator site''*** where transcription stops.

- **Sigma factor (σ)** recognises the promoter site of DNA. With the help of sigma factor RNA polymerase enzyme attached to a specific site of DNA called ***''Promoter site''***.
- In prokaryotes before the 10 bp from ***''Starting point''*** a sequence of 6 base pairs (TATAAT) is present on DNA, which is called ***''Pribnow box''***.
- In eukaryotes before the 20 bp from ***''Starting point''*** a sequence of 7 base pairs (TATAAAA) or (TATATAT) is present on DNA which is called **"TATA box** or **Hogness box"**
- RNA polymerase enzyme breaks H–bonds between two DNA strands and separates them. One of the strand takes part in transcription. **Transcription proceeds in 5' → 3' direction.**



Process of Transcription in Bacteria

- Ribonucleoside triphosphate come to lie opposite complementary nitrogen bases of anti sense strand. These Ribonucleotides present in the form of triphosphate ATP, GTP, UTP and CTP. When they are used in transcription, pyrophosphatase hydrolyse two phosphates from each activated nucleotide. This releases energy. This energy is used in the process of transcription.

(2) Elongation:

- RNA polymerase enzyme establishes phosphodiester bond between adjacent ribonucleotides. Sigma factor separates and core enzyme moves along the anti sense strand till it reaches terminator site.

(3) Termination:

- When RNA polymerase enzyme reaches at terminator site, it separates from DNA template.
- At terminator site on DNA, palindromic sequence are present.
- In most cases RNA polymerase enzyme can recognise the ***'Terminator site'*** and stop the synthesis of RNA chain, but in prokaryotes, it recognises the terminator site with the help of **Rho factor (ρ factor)**.
- Rho (ρ) factor is a specific protein which helps RNA polymerase enzyme to recognise the terminator site.
- In bacteria, since the mRNA does not require any processing to become active, and also since transcription and translation take place in the same compartment (there is no separation of cytosol and nucleus in bacteria), many times the translation can begin much before the mRNA is fully transcribed. Consequently, the transcription and translation can be coupled in bacteria

DETECTIVE MIND

At termination site palindromic sequence is present. When RNA, from this sequence folds upon it self & forms a loop known as hair pin loop

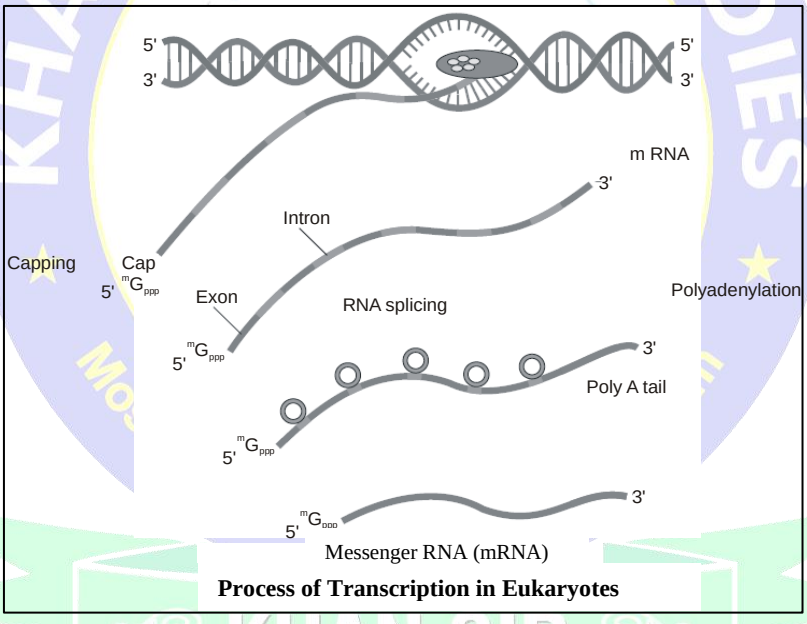
POST-TRANSCRIPTIONAL PROCESSING:

- It occurs in nucleus of eukaryotes. In eukaryotes RNA produced after transcription is not suitable for translation and it undergoes processing and converted into mRNA.
- In eukaryotes most of genes are **Split genes.** Gene which contains non functional part along with functional part is known as **split gene**.
- Non functional part (Intervening sequences) is called **intron** and functional part (Coding sequins or expressed sequences) is called **exon.** By transcription split gene produces a RNA which contains coding and non coding sequence and called **hn RNA (Heterogeneous nuclear RNA)**. This hnRNA is unstable.

- Now 7 methyl guanosine triphosphate is added to its 5' end, and a cap like structure is formed. It is called **capping** and 200-300 nucleotides of adenylic acid are added to its 3' end, which is called **tailing or polyadenylation**, Now it becomes stable.
- By the process of **RNA splicing** hnRNA produces functional m-RNA that is exonic RNA.
- Some specific proteins are also helpful in RNA - splicing called **'Small nuclear ribonucleoprotein' or 'SnRNP'.** These SnRNP proteins combine with some other proteins and SnRNA to form **spliceosome complex**.
- This spliceosome complex uses energy of ATP to cut the RNA, releases the non-coding part and joins the coding-part ny RNA ligase to produce functional RNA.
- Non coding part of hnRNA remained inside the nucleus and not translated into protein. Only coding part moves from nucleus to cytoplasm and translated into protein.

SPOT LIGHT

- Mostly Eukaryotic genes are example of split gene, **but gene which forms histone and interferon protein are non split gene.** It contains only and only exonic part.
- Most of prokaryotic genes are example of non- split gene.
- The split gene represent an ancient (primitive) feature of gene.
- Presence of intron is a primitive character and is reminiscent of antiquity.
- The splicing process represents the dominance of **RNA world**.



Process of Transcription in Eukaryotes

GENETIC CODE

- During replication and transcription a nucleic acid was copied to form another nucleic acid. Hence, these processes are easy to conceptualise on the basis of complementarity.
- The process of translation requires transfer of genetic information from a polymer of nucleotides to a polymer of amino acids.
- Neither does any complementarity exist between nucleotides and amino acids, nor could any be drawn theoretically.
- There existed ample evidences, though, to support the notion that change in nucleic acids (genetic material) were responsible for change in amino acids in proteins.
- This led to the proposition of a genetic code that could direct the sequence of amino acids during synthesis of proteins.

- The relationship between the sequence of amino acids in a polypeptide chain and nucleotide sequence of DNA or m–RNA is called **genetic code.**
- The term genetic code given by **George Gamow** and they are discovered by **Nirenberg**, **Mathai** and **Khorana**. There occur 20 types of amino acids which participate in protein synthesis.
- DNA contains information for the synthesis of any types of polypeptide chain. In the process of transcription, information is transferred from DNA to m–RNA in the form of complementary N_2–base sequences.
- The main problem of genetic code was to determine the exact number of nucleotide in a codon which codes for one amino acid. There are four types of N_2–bases in m–RNA (A, U, G, C) for 20 types of amino acids.
- If genetic code is ***singlet*** i.e. codon is the combination of only one nitrogen base, then only four codons are possible A, C, G and U. These are insufficient to code for 20 types of amino acids. (Singlet code = $4^1 = 4 \times 1 = 4$ codons)
- If genetic code is ***doublet*** (i.e. codon is the combination of two nitrogen bases) then 16 codons are formed. (Doublet code = $4^2 = 4 \times 4 = 16$ codons.) 16 codons are insufficient for 20 amino acids.
- **George Gamow**, a physicist, argued that since there are only 4 bases and if they have to code for 20 amino acids, the code should constitute a combination of bases.
- He suggested that in order to code for all the 20 amino acids, the code should be made up of three nucleotides. This was a very bold proposition, because a permutation combination of 4^3 ($4 \times 4 \times 4$) would generate 64 codons; generating many more codons than required.
- Providing proof that the codon was a triplet, was a more daunting task. The chemical method developed by **Har Gobind Khorana** was instrumental in synthesising RNA molecules with defined combinations of bases (homopolymers and copolymers).
- **Marshall Nirenberg's** cell-free system for protein synthesis finally helped the code to be deciphered.
- Severo Ochoa enzyme (polynucleotide phosphorylase) was also helpful in polymerising RNA with defined sequences in a template independent manner (enzymatic synthesis of RNA).
- Finally a checker-board for genetic code was prepared which is given in Table

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

Table: The Codons for the Various Amino Acids

CHARACTERISTICS OF GENETIC CODE:

(i) Triplet in Nature:

- A codon is composed of three adjacent nitrogen bases which specifies the one amino acid in polypeptide chain.

For Example:-

- In mRNA if there are total 90 N_2 – bases, then this mRNA determines 30 amino acids in polypeptide chain.

(ii) Universality :

- The genetic code is nearly universally. The same genetic code is present in all kinds of living organism including, bacteria, unicellular and multicellular organisms.

For Example:- From bacteria to human UUU would code for phenylalanine (phe)

Exception:- Mitochondrial codons and some protozoan codons.

(iii) Non – Ambiguous :

- Genetic code is non ambiguous i.e. one codon specifies only one amino acid and not any other.
- In this case one codon never code two different amino acids. Exception GUG codon which codes both valine and methionine amino acids.

(iv) Non – Overlapping:

- In DNA/mRNA codes are present one after the other and they never overlaps.

(v) Comma less :

- The codon is read in m-RNA in contiguous fashion There is no punctuation (comma) between the adjacent codon i.e. each codon is immediately followed by the next codon.
- If a nucleotide is deleted or added, the whole genetic code read differently.

DETECTIVE MIND

- A polypeptide chain having 50 amino acids shall be specialized by a linear sequence of 150 nucleotides. If a nucleotide is added in the middle of this sequence, the first 25 amino acids of polypeptide will be same but next 25 amino acids will be different.

(vi) Degeneracy of Genetic code :–

- There are 64 codons for 20 types of amino acids, so most of the amino acids (except two) can be coded by more than one codon. Single amino acid coded by more than one codon is called ***''Degeneracy of genetic code''***. This incident was discovered by **Baurnfield** and **Nirenberg**.
- Only two amino acids **Tryptophan** and **Methionine** are specified by single codon.

  UGG for Tryptophan

  AUG for Methionine
- All the other amino acids are specified or coded by 2 to 6 codons.
- Leucine, serine and arginine are coded or specified by 6–codons.
- **Leucine** = CUU, CUC, CUA, CUG, UUA & UUG
- **Serine** = UCU, UCC, UCA, UCG, AGU, AGC
- **Arginine** = CGU, CGC, CGA, CGG, AGA, AGG
- Degeneracy of genetic code is related to third position (3′ – end of triplet codon) of codon. The third base is described as **''Wobbly base''**.

(vii) Chain Initiation and Chain Termination Codon :–

- Polypeptide chain synthesis is signalled by two initiation codons **AUG or GUG**.
- AUG codes methionine amino acid in eukaryotes and in prokaryotes AUG codes N–formyl methionine.
- Some times GUG also functions as start codon it codes for valine amino acid normally but when it is present at starting position it codes for methionine amino acid.

- Out of 64 codons **3–codons are stopping** or nonsense or termination codon.
- Nonsense codons do not specify any amino acid.

UAA (Ochre)
UAG (Amber) } **Non-Sense Codons or Stop codons**
UGA (Opal)

- So only 61 codons are sense codons which specify 20 amino acids.

DETECTIVE MIND

WOBBLE HYPOTHESIS

- It was propounded by CRICK. All those codes which codes for same amino acids are differ only at 3rd base pair (wobble base) can be recognised by a common tRNA

TRANSLATION (PROTEIN SYNTHESIS)

Translation refers to process of polymerisation of amino acids to form polypeptide

♦ Translation Machinery

Translation requires a machinery which consists of ribosome, mRNA, tRNAs, aminoacyl tRNA synthetase (enzyme that helps in combining amino acid to particular tRNA) and amino acids.

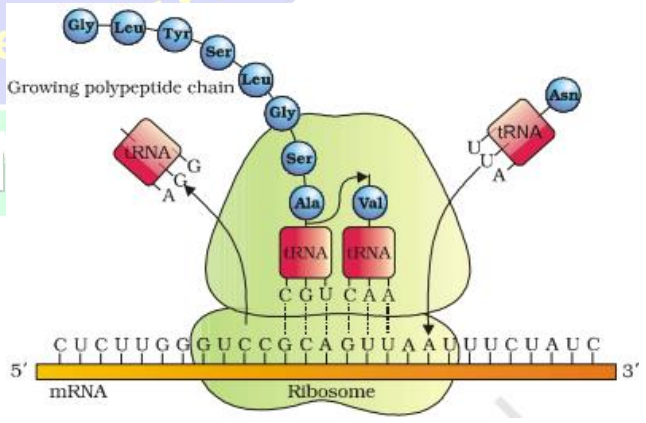
- **Initiator tRNA-** It is a specific tRNA for the process of initiation and there are no tRNAs for stop codons.
- **Ribosome-** The ribosome consists of structural RNAs and about 80 different proteins. It is responsible for protein synthesis. Ribosome exists as two subunits in its inactive stage.
- **Small subunit-** When the small subunit encounters an mRNA, translation of mRNA to protein begins.
- **Large subunit-** It consists of two sites, where amino acids can bind to and be close to each other for the formation of **peptide bond**. Ribosome also acts as a catalyst (23S rRNA in bacteria is the enzyme, **ribozyme)** for peptide bond formation.
- **Translational unit-** It is the sequence of RNA flanked by the start codon (AUG) and the stop codon in mRNA. It codes for the polypeptide to be produced.
- **Untranslated Regions-** (UTRs) These are some additional sequences in mRNA which are not translated. These are present at both the ends, i.e. at 5' end (before start codon) and at 3' end (after stop codon). These improve the efficiency of translation process.

♦ Stages of Protein Synthesis

Synthesis of proteins takes place in three stages which are as follows

1. **Initiation**: In prokaryotes, initiation requires ribosome (large and small subunits), mRNA, initiation tRNA and three Initiation Factors (IFs). For initiation to take place, the ribosome first binds to mRNA at the start codon (AUG) that is recognized only by the initiator tRNA.

- The formation of peptide bond requires energy and in first phase, the amino acids are activated in the presence of ATP and linked to their cognate tRNA by a process known as **charging of tRNA or aminoacylation of tRNA**. In the presence of ATP and mg^{2+}, amino acids become activated by binding with **aminoacyl tRNA synthetase** enzyme.



Translation

- The amino acids- AMP- enzyme complex is called an **activated amino acid.**

2. **Elongation of polypeptide chain:** In this step, another charged aminoacyl tRNA complex binds to the A-site of the ribosome.

- A peptide bond forms between carboxyl group (–COOH) of amino acid at P-site and amino group ($-NH_3$) of amino acid at A-site in a reaction catalyzed by the enzyme **peptidyl transferase.**
- During this stage, ribosome moves from one codon to another codon along the mRNA in the 5'→3' direction. Amino acids are then added one-by-one in the sequence of codons and translated into a polypeptide sequences, dictated by DNA and represented by mRNA.

3. **Termination of polypeptide:** When the A-site of ribosome reaches a termination codon, then no tRNA binds to the A-site of ribosome. At the end, a release factor binds to the stop codon and which terminates translation and releases the complete polypeptide from the ribosome.

REGULATION OF GENE EXPRERSION

- It is a molecular mechanism by which the expression of any gene is regulated as per requirement of cell.
- Regulation of gene expression refers to a very broad term that may occur at various levels. Considering that gene expression results in the formation of a polypeptide, it can be regulated at several levels.
- The genes in a cell are expressed to perform a particular function or a set of functions. For example, if an enzyme called **beta-galactosidase** is synthesised by *E.coli*, it is used to catalyse the hydrolysis of a disaccharide, lactose into galactose and glucose; the bacteria use them as a source of energy.
- Hence, if the bacteria do not have lactose around them to be utilised for energy source, they would no longer require the synthesis of the enzyme beta-galactosidase.
- Therefore, in simple terms, it is the metabolic, physiological or environmental conditions that regulate the expression of genes.
- The development and differentiation of embryo into adult organisms are also a result of the coordinated regulation of expression of several sets of genes.

DETECTIVE MIND

One the basis of requirement of expression, genes are of two types-

(1) **Constitutive genes (House-keeping genes):** These genes are expressed constantly, because their products are constant needed for cellular activity. e.g. genes for glycolysis, gene of ATPase enzyme.

(2) **Non-constitutive genes (Smart gene or Luxary gene):** These genes remain silent and are expressed only when the gene product is needed. They are switched 'on' or 'off' according to the requirement of cellular activities. Non-constitutive genes are of two types; **inducible** and **repressible**.

- The **inducible genes** are switched-on in presence of a chemical substance called **inducer**, required for the functioning of gene activity. The **repressible genes** continue to express themselves till a chemical, often an end product of the metabolism inhibits or represses their activity.
- The mechanism which stimulates the expression of certain genes and inhibits that of others is called **regulation of gene expression**.
- It is possible only if the organism has a mechanism of regulating gene activity by allowing some to function and others to restrain their activity through switching on and switching off system. This means, the genes are turned 'on' or 'off' as per requirement.
- A set of genes is 'switched on' when enzymes are required to metabolise a new substrate. The enzymes produced by these genes metabolise the substrate.
- The molecules of metabolite that come to switch on of the genes are termed as **inducers** and the phenomenon is called **induction**.

(I) GENE REGULATION IN PROKARYOTES: -

The *Lac* Operon Concept: –

- In 1961, a geneticist **Francis Jocob** and **Jacques Monod** at the Pasteur Institute in Paris, proposed a mechanism called **operon model** for the regulation of gene action in *E.coli.*
- *An operon is a part of genetic material or DNA, which acts as a single regulated unit having one or more* (polycistronic)*structural genes-an operator gene, a promoter gene, a regulator gene.*

  Examples of operon- lac operon, trp operon, ara operon, his operon, val operon etc.

  An inducible operon system normally remains in switched off condition and begins to work only when the substance to be metabolised by it, is present in the cell. Inducible operon system generally occurs in catabolic pathways. *e.g.* Lac operon of *E. coli.*

  Active repressor + inducer = inactive repressor

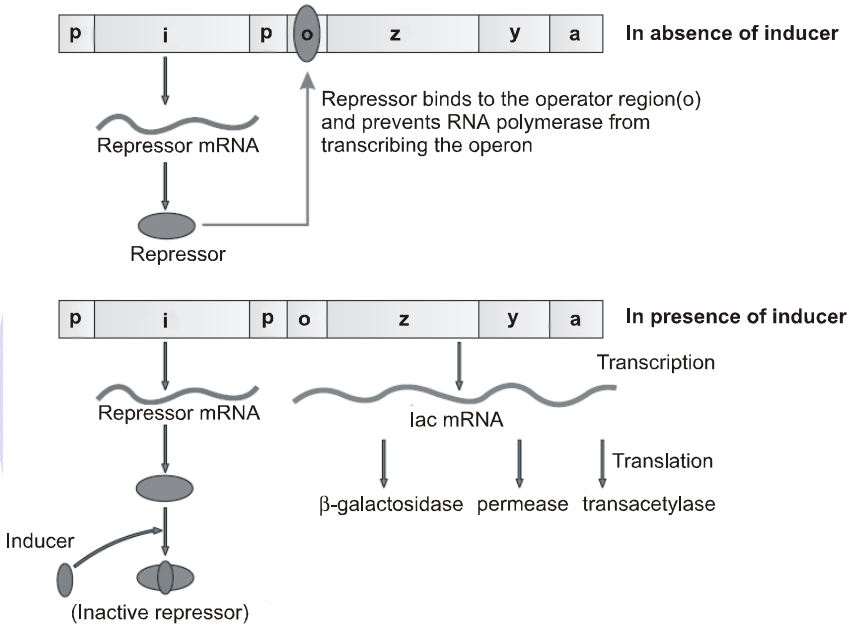


Fig.: The *Lac* Operon

An inducible operon system consists of four types of genes

(i) Structural genes : These genes synthesise mRNAs, which in turn synthesise polypeptide or enzyme over the ribosomes. An operon may have one or more structural genes. Each structural gene of an operon is called cistron. The lac operon (lactose operon) of *Escherichia coli* contains three structural genes (z, y and a).

- These genes occur adjacent to each other and thus are linked.
- They transcribe a polycistronic mRNA molecule (a single stretch of mRNA covering all the three genes), that helps in the synthesis of three enzymes-**β galactosidase** (breaks lactose into glucose and galactose), **lactose permease** (helps in entry of lactose in cell from outside) and **transacetylase** (transfers an acetyl group from acetyl Co A to β galactosidase).

(ii) Operator gene : It lies adjacent to the structural genes and directly controls the synthesis of *mRN*A over the structural genes. It is switched off by the presence of a repressor. An inducer can take away the repressor and switch on the gene that directs the structural genes to transcribe.

(iii) Promoter gene :

- Present just before the operator gene.

- This gene is the site for initial binding of RNA polymerase. When the operator gene is turned on, the enzyme RNA polymerase moves over it and reaches the structural genes to perform transcription.

(iv) Regulator gene (i-gene): It produces a repressor that binds to operator gene and stops the working of the operator gene i-gene is constitutive gene.

- **Repressor :** It is a protein, produced by the regulator gene. It binds to the operator gene so that the transcription of structural gene stops. Repressor has two allosteric site (1) operator gene (2) effective molecule (inducer/corepressor)
- **Inducer :** It is a chemical (substrate, hormone or some other metabolite) which after coming in contact with the repressor, forms an **inducer repressor complex**. This complex cannot bind with the operator gene, which is thus switched on. The free operator gene allows the structural gene to transcribe mRNA to synthesise the enzymes.
- **Mechanism:-**
  - The inducer for lac operon of *Escherichia coli* is lactose (in fact allolactose an isomer of lactose). When the sugar lactose is added to the culture of *E. coli,* a few molecules of lactose gets into the bacterial cells by the action of the enzyme permease.
  - A small amount of this enzyme is present in the cell even when the operon is not working. The lactose then induces the operon.
  - The I gene synthesises the repressor of the operon .
  - Repressor binds to the operator region of the operon preventing RNA polymerase from transcribing the operon.
  - However when inducer such as lactose or allolactose (real inducer) is present, it binds with the repressor and inactivates it
  - Now the repressor is inactivated, RNA polymerase is allowed access to the promoter and transcription proceeds.

- The **repressor molecule has key role in regulation** of lac-operon. Repressor molecule may be active or inactive. Active repressor may be rendered inactive by addition of an inducer while the inactive repressor can be made active by addition of a co-repressor.
- Because the product of regulator gene the repressor act by **shutling** off the transcripition of structural gene the operon model, as originally proposed by Jocob & Monad is referred as –negative control system.

(II) GENE REGULATION IN EUKARYOTES:-

- **In eukaryotes, the regulation could be exerted at several levels:**
  (i) transcriptional level (formation of primary transcript),
  (ii) processing level (regulation of splicing),
  (iii) transport of mRNA from nucleus to the cytoplasm,
  (iv) translational level.

MUTATION

- Sudden heritable change in genetic material of an organism or mutation is a phenomenon which results in alteration of DNA sequences and consequently results in changes in the genotype and the phenotype of an organism is called as **Mutation.**
- Mutation are discontinuous source of variation.

- Frequency of mutation at present is 1×10^{-6} (1 cell in : 1 million-cell). But it will increase in future due to pollution and destruction of ozone layer.
- Mutation and recombination – source of variations in DNA

DETECTIVE MIND

- Mutation word was given by **Hugo De Vries**. **De Vries** studied mutations in the plant ***Oenothera lamarckiana*** (evening primrose). It is a hybrid plant. De Vries proposed **Mutation theory of evolution**.
- Those mutation are only heritable which occur in **germinal cell** of an organism. While somatic mutations are non heritable. **Somatic mutations are also heritable in vegetative propagated plants.**

TYPES OF MUTATION:

1. Chromosomal mutation
2. Gene mutations

1. Chromosomal mutations : Change in number or structure of chromosome.

Types of chromosomal mutation

(I) Heteroploidy/Genomatic mutation → Change in chromosome number.

(II) Chromosomal aberration → Change in structure of chromosome.

(I) Heteroploidy / Genomatic mutation :

- Change in number of sets or chromosomes in sets. Two types -

(A) **Euploidy →** Change in number of sets.

(B) **Aneuploidy →** Change in number of chromosome in set.

(A) Euploidy :

- Change in number of sets of chromosome.

  Monoploidy (x) - Presence of one set of chromosome.

  Diploidy (2x) - Presence of two set of chromosome.
- **Polyploidy** - Presence of more than two sets of chromosomes failure of cytokinesis after telophase stage of cell division results in an increases in a whole let of chromosomes in an organism and this phenomenon is known as polyploidy common in plants.

  It may be : Triploidy (3x), Tetraploidy (4x), Pentaploidy (5x), Hexaploidy (6x) etc.

(B) Aneuploidy:

Failure of segregation of chromatin during cell division cycle results in the gain or loss of a chromosome (s), called aneuploidy.

Types of Aneuploidy :

(i) Hypoaneuploidy (loss)

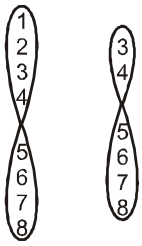
- **2n – 1 = monosomy** :- (Loss of one chromosome in one set). E.g. :- By Turners' syndrome
- **2n – 1 – 1 = double monosomy** (Loss of one chromosome from each set, but these are non homologus.)
- **2n – 2 = Nullisomy** (loss of two homologus chromosomes)

(ii) Hyperaneuploidy (add.)

- **2n + 1 = Trisomy :** Addition of one chromosome in one set. E.g. :- Down's syndrome

- **2n + 1 + 1 = Double Trisomy:** Addition of one chromosome in each set.
- **2n + 2 = Tetrasomy:** Addition of two chromosome in one set.
- Cause of aneuploidy is chromosomal non-disjunction means chromosomes fail to separate during meiosis.

(II) Chromosomal Aberrations: Change in structure of chromosome.



(A) Deletion: Loss of a part or segment of chromosome which leads to loss of some genes is called as deletion.

(B) Inversion: Breakage of chromosomal segment but reunion on same chromosome in reverse orders. It leads to change in distance between genes on chromosome or sequence of genes on chromosome so crossing over is affected.

(C) Duplication: Occurence of a chromosomal segment twice on a chromosome.

- **Example :** In *Drosophila* **"Bar eye character"** is observed due to duplication in X-chromosome. Bar eye is a character where eyes are narrower as compared to normal eye shape.

(D) Translocation : In this, a part of the chromosome is broken and may be joined with non homologous chromosome. This is also known as **Illegitimate** crossing over (illegeal crossing over)

2. Gene Mutation or point mutation:

Mutation arises due to change in a single base pair of DNA.

Two types: –

A. Substitution **B. Frame shift mutation.**

A. Substitution:

Replacement of one nitrogenous base by another nitrogenous base is called as substitution.

- It causes change in one codon in genetic code which leads to change in one amino acid in structure of protein.

eg. **Sickle cell anaemia** (In it, a change of single base pair in the gene for **beta globin chain** that results in change of amino acid residue **glutamate to valine**)

B. Frame shift mutation/Gibberish mutation:

Loss or addition of one or rarely more than one nitrogenous bases in structure of DNA.

Frame shift mutation is of two types.

(i). Addition (insertion):

- Addition of one or rarely more than one nitrogenous bases in structure of DNA.

(ii). Deletion :

- Loss of one or rarely more than one nitrogenous bases in structure of DNA.

DETECTIVE MIND

- Due to frame shift mutation complete reading of genetic code is changed. It leads to change in all animo acids in structure of protein so a new protein is formed which is completely different from previous protein.
- So frame shift mutations are more harmful as compared to substitution. e.g. **Thallesemia (lethal genetic disorder)**
- Colchicine inhibits spindle formation thus results in polyploidy.

MUTAGENS

➢ Mutagens are those substances which cause mutations :-

They are of two types :-

1. Physical mutagens **2. Chemical mutagens**

1. **Physical mutagens :** Mainly includes radiations. Radiations are of two types-

   (i) Ionising :- α, β, γ, X-rays

   (ii) Non ionising : UV rays.

2. **Chemical mutagens :** Chemical mutagens are more harmful than radiations because body is not protected against chemicals. Source of chemical mutagens are food, air and water.

➢ Effect of radiation is localised, while chemical mutagens spread in complete body through blood circulation and when they reach in gonads they cause germinal mutation.

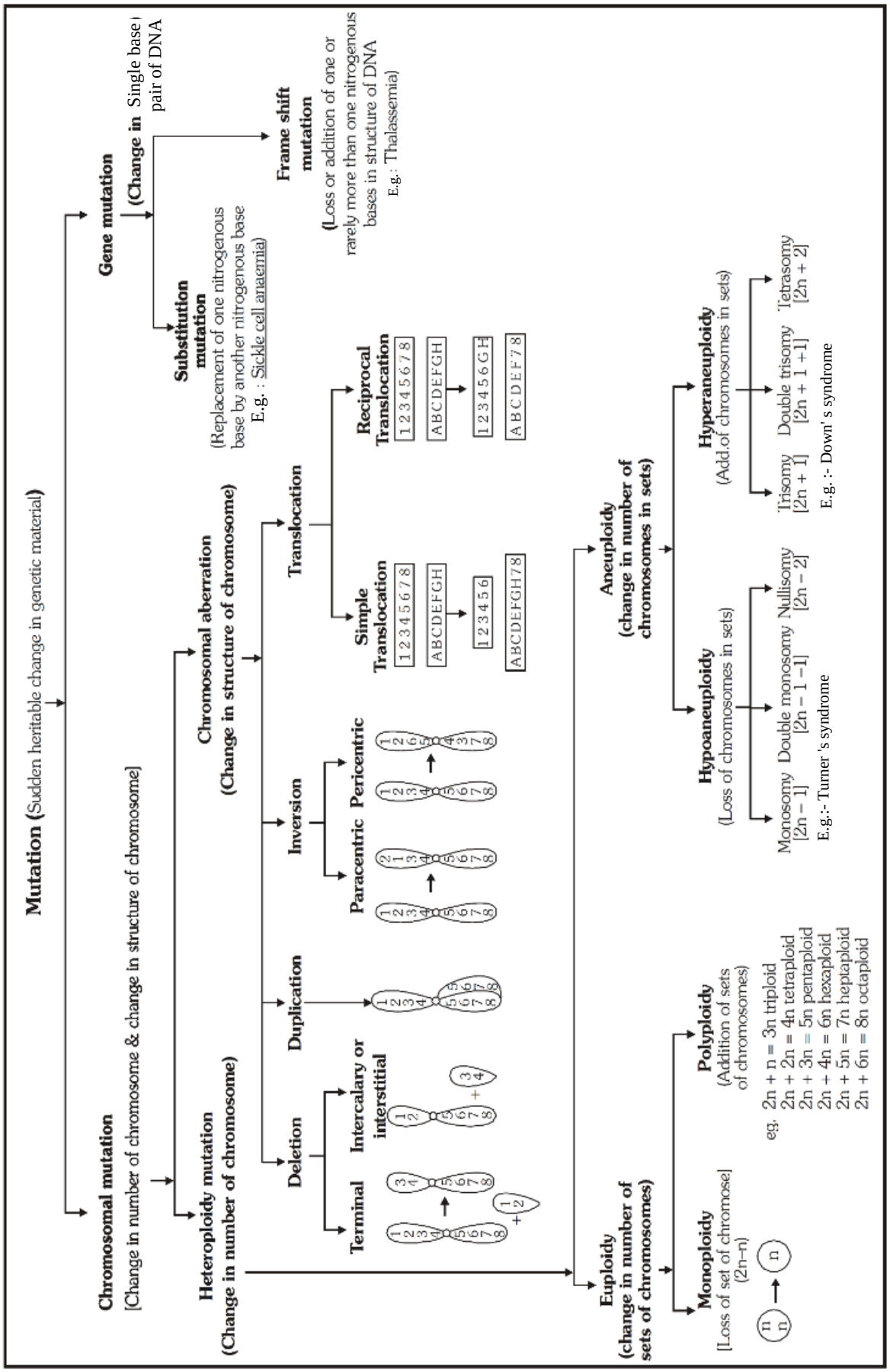
➢ Some chemical mutagens are :

(i) Mustard gas (first identified Chemical Mutagens)

(ii) Nitrous acid (HNO_2)

(iii) Alkylating agents

(iv) Colchicine


Mutation (Sudden heritable change in genetic material)
Gene mutation
(Change in Single base pair of DNA
Substitution mutation
(Replacement of one nitrogenous base by another nitrogenous base
E.g. : Sickle cell anaemia)
Frame shift mutation
(Loss or addition of one or rarely more than one nitrogenous bases in structure of DNA
E.g.: Thalassemia)
Chromosomal mutation
[Change in number of chromosome & change in structure of chromosome]
Heteroploidy mutation
(Change in number of chromosome)
Chromosomal aberration
(Change in structure of chromosome)
Deletion
Terminal
Intercalary or interstitial
Duplication
Translocation
Inversion
Paracentric
Pericentric
Simple Translocation
Reciprocal Translocation
1 2 3 4 5 6 7 8
A B C D E F G H
1 2 3 4 5 6
A B C D E F G H 7 8
1 2 3 4 5 6 G H
A B C D E F 7 8
Euploidy
(change in number of sets of chromosomes)
Monoploidy
[Loss of set of chromose]
(2n–n)
n → n
Polyploidy
(Addition of sets of chromosomes)
eg. 2n + n = 3n triploid
2n + 2n = 4n tetraploid
2n + 3n = 5n pentaploid
2n + 4n = 6n hexaploid
2n + 5n = 7n heptaploid
2n + 6n = 8n octaploid
Aneuploidy
(change in number of chromosomes in sets)
Hypoaneuploidy
(Loss of chromosomes in sets)
Monosomy [2n – 1]
Double monosomy [2n – 1 –1]
Nullisomy [2n – 2]
E.g.:- Turner 's syndrome
Hyperaneuploidy
(Add.of chromosomes in sets)
Trisomy [2n + 1]
Double trisomy [2n + 1 +1]
Tetrasomy [2n + 2]
E.g. :- Down 's syndrome

HUMAN GENOME PROJECT

- Genetic make-up of an organism or an individual lies in the DNA sequences. If two individuals differ, then their DNA sequences should also be different, at least at some places. These assumptions led to the quest of finding out the complete DNA sequence of human genome.
- With the establishment of genetic engineering techniques where it was possible to isolate and clone any piece of DNA and availability of simple and fast techniques for determining DNA sequences, a very ambitious project of sequencing human genome was launched in the year 1990.
- **Human Genome Project** (HGP) was called a mega project. HGP was a 13 year project that was coordinated by the **US Department of energy and the National institute of Health.**
- Human genome is said to have approximately 3×10^9 bp, and if the cost of sequencing required is US $ 3 per bp (the estimated cost in the beginning), the total estimated cost of the project would be approximately 9 billion US dollars.
- Further, if the obtained sequences were to be stored in typed form in books, and if each page of the book contained 1000 letters and each book contained 1000 pages, then 3300 such books would be required to store the information of DNA sequence from a single human cell.
- HGP was closely associated with the rapid development of a new area in biology called as **Bioinformatics.**

Goals of HGP:

Some of the important goals of HGP are as follows:

(i) Identify all the genes (Approxmately 20000 – 25000 genes in human DNA.) in human DNA.
(ii) Determine the sequences of the 3 billion chemical base pairs that make up human DNA.
(iii) Store this information in databases.
(iv) Improve tools for data analysis.
(v) Transfer related technologies to other sectors, such as industries.
(vi) Address the ethical, legal, and social issues (ELSI) that may arise from the project.

- During the early years of the HGP, the welcome trust (U.K) became a major partner. Japan, France Germany and China also provided additional contributions.
- The project was completed in 2003. Knowledge about the effects of DNA variations among individuals can lead to revolutionary new ways to diagnose, treat and someday prevent the thousands of disorders that affect human beings. Besides providing clues to understanding human biology, learning about non-human organisms, DNA sequences can lead to an understanding of their natural capabilities that can be applied toward solving challenges in health care, agriculture, energy production, environmental remediation.
- Many non-human model organisms, such as bacteria, yeast, ***Caenorhabditis elegans*** (a free living non-pathogenic nematode), ***Drosophila*** (the fruit fly), plants (rice and ***Arabidopsis***), etc., have also been sequenced.

Methodologies :

- The methods involved **two** major approaches.

(1) **Expressed Sequence Tags** (ESTs) - Identifying all the genes that expressed as RNA.
(2) **Sequence Annotation -** The blind approach of simply sequencing the whole set of genome that contained all the coding and non-coding sequence, and later assigning different regions in the sequence with functions.

- For sequencing, the total DNA from a cell is isolated and converted into random fragments of relatively smaller sizes (recall DNA is a very long polymer, and there are technical limitations in sequencing very long pieces of DNA) and cloned in suitable host using specialised vectors.
- The cloning resulted into amplification of each piece of DNA fragment so, that is subsequently could be sequenced with ease.
- The commonly used hosts were bacteria and yeast, and the vectors were called as **BAC** (bacterial artificial chromosomes), and **YAC** (yeast artificial chromosomes).
- The fragments were sequenced using automated DNA sequencers that worked on the principle of a method developed by **Frederick Sanger**.
- These sequences were then arranged based on some **overlapping regions** present in them. This required generation of overlapping fragments for sequencing.

- Alignment of these sequences was humanly not possible. Therefore, specialised computer based programmes were developed.
- These sequences were subsequently annotated and were assigned to each chromosome. The sequence of chromosome I was completed only in May 2006 (this was the last of the 24 human chromosomes -22 autosomes and X and Y- to be sequenced).

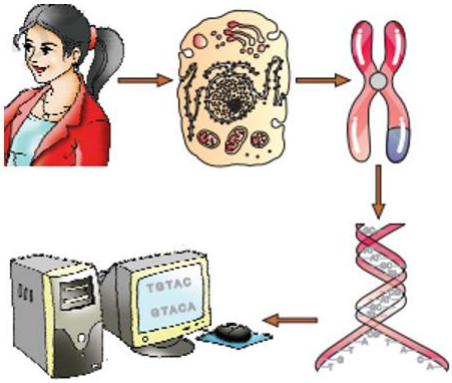


Figure : - A representative diagram of human genome project

- Another challenging task was assigning the genetic and physical maps on the genome. This was generated using information on polymorphism of restriction endonuclease recognition sites, and some repetitive DNA sequences known as microsatellites.

SPOT LIGHT

- **Frederick Sanger** is also credited for developing method for determination of amino acid sequences in proteins.

SALIENT FEATURES OF HUMAN GENOME:

Some of the salient observations drawn from human genome project are as follows:

(i) The human genome contains **3164.7 million** nucleotide bases.

(ii) The average gene consists of **3000 bases**, but sizes vary greatly, with the largest known human gene being **dystrophin at 2.4 million bases**.

(iii) The total number of genes is estimated at 30,000-much lower than previous estimates of 80,000 to 1,40,000 genes. Almost all (99.9 per cent) nucleotide bases are exactly the same in all people.

(iv) The functions are unknown for over 50 per cent of discovered genes.

(v) Less than 2 per cent of the genome codes for proteins.

(vi) Repeated sequences make up very large portion of the human genome.

(vii) Repetitive sequences are stretches of DNA sequences that are repeated many times, sometimes hundred to thousand times. They are thought to have no direct coding functions, but they shed light on chromosome structure, dynamics and evolution.

(viii) Chromosome 1 has most genes (2968). and the Y has the fewest (231).

(ix) Scientists have identified about 1.4 million locations where single-base DNA differences **(SNPs- single nucleotide polymorphism,** pronounced as 'snips') occur in humans, This information promises to revolutionise the processes of finding chromosomal locations for disease-associated sequences and tracing human history.

APPLICATIONS AND FUTURE CHALLENGES:

(1) The genome project is being compared to the discovery of antibiotics.

(2) Efforts are in progress to determine genes that will revert cancerous cells to normal.

(3) One of the greatest impacts of having the HG sequence may well be enabling radically new approach to biological research.

DNA FINGER PRINTING / DNA TYPING / DNA PROFILING/ DNA TEST

- It is technique to identify a person on the basis of his/her DNA specificity.
- This technique was invented by sir **Alec. Jeffery** (1984). In India DNA Finger printing has been started by **Dr. V.K. Kashyap** & **Dr. Lal Ji Singh.**
- DNA of human is almost the same for all individuals but very small amount that differs from person to person that forensic scientists analyse to identify people.
- DNA fingerprinting is a very quick way to compare the DNA sequences of any two individuals.
- DNA fingerprinting involves identifying differences in some specific regions in DNA sequence called as **repetitive DNA,** because in these sequences, a small stretch of DNA is repeated many times. These repetitive DNA are separated from bulk genomic DNA as different peaks during density gradient centrifugation. The bulk DNA forms a major peak and the other small peaks are referred to as satellite DNA.
- Depending on base composition (A : T rich or G:C rich), length of segment, and number of repetitive units, the satellite DNA is classified into many categories, such as micro-satellites, mini-satellites etc. These sequences normally do not code for any proteins, but they form a large portion of human genome. **These sequence show high degree of polymorphism and form the basis of DNA fingerprinting.**
- Since DNA from every tissue (such as blood, hair-follicle, skin, bone, saliva, sperm etc.), from an individual show the same degree of polymorphism, they become very useful identification tool in forensic applications.
- Further, as the polymorphisms are inheritable from parents to children, DNA fingerprinting is the basis of paternity testing, in case of disputes Polymorphism in DNA sequence is the basis of genetic mapping of human .
- **Polymorphism** (variation at genetic level) arises due to mutations.
- New mutations may arise in an individual either in somatic cells or in the germ cells (cells that generate gametes in sexually reproducing organisms). If a germ cell mutation does not seriously impair individual's ability to have offspring who can transmit the mutation, it can spread to the other members of population (through sexual reproduction).
- Allelic sequence variation has traditionally been described as a DNA polymorphism if more than one variant (allele) at a locus occurs in human population with a **frequency greater than 0.01.** In simple terms, if an **inheritable mutation** is observed in a population at high frequency, it is referred to as **DNA polymorphism.**
- The probability of such variation to be observed in noncoding DNA sequence would be higher as mutations in these sequences may not have any immediate effect/impact in an individual's reproductive ability. These mutations keep on accumulating generation after generation, and form one of the basis of variability/polymorphism.
- Polymorphisms are the key of DNA typing. Polymorphisms are most useful to forensic scientist. It is consist of variation in the length of DNA at specific loci is called Restricted fragment. It is most important segment for DNA test made up of short repetitive nucleotide sequences, these are called **VNTRs (Variable Number of Tandem Repeat).**
- VNTR's also called minisatellites were discovered by Alec Jeffery. Restricted fragments consist of hypervariable repeat region of DNA having a basic repeat sequence of 11-60 bp and flanked on both sites by restriction site.
- The number and position of minisatellites or VNTR in restriction fragment is different for each DNA and length of restricted fragment depends on number of VNTR.
- Therefore, when the genome of two people are cut using the same restriction enzyme the length of fragments obtained is different for both the people.
- These variations in length of restricted fragment is called RFLP or Restriction fragment length polymorphism.
- Restriction Fragment Length Polymorphism distributed throughout human genomes are useful for DNA Finger printing.

- DNA Fingerprint can be prepared from extremely minute amount of blood, semen, hair bulb or any other cell of the body.
  DNA content of 1 - Microgram is sufficient.

TECHNIQUE OF DNA FINGER PRINTING INVOLVES THE FOLLOWING MAJOR STPES.

- It includes:-

(1) isolation of DNA,
(2) digestion of DNA by restriction endonucleases,
(3) separation of DNA fragments by electrophoresis,
(4) transferring (blotting) of separated DNA fragments to synthetic membranes, such as nitrocellulose or nylon,
(5) hybridisation using labelled VNTR probe, and
(6) detection of hybridised DNA fragments by autoradiography

1. Extraction :

- DNA is extracted from the cell by cell lysis. If the content of DNA is limited then DNA can be amplified by Polymerase chain reaction (PCR). This process is known as amplification.

2. Restriction Enzyme Digestion :

- Restriction enzyme cuts DNA at specific 4 or 6 base pair sequences called restriction site.
  Hae III (*Haemophilus aegyptius*) is most commonly used enzyme. It cuts the DNA, every where the bases are arranged in the sequence GGCC. These restricted fragments are transferred to Agarose Polymer gel.

3. Gel Electrophoresis :

- Gel electrophoresis is a method that separates macromolecules-either nucleic acid or proteins-on the basis of size, electric charge.
- Gel electrophoresis refers to the technique in which molecules are forced across a span of gel, motivated by an electrical current.
- By the gel electrophoresis these restricted fragments move towards the positive electrode (anode) because DNA has –ve electric charge (PO_4^{-3}).
- Smaller Fragment more move towards the positive pole due to less molecular weight.
- So after the gel electrophoresis DNA fragment arranged according to molecular weight.
- These separated fragments can be visualized by staining them with a dye that fluoresces ultraviolet radiation.

4. Southern transfer / Southern blotting :

- The gel is fragile. It is necessary to remove the DNA from the gel and permanently attaches it to a solid support. This is accomplished by the process of Southern blotting.
- The DNA is transferred by the process of blotting to a sheet of nylon or nitrocellulose. The nylon acts like an ink blotter and "blots" up the separated DNA fragments, the restriction fragments, invisible at this stage are irreversibly attached to the nylon membrane the "blot".
  This process is called Southern blot by the name of **Edward Southern (1970).**

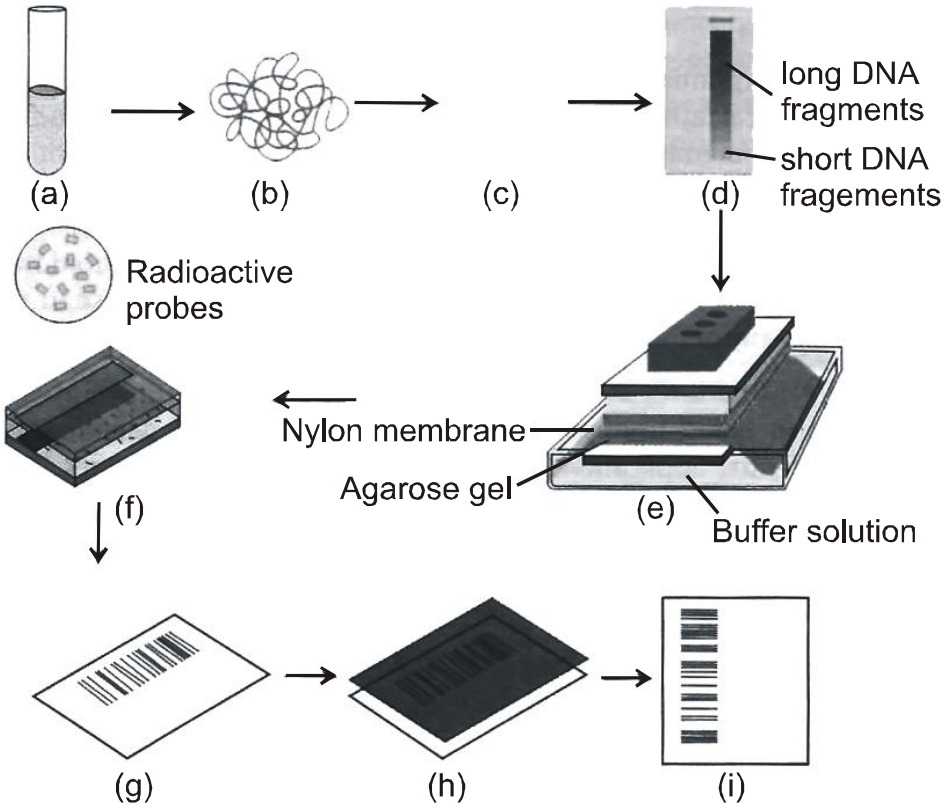
5. Hybridization :

- To detect VNTR locus on restricted fragment, we use single stranded **Radioactive (P^{32}) DNA** probe which have the base pair sequences complimentary to the DNA sequences at the VNTR locus.
- Labelled Probes are attached with the VNTR loci of restricted DNA Fragments, this process is called Hybridization.

6. Autoradiography :

- Nylon membrane containing radioactive probe is exposed to X-ray. Specific bands appear on X-ray film. These bands are the areas where the radioactive probe bind with the VNTR.
- This appears the specific restricted fragment length pattern. This length pattern is different in different individual. This is called **Restricted Fragment length Polymorphism (RFLP)**.

- These allow analyzer to identify a particular person DNA, the occurance and frequency of a particular genetic pattern contained in this X-ray film. These x-ray film called DNA signature of a person which is specific for each individual.
- The probability of two unrelated individual having same pattern of location and repeat number of minisatellite (VNTR) is one in ten billion (world population 6.1 billion)



DNA fingerprinting (a) Blood sample, (b) Extraction of DNA from blood sample (c) Cutting of DNA by restriction enzyme, (d) Separation of DNA fragments by electrophoresis, (e) Transmission of DNA bands on nylon membrane by southern Blotting, (f) Binding of radioactive probes to specific DNA pattern on membrane, (g) Washing of excess of DNA, (h) X-ray to detect radioactive pattern and (i) Autoradiograph.

APPLICATION OF DNA FINGER PRINTING :

1. **Paternity tests-** The major application of DNA finger printing is in determining family relationships.
2. **Identification of the criminal-** The DNA fingerprint of the person matching the one obtained from sample collected from scene of crime can give a clue to the actual criminal.

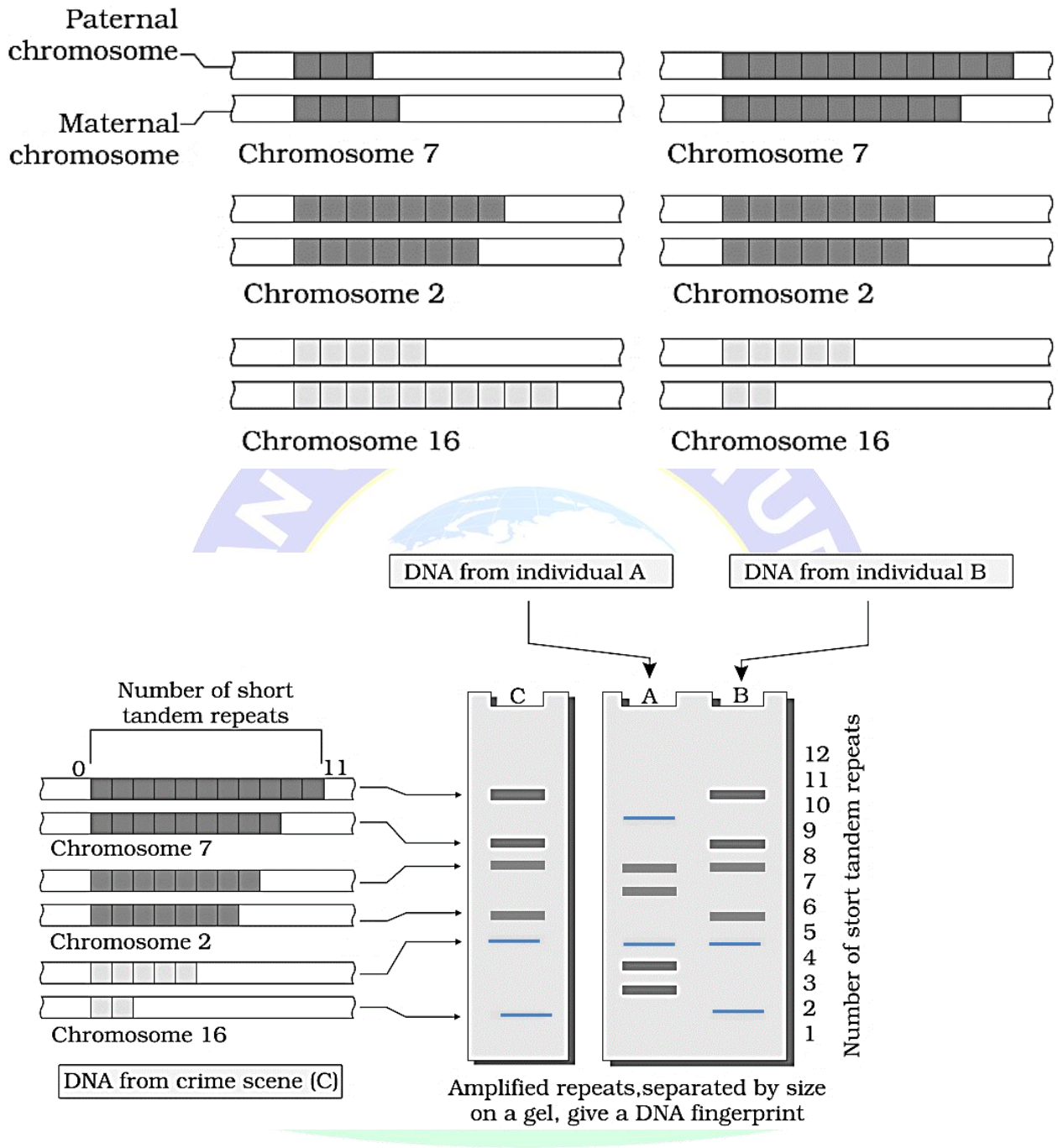


Figure: - Schematic representation of DNA finger printing: few representative chromosomes have been shown to contain different copy number of VNTR. The two alleles (Paternal & maternal) of a chromosome also contain different copy number of VNTR. It is clear that the bending pattern of DNA from crime scene matches with B and not with A,