# MOLECULAR BASIS OF INHERITANCE

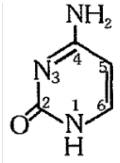
## NUCLEIC ACIDS

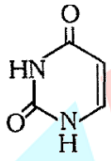
- F. Meischer discovered nucleic acid in nucleus of pus cell and called it "nuclein". The term nucleic acid was coined by "Altman."
- Nucleic acids are polymer of nucleotides.
  Example : DNA and RNA.
  Nucleotide = Nitrogen base + pentose sugar + phosphate
  Nucleoside = Nitrogen base + pentose sugar.

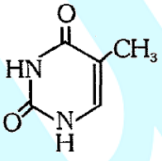
**A. Nitrogen base :**

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

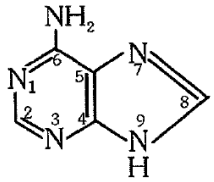
1. **Pyrimidines** - Consist of one pyrimidine ring. Skeleton of ring composed of two nitrogen and four Carbon atoms. e.g. Cytosine, Thymine and Uracil

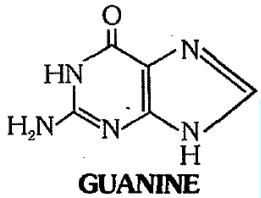

**CYTOSINE**


**URACIL**


**THYMINE**

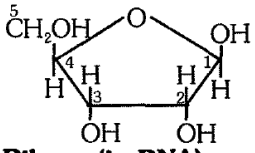
2. Purines - Consist of two rings i.e. one pyrimidine ring (2N + 4C) and one imidazole ring (2N + 3C) e.g. Adenine and Guanine.


**ADENINE**

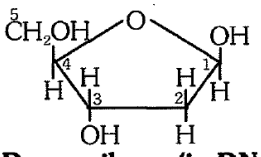

**GUANINE**

**B. Pentose Sugar (Number of Carbon = 5) :-**

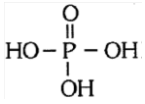
(1)


**Ribose (in RNA)**

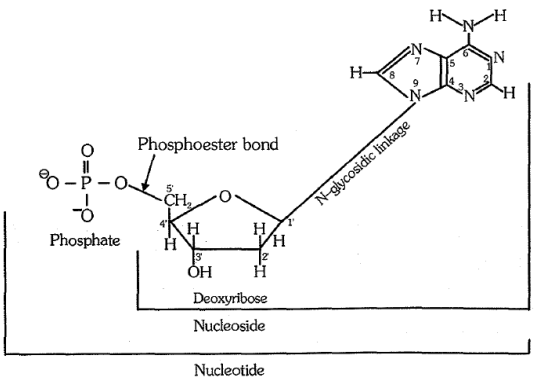
(2)


**Deoxy ribose (in DNA)**

C. Phosphate :
- Acidic
- Negative charged



Nitrogen base forms bond with first carbon of pentose-sugar 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.

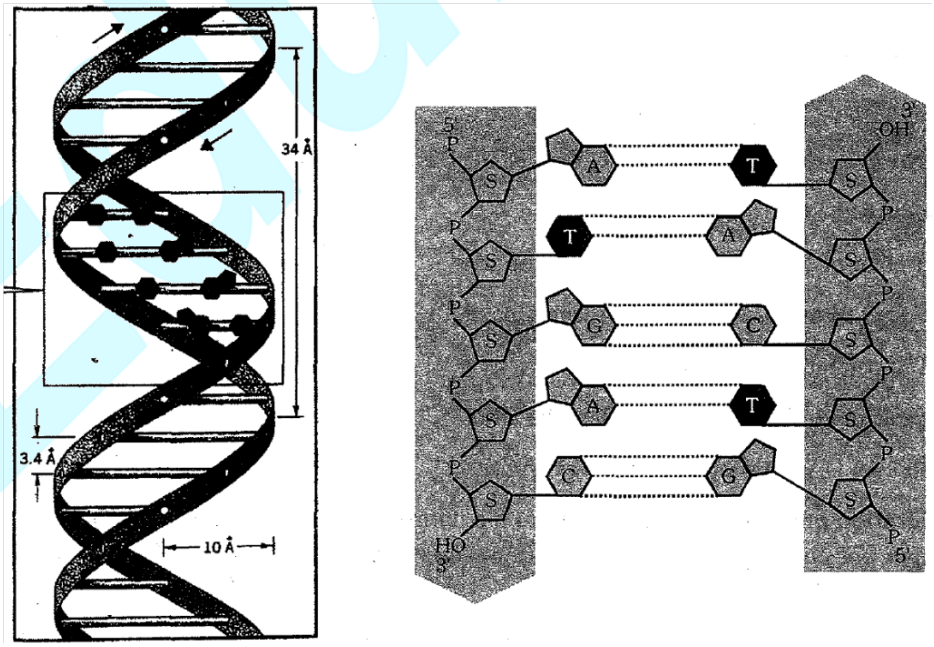


## Types of Nucleosides and Nucleotides

1. Adenine + Ribose = Adenosine
   Adenosine + Phosphate = Adenylic acid (AMP)
2. Adenine + Deoxyribose = Deoxy adenosine
   Deoxy adenosine+ P = Deoxy adenylic acid (dAMP)
3. Guanine + Ribose = Guanosine
   Guanosine + P = Guanylic acid (GMP)
4. Guanine + Deoxyribose = Deoxy guanosine
   Deoxy guanosine+ P = Deoxy guanylic acid (dGMP)
5. Cytosine + Ribose = Cytidine
   Cytidine+ P =_Cytidylic acid (CMP)
6. Cytosine +Deoxyribose = Deoxycytidine
   Deoxycytidine + P = Deoxycytidylic acid (dCMP)
7. Uracil + Ribose = Uridine
   Uridine + P = Uridylic acid (UMP)
8. Thymine + Deoxyribose = Deoxy thymidine
   Deoxythyrnidine + P = Deoxythymidylicacid (dTMP)

## DNA

- DNA as an acidic substance present in nucleus was first identified by Friedrich Meischer in 1869 .
- DNA term was given by - Zacharis
- DNA is long polymer of deoxyribonucleotides .
- DNA is negatively charged .
- In DNA pentose sugar is deoxyribose sugar and four types of nitrogen bases A,T,G,C
- Wilkins and Franklin studied DNA molecule with the help of X-Ray crystallography.



- With the help of this study, Watson and Crick (1953) proposed a double helix mole! 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.
- 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 are held together by hydrogen bonds. These hydrogen bonds are present between nitrogen bases of both strand.
- Adenine binds to thymine by two hydrogen bonds and cytosine binds to guanine by three hydrogen bonds.
- In a DNA molecule one purine always pairs with a pyrimidine. 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
- Chargaff's equivalency rule - In a double stranded DNA amount of purine nucleotides is equals to amount of pyrimidine nucleotides.
  Purine = Pyrimidine
  [A] + [G] = [T] + [C]
  $\frac{[A]+[G]}{[T]+[C]} = 1$
- Base ratio = $\frac{A+T}{G+C}$ = constant for a given species. i.e. species specific.
- 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
- Melting point of DNA depends on G - C contents.
  More G - C contents means higher melting point.
  $T_m$ = Temperature of melting.
  $T_m$ of prokaryotic DNA > $T_m$ of Eukaryotic DNA
- DNA absorbs U.V. rays of 2600A wavelength.

- **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 attached and recoiled to each other. It is called renaturation of DNA

**Configuration of DNA Molecule :-**

- 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.

Two chains have anti-parallel polarity. It means, if one chain has the polarity 5'→3', the other has 3'→5.

- Distance between two successive steps is 3.4 $A^0$. In one complete turn of DNA molecule there are such 10 steps (10 pairs of nitrogen bases). So the length of one complete turn is 34 $A^0$. This is called helix length.
- Diameter of DNA molecule i.e. distance between phosphates of two strands is 20$A^0$.
- Each step of ascent is represented by a pair of bases. At each step of ascent, the strands turns 36°.
- In nucleus of eukaryotes the DNA is associated with histone protein to form nucleoprotein.
- Bond between bNA and Histone is salt linkage ($Mg^{+2}$).
- 

| | |
|---|---|
| $\phi \times 174$ (bacteriophage) [Single stranded] | 5386 Nucleotides |
| $\lambda$ bacteriophage | 48502 base pair |
| E.coli | $4.6 \times 10^6$ base pair |
| Human | $6.6 \times 10^9$ base pair |

**Type of DNA**

On the basis of direction of twisting, there are two types of DNA.

**1. Right Handed DNA –**

| DNA | Helix Length | No. of base pairs | Distance between two pairs | Diameter |
|---|---|---|---|---|
| 'A' | 28 $A^0$ | 11 pairs | 2.56 $A^0$ | 23 $A^0$ |
| 'B' | 34 $A^0$ | 10 pairs | 3.4 $A^0$ | 20 $A^0$ |
| 'C' | 31 $A^0$ | 9.33 pairs | 3.32 $A^0$ | 19 $A^0$ |
| 'D' | 24.24 $A^0$ | 8 pairs | 3.03 $A^0$ | 19 $A^0$ |

2. Left handed DNA :-

Anticlockwise twisting e.g. Z-DNA- discovered by Rich. Phosphate and sugar backbone is zig-zag.

Helix length - 45.6 $A^0$
Diameter - 18.4$^0$
No. of base pairs - 12 (6 dimmers)
Distance between 2 base - pairs - 3.75$A^0$

- Palindromic DNA-Wilson and Thomas

$\longrightarrow$

CC GG TA CC GG
GG CC AT GG CC

$\longleftarrow$

Sequence of nucleotides same from both ends.

## PACKAGING OF DNA HELIX

The average distance between the two adjacent base pairs of 0.34 nm ($0.34 \times 10^{-9}$ m or 3.4 Å). Length of DNA for a human diploid cell is 6.6 x $10^9$ bp × $0.34 \times 10^{-9}$ m/bp = 2.2 m. The length is far greater than the dimension of a typical nucleus (approximately $10^{-6}$ m).

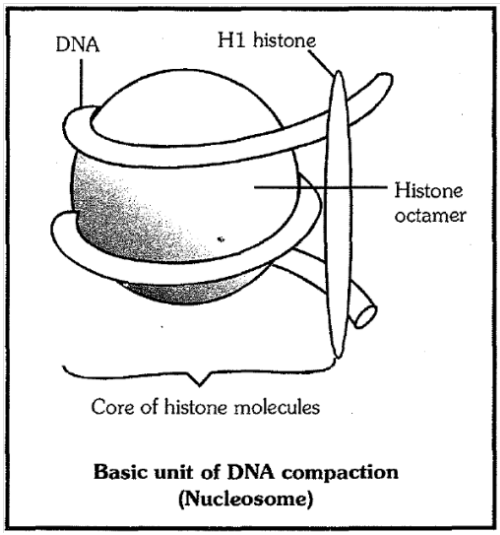
The number of base pairs in Escherichia coli is $4.6 \times 10^6$. The total length is 1.36 mm. The long sized DNA accommodated in small area (about 11.1m in E. coli) only through packing or compaction. DNA is acidic due to presence of large number of phosphate group. Compaction occurs by folding acid attachment of DNA with basic proteins, polyamine in prokaryotes and histone in eukaryotes.

**DNA packaging in Prokaryotes :** DNA is found in cytoplasm in supercoiled state. The coils are maintained by non histone basic protein like polyamines. This compact structure of DNA is called nucleoid or genophore.

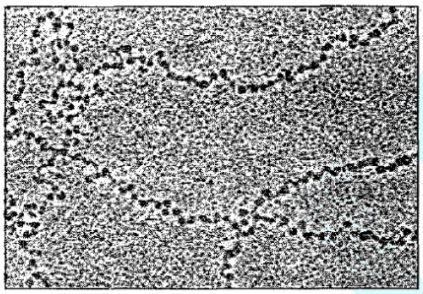
**DNA packaging in Eukaryotes :** It is carried out with the help of lysine and ariginine rich basic proteins called histone. The unit of compaction is nucleosome. There are five types of histone proteins $H_1$ $H_2A$, $H_2B$. $H_3$ and $H_4$. Four of them occur in pairs to produce histone octamer (2 copies of each- $H_2A$, $H_2B$, $H_3$ and $H_4$, called nubody or core of nucleosome. Their positively charged ends are directed outside. The negatively charged DNA is wrapped around the positively charged histone octamer to form a structure called nucleosome. A typical nucleosome contains 200 bp of DNA Helix. DNA present between two adjacent nucleosome is called linker DNA. It is attached to $H_1$ histone protein. Length of linker DNA varies from species to species. Nucleosome chain gives a beads on string appearance under electron microscope. The nucleosomes furthers coils to form solenoid. It has diameter of 30 nm as found in chromatin. The beads on string structure in chromatin is packaged to form chromatin fibres that are further coiled and condensed at metaphase stage of cell division to form chromosomes. The packaging at higher level requires additional set of proteins (acidic) that collectively are referred to as non-histone chromosomal (NHC) proteins.

Non-Histone chromosomal proteins are of three types :

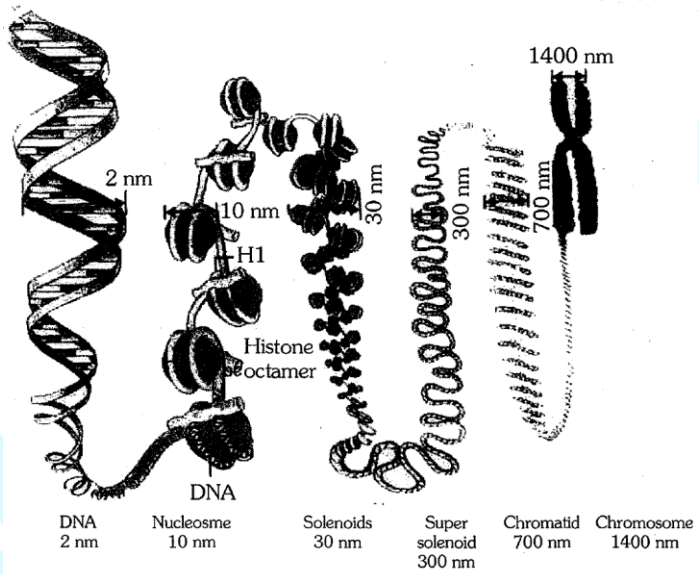
(i) Structural NHC protein
(ii) Functional NHC protein e.g., DNA polymerase, RNA polymerase
(iii) Regulatory NHC protein




**Basic unit of DNA compaction (Nucleosome)**



**EM picture - "Beads-on-String**




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

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

## THE SEARCH FOR GENETIC MATERIAL

The experiments given below prove that DNA is the genetic material.

**(I) Evidence from bacterial transformation :** The transformation experiments conducted by Frederick Griffith in 1928, are of greater importance in establishing the nature of genetic material. He used two strains of bacterium Diplococcus or Streptococcus pneumoniae or Pneumococcus i.e., S-III and R-ll.

(i) Smooth (S) or capsulated type which have a mucous coat and produce shiny colonies. These bacteria are virulent and cause pneumonia.

(ii) Rough (R) or non-capsulated type in which mucous coat is absent and produce rough colonies. These bacteria are nonvirulent and do no.t cause pneumonia.

The experiment can be described in following four steps :

(a) Smooth type bacteria were injected into mice. The mice died as a result of pneumonia caused by bacteria.

Live S strain $\rightarrow$ injected into mice $\rightarrow$ Mice died

(b) Rough type bacteria were injected into mice. The mice lived and pneumonia was not produced

Live R strain $\rightarrow$ injected into mice $\rightarrow$ mice lived

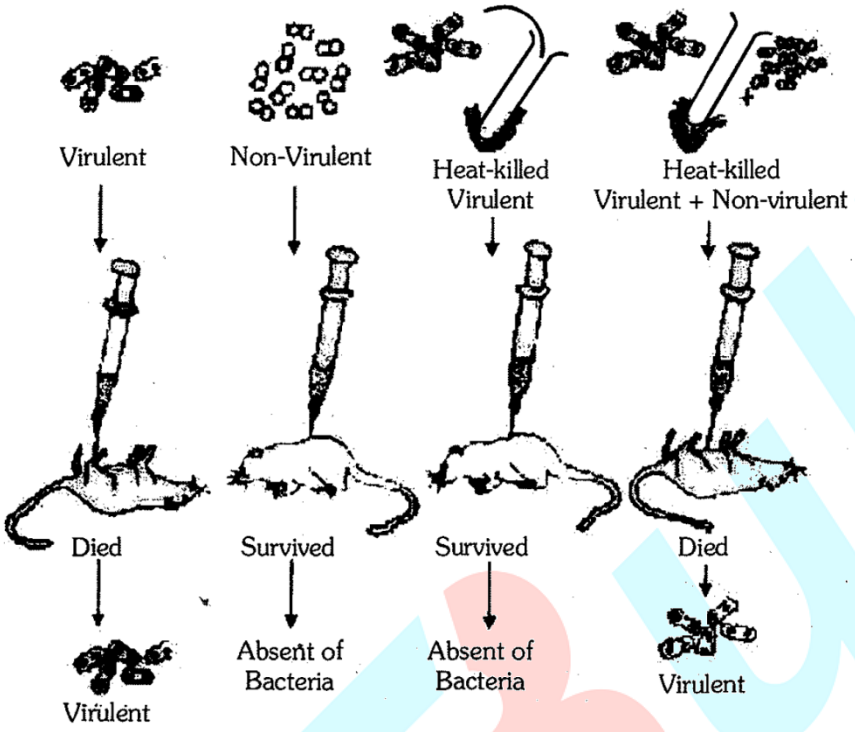
(c) Smooth type bacteria which normally cause disease were heat killed and then injected into the mice. The mice lived and pneumonia was not caused.

S strain (heat killed) $\rightarrow$ Injected into mice $\rightarrow$ Mice lived

(d) Rough type bacteria (living) and smooth type heat-killed bacteria (both known not to cause disease) were injected together into mice. The mice died due to pneumonia and virulent smooth type living bacteria could also be recovered from their bodies.

S strain (heat killed) + R strain (living) $\rightarrow$ injected into mice $\rightarrow$ Mice died

He concluded from fourth step of the experiment that some rough bacteria (nonvirulent) were transformed into smooth type of bacteria (virulent). This occurred perhaps due to absorption of some transforming substance by rough type bacteria from heat killed smooth type bacteria. This transforming substance from smooth type bacteria caused the synthesis of capsule which resulted in production of pneumonia and death of mice. Therefore, transforming principle appears to control genetic characters (for example, capsule as in this case). However, the biochemical nature of genetic material was not defined from this experiments.

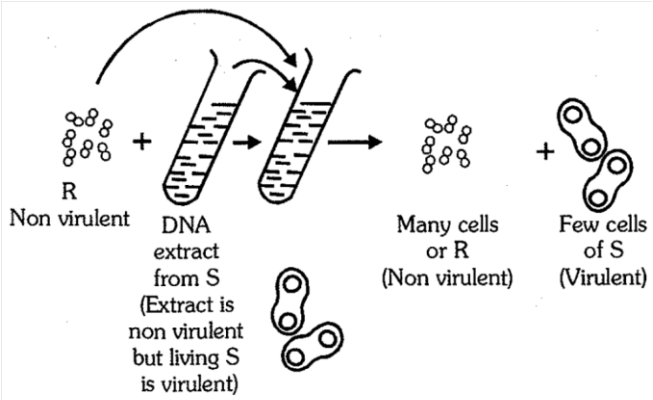



**Bacterial transfomation experiments conducted by Griffith**

**Biochemical characterisation of Transforming Principle :**

Later, Avery, Macleod and McCarty (1944) repeated the experiment in vitro to identify the biochemical nature of transforming substance. They proved that this substance is DNA Pneumococcus bacteria cause disease when capsule is present. Capsule production is under genetic control.

In the experiments, rough type bacteria (non-capsulated and non-virulent) were grown in a culture medium to which DNA extract from smooth type bacteria (capsulated and virulent) was added. Later, the culture showed the presence of smooth type bacteria also in addition to rough type. This is possible only if DNA of smooth type was absorbed by rough type bacteria which developed capsule and became virulent. This process of transfer of characters of one bacterium to another by taking up DNA from solution is caleld transformation. When DNA extract was treated with DNase (an enzyme which destroys DNA), transformation did not occur. The transformation occurs when proteases and RNases were used. This clearly shows that DNA is the genetic material.



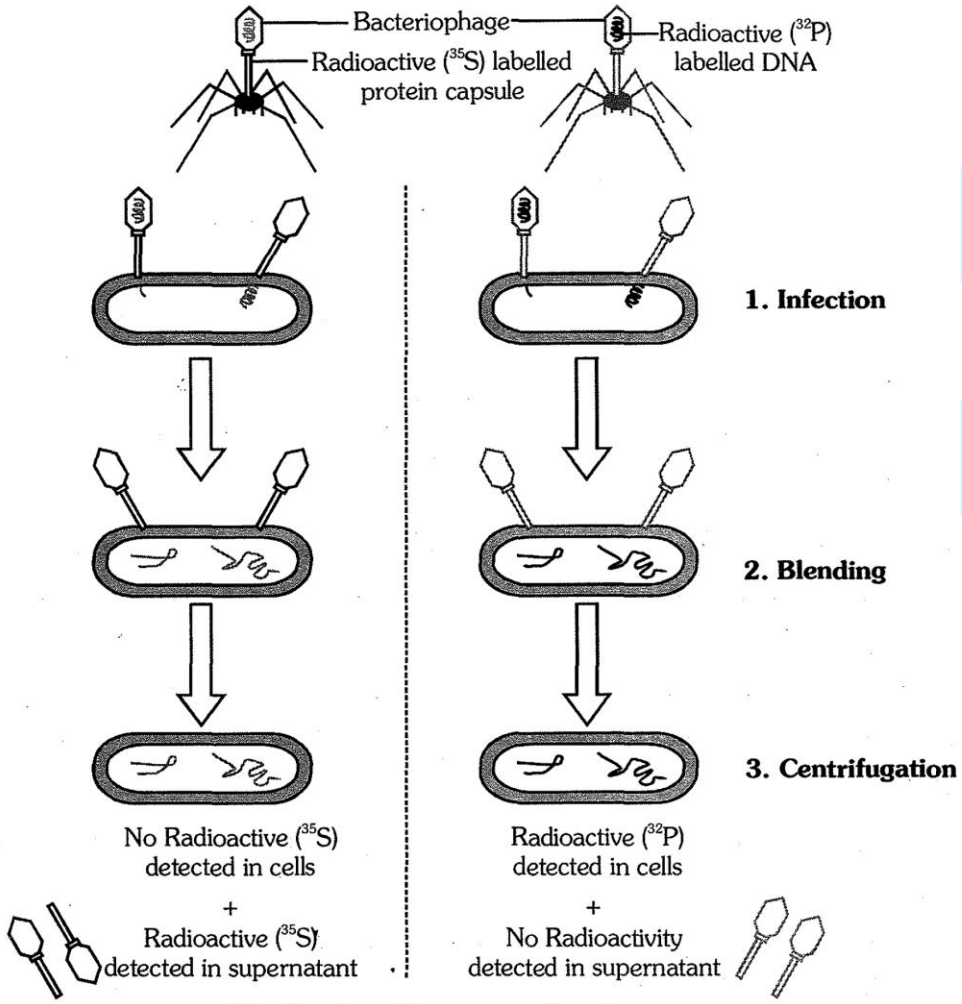

**In vitro experiment of Avery and others demonstrating that DNA is genetic material**

**(II) Evidence from experiments with bacteriophage :** $T_2$ bacteriophage is a virus that infects bacterium Escherichia coli and multiplies inside it. $T_2$ phage is made up of DNA and protein coat. Thus, it is the most suitable material to determine whether DNA or protein contains information for the production of new virus (phage) particles. Hershey and Chase (1952) demonstrated that only DNA of the phage enters the bacterial cell and therefore, contains necessary genetic information for the asssembly of new phage particle.

The functions of DNA and proteins could be found out by labelling them with radioactive tracers. DNA contains phosphorus but not sulphur. Therefore, phage DNA was labelled with $P^{32}$ by growing bacteria infected with phages in culture medium containing $^{32}PO_4$. Similarly, protein of phage contains sulphur but no phosphorus. Thus, the phage protein coat was labelled with $S^{35}$ by growing bacteria infected with phages in another culture medium containing $^{35}SO_4$. After the formation of labelled phages. Three steps were followed, i.e., infection, blending, centrifugation.

1. **Infection:** Both type of labelled phages were allowed to infect normally cultured bacteria in separate experiments .
2. **Blending :** These bacteria cells were agitated in a blender to break the contact between virus and bacteria.
3. **Centrifugation :** The virus particles were separated from the bacteriam 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 on the outside of bacteria cells (i.e., in the medium). Labelled DNA was also found in the next generation or phage. This clearly showed that only DNA enters the bacterial host and not the protein. DNA therefore, is the infective part of virus and also carries all the genetic information. This provided the unequivocal proof that DNA is the genetic material.




**The Hershey-Chase experiment**

Following are the properties and functions which should be fulfilled by a substance if it is to qualify as genetic material.

(1) The genetic material should be able to generate its own kind (replication). 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 fulfill first criteria itself.

(2) It should chemically and structurally be stable. 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

(3) The genetic material should also be capable of undergoing mutations- 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.

(4) The genetic material should be able to transmit faithfully to the next generation, as Mendelian characters. 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 protem synthesising machinery has evolved around RNA

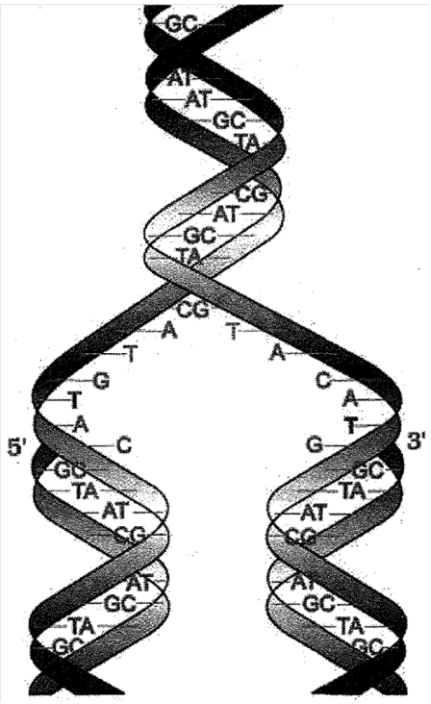
The above discussion indicate that both RNA and DNA can function as genetic materiaL but DNA being more stable is preferred for storage of genetic information. For the transmission of genetic information, RNA is better.

## RNA WORLD

RNA was the first genetic material. There are evidences 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 (e.g., splicing)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. RNA is adapter, structural molecule and in some cases catalytic. Thus RNA is better material for transmission of information.

## DNA REPLICATION

- While proposing the double helical structure for DNA, Watson and Crick had immediately proposed a scheme for replication of DNA To quote their original statement that is as follows:
- "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material" (Watson and Crick, 1953).
- The scheme suggested that the two strands would separate and act as a template for the synthesis of new complementary strands. After the completion of replication, each DNA molecule would have one parental and one newly synjhesised strand. This scheme was termed as semiconservative DNA replication
- D.N.A. capable of self duplication.
- All living beings have the capacity to reproduce because of this characteristic of D.N.A.



- D.N.A replication takes place in "S- Phase" of the cell cycle. At the time of cell division, it divides in equal parts in the daughter cells.

## SEMI CONSERVATIVE MODE OF DNA REPLICATION

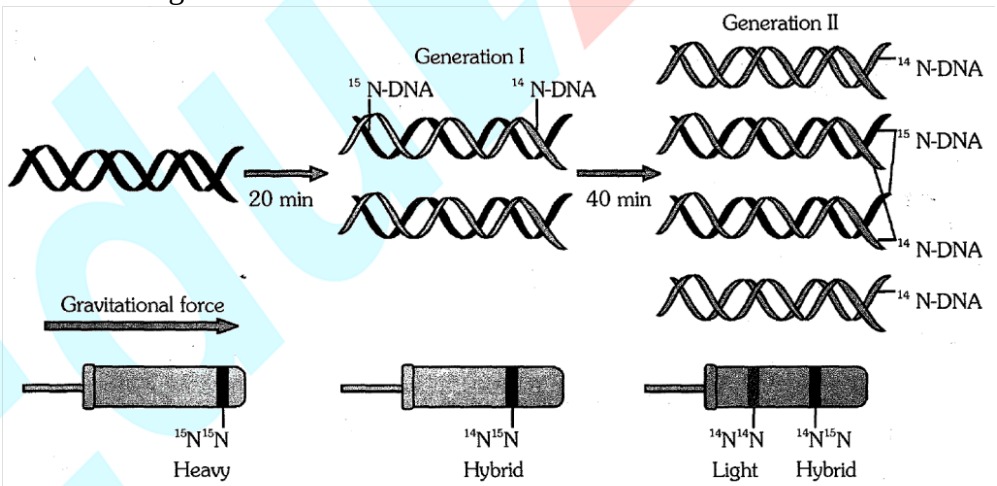
Semi conservative mode of D.N.A. replication was first proposed by Watson & Crick. Later on it was experimentally proved by Meselson & Stahl (1958) in E- Coli and Taylor in Viciafaba (1958). To prove this method, Taylor used Radiotracer Technique in which Radioisotopes (tritiated thymidine = $_1H^3$ were used. Meselson and Stahl used heavy isotope of nitrogen ($N^{15}$).

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 synthesized 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 $^{14}N$ is not a radioactive isotope, and it can be separated from MN only based on densities).

(ii) Then they transferred the cells into a medium with normal HNH1Cl 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.




**Separation of DNA by Centrifugation**

## MECHANISM OF DNA REPLICATION

The following steps are included in DNA replication :-

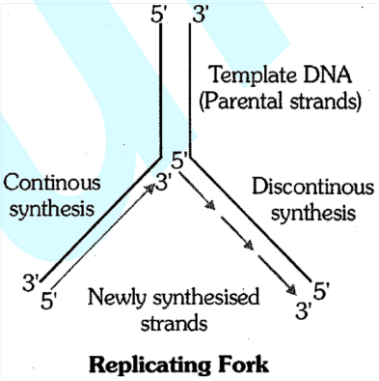
(1) Unzipping (Unwinding) :-

- 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 occur 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 Heficase (= Swivelase). In the process of unzipping $Mg^{+2}$ act as cofactor.
- SSB (single stranded DNA binding protein) prevents the reformation of H-bonds .
- To poisomerase (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 Jew hours in Eukaryotes.

**(2) Formation of New Chain :-**

- 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 Nuceloplasm. In the nucleoplasm, Nucleotides are present in the form of triphosphates like dA1P, dG1P, dcrP, dTIP etc.



**Replicating Fork**

- During replication, the 2 phosphate groups of all nucleotides are separated. In this process energy is yeilded which is consumed in DNA replication.
- Energetically replication is a very expensive process. Daoxyribonucleoside triphosphase serve dual purposes in addition to acting as substrates they provide energy for polymerisation.
- The formation of new chain always takes place in 5'- 3' direction. As a result of this, one chain of D.N.A. 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 Polymerase. In prokaryotes it is of 3 types:

**(l)** **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 gap.lt is also known as DNA-repair enzyme.

**(2)** **DNA - Polymerase II** :- It is least reactive in replication process. It is also helpful in DNA-repairing in absence of DNA-polymerase-1 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.

♦ 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.

**Special Point :**

- All DNA polymerase I, II and III enzymes have 5'-3" polymerisation activity and 3'-5" exonuclease activity.
- DNA polymerase I also has 5'-3" exonuclease activity.
- Any failure in cell division after DNA replication result into polyploidy.
- Difference between DNAs and DNase is that DNAs menas many DNA and DNase means DNA digestive enzymes.

## BEGINNER'S BOX-1

1. nu body of nucleosome consists of
   (1) $H_1$ and $H_2A$ (2) $H_2A$ and $H_2B$
   (3) $H_3$ and $H_4$ (4) Both (2) and (3)

2. Radioactive element used to label DNA of bacteriophage in Wareing-blender experiment of Hershey and Chase was:-
   (1) $S^{35}$ (2) $P^{32}$ (3) $N^{15}$ (4) $C^{14}$

3. Bonding between deoxyribose and base in pyrimidine nucleoside molecule is:-
   (1) 1'-1' glycosidic linkage (2) 1'-6' glycosidic linkage
   (3) 1 '-9' glycosidic linkage (4) 1'-4' glycosidic linkage

4. $T_m$ (melting temperature) value of DNA is high when it contains
   (1) A + T > G + C (2) G + C > A + T
   (3) A + T = G + C (4) A + G = T + C

5. Select an incorrect statement regarding RNA molecule :
   (1) It has highly reactive 2'-OH group
   (2) It shows higher rate of mutation than DNA
   (3) It is genetic material in some viruses
   (4) It follows Chargaff rule

6. In Meselson and Stahl's experiment, heavy isotope 15N was used in the form of
   (1) $^{15}NaNO_3$ (2) $^{15}NH_4Cl$ (3) $^{15}KNO_3$ (4) $^{15}NH_4NO_3$

7. Assuming that 50 heavy (i.e. containing $N^{15}$) DNA molecules replicated twice in a medium containing $N^{14}$, we expect
   (1) 100 half heavy and half light and 150 light DNA molecules
   (2) 100 half heavy and half light and 100 light DNA molecules
   (3) 50 heavy and 150 light DNA molecules
   (4) 50 heavy and 100 light DNA molecules

8. The enzyme which shows polymerising activity in 5'→3' direction is :
(1) DNA polymerase III (2) DNA polymerase II
(3) DNA polymerase I (4) All of these

9. DNA polymerase I is involved in:
(1) Removal of RNA primer (2) Filling of gap
(3) Joining of okazaki fragments (4) Both (1) and (2)

10. DNA replication in lagging strand of most of the eukaryotic organisms is:-
(l) Conservative and continous (2) Semi-conservative but discontinous
(3) Consenrvative and semi-discontinous (4) Semi-conservative but continous

## RIBO NUCLEIC ACID (RNA)

Structure of RNA is fundamentally same as DNA, but there are some differences. The differences are as follows:-

(1) In place of De-oxyribose sugar in DNA, there is present Ribose sugar in RNA.
(2) In place of nitrogen base Thymine in DNA, there is present uracil in RNA.
(3) RNA is made up of only one polynucleotide chain i.e. R.N.A. is single stranded.

**Exception :-**

RNA found in Reo - virus is double stranded, i.e. it has two polynucleotide chains.

**Types of RNA :**

1. **Genetic RNA or Genomic RNA-** In some viruses RNA works as genetic material and it transfers informations from., one generation to next generation.
eg. Reo virus, TMV, QB, bacteriophage.
2. **Non-genetic RNA - mainly of 3 types -**
(1) r- RNA (2) t- RNA (3) m- RNA

♦ RNA functions as adapter, structural and in same 0 cases as a catalyst (Ribozyme)

**(1) Ribosomal RNA (r - RNA) :-**

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

**Function:-**

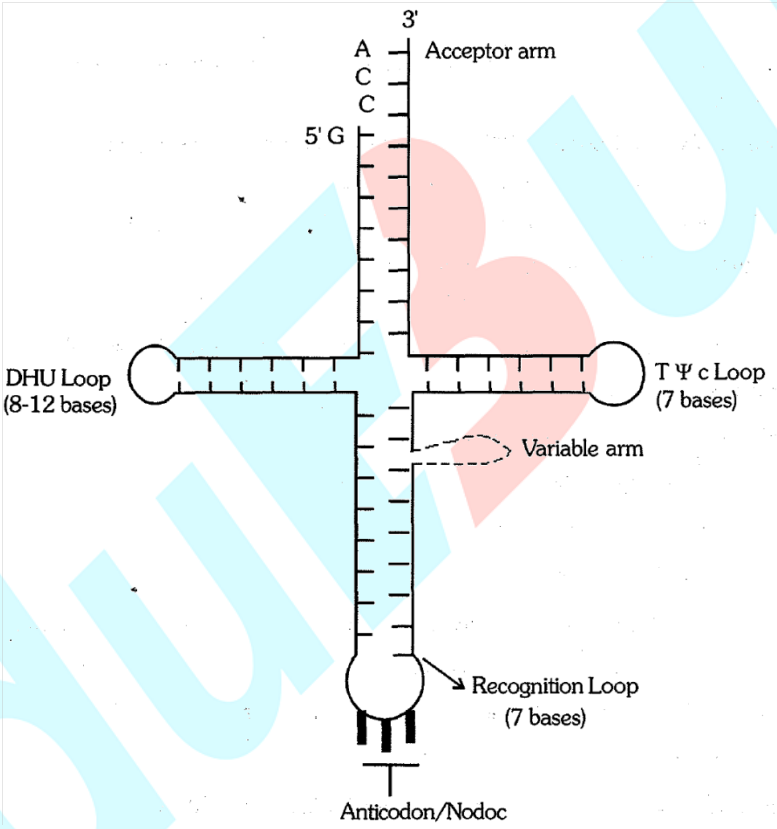
- At the time of protein synthesis, r-RNA provides attachement site to t-RNA and m-RNA and attaches them on the ribosome.
- It attaches t-RNA to the larger subunit of ribosome and m-RNA to smaller subunit of ribosome.

**(2) Transfer - RNA (t-RNA) :-**

- It is 10-15% of total RNA
- It is synthesized in the nucleus by DNA ·
- It is also known as soluble RNA (sRNA)
- It is also known as 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 oft- RNA is most complicated.




Holley presented Clover leaf model of its structure. In two dimensional structure the t-RNA appears clover leaf like but in three dimensional structure (by Kim) it appears inverted L-shaped.

The structure of tRNA looks like a clover leaf but in actual structure, the tRNA is a compact molecule which looks like inverted 'L'.

- There are present three nucleotides in a particular sequence at 3' end of t - RNA and that sequence is CCA.
- All the 5' ends i.e. last ends are having G (guanine).
- 3' end is known as Acceptor end.
- t-RNA accepts amino acids at acceptor points. Amino acid binds to 3' end by its - COOH group.

- The molecule of t - RNA 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:-**

There are some abnormal nitrogenous bases in the loops, that is why hydrogen bonds are not formed.

e.g. (i) Inosine (I) (ii) Pseudouracil (Ψ) (iii) Dihydrouridine (DHU)

**(A) T Ψ C Loop or Attachment loop :-**

This loop connects t- RNA to the larger subunit of ribosome.

**(B) Recognition Loop (Anticodon loon) :-**

- This is the most specific loop of t-RNA and different types of t-RNA are different due to this loop. There " is a specific sequence of three rucleotides called Anticodon, is present at the end of this loop.
- t-RNA recognizes its place on m- RNA with the help of Anticodon.
- The anticodon of t-RNA recognises its complimentary seguence on rn-RNA. This complimentary seguence is known as codon.

**(C) 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 aminoacid to the 3' end oft-RNA.
- There are 20 types of enzymes for 20 types of aminoacids.
- The function "of DHU loop is to recognize this specific Aminoacyl synthetase enzyme.

**(3) Messenger RNA (m -RNA) :-**

- The m - RNA is 1 - 5% of the cell's total RNA.
  The name m-RNA was given by Jacob and Monad.
- The m - RNA is produced by genetic DNA in the nucleus.
- It is least stable RNA.

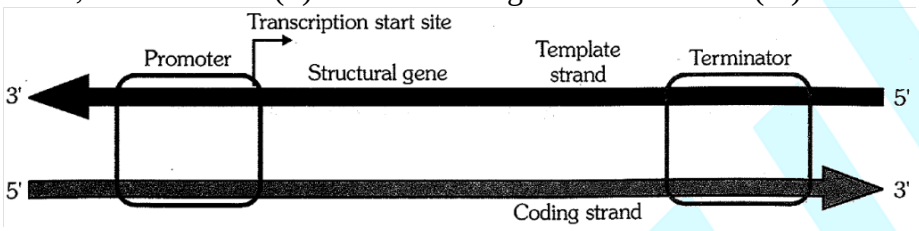
## TRANSCRIPTION

- 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" or "Template strand".
- 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 arninoacid in these protein would be different 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.
- 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).
- The segment of DNA which contains signal for the s~thesis of one polypeptide is known as "Cistron " .
- 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 hn-RNA synthesis (Precursor of m-RNA)
- RNA polymerase-III for t-RNA 5s rRNA SnRNA synthesis.

♦ Prokaryotes have only one type of RNA polymerase which synthesizes all types of RNAs .
♦ RNA polymerase (Core enzyme) of E. Coli has five polypeptide chains 13, 13: a, a and ro .
♦ σ polypeptide chain is also known as σ factor (sigma factor).
♦ Core enzyne + Sigma factor ⇒ RNA Polymerase
(β, β; α, α, ω) (σ)

A transcription unit in DNA is defined primarily by three regions in the DNA :-
(i) A promoter, (ii) The structural gene (iii) A terminator




Following steps are present in transcription -

**(1) INITIATION :-**

♦ DNA has a "Promoter site" where RNA polymerase binds 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 $N_2$ base from. "Starting point" a sequence of 6 base pairs (TATMT) is present on DNA, which is called "Pribnow box".
♦ In eukaryotes before the 20 N2 base from "Starting point" a sequence of 7 base pairs (TATMM) or (TATATAT) is present on DNA which is called "TATA box or Hogness box"
♦ At start point RNA polymerase enzyme breaks H-bonds between two DNA strands and separates them.
♦ One of them strand takes part in transcription. Transcription proceeds in 5'→3' direction.
♦ 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.

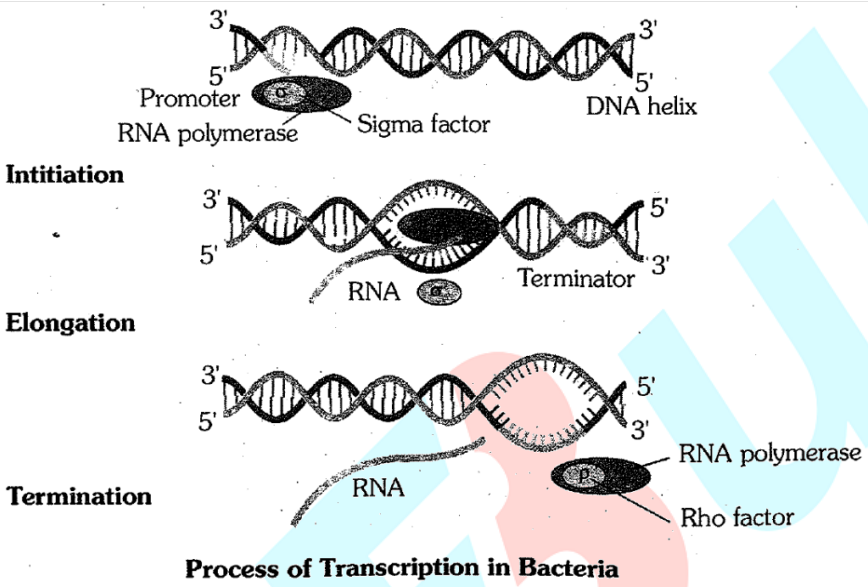
**(2) ELONGATION :-**

♦ RNA polymerase enzyme establishes phosphodiester bond between adjacent ribonucleotides.
♦ Sigma factor separates and RNA polymerase 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 templet.

- 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 (p factor).
- Rho (ρ) factor is a specific protein which helps RNA polymerase enzyme to recognise the terminator site.




**Process of Transcription in Bacteria**

## SPLIT GENE

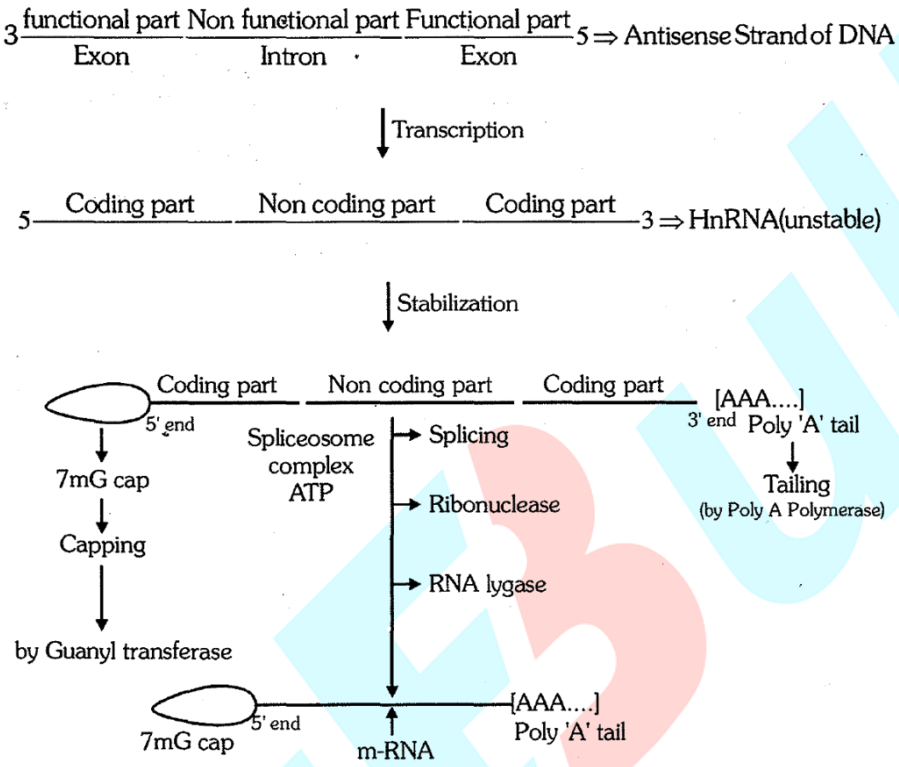
Discovered by sharp and Roberts. They were awarded Nobel Prize in 1993. Gene which contains non functional part along with functional part is known as split gene. Non functional part is called intron and functional part is called exon. By transcription split gene produces a RNA which contains coding and non coding sequence and called hn RNA (Heterogenous nuclear RNA). This hn RNA is unstable. Now 7 methyl guanonsine 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 poly 'A' tail, Now it becomes stable. By the process of RNA splicing hn-RNA produces functional m-RNA that is exonic RNA In RNA splicing non coding parts is removed with the help of spliceosome enzyme and coding part join together with the help of RNA ligase. Some specific proteinsre also helpful in RNA- splicing called 'Small nuclear ribonucleoprotein' or 'SnRNP' or 'Snurps'. 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 to produce functional RNA Non coding part of hn RNA remained inside the nucleus and not translated in to protein. Only coding part moves from nucleus to cytoplasm and gets translated into protein.

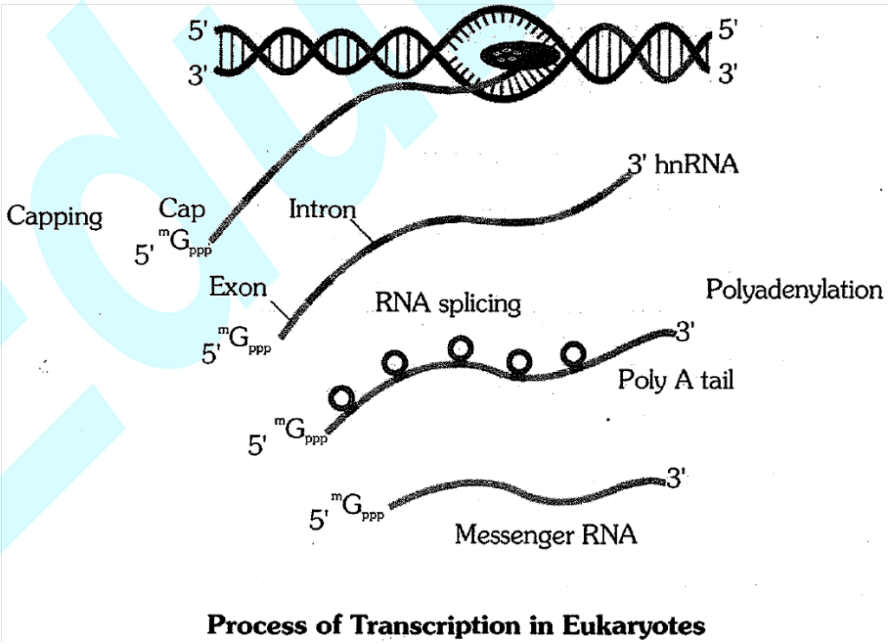
Mostly Eukaryotic genes are example of split gene, but gene which forms histone and interferon protein are non split gene.

Mostly prokaryotic genes are example of non split gene.

- In euckaryotes after transcription splicing process also occured.

- The split gene represent an ancient (primitive) feature of gene.
- Presence of intron is a primitive character.
- The splicing process represent the dominance of RNA world .



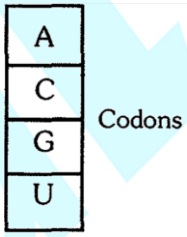

**Process of Transcription in Eukaryotes**

## GENETIC CODE

- Term Given by George Gamow.
- Discovered by Nirenberg, Matthaei and Khorana.
- The relationship between the sequence of amino acids in a polypeptide chain and nucleotide sequence of DNA or m-RNA is called genetic code.
- 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 transfered from DNA to m-RNA in the form of complementary $N_2$-base sequences.
- m-RNA contains code for each amino acid and it is called codon. A codon is the nucleotide sequence on m-RNA which codes for particular amino acid ; wherease the genetic code is the sequence of nucleotides on m-RNA molecule, which contains information for the synthesis of polypeptide chain.

**Triplet Code :-**

- The main probrem 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 N2-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 ammonia acids.
  - Single code = $4^1 = 4 \times 1 = 4$ codons




**Single code : 4 × 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 acid

| | | | |
|---|---|---|---|
| AA | AC | AG | AU |
| CC | CA | CG | CU |
| GG | GA | GC | GU |
| UU | UA | UG | UC |

**Doublet Code : 4 × 4 = 16 codons**

- Gamow (1954) pointed out the possibility of three letters code ITriplet code).
- Genetic code is triplet i.e. one codon consists of three nitrogen bases
  Triplet code = 43 = 4 x 4 x 4 = 64 codons
- In this case there occurs 64 codons hi dictionary of genetic code.
- 64 codons are sufficient to code 20 types of amino acids.
- 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).

| 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 Gin | CGA Arg | A |
| | CUG Leu | CCG Pro | CAG Gin | 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 |

**Triplet 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 one amino acid in polypeptide chain.

  **For Ex.:**

  - In m-RNA if there are total 90 $N_2$ - bases.
  - Then this m-RNA determines 30 amino acids in polypeptide chain.
  - In above example, number of nitrogen bases are 90 so codons $\Rightarrow$ 30 and 30 codons decide 30 amino acids in polypeptide chain.

**(ii) Universality :-**

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

**(iii) Non - Ambiguous and specific unamfiguous :-**

- 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 for two different amino acids. Exception GUG codon which codes both valine and methionine amino acids.

**(iv) Non - Overlapping :-**

A nitrogen base is a constituent of only one codon.

**(v) Comma less :-**

- 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.
- 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 Baumfield 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 thirc;l position (3 '- end of triplet codon) of codon. The third base is described as "Wobbly base".

**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 acid.

WOBBLE HYPOTHESIS

- It was propounded by CRICK
- Normally an anticodon recognises only one codon, but sometimes an anticodon recognises more than one codon. This is known as Wobbling. Wobbling normally occurs for third nucleotide of codon.
- For e.g. anticodon AAG can recognise two codons i.e. UUU and UUC, both stands for phenyl alanine.

**Types of m-RNA-** m-RNA is of 2 types-

**(1)** **Monocistronic** - The m - RNA in which genetic signal is present for the formation of only one polypeptide chain eg. Eukaryotes.

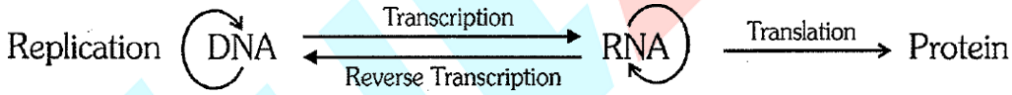
**(2)** **Polycistronic :-** The m-RNA in which genetic signal is present for the formation of more than one polypeptide chains' eg. Prokaryotes .

- Non sense codons are found in middle· position in polycistronic m-RNA

## CENTRAL DOGMA

- Central dogma was given by Crick.
- The formation (production) of m – RNA from DNA and then synthesis of protein from it, is known as central Dogma.


- It means, it includes transcription and translation.




**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.
- 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 e-DNA (Complimentary DNA).

## TRANSLATION (Protein Synthesis)

**(1)** **Activation of Amino acid :-**

- 20 types of amino acids participate in protein synthesis.
- Amino acid reacts with ATP to form "Amino acyl AMP enzyme complex", which is also known as 'Activated Amino acid'.

Amino acid + ATP $\xrightarrow[t-RNA\ synthetase]{Amino\ acyl}$ Amino acyl AMP-enzyme complex + PP

- This reaction is catalyzed by a specific 'Amino acyl t-RNA synthetase' enzyme.
- There is a separate 'Amino acyl t-RNA synthetase' enzyme for each kind of amino acid.

**(2)** **Charging oft-RNA (Loading of t-RNA) :-**

Specific activated amino acid is recognised by its specific t-RNA.

Now amino acid attaches to the 'Amino acid attachment site' of its specific t-RNA and AMP and enzyme are separated from it.

Amino acyl AMP-enzyme complex + t-RNA $\rightarrow$ Amino acyl t-RNA complex + AMP+ enzyme

♦ Amino acyl t-RNA complex is also called 'Charged t-RNA'.

♦ Now Amino acyl t-RNA moves to the ribosome for protein synthesis.

**(3) Translation :- 3 steps -**

(A) Initiation of polypeptide chain . :-

- In this step 30s and 50s sub units of ribosome, GTP, $Mg^{+2}$, charged t-RNA, m-RNA and some initiation factors are required.
- In prokaryotes there are three initiation factors present- IFl, IF2, IF3.
- Initiation factors are specific protein.
- GTP and initiation factors promote the initiation process.
- In prokaryotes with the help of "S D sequence" (Shine-Delgamo sequence) m-RNA recognizes the smaller sub unit of ribosome. A sequence of 8 $N_2$ base is present before the 4 – 12 $N_2$ base of initiation codon on mRNA, called "SD sequence". In Smaller subunit of ribosome, a complementary sequence of "SD sequence" is present on 16s rRNA, which is called "Anti Shine-Delgarno sequence" (ASD sequence)
- With the help of 'SD' and 'ASD sequence' mRNA recognises the smaller sub unit of ribosome.
- While in eukaryotes, smaller sub unit of ribosome is recognised by "7mG cap".
- In eukaryotes, 18s rRNA of smaller sub unit has a complementary sequence of "7mG cap".

  30s sub unit + m-RNA $\xrightarrow[Mg^{+2}]{IF3}$ 30s m-RNA -complex" reacts

- The “30s m-RNA – complex” reacts with 'Formyl methionyl t-RNA- complex' and 30s mRNA - formyl methionyl t-RNA - complex" is formed. This t-RNA attaches with codon part of m-RNA. A GTP molecule is required.

  30s m-RNA – complex + Formyl methionyl t-RNA-complex
  GTP $\downarrow$ IF2, IF3 $Mg^{+2}$
  30s m-RNA formyl methionyl t-RNA – complex

- Now larger sub unit of ribosome (50s sub unit) joins this complex. The initiation are factor released and complete 70s ribosome is formed.
- In larger sub unit of ribosome there are three sites for t-RNA-
  'P' site = Peptidyl site.
  'A' site = Amino acyl site.
  'E' site = Exit site
- Starting codon of m-RNA is near to 'P' site of ribosome, so t-RNA with formyl methionine amino acid first attaches to 'P' site of ribosome and next codon of m-RNA is near to 'A' site of ribosome. So next newt-RNA with new amino acid always attach at 'A' site of ribosome but in initiation step 'A' site is empty.

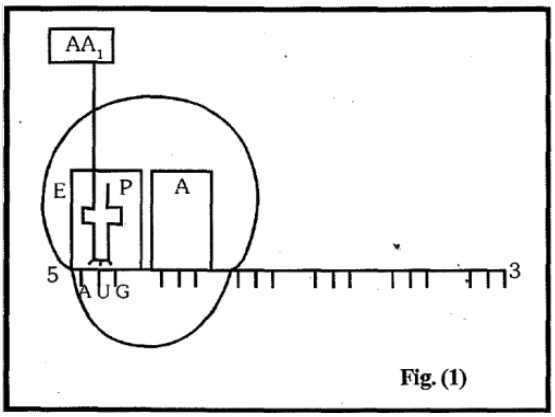


Fig. (1)

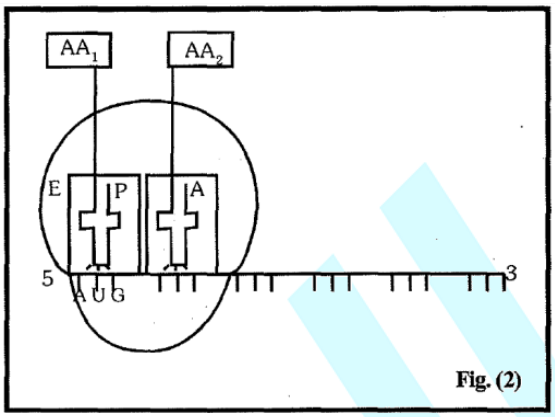


Fig. (2)

(B) Chain - Elongation :-

- New tRNA with new amino acid is attaches at 'A' site of ribosome.
- The link between amino acid of 'P' site of t-RNA is broken and t-RNA of P-site is discharged so COOH of P-site A.A. becomes free.
- Now peptide bond takes place between - COOH group of P site amino acid and - $NH_2$ group of A-site amino acid.
- 23-s rRNA induces the formation of peptide bond. This r-RNA acts as an enzyme so it is called "Ribozyme' '.
- After formation of peptide bond t-RNA of P site released from ribosome via E-site and dipeptide attaches with A site.
- Now t-RNA of A site is transferred to P site and $\therefore$ A site becomes empty.
- Now ribosome slides over m-RNA strand in 5'$\rightarrow$3' direction. Due to sliding of ribosome on m-RNA, new codon of m-RNA continuously available at A site of ribosome and according to new codon of m-RNA new amino acid attaches in polypeptide chain.
- Translocase enzyme is helpful in movement of ribosome (translocation). GTP provides energy for sliding of ribosome.
- In elongation process some protein factors are also helpful, which are known as 'Elongation factors'.
- In prokaryotes three 'Elongation factors' are present - EF-Tu, EF-Ts, EF-G.

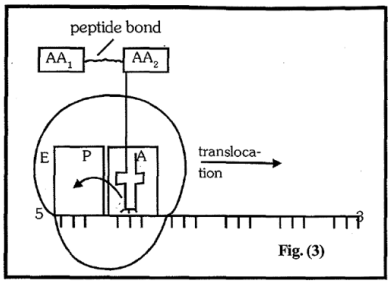


Fig. (3)

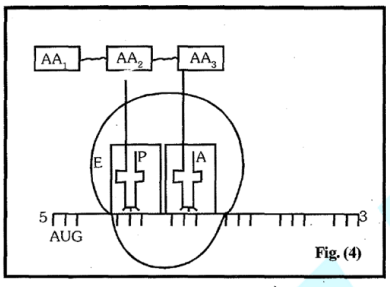


Fig. (4)

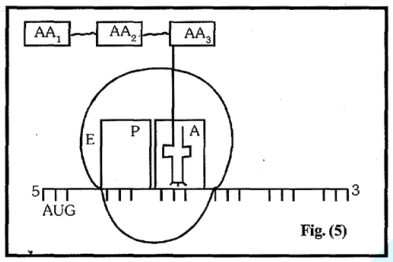


Fig. (5)

**(C) Chain - Termination :-**

- Due to sliding of ribosome over m-RNA when any Nonsense codon (UAA, UAG, UGA) available at A site of ribosome, then polypeptide chain terminates.
- The linkage between the last t-RNA and the polypeptide chain is broken by three release factor called RFl, RF2, RF3 with the help of GTP.

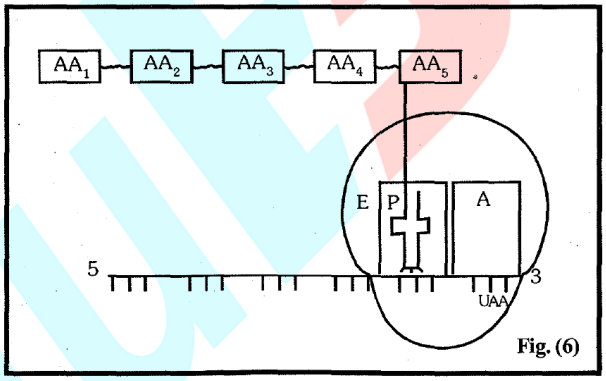


Fig. (6)

- An mRNA also have some additional sequences that are not translated and are referred as untranslated regions (UTR). The UTRs are present at both 5'end (before start codon) and at 3'end (after stop codon).
- The UTR(untranslated regions) present on mRNA are required, for efficient translation process (by recognizing the smaller subunit of ribosome by mRNA)

**Some Inhibitors of Bacterial Protein Synthesis :**

| Antibiotic | Effect |
|---|---|
| Tetracycline | Inhibits binding of amino-acyl tRNA to ribosome |
| Streptomycin | Inhibits initiation of translation and causes misreading |
| Chloramphenicol | Inhibits peptidyl transferase and so formation of peptide bonds |
| Erythromycin | Inhibits translocation of ribosome along mRNA |
| Neomycin | Inhibits interaction between tRNA and mRNA |

## BEGINNER'S BOX-2

1. Unidirectional flow of information called central dogma was given by
(1) F.H.C. Crick (2) Temin (3) Baltimore (4) Dulbecco

2. In eukaryotes, RNAPIII catalyses the synthesis of
(1) All rRNA and tRNA (2) mRNA, HnRNA and SnRNA
(3) 5S rRNA, tRNA and SnRNA (4) 28S, 18S and 5S rRNA

3. The core enzyme requires a factor for termination of RNA synthesis at some sites. This is known as
(1) Sigma factor (2) Rho factor (3) Gamma factor (4) Alpha particle

4. If one strand of DNA has the base sequence ATCCACGACTAG and the second strand undergoes transcription what would be the base sequence on mRNA ?
(1) TACGTGcTGATC (2) ATCCACGACTAG
(3) AUCCACGACUAG (4) AUGCACGACTAG

5. During protein synthesis, amino acid gets attached to tRNA with the help of
(1) mRNA (2) Arninoacyl synthetase
(3) Ribosome (4) rRNA

6. The first amino acid in any polypeptide chain of prokaryote is always
(1) Formulated methionine (2) Formulated arginine
(3) Lysine (4) Methionine

7. Which site of a tRNA molecule forms hydrogen bonds with mRNA molecule ?
(1) Codon (2) Anticodon
(3) 5' end of the t-RNA molecule (4) 3' end of the t-RNA molecule

8. To code the 50 arninoacids in a polypeptide chain, what will be the minimum number of nucleotides in its citron?
(1) 50 (2) 153 (3) 306 (4) 309

9. A single anticodon can recognize more than one codon of m-RNA. This phenomenon is termed as
(1) Richmond and Lang effect (2) Gene flow hypothesis

(3) Wobble hypothesis (4) Transposability

10. The genetic code is called a degenerate code because
(1) One codon has many meanings (2) More than one codon has the same meaning
(3) One codon has one meaning (4) There are 64 codons present

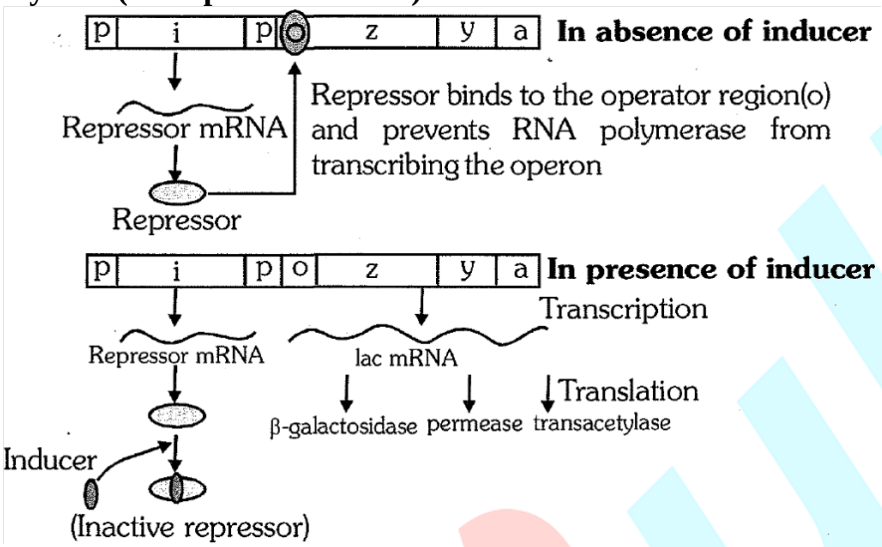
## REGULATION OF GENE EXPRESSION

- **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.
- **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. Such type of inhibition is called feed back inhibition or feed back repression.
- 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.
- Similarly, certain genes which are in their 'switch on' state, continue to synthesise a metabolite till the later is produced in amount more than required or else, it is supplied to the cell from outside. In other words, certain genes continue to express themselves till the end product of inhibits or repress their expression. Inhibition by an end product is known as 'feed back repression'.
- 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. In eukaryotes, the regulation could be exerted at
  (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.

### OPERON CONCEPT

- In 1961, two French microbiologist Francis Jocob and Jacques Monad 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 structural genes-an operator gene, a promoter gene, a regulator gene.
- Operons are of two types (i) inducible (ii) repressible.

**1. Inducible System (Lac operon of E. coli)**




♦ 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
An inducible operon system consists of four types of genes

**(i)** **Structural genes** - These genes synthesise mRNAs, which in tum 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 (2, 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-ft 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 mRNA 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** - 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** - It produces a repressor that binds to operator gene and stops the working of the operator 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 binding 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.

The inducer for lac operon of Escherichia coliis 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. These few lactose molecules are then converted into an active form which acts as an inducer and binds to the repressor protein. The inducer repressor complex fails to join with the operator which is turned on. The three genes are expressed as three enzymes to metabolise lactose. Allolactose is real inducer of lac operon.

**2. Repressible System (Tryptophan operon of E. coli.)**

♦ A repressible operon system is normally in it's switch on state and continue to synthesise a metabolise till the latter is produced in amount more than required, or else it becomes available to the cell from outside. Repressible operon system is commonly found in anabolic pathway. e.g. Tryptophan operon of E. coli.

[Inactive repressor + co-repressor = acitve repressor]

**Tryptophan operon** of Escherichia coliis an example of repressible system. It consists of the following:

**(i)** **Structural genes.** These genes are meant for transcription of mRNA, which in turn synthesise enzyme. Tryptophan operon has five structural genesE, D, C, Band A. They lie in continuation and synthesise enzymes for five steps of tryptophan synthesis.

(ii) **Operator gene (trp O).** It lies adjacent to the structural genes and controls the functioning of the structural genes. Normally, it is kept switched on, because the apo-repressor produced by the regulator gene does not bind to it. The operator gene is switched off when a co-repressor is available along with apo-repressor.

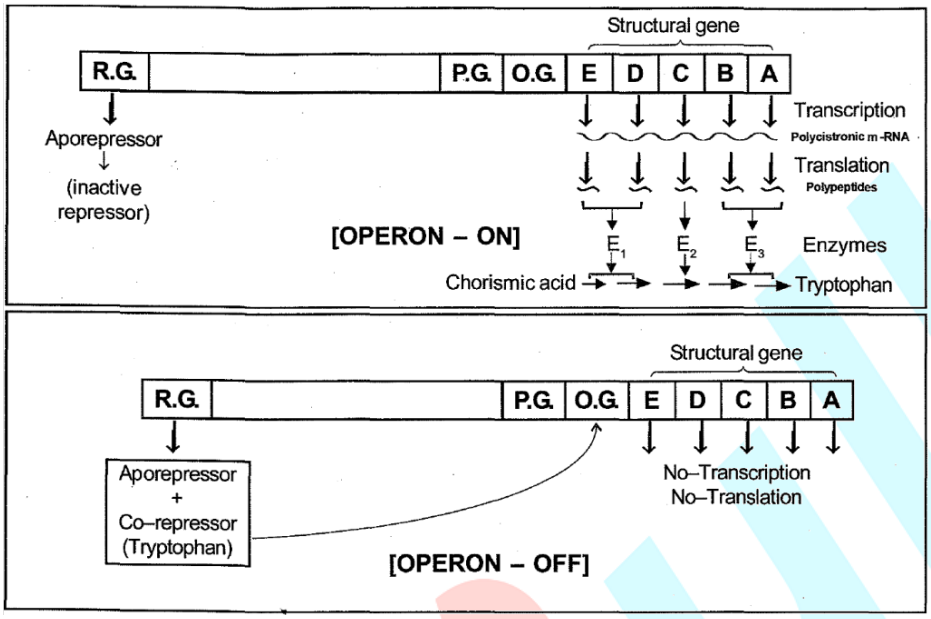
**(iii)** **Promoter gene (trp P).** It marks the site at which the RNA polymerase enzyme binds. When the operator gene is switcped on, it moves from promotor gene to structural genes for transcription.

**(iv)** **Regulator gene (trp R).** It produces a regulatory protein called apo-repressor for (Inactive repressor) possible blocking the activity of operator gene.

**(v)** **Apo-repressor.** It is a regulatory protein synthesised by regulator gene. When a co-repressor substrate is available in the cell, the apo-repressor combines with the co-repressor to form a apo-repressor co-repressor complex. This complex binds with the operator gene and switches it off. Presence of apo-repressor alone, the operator gene is kept switched on because, by itself the apo-repressor is unable to block the working of operator gene.

**(vi)** **Co-repressor.** It is an end product of reactions catalysed by enzymes produced by the structural genes.

In the presence of tryptophan some molecules of tryptophan act as co-repressor, co-repressor-bind with inactive repressor. co-repressor repressor complex bind with-operator region and prevent the binding of RNA polymerase to the promoter, the trp-operon is off.




The repressor molecule has key role in regulation of lac-operon. Repressor molecule 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.

## MUTATION

- Sudden heritable change in genetic material of an organism is called as Mutation.
- Mutation are source of discontinuous variation.
- Only those mutation are heritable which occur in germinal cell of an organism. While somatic mutations are non heritable. Somatic mutations are also heritable in vegetative propagated plants.
- Mutation word was given by Hugo De Vries.
- De Vries studied mutations in the plant Oenothera lamarckiana (evening primrose).
- Mutation was first observed by Seth Wright. He observed some short legged sheep (Ancon) variety in a population of long legged sheep.
- Beadle and Tatum induced mutations in Neurospora by the help of U. V. rays. or X-rays.

Wild Neurospora $\xrightarrow{\text{U.V.rays}}$ U. V. rays Mutant Neurospora.
(PROTOTROPH) (AUXOTROPH)

Normal-Neurospora can be grown in minimal medium because Neurospora can make all essential nutrients required for it. This is known as Prototroph.

Mutant Neurospora doesn't has capability to grow in minimal medium because due to mutation it loses those genes which codes for the enzyme that helps to prepare some special nutrients for it. They gave "one gene-one enzyme" concept. This form is known as Auxotroph .

- M.S. Swaminathan induced mutations in wheat by the help of y-rays to obtain good varieties for eg. Sharbati Sonora, Pusa Lerma. Swaminathan established y garden in IARI-New Delhi (Pusa Institute).

**Types of mutation :**

**i.** CHROMOSOMAL MUTATION
**ii.** GENE MUTATIONS

**(I) Chromosomal Mutations :**

- Change in number or structure of chromosome.
  Types of chromosomal mutation
  (1) Heteroploidy/Genomatic mutation → change in chromosome number.
  (2) Chromosomal aberration → change in structure of chromosome.

**1. Heteroploidy / Genomatic mutation**

- Change in number of one or few chromosomes in a set or number of entire set of chromosome.
  It is of two types :
  (i) Euploidy → Change in number of chromosome sets.
  (ii) Aneuploidy → Change in number of chromosome in a set.

**Euploidy:**

- Change in number of sets of chromosome i.e. either loss or addition of sets of chromosomes .
- Monoploidy (x)- Presence of one set of chromosomes .
- Diploidy (2x) -Presence of two sets of chromosomes .
- Polyploidy - Presence of more than two sets of chromosomes.
  It maybe:-
  T riploidy (3x)
  Tetraploidy (4x)
  Pentaploidy (5x)
  Hexaploidy (6x)
  Heptaploidy (7x)
  Octaploidy (8x)
- Polyploid plants with even number of sets are always fertile, reproduce sexually a11d form seeds .
- Polyploid plants with odd number of sets are always sterile don't reproduce by sexual reproduction, They don't produce seeds but they may produce seedless fruits by parthenocarpy. eg. Banana and seedless grapes.

Example

**Triticale :** **Triticum aestivum** (6x = 42) X Rye **(Secale Cereale)** (2x = 14)

↓

Triticale 4x = 28 — Sterile

↓ Colchicine

Triticale 8x = 56 — Fertile

**Aneuploidy :**

Loss or addition of chromosomes in a set of chromosomes.

**Types of Aneuploidy :**

**(1) Hypoaneuploidy (loss)**

- 2n- 1 = Monosomy :- (loss of one chromosome in one set).
- 2n -1 - 1 =Double monosomy (loss of one chromosome from each set, but these are non homologus.)
- 2n - 2 = Nullisomy Ooss of two homologus chromosome)

**(2) Hyperaneuploidy (addition)**

- 2n + 1 = Trisomy : addition of one chromosome in one set.
- 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 nondisjunction means chromosomes fail to separate during meiosis .
- Chances of aneuploidy are more in higher age female due to less activity of oocyte, so chances of syndrome increase in children who are born from higher age female.

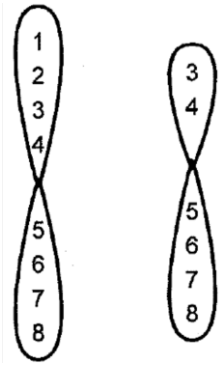
**2. Chromosomal Aberrations :**

Change in structure of chromosome.

**(i) Deletion :**

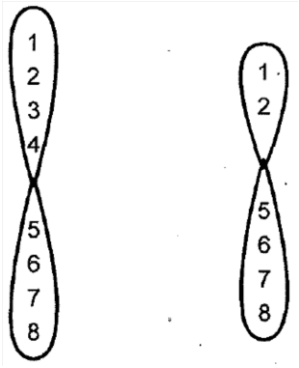
Loss of a part or segment of chromosome which leads to loss of some gene is called as deletion. It is of 2 types :-

**(a) Terminal deletion** - Loss of chromosomal segment from one or both ends.



eg. The cry -du-chat syndrome is an example of terminal deletion in short arm of 5th chromosome.

**(b) Intercalary deletion -** Loss of chromosomal part between the ends.

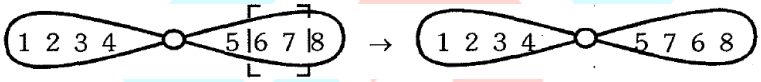



**(ii) Inversion :**

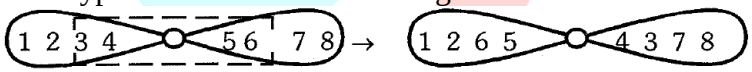
Breakage of chromosomal segment but reunion on same chromosome in reverse orders. It leads to change in distance between genes on chromosome or sequence of genes on chromosome so crossing over is affected.

It is of 2 types :-

**(a) Para centric -** If inversion occur only in one arm and inverted segment does not include centromere.




**(b) Pencentric -** In this type of inversion inverted segment include centromere.




**(iii) 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.

A B C D E F  A B C D

⟶

A B C D E F  A B C D E F E F

**(iv) 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)

Types of translocation-

**(A) Simple Translocation →** When a chromosomal segment breaks and attached to the terminal end of a non- homologous chromosome.

1 2 3 4 5 6  1 2 3 4

⟶

A B C D E F  A B C D E F 5 6

**(B) Reciprocal Translocation→** Exchange of segments between two non-homologous chromosome.

1 2 3 4 5 6  1 2 3 4 E F

A B C D E F ⟶ A B C D 5 6

eg. Chronic myloid leukemia [C M L] is a: type of blood cancer. This disease is a result of reciprocal transiocation between 22 and 9 chromosome.

**Note :** If exchange of segments takes place in between homologous chromosomes then it is called crossing over.

**(II) Gene Mutation or point mutation**

Two types:-

1. Substitution
2. 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 anemia
- Change may not occur some time because for one animo acid more than one type of codons are present.

**Substitution is of two types :-**

**1. Transnion :**

Replacement of one purine by another purine or replacement of pyrimidine by another pyrimidine.

**2. Transversion :**

Replacement of purine by pyrimidine or pyrimidine by purine is called transversion.

**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

**1. Addttion**

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

**2. Deletion**

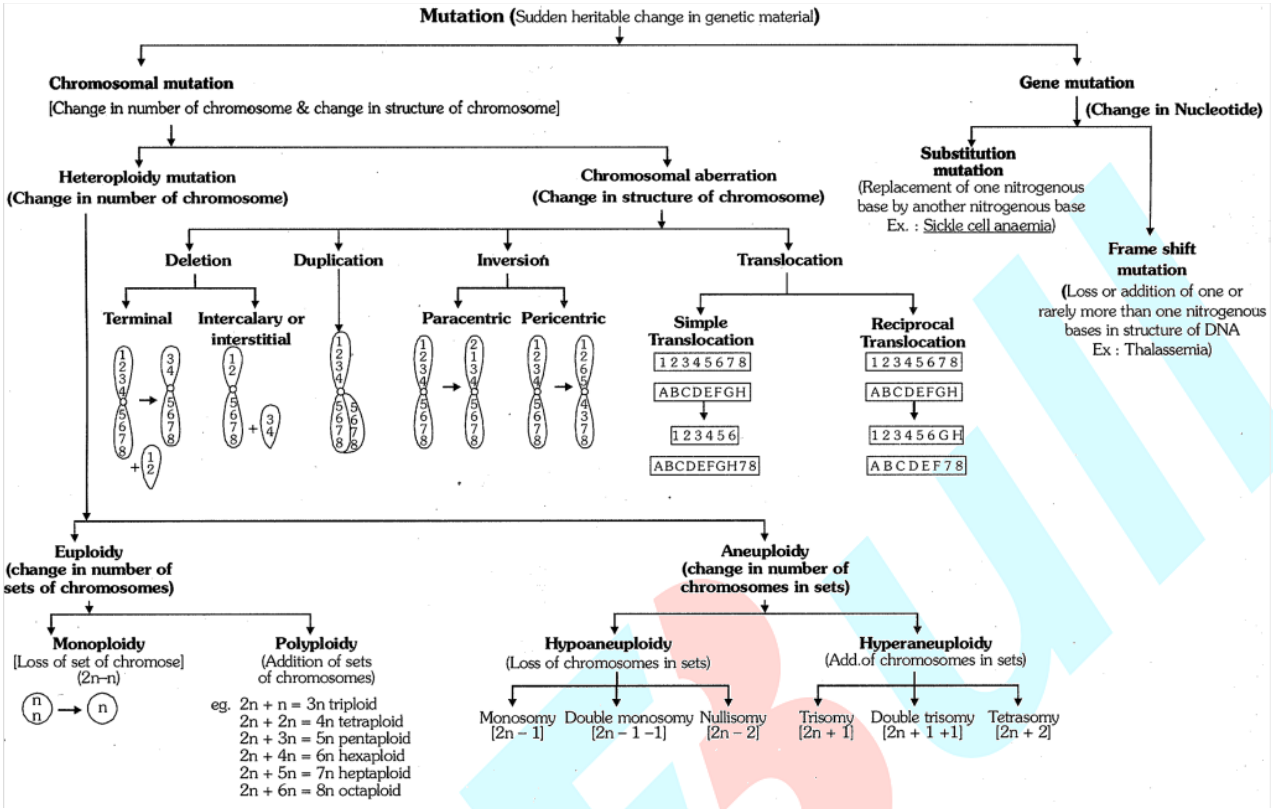
Loss of one or rarely more than one nitrogenous bases in structure of DNA.

Due to frame shift mutation complete reading of genetic code is changed. It leads to change in all animo adds 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.
  eg : Thalassemia (lethal genetic disorder)

**MUTAGENS:**

- Mutagens are those substances which cause mutations.
  **Non ionising :-** U. V. rays.


Mutation (Sudden heritable change in genetic material)
Chromosomal mutation
[Change in number of chromosome & change in structure of chromosome]
Gene mutation
(Change in Nucleotide)
Substitution mutation
(Replacement of one nitrogenous base by another nitrogenous base
Ex. : Sickle cell anaemia)
Frame shift mutation
(Loss or addition of one or rarely more than one nitrogenous bases in structure of DNA
Ex : Thalassemia)
Heteroploidy mutation
(Change in number of chromosome)
Chromosomal aberration
(Change in structure of chromosome)
Deletion
Duplication
Inversion
Translocation
Terminal
Intercalary or interstitial
Paracentric
Pericentric
Simple Translocation
Reciprocal Translocation
12345678
ABCDEFGH
123456
ABCDEFGH78
123456GH
ABCDEF78
Euploidy
(change in number of sets of chromosomes)
Aneuploidy
(change in number of chromosomes in sets)
Monoploidy
[Loss of set of chromose]
(2n–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
Hypoaneuploidy
(Loss of chromosomes in sets)
Hyperaneuploidy
(Add.of chromosomes in sets)
Monosomy [2n – 1]
Double monosomy [2n – 1 –1]
Nullisomy [2n – 2]
Trisomy [2n + 1]
Double trisomy [2n + 1 +1]
Tetrasomy [2n + 2]

## GOLDEN KEY POINTS

- Mostly mutations are harmful.
- Sometimes they are lethal which leads to death of organisms.
- But sometimes they are beneficial which are used to obtain good varieties of plants and animals. It is called as Mutation Breeding.
- Mostly mutations are recessive and they never eliminate from a population.

**Forward and Backward Mutation :**

- Wild gene ⇕ Forward / Backward ⇕ Mutant gene

**Muton (unit of mutation) :**

- Smallest part of DNA which undergoes mutation.
- It is one nucleotide.

**Mis-sense mutation :-**

- When a nucleotide change in genetic code cause the change of one amino acid of a polypeptide chain it is called mis-sense mutation.

**Non-sense mutation :-**

- When a nucleotide change in one codon causes termination of polypeptide synthesis by producing non-sense codon.

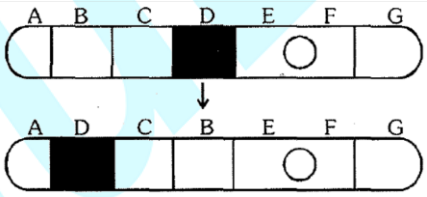
**Same sense codon:-**

- A change in one nucleotide in a codon does not change amino acid in polypeptide chain, because both codons code same amino acid.

## BEGINNER'S BOX-3

1. (2n-1) condition of chromosomes is called :-
   (1) tetrasomy (2) trisomy (3) monosomy (4) nullisomy

2. Given below is the representation of ·a kind of chromosomal mutation. It is :-




   (1) Deletion (2) Inversion (3) Duplication (4) Reciprocal translocation

3. Which of the following has normal sex chromosome complements.
   (1) Down's syndrome (2) Klinefelter's syndrome
   (3) Super female (4) Turner's syndrome

4. Addition or deletion of a nitrogenous base causes
   (1) Frameshift mutation (2) Inversion
   (3) Transformation (4) Translocation

5. Trisomy of which chromosome is involved in Down's syndrome.
   (1) $8^{th}$ (2) $13^{th}$ (3) $21^{st}$ (4) $22^{nd}$

## 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 analyze to identify people.
  These differences are called Polymorphism (many forms) and 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 fragment 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 fragmnt is different for each DNA and length of restricted fragment is depend 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. Restrtction Fragment Length Polymorphism distributed throughout human genomes are useful for DNA Fmger 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.**

1. **Extraction-** DNA 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 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 restricated fragment 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. Activated electrodes at either end of the gel provides the driving force. A molecule's properties determine, how rapidly an electric field can move the molecule through a gelatinous medium .
- Nowadays the most commonly used matrix is agarose which is a natural polymer extracted from sea weeds. The DNA fragments separate (resolve) according to their size through sieving effect provided by the agarose gel.
- Many important biological molecules such as amino acids, peptides, proteins, nucleotides, and nucleic acids posses ionisable groups and, therefore, at any given pH, exist in solution as

electrically charged species either as cation (+) or a.nions (–). Depending on the nature of the net charge, the charged particles will migrate either to the cathode or to the anode.

- 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 first step is to denature the DNA in the gel which means that the double-stranded restriction fragments are chemically separated into the single stranded form.

The DNA then is transferred by the process of blotting to a sheet of nylon. 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. Commonly we use a combination of at least 4 to 6 separate DNA probes.

Labelled Probes are attached with the VNTR loci of restricted DNA Fragments, this process is called Hybridization.

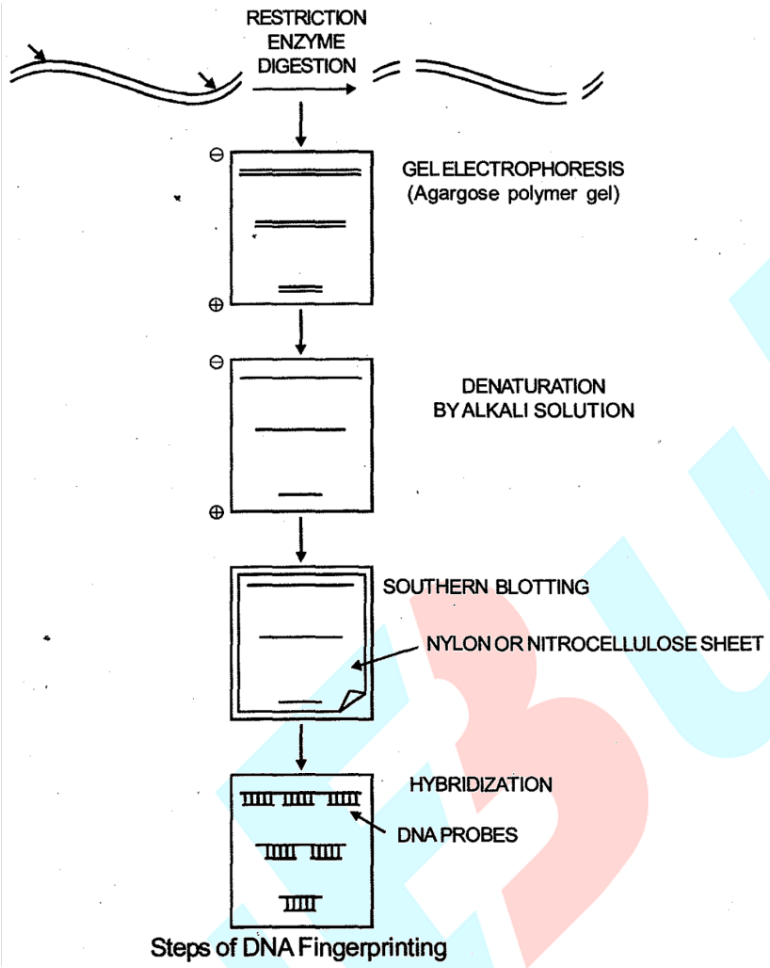
**6. Autoradiography :** Nylon membrane containing radio active probe exposed to X-ray. Specific bands appear on X-ray film. These bands are the areas where the radioactive probe bind with the VNTR.

- 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) _

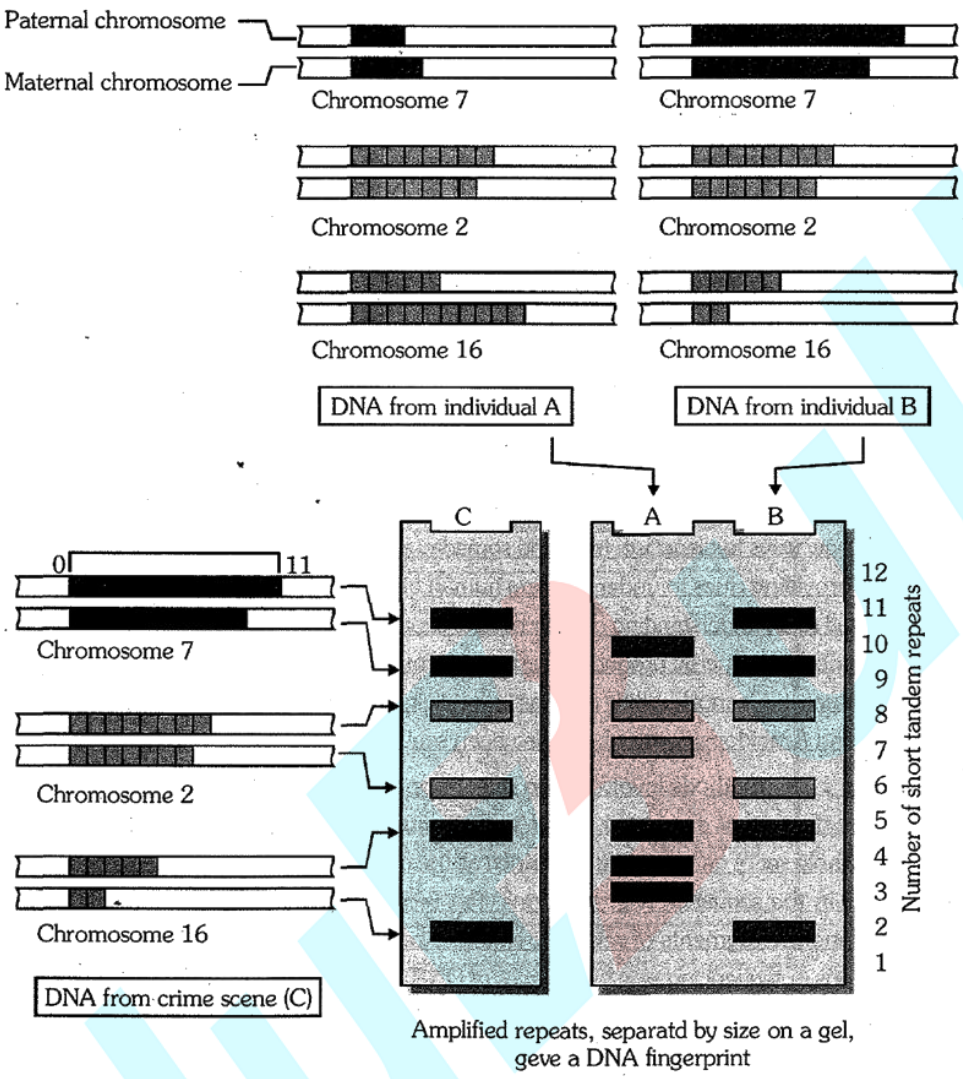
In India the centre for DNA finger printing and diagnosis (CDFD - center for DNA finger printing & diagnosis) located at Hyderabad.



Steps of DNA Fingerprinting

**Application of DNA Finger printing**

1. **Paternity tests.** The major application of DNA finger printing is in determining family relationships. For identifying the true (biological) father, DNA samples of Child, mother and possible fathers are taken and their DNA finger prints are obtained. The prints of child DNA match to the prints of biological parents.
2. **Identification of the criminal.** DNA finger printing has now become useful technique in forensic (crime detecting) science, specially when serious crimes such as murders and rapes are involved. For identifying a criminal, the DNA fingerprints of the suspects from blood or hair or semen picked up from the scene of crime are prepared and compared. 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.




**Schematic representation of DNA fingerprinting : Few representative chromosomes have been shown to contain different copy number of VNTR. For the sake of understanding colour schemes have been used to trace the origin of each band in the gel. The two alleles (paternal and maternal) of chromosome also contain different copy numbers of VNTR. It is clear that the banding pattern of DNA from crime sceme matches with individual B, and not with A.**

# 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 geome was launched in the year 1990.

**Human Genome Project** (HGP) was called a mega project. You can imagine the magnitude and the requirements . for the project if we simply define the aims of the project as follows :

Human genome is said to have approximately $3 \times 10^9$ 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 aproximately 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 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) full Improve tools for data analysis.
(v) Transfer related technologies to other sectors, such as industries.
(vi) Address the ethical, legai, and social issues (ELSI) that may arise from the project.

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 freeliving 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, sequendng the whole set of genome that contained all the coding and non-coding sequence, and later assigning different regions irt 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. (Remember, Sanger is also credited for developing method for determination of amino acid sequences in proteins). These sequences were then arranged based on some overlapping regions present in them. This required generation of overlapping fragments for seguendng. Alignment of these Sequences was humanly not possible. Therefore, specialised computer based progran1mes were developed. These sequences were subsequently am1otated 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). 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.

**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 iliought to have no direct coding functions, but iliey 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 ilie process of finding chromosomal locations for disease-associated sequences and tracing human history.

| Organisms | Base pair | Gene No. |
|---|---|---|
| Bateriophage | 10,000 | -----/---- |
| Lily | 106 Billion B.P. | ----- |
| *E.coli* | 4.7 million B.P. | 4,000 |
| *S. cerevisiae* | 12 Million B.P. | 6,000 |
| *D. melangaster* | 180 Million B.P. | 13,000 |
| *Caenorhabditis elegans* | 97 Million B.P. | 18,000 |
| Human | 3 Billion B.P. | 30,000 |

(a) First prokaryotes in whidl romplete genome WC/3 sequenred is Haemophilus influenzae.

(b) First Eukaryote in which complete genome was sequenced is Saccharomyces cerviceae (Yeast).

(c) First plant in which complete genome was sequenced is Arabidopsis thaliana (Small mustard plant).

(d) First animal in which complete genome was sequenced is Caenorhabditis elegans (Nematode).

♦ β - globin and insulin gene are less than 10 kilo base pair T.D.F. gene is the smallest gene (14 base pair) and Duchenne muscular Dystrophy gene is made up of 2400 kilo base pair.(Longest gene)

## BEGINNER'S BOX-4

1. DNA finger printing involves identifying differences in some specific
(1) Repetitive DNA (2) Non repetitive DNA
(3) Selfish DNA (4) All of the above

2. Which of the following is produced by &Coli in the lactose operon.
(1) B galactosidase (2) Transacetylase
(3) Permease (4) All of the above

3. Maximum number of gene present on which chromosome number in human.
(1) $1^{st}$ (2) X (3) Y (4) $10^{th}$

4. In lac operon RNA polymerase binds with
(1) Promoter gene (2) Operator gene
(3) Structural gane (4) Regulator gene

5. Fill the gap in following statement
Human genome have approximately _____ and the cost of sequencing was ____ per base pair.
(1) $4 \times l0^9$ bp, 9 billionUS dollars (2) 9 billion US dollars, $4 \times 19^9$ bp
(3) $3 \times 10^9$ bp, 3 US dollars (4) 4.7 million bp, 9 billion US dollars

## ANSWER KEY

### BEGINNER'S BOX-1

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1.** | (1) | **2.** | (2) | **3.** | (1) | **4.** | (2) | **5.** | (4) | **6.** | (2) | **7.** | (2) |
| **8.** | (4) | **9.** | (4) | **10.** | (2) | | | | | | | | |

### BEGINNER'S BOX-2

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1.** | (1) | **2.** | (3) | **3.** | (2) | **4.** | (3) | **5.** | (2) | **6.** | (1) | **7.** | (2) |
| **8.** | (3) | **9.** | (3) | **10.** | (2) | | | | | | | | |

### BEGINNER'S BOX-3

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **1.** | (3) | **2.** | (2) | **3.** | (1) | **4.** | (1) | **5.** | (3) |

### BEGINNER'S BOX-4

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **1.** | (1) | **2.** | (4) | **3.** | (1) | **4.** | (1) | **5.** | (3) |