BYJU'S Classes Notes

Molecular Basis of Inheritance

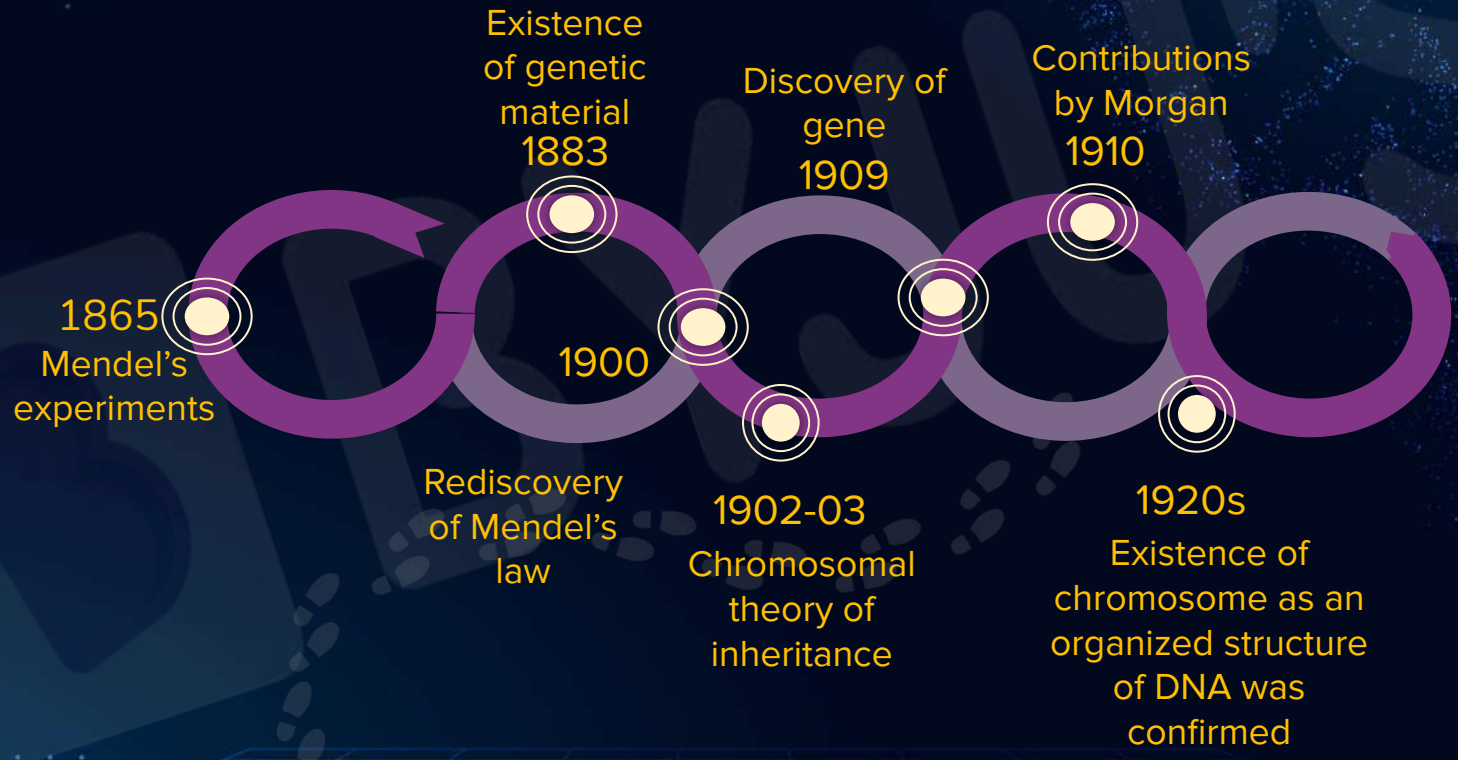
Griffith Experiment, Hershey and Chase Experiment, Avery, MacLeod and McCarty Experiment

Key Takeaways

1. Griffith’s experiments
   - Transforming principle
2. Avery, MacLeod and McCarty experiment
3. Hershey and Chase’s experiment

Summary

Search for Genetic Material



Recall: Mendel's Conclusions

1 Factors come in pairs

Pod	Stem	Seed	Flower
Inflated (full)	Axial flowers	Round	Violet
Green	Tall (6-7 ft)	Yellow	
Constricted	Terminal flowers	Wrinkled	White
Yellow	Dwarf (3/41 ft)	Green	

2 Factors segregate during gamete formation

Heterozygous

Parents: Yy

Gametes: Y, y

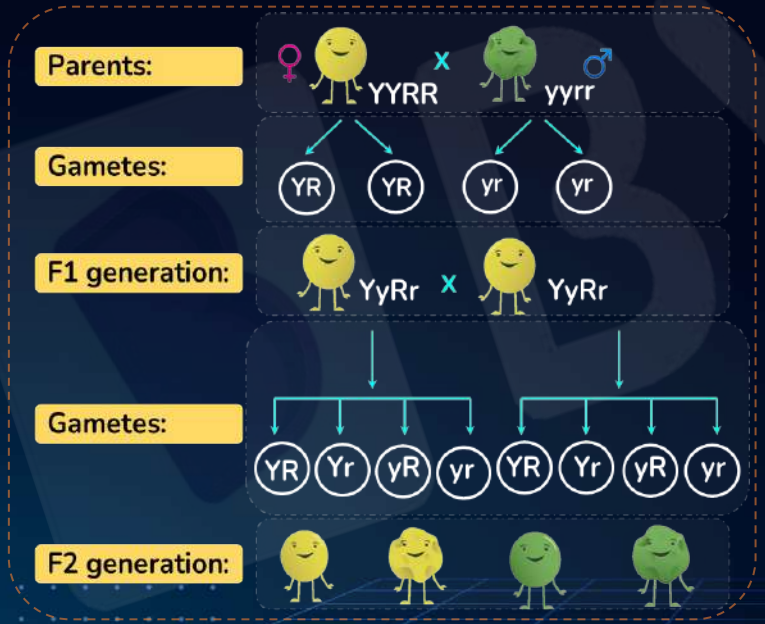
Homozygous

Parents: YY

Gametes: Y, Y

Recall: Mendel's Conclusions

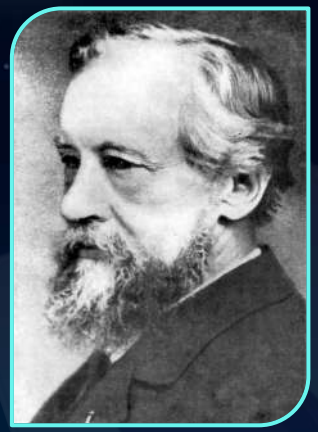
3 Factors segregate independently of each other



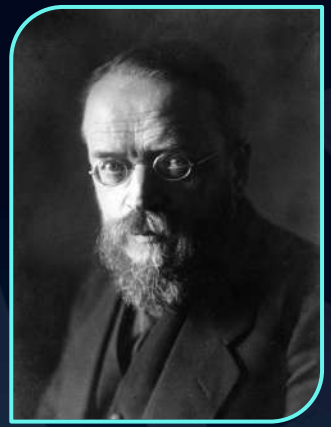
Mendel's conclusions

- Factors come in pairs
- Factors segregate during gamete formation
- Factors segregate independently of each other

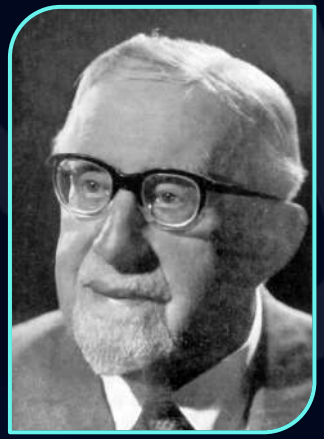
Recall: Resurgence of Genetics



Hugo de Vries



Carl Correns



Erich Von Tschermak

With improvement in technology, scientists discovered that chromosomes separate during cell division, specifically mitosis.

Mendel's work was rediscovered by them in 1900 because of the improvement in technology.

Recall: Sutton and Boveri Theory of Inheritance

Walter Sutton

Theodore Boveri

Chromosomal theory of inheritance

- Chromosomes come in pairs
- Chromosomes segregate during gamete formation
- Chromosomes segregate independently of each other

They had given three laws on the basis of their experiments and together all these constituted the chromosomal theory of inheritance.

Recall! Sutton and Boveri Theory of Inheritance

Inheritance of chromosomes = Inheritance of factors

Hypothesis (by Sutton and Boveri)
Mendel's factors are present on specific locations of the chromosomes.

The term gene was coined by Wilhelm Johannsen. It is the fundamental physical and functional unit of heredity.

Recall! Contributions by Morgan

T H Morgan

He gave the confirmation to the chromosomal theory of inheritance with his experiments on *Drosophila*.

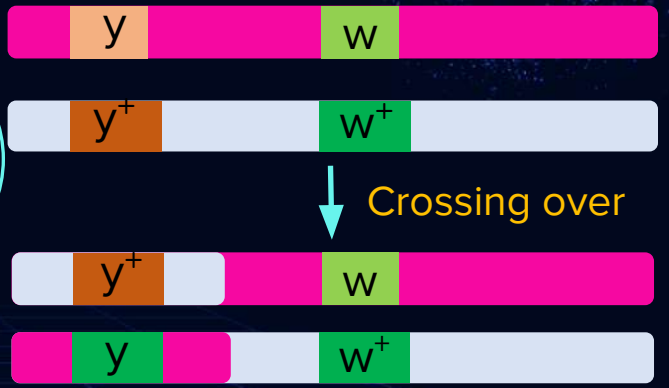
Recall: Contributions by Morgan

Linkage

Linkage is association of two genes located on the same chromosome. It describes the probability of the two genes being inherited together.

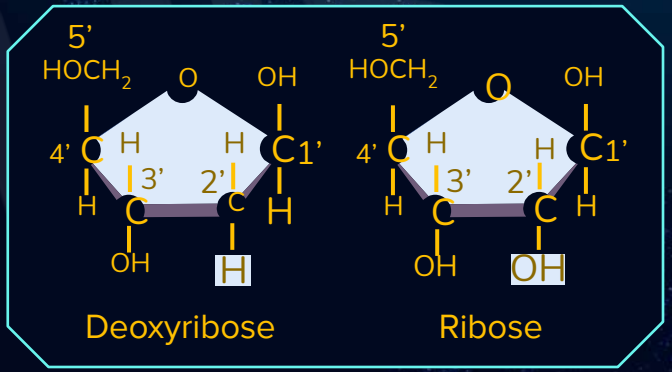
Recombination

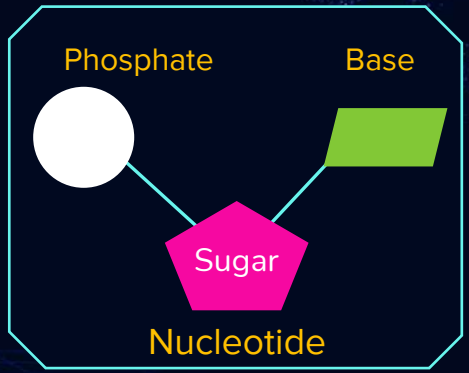
Crossing over leads to new alleles in offspring and the combination of these alleles is called recombination.



Chromosomes and DNA

- By 1920 the existence of chromosomes as an organized structure of DNA was confirmed.
- It was also known that, when a phosphate group is attached to a nucleoside it becomes a nucleotide.
- When the sugar is ribose, then the nucleotide is ribonucleic acid or RNA, and when the sugar is deoxyribose then it is called deoxyribose nucleic acid or DNA.



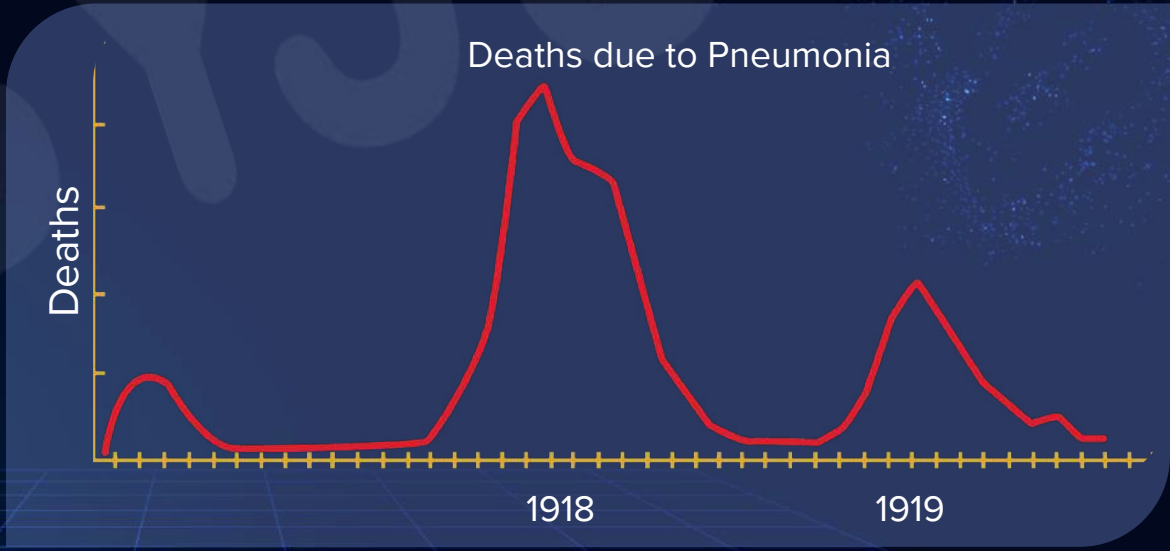


Griffith's Experiments

Frederick Griffith

- In 1928, a scientist, Frederick Griffith began experiments with *Streptococcus pneumoniae.*
- During that time, the deaths due to pneumonia increased.
- Pneumonia was a serious cause of death in the wake of the post-WWI Spanish influenza pandemic.

- Griffith was studying the possibility of creating a vaccine.
- He started conducting experiments on *Streptococcus pneumoniae* (bacterium responsible for pneumonia).



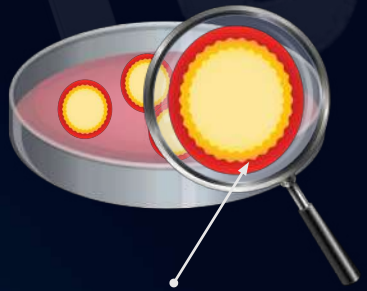
Griffith's Experiments

- The bacteria *Streptococcus pneumoniae* (pneumococcus) was grown on a culture plate.
- Some produce smooth shiny colonies (S) while others produce rough colonies (R).
- This is because the S strain bacteria have a mucous (polysaccharide) coat, while R strain do not.
- S strain or smooth bacteria are virulent, whereas R strain bacteria are non-virulent.

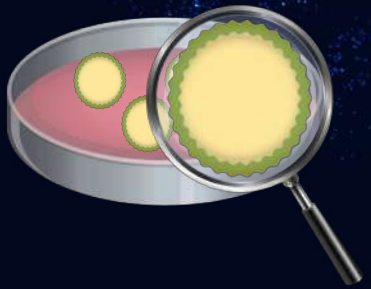
Streptococcus pneumoniae

- Smooth colonies
- S strain
- Virulent

- Rough colonies
- R strain
- Non-virulent



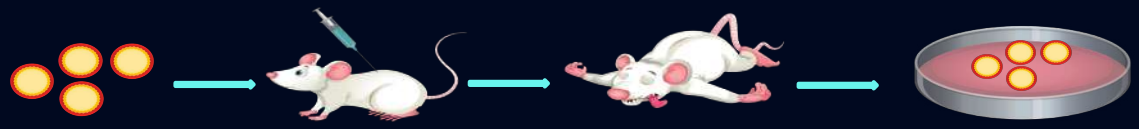
Polysaccharide coat



No polysaccharide coat

Griffith's Experiments

Step 1



Live S strain → Live mice → Mice die → S strain bacteria isolated from dead mice

He injected mice with live S strain of bacteria which are virulent and the mice died, then he isolated S strain bacteria from the dead mice.

Step 2

Live R strain → Live mice → Mice live → No living bacteria was isolated from mice

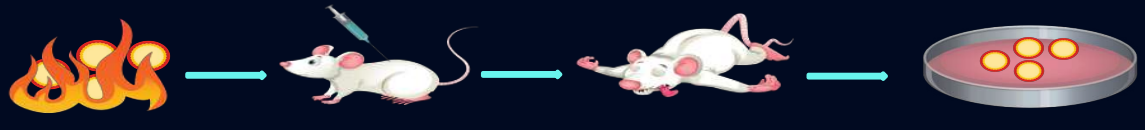
In the next step, he injected mice with live R strain of bacteria which are non-virulent and the mice survived, but this time, no living bacteria were isolated from the mice.

Griffith's Experiments

Step 3

Heat killed S strain → Live mice → Mice live → No living bacteria was isolated from mice

Then in the third step, he injected mice with S strain of bacteria which were heat killed. The mice survived and no living bacteria were isolated from the mice.



Heat killed S strain + Live R strain → Live mice → Mice die → S type bacteria was isolated from dead mice

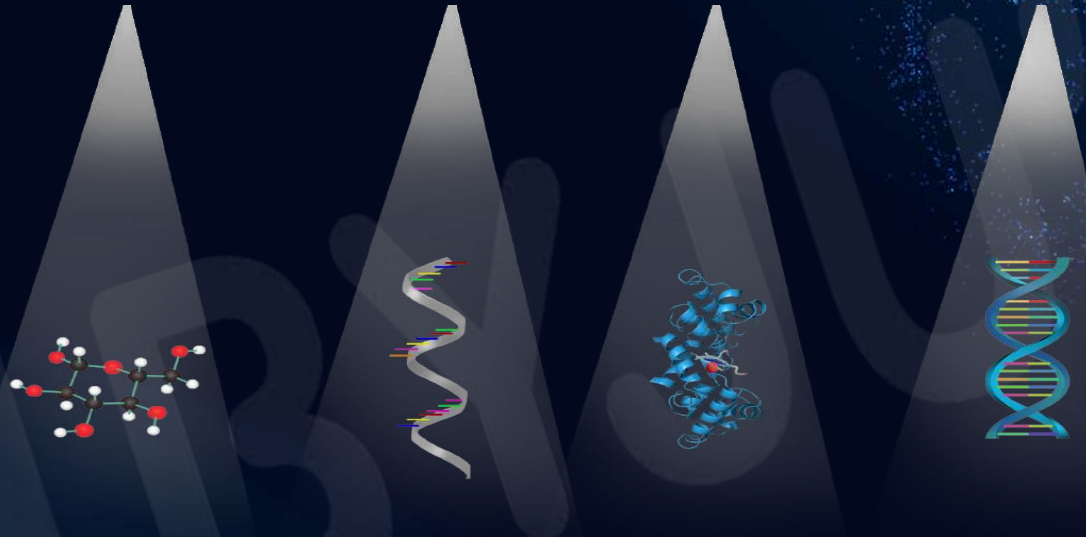
Finally, he injected both heat killed S strain and live R strain. As a result the mice died and Griffith was able to isolate S strain bacteria from the dead mice.

Griffith's Experiments

Transforming principle

- When S strain and R strain bacteria are mixed, the non-virulent R strain of bacteria is transformed into the virulent S strain bacteria.
- This process is called transformation, and the molecule with which it was happening is called the transforming principle.
- This, 'transforming principle', was being transferred from the heat-killed S strain.
- This had enabled the R strain to synthesise a smooth polysaccharide coat and become virulent. Griffith concluded that this must be due to the transfer of the genetic material.

Griffith's Experiments

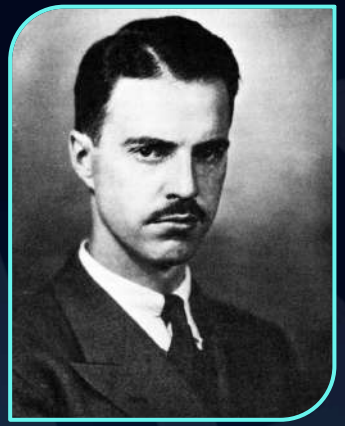


Carbohydrates RNA Proteins DNA

- The Griffith experiment was a turning point towards the discovery of genetic material.
- However, it failed to explain the biochemistry of genetic material.
- The exact biochemical nature of genetic material was unknown. It could be proteins, carbohydrates, RNA or DNA.

Avery, MacLeod and McCarty Experiment

Oswald Avery



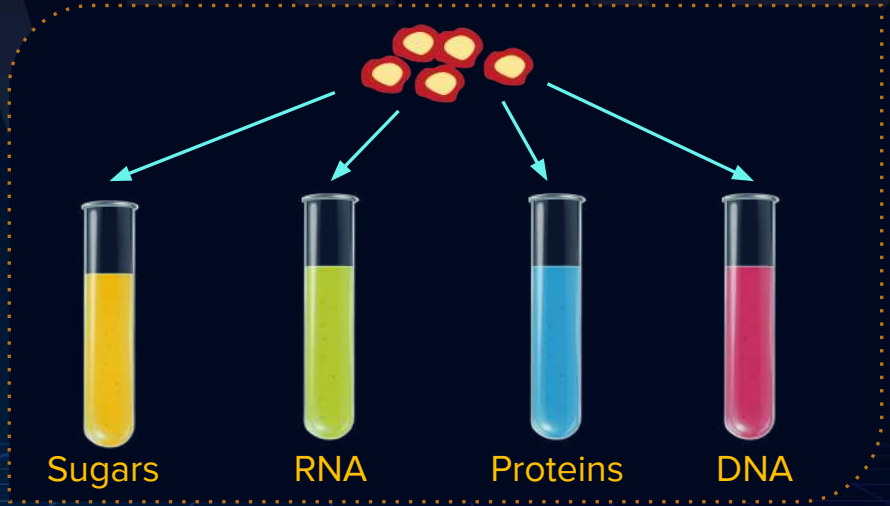
Colin MacLeod

Maclyn McCarty

- In 1944, a group of scientists, Oswald Avery, Colin MacLeod and Maclyn McCarty continued the Griffith experiment in search of the biochemical nature of the hereditary material.
- They purified biochemicals, carbohydrates, proteins, DNA and RNA, from the heat-killed S cells.

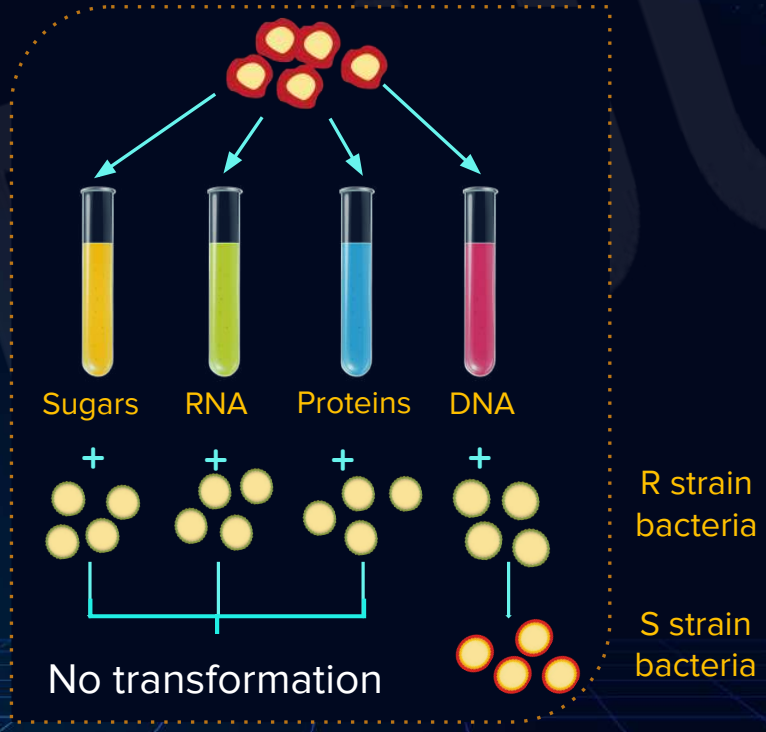
Avery, MacLeod and McCarty Experiment

- They purified biochemicals, carbohydrates, proteins, DNA and RNA, from the heat-killed S cells.
- To those solutions, they added the live R strain bacteria.
- The sugars/carbohydrates, RNA and proteins showed no transformation.

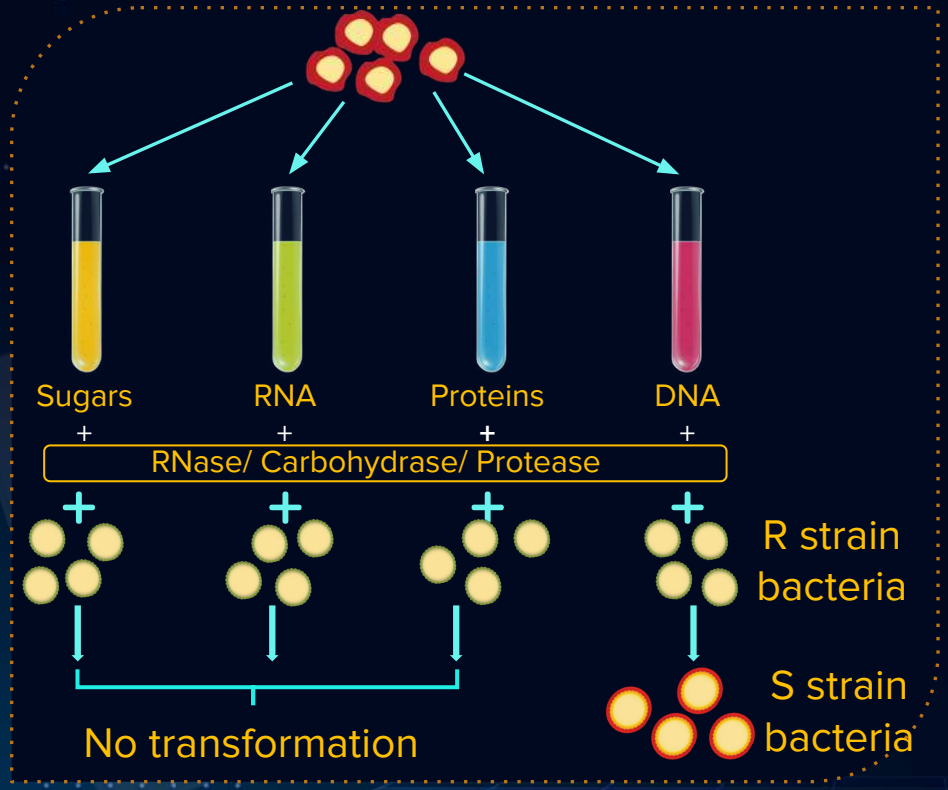


Avery, MacLeod and McCarty Experiment

The one with DNA in it transformed the R strain bacteria into the S strain. It shows that DNA is the transforming principle.



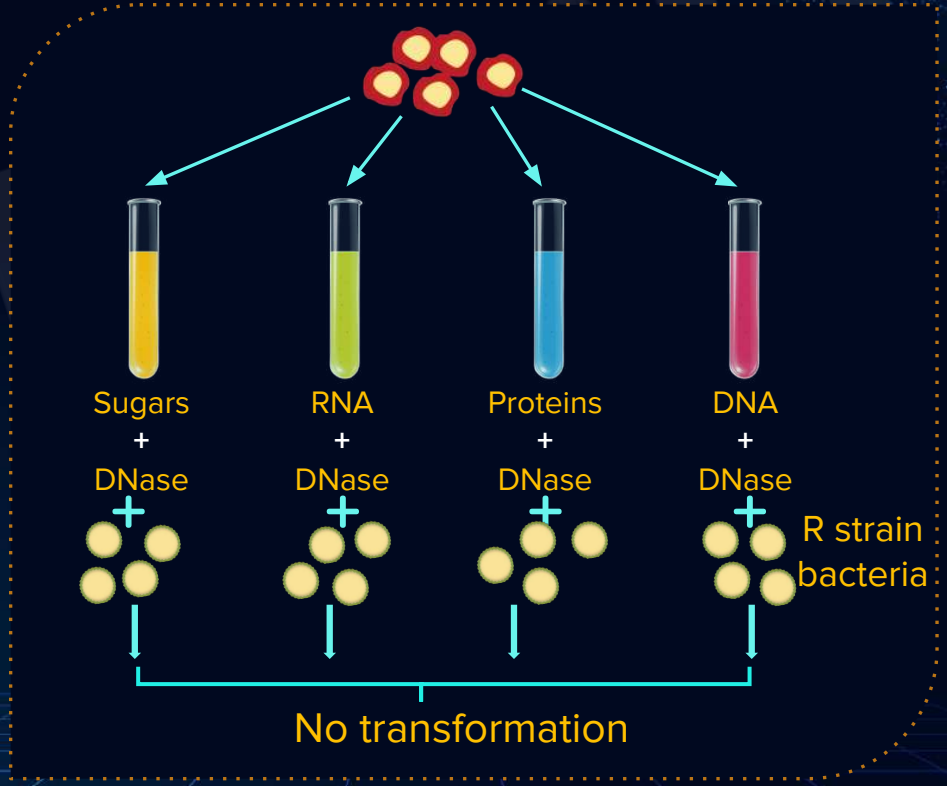
Avery, MacLeod and McCarty Experiment



- To confirm that DNA is the “transforming principle”.
- They added either:
  - carbohydrase (the enzyme which breaks down carbohydrates in all the solutions).
  - RNase (the enzyme which degrades RNA molecules).
  - Protease (the enzyme that breaks down proteins).
- Still, the solution with DNA caused transformation of R strain to the virulent S strain.

Avery, MacLeod and McCarty Experiment

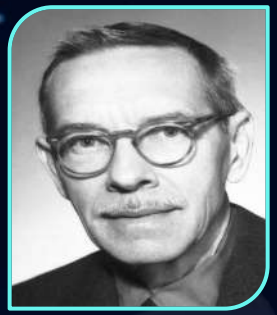
But when they added DNase in all the solutions, none of the solutions showed transformation. This proved that DNA is the genetic material.



Transforming Principle

- They concluded that DNA is the hereditary material, but not all biologists were convinced.
- At the time, the belief was that DNA was just a chain of four repeating nucleotides, and thus was not complex enough to be genetic material.
- On the other hand, proteins were known to be very diverse.
- Some scientists believed that the active component of the transforming substance was a protein associated with nucleic acids in the chromosomes.
- And that Avery and his colleagues were not able to purify the DNA properly and thus suggested that the DNA could have been contaminated by traces of protein.

Hershey and Chase Experiment

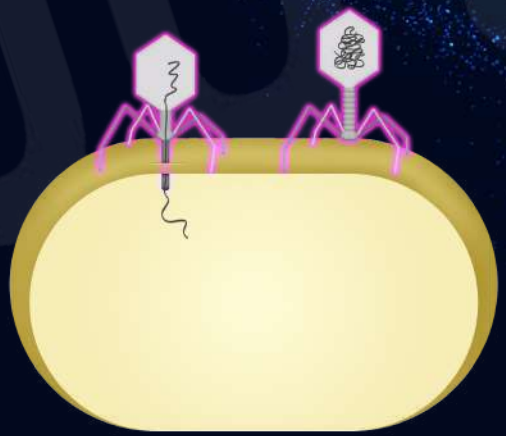


Alfred Hershey

Martha Chase

In 1952, Alfred Hershey and Martha Chase worked on viruses that infect bacteria called bacteriophages.

Transduction: The process by which DNA is introduced into a cell by a virus.



Hershey and Chase Experiment

Step 1

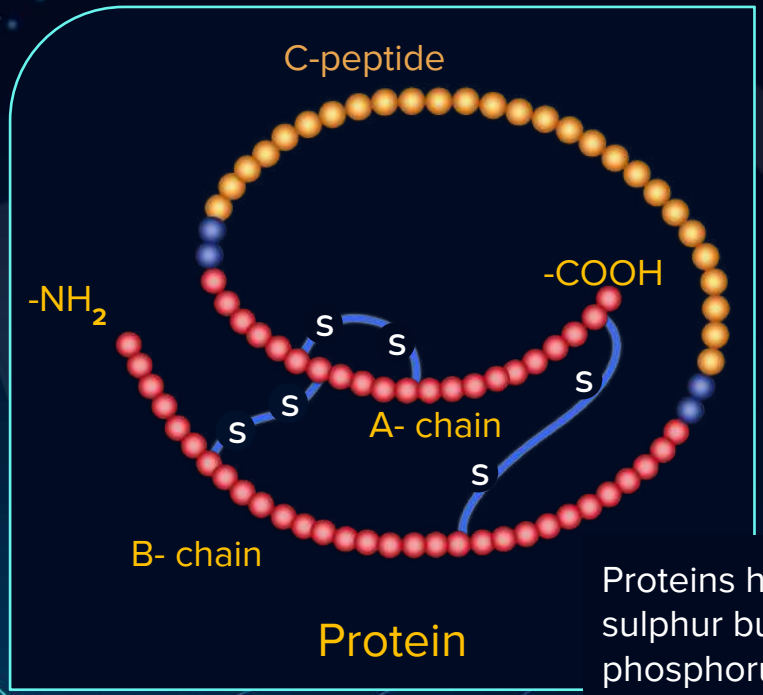
Medium with radioactive sulfur (^{35}S)

Medium with radioactive phosphorus (^{32}P)

- They grew some viruses on a medium that contained radioactive phosphorus and some others on medium that contained radioactive sulfur.
- Viruses grown in the radioactive phosphorus contained radioactive DNA but not radioactive protein.
- Viruses grown in the radioactive sulfur contained radioactive protein but not radioactive DNA.

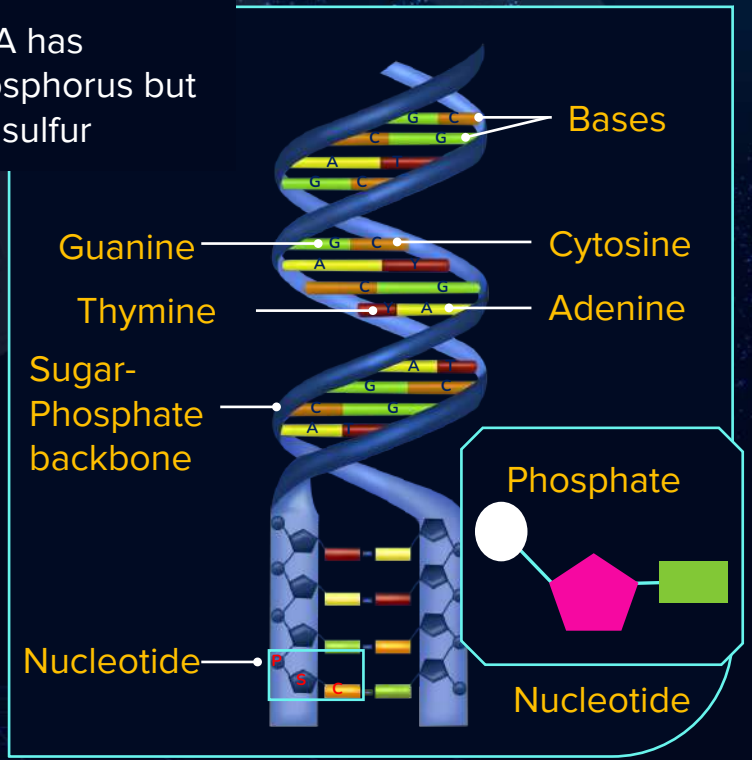
Hershey and Chase Experiment

Reasons for growing viruses on radioactive sulfur and phosphorus.



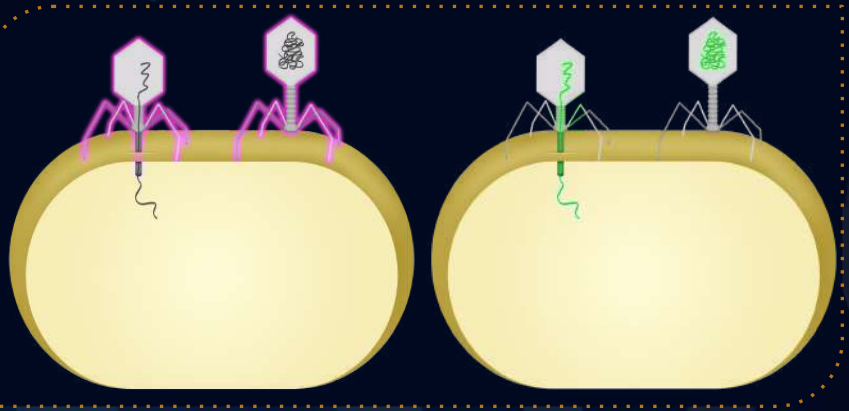
Proteins have sulphur but not phosphorus

DNA has phosphorus but not sulfur



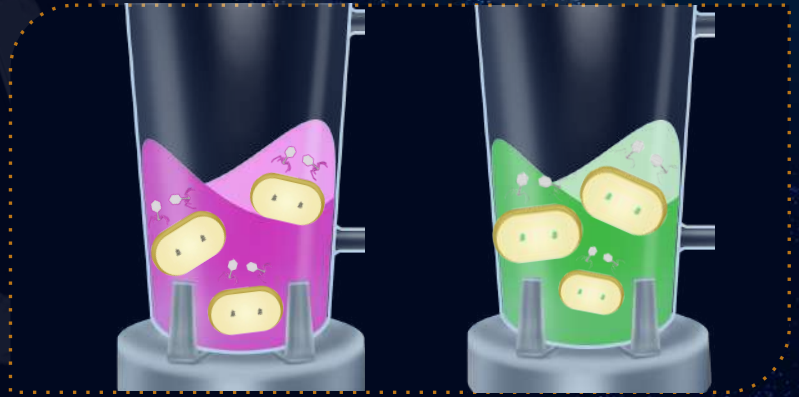
Hershey and Chase Experiment

Step 2



Radioactive phages were allowed to attach to *E. coli* bacteria and infect them.

Step 3



Bacteria were agitated in blender and protein coats were removed.

Hershey and Chase Experiment

Step 4

- The virus particles were separated from the bacteria by spinning them in a centrifuge.
- Centrifugation separated the bacterial cells and the viruses into two different levels as bacterial cells are heavier than virus coats.

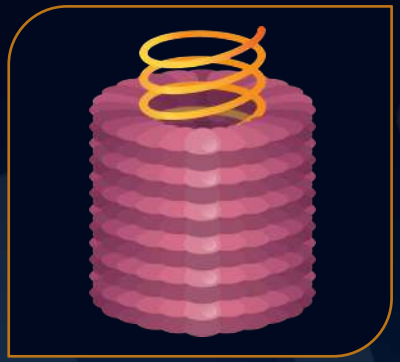
Hershey and Chase's Experiment

Step 4

- Results : The bacteria which was transfected with radioactive DNA viruses were radioactive, indicating that DNA was the material that passed from the virus to the bacteria.
- However, bacteria that were infected with viruses containing radioactive proteins were not radioactive.
- This showed that proteins did not enter the bacteria from the viruses.
- Hence, it was proved that DNA is the genetic material.

The Genetic Material

RNA viruses



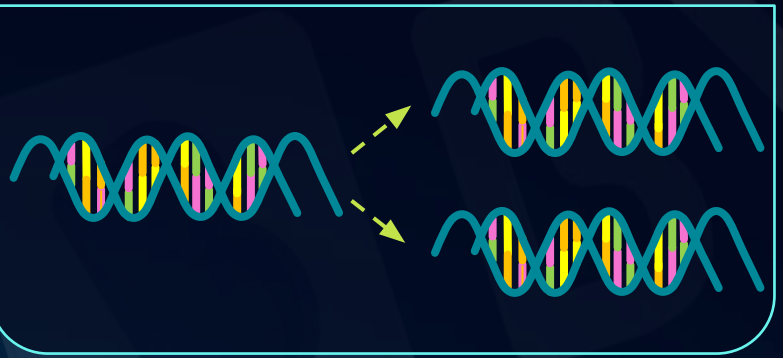
Tobacco mosaic virus

Coronavirus

- DNA is the genetic material of most organisms
- it subsequently became clear that in some viruses, RNA is the genetic material (for example, Tobacco mosaic viruses, Coronavirus, etc.).

Ideal Genetic Material

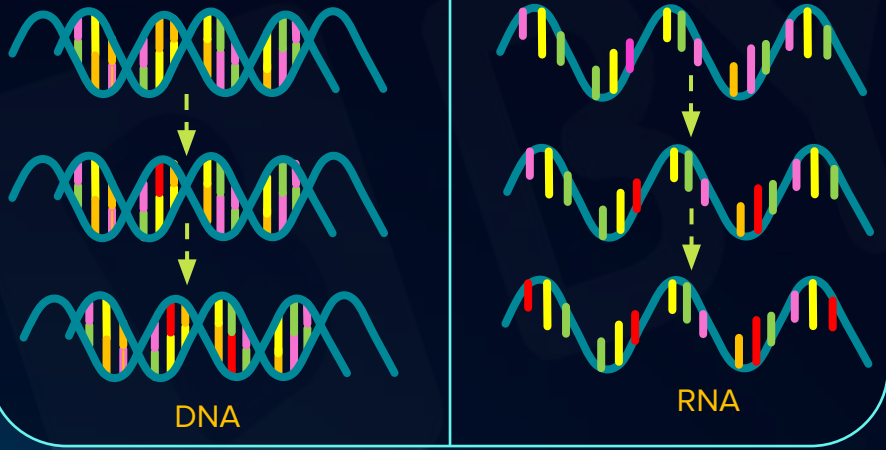
- It should be able to replicate



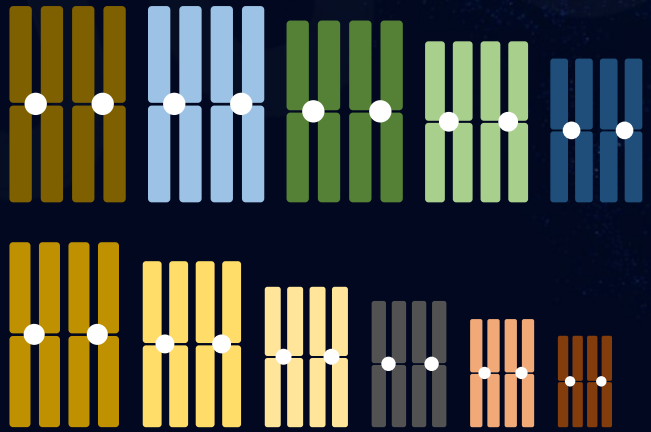
- It should be chemically and structurally stable

Ideal Genetic Material

- It should provide scope for slow changes (mutation) required for evolution



- It should be able to express itself in the form of ‘Mendelian characters’



Summary

Griffith experiment

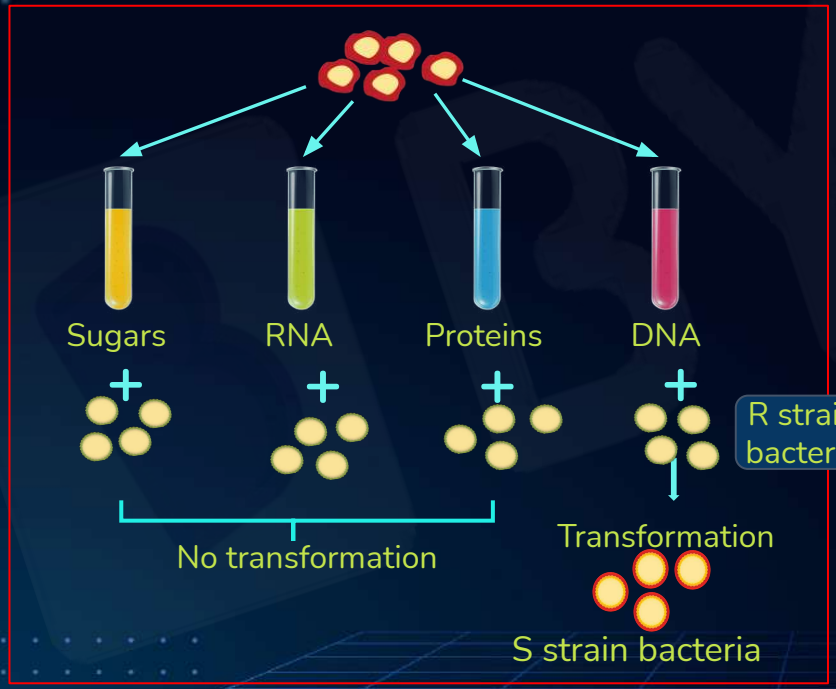
Live S strain	Live R strain	Dead S strain	Live R strain + dead S strain
Mouse dies	Mouse lives	Mouse lives	Mouse dies



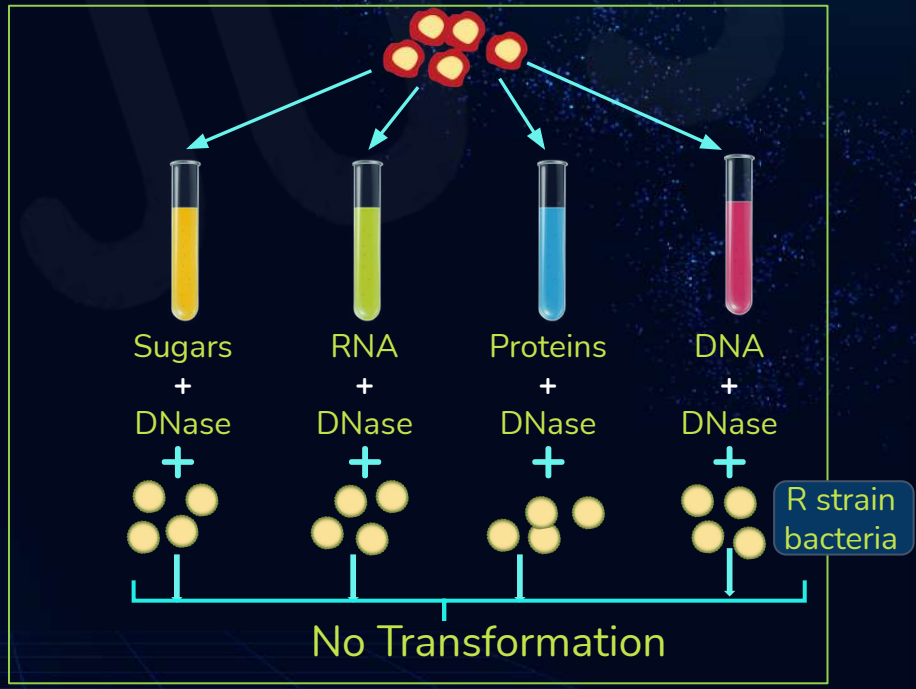

Summary

Avery, MacLeod and McCarty experiment

R strain got transformed to the S strain in the presence of DNA alone, and not when sugars, RNA or proteins were present.



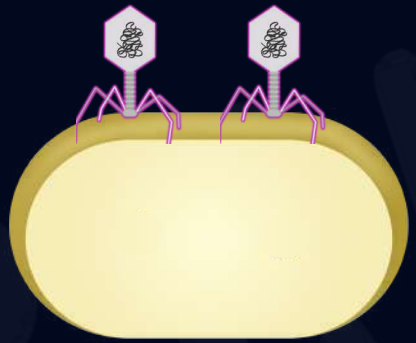
Upon addition of DNase, the transformation did not occur, suggesting that the transforming principle was indeed DNA.



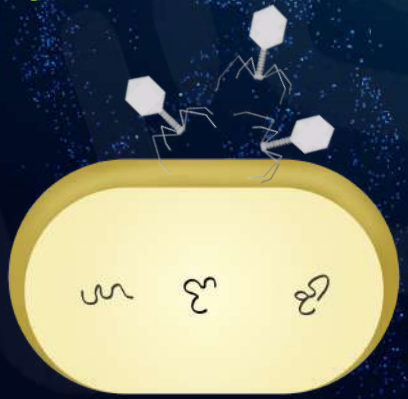
Summary

Hershey and Chase experiment

Phage with radioactive labelled protein coat

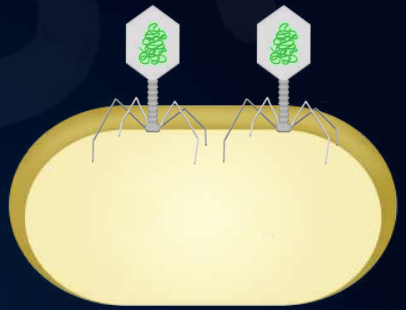


Phages infect bacteria

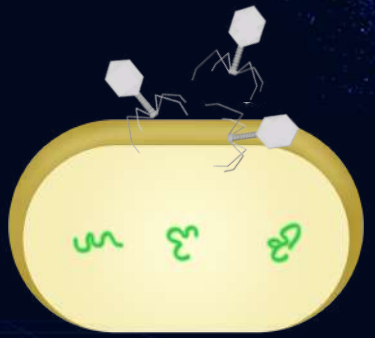


No radioactivity in the cell

Phage with radioactive labelled DNA



Phages infect bacteria



Radioactivity in the cell

Summary

Ideal genetic material

- It should be able to replicate.
- It should be chemically and structurally stable.
- It should provide scope for slow changes (mutation) required for evolution.
- It should be able to express itself in the form of 'Mendelian Characters'.

BYJU'S Classes Notes

Molecular Basis of Inheritance

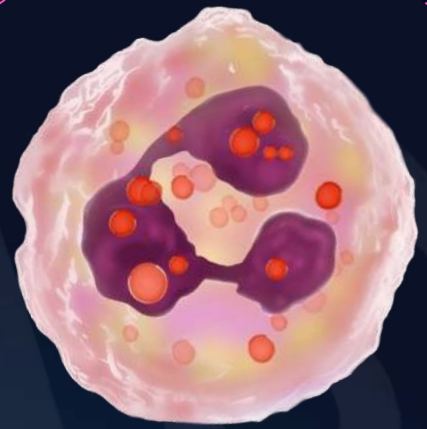
Discovery of Nucleic Acids, Double Helical Model of DNA and Types of DNA

Key Takeaways

1. Discovery of nucleic acids
2. Double helical model of DNA
3. Types of DNA

Summary

Recall! White Blood Cells

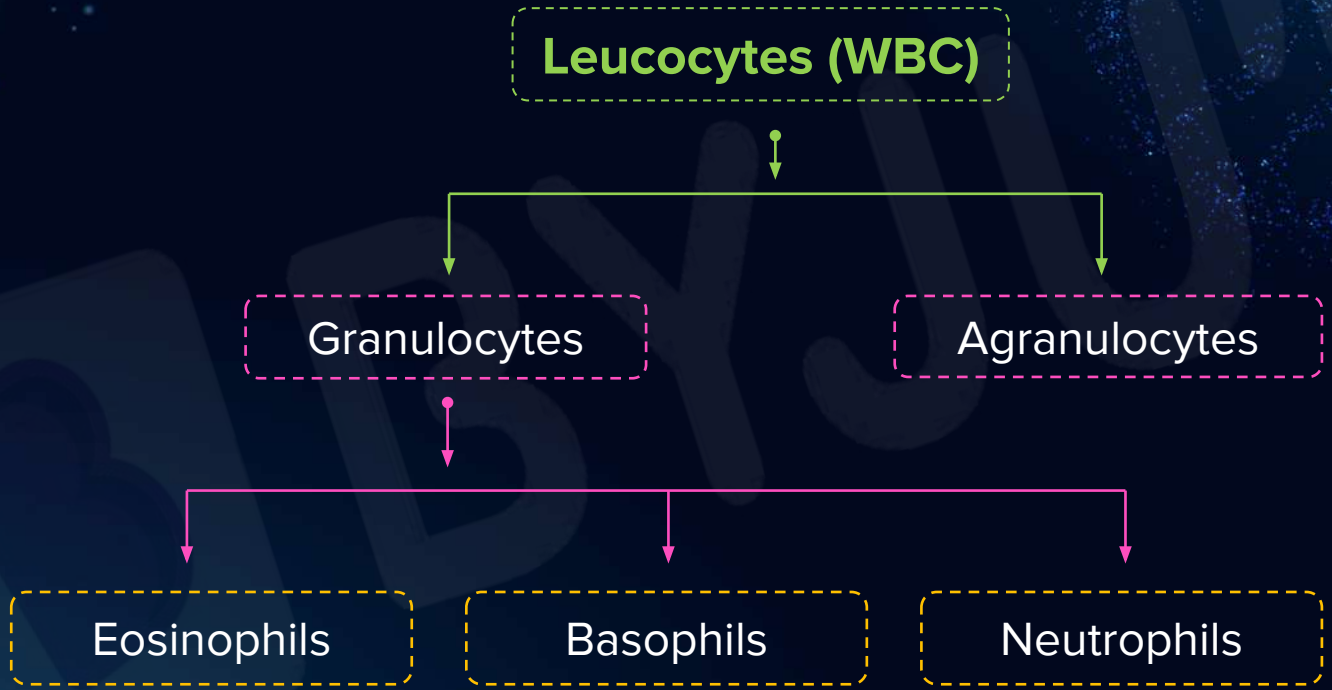


Nucleated

Roughly spherical

- Short lived cells
- Protects from infectious microorganisms

Recall! White Blood Cells



Recall! Nitrogenous Base

Nitrogenous bases

Nitrogen - containing compound with a ring structure

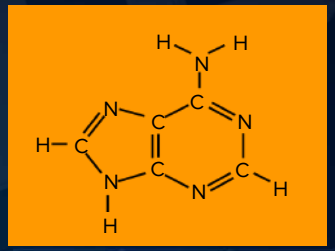
Purines

Purine base – have double ringed structure

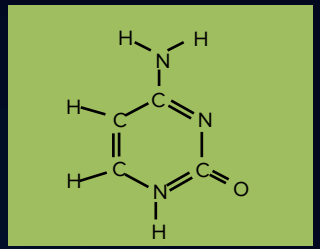
Pyrimidines

Pyrimidine base – have single ring structure

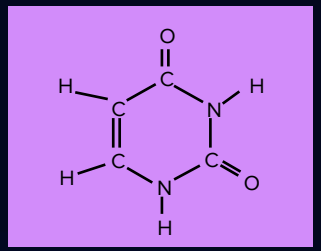
Guanine



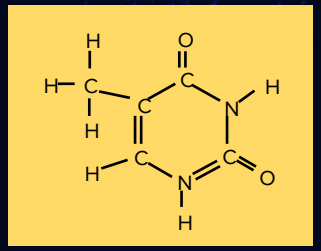
Adenine



Cytosine

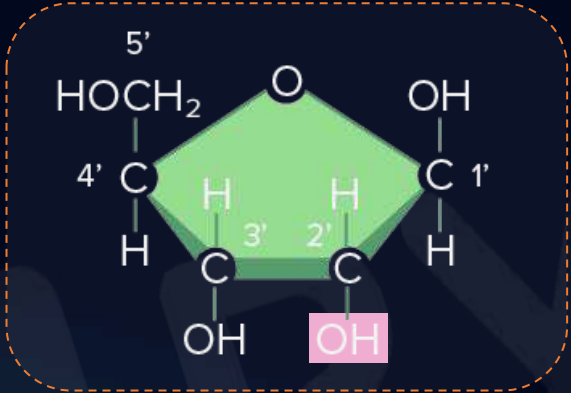


Uracil

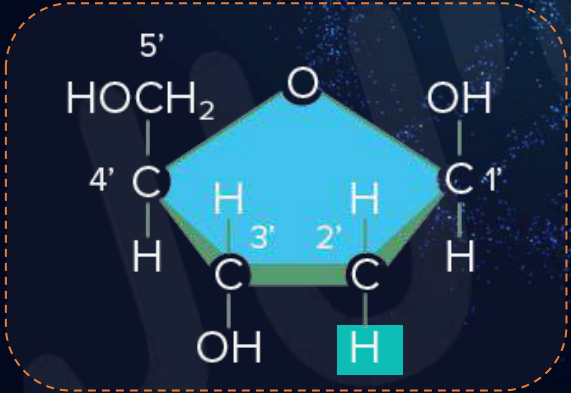


Thymine

Recall! Ribose - Pentose Sugar



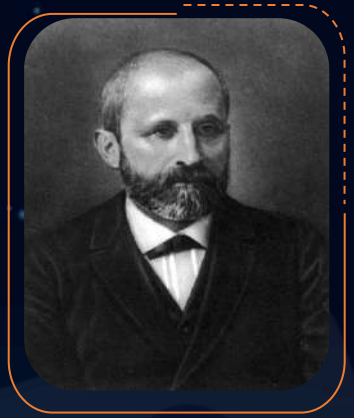
Ribose



Deoxyribose

- Ribose is a pentose sugar - 5 carbon sugar
- Ribose has -OH group at 2' position
- Deoxyribose has -H group at 2' position

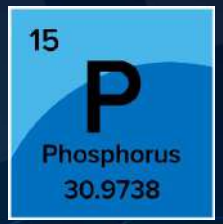
Discovery of Nucleic Acids



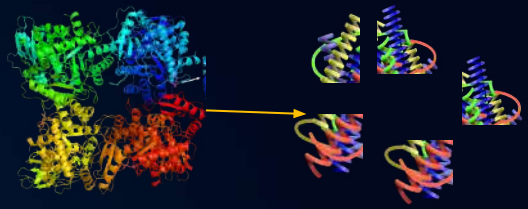
Friedrich Miescher

- Friedrich Miescher, a Swiss biologist in 1869, identified an acidic substance inside the nucleus of white blood cells.
- He accidentally found this compound when he was trying to isolate proteins.
- He saw that this new compound was neither a protein nor a lipid or carbohydrate.
- He also found that, it was acidic in nature, had high phosphorus content and was resistant to protein digestion.
- As it was isolated from nucleus, he gave the term nuclein.
- The term nuclein later became nucleic acids.

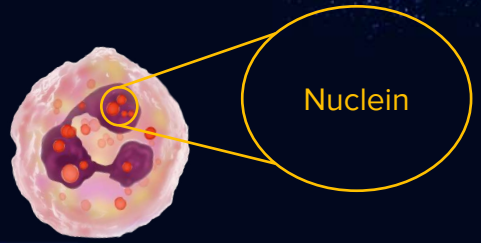
Acidic



High Phosphorus



Resistant to Protein digestion

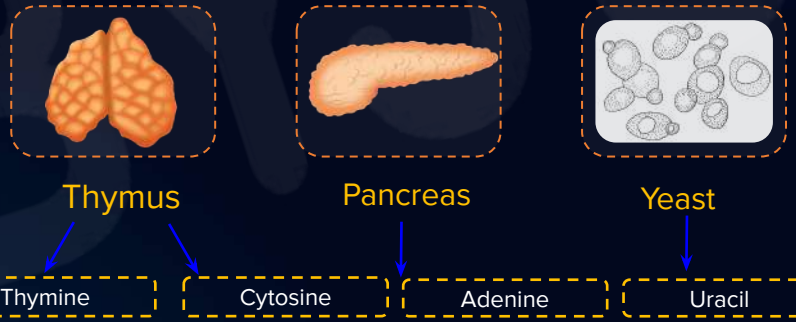


White Blood Cell

Discovery of Nucleic Acids

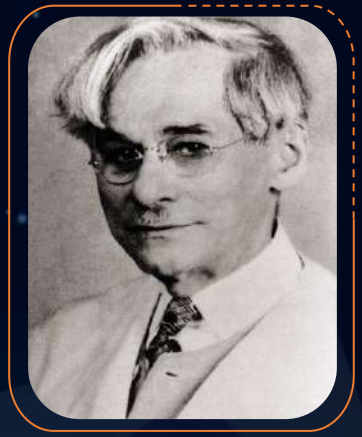
Kossel Albrecht

- Kossel Albrecht, a German scientist in 1910, found that nuclein is made of nitrogen containing bases
- He worked with animal organs and isolated the nucleotides from the same.
- He named them- thymine, cytosine, adenine, guanine and uracil.



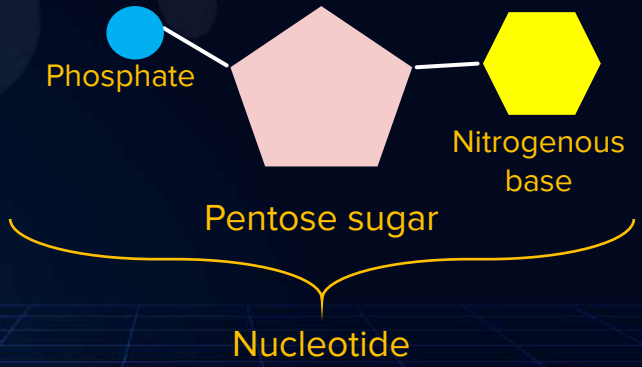
Albrecht Kossel, was awarded a Nobel Prize in 1910 for his work on discovery of nitrogenous base.

Discovery of Nucleic Acids



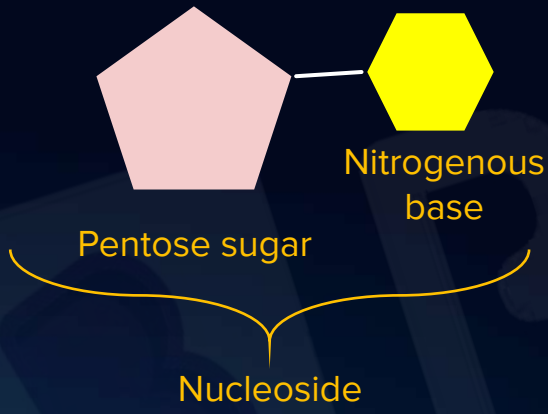
Phoebus Levene

- Phoebus Levene in 1919, was the first to come up with the three major components of nucleic acids i.e phosphate, sugar and nitrogenous base.
- He called this as nucleotide.
- He was the first to discover that the sugar component of RNA is ribose while that of DNA is deoxyribose.
- He proposed a tetranucleotide structure, in which the nucleotides were always linked in the same order.

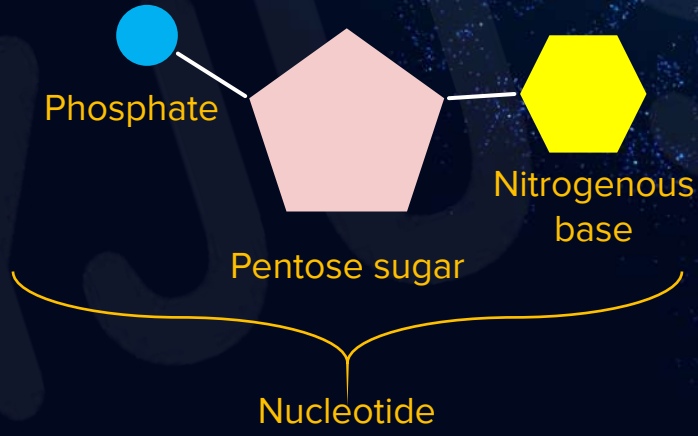


Discovery of Nucleic Acids

Nucleoside = Sugar + Nitrogenous base

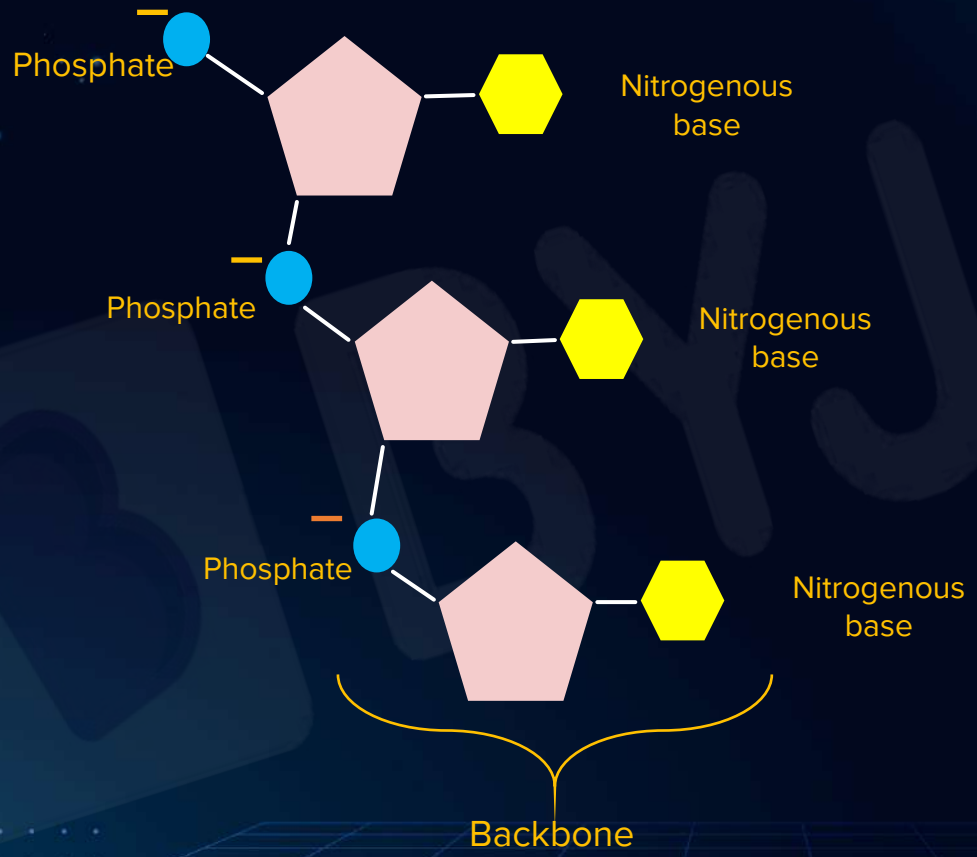


Nucleotide = Sugar + Nitrogenous base + phosphate



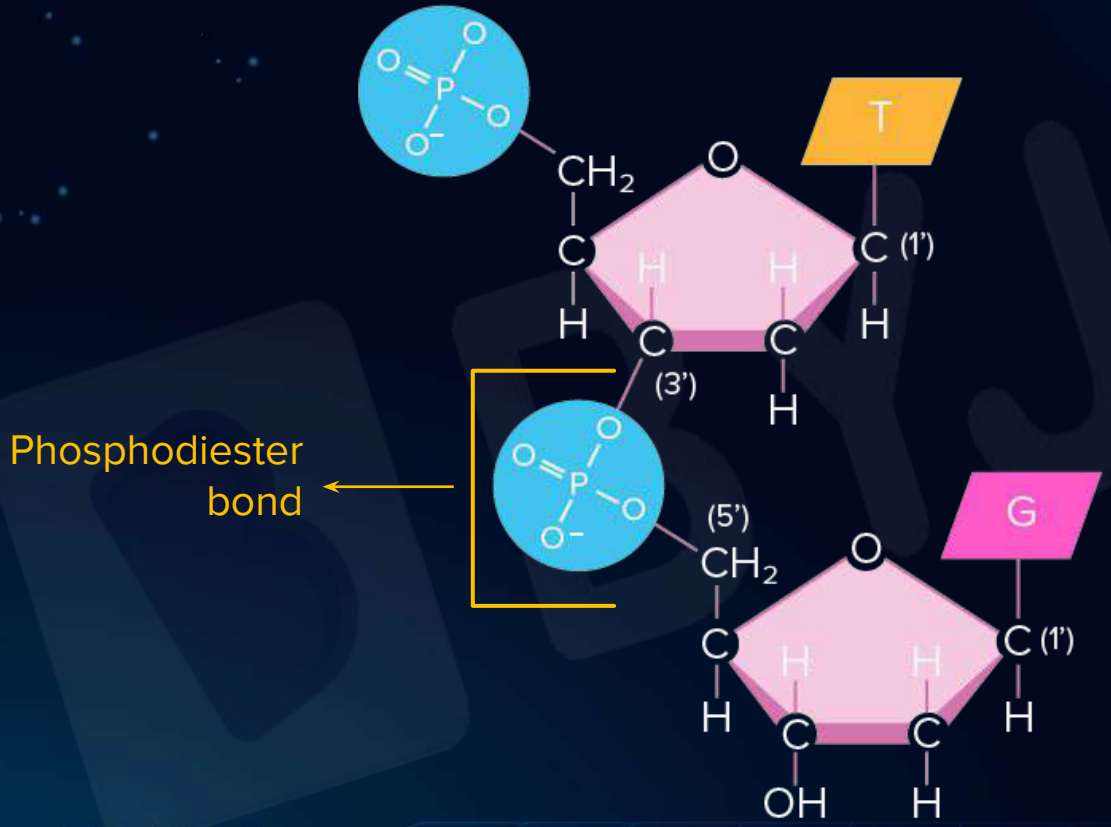
- Levene described nucleoside as nitrogenous bases with sugar.
- Note that, the difference between nucleoside and nucleotide is the presence of phosphate group in nucleotides.
- Levene also hypothesised that nucleic acids are chain of repeating units of these nucleotides.

Discovery of Nucleic Acids



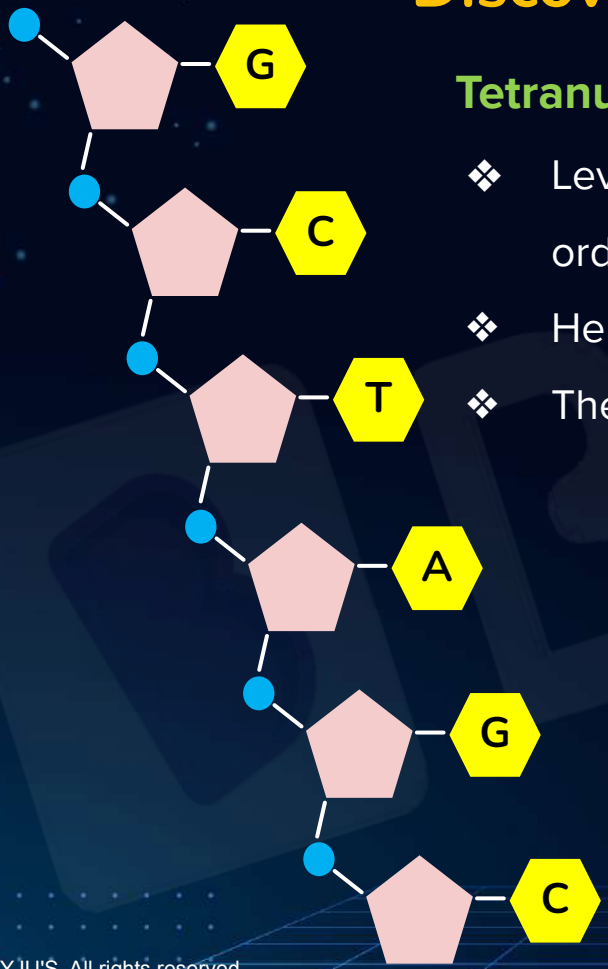
- These nucleotides are linked to one another through phosphate group.
- The phosphate backbone of DNA is negatively charged due to the bonds created between the phosphorous atoms and the oxygen atoms.
- Thus backbone of the nucleic acid is made of sugar - phosphate - sugar.
- DNA has negative charge, due to the negatively charged phosphate groups.

Discovery of Nucleic Acids



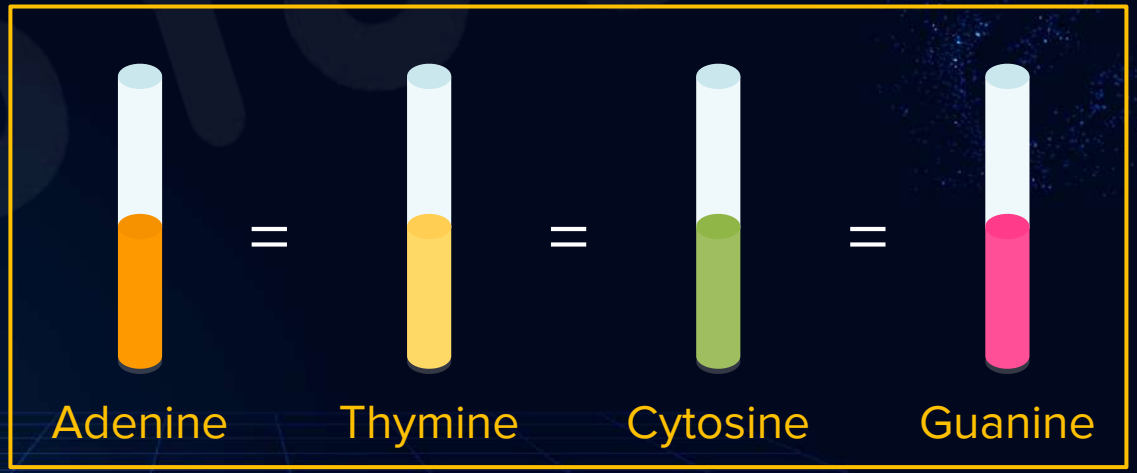
- Phosphate group links the 3'-carbon of one sugar of one nucleotide to the 5'-carbon of the sugar of the succeeding nucleotide through an ester bond.
- As there is one such ester bond on either side, it is called phosphodiester bond.

Discovery of Nucleic Acids



Tetranucleotide structure

- Levene proposed that nucleotides are always linked in the same order - (G-C-T-A-G-C-T-A).
- He also proposed that DNA has equal amounts of A, T, G, C.
- These propositions were later found to not be true.



Discovery of Nucleic Acids

Erwin Chargaff

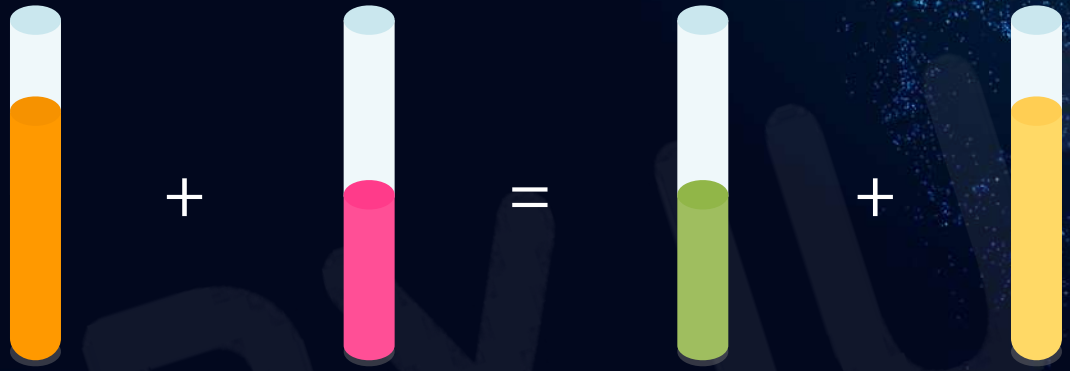
- Erwin Chargaff in 1948, noticed that
  - amount of adenine = amount of thymine
  - amount of cytosine = amount of guanine
- This was later known as Chargaff's rule.
- He proved that A+G = C+T
  - This meant number of purine bases is equal to number of pyrimidine bases.
- Hence, Chargaff's explanation proved Levene's tetranucleotide theory to be wrong.

Discovery of Nucleic Acids

Composition of **A**, **G**, **T**, **C** in DNA varies from one species to another

					
A	30.9	28.8	32.8	29.3	27.3
T	29.4	29.2	32.1	29.3	27.1
G	19.9	20.5	17.7	20.5	22.7
C	19.8	21.5	17.3	20.7	22.8

Discovery of Nucleic Acids



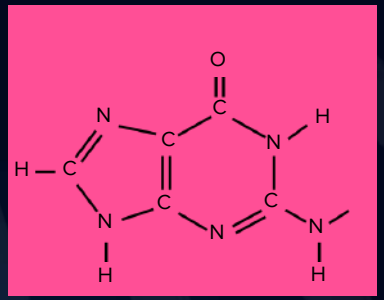
Purines Pyrimidines

Chargaff's Rule :

- In DNA, Adenine = Thymine; Cytosine = Guanine
- A + G = C + T
- Total number of Purines = Total number of Pyrimidines

Chargaff's rule applies only for DNA. it is not applicable for RNA.

Did you know ?



Guanine

- Guanine - It is the name based from where it had first been discovered.
- It was isolated from the excreta of sea birds called guano.

For double-stranded DNA, which of the following base ratios always equals 1?

a) (A+T) / (G+ C)

b) (A+ G) / (C+ T)

c) C / G

d) (G +T) / (A +C)

For double-stranded DNA, which of the following base ratios always equals 1?

a) (A+T) / (G+ C)

b) (A+ G) / (C+ T)

c) C / G

d) (G +T) / (A +C)

In sea urchin DNA, 17% of the bases were shown to be cytosine. The percentages of the other three bases in this DNA are?

a) G= 15%, A= 16.5%, T= 32.5%

b) G=15%, A= 35%, T= 35%

c) G= 8.5%, A= 50%, T= 24.5%

d) G= 34%, A= 24.5%, T= 24.5%

Solution

From the given information,
Cytosine= 15%,

According to the Chargaff's Rule,
A+T = G+C
A= T and C= G

Therefore, C= G= 15%

Total number of bases = 100%
i.e. A (%) + T (%) + G (%) + C (%) = 100%

Now C + G= 15 + 15 = 30 %
A + T + 30 = 100
A + T = 100 – 30 = 70
A + T = 70

Since, A = T

$\frac{70}{2}$ = 35 %

A = T = 35%

Therefore, the final composition is,

A= 35%, T= 35%, G= 15%, C= 15%

In sea urchin DNA, 17% of the bases were shown to be cytosine. The percentages of the other three bases in this DNA are?

a) G= 15%, A= 16.5%, T= 32.5%

b) G=15%, A= 35%, T= 35%

c) G= 8.5%, A= 50%, T= 24.5%

d) G= 34%, A= 24.5%, T= 24.5%

The percentage of cytosine in a double-stranded DNA is 17. What is the percentage of adenine in that DNA ?

a) 17

b) 45

c) 33

d) 66

Solution

From the given information,
Cytosine= 17%,

According to the Chargaff's Rule,
A+T = G+C
A= T and C= G

Therefore, C= G= 17%

Total number of bases = 100%
i.e. A (%) + T (%) + G (%) + C (%) = 100%

Now C + G= 17 + 17 = 34 %
A + T + 34 = 100
A + T = 100 – 34 = 66
A + T = 66

Since, A = T

$\frac{66}{2}$ = 33 %

A = T = 33%

Therefore, the final composition is,

A= 33%, T= 33%, G= 17%, C= 17%

The percentage of cytosine in a double-stranded DNA is 17. What is the percentage of adenine in that DNA ?

a) 17

b) 45

c) 33

d) 66

A segment of DNA has 120 adenine and 120 cytosine bases. The total number of nucleotides present in the segment is?

a) 120

b) 240

c) 60

d) 480

A segment of DNA has 120 adenine and 120 cytosine bases. The total number of nucleotides present in the segment is?

a) 120

b) 240

c) 60

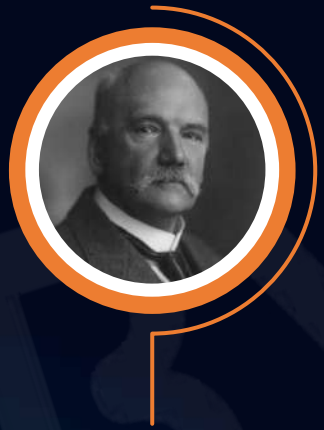
d) 480

The Double Helix Model

Friedrich Miescher

1869

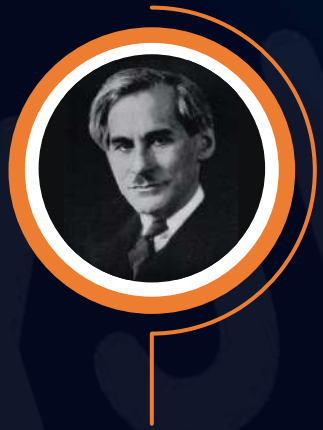
Miescher discovers nuclein (DNA) in the nuclei of white blood cells



Kossel Albrecht

Late 1800's

Kossel determines that DNA contains nitrogenous bases



Phoebus Levene

1910

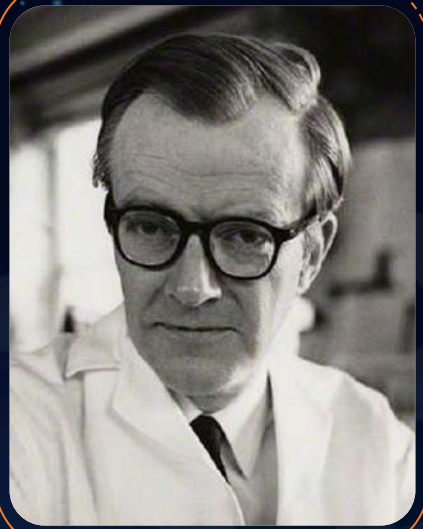
Levene proposes tetranucleotide theory

Erwin Chargaff

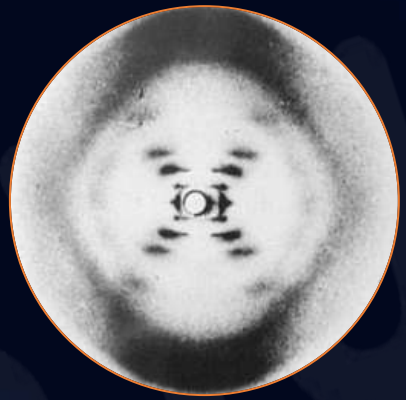
1948

Chargaff and colleagues discover regularity in base ratios of DNA

The Double Helix Model



Maurice Wilkins

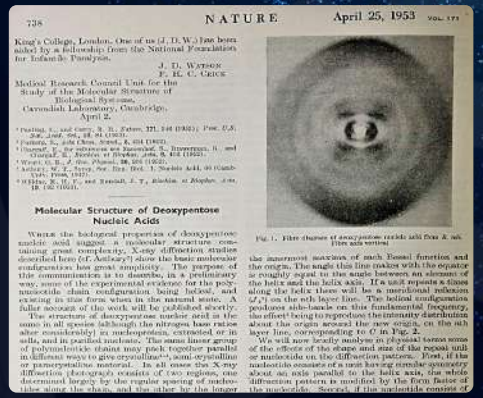


X- ray photo of DNA

Rosalind Franklin

- Scientists Maurice Wilkins and Rosalind Franklin studied the structure of DNA using X-ray diffraction.

The Double Helix Model



- Molecular building models - Watson and Crick, were able to investigate the structure by using all available information about the chemistry of DNA and construct molecular models.
- Watson and Crick tested various structures by building models made of wire and metal plates.
- With their models, they were able to see whether a structure was compatible with chemical principles and with the X-ray images.
- They finally came up with the 3-D structure of the DNA, which they called the 'The Double-Helix model' also known as 'Watson and Crick Model'.

The Double Helix Model

In 1962, Francis Crick, James Watson and Maurice Wilkins were awarded with the Nobel Prize for their discovery of the DNA structure

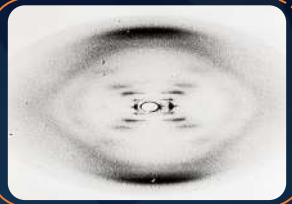
Francis Crick

James Watson

Maurice Wilkins

Did you know ?

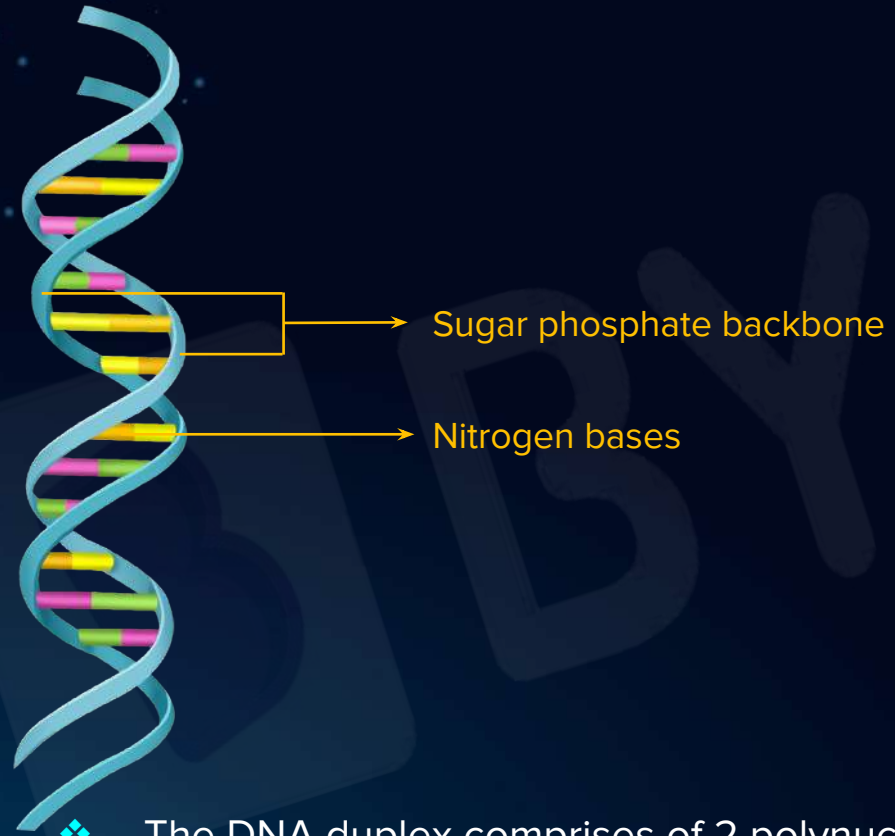
Rosalind Franklin

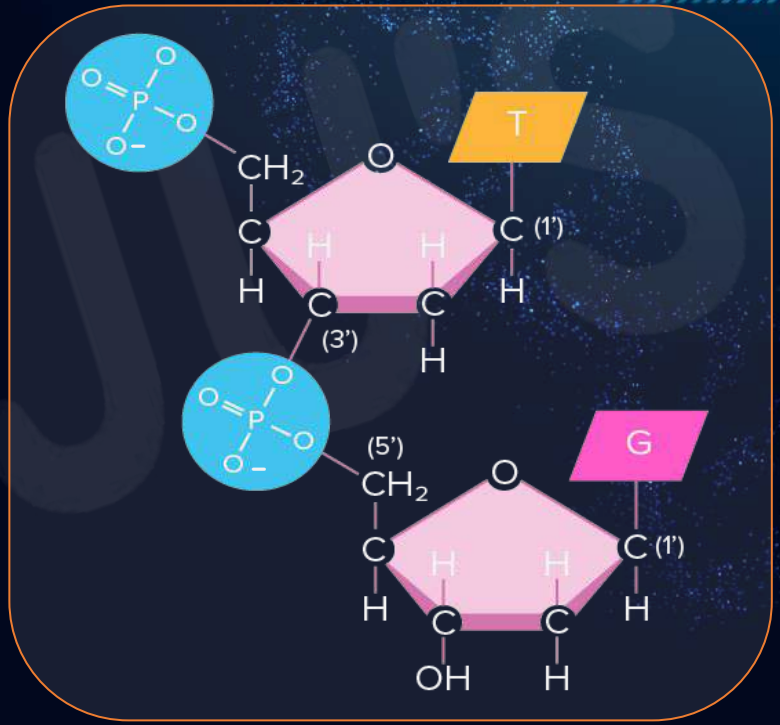


X- ray photo of DNA (Photo 51)

- Rosalind Franklin is the only woman involved in solving the DNA mystery .
- The photo shown is famously known as Photo 51, which was taken by Rosalind Franklin.
- This was an important contribution towards unravelling the structure of DNA.
- Sadly, she died because of ovarian cancer in 1957 and could not be considered for the Nobel prize, as the Nobel committee does not consider awarding candidates after death, for the award.
- However, her contribution in deciphering the DNA structure is significant and praise-worthy.

The Double Helix Model

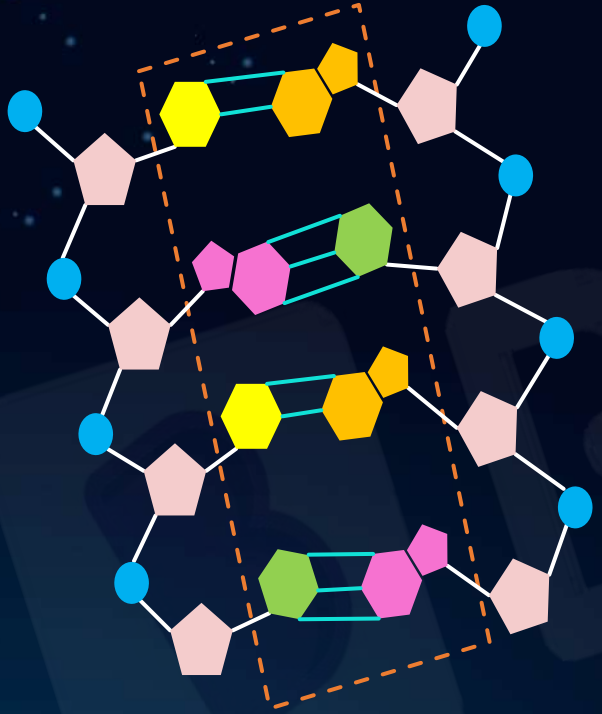




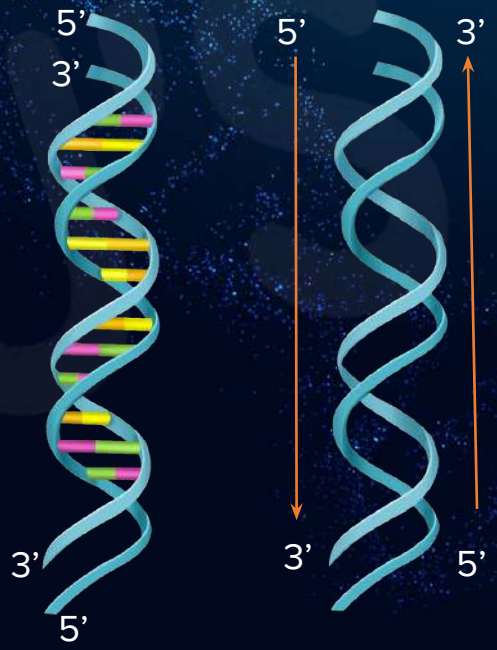
Sugar phosphate chains as backbone

- The DNA duplex comprises of 2 polynucleotide chains with a sugar phosphate backbone and nitrogenous bases projecting inside.

The Double Helix Model

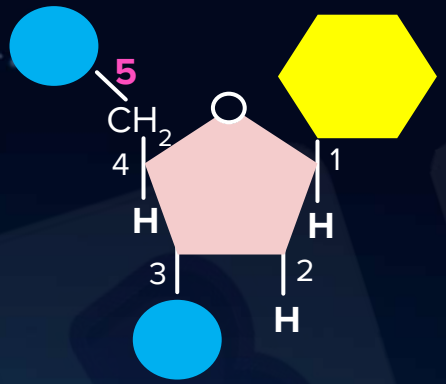


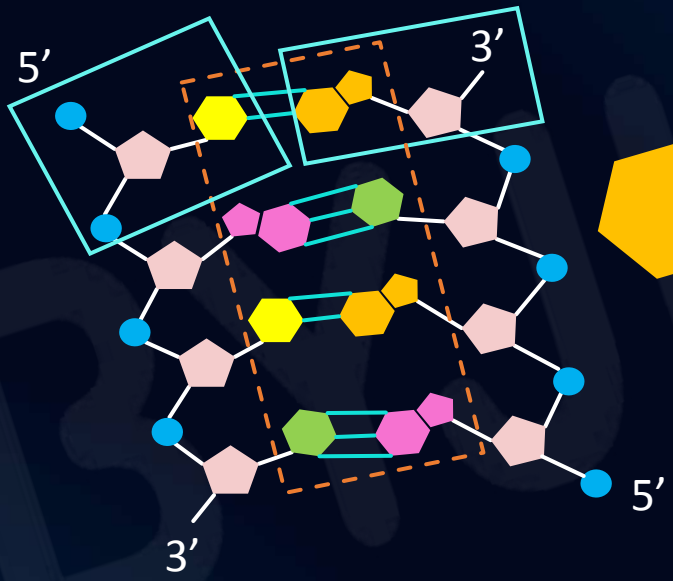
- The 2 chains are coiled in right - handed fashion.

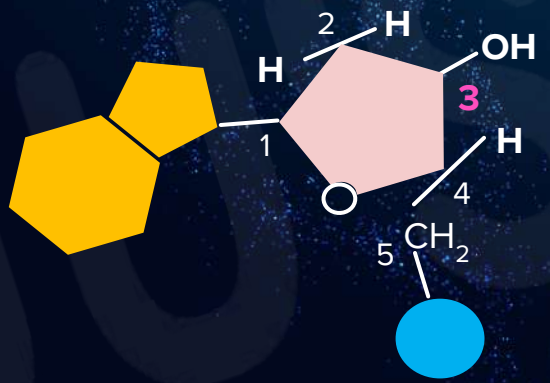


- The nitrogenous bases project inside the sugar phosphate backbone.

The Double Helix Model

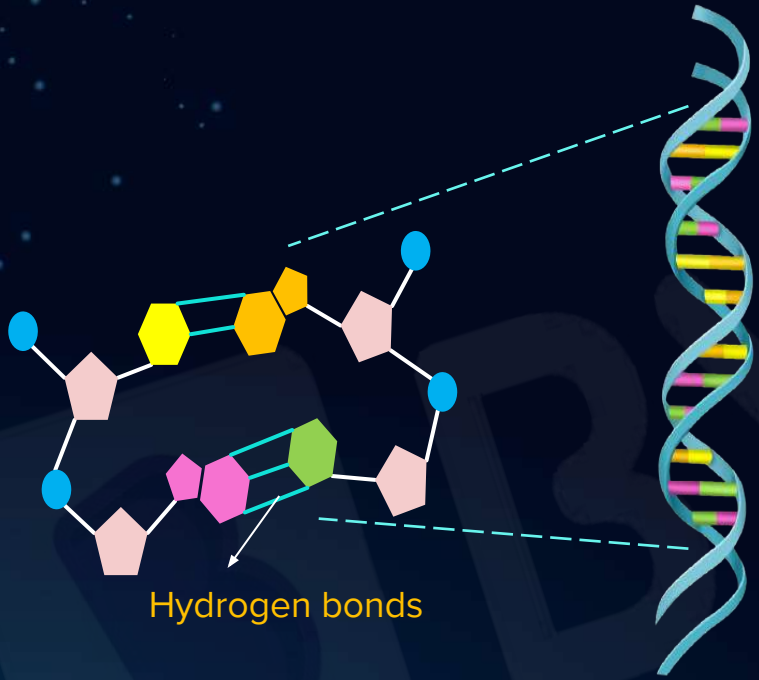






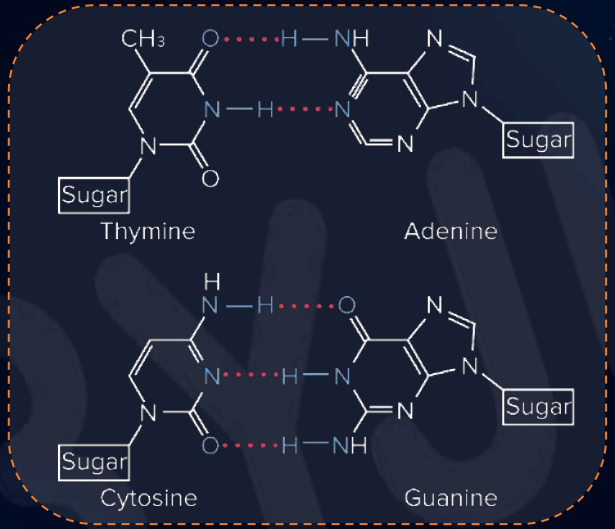
- The polynucleotide has, at one end, a free phosphate moiety at 5' -end of sugar, which is referred to as 5’-end of polynucleotide chain.
- Similarly, at the other end of the polymer the sugar has a free OH of 3'C group which is referred to as 3' -end of the polynucleotide chain.
- Since both the ends run opposite to each other, they are called as antiparallel in polarity.

The Double Helix Model



- The bases of the 2 strands are paired through hydrogen bonds forming base pairs
- A purine always pairs with a pyrimidine

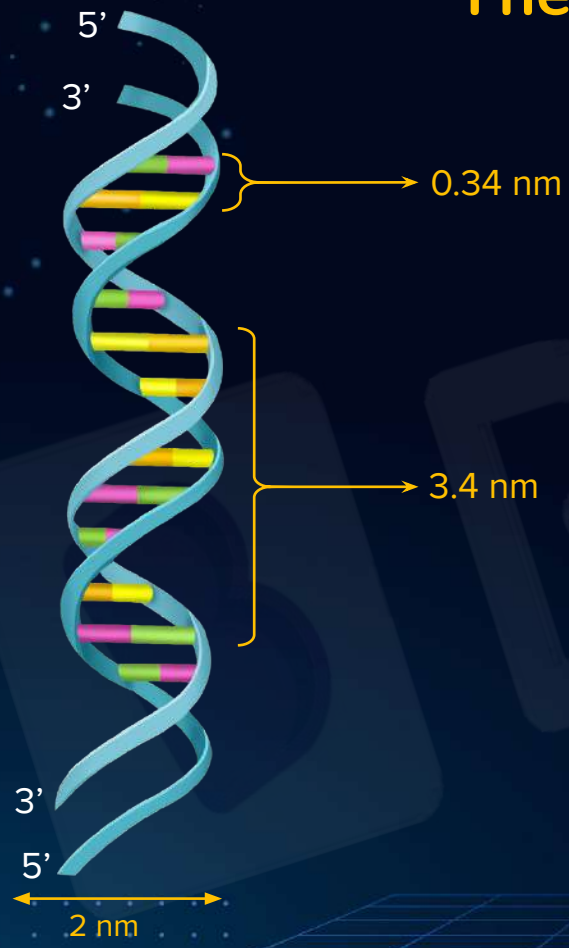
The Double Helix Model



- Adenine pairs with Thymine through 2 hydrogen bonds
- Cytosine pairs with Guanine through 3 hydrogen bonds,
- Hence, C-G is stronger i.e. difficult to break as compared to A-T.

C ≡ G is stronger than A = T

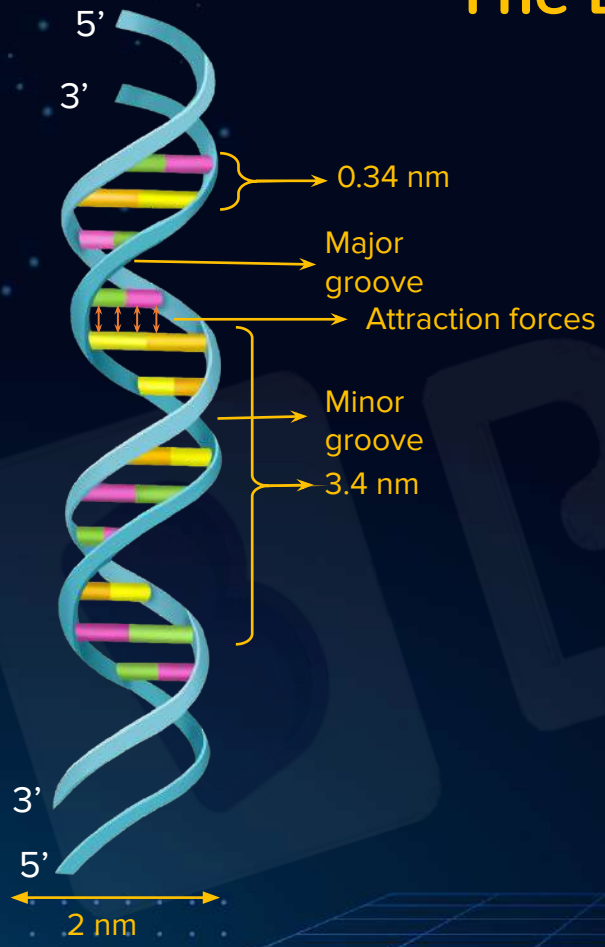
The Double Helix Model



Features

- Helix pitch is 3.4nm (height of one complete helix turn)
- Roughly 10bps in each turn
- Helix diameter is roughly 2nm
- Distance between the bps in the helix (also called helical rise) is 0.34nm

The Double Helix Model



Features

- Major groove occurs where the backbones are far apart.
- Minor groove occurs where they are close together.
- The grooves twist around the molecule on opposite sides
- The plane of one base pair stacks over the other in double helix
- Attractive forces hold them together, adding to stability

Types of DNA

A - DNA

- Dehydrated form of B-DNA
- Right handed helix
- It has 11 base pairs
- Similar to B-DNA in structure and orientation

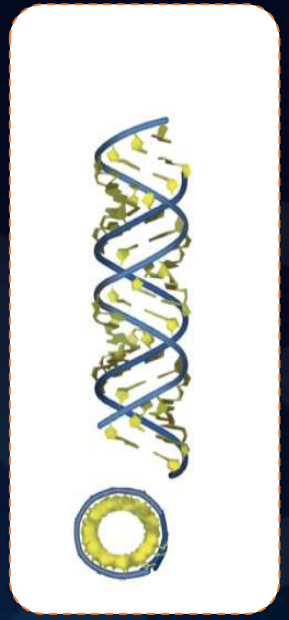
B - DNA

- Most common and predominant type
- Right handed helix
- Bases nearly perpendicular to the helical axis
- It has 10 base pairs
- Believed to play some role in gene regulation

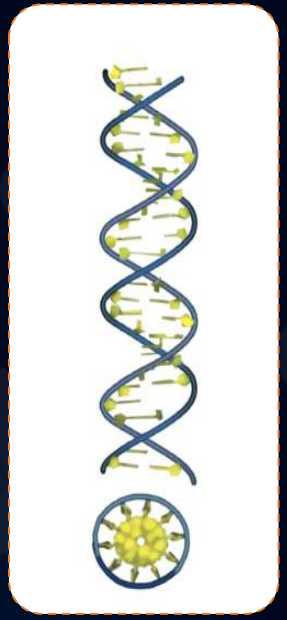
Z - DNA

- Found in very less amounts
- Left handed helix
- It has 12 base pairs
- Bases are arranged in zig-zag manner

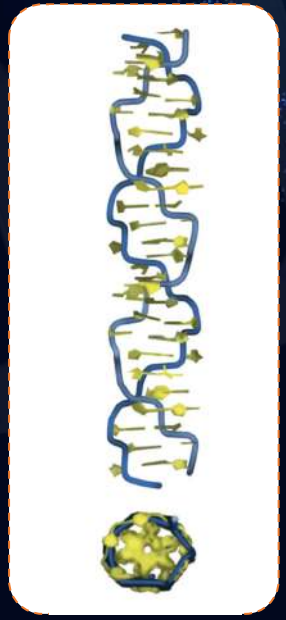
Types of DNA



A - DNA



B - DNA



Z - DNA

3 types of DNA

Characteristics	A - DNA	B - DNA	Z - DNA
Helix type	Right handed	Right handed	Left handed
Helical pitch	2.86 nm	3.4 nm	4.4 nm
Helical rise	0.29 nm	0.34 nm	0.74 nm
Helix diameter	2 nm	2.6 nm	1.8 nm
No of bps per turn	~ 11	~ 10	~ 12
Major groove	Narrow and deep	Wide and deep	Flat
Minor groove	Wide and shallow	Narrow and deep	Narrow and deep

Which one of the following is not applicable to RNA?

a) Heterocyclic Nitrogenous Bases

b) Chargaff's Rule

c) Complementary Base Pairing

d) 5'-phosphoryl and 3'-hydroxyl ends

Which one of the following is not applicable to DNA?

a) Heterocyclic Nitrogenous Bases

b) **Chargaff's Rule**

c) Complementary Base Pairing

d) 5'-phosphoryl and 3'-hydroxyl ends

Summary

Solving the DNA mystery

Friedrich Miescher

1869
Miescher discovers nuclein (DNA) in the nuclei of white blood cells

Kossel Albrecht

Late 1800's
Kossel determines that DNA contains nitrogenous bases

Phoebus Levene

1910
Levene proposes Tetranucleotide theory

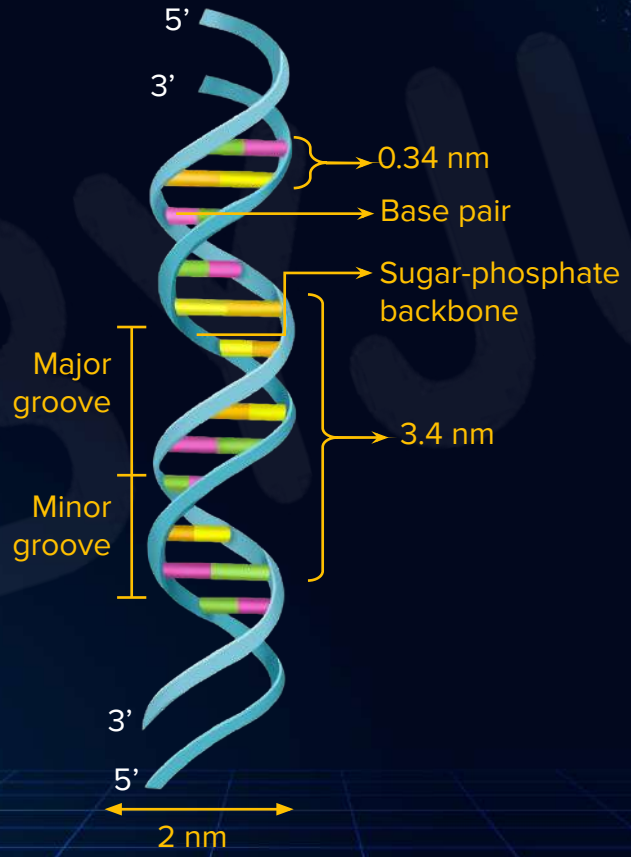
Erwin Chargaff

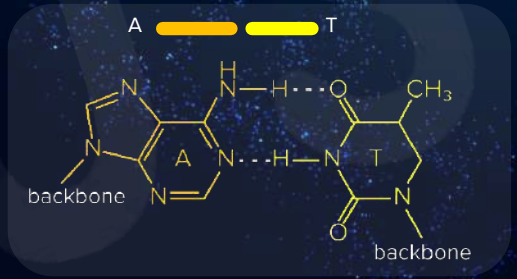
1948
Chargaff and colleagues discover regularity in base ratios of DNA

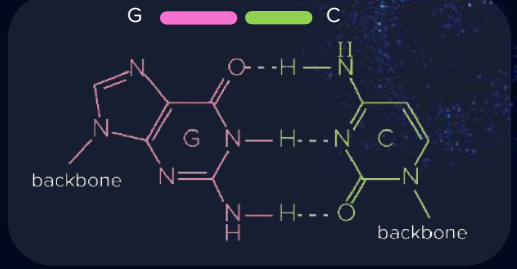
Summary

Solving the DNA mystery

James Watson and Francis Crick

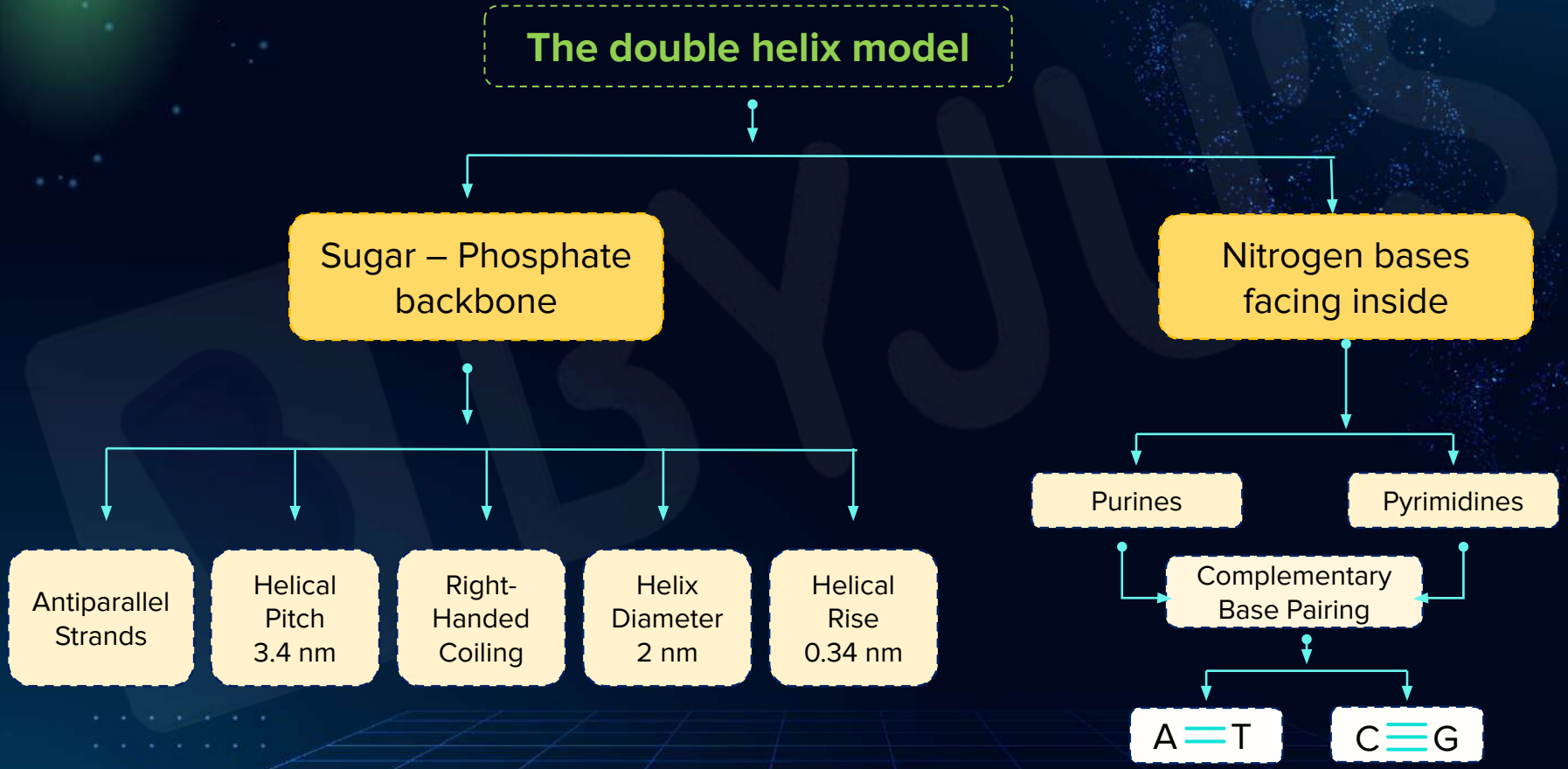






Summary

Summary



BYJU'S Classes Notes

Molecular Basis of Inheritance

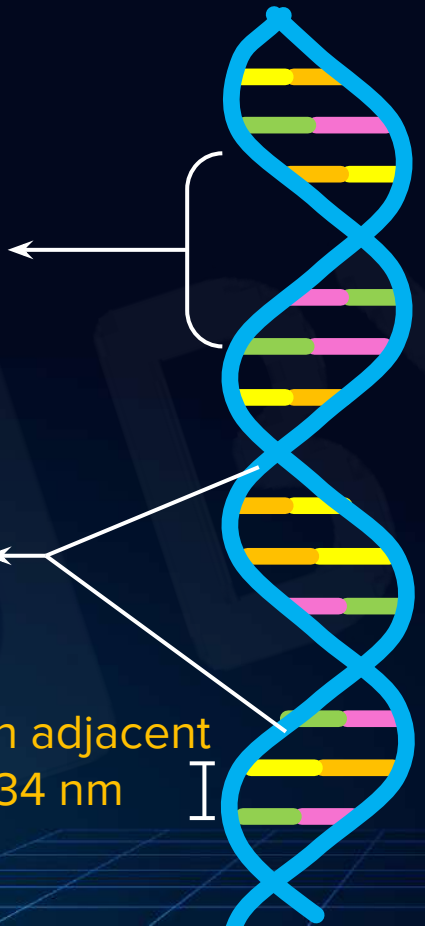
Packaging of DNA, Types of Chromatin, Characteristics of DNA, RNA World Hypothesis

Key Takeaways

1. Packaging of DNA
2. Types of chromatin
3. Key characteristics of RNA
4. DNA vs RNA

Summary

Recall! DNA



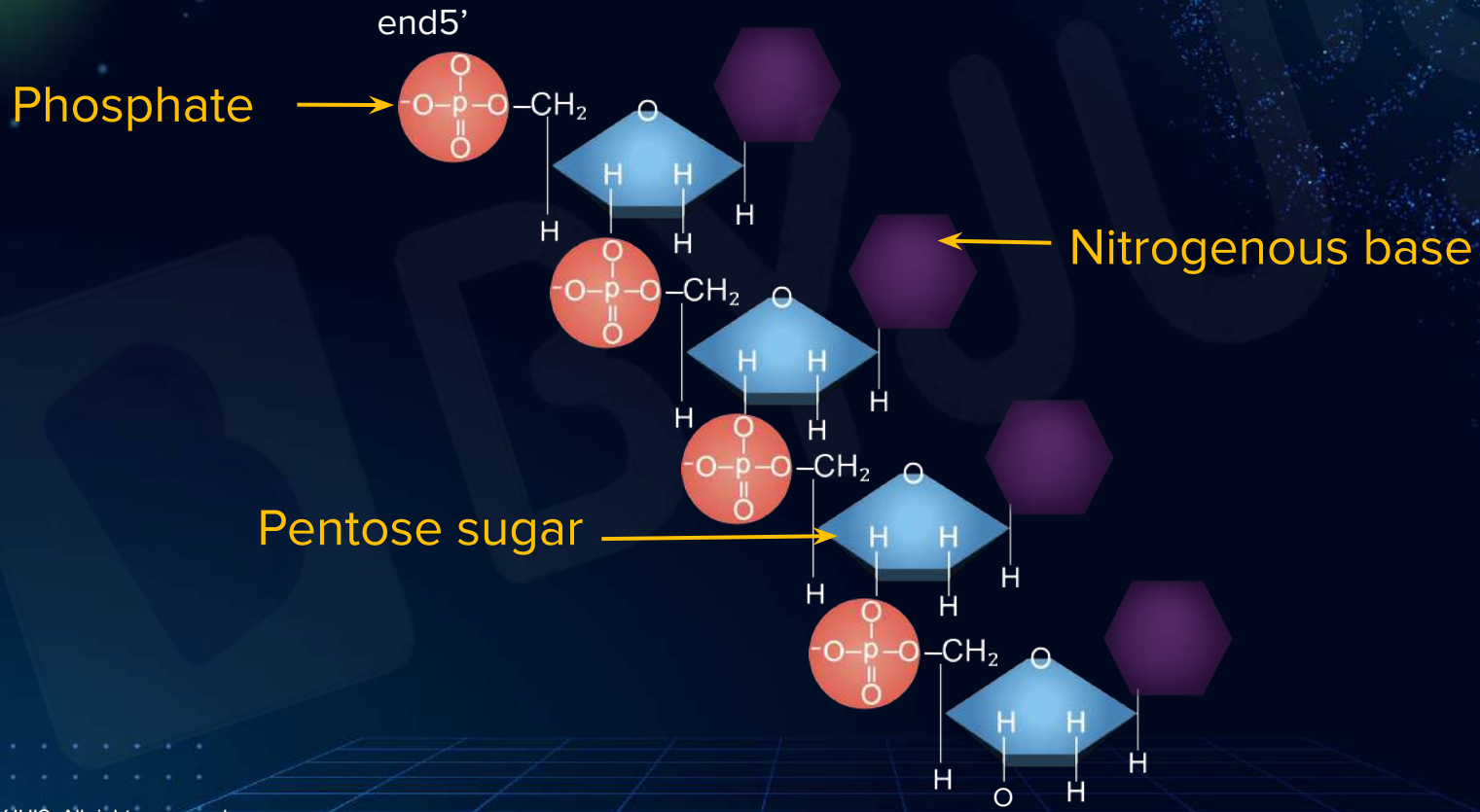
Plectonemic

- DNA is plectonemic in nature.
  - The two strands of DNA cannot be separated without completely unwinding them.
- The two chains of DNA have antiparallel polarity.
- The whole double strand is spirally coiled upon itself around a common axis in a right handed manner.

Recall! DNA

- Pitch (height of one complete helix turn) is 3.4 nm
- One turn has 10 base pairs.
- Hence the distance between adjacent base pairs is 34 / 10 = 0.34 nm

Recall! DNA Duplex

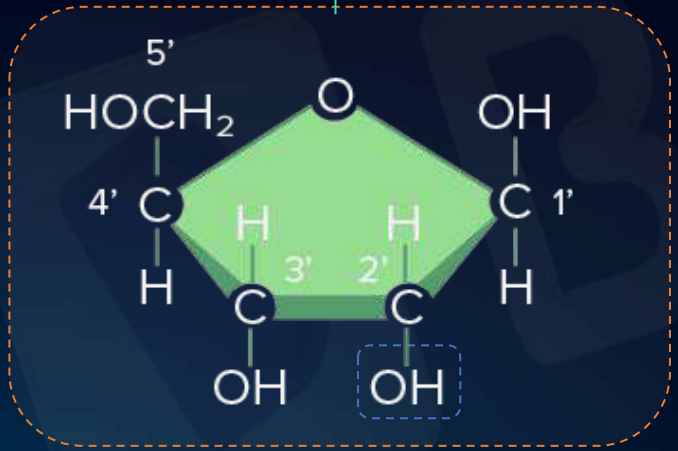


Recall! DNA Duplex

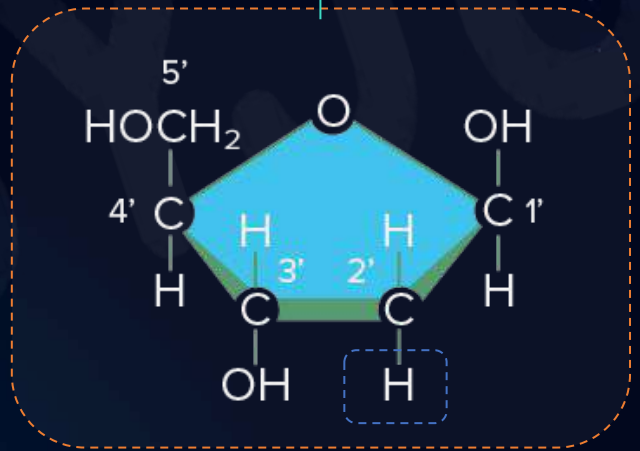
Pentose sugar

Pentose sugar

Ribose sugar of DNA has hydrogen atom at 2' carbon whereas in RNA the hydroxyl group is present at 2' carbon.

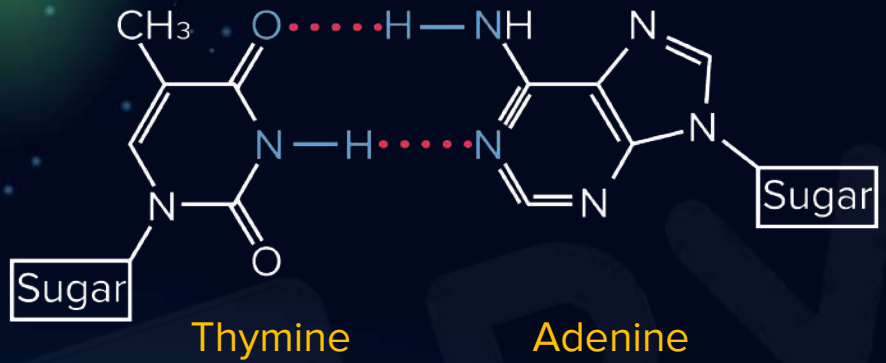


Ribose

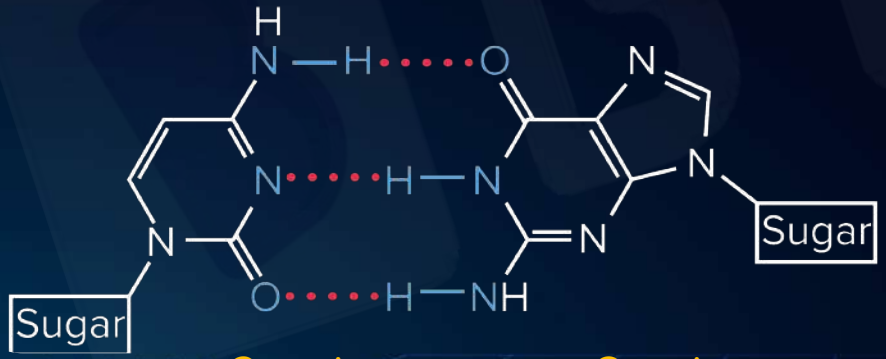


Deoxyribose

Recall! DNA Duplex



Thymine Adenine



Cytosine Guanine

- The nitrogenous bases (ATGC) of DNA strands are paired through hydrogen bond (H-bonds) forming base pairs (bp).
- Purines (adenine and guanine) form hydrogen bond with pyrimidines (thymine and cytosine).
- Thymine pairs with adenine and guanine pairs with cytosine.

DNA of Prokaryotes

Base pairs in a cell = 4.6×10^6

Distance between adjacent base pairs = 0.34 nm

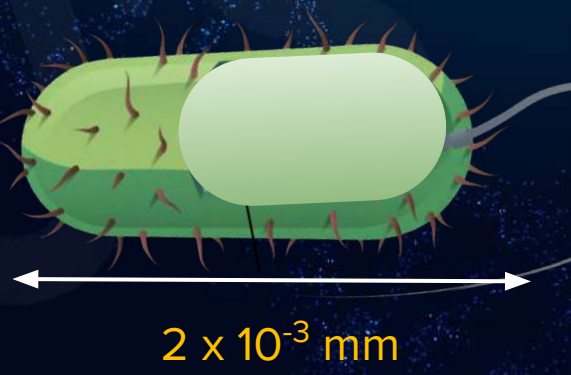
Total length of DNA = $4.7 \times 10^6 \times 0.34 \times 10^{-9}$ m

= 1.6×10^{-3}m

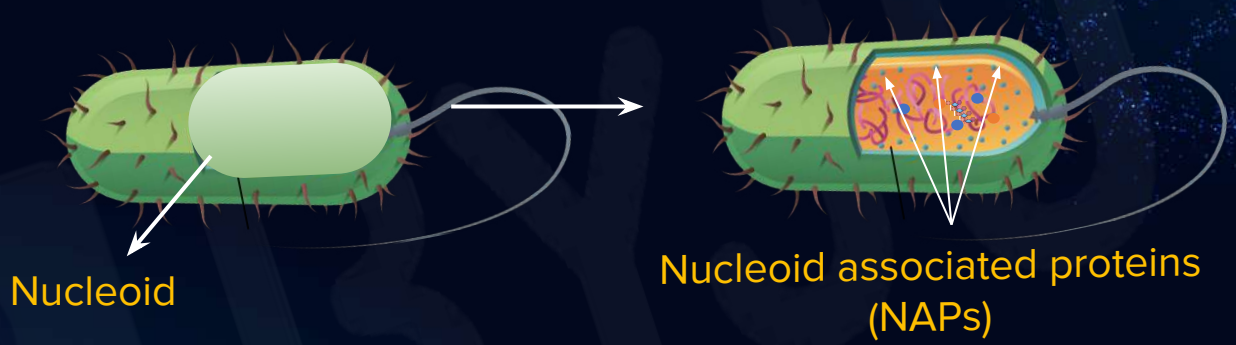
= 1.6 mm

DNA of Prokaryotes

- The prokaryotic DNA is not enclosed within a membrane.
- The size of the cell is between 1 to 2 micrometers.
- Length of the DNA is 1.6 mm.
- The DNA is 80 times the length of the bacteria.



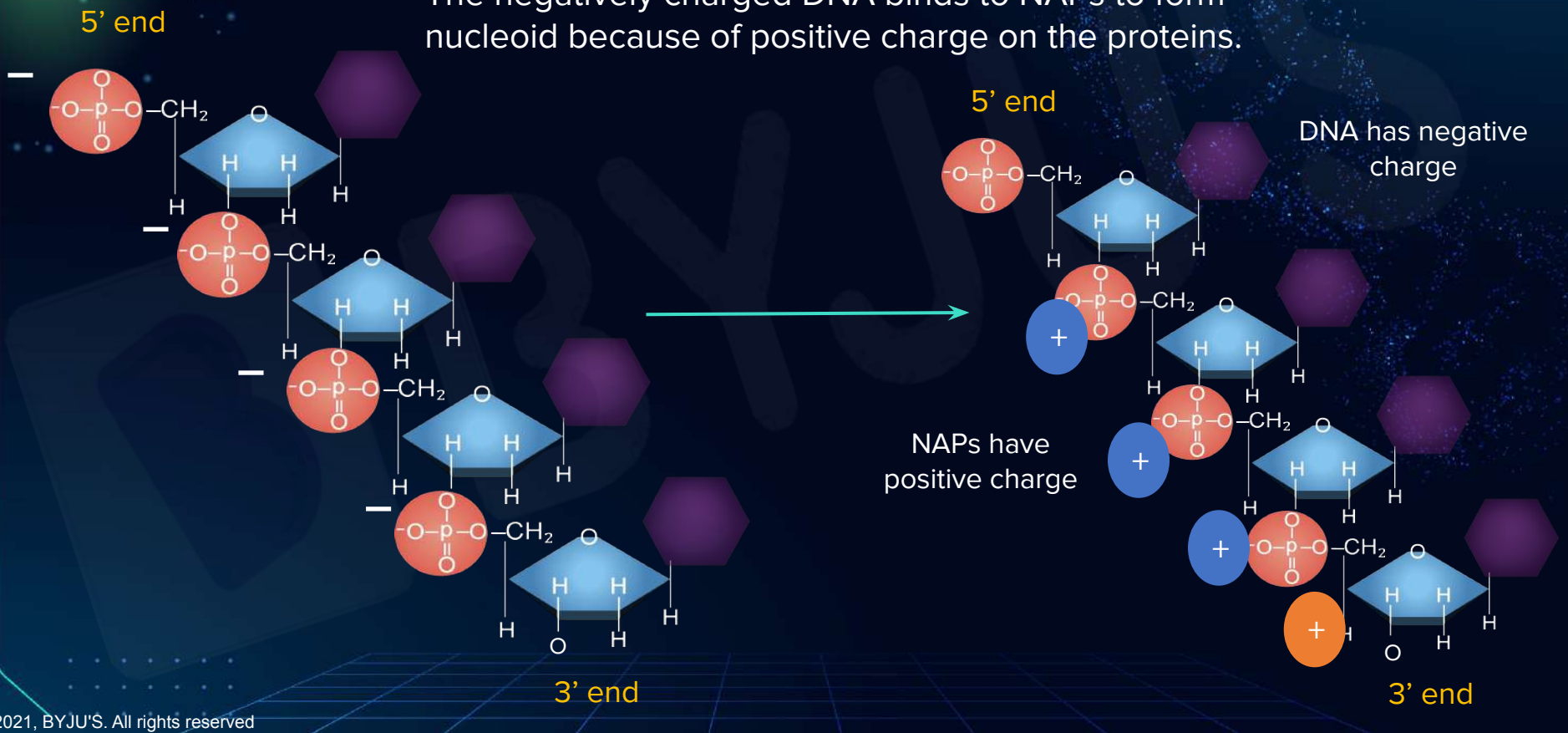
DNA Packaging in Prokaryotes



- Genomic DNA in prokaryotes is organized in large loops held by special proteins called NAPs.
- This entire structure constitutes the nucleoid.

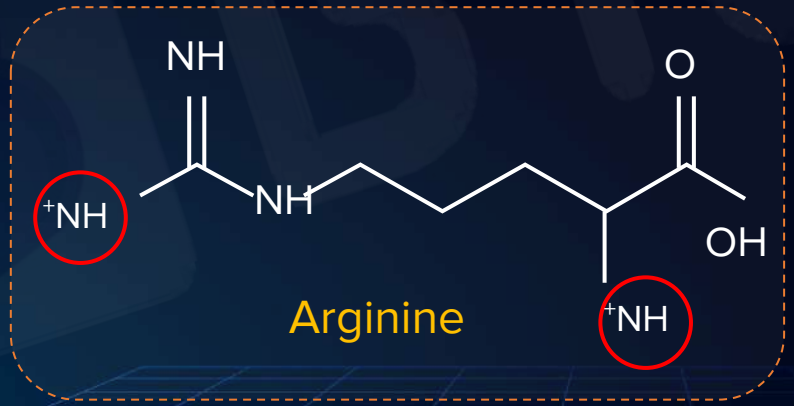
Recall! DNA Charge

The negatively charged DNA binds to NAPs to form nucleoid because of positive charge on the proteins.



DNA Packaging in Prokaryotes

NAPs have positive charge



- NAPs are rich in the basic amino acid residues lysine and arginine.
- Both the amino acid residues carry positive charges in their side chains. Hence, NAPs are positively charged.
- Prokaryotic chromosomes fit the DNA in the cell by the supercoiling of DNA which are held together by positively charged NAPs.

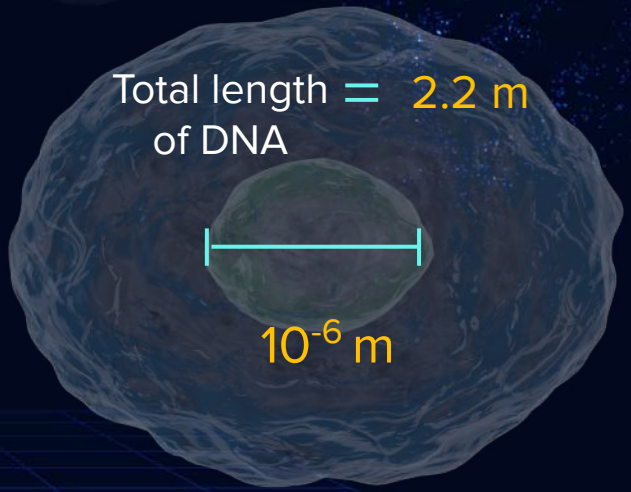
DNA of Eukaryotes

Base pairs in a cell = 6.6×10^6

Distance between adjacent base pairs = 0.34 nm

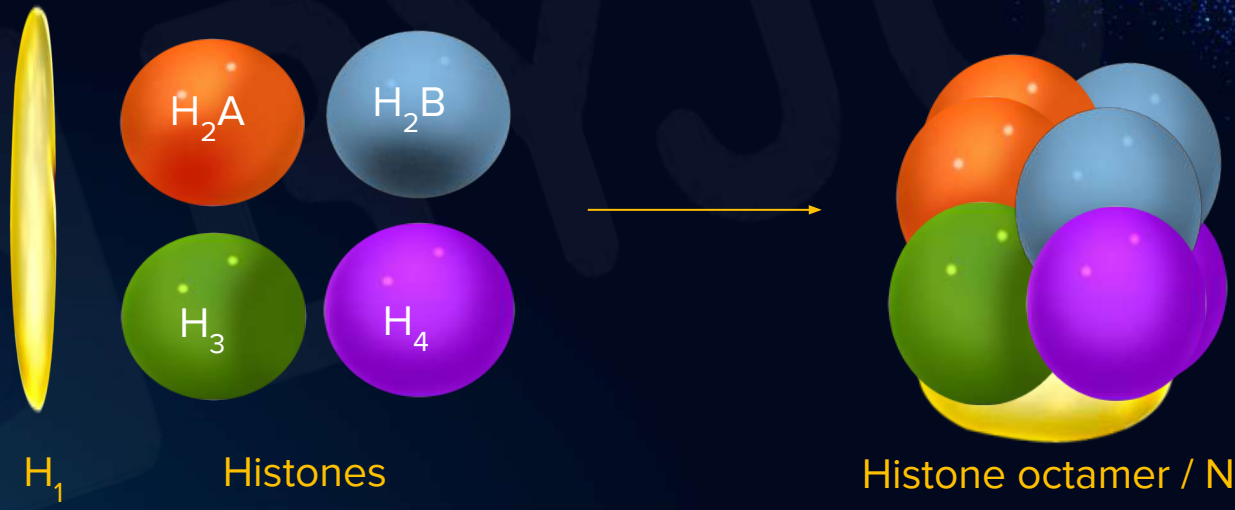
Total length of DNA = $6.6 \times 10^6 \times 0.34 \times 10^{-9}$ m

= 2.24 m



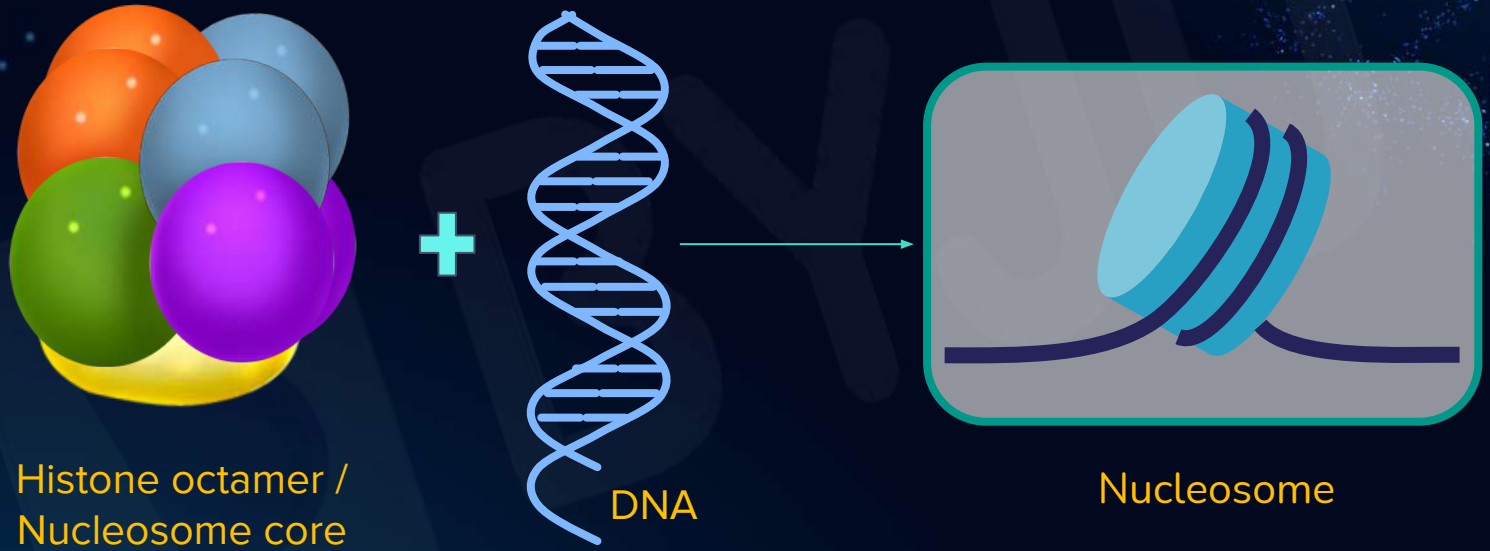
DNA Packaging in Eukaryotes

- In eukaryotes, the proteins involved in packaging are called histones.
- Histones are organised to form an unit of eight molecules called histone octamer.
- H_2A, H_2B, H_3, H_4 occur in pairs.



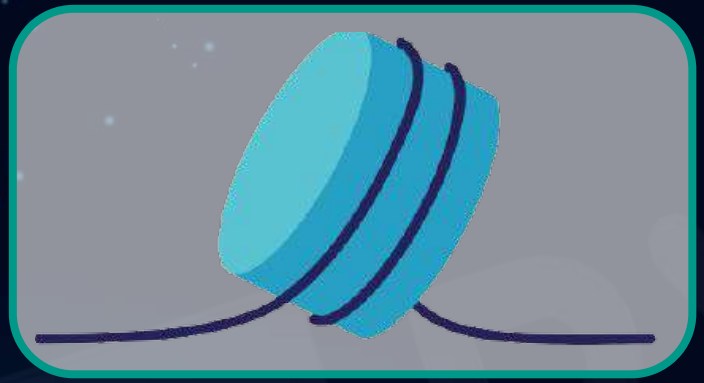
DNA Packaging in Eukaryotes

DNA + histone octamer = nucleosome

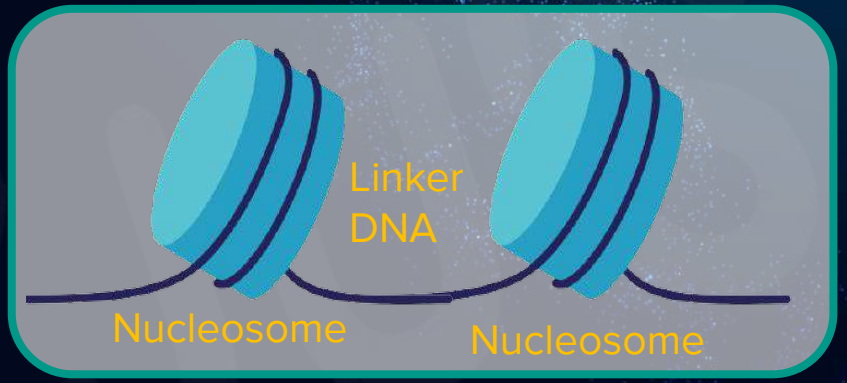


Histone octamer binds and wraps approximately 1.7 turns of DNA/ 200 base pairs of DNA.

DNA Packaging in Eukaryotes



Nucleosome

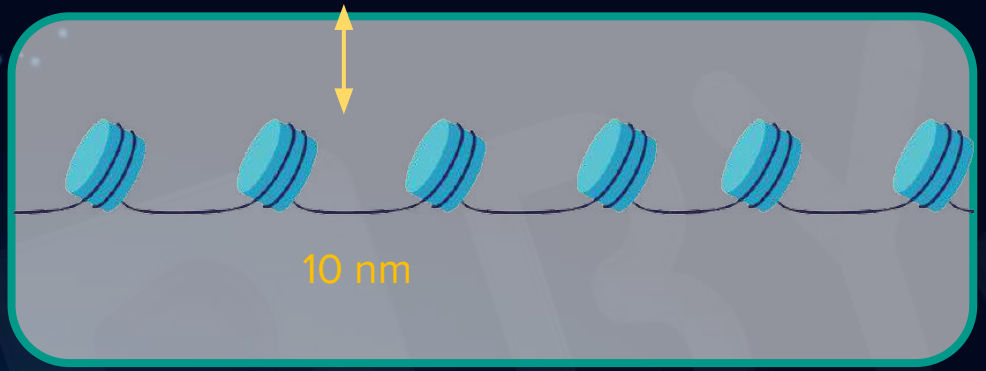


Chromatosome

- It is stable due to hydrogen bonding between:
  - -OH groups in amino acids of histone peptide
  - phosphodiester backbone of DNA

- Small segment of DNA connecting two adjacent nucleosomes is called interbead or linker DNA
  - The H_1 protein of the histone octamer is attached over the linker DNA.
  - Nucleosome and linker DNA together constitute the chromatosome.

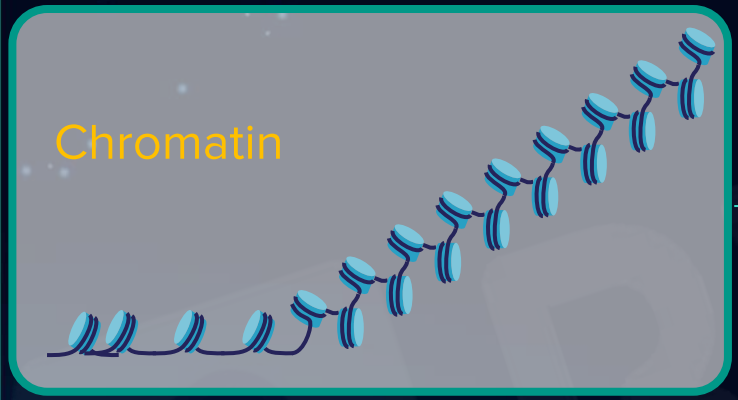
DNA Packaging in Eukaryotes

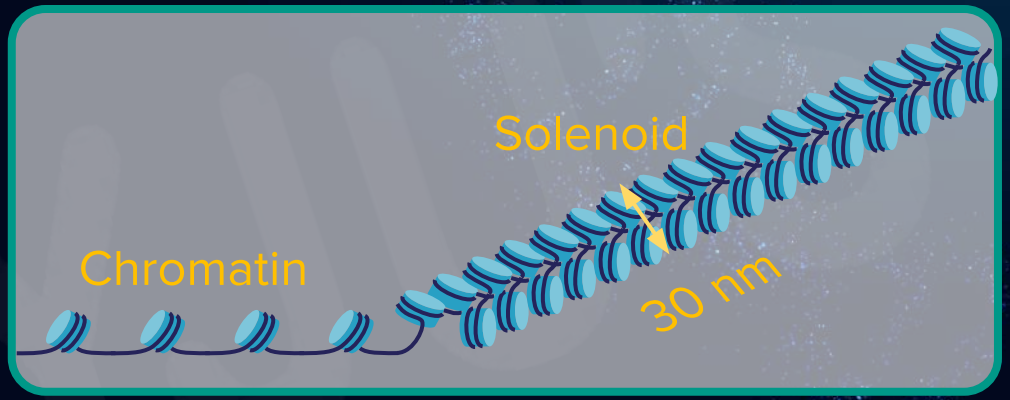


Chromatin

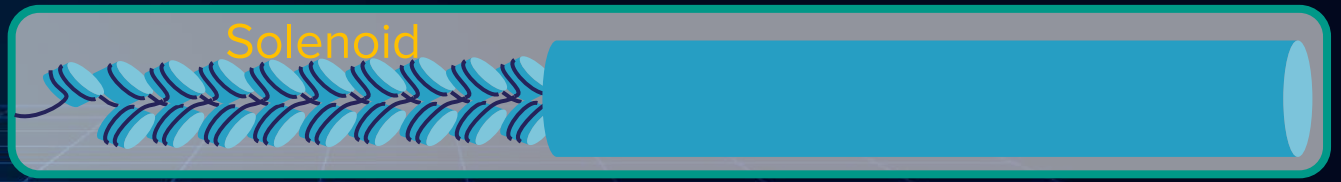
- Nucleosomes are seen as beads on thread-like structures in the nucleus.
- These structures are known as chromatin because they are seen as coloured bodies when stained.
- They are 10 nm in diameter.
- It gives a beads on string appearance under an electron microscope.

DNA Packaging in Eukaryotes





Chromatin further condenses to form a solenoid that is a 30 nm structure. This is more condensed than chromatin.

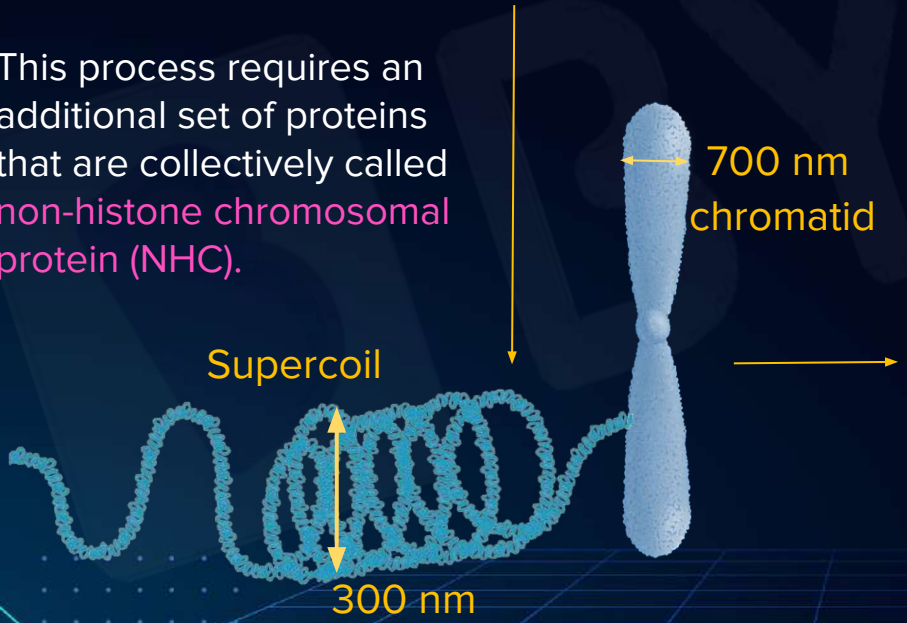


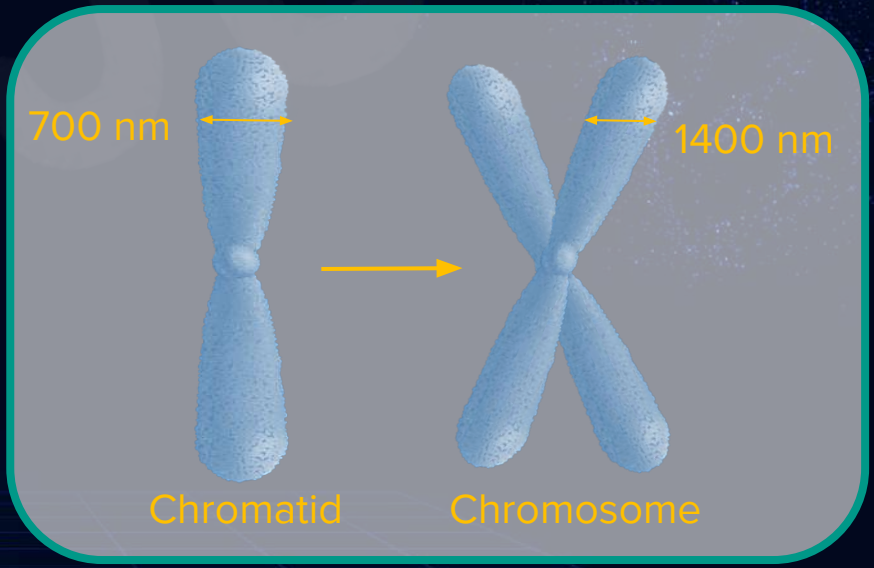
DNA Packaging in Eukaryotes

Finally solenoid supercoils to form the chromatid that is 700 nm in diameter.

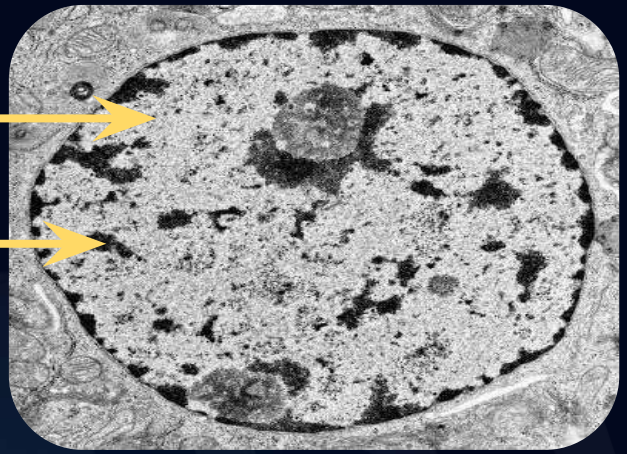
During metaphase, the chromosome is in the most coiled state at 1400 nm.

This process requires an additional set of proteins that are collectively called non-histone chromosomal protein (NHC).



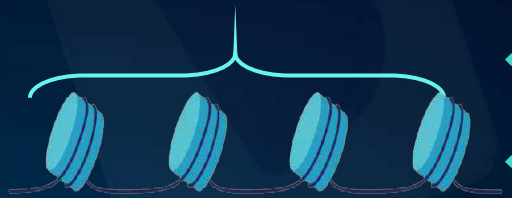


Type of Chromatin Fibers



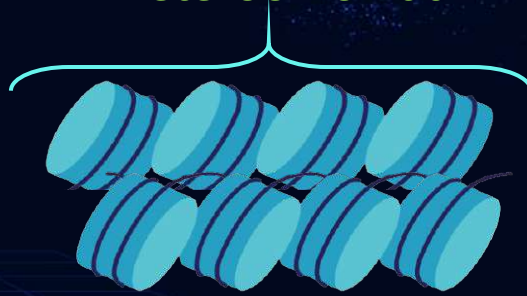
- When nucleus is stained some parts of the chromatin are lightly stained whereas others are stained darker.
- Lightly stained fibres are called euchromatin.
- Dark stained are called heterochromatin.

Euchromatin



- Loosely packed region
- Stains light
- Transcriptionally active

Heterochromatin



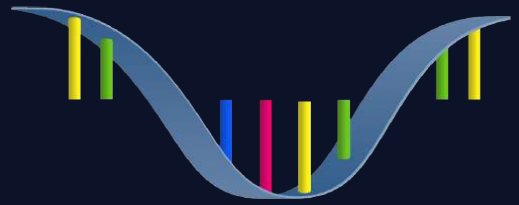
- Densely packed region
- Stains dark
- Transcriptionally inactive

Recall : Nucleic Acids

Nucleic acids

DNA
Deoxyribonucleic acid

RNA
Ribonucleic acid



Key Characteristics of RNA

- RNA is found in almost all living organism today.

Plantae

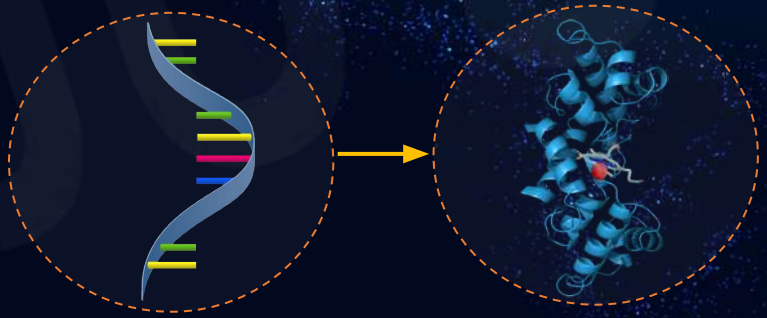
Animalia

Monera

Fungi

Protista

- RNA can directly produce proteins

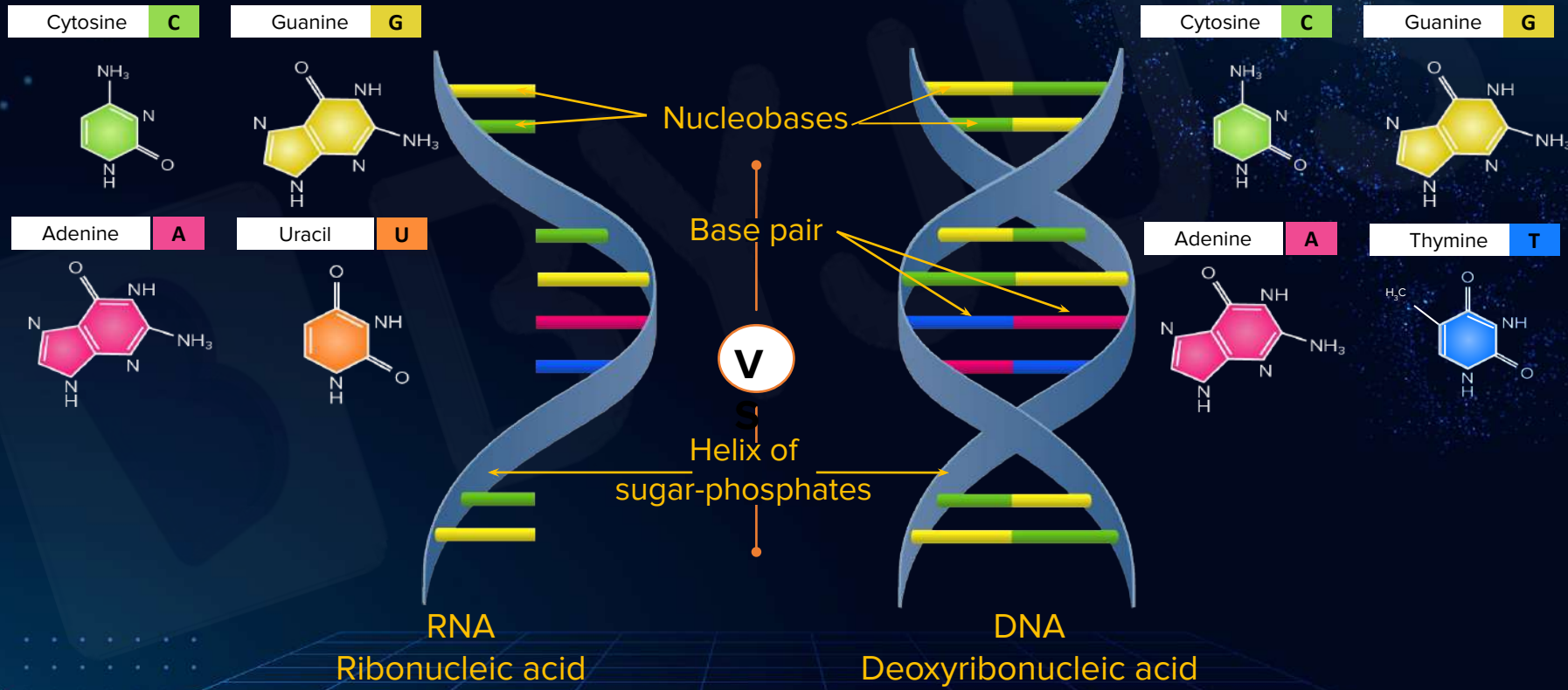


RNA

Proteins

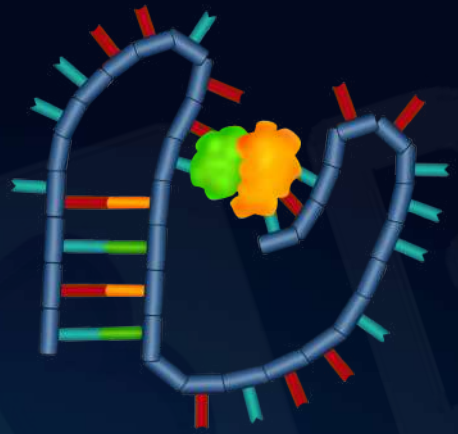
Key Characteristics of RNA

- It is structurally similar to DNA



Key Characteristics of RNA

- RNA is able to act as an enzyme



- Important biochemical reactions in living systems are catalysed by RNA catalysts and hence called as ribozymes.

- RNA is able to cut out unnecessary parts

- RNA can cut out the junk part

RNA World Hypothesis

RNA = First genetic material

Life on Earth began with a simple RNA molecule that could copy itself without help from other molecules.

DNA vs RNA

DNA	RNA
1	1

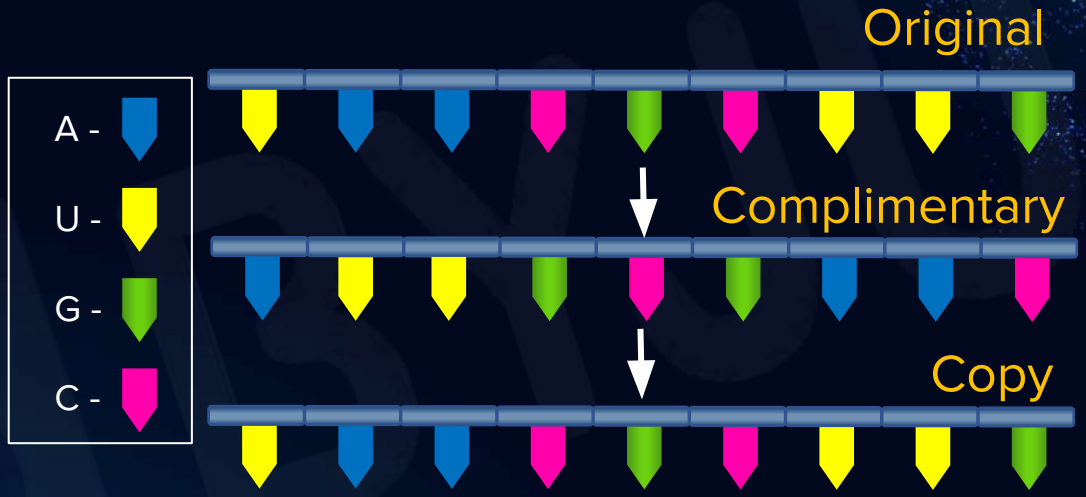
Generates replica

Because of rule of base pairing and complementarity, both the nucleic acids (DNA and RNA) have the ability to direct their duplications by the process of replication.

In case of RNA -

- A single strand of RNA is used as a template to create a second complementary strand by base pairing.
- This complementary strand is then used to create another complementary strand which becomes the copy of the original.

DNA vs RNA



DNA vs RNA

DNA	RNA
1	0.5

Provides the scope for slow changes required for evolution

- These slow changes are facilitated by mutation.
- Both DNA and RNA are able to mutate.
- However, mutations in DNA are slow and minor which provides the organism with better chances of survival while also incorporating changes and evolving.

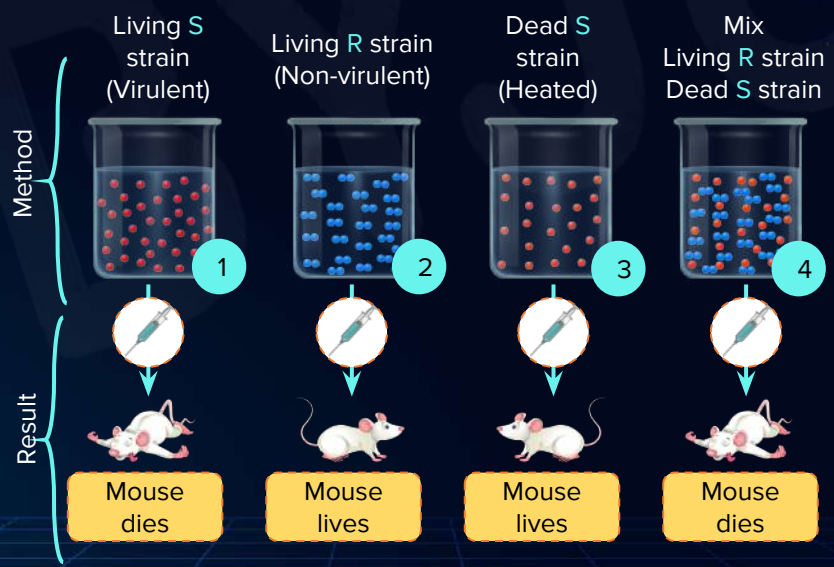
- In RNA, mutations are faster and often major. Thus, organisms with RNA often are at a risk of going extinct because of lethal mutations.
- Organism with RNA, hence, mostly have shorter life spans and evolve faster.

DNA vs RNA

DNA	RNA
1	0

Stable chemically and structurally

DNA is much more stable than RNA



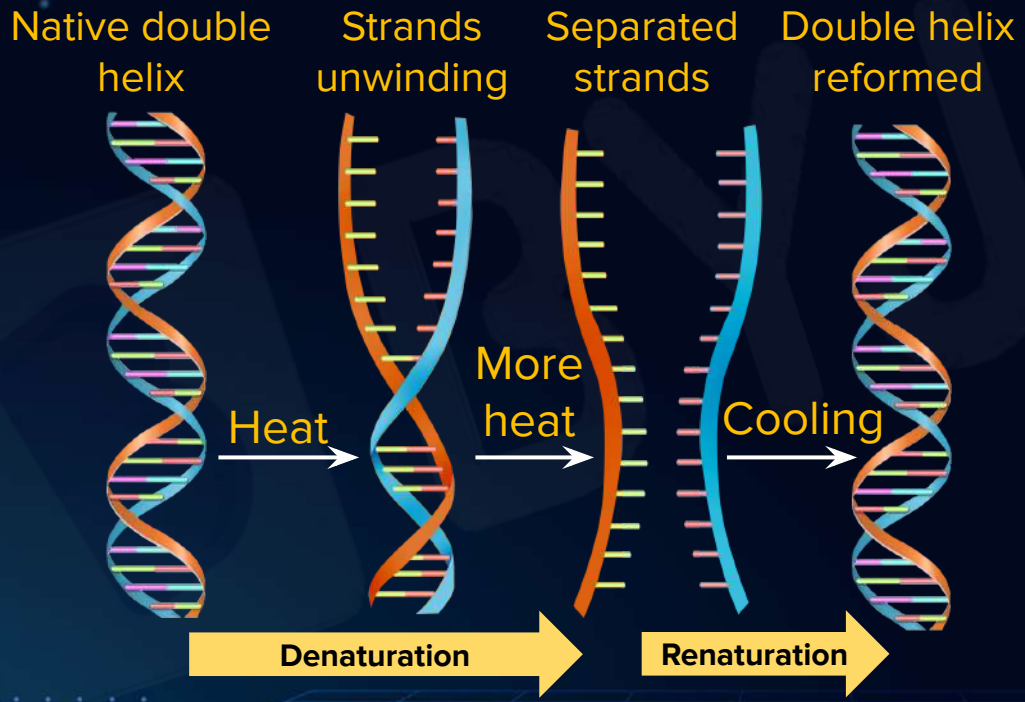
DNA vs RNA

DNA is much more stable than RNA

- In Griffith's 'transforming principle' showed that heat, which killed the bacteria, at least did not destroy some of the properties of genetic material.
- This suggested that genetic material was much more stable and resistant to damage than other materials of the cell.
- There are 3 reasons to why DNA is more stable than RNA

DNA vs RNA

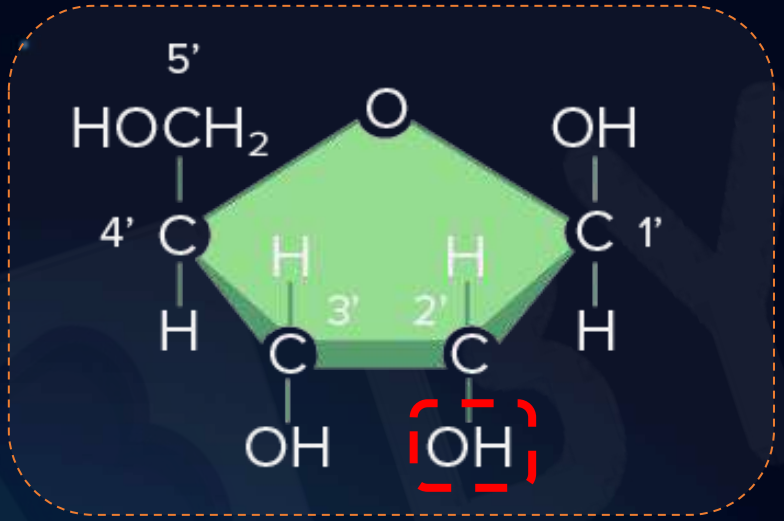
Reason 1 - Double stranded nature of DNA



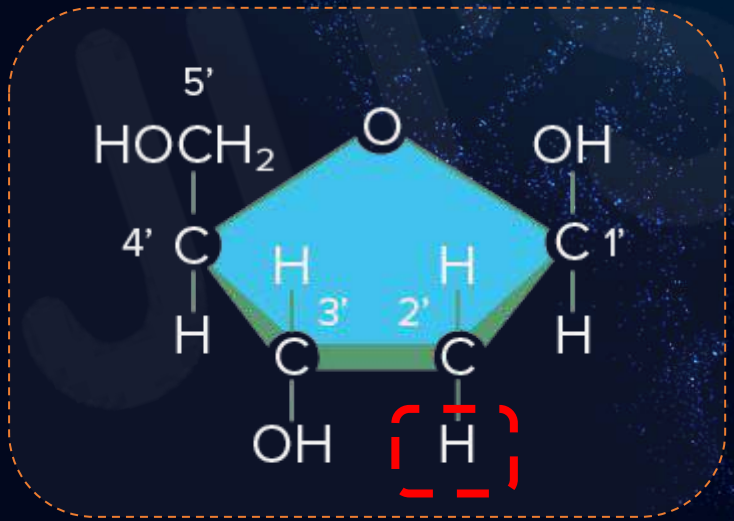
- DNA's double helical structure has the ability to get back to its original state after extreme heat is removed.
- Hence, it can survive heat, making it extremely stable.
- Since, DNA is double stranded and has a complementary strand, it further resists changes by evolving a process of repair.

DNA vs RNA

Reason 2 - Structural difference



Ribose

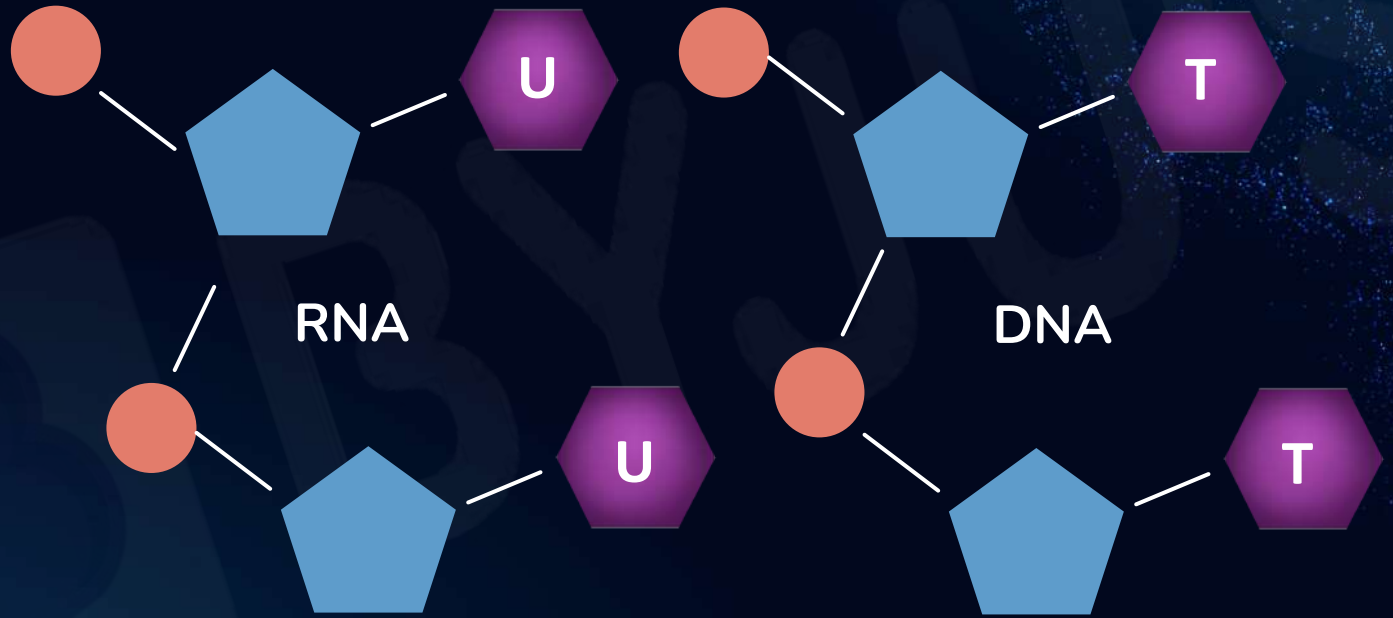


Deoxyribose

- 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, which means less stable.

DNA vs RNA

Reason 3 - Difference in nucleobases



- Thymine, which is found in DNA, is more stable than Uracil, found in RNA, because of the presence of extra methyl group.
- Hence, DNA becomes more stable than RNA.

DNA vs RNA

DNA	RNA
1	1

Express itself in the form of 'Mendelian characters'

- Both DNA and RNA are able to change the phenotype of individuals.
- Both of them can be passed on from generation to generation i.e. they can be inherited.

- RNA can directly code for the synthesis of proteins.
- Hence, it can easily express the characters.
- DNA, however, is dependent on RNA for the synthesis of proteins.

Summary

Packaging of DNA

In prokaryotes

- Prokaryotes don't have a nucleus.
- Genomic DNA in prokaryotes is organized in large loops held by special proteins called NAPs.

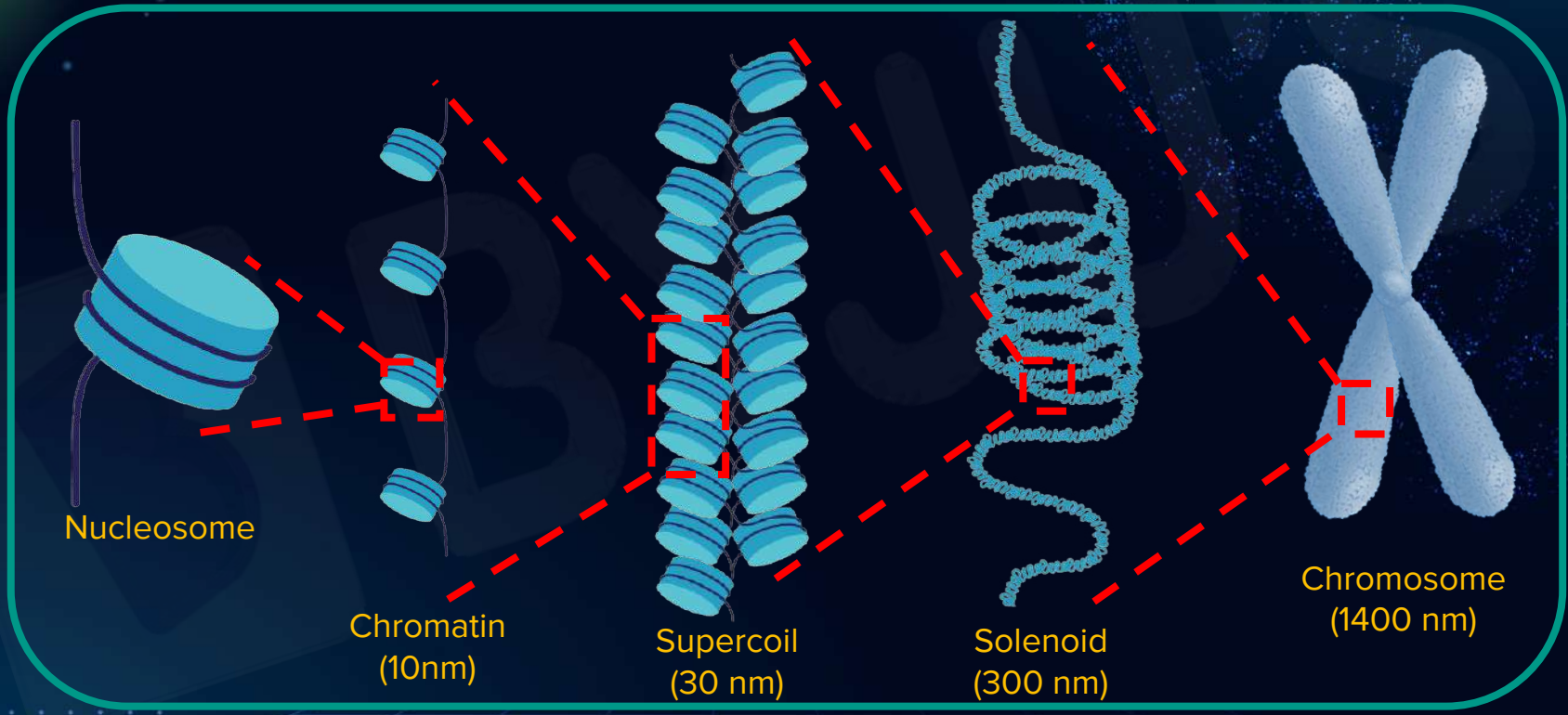
- Basically, prokaryotic chromosomes fit the DNA in the cell by supercoiling of DNA that is held close together by NAPs.

In eukaryotes

- Eukaryotes have a well-defined nucleus.
- Genomic DNA in eukaryotes is wrapped around an eight molecule structure called histone octamer to form nucleosome.

- Nucleosomes are connected by linker DNA and go a few more stages of packaging to transform into chromosomes.

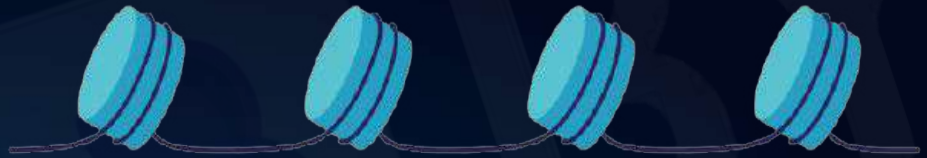
Summary



Summary

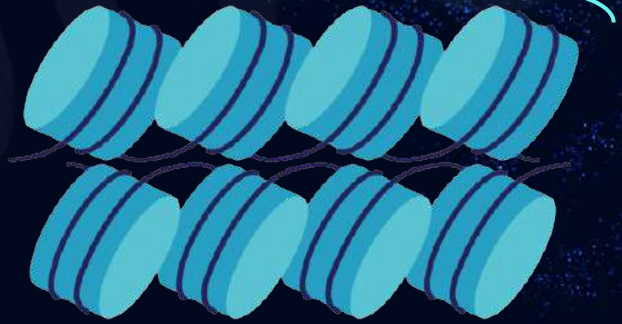
Types of chromatin

Euchromatin



- Loosely packed region

Heterochromatin



- Densely packed region

Summary

DNA		RNA
1	Generates replica	1
1	Provides scope for slow changes required for evolution	0.5
1	Stable chemically and structurally	0
1	Express itself in the form of 'Mendelian characters'	1
4	Final score	2.5

Thus, DNA is better genetic material than RNA

BYJU'S Classes Notes

Molecular Basis of Inheritance

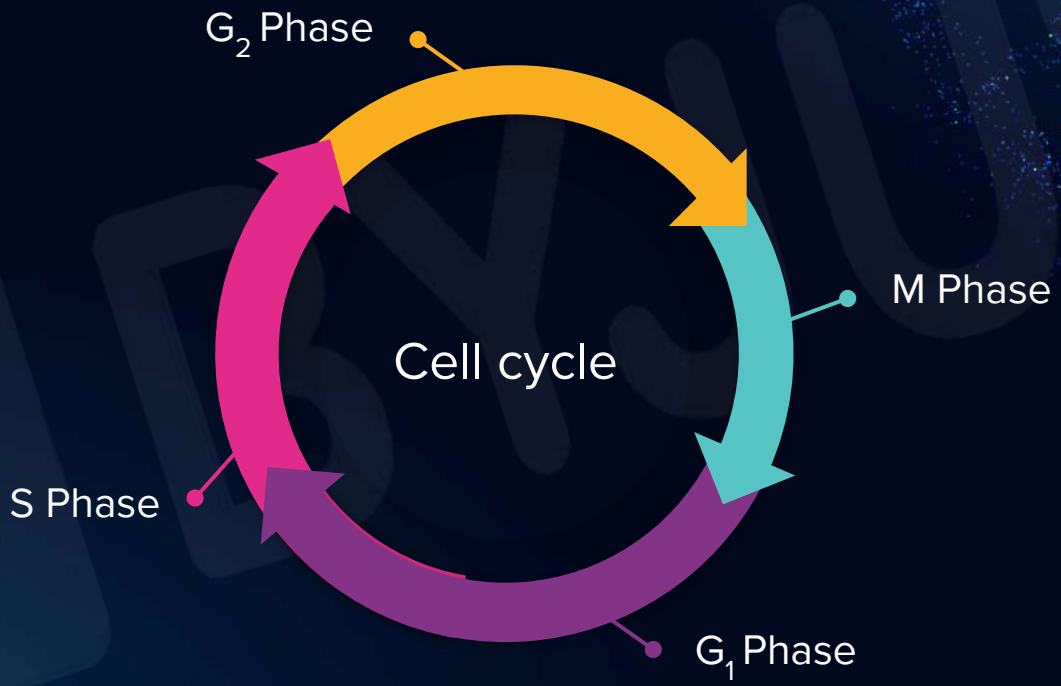
DNA Replication, Meselson and Stahl’s Experiment, Taylor’s Experiment, Process of DNA Replication

Key Takeaways

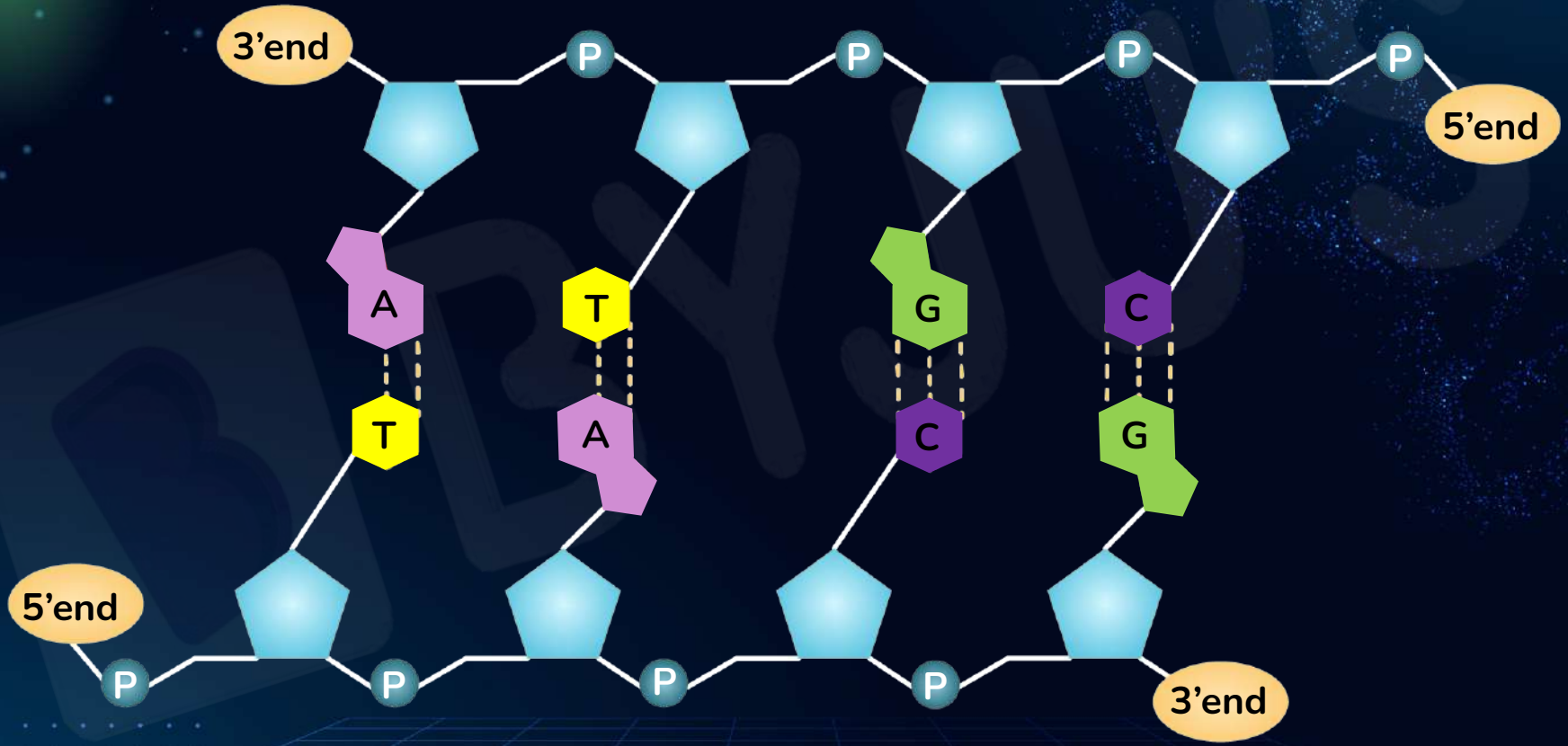
1. Semi-conservative DNA replication
   - Meselson and Stahl's experiment
   - Taylor's experiment
2. Process of replication

Summary

Recall! Cell Cycle



Recall! Antiparallel DNA Strands

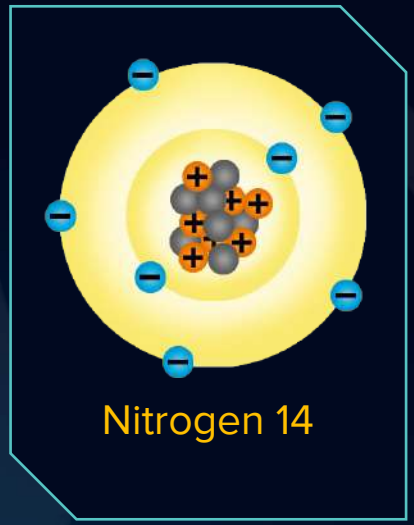


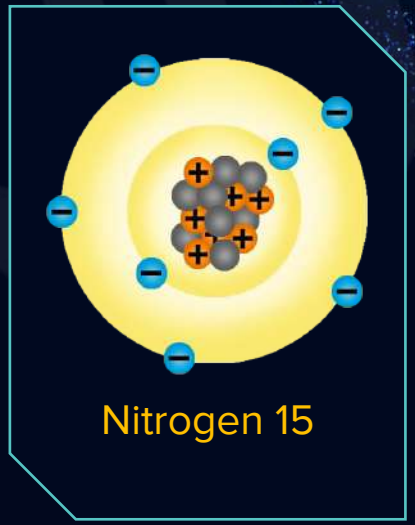
Recall! Antiparallel DNA Strands

- DNA has 2 main strands. Each strand consists of a phosphate backbone.
- It basically consists of the 5’ carbon of the ribose sugar having a phosphate group which is then connected to the 3’ carbon having free-OH group of the next deoxyribose sugar.
- So one strand goes from 5’ to 3’.
- The other strand of DNA, goes from 3’ to the 5’.
- Thus, DNA strands are antiparallel to each other.

Recall! Isotopes

Isotopes are the atoms of a chemical element having same atomic number but different mass number.





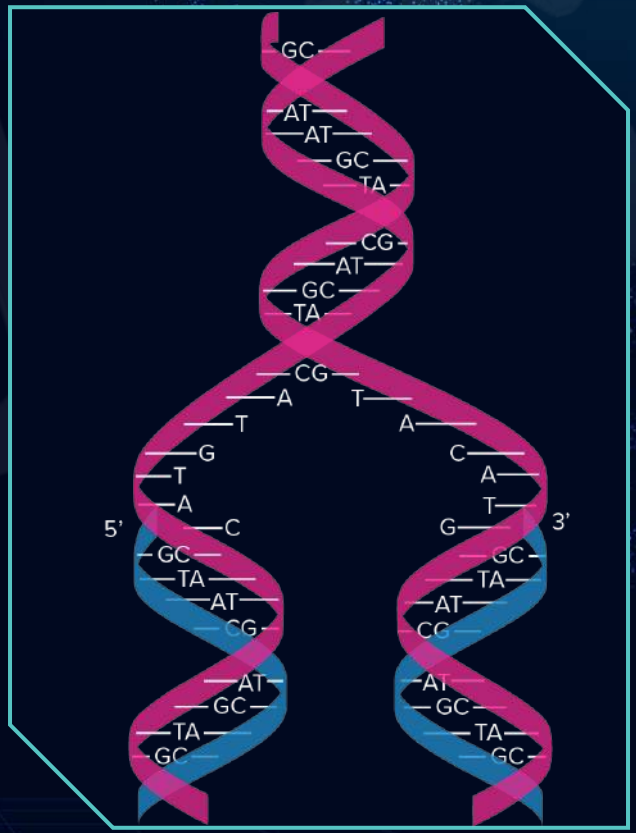
^{14}N contains 7 neutrons

^{15}N contains 8 neutrons

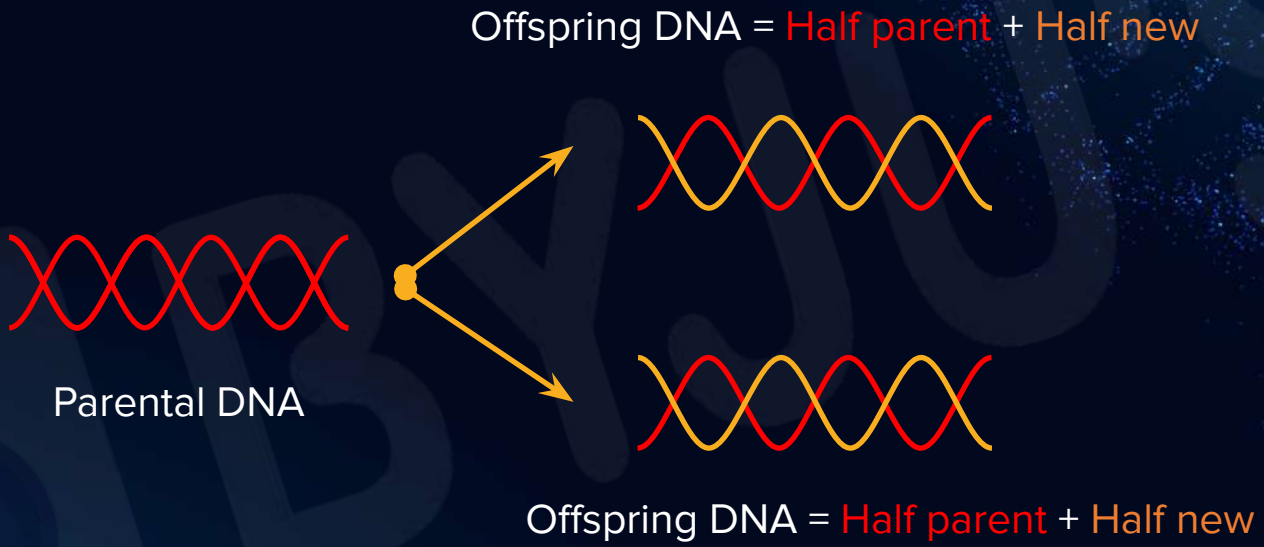
Semi-conservative DNA Replication

Replication

- A process of copying and duplicating the genetic material (DNA).
- Watson and Crick were believers of semi-conservative DNA replication.



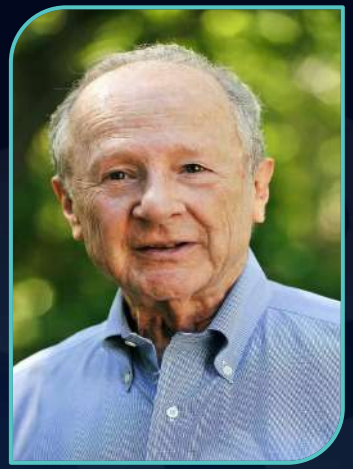
Semi-conservative DNA Replication



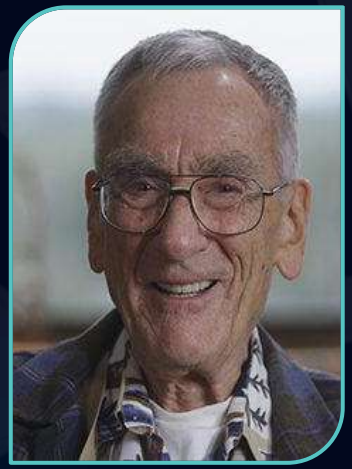
- In the semi-conservative DNA replication, each daughter cell receives 1 half of the each parental strand that acts as a template strand and new offspring DNA will be complementary to the parental DNA.

Semi-conservative DNA Replication

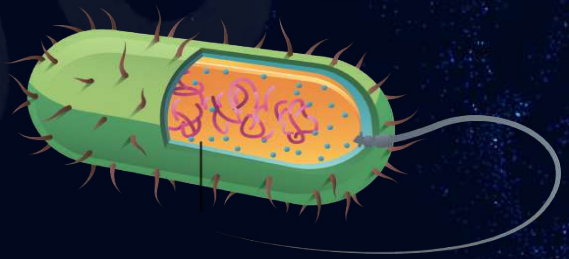
Meselson and Stahl's Experiment



Matthew Meselson



Franklin Stahl



Escherichia coli

- Matthew Meselson and Franklin Stahl worked on *Escherichia coli* (*E. coli*) in 1958.

Meselson and Stahl's Experiment

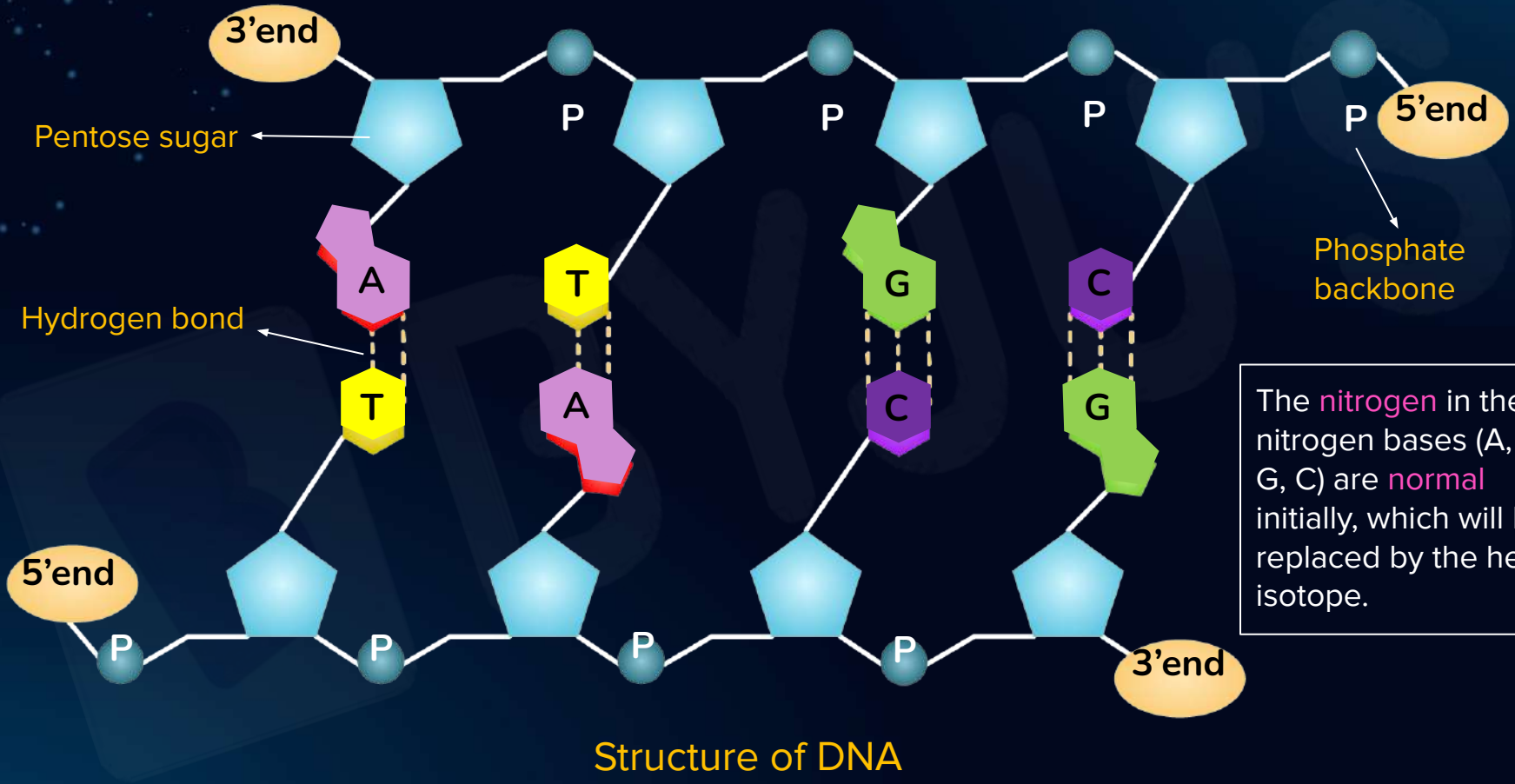
Step 1: Grow *E. coli* in $^{15}NH_4Cl$ Media

$^{15}NH_4Cl$ Media

E. coli

- First normal bacteria is grown in culture containing $^{15}NH_4Cl$.
- ^{15}N is heavy isotope of Nitrogen.
- It is not a radioactive nitrogen, but is a heavier isotope of nitrogen.

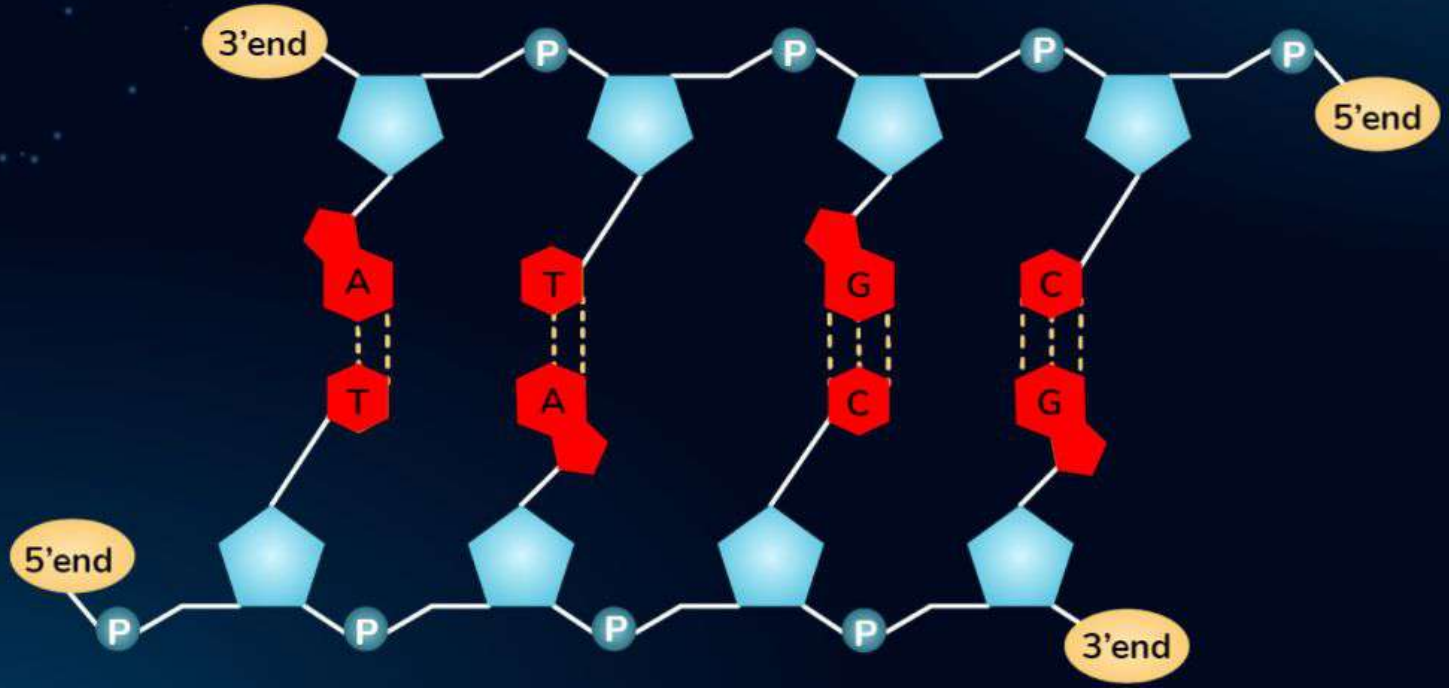
Meselson and Stahl's Experiment



Structure of DNA

The nitrogen in the nitrogen bases (A, T, G, C) are normal initially, which will be replaced by the heavy isotope.

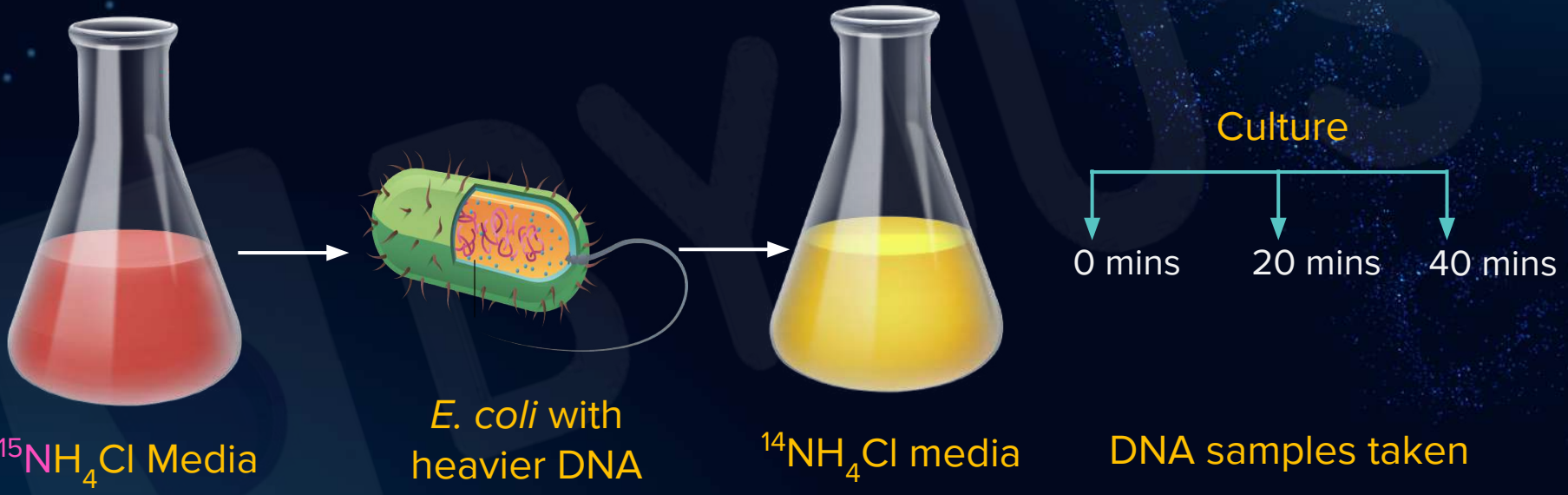
Meselson and Stahl's Experiment



- The nitrogens in all the nitrogen bases would be replaced with heavier ^{15}N as the bacteria grow.
- As a result, DNA molecule will have all the nitrogenous bases containing the heavier ^{15}N.

Meselson and Stahl's Experiment

Step 2: Transfer of *E. coli* with heavier DNA (^{15}N) into regular $^{14}NH_4Cl$ media



- The *E. coli* from the heavier ^{15}N media are transferred into the regular ^{14}N media and the *E. coli* are allowed to grow.
- Samples were collected after every 20 mins as they divide after every 20 mins.

Meselson and Stahl's Experiment

Step 3: CsCl centrifugation of the DNA samples

CsCl centrifugation is a technique used for the separation of particles from a solution according to

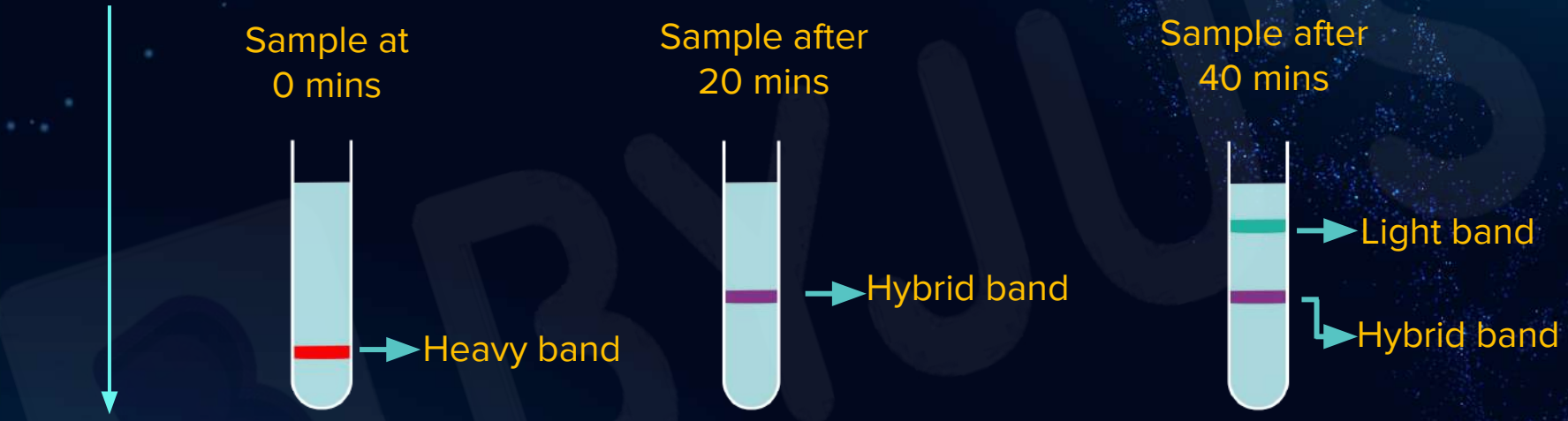
- ✔ Size
- ✔ Shape
- ✔ Density
- ✔ Viscosity of the medium

Meselson and Stahl's Experiment

Step 3: CsCl centrifugation

After centrifuging the DNA samples in CsCl, the heavier DNA will settle down as heavier bands while the lighter DNA samples will get suspended in the middle.

Meselson and Stahl's Experiment



The 0 min sample had just a heavy band at the bottom

The 20 min sample had a hybrid band

The 40 min sample had 2 bands:

- 1 that was the same as the 20 min sample i.e. the hybrid band
- New band much lighter than the bands seen before.

Meselson and Stahl's Experiment - Results

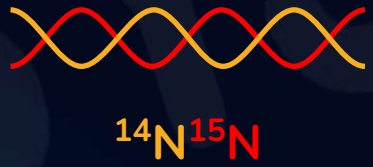
Heavy DNA band

- Contains both the DNA strands with the heavier ^{15}N

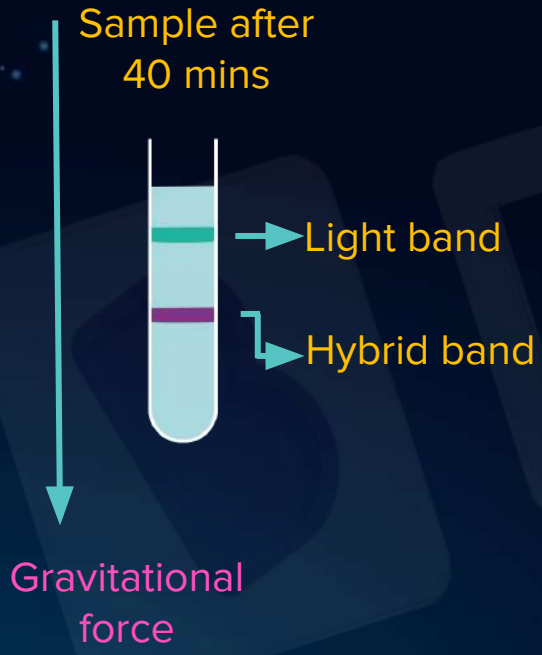
Meselson and Stahl's Experiment - Results

Hybrid DNA band

- Contains one strand of lighter ^{14}N and another of heavier ^{15}N



Meselson and Stahl's Experiment - Results



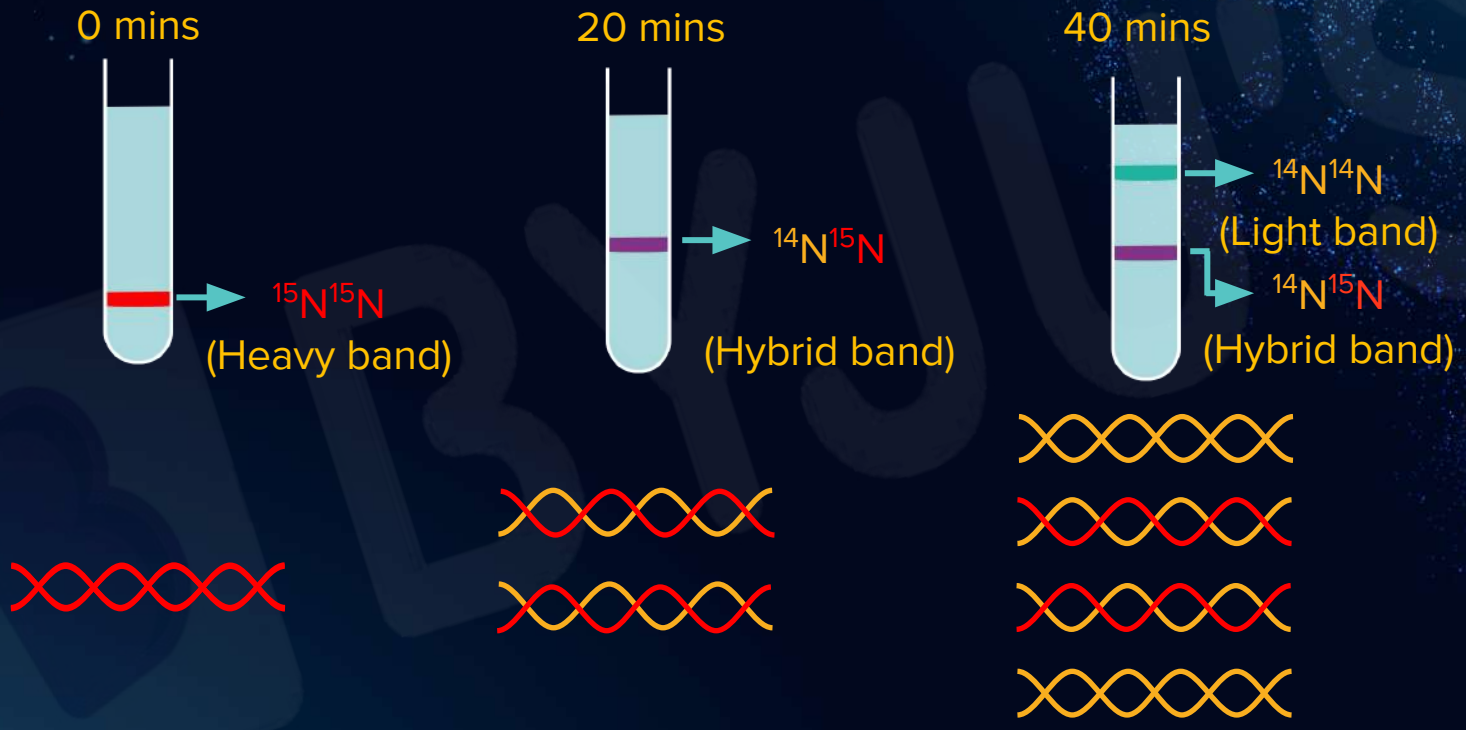
Light band

- Contains both strands of lighter ^{14}N

Hybrid DNA band

- Contains one strand of lighter ^{14}N and another of heavier ^{15}N

Meselson and Stahl's Experiment - Results



Meselson and Stahl's Experiment - Conclusions

- They concluded that the DNA strands separate and each makes a copy of itself, so that the daughter molecule comprises of one old and one new strand.
- This implies that the newly synthesised DNA obtained one of its strands from the parent strand.
- Thus DNA replication is semi-conservative in nature, as proved by this experiment.

Taylor’s Experiment

Dr. Herbert Taylor

Faba bean plant

- Taylor and his colleagues performed similar experiment on the root tip cells of faba beans plant.

Taylor's Experiment

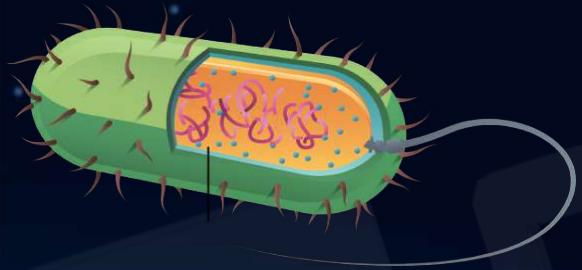
Faba bean plant

Radioactive thymidine

- They immersed the root tips of faba bean plant in the radioactive thymidine solution and grew it for a while.
- In the first generation, both chromatids were labeled.
- In the second generation of cell division, one chromatid of each chromosome was labeled and one was normal.
- This demonstrated semi-conservative mode of replication.

Semi-Conservative Replication

Meselson & Stahl



E. coli (Prokaryotes)

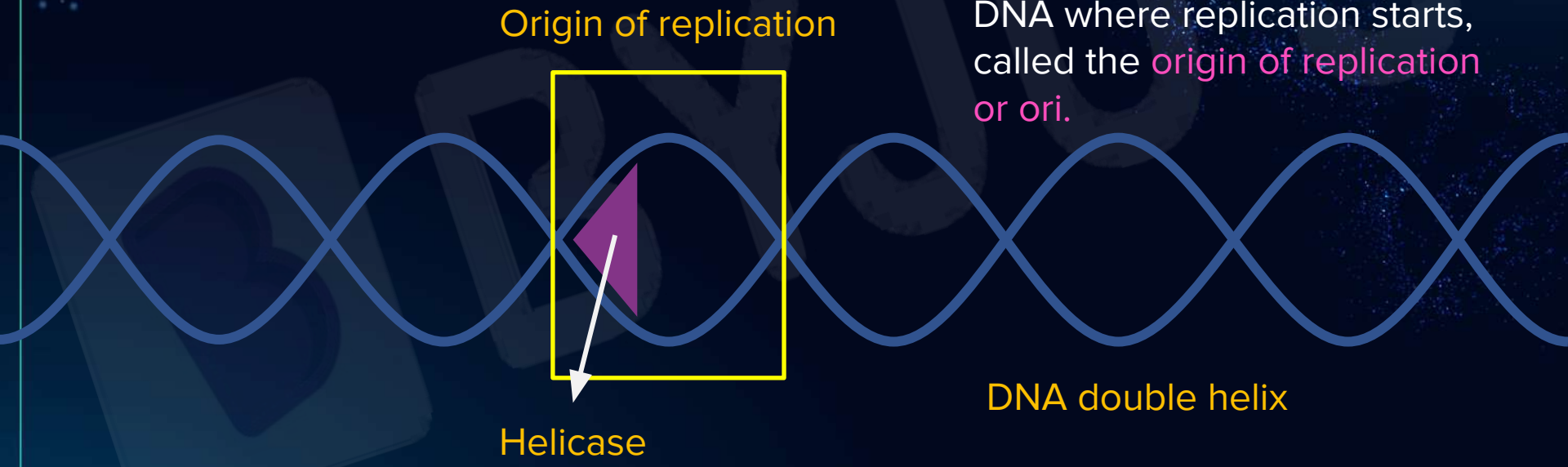
Herbert Taylor

Faba beans (Eukaryotes)

In conclusion, Meselson and Stahl proved semiconservative replication in prokaryotes, while Taylor did the same for eukaryotes.

Process of DNA Replication

- Replication cannot begin randomly in any place in DNA.
- There are specific regions of DNA where replication starts, called the origin of replication or ori.



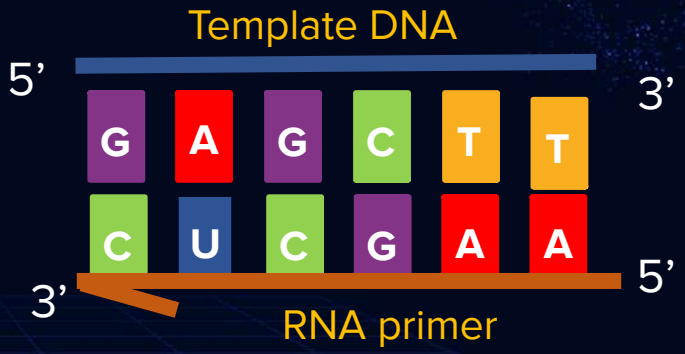
- Helicase helps unwind the DNA

Process of Replication

RNA primer

- 5 – 10 nucleotide long RNA fragment .

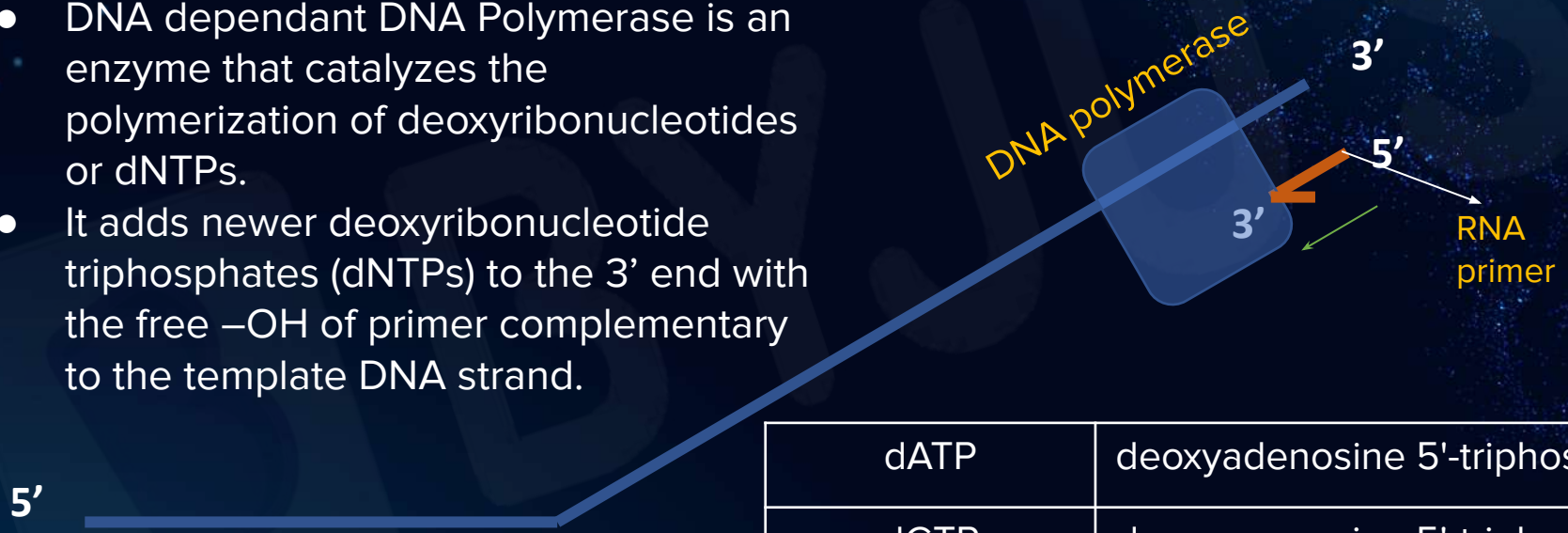
- RNA primer is complementary to the template DNA.
- Instead of thymine which is complementary to the adenine, uracil is present in RNA.



Process of Replication

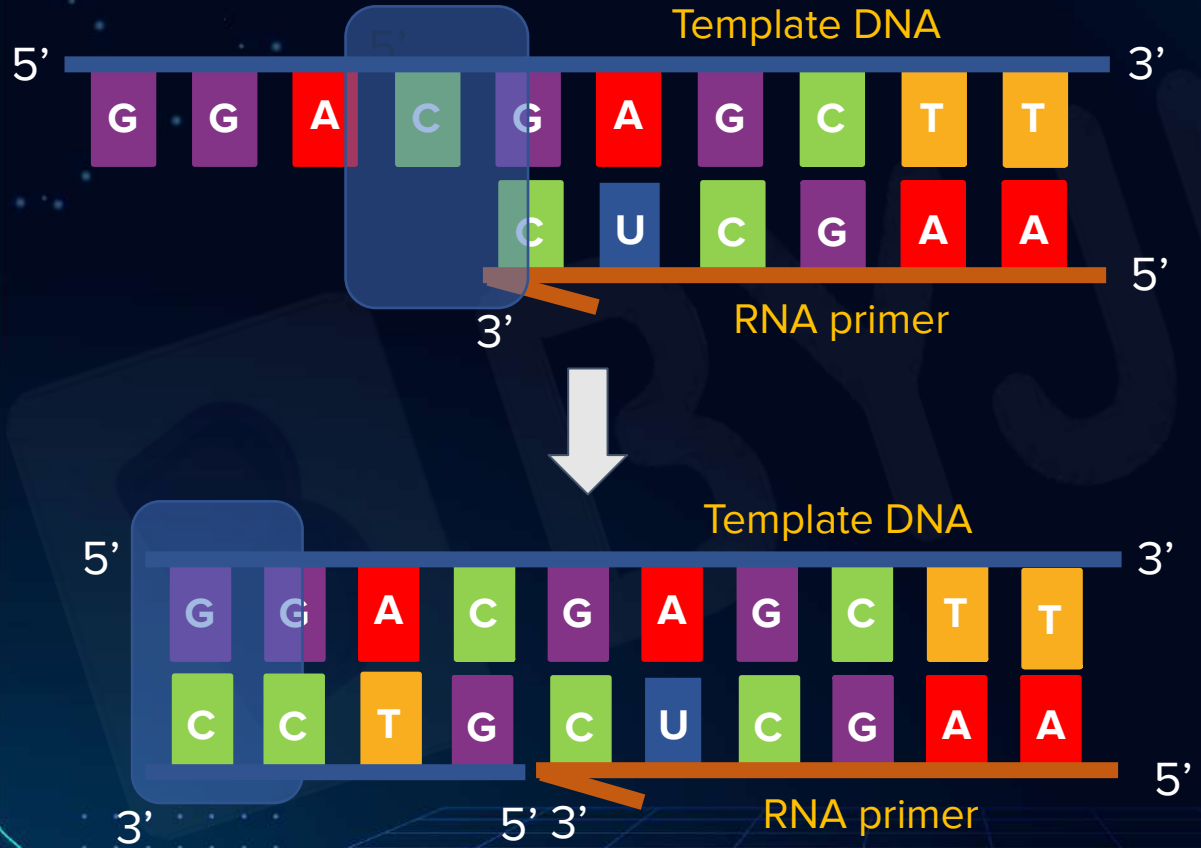
DNA polymerase

- DNA dependant DNA Polymerase is an enzyme that catalyzes the polymerization of deoxyribonucleotides or dNTPs.
- It adds newer deoxyribonucleotide triphosphates (dNTPs) to the 3’ end with the free –OH of primer complementary to the template DNA strand.



dATP	deoxyadenosine 5'-triphosphate
dGTP	deoxyguanosine 5'-triphosphate
dTTP	deoxythymidine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate

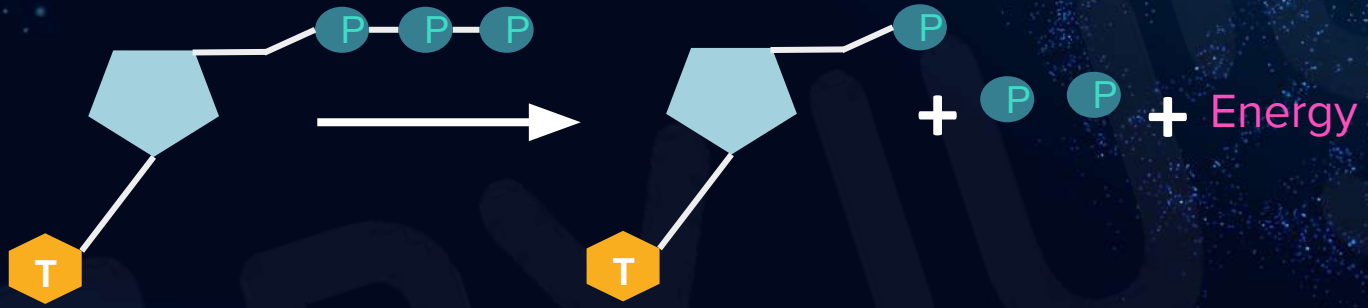
Process of Replication



- A small strand of RNA is synthesized with RNA primase enzyme.
- Synthesized RNA is called RNA primer.
- It has 4-12 nucleotides.
- RNA primase initiates DNA synthesis by synthesizing short-lived oligonucleotides by providing 3-OH group for joining nucleotides.

Process of Replication

Where does the energy for polymerization come from?

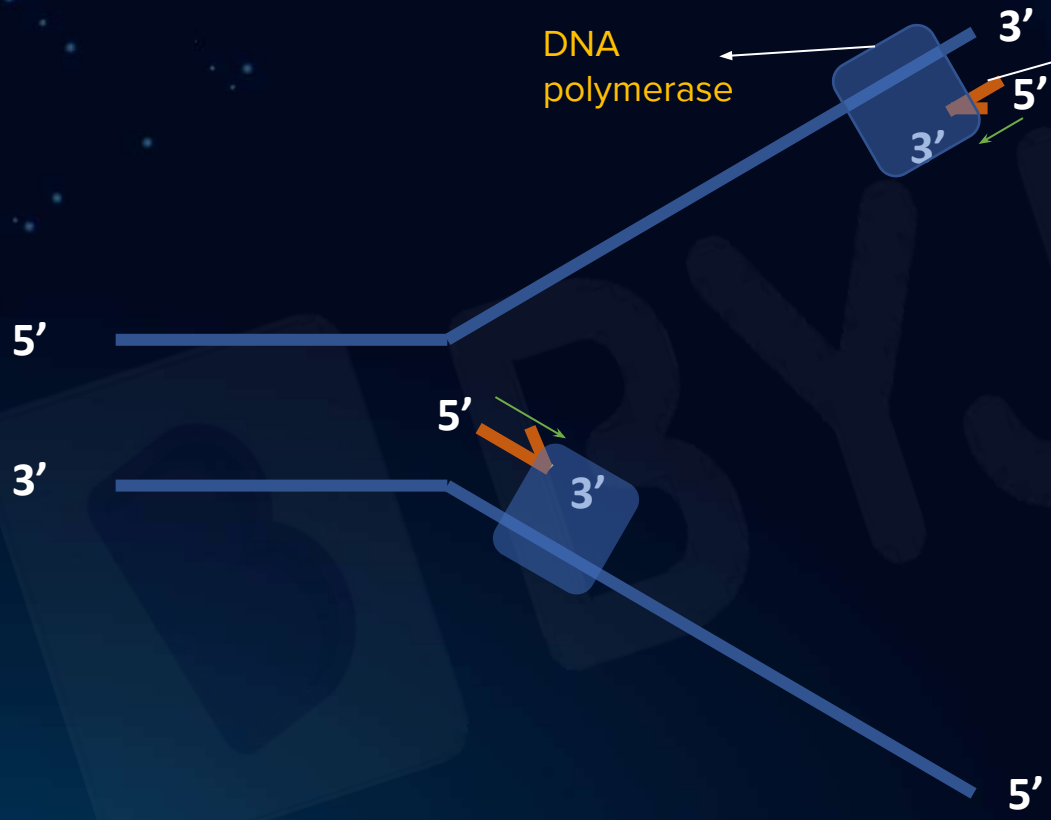


dTTP
Deoxythymidine
triphosphate

dTMP
Deoxythymidine
monophosphate

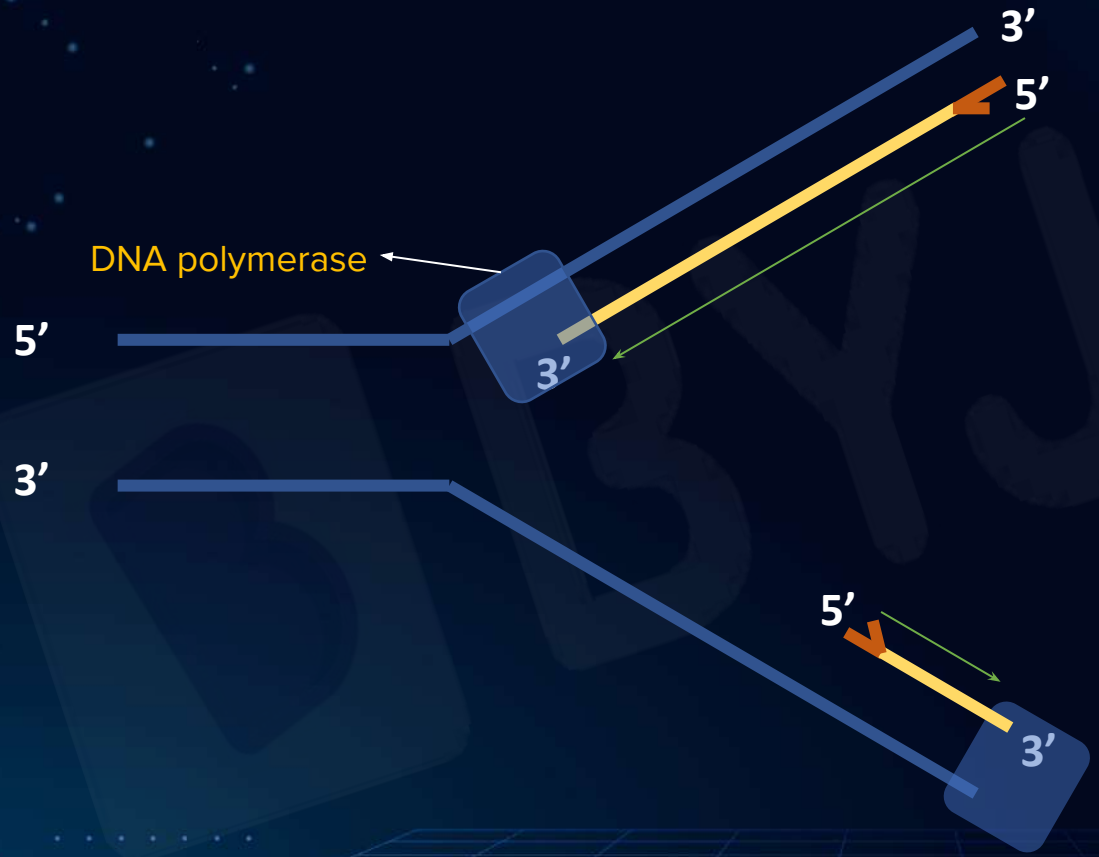
- Deoxyribonucleotide triphosphate (dNTP) is the form of nucleotides which acts as substrates and provides energy for polymerisation.
- It has 3 main parts, it has the deoxyribose, 3 phosphates held together with energy rich bonds and 1 nitrogen base (Thymidine), thus making this deoxythymidine triphosphate.
- When this gets added to the DNA strand, the bonds between the 2 phosphate groups and the dNTP is broken, releasing a lot of energy which is utilized to add the dNMP that is left behind to the growing DNA strand.

Process of Replication



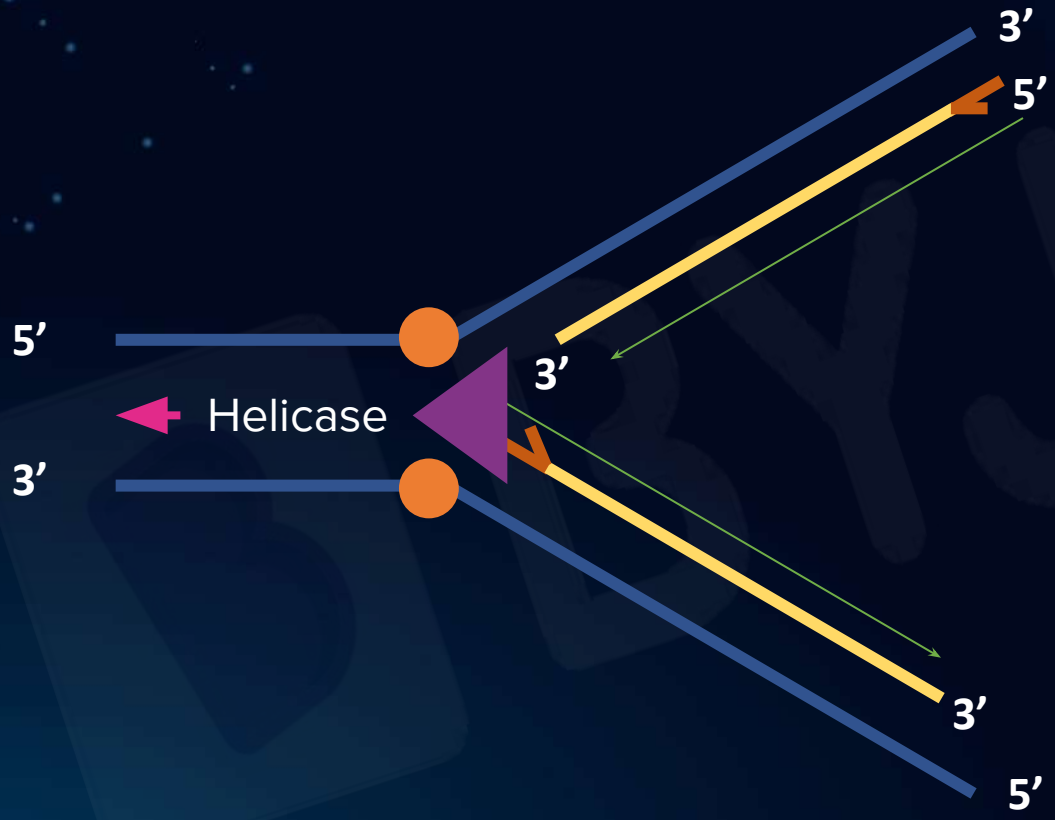
- Enzyme helicase acts over *ori* site and unwinds the two strands of DNA resulting in formation of Y-shaped structure called replication fork.
- When the primer appears on the top strand, another primer appears at the same time at the bottom strand as well.
- Simultaneously, DNA polymerase enzyme, like the one on the upper strand, also adds nucleotides only at the 3’ end of the primer.

Process of Replication



- So the DNA strand extends away from the primer in the 3’ direction.
- This continues until it reaches the end of the strand.

Process of Replication

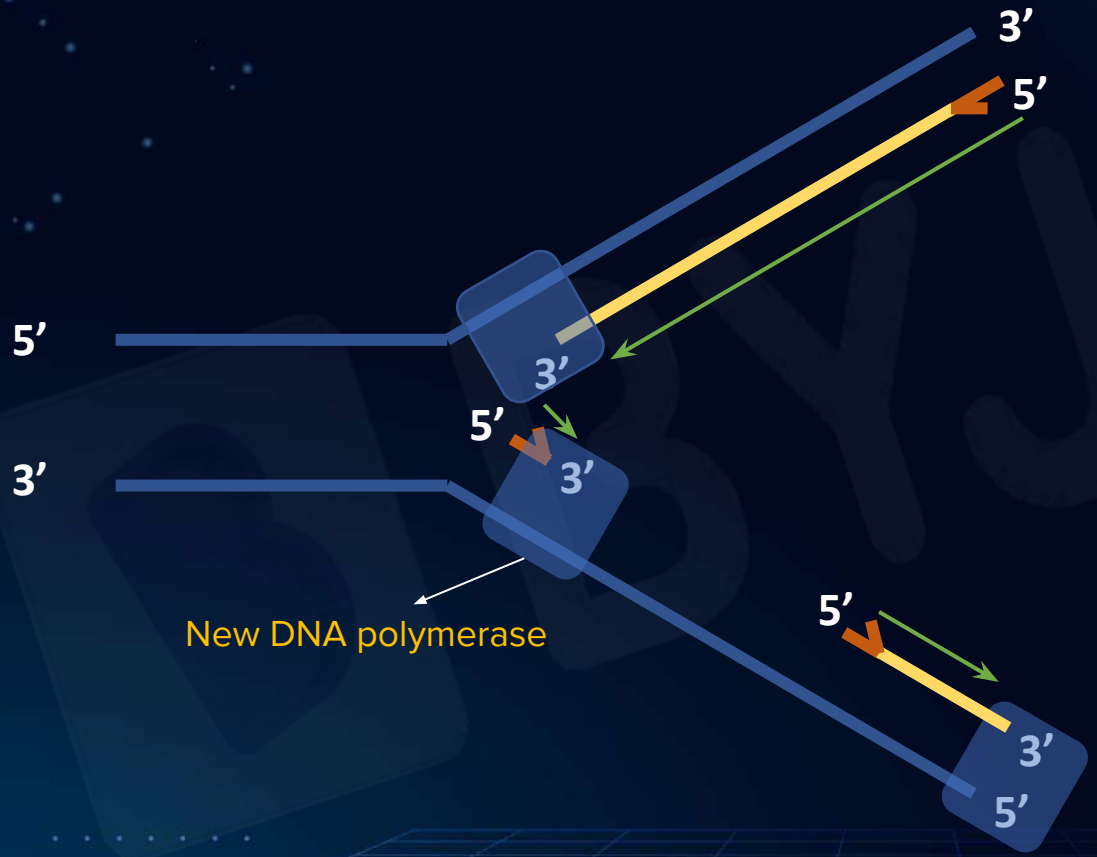


- The helicase moves leftwards

Process of Replication

- In case of the upper strand, the polymerase continues in the 3’ direction and it continues till it reaches the end of the replication fork.
- DNA polymerase can only move towards the 3’ direction. So there is no way for the new DNA strand at the lower end to extend towards the replication fork.

Process of Replication

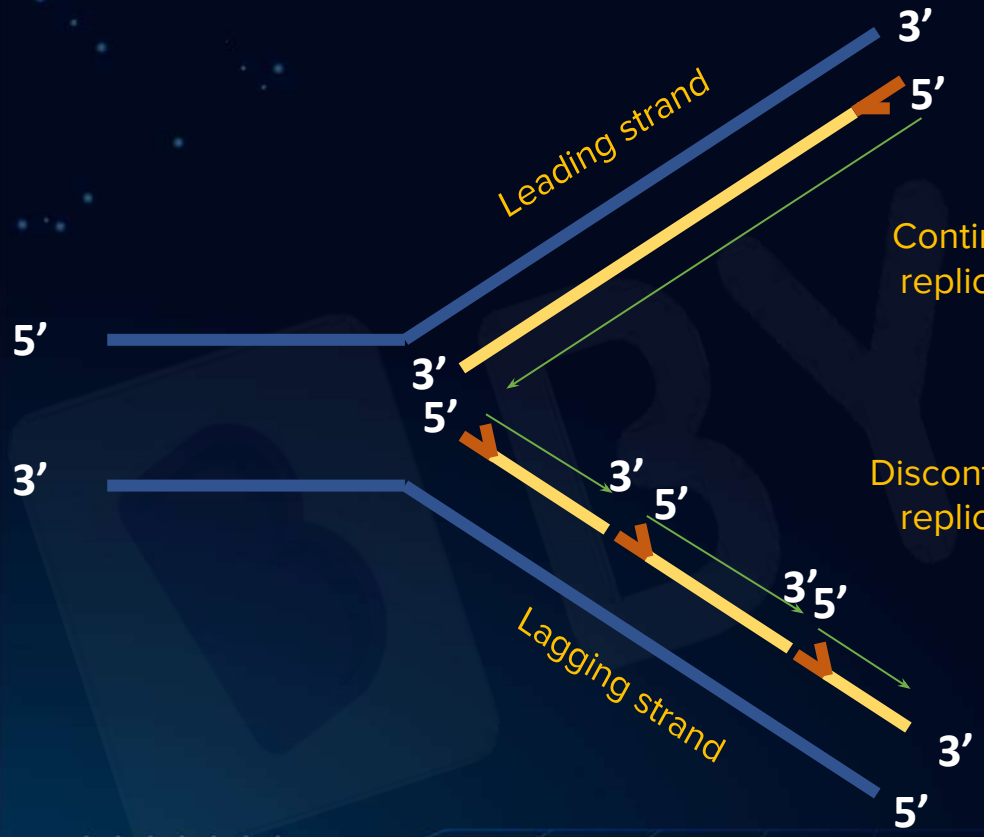


- One more primer gets attached to the lagging strand.
- Another DNA polymerase appears on the primer. It starts adding nucleotides in the 3’ direction.

Process of Replication

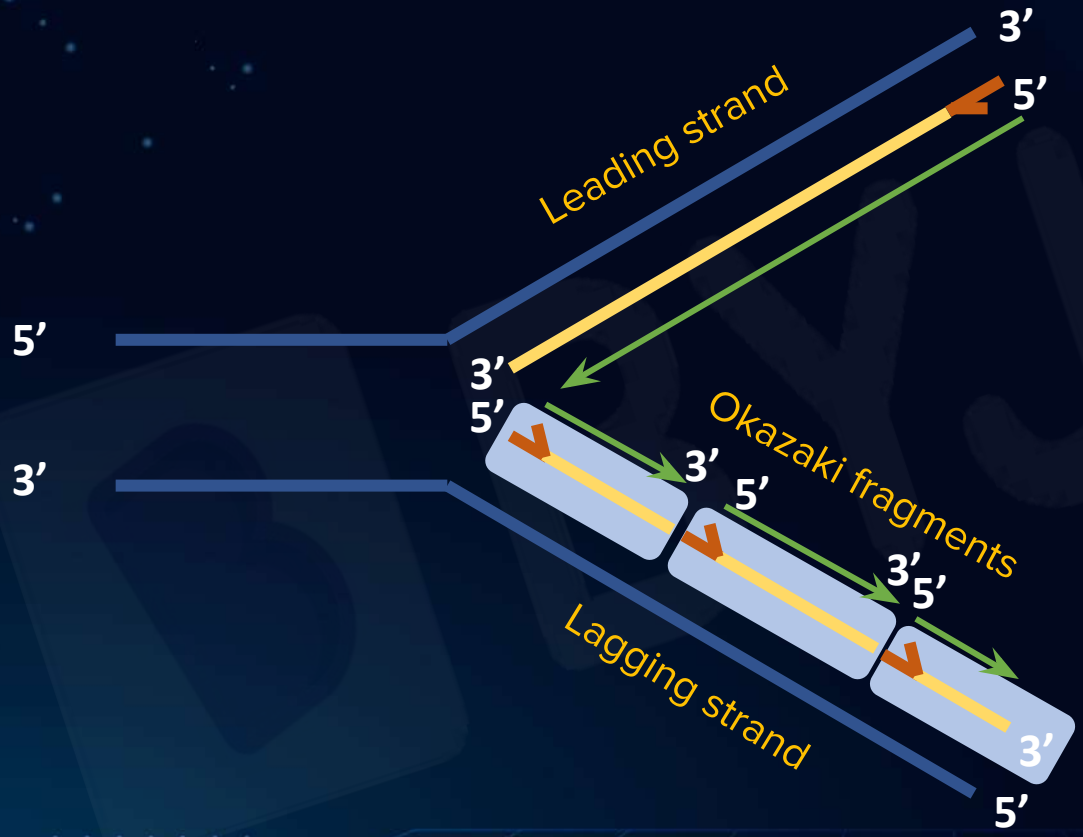
- Since the DNA polymerase cannot continue ahead, it disappears.
- On the lower strand, there is yet another primer and short DNA fragments are synthesized.
- Hence, the lower strand has DNA synthesized in several fragments.

Process of Replication



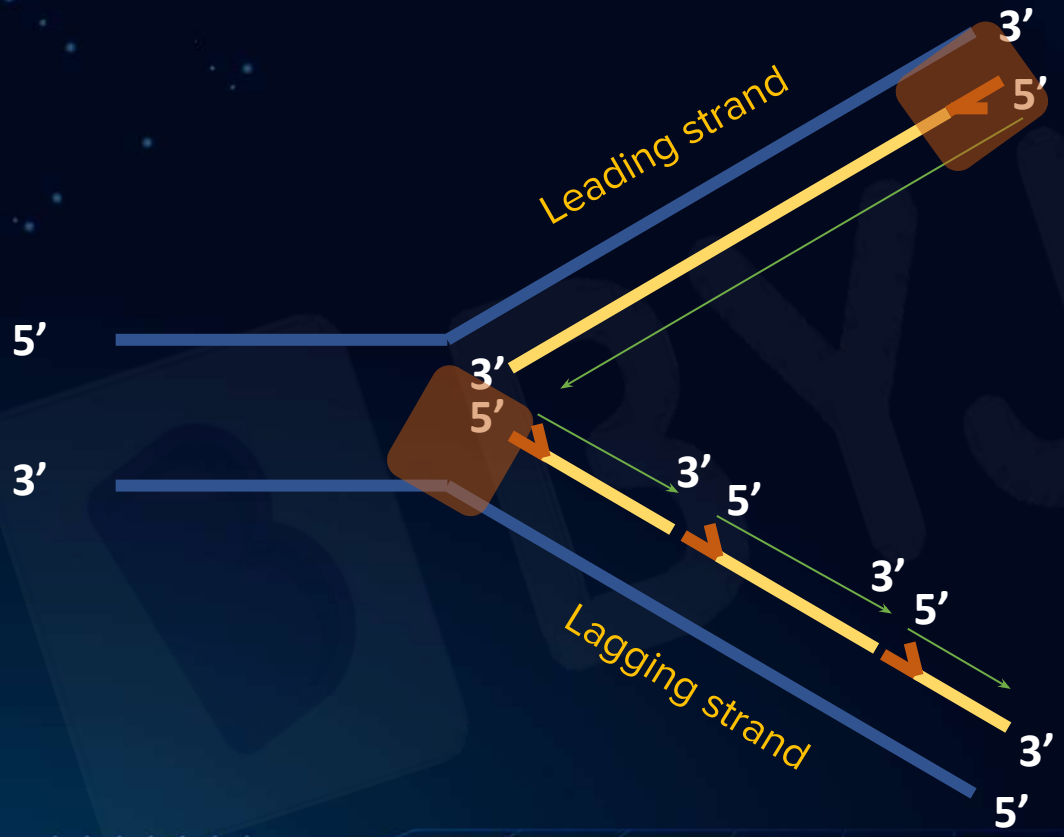
- The upper strand which is oriented from the 3’ to 5’ direction towards the replication fork, is called the leading strand. This is called so since the new strand being synthesized complementary to it is continuous.
- The lower strand oriented from the 5’ to 3’ direction towards the replication fork, is called the lagging strand, since the new strand that’s synthesized complementary to it is discontinuous, there are a lot of breaks.

Process of Replication



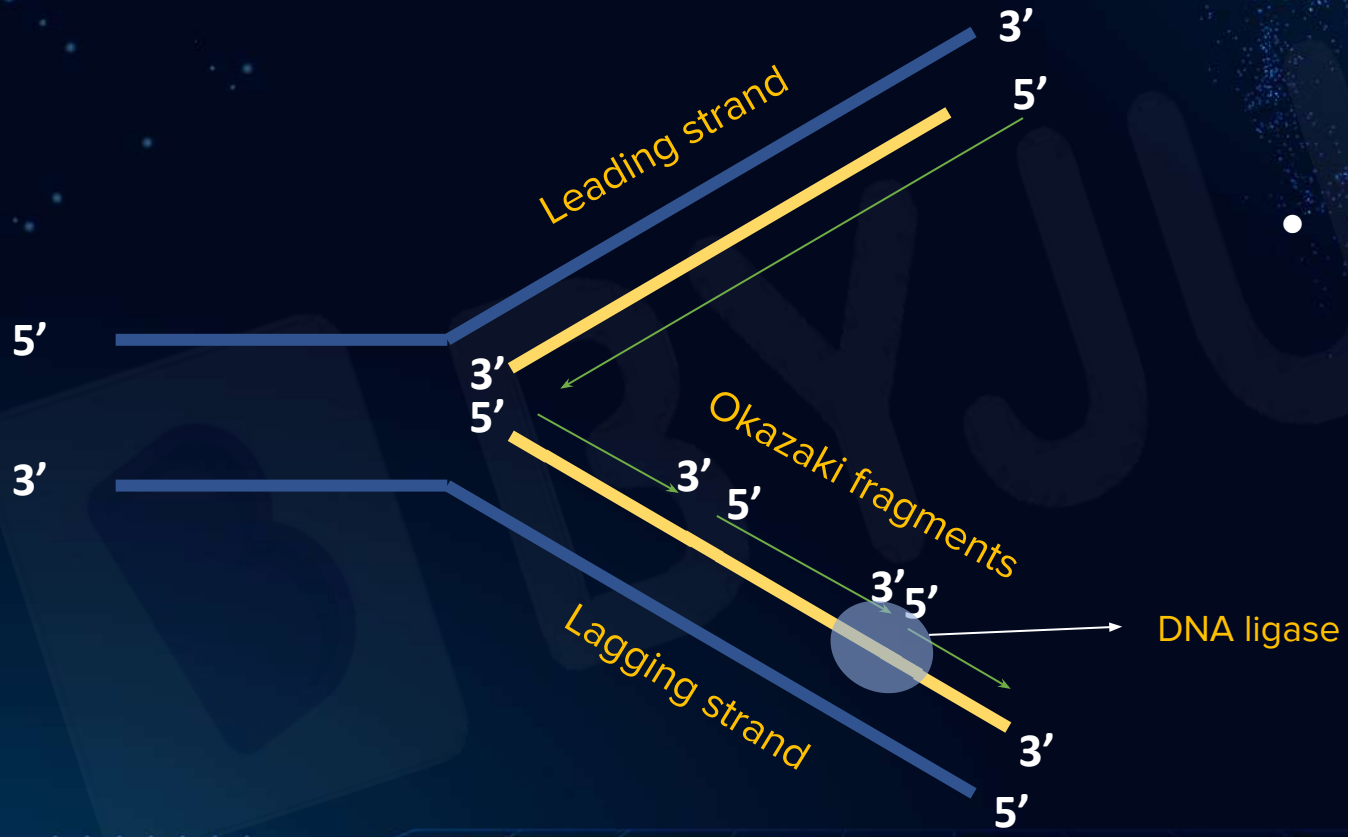
- Since the strand complementary to the lagging strand is discontinuous, instead of one large DNA fragment, many smaller DNA fragments are formed.
- These are called Okazaki fragments.

Process of Replication



- Since the lagging strand has RNA primers, which are different from the DNA nucleotides, the RNA primers are removed.

Process of Replication



- Finally, the gaps between the Okazaki fragments are joined together with the help of a DNA ligase enzyme.

Summary

Process of replication

Starts at origin of replication.

DNA helicase separates the two strands forming replication fork.

Primase binds at the replication fork.

Primase synthesises RNA primers.

DNA polymerase uses the primer to synthesise the two strands.

When using the leading strand as a template DNA polymerase adds nucleotides to the new strands continuously.

When using the lagging strand as a template, DNA polymerase adds nucleotides in short stretches called okazaki fragments to overcome the directionality problem (As the strand runs in 5'-3' direction and the enzyme can only run in 5'-3' direction).

The RNA primers are removed and the DNA sequences are synthesised by another DNA polymerase.

DNA ligase fills in the small gaps between all the fragments.

BYJU'S Classes Notes

Molecular Basis of Inheritance

Transcription, RNA World, Replication vs Transcription

Key Takeaways

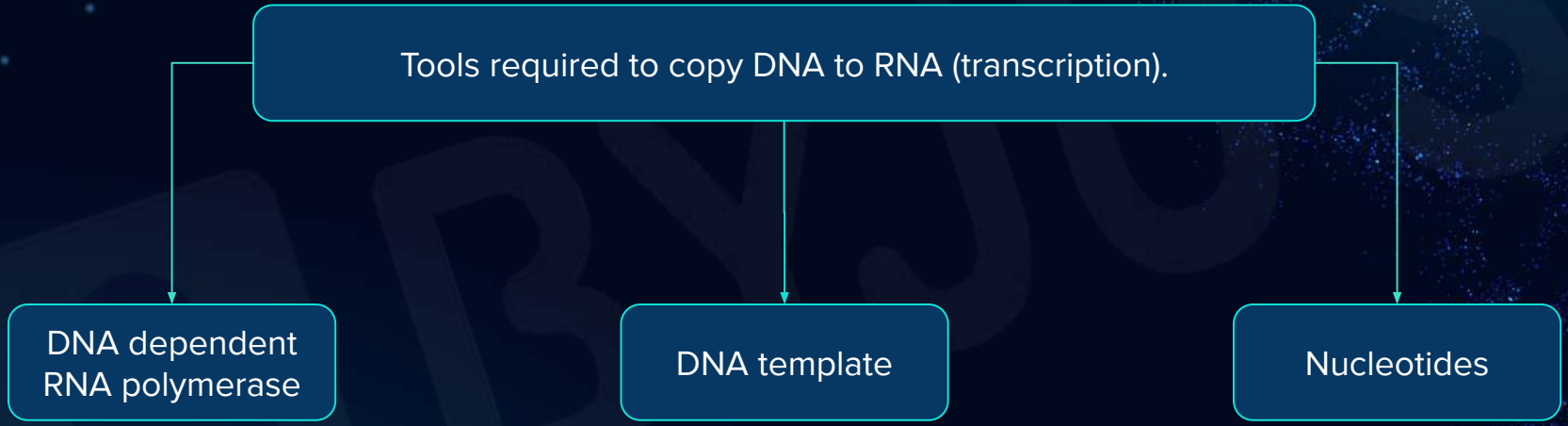
1. Transcription
   - Tools of transcription
   - Prokaryotic transcription
2. RNA world
   - Types of RNA
3. Replication vs transcription

Summary

Transcription

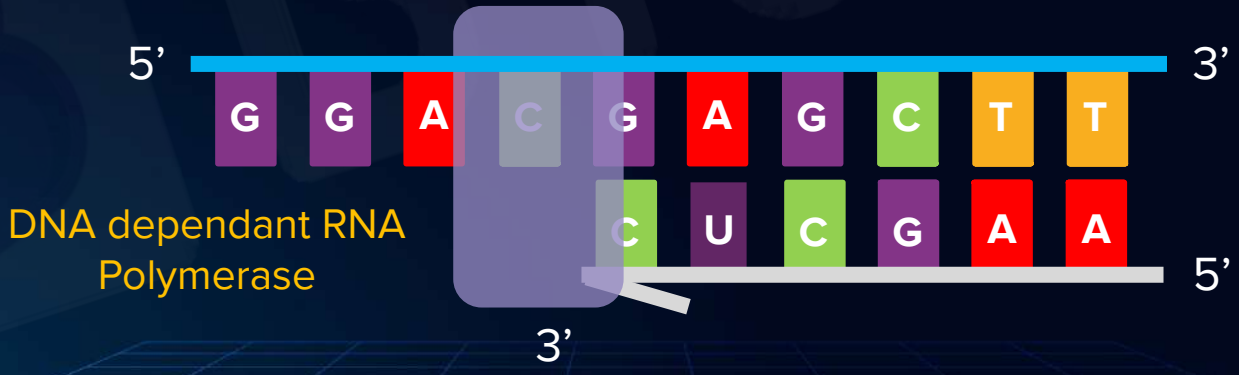
- The process of copying genetic information from DNA into RNA is called transcription.
- DNA has the information while proteins carry out the functions.
- So, this information from DNA is encoded and sent to other parts of the cell through a messenger as DNA cannot be sent out from the nucleus.
- RNA acts like this messenger. It relays the message from DNA to proteins.

Tools for Transcription



DNA Dependent RNA Polymerase

- Transcription is carried out by an enzyme known as DNA dependent RNA polymerase.
- DNA dependent in the name is simply because the enzyme uses DNA as a template.
- RNA polymerase catalyses transcription in the direction 5’ to 3’.
- RNA polymerase does not require a primer to initiate RNA synthesis.
- RNA polymerase adds uracil in place in thymine.



DNA Template

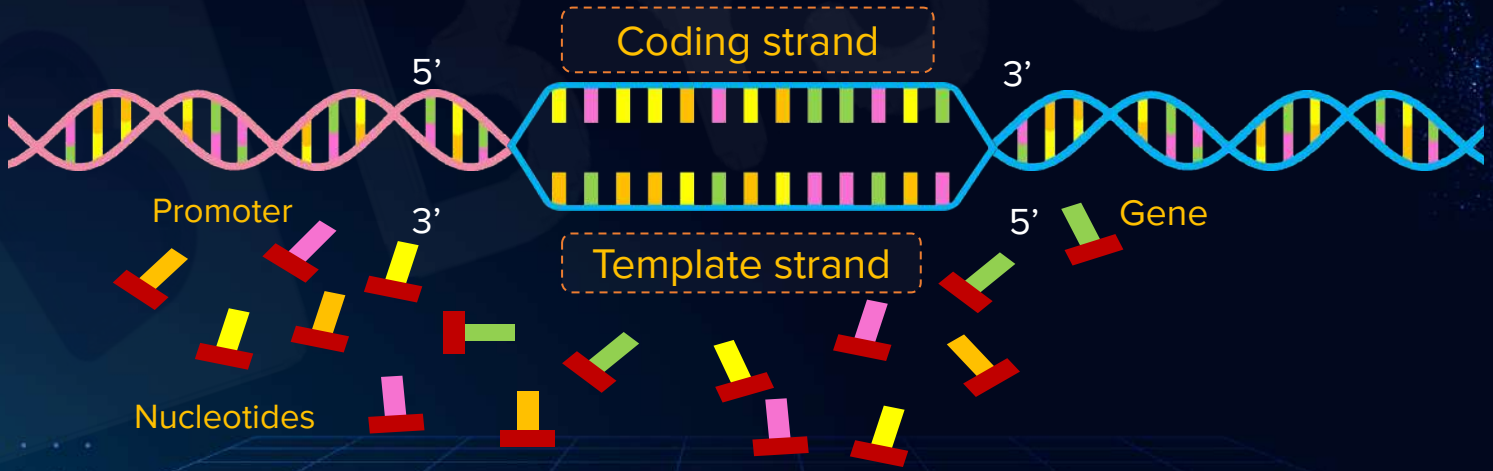
- A DNA segment from the template strand gets copied into RNA during the process.
- This segment of DNA getting transcribed is called a gene.
- In transcription, only one strand of DNA gets copied into RNA.
- The strand of DNA getting copied is called the template strand.

DNA Template

- RNA polymerase can add nucleotides from the 5’ to 3’ direction.
- Hence, the 3’-5’ DNA strand is called the template strand while the other 5’-3’ DNA strand is known as the non template strand.
- RNA sequence is similar to sequence of non template strands. The only thing that differs is T is replaced by U in RNA.
- 5’-3’ DNA strand is also called coding strand while 3’-5’DNA strand also called non coding strand.

Nucleiotide

- The third requirement during transcription is that of nucleotides.
- RNA is made of four nucleotides - Adenine (A), Guanine (G), Cytosine (C) and Uracil (U).
- RNA polymerase picks up the nucleotides and starts adding them with DNA as a template based on the complementarity rule.



Transcription of One Strand

In transcription only one strand is transcribed to mRNA

Reasons

Complications in genetic information transfer

- If the both strands act as a template and transcribe, then they would code for RNA molecules with different sequences from the same DNA molecule.
- In turn, if these RNAs code for proteins, the two RNA sequences would code for two different proteins.
- Hence, one segment of the DNA would be coding for two different proteins complicating the genetic information transfer machinery.

Complementary RNA molecules

- The two RNA molecules produced would be complementary to each other, hence would form a double stranded RNA.
- This would prevent RNA from being converted into protein.

Transcription Units

Gene along with other necessary sequences are combinedly known as transcription units.

Transcription unit mainly consists of three regions of DNA: structural gene, promoter and terminator sequence.

Structural genes

a) Structural genes have the nucleotide sequence which codes for RNA molecules.

b) The final RNA molecule will have a sequence complementary to that region.

Promoter

a) Promoter is the region which is present upstream of the coding strand of the gene at the 5’ end.

b) The DNA sequence to which RNA polymerase binds to initiate transcription of a gene.

Terminator sequence

a) Terminator region is present downstream or 3’ end of coding strand of gene.

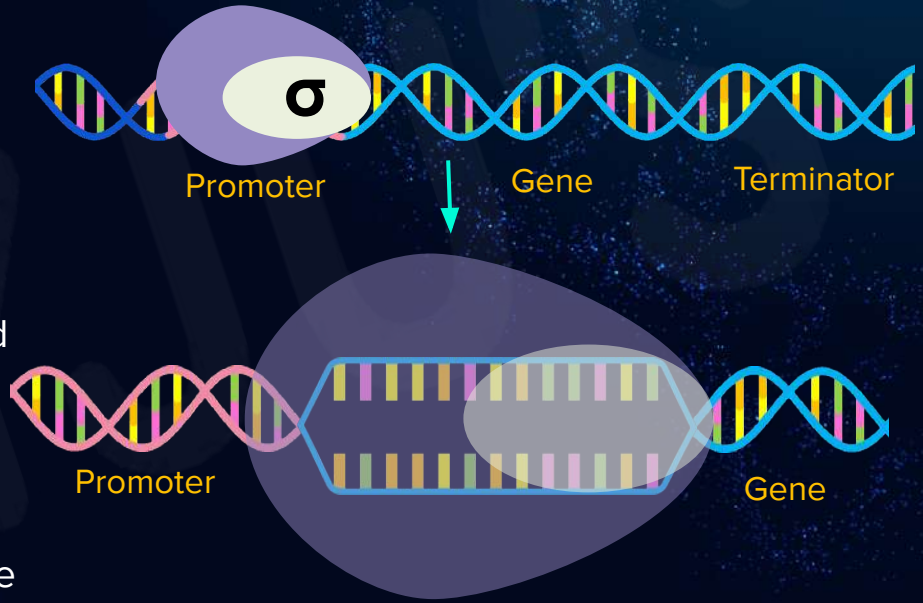
b) When RNA polymerase reaches this site, transcription ends and RNA is released.

Transcription in Prokaryotes

Transcription steps

Initiation

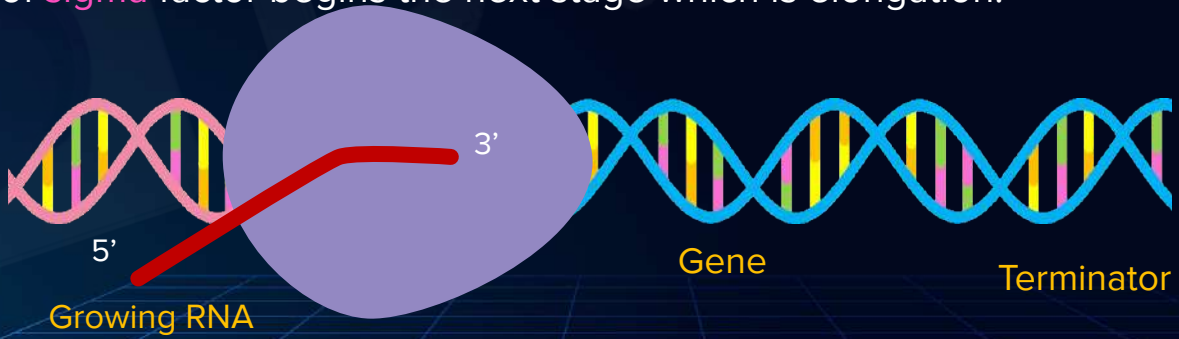
- Transcription requires DNA as a template and RNA polymerase.
- However, to initiate the transcription, RNA polymerase requires one more element - initiation or sigma factor.
- The initiation factor or sigma factor directs the RNA polymerase towards the promoter region and helps in binding to the promoter.
- RNA polymerase and sigma factor move forward toward the genes and unwind the DNA strands around the transcription start site.



Transcription in Prokaryotes

Elongation

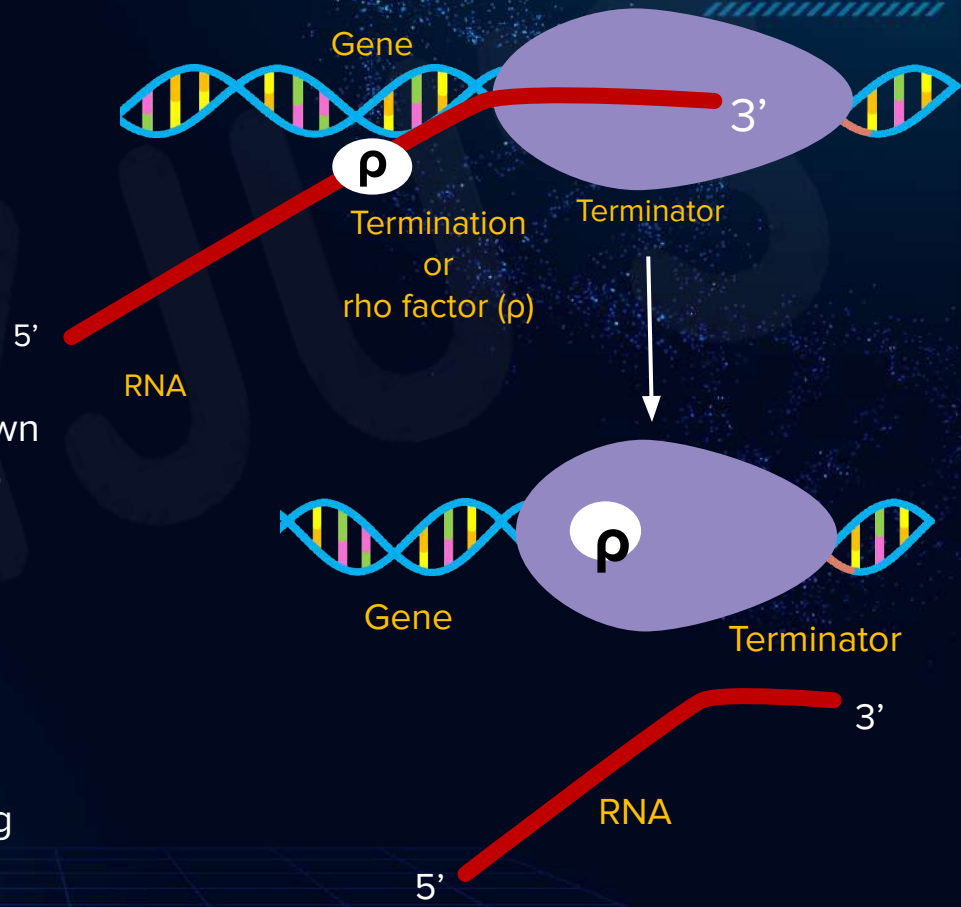
- During elongation, RNA synthesis continues.
- As RNA polymerase moves across, only a small part of RNA is attached to the template strand, the other end of growing RNA is free and detached from the DNA template.
- Size of RNA eventually grows as and when new nucleotides are added.
- The polymerase moves ahead, the unwound DNA gets rewinded.
- After adding a few bases, the sigma factor is released from RNA polymerase.
- Release of sigma factor begins the next stage which is elongation.



Transcription in Prokaryotes

Termination

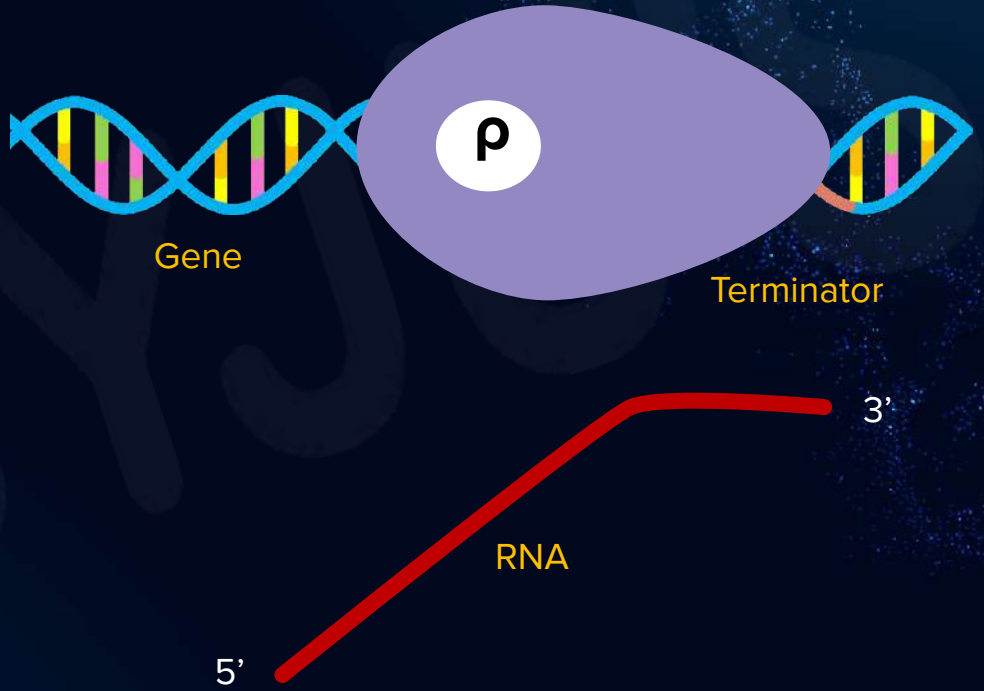
- Termination begins when another factor known as termination factor or rho (ϱ) factor binds to RNA.
- Once bound to RNA, the Rho (ϱ) factor starts ascending upwards and finally reaches the region where RNA is bound to the DNA template strand.
- After reaching this region, the rho (ϱ) factor unwinds the DNA-RNA complex, thus causing the release of RNA from RNA polymerase.



Transcription in Prokaryotes

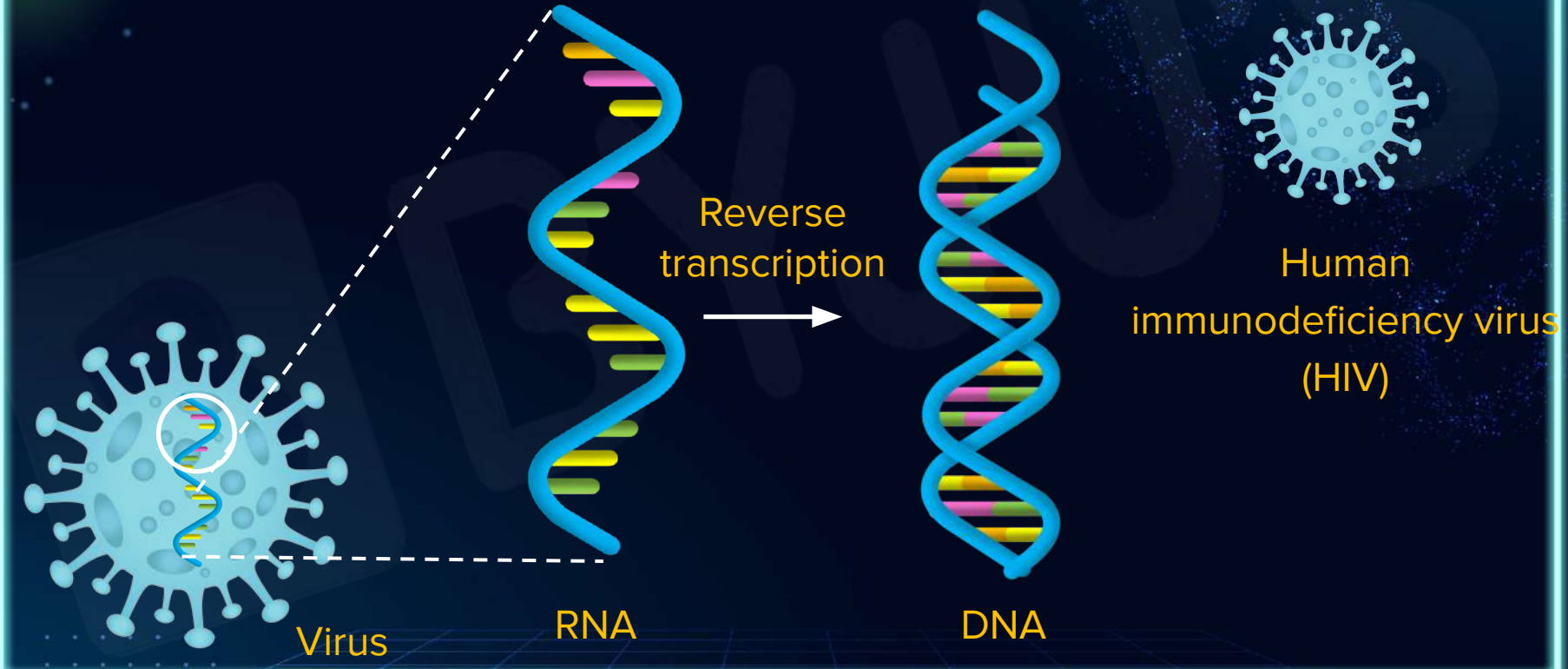
Termination

- Once the RNA is released, RNA polymerase and rho (ρ) factor comes off of DNA strand.
- DNA gets rewinded back into helical structure. This marks the end of RNA transcription.



Did You Know?

Some viruses have the ability to transcribe RNA into DNA.
This process is known as reverse transcription.



Transcription termination requires .....

a) Initiation Factor

b) Sigma Factor

c) rho Factor

d) All of the above

Transcription termination requires .....

a) Initiation Factor

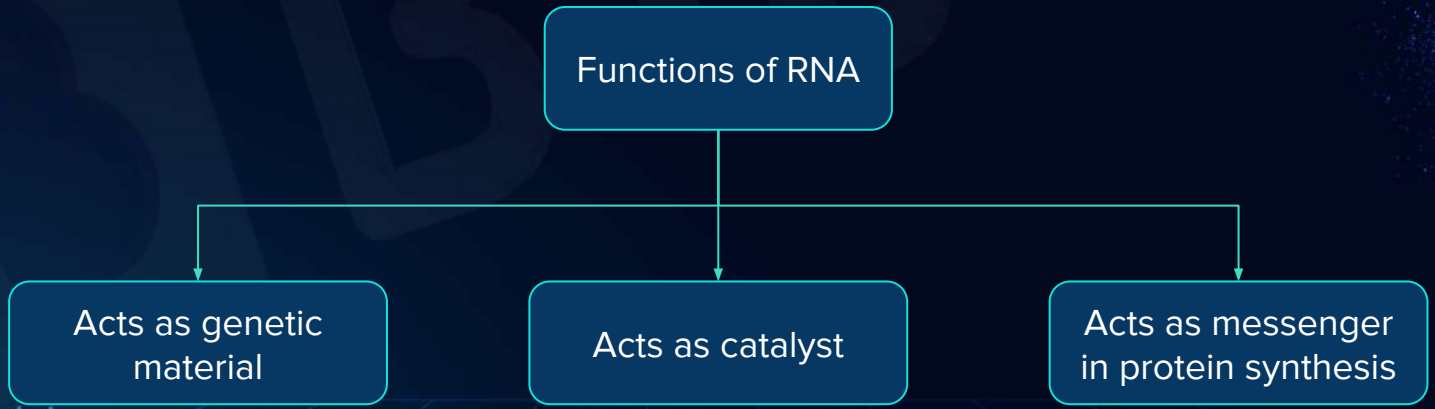
b) Sigma Factor

c) rho Factor

d) All of the above

RNA World

- Ribonucleic acid (RNA) is a nucleic acid which is directly involved in protein synthesis.
- It is an important nucleotide with long chains of nucleic acid present in all living cells.
- RNA contains the sugar ribose, phosphates, and the nitrogenous bases adenine (A), guanine (G), cytosine (C), and uracil (U).



RNA World

Types of RNA

mRNA

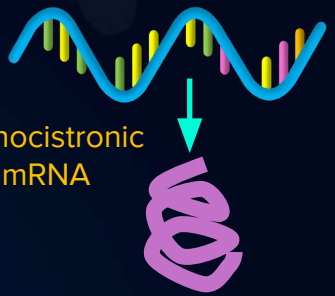
- Carries message from DNA
- Template for protein synthesis
- Can have one or more cistron (a segment of DNA coding for a polypeptide)

mRNA

Cistron : Nucleotide sequence which codes for single protein

Monocistronic

- Contains single cistron
- Found in eukaryotes

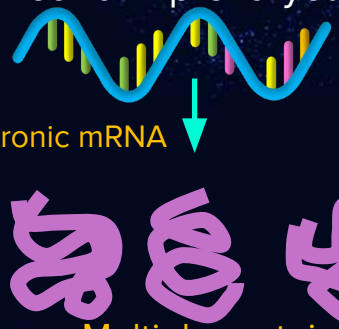


Monocistronic mRNA

Single protein

Polycistronic

- Contains multiple cistrons
- Found in prokaryotes



Polycistronic mRNA

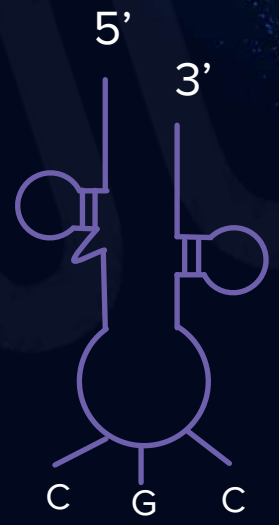
Multiple proteins

RNA World

Types of RNA

tRNA

- Carries correct amino acids to site of protein synthesis.

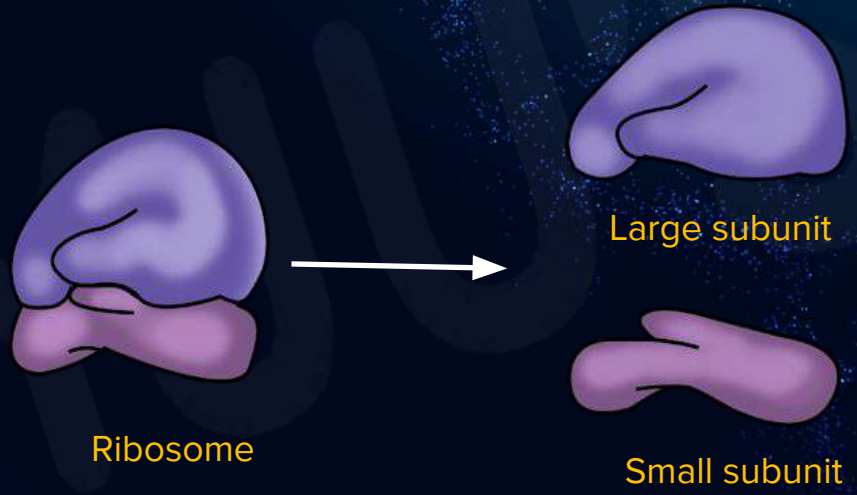


Clover leaf shape structure

RNA World

Types of RNA

- rRNA
  - Present in ribosomes.
  - Helps in catalysing protein synthesis.



Which of these have clover leaf structure ?

a) mRNA

b) tRNA

c) rRNA

d) DNA

Which of these have clover leaf structure ?

a) mRNA

b) tRNA

c) rRNA

d) DNA

Replication vs Transcription

Replication	Transcription
Entire genomic DNA is copied	Only a segment of DNA (gene) is copied
Both the strands of DNA are copied	Only the template strand is copied
Needs DNA dependent DNA polymerase	Needs DNA dependent RNA polymerase
Requires primers	Does not require primers
Product is double stranded DNA	Product is single stranded RNA

Summary

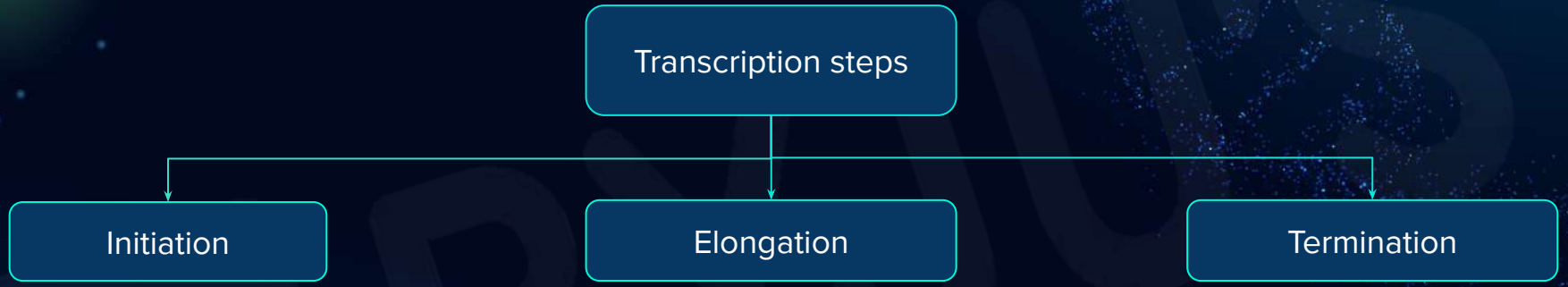
Tools required to copy DNA to RNA (transcription).

- DNA dependent RNA polymerase
- DNA template
- Nucleotides

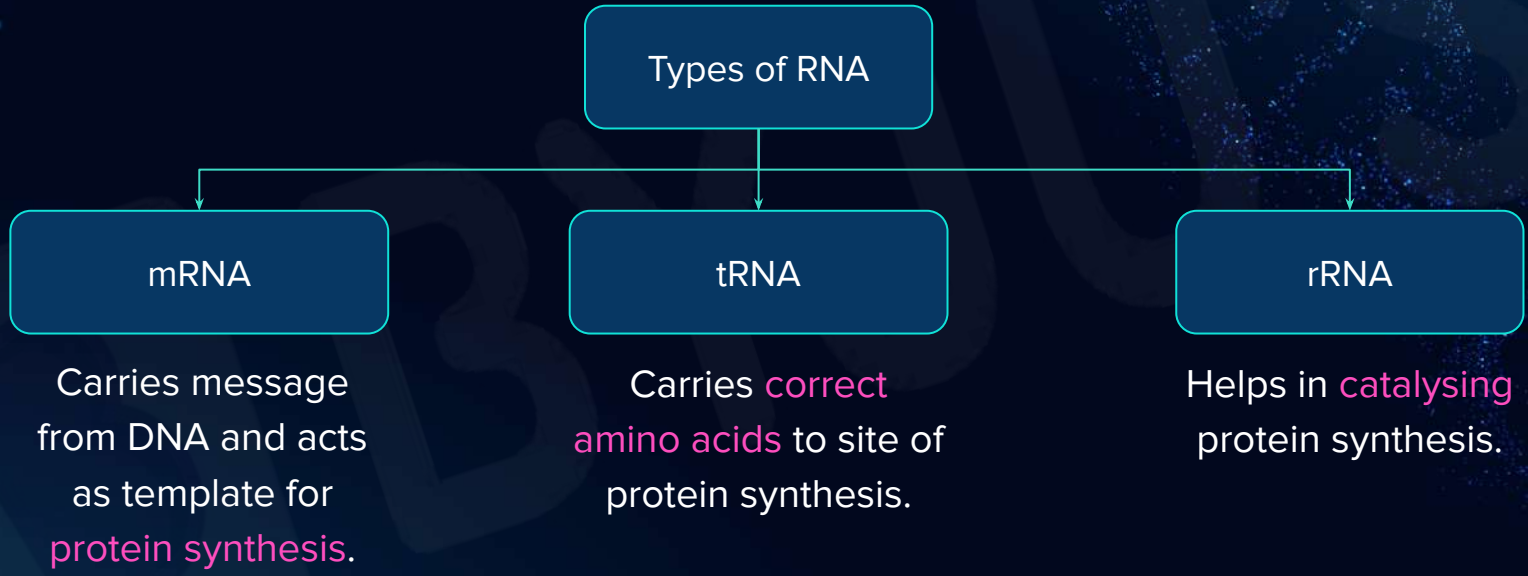
Transcription unit mainly consists of three regions of DNA: structural gene, promoter and terminator sequence.

- Structural genes
- Promoter
- Terminator sequence

Summary



Summary



BYJU'S Classes Notes

Molecular Basis of Inheritance

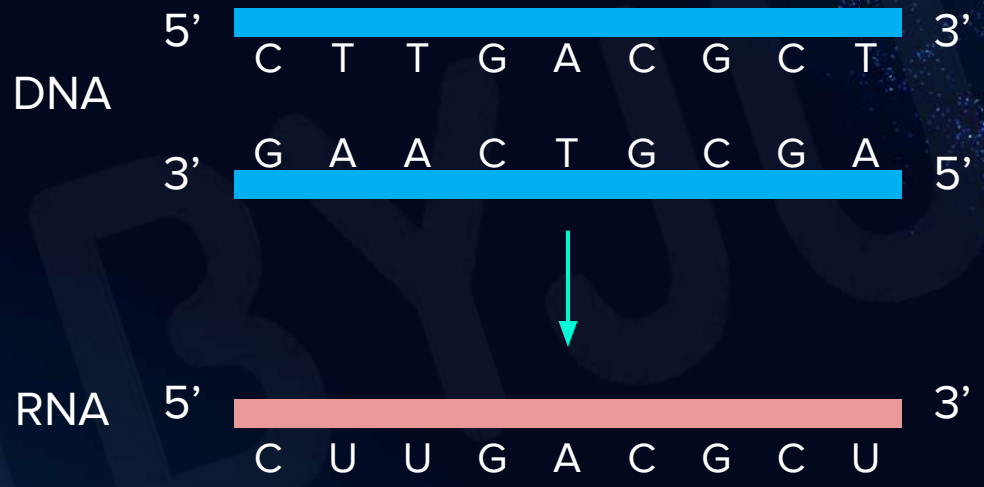
Eukaryotic Transcription, Post-transcriptional Modifications

Key Takeaways

1. Eukaryotic transcription
   - Initiation
   - Elongation
   - Termination
2. Post-transcriptional modifications
   - Splicing
   - Capping
   - Tailing

Summary

Recall! Transcription



- The process of copying genetic information from DNA into RNA is called transcription.

Recall! Prokaryotic Transcription

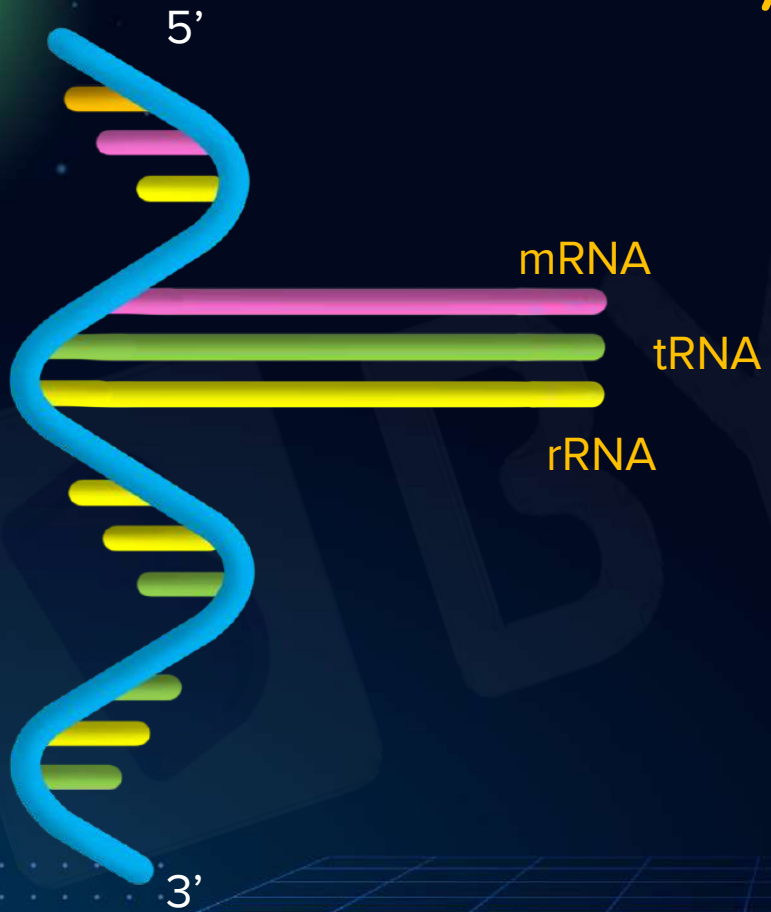
- RNA polymerase along with sigma factor attaches to the DNA molecule and recognises a promoter sequence, and initiates new mRNA synthesis.
- RNA polymerase encounters a terminator sequence after elongation, thus causing the release of RNA from RNA polymerase with the help of rho factor.

Recall! RNA Polymerase

DNA dependent RNA polymerase:

- Catalyses in the direction 5’ → 3’
- Does not require primer to initiate RNA synthesis
- Adds uracil instead of thymidine

Recall! Types of RNA



rRNA or ribosomal RNA

- Present in ribosomes
- Helps in catalysing protein synthesis

Recall! Ribosomes

Ribosome

- Synthesise proteins
- Made of RNA and proteins

Large subunit

28S

5.8S

5S

Small subunit

18 S

- These three different rRNA molecules together form the large subunit.

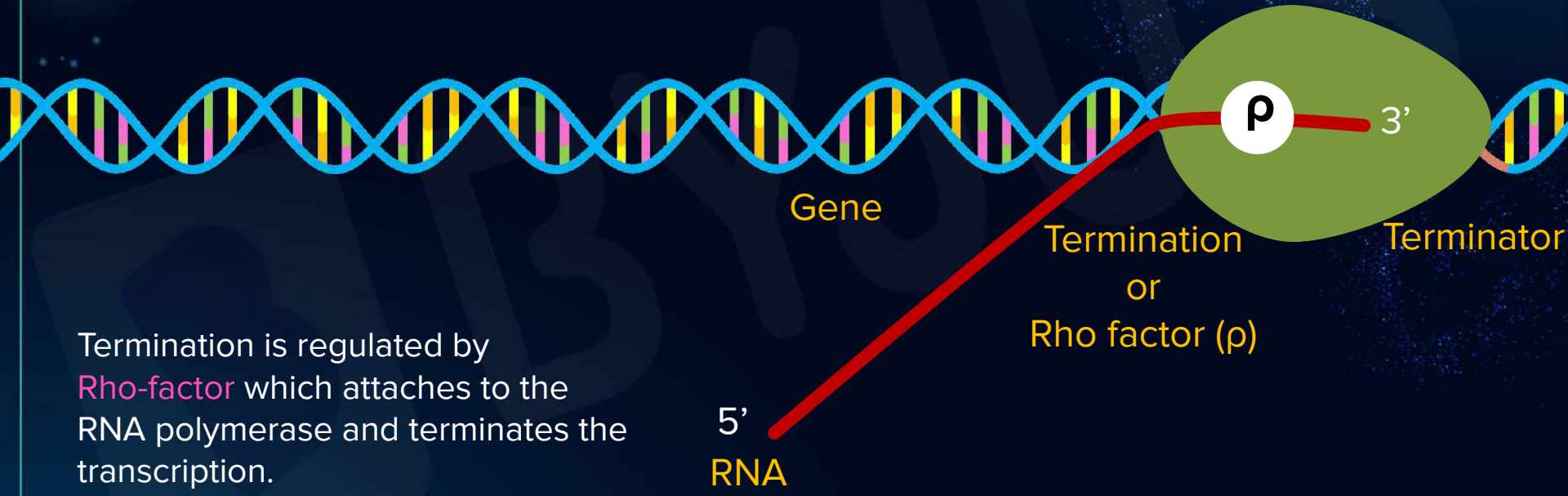
- rRNA forms the ribosomal units
- Ribosome has two subunits - large subunit and small subunit

Recall! Transcription Unit

Transcription unit

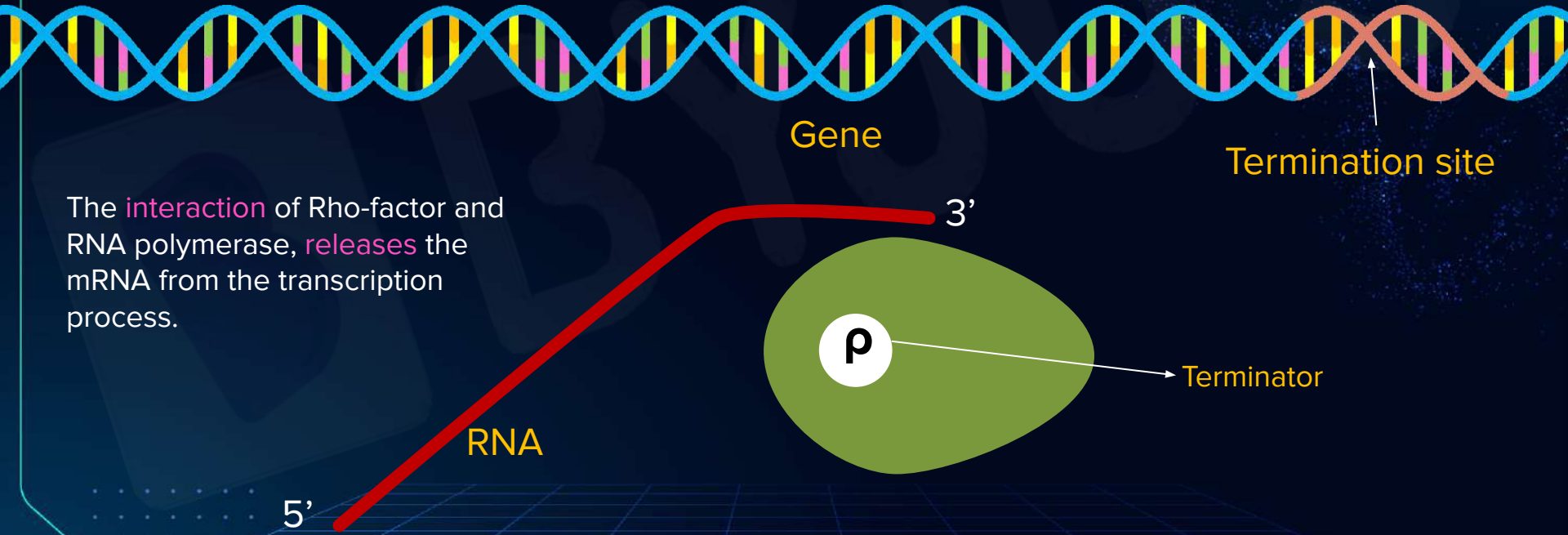
- Gene
- Promoter sequence
- Terminator sequence

Recall! Termination in Prokaryotes



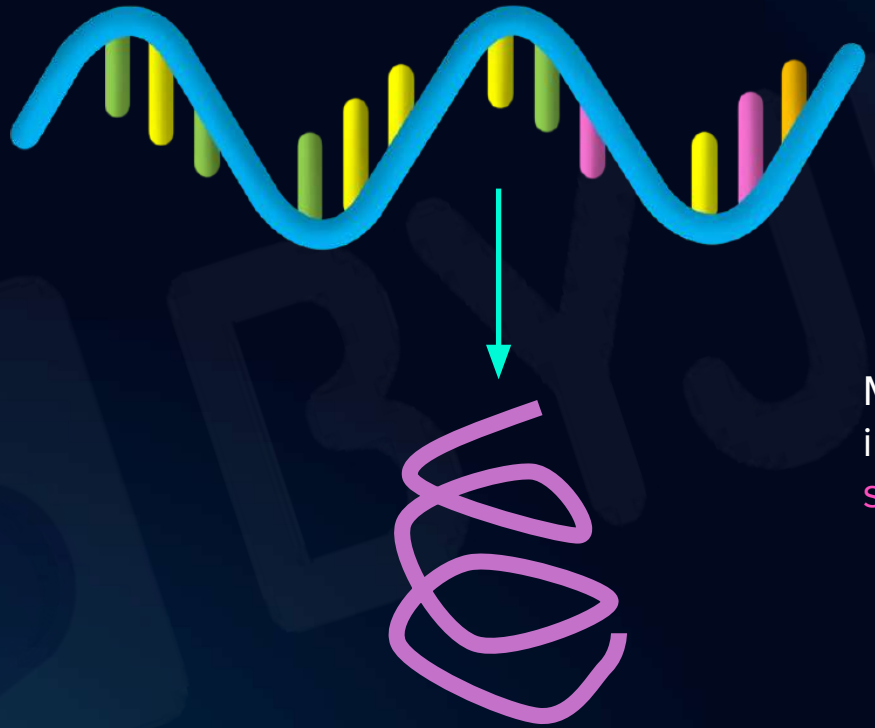
Termination is regulated by Rho-factor which attaches to the RNA polymerase and terminates the transcription.

Recall! Termination in Prokaryotes



The interaction of Rho-factor and RNA polymerase, releases the mRNA from the transcription process.

Recall! Monocistronic mRNA



Monocistronic mRNA
(Eukaryotes)

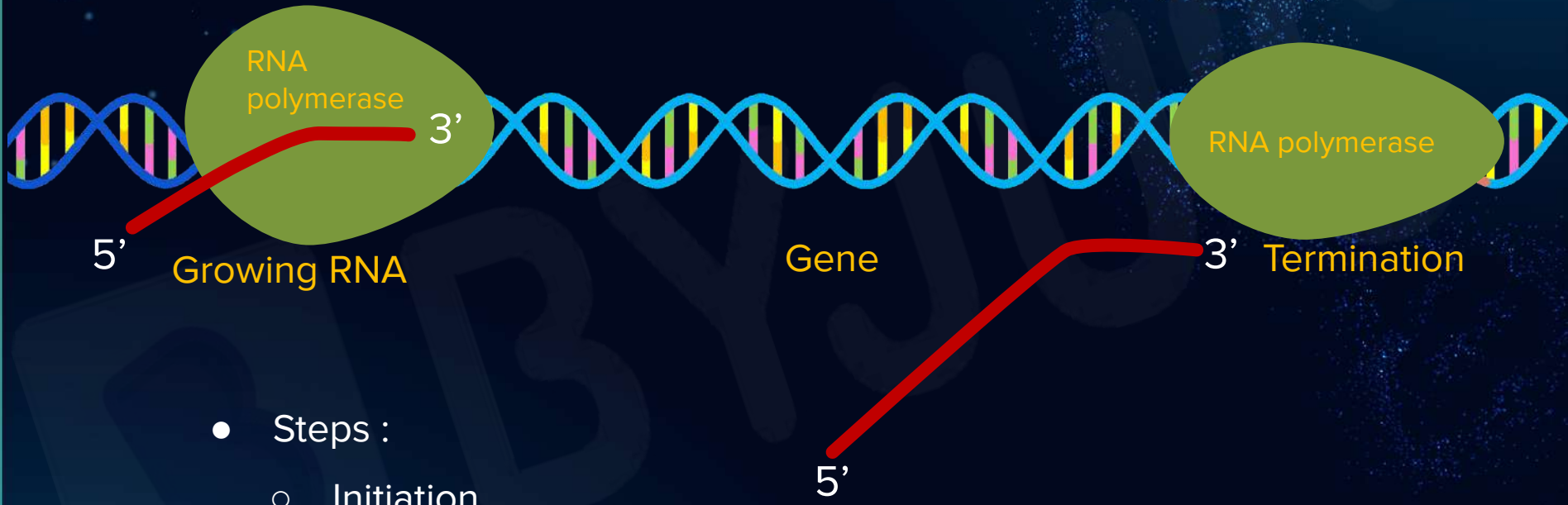
Monocistronic mRNA has information for synthesis of single protein only.

Single protein

Eukaryotic Transcription

- The process of copying genetic information from DNA into RNA is termed as transcription.
- It occurs in the nucleus of eukaryotic cells.
- Each type of RNA is synthesised by a different RNA polymerase.
- Transcription is not immediately followed by translation because transcription and translation sites differ.
- It produces heterogeneous RNA (hnRNA) as these newly formed RNAs contain many introns (non-coding regions).
- Instability of the mRNA makes it compulsory for RNA to be processed through capping, polyadenylation and splicing.

Eukaryotic Transcription



- Steps :
  - Initiation
  - Elongation
  - Termination

Eukaryotic RNA Polymerase

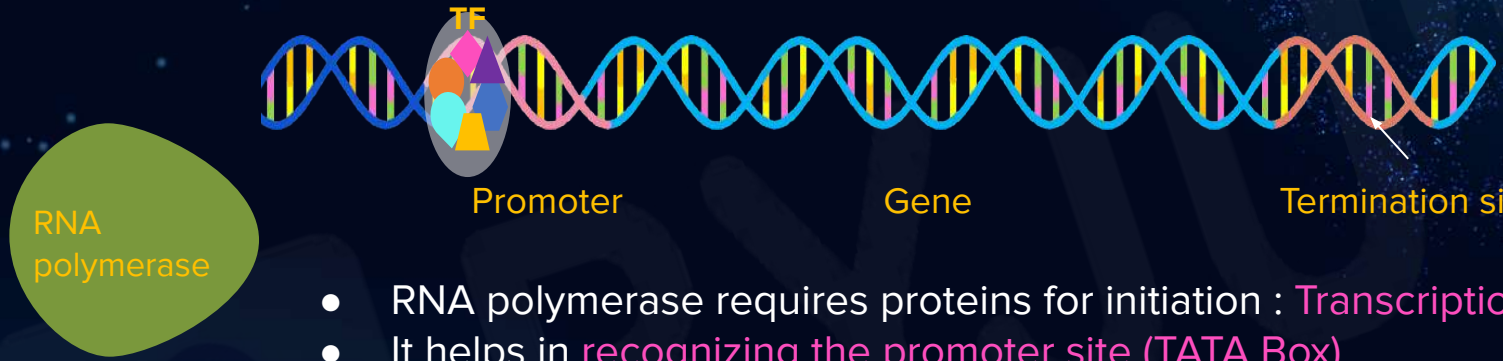
- Eukaryotes have 3 RNA polymerases depending on the type of RNA to be transcribed.

RNA polymerase	Type of RNA
RNA polymerase I	rRNA (28S, 18S and 5.8S)
RNA polymerase II	mRNA
RNA polymerase III	tRNA, 5S rRNA, snRNA

- Apart from the 3 RNA’s (mRNA, tRNA and rRNA), there is a 4th type of RNA found only in eukaryotes called snRNA.
- It is involved in the processing of pre-messenger RNA (pre-mRNA) into mature mRNA.

Eukaryotic Transcription

Initiation

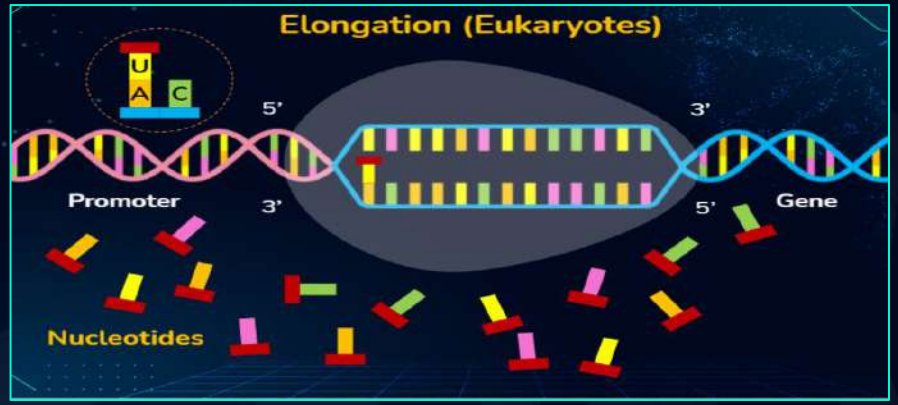


- RNA polymerase requires proteins for initiation : Transcription factors (TF).
- It helps in recognizing the promoter site (TATA Box).

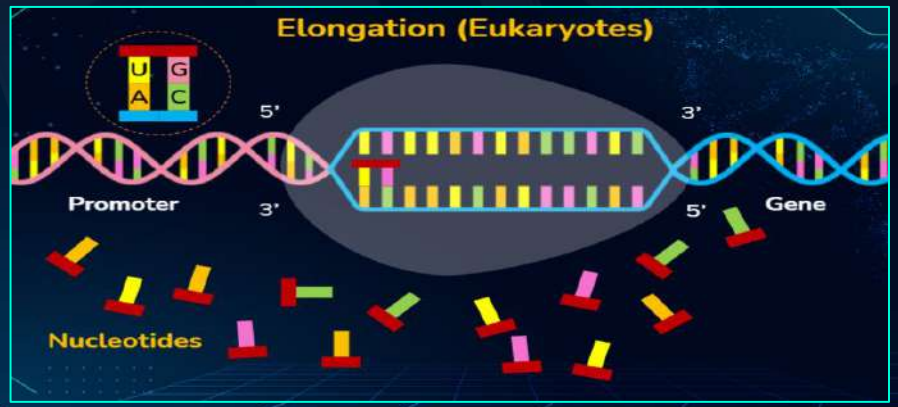
- Once the TF recognizes the promoter region, the RNA polymerase binds it and starts the transcription process.

Eukaryotic Transcription

Elongation



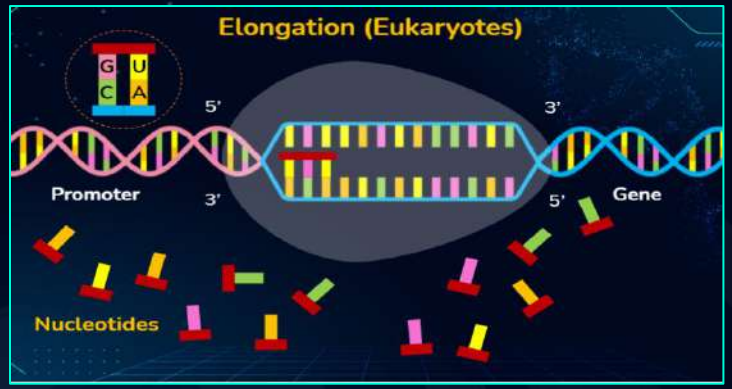
1. In RNA, RNA polymerase adds uracil over adenine.



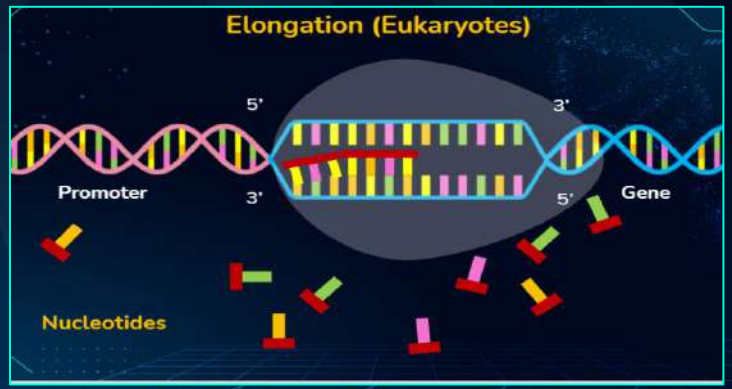
2. Next, nucleotide on coding strand is cytosine; so RNA polymerase picks up guanine and adds it next to the uracil.

Transcription (Eukaryotes)

Elongation



3. Third nucleotide in coding strand is adenine. So third base added on RNA will be uracil again.

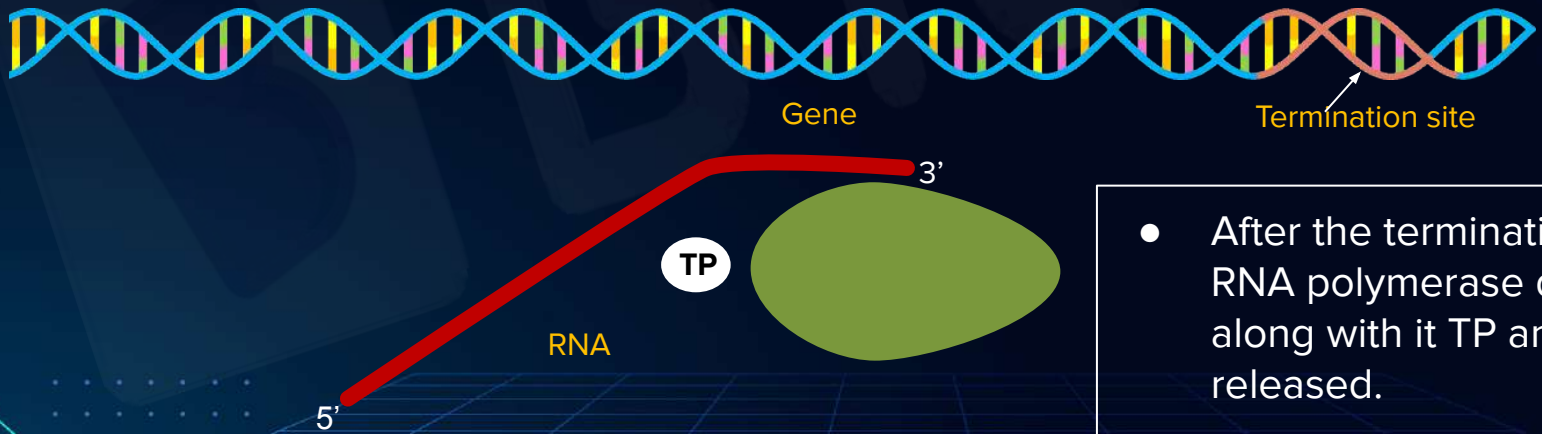


4. Similarly, RNA polymerase keeps adding bases as per base complementarity.
Size of RNA eventually grows as and when new nucleotides are added.

Transcription (Eukaryotes)

Termination

- A termination protein stops the transcription process.



- After the termination, the RNA polymerase dissociates, along with it TP and RNA are released.

Post – Transcriptional Modifications

RNA formed :

- Called primary transcript or nascent RNA
- Non-functional
- Undergoes modifications to form functional mRNA, rRNA, tRNA
- Monocistronic

Post – Transcriptional Modifications

RNA formed :

- The primary transcript of mRNA is called hnRNA or heterogeneous nuclear RNA.
- Since it has not undergone modification, it is longer than the functional mRNA.

Post – Transcriptional Modifications

- Splicing
- Capping
- Tailing

Splicing

- The first step is the splicing .
  - It is the process of removal of extra segments from the RNA transcript.

hnRNA

- The primary transcript in eukaryotes has introns and exons.
- Introns are noncoding/ non functional/ intervening sequences in the RNA.
- Exons are the coding/functional sequence in the RNA that code for proteins.

Did You Know?

- Introns were earlier called junk DNA.
- But some believed there must be reason for preserving introns over evolutions though they did not serve any purpose.
- That is when scientists started scrutinizing introns and they are shown to be useful in a number of ways but they do not encode proteins.

Splicing

- Splicing involves removal of the introns and joining of the exons.
- The splicing reaction is catalyzed by a large protein complex called the spliceosome (snRNA+proteins).

- This leads to the bending of the introns and hence, brings exons close to each other.

Splicing

hnRNA

- Removal of introns and joining of exons is mediated by spliceosome (snRNA+proteins).

hnRNA

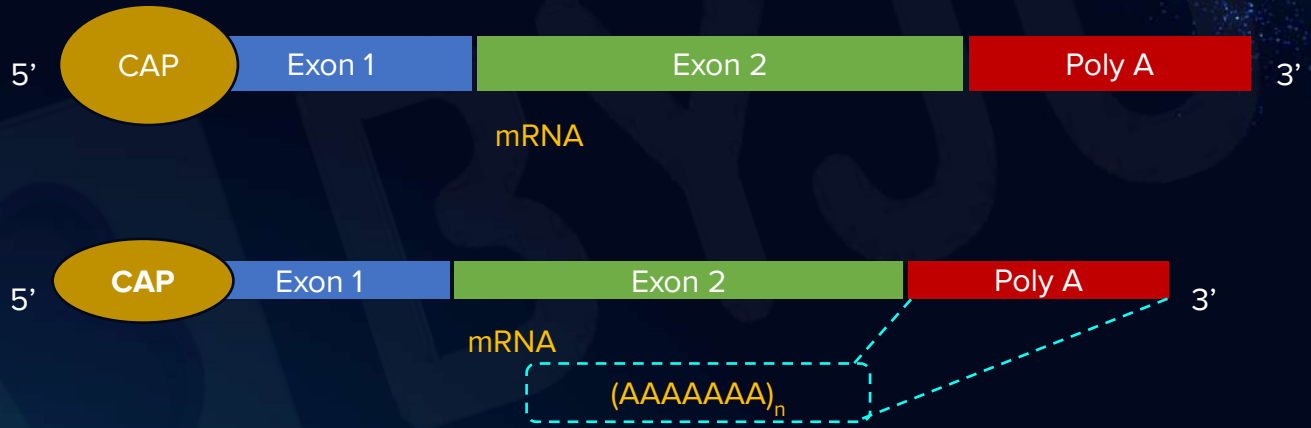
Capping

- It is the addition of an unusual nucleotide on 5’ end of the mRNA.

- Methylated guanosine triphosphate (mGppp), is an example of unusual nucleotide.
- This capping mainly provides protection to the mRNA against phosphatases and other nucleases.

Tailing

- It is the addition of a poly A tail at 3’ end.
- In tailing, adenylate residues (200-300) are added at 3'-end in a template independent manner.

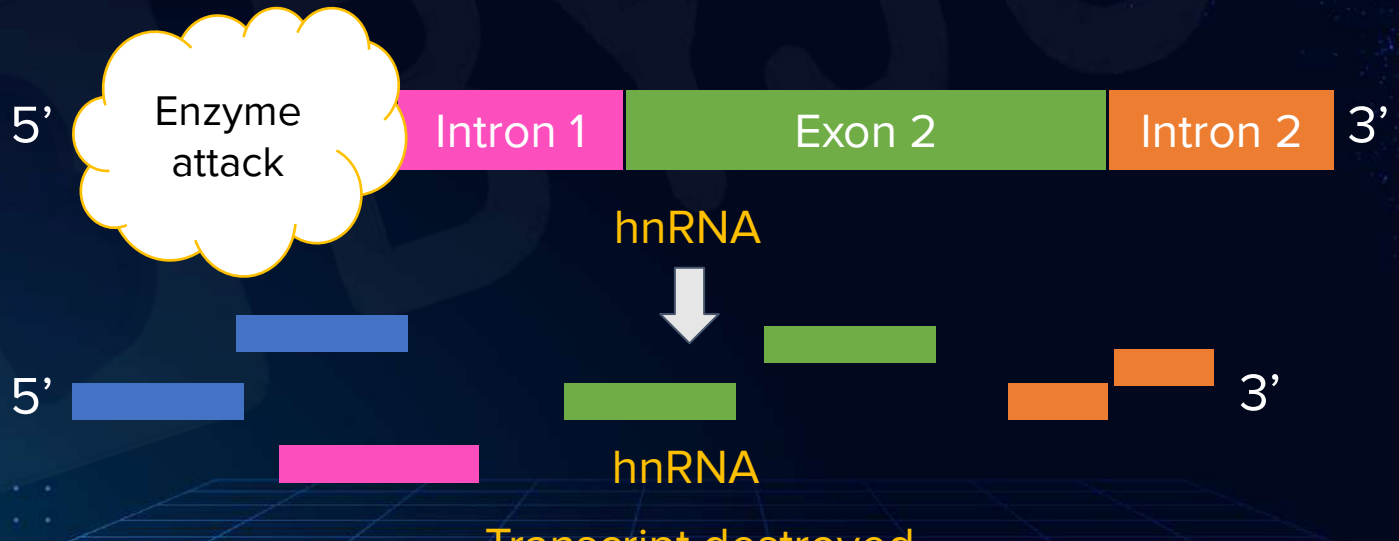


- After all these modifications, finally the hnRNA becomes a functional mRNA.
- Following this, mRNA leaves the eukaryotic nucleus.

Post – Transcriptional Modifications

Advantages:

- Capping and tailing protect transcript from enzyme attack.
- Modifications help the RNA molecule to be recognized by molecules that mediate RNA translation into proteins.



Prokaryotic v/s Eukaryotic Transcription

Prokaryotic transcription	Eukaryotic transcription
Occurs in cytoplasm	Occurs in nucleus
1 RNA polymerase	3 RNA polymerases
Needs sigma factor for initiation	Needs transcription factors for initiation
RNA formed is polycistronic	RNA formed is monocistronic
No modification required	Involves post-transcriptional modifications

What will be present in a processed eukaryotic RNA?

a) Introns

b) Exons

c) Both A and B

d) None of the above

What will be present in a processed eukaryotic RNA?

a) Introns

b) Exons

c) Both A and B

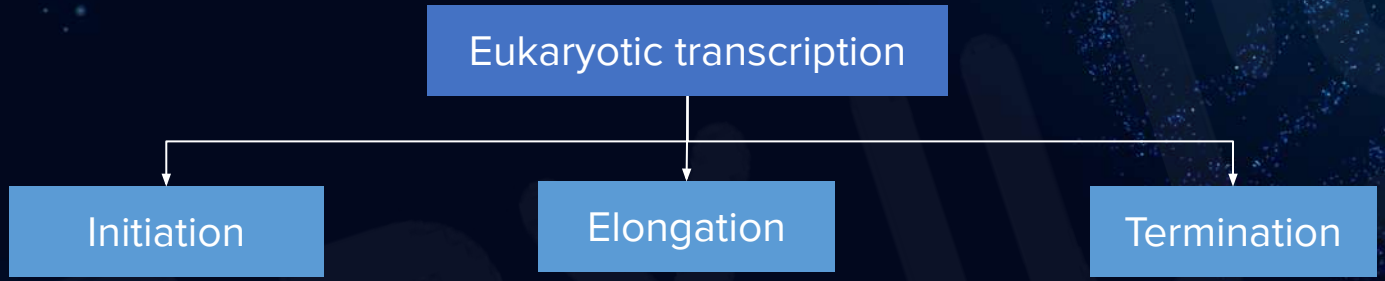
d) None of the above

Summary

RNA polymerase

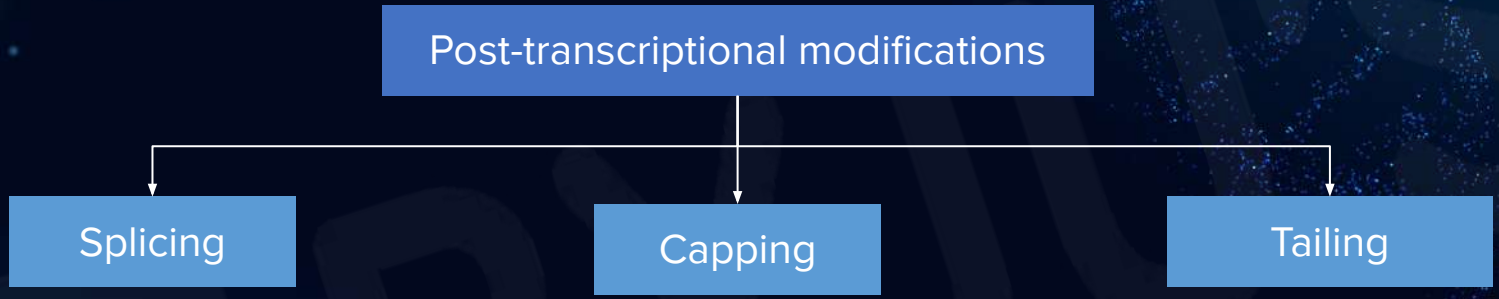
RNA polymerase	Type of RNA
RNA polymerase I	rRNA (28S, 18S and 5.8S, except 5S rRNA)
RNA polymerase II	mRNA
RNA polymerase III	tRNA, 5S rRNA, snRNA

Summary



- RNA polymerase requires proteins for initiation called transcription factors (TF).
- RNA polymerase keeps adding bases as per base complementarity.
- Termination protein stops the transcription process.
- RNA polymerase dissociates along with TP and RNA is released.

Summary



- Removal of introns and joining of exons
- Addition of an unusual nucleotide at 5’ end
- Adenylate residues (200-300) are added at 3'-end in a template independent manner

BYJU'S Classes Notes

Molecular Basis of Inheritance

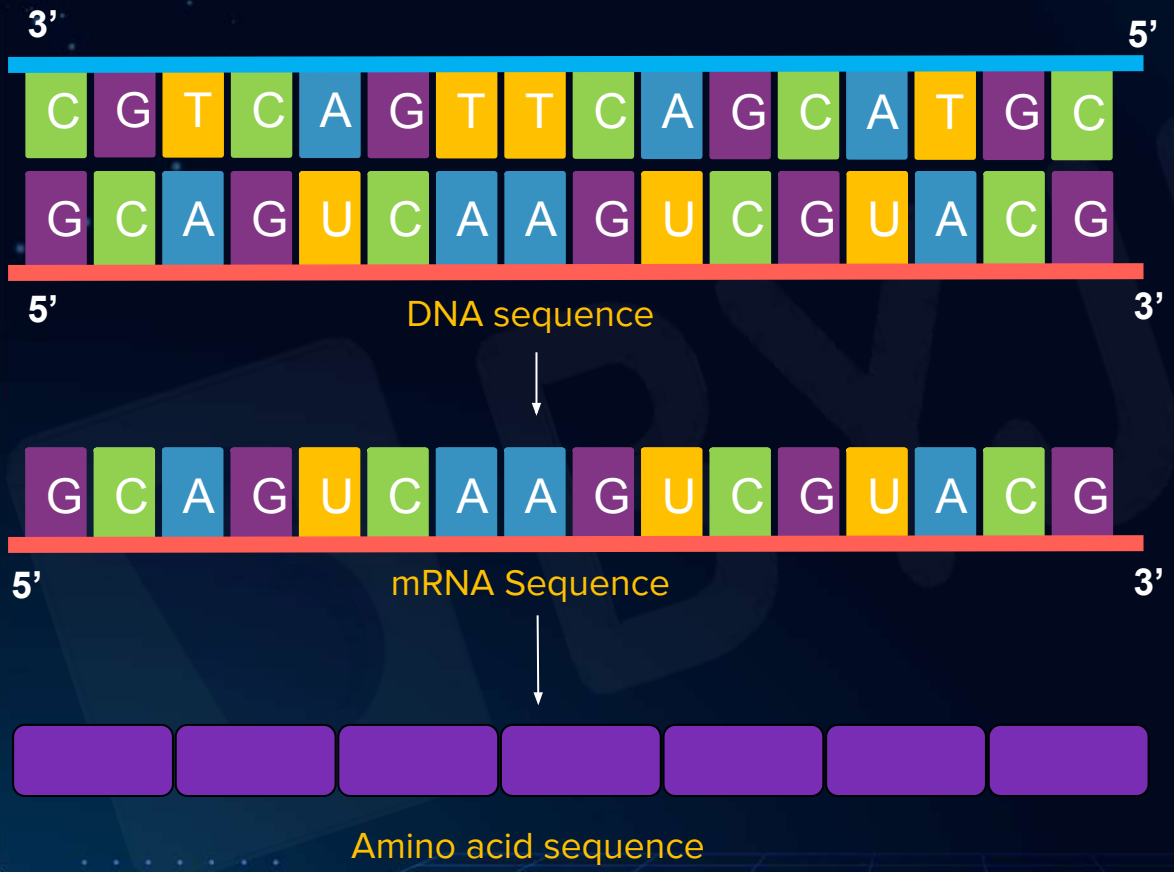
Discovery and Salient Features of Genetic Code, Mutations and Types

Key Takeaways

1. Genetic code
   - Discovery
   - Salient features
2. Mutations
   - Types

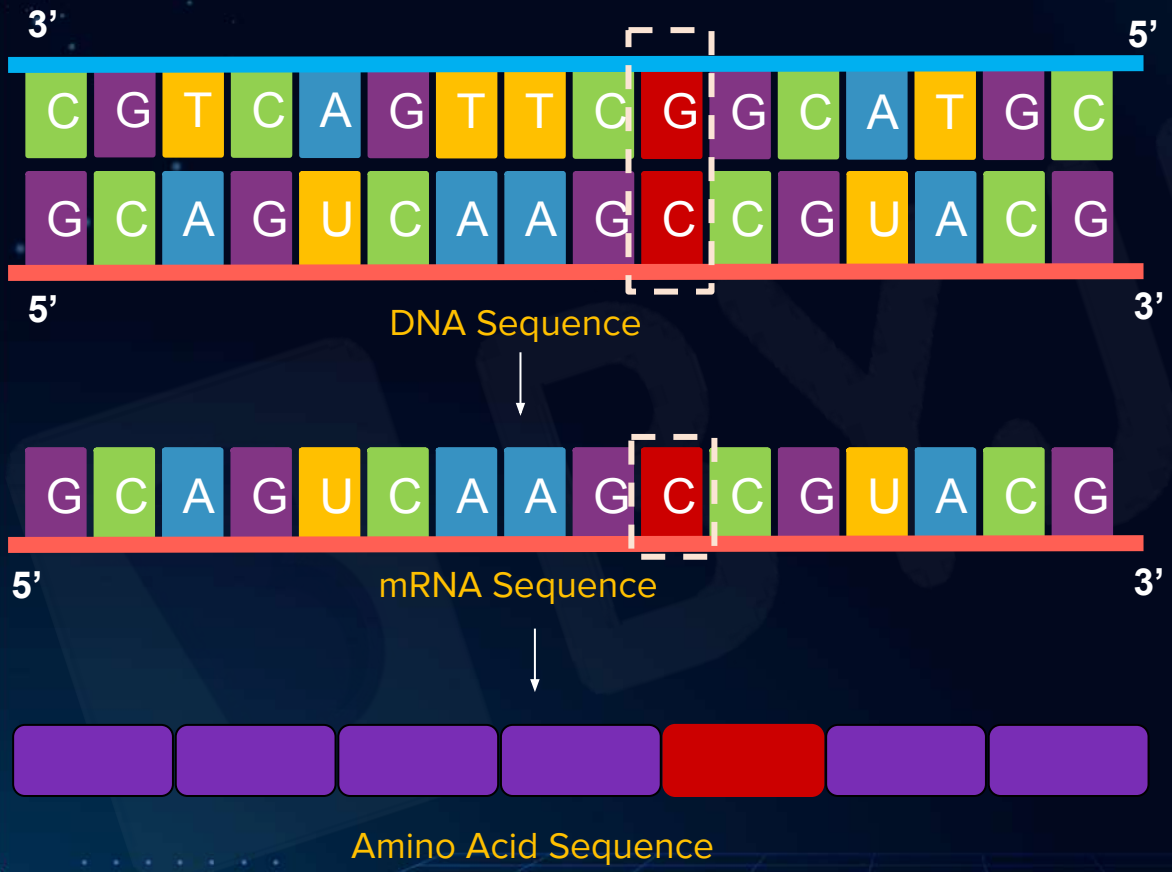
Summary

DNA to Proteins



- Nucleus in the cell contains the DNA.
- DNA is transcribed into mRNA, which is then sent out to the cytoplasm.
- The message in the RNA is then decoded to form proteins(amino acid sequence).

DNA to Proteins



- Scientists observed that a change in the DNA nucleotide sequence would lead to a change in the mRNA sequence which in turn would change the amino acid sequence of the protein too.
- They figured that there might be some relationship between the nucleotide sequence of DNA, mRNA and amino acid sequence of proteins.

Nucleotides to Amino Acids

G	U	A	C
Guanine	Uracil	Adenine	Cytosine

4 Bases

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

20 Amino acids

George Gamow

- It was George Gamow who first attempted to solve the problem of how nucleotide sequence might control the amino acid sequence of proteins.
- He made a wild guess that there are 20 amino acids which turned out to be correct.
- He argued that amino acids must constitute a combination of bases as there are just 4 bases and 20 amino acids.

Nucleotides to Amino Acids

If it is considered that:

- 1 base = 1 amino acid, then each base would code for 1 amino acid.
- So there would be 4 different amino acids.

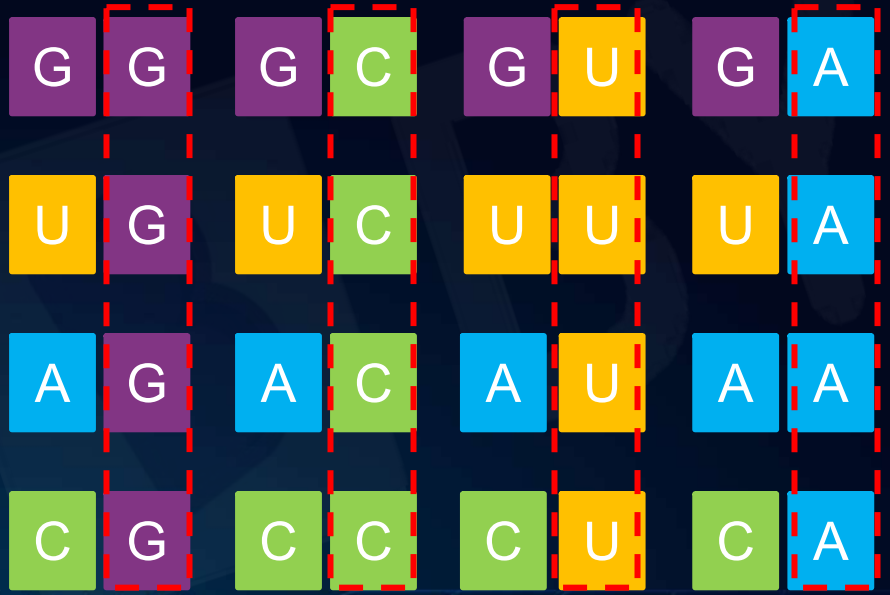
4 combinations < 20 amino acids

- Hence, it was considered that 1 letter code would be insufficient to code for 20 amino acids.

Nucleotides to Amino Acids

If it is considered that:

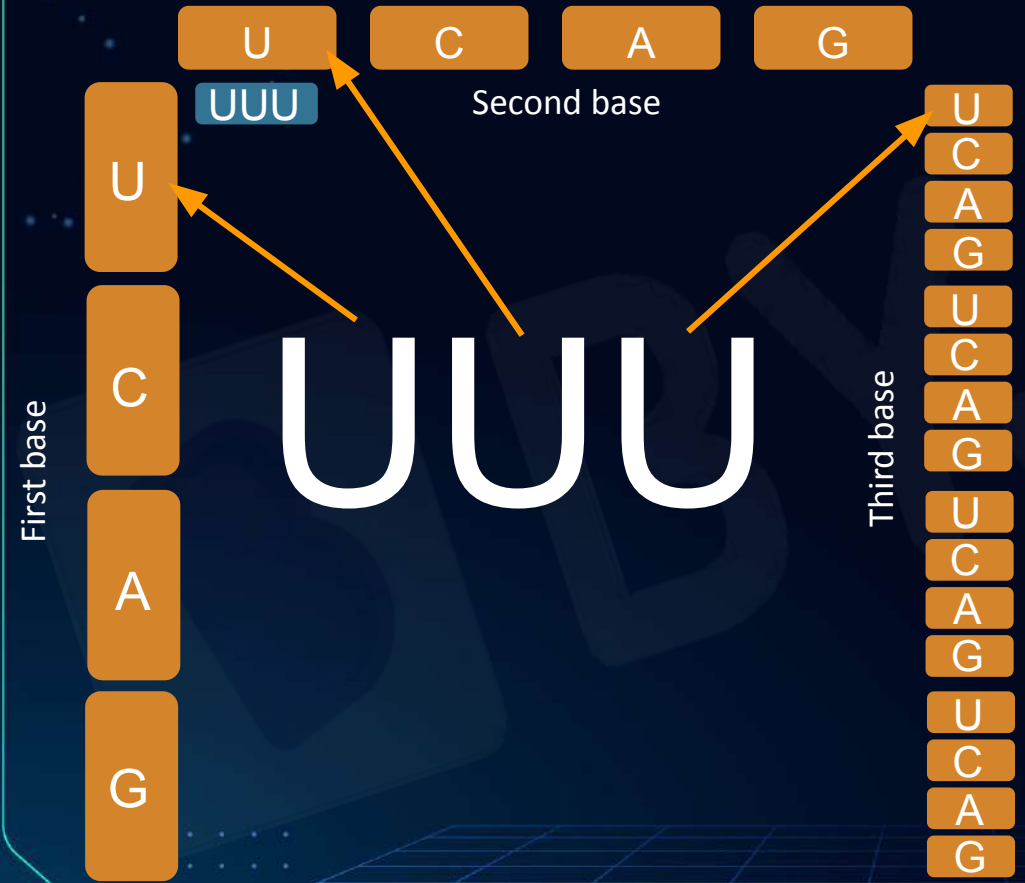
- 2 base = 1 amino acid, then two bases would code for 1 amino acid.
- So, there would be 16 possible combinations which would code for 16 different amino acids.



16 combinations < 20 amino acids

- In the second position of first column G is fixed and bases in first column are different, then 4 combinations are formed.
- Similarly if we look at 4 different columns then 16 combinations are formed.
- Hence it was considered that 2 letter code would be insufficient to code for 20 amino acids.

Nucleotides to Amino Acids



If it is considered that:

- 3 base = 1 amino acid, then three base would code for 1 amino acid.
- For example, considering the first three bases from each header, we get a set of 3 bases or triplet of UUU.

Nucleotides to Amino Acids

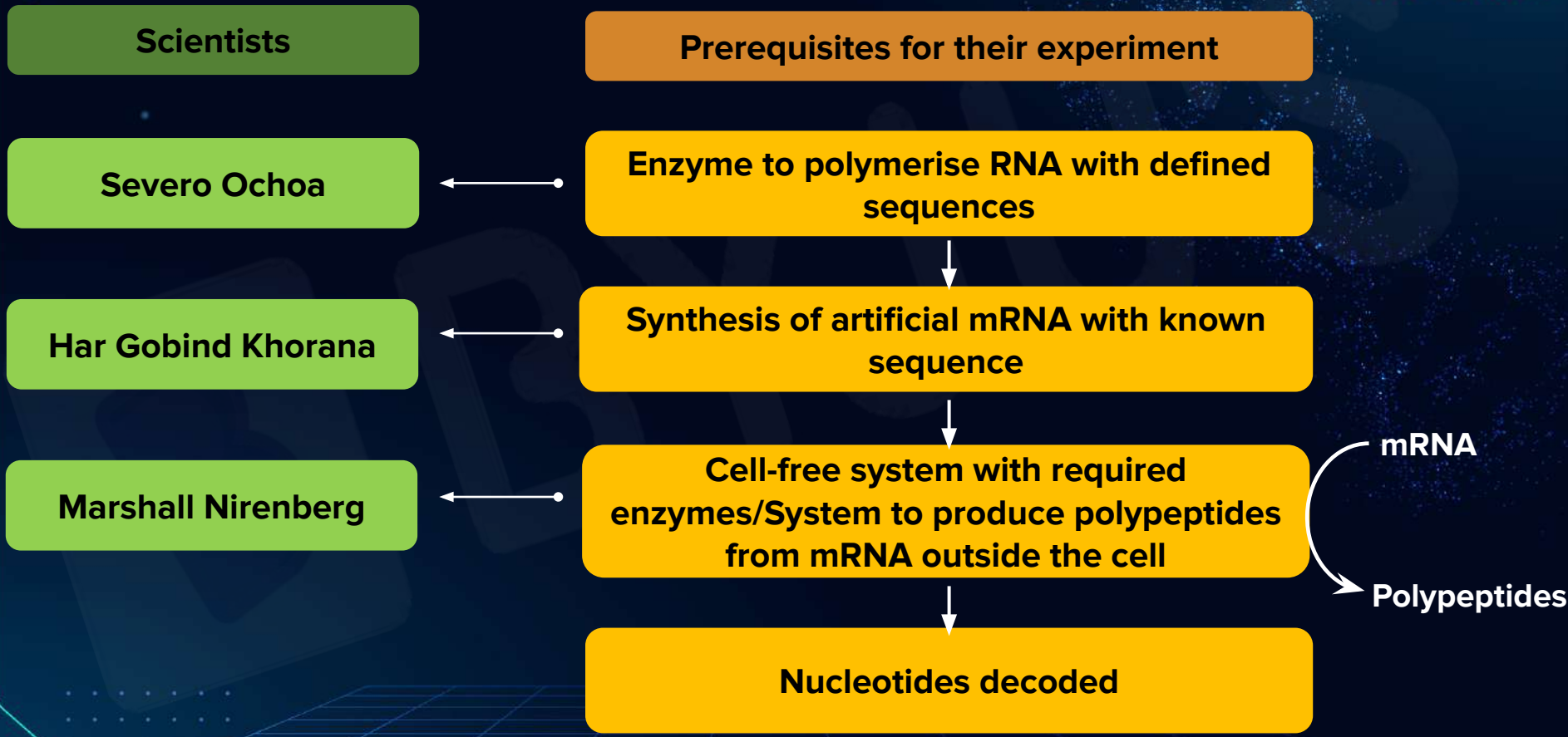
	U	C	A	G	
U	UUU	UCU	UAU	UGU	U
	UUC	UCC	UAC	UGC	C
	UUA	UCA	UAA	UGA	A
	UUG	UCG	UAG	UGG	G
C	CUU	CCU	CAU	CGU	U
	CUC	CCC	CAC	CGC	C
	CUA	CCA	CAA	CGA	A
	CUG	CCG	CAG	CGG	G
A	AUU	ACU	AAU	AGU	U
	AUC	ACC	AAC	AGC	C
	AUA	ACA	AAA	AGA	A
	AUG	ACG	AAG	AGG	G
G	GUU	GCU	GAU	GGU	U
	GUC	GCC	GAC	GGC	C
	GUA	GCA	GAA	GGA	A
	GUG	GCG	GAG	GGG	G

- Similarly If all other bases are considered then there will be 64 combinations.

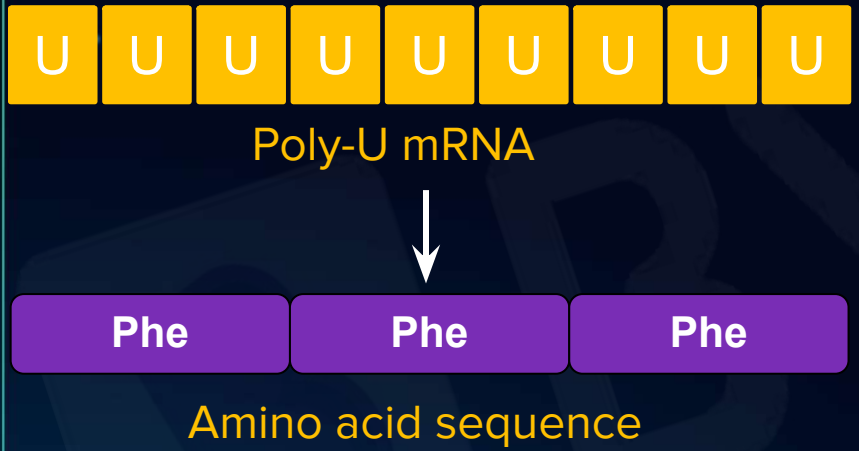
64 combinations > 20 amino acids

- Hence, it was considered that 3 letter code (triplets) would be sufficient to code for 20 amino acids.

Discovery of Genetic Code

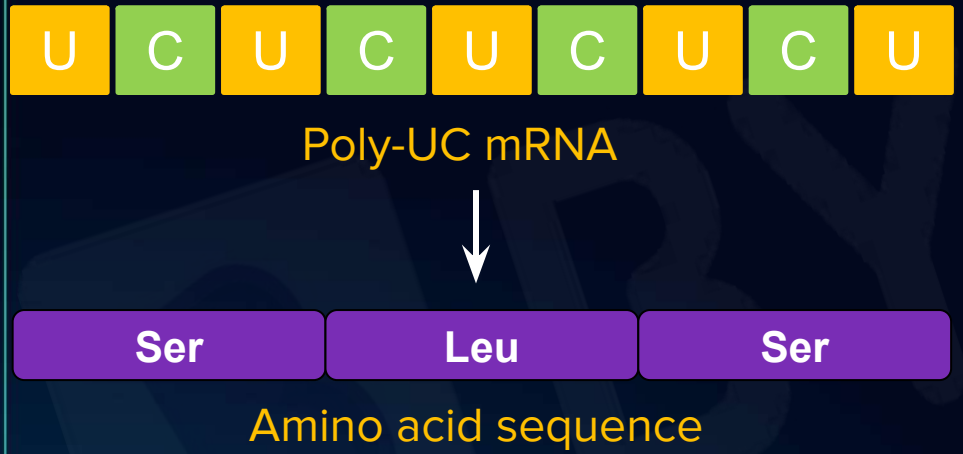


Discovery of Genetic Code



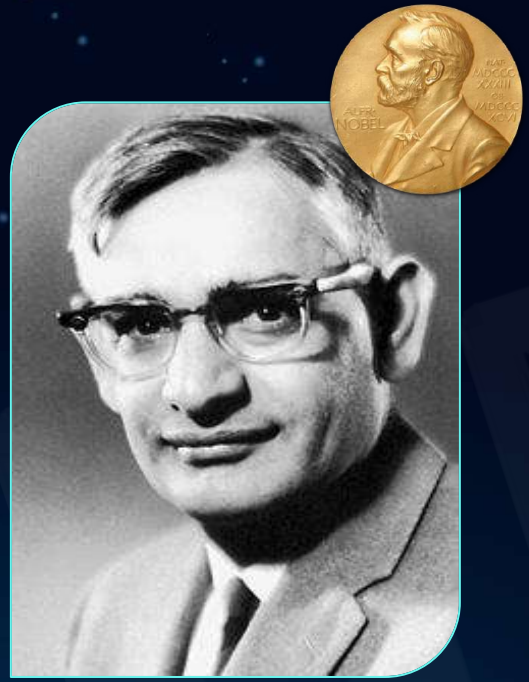
- Nirenberg first synthesized an mRNA molecule consisting only of the nucleotide uracil (called poly-U).
- He found that the polypeptides made consisted exclusively of the amino acid phenylalanine.
- Because the only triplet in poly-U mRNA is UUU, Nirenberg concluded that UUU might code for phenylalanine.
- He performed the experiments with A, G, C as well to find the corresponding amino acids.

Discovery of Genetic Code

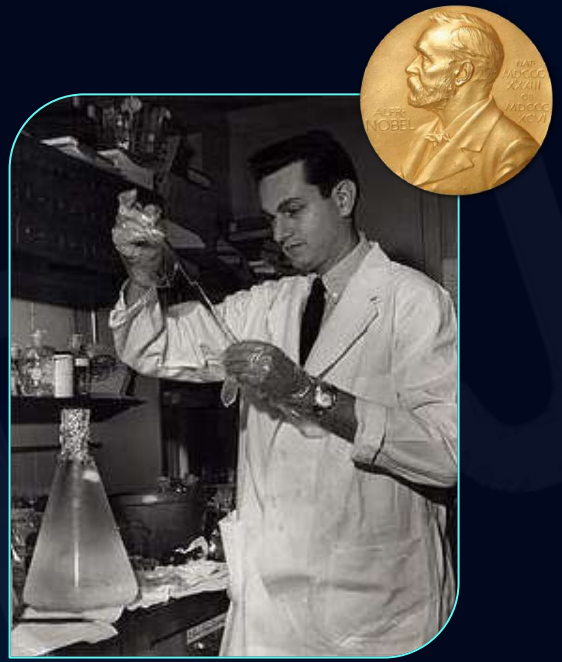


- Har gobind khorana extended Nirenberg's experiment.
- Khorana generated a poly-UC mRNA that resulted in a polypeptide with an alternating pattern of serine and leucine amino acids.
- Since there were only two possible combinations on triplets, either UCU or CUC, they concluded that one of these would code for leucine and the other for serine.

Discovery of Genetic Code



Har Gobind Khorana



Marshall Nirenberg

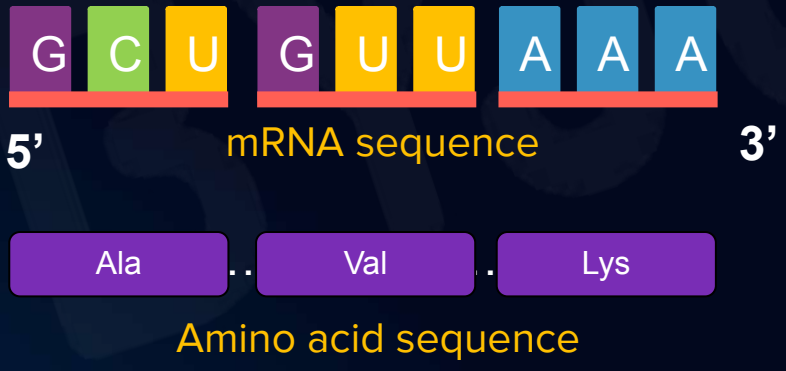
- They received nobel prize for medicine in 1930 for decoding, what each of 64 nucleotides would give rise to.

Salient Features of Genetic Code

Salient Features of Genetic Code

Triplet Codon

- Three nitrogenous bases form a codon.



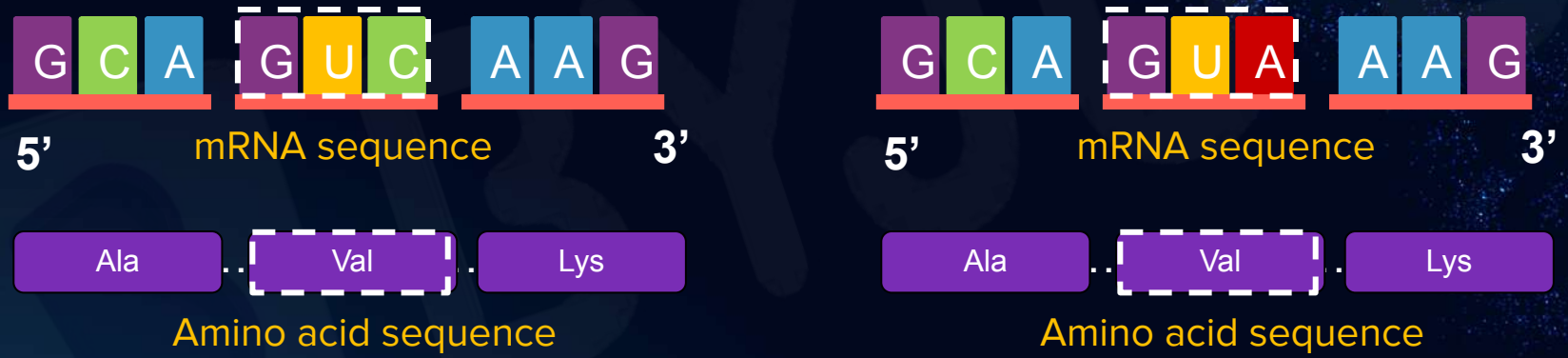
UUU Phe	UCU Ser	UAU Tyr	UGU Cys
UUC	UCC	UAC	UGC
UUA Leu	UCA	UAA	UGA
UUG	UCG	UAG	UGG Trp
CUU	CCU Pro	CAU His	CGU Arg
CUC	CCC	CAC	CGC
CUA	CCA	CAA Gln	CGA
CUG	CCG	CAG	CGG
AUU Ile	ACU Thr	AAU Asn	AGU Ser
AUC	ACC	AAC	AGC
AUA	ACA	AAA Lys	AGA
AUG Met	ACG	AAG	AGG
GUU Val	GCU Ala	GAU Asp	GGU Gly
GUC	GCC	GAC	GGC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

- Triplet codon for all the 20 amino acids are given.

Salient Features of Genetic Code

Degenerate

- Some amino acids are coded by more than one codon.



- Both GUC and GUA code for Valine and there is no change in the amino acid sequence.

Salient Features of Genetic Code

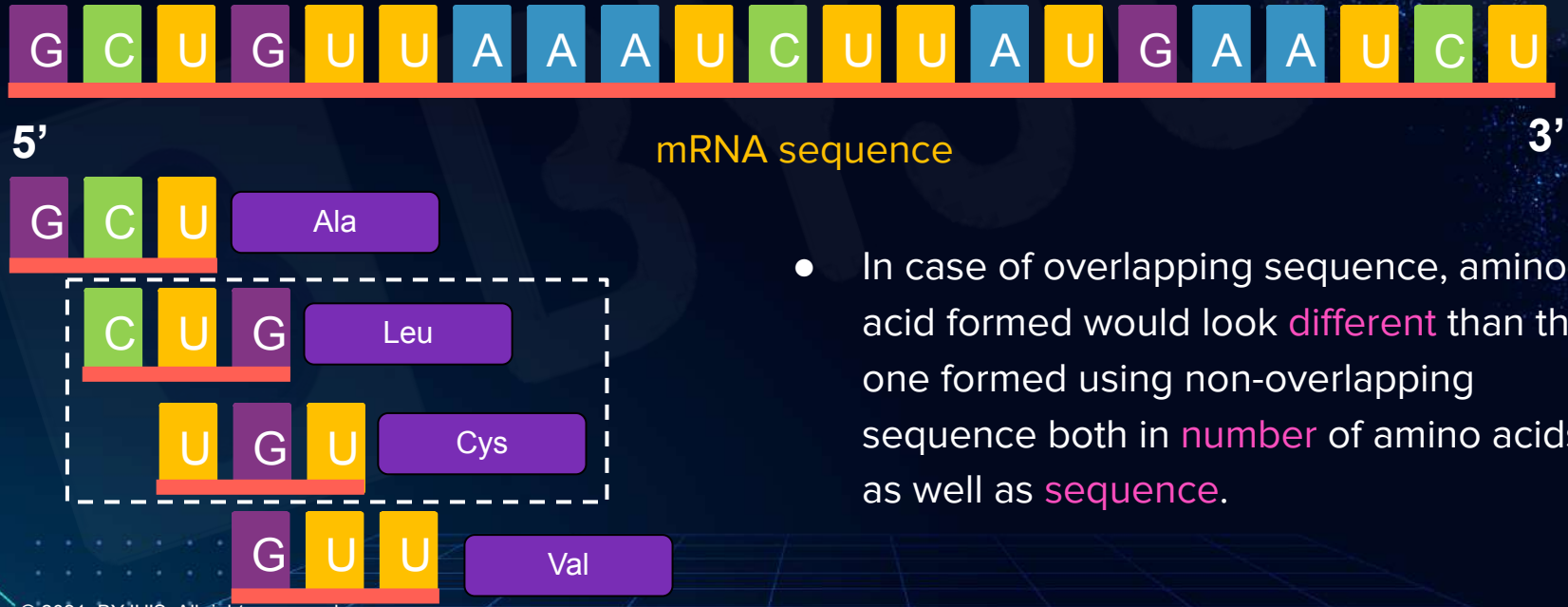
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
UUA	Leu	UCA	Ser	UAA		UGA	
UUG	Leu	UCG	Ser	UAG		UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC	Ile	ACC	Thr	AAC	Asn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

- There are various codons that code for same amino acid.

Salient Features of Genetic Code

Non-overlapping

- Codons do not overlap over each other. They are discrete.

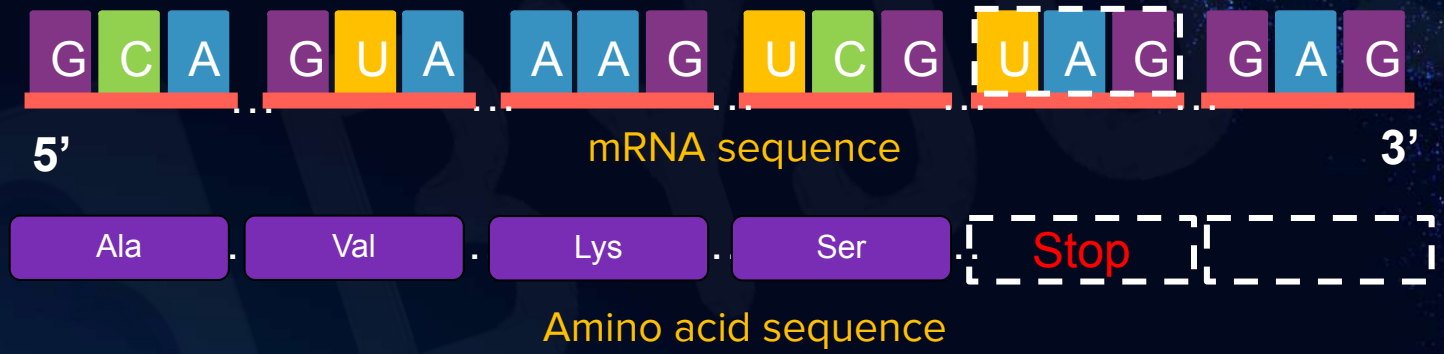


- In case of overlapping sequence, amino acid formed would look different than the one formed using non-overlapping sequence both in number of amino acids as well as sequence.

Salient Features of Genetic Code

Stop codon

- Three codons do not code for any amino acid and hence function as stop codons.



- In the above mRNA sequence 5th codon i.e, UAG acts as stop codon.
- So in the sequence, the amino acids for the first 4 codons would be produced and the polypeptide would be terminated immediately.

Salient Features of Genetic Code

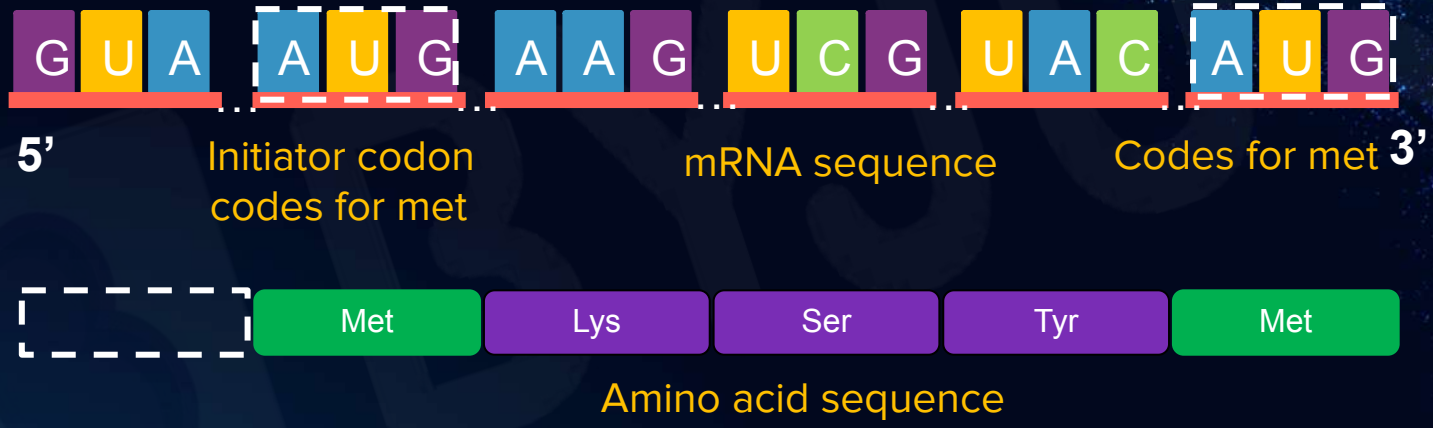
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC	Ile	ACC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

- There are 3 stop codons.

Salient Features of Genetic Code

Dual nature

- AUG codes for methionine and also functions as initiator codon.

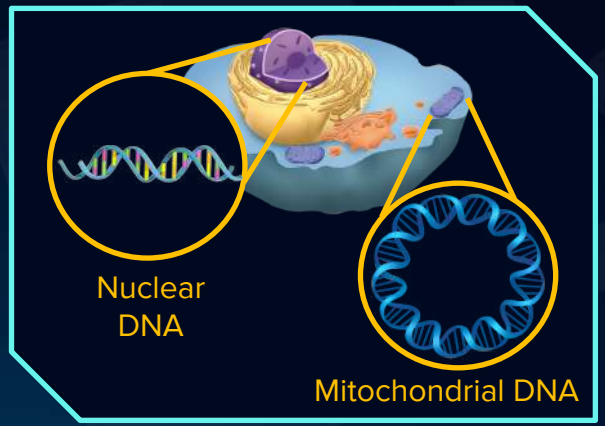


- In an mRNA, amino acid sequence starts only after first AUG is encountered.
- So, the first AUG of the mRNA performs dual functions and signals the start of polypeptide synthesis as well as codes for methionine.

Salient Features of Genetic Code

Universal

- A codon codes for the same amino acid across all living organisms and viruses.
- For example, UUU codes for phenylalanine in all organisms including viruses.
- There are some exceptions such as nuclear and mitochondrial DNA.



- In a cell, if the nuclear DNA transcribes an mRNA sequence with the codons AGA and AGG, then these codons code for arginine.

AGA	Arg
AGG	Arg

- But mitochondrial DNA transcribes an mRNA sequence with the codons AGA and AGG, then these codons act as stop codons.

AGA	Stop
AGG	Stop

Salient Features of Genetic Code

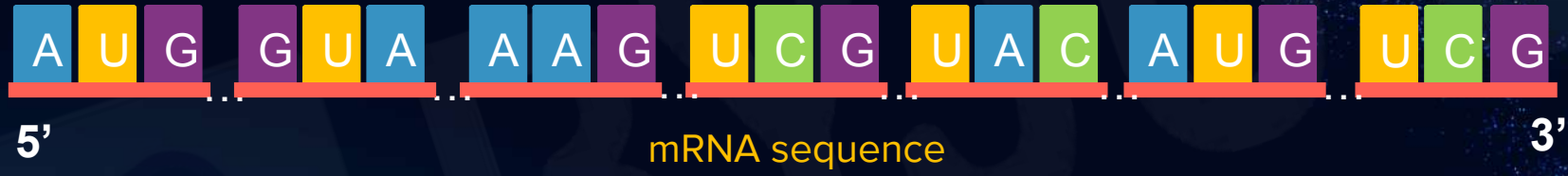
Non-ambiguous

- A codon codes for a specific amino acid.
- UUU always codes for phenylalanine and never codes for any other amino acid.

Salient Features of Genetic Code

Contiguous

- There are no punctuations between codons in an mRNA.



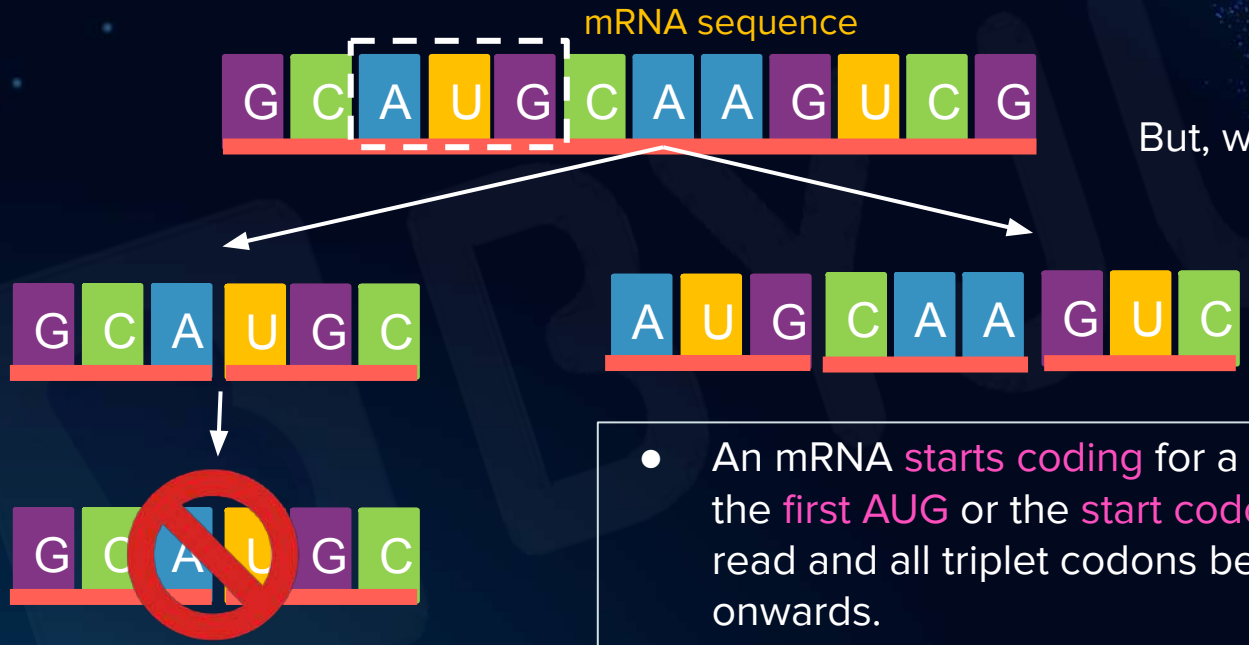
mRNA sequence

5’ AUGGUAAAGUCGUACAUGUCG 3’

- mRNA sequence is divided into 3 bases each for our convenience and is always read in triplets.

Reading an mRNA Sequence

mRNA sequence is always read in triplets



But, where do the triplets start?

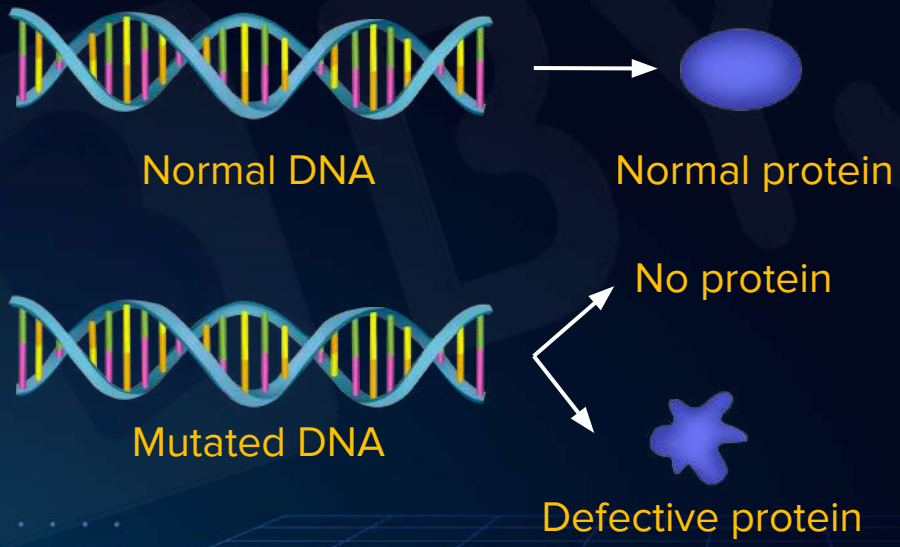
Starting from GCA would not work

- An mRNA starts coding for a protein only after the first AUG or the start codon on the mRNA is read and all triplet codons begin from there onwards.
- Division of set of nucleotides into non-overlapping triplets is known as reading frame.

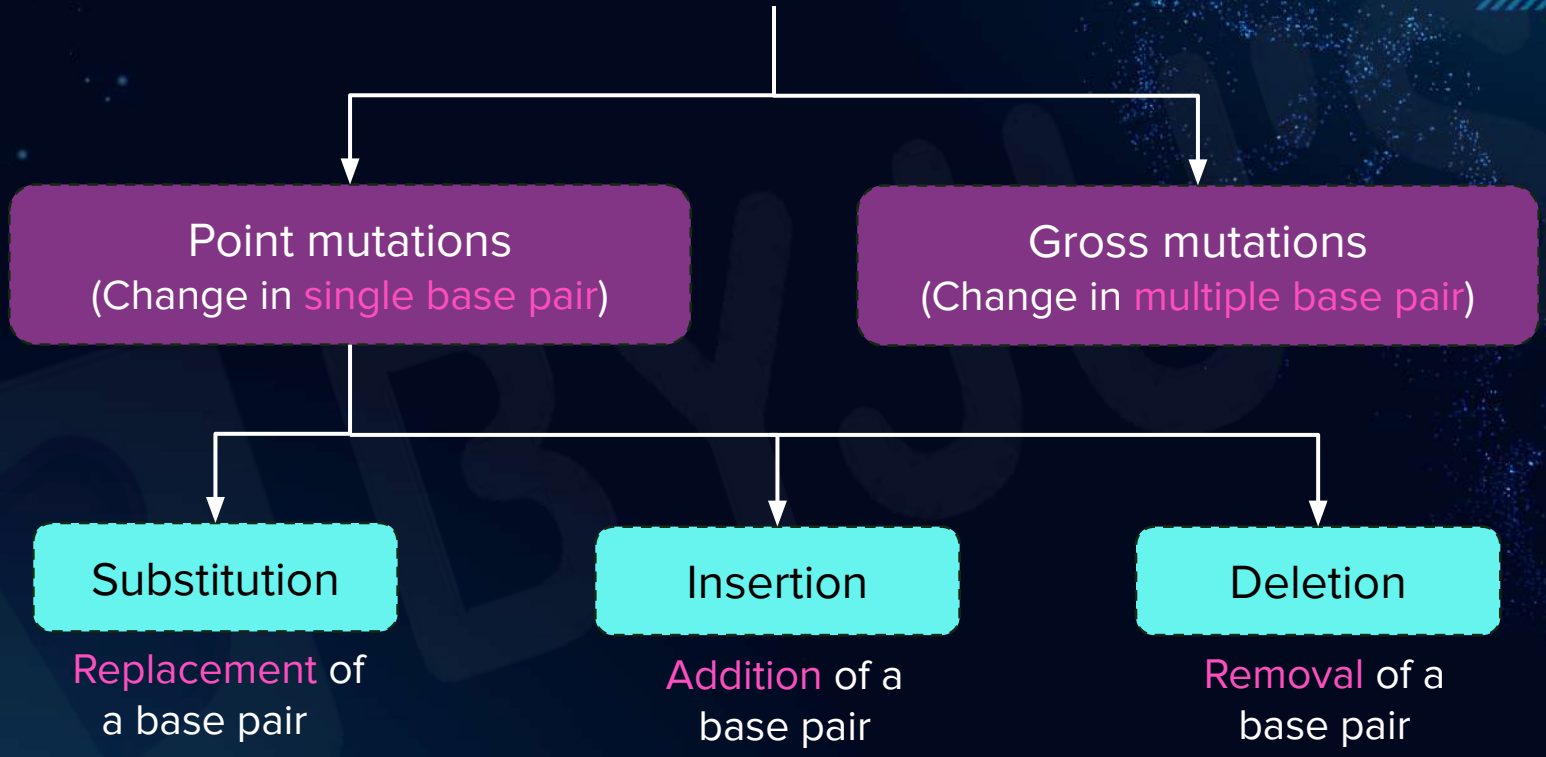
Mutations

It is a phenomenon that results in:

- Alteration of DNA sequences
- Changes in the genotype and phenotype
- UV rays causes mutation in the DNA sequence which may cause cancer.



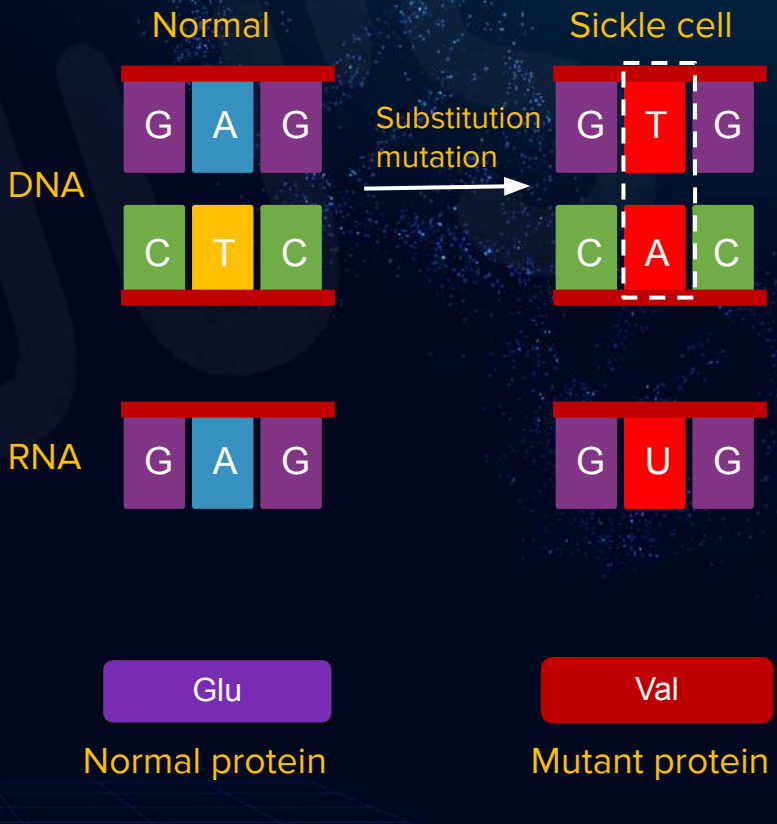
Types of mutations



Mutations

Substitution mutation

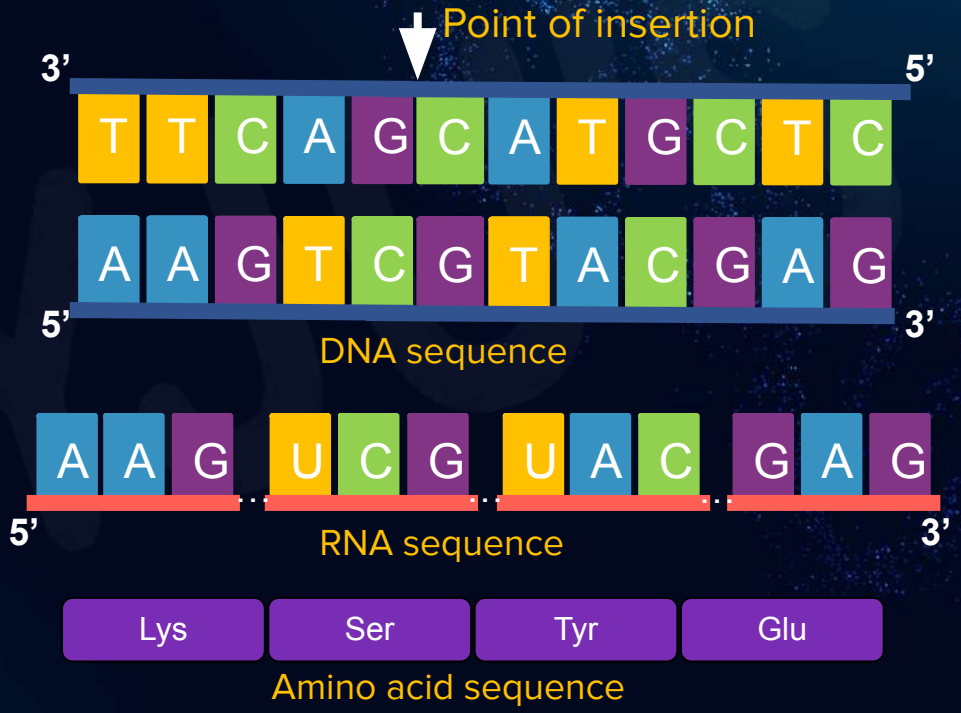
- Sickle cell anemia is an example of substitution mutation.
- In this condition, the structure of Hb drastically changes and thereby causing change in the RBC shape.
- In this, A is replaced with T.



Mutations

Insertion mutation

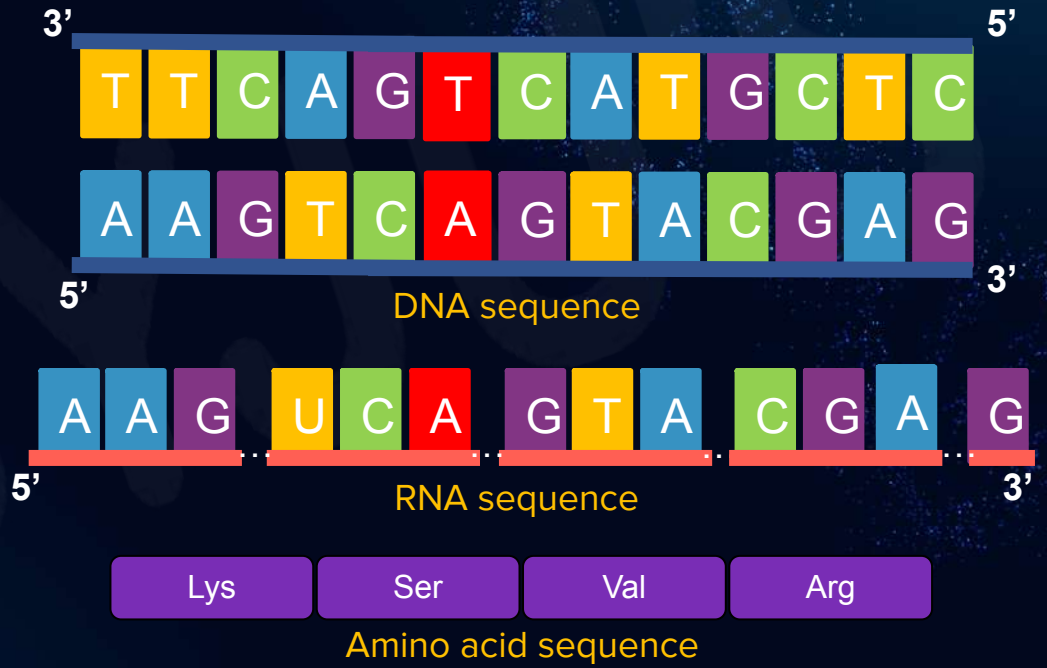
- TA base pair is inserted into the DNA sequence.
- This type of mutation is also known as frameshift mutation.



Mutations

Insertion mutation

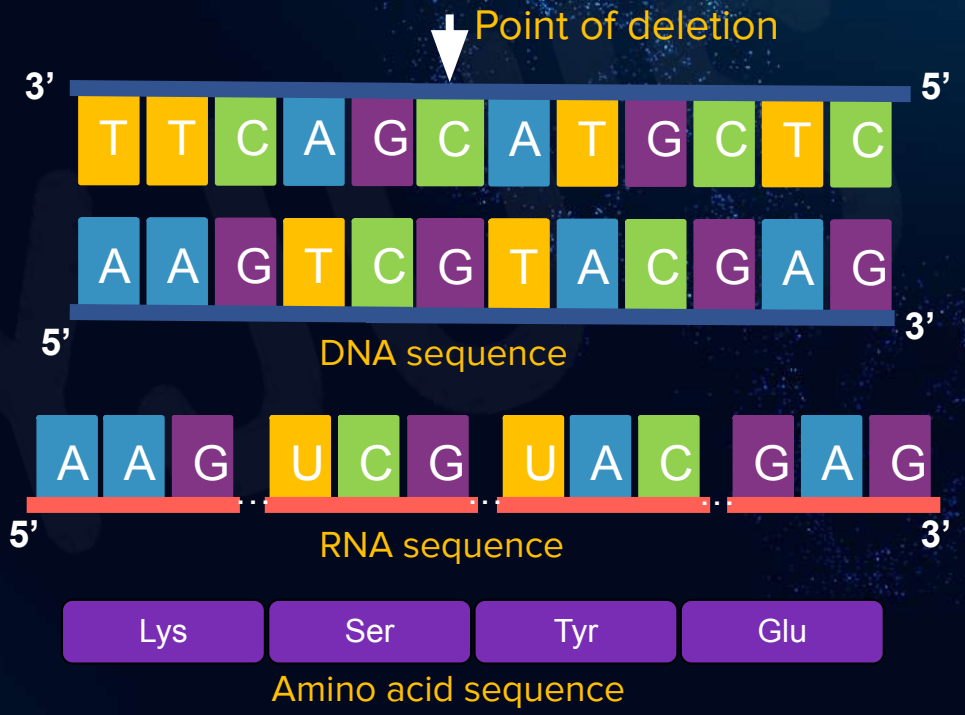
- Due to frameshift, amino acid sequence also changes.



Mutations

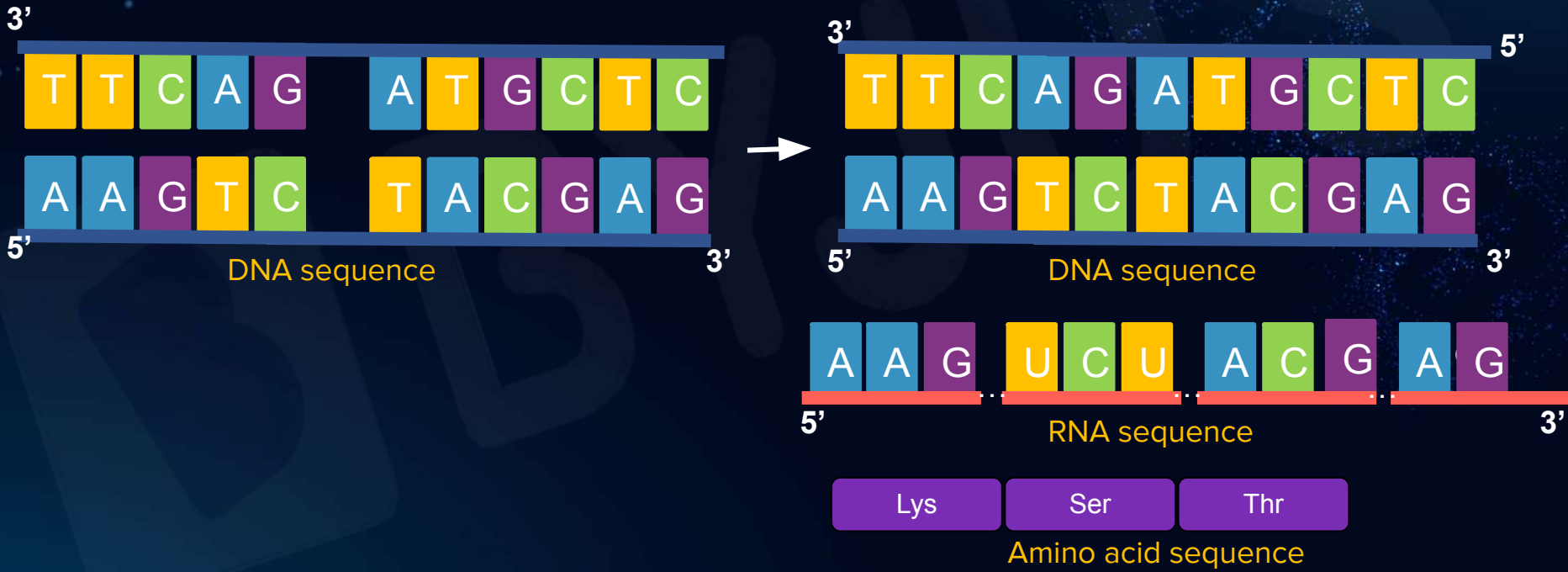
Deletion mutation

- CG base pair is deleted from the DNA sequence.
- This type of mutation is also known as frameshift mutation.



Mutations

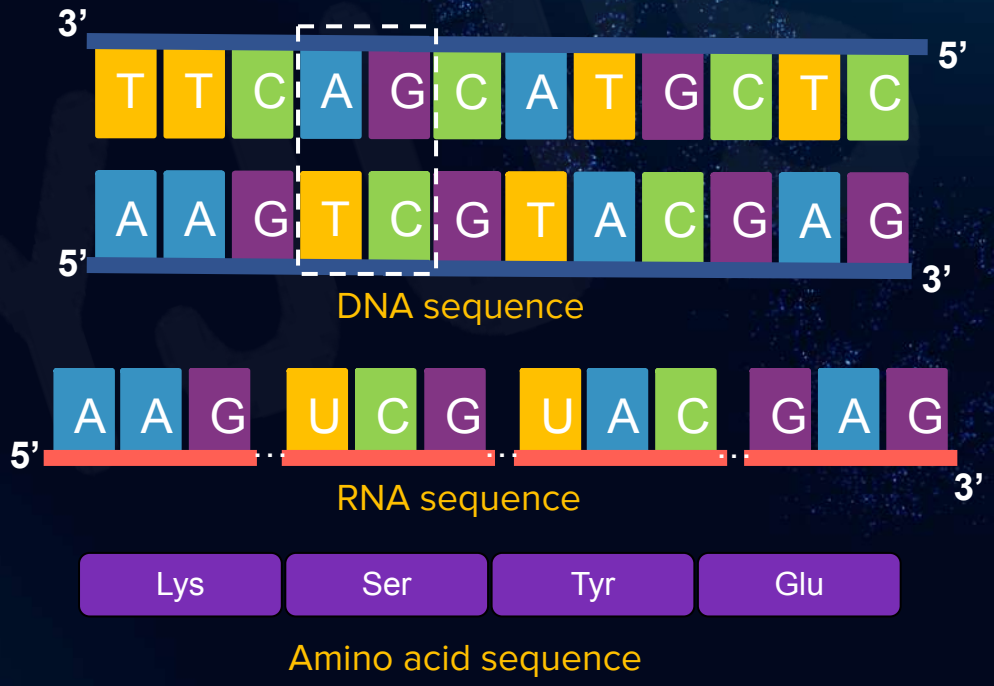
Deletion mutation



Mutations

Gross mutation

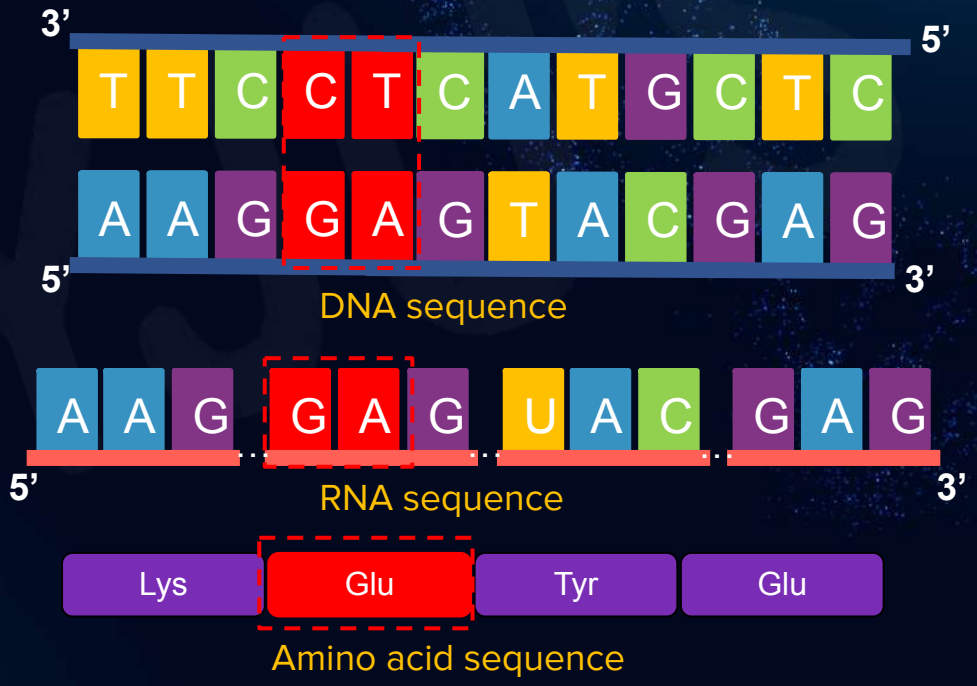
- Change in multiple base pairs lead to gross mutation.
- AG is mutated to CT and since there is complementary pairing, nucleotide opposite to CT also changes.



Mutations

Gross mutation

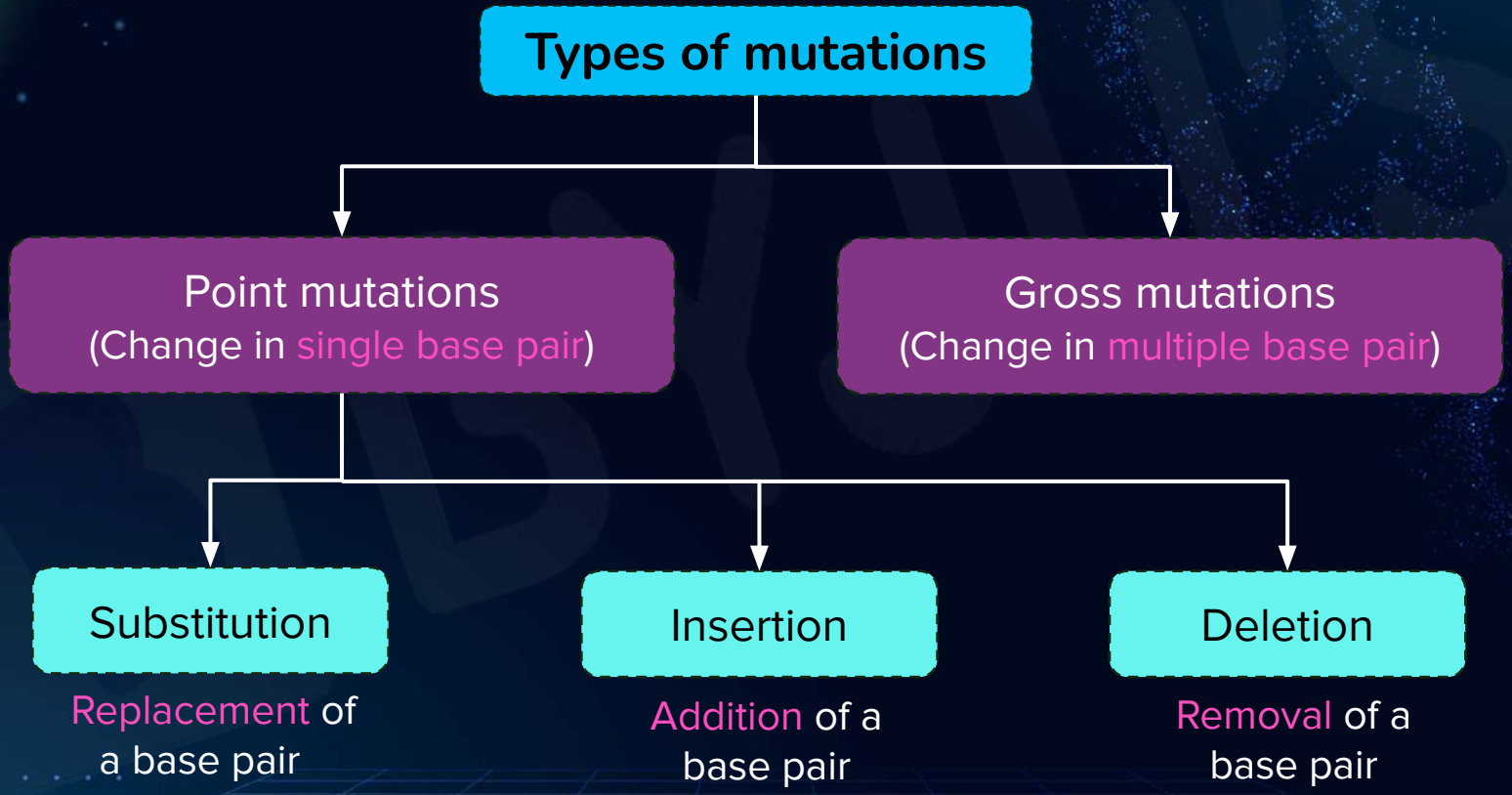
- So when the mRNA is created, the 4the codon changes from UCG to GAG and because of this the amino acid sequence also changes from Serine to Glutamic acid.



Summary

Triplet Codon	Three nitrogenous bases form a codon.
Degenerate	Some amino acids are coded by more than one codon.
Non-overlapping	Codons do not overlap with each other. They are discrete.
Stop Signal	Three codons do not code for any amino acid and hence function as stop codons.
Dual Nature	AUG functions as an initiator codon as well as codes for methionine
Universal	A codon codes for the same amino acid across all living organisms and viruses.
Unambiguous	A codon is specific to only one amino acid.
Contiguous	There are no punctuations between codons in an mRNA

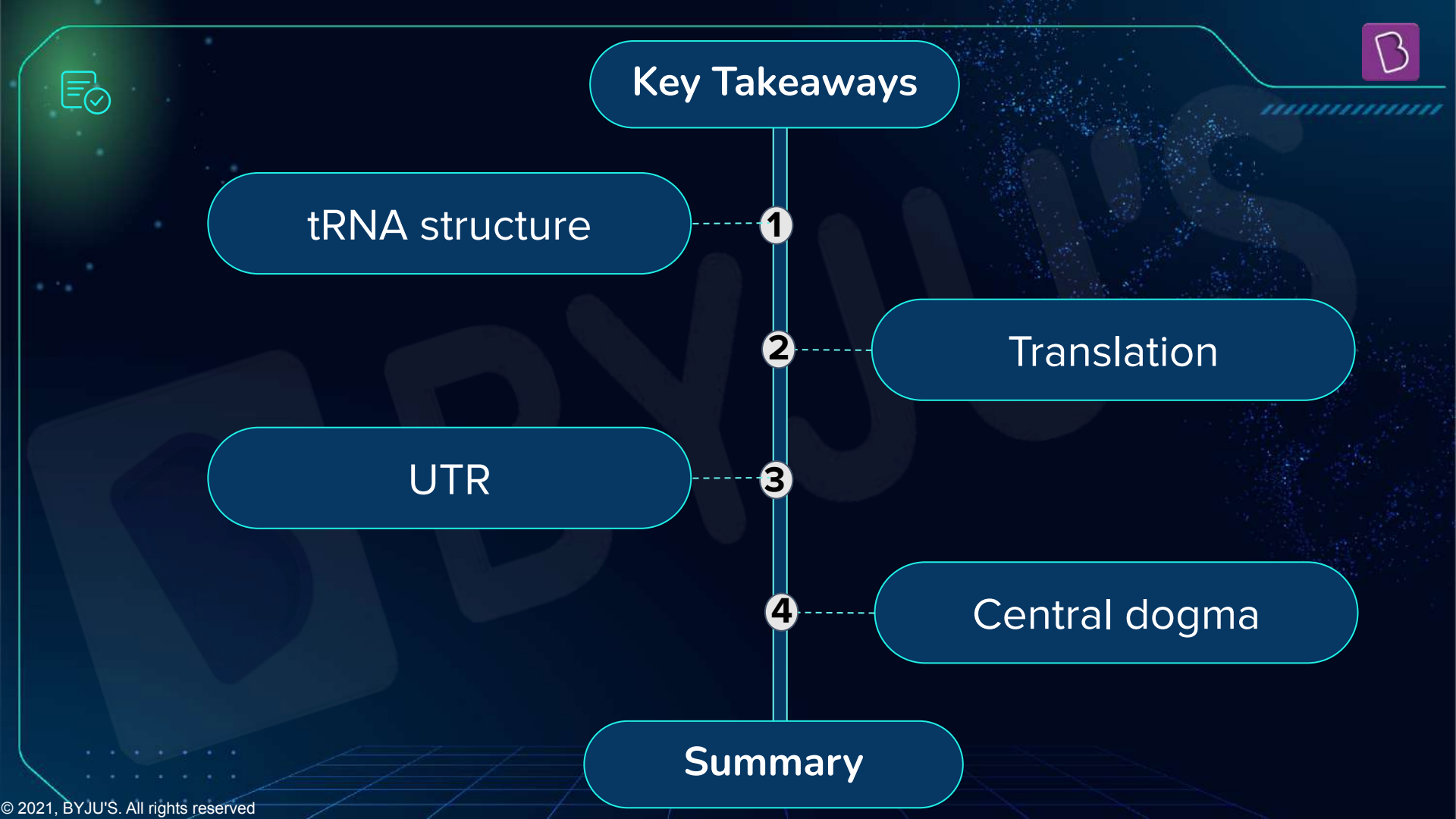
Summary


Types of mutations
Point mutations
(Change in single base pair)
Gross mutations
(Change in multiple base pair)
Substitution
Insertion
Deletion
Replacement of a base pair
Addition of a base pair
Removal of a base pair

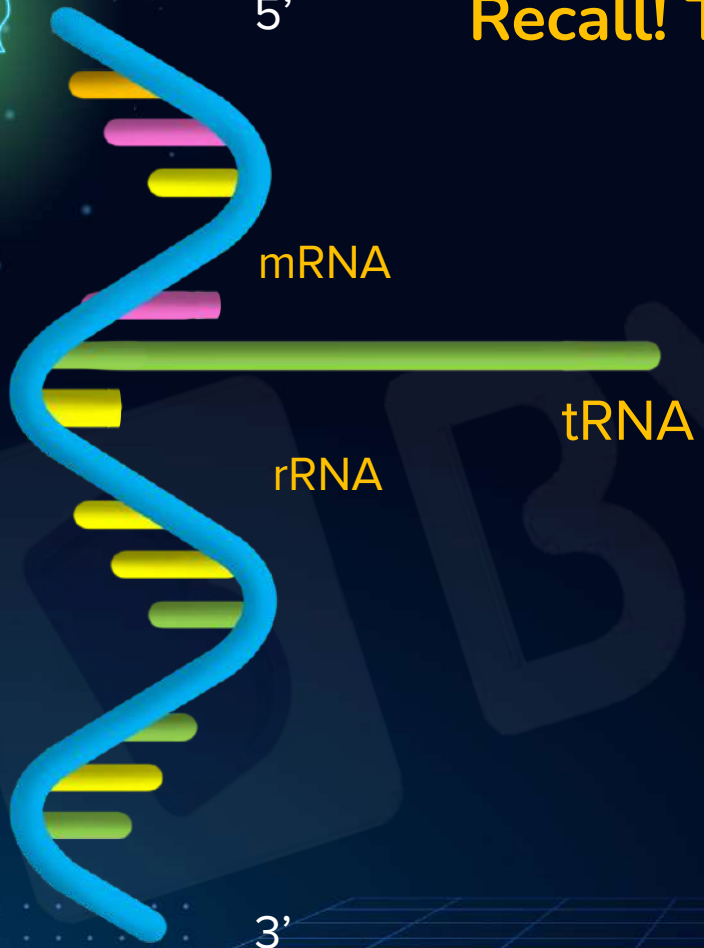
BYJU'S Classes Notes

Molecular Basis of Inheritance

Transfer RNA, Translation, UTR, Central Dogma



Recall! Types of RNA



tRNA or transfer RNA

- tRNA carries specific amino acids in cytoplasm to the site of protein synthesis i.e. ribosomes
- Smallest amongst the various types of RNA
- Previously known as sRNA or soluble RNA
- Soluble in 1M of NaCl

Recall! Types of RNA

Types of RNA	Found in	Location	Function
rRNA- Ribosomal RNA	Both prokaryotes and eukaryotes	Cytoplasm	Forms the structural and functional component of the ribosome
tRNA-Transfer RNA	Both prokaryotes and eukaryotes	Cytoplasm	Carries amino acids to ribosomes during protein synthesis
mRNA-Messenger RNA	Both prokaryotes and eukaryotes	Cytoplasm and nucleus	Carries genetic code from DNA in the nucleus to the site of protein synthesis (in cytoplasm)
hnRNA-Heterogenous nuclear RNA	Eukaryotes	Nucleus	Forms the primary transcript for the mature RNA
snRNA-Small nuclear RNA	Eukaryotes	Nucleus	Takes part in processing of pre-mRNA (hnRNA) in the nucleus

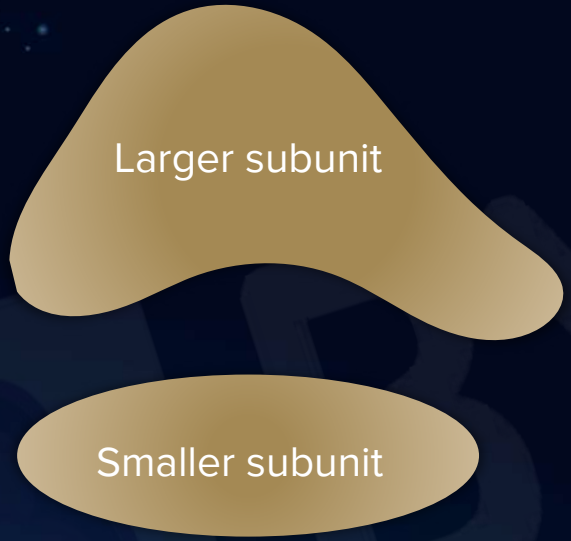
Recall! Codon

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

- Three nitrogenous bases together code for an amino acid.

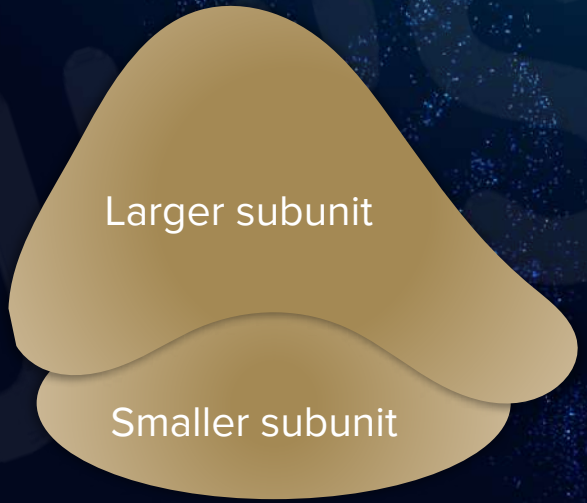
- These triplets, present in the mRNA are called codons.

Recall! Ribosomes



Inactive ribosome

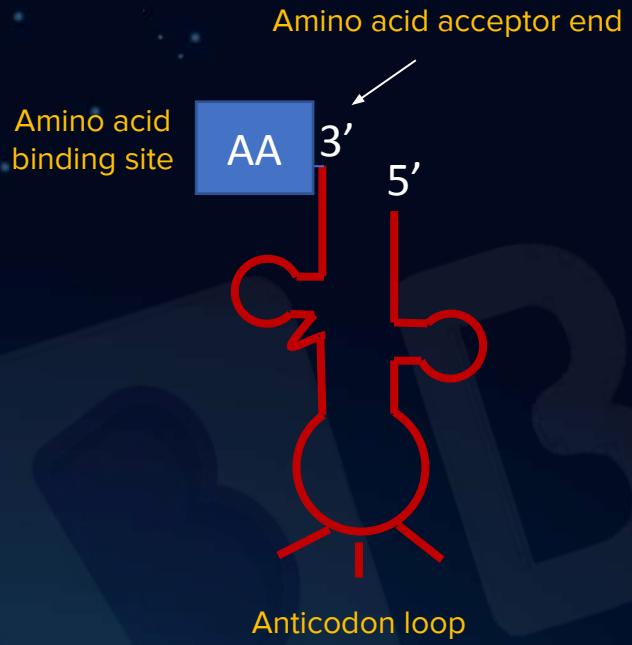
- When protein production is not carried out, the subunits are separate.



Active ribosome

- When protein production is carried out, the subunits get attached.

tRNA : Structure

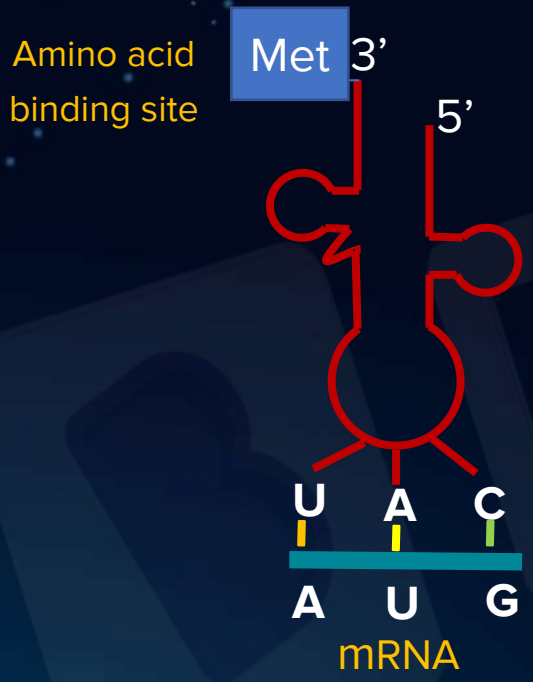


Diagrammatic representation of structure of t-RNA

tRNA or transfer RNA structure

- It has a 3’ and 5’ end.
- It is non-linear, clover leaf shaped structure.
- Its actual 3D structure looks like a inverted letter ‘L’.
- 3’ end has a free -OH group.
- Hence, the amino acid which has to be transferred binds to this site.
- It is called amino acid binding site.
- This end of tRNA is called amino acid acceptor end.
- The lower end of the tRNA is called the anticodon loop.

t-RNA : Structure

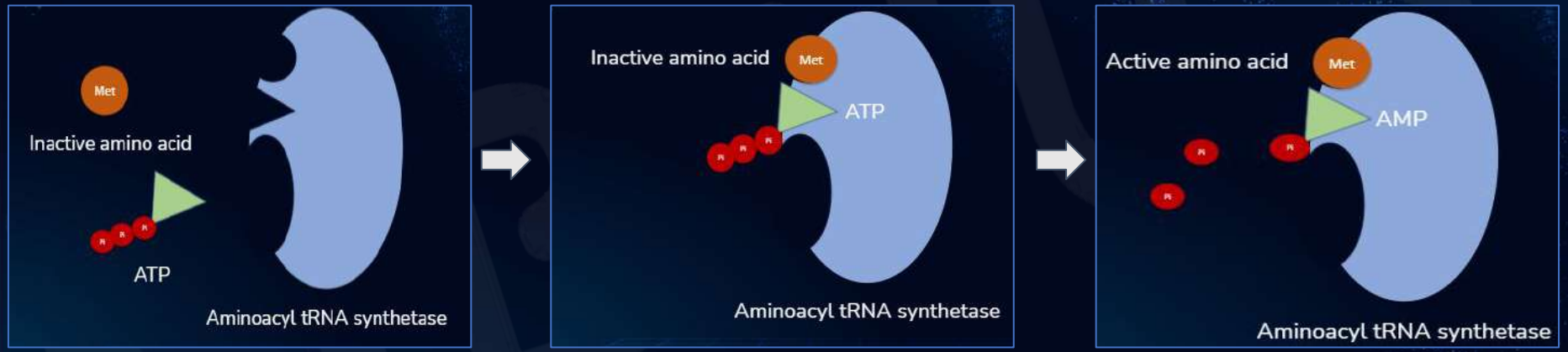


Anticodon loop and its functions

- Anticodon loop in tRNA helps to identify the codon for the corresponding amino acid it is carrying.
- Triplets which are complementary to codon are called anticodons.
- This anticodon for the corresponding amino acid, that a specific tRNA carries, is present on the anticodon loop.
- For example - The tRNA that carries the amino acid called methionine will have the anticodon UAC.
- As it comes in the contact with AUG codon on mRNA, it will bind to that site and deliver methionine.
- Hence the tRNA is called an adapter molecule as it acts as connecting link between amino acids (AAs) and mRNA.

tRNA Charging

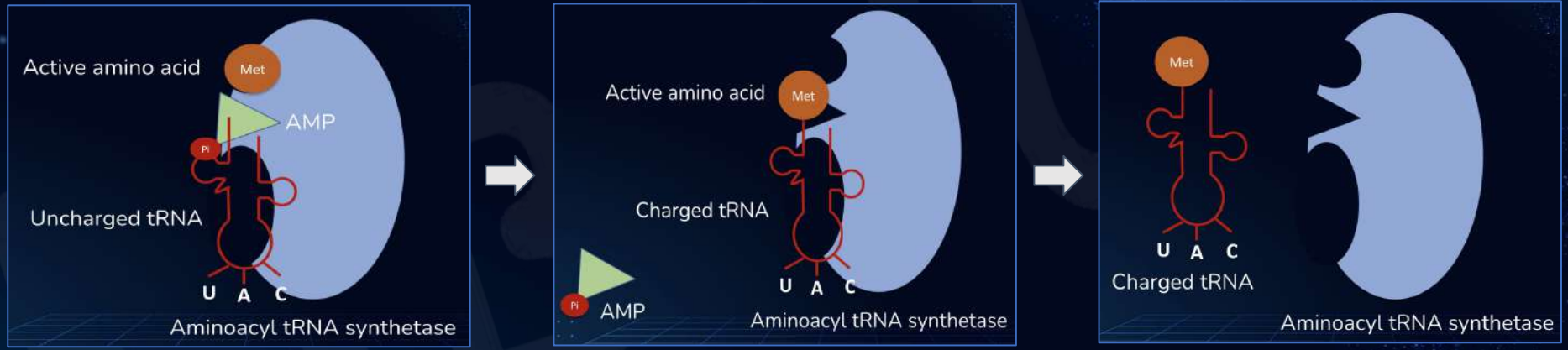
1 Activation of amino acids



- Inactive free amino acid in cytoplasm (methionine in this case) binds to the cognate aminoacyl tRNA synthetase.
- Energy for this process is provided by the splitting ATP.
- ATP loses two Pi to become AMP.
- Hence, a complex called aminoacyl adenylate synthetase complex is formed.

tRNA Charging

2 Charging of tRNA



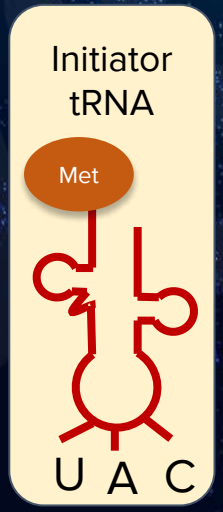
- The cognate uncharged tRNA binds to the complex.
- The enzyme transfers amino acid to the tRNA and releases the AMP.
- tRNA is now charged (bound to amino acid), dissociates from enzyme and can take part in translation.

Steps of Translation

Steps of Translation

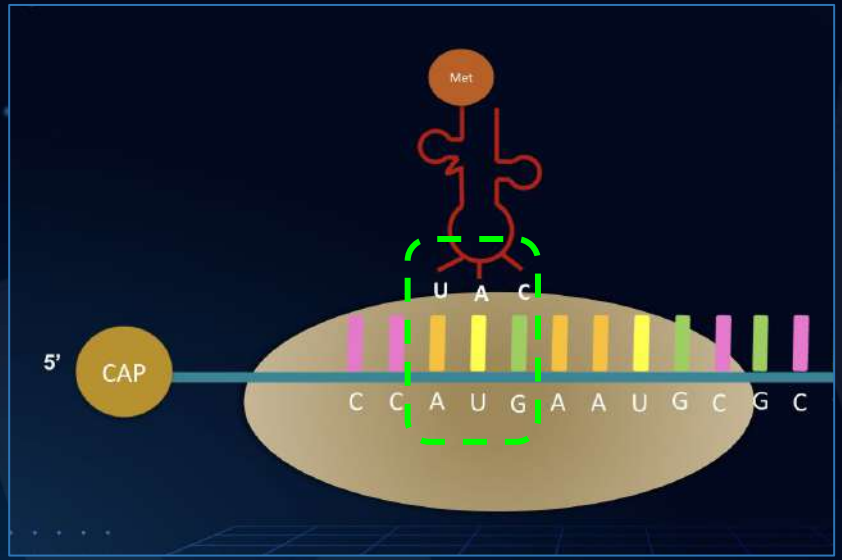
1 Initiation

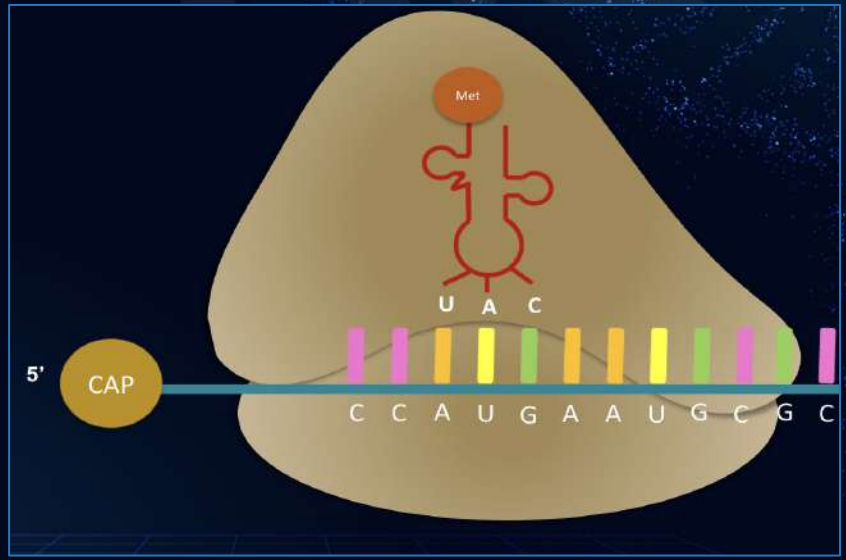
- Translation always starts at the start codon AUG present on the mRNA (near the 5’ end).
- The initiator tRNA has an anticodon UAC and is charged with methionine amino acid.



Steps of Translation

1 Initiation





- The small ribosomal subunit encounters the mRNA and positions near the start codon.
- Then, the initiator tRNA comes and binds to the AUG codon.

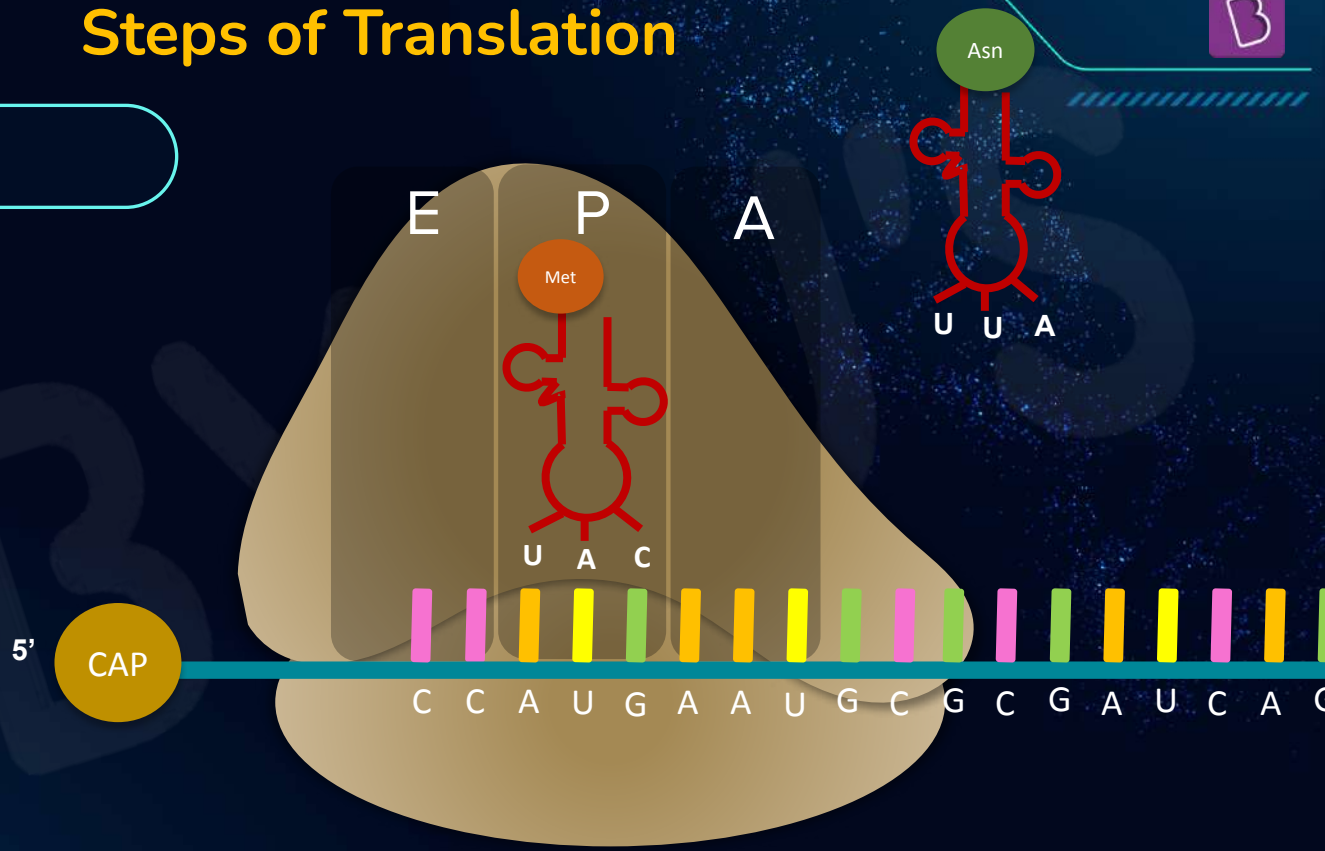
- After the formation of initiation complex, the large ribosomal subunit joins with the small subunit and the process moves ahead.

Steps of Translation

2 Elongation

There are three sites in a ribosome:

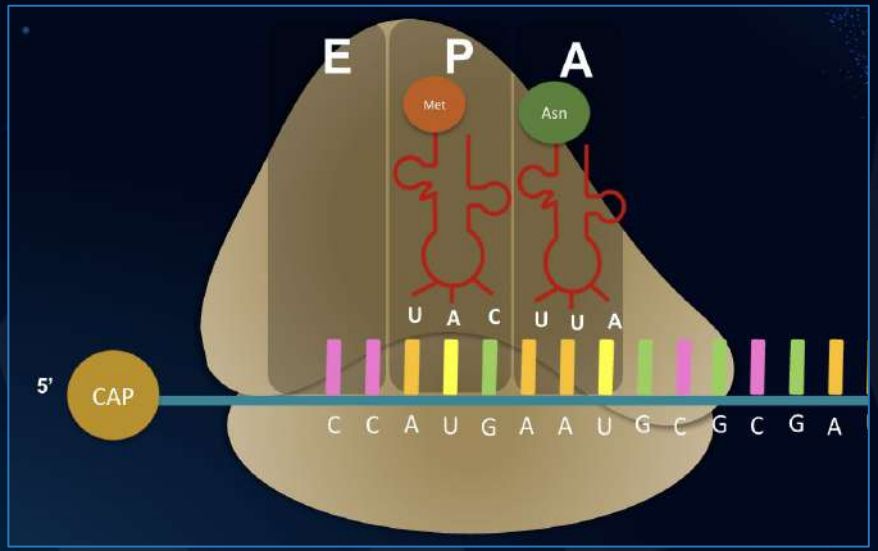
- A (Acceptance)- where the new tRNA enters
- P (Peptidyl)- where the peptide bond is formed
- E (Exit)- where the tRNA (without amino acid) leaves the ribosome

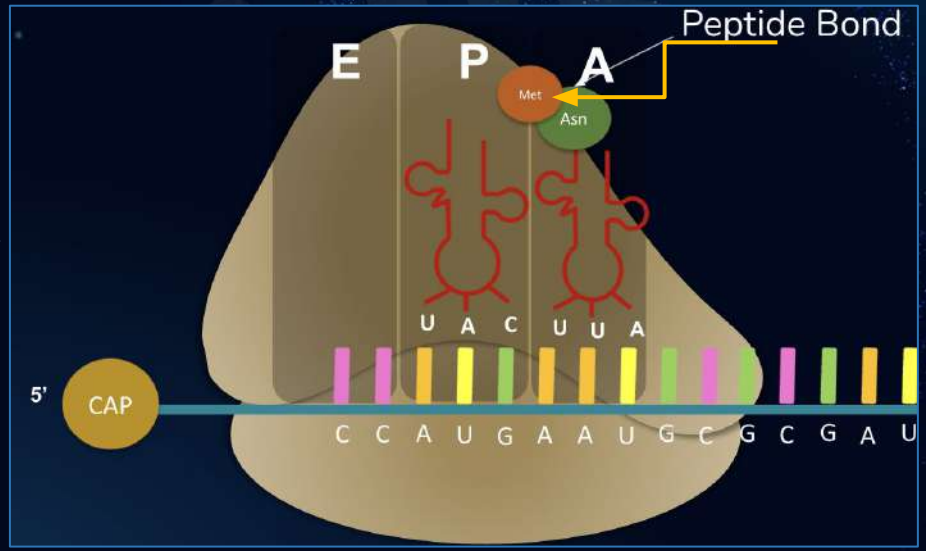


NOTE: Only the initiator tRNA starts from the P site instead of A site.

Steps of Translation

2 Elongation



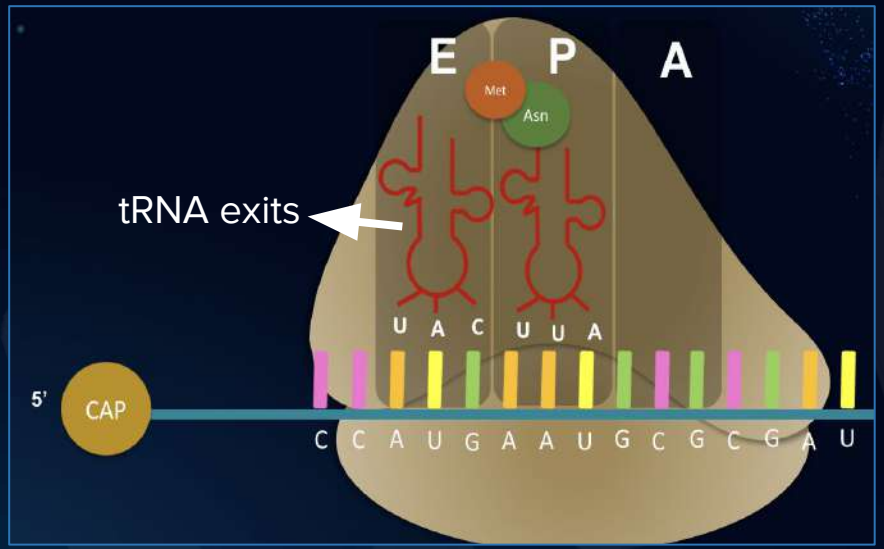


- During elongation, the next charged cognate tRNA comes in to the A site.
- Ribosome acts as catalyst for polymerisation of amino acids.

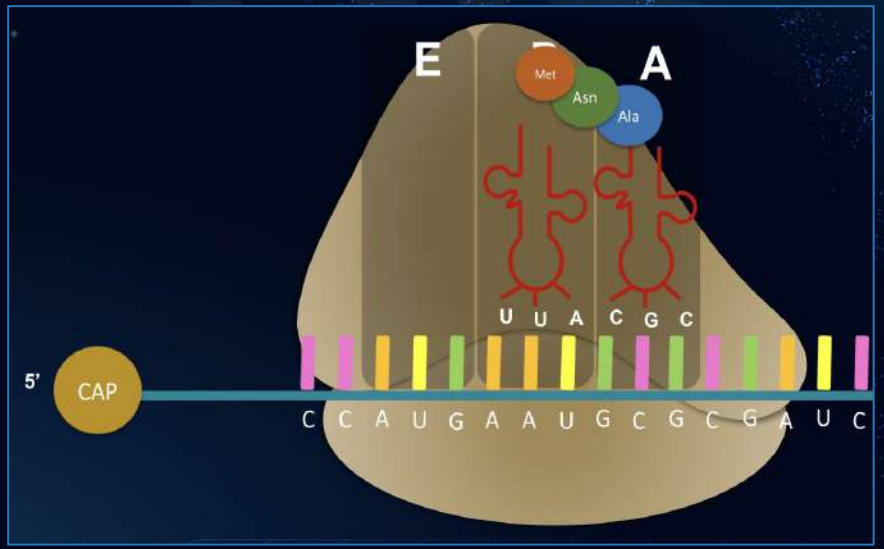
- Peptide bond formation occurs between the first and the second amino acid.
- 23S rRNA catalyses this reaction in bacteria. Hence it is called 'ribozyme'.

Steps of Translation

2 Elongation



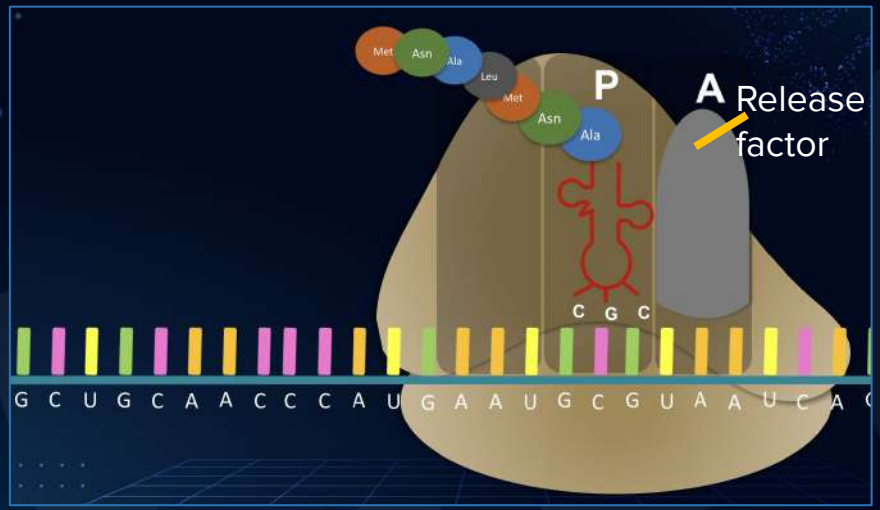
- The whole ribosomal unit moves one whole codon such that the A site becomes vacant.
- The tRNA which was in the P site exits the ribosome through E site.

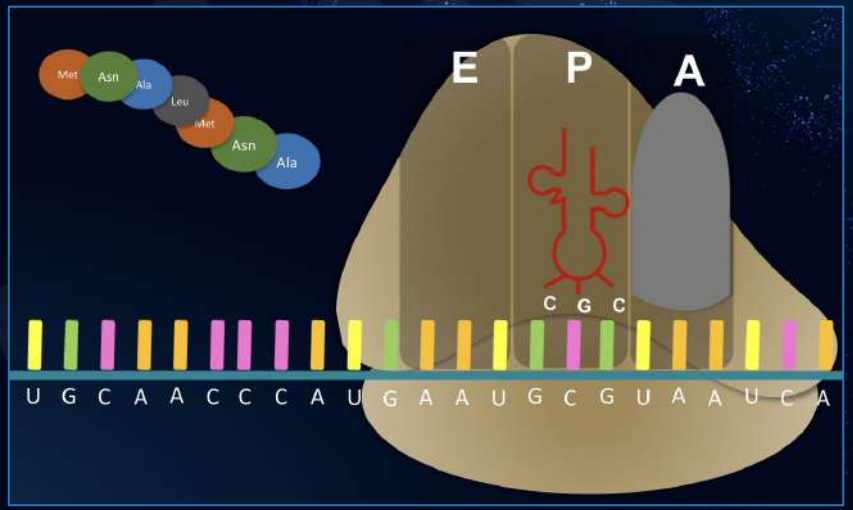


- The A site accepts the next aminoacyl tRNA.
- The polymerisation of amino acids continues.

Steps of Translation

3 Termination

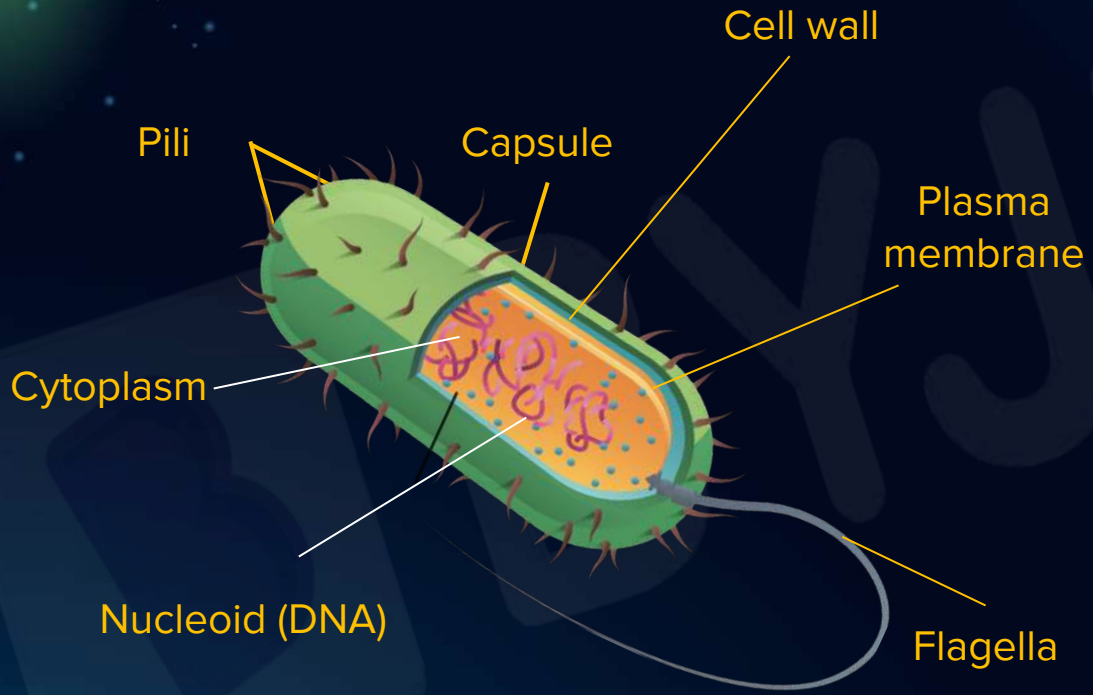




- After elongation, the ribosome reaches the stop codon.
- There is no tRNA specific to the stop codon, instead there is a release factor.

- Release factor identifies the stop codon and attaches at the A site.
- It releases the amino acid chain from the tRNA.
- Then all the parts disassociate, ending the translation process.

Did You Know?



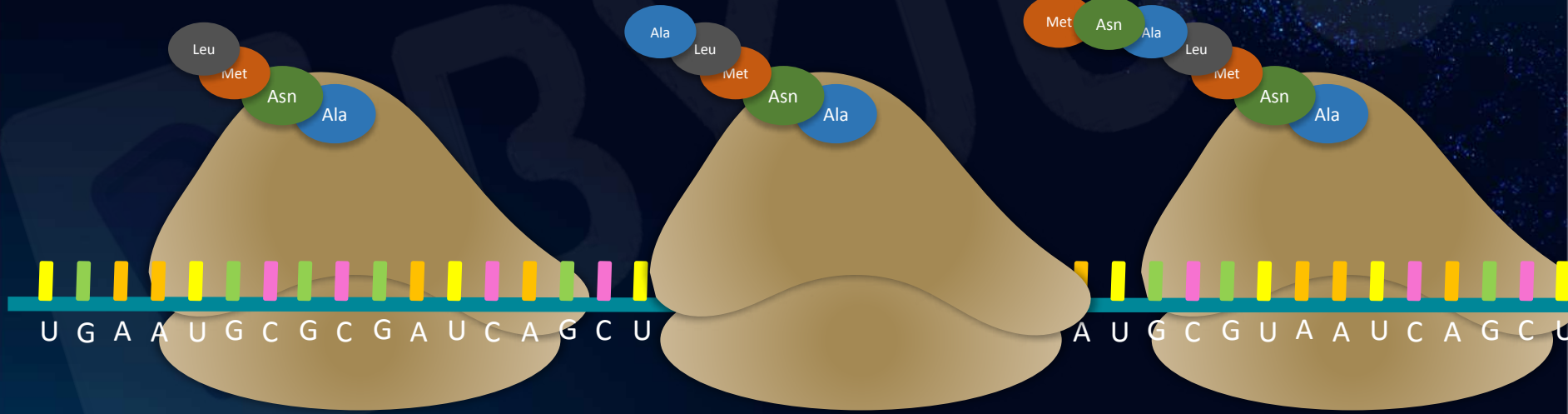
- In prokaryotes, due to the lack of nucleus, translation happens along with transcription!

Untranslated Regions (UTRs)

- A translational region in mRNA is the sequence from the start codon (AUG) to the stop codon. The codons in this region code for a polypeptide.
- There are some additional sequences in the mRNA that are not translated and are known as untranslated regions (UTRs).
- They are present at both 5' -end (before the start codon) and at 3' -end (after the stop codon).
- They are important for efficient translation.

Did You Know?

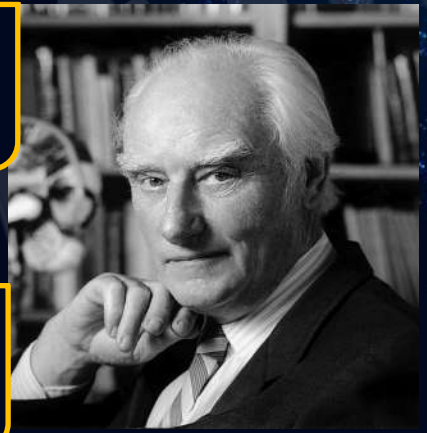
- There can be multiple ribosomes reading the same mRNA at the same time.
- This structure is called polysome or polyribosome.



Central Dogma

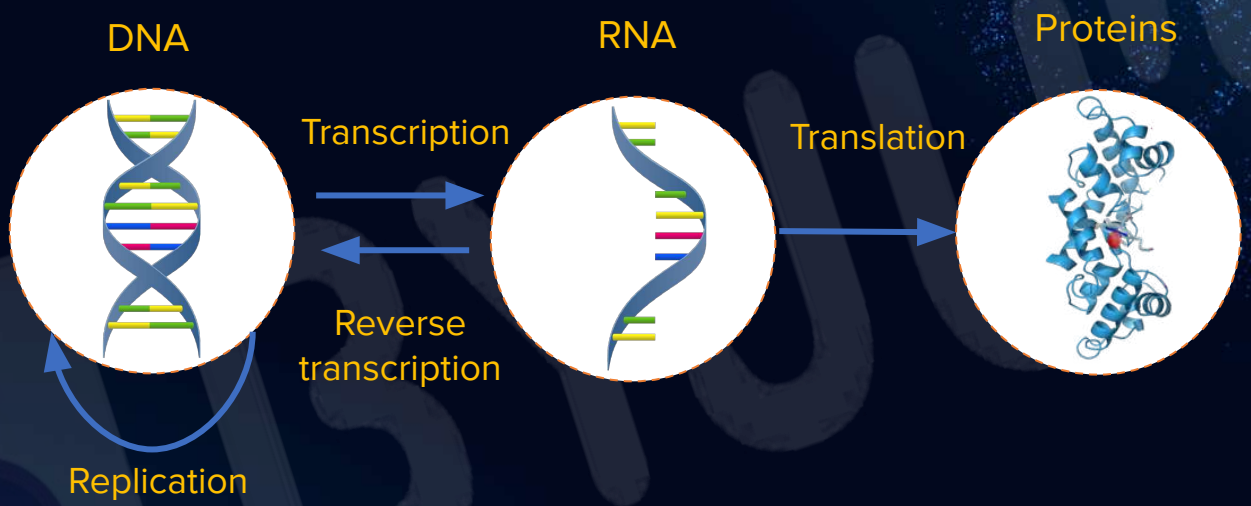
Central dogma of molecular biology was proposed by Francis Crick

It involves the conversion of DNA into a functional product



Francis Crick

Central Dogma

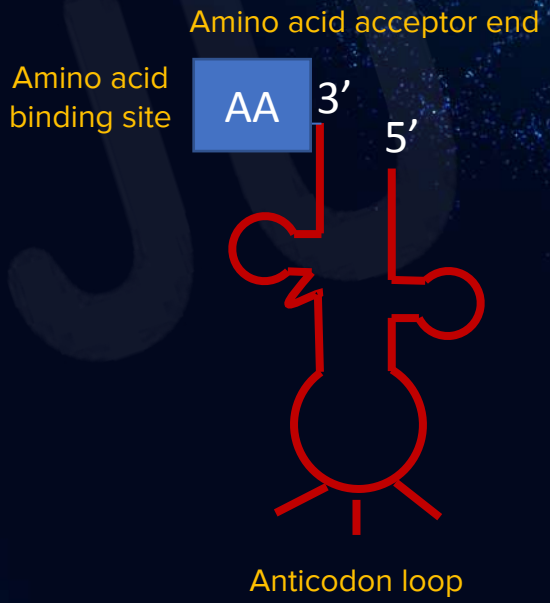


- The central dogma illustrates the flow of genetic information in cells.
- The genetic information in DNA is transcribed into RNA and then translated to form proteins.
- DNA duplicates by replication and it is inherited.
- Viruses can synthesise DNA by taking RNA as a template by the process of reverse transcription.

Summary

tRNA or transfer RNA structure

- Anticodon loop helps to identify the codon for the corresponding amino acid it is carrying, in the mRNA.



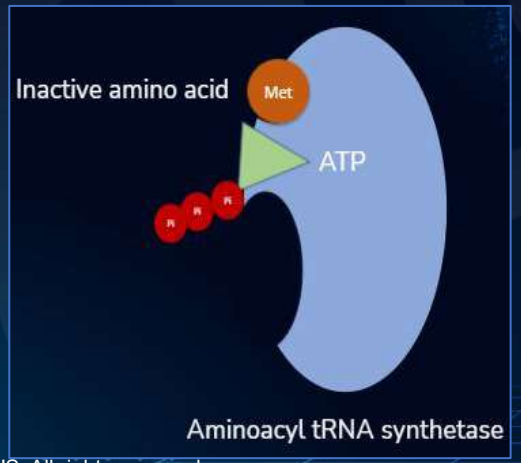
Diagrammatic representation of structure of tRNA

Summary

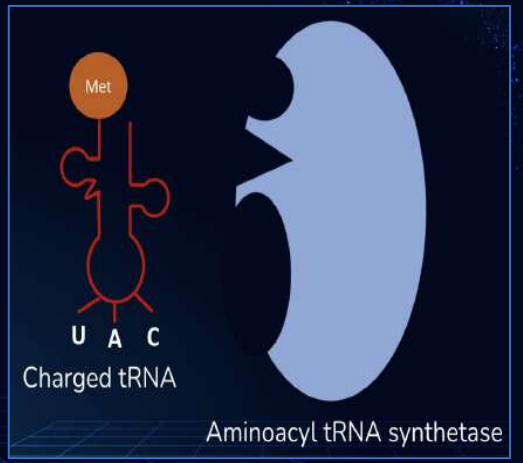
tRNA charging

- Inactive free amino acid binds to aminoacyl tRNA synthetase, with the help of energy from ATP, to form a complex.
- The uncharged tRNA arrives and binds to the complex.
- The enzyme transfers amino acid to the tRNA and releases the AMP, thus charging the tRNA for translation.

1 Activation of amino acids



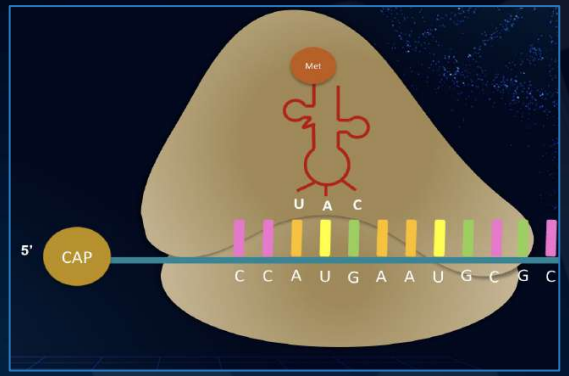
2 Charging of tRNA



Summary

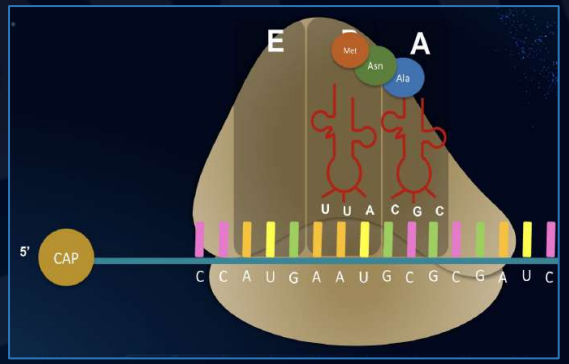
Steps of translation

1 Initiation



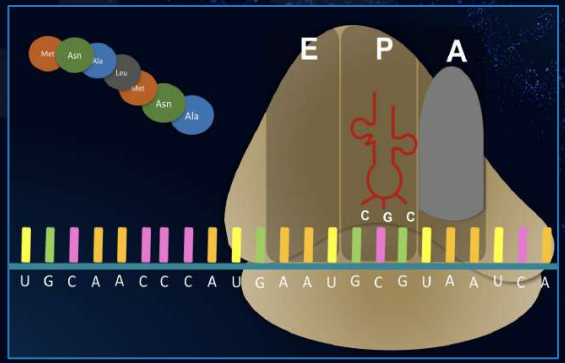
- Assembly of mRNA, ribosome and the initiator tRNA

2 Elongation



- Polymerisation of amino acids

3 Termination



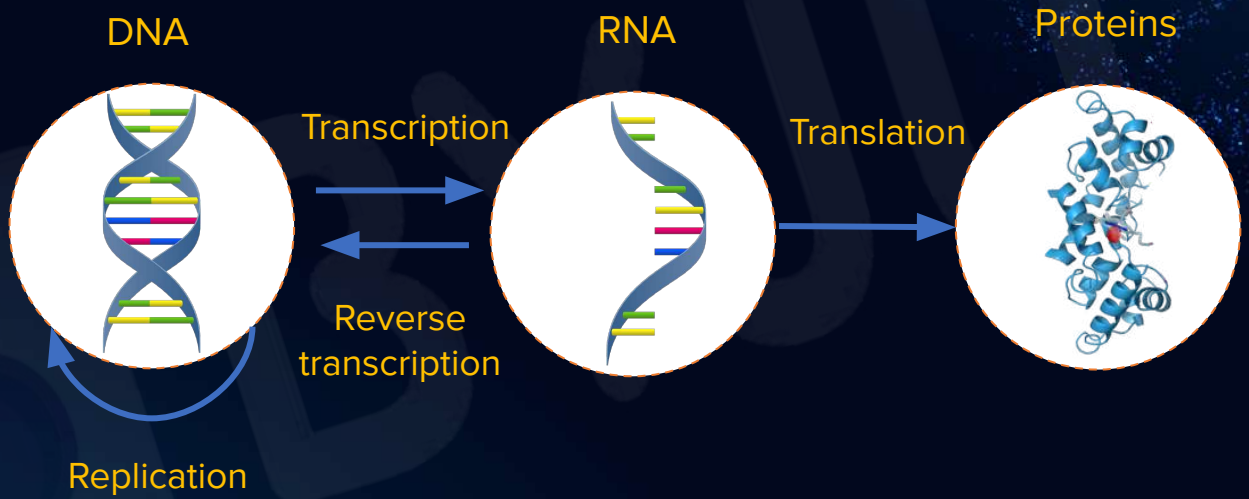
- Release of the polypeptide and disassembly of ribosomes and tRNA

Summary

Untranslated region (UTR)

Summary

Central dogma



BYJU'S Classes Notes

Molecular Basis of Inheritance

Gene Expression, Types of genes, Lac operon

Key Takeaways

1. Prokaryotic gene regulation
2. Lac operon
3. Eukaryotic gene regulation

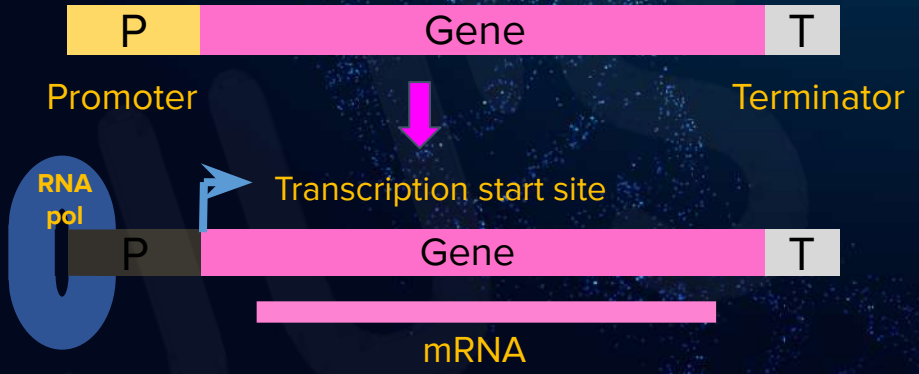
Summary

Recall! Central Dogma

- DNA is the blueprint of all the information.
- DNA recreates itself by the process of replication.
- The RNA is formed from DNA by the process of transcription.
- Information in RNA is decoded and translated into protein.

Recall – Transcription Unit

- Transcriptional unit is present in the DNA.
- It consists of promoter, gene and terminator region.



Promoter

- Present upstream (i.e. before) the gene where RNA polymerase comes and binds.

Gene

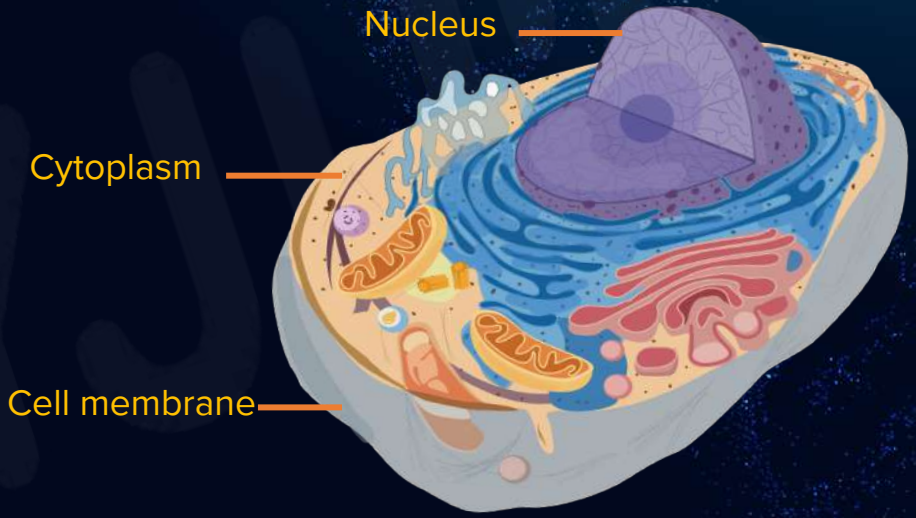
- Gene is part of DNA which codes for the mRNA and hence protein.

Terminator

- Terminator is present downstream i.e. after the gene.
- As soon as RNA polymerase reaches the terminator region it falls off and transcription stops.

Recall - Eukaryotic Gene Regulation

- Since the process of gene expression in eukaryotes has multiple steps hence there are multiple stages at which gene expression can be regulated.



Eukaryotic cell

Gene Expression

It is the process by which genetic information stored in the DNA is converted into protein within the cell.

- Information of protein formation is present in the DNA in the form of genes.
- These genes are expressed into protein.
- Hence, this process is also known as gene expression.

Regulation of gene expression

- Regulation of gene expression is also known as gene regulation.

It is a process of turning gene expression on or off.

Prokaryotic Gene Regulation

- Even the simplest of the organism, such as prokaryotes do not express all its gene every time.
- Prokaryotes also adapt to their environment immediately for their survival.
- They do so by regulating their gene expression and producing only required proteins for the required circumstances.
- This helps them to conserve energy which is used for other important functions of reproduction and multiplication.

Prokaryotic Gene Regulation

- There are two types of genes based on gene expression:
  - Constitutive genes are expressed continuously.
  - Non – constitutive genes are not expressed continuously.

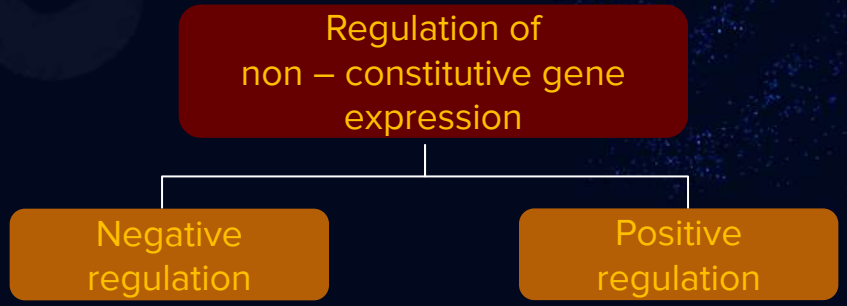
Constitutive genes	Non – constitutive genes
• Expressed all the time • Also called housekeeping genes • Proteins translated from these genes always required by the cell • Expression of such genes is not regulated	• Not expressed all the time • Also called luxury genes • Proteins translated from such genes required occasionally by the cell • Expression of such genes is regulated

Prokaryotic Gene Regulation

Non-constitutive genes

- Expression of non - constitutive genes is regulated by regulatory proteins.
- They are also called accessory proteins.
- They affect the ability of RNA polymerase to recognize the transcriptional start site.

Negative regulation	Positive regulation
• If regulatory protein affects RNA polymerase activity negatively, then it is called negative regulation.	• If regulatory protein affects RNA polymerase activity positively, then it is called positive regulation.



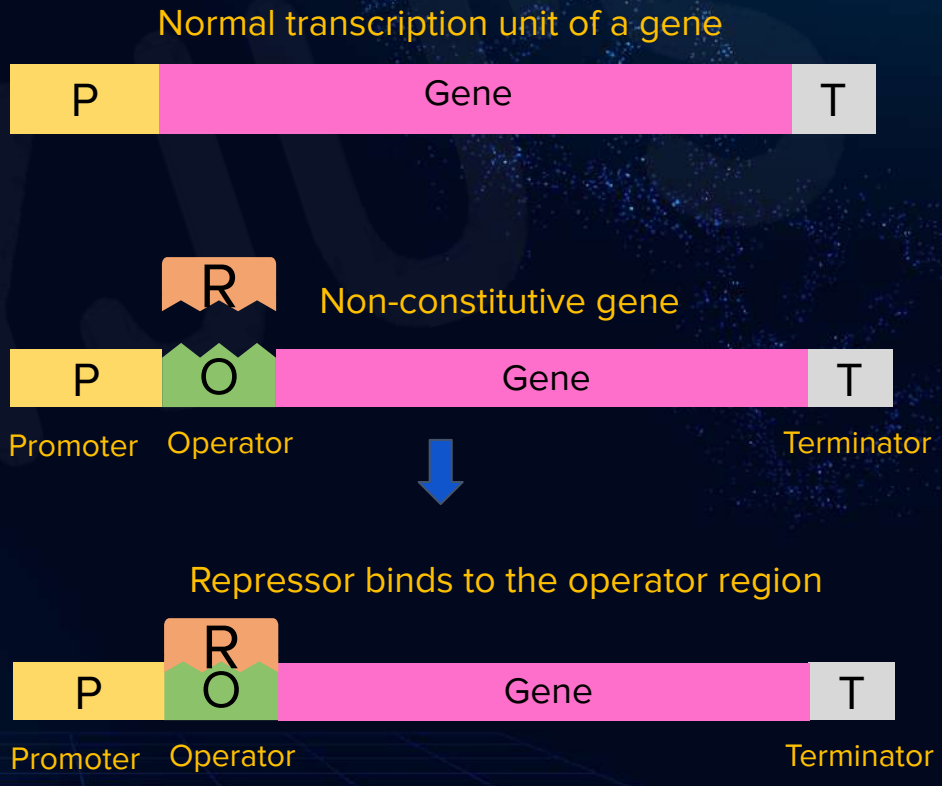
Prokaryotic Gene Regulation

Non-constitutive genes

- Non – constitutive genes have transcriptional unit with an operator region between promoter and gene.
- Operator region is very important region for regulation of gene expression because repressor protein comes and binds at the operator region.

Repressor protein

- ❖ It is a kind of regulatory protein.
- ❖ It is produced constitutively (actively).
- ❖ it binds to the operator region.

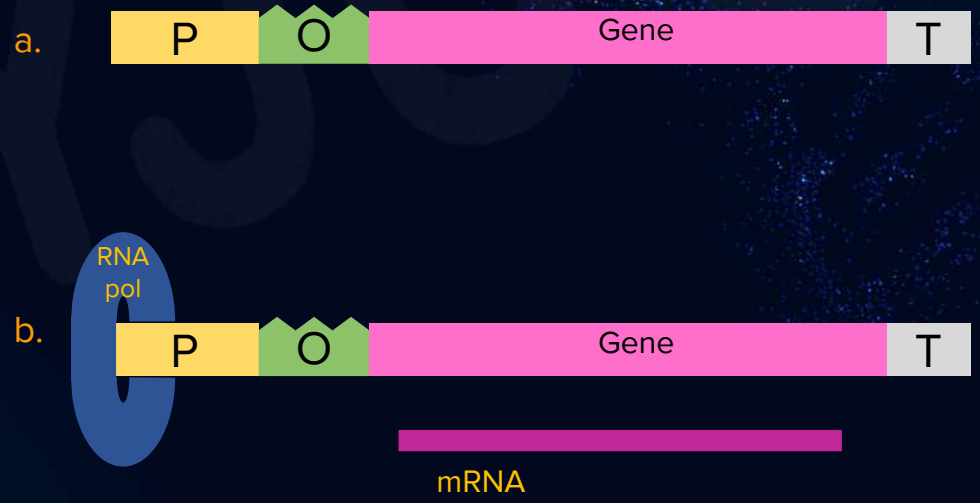


Prokaryotic Gene Regulation

There can be 3 possible scenarios in case of gene regulation in non-constitutive genes.

Case 1 : Repressor not bound to operator region

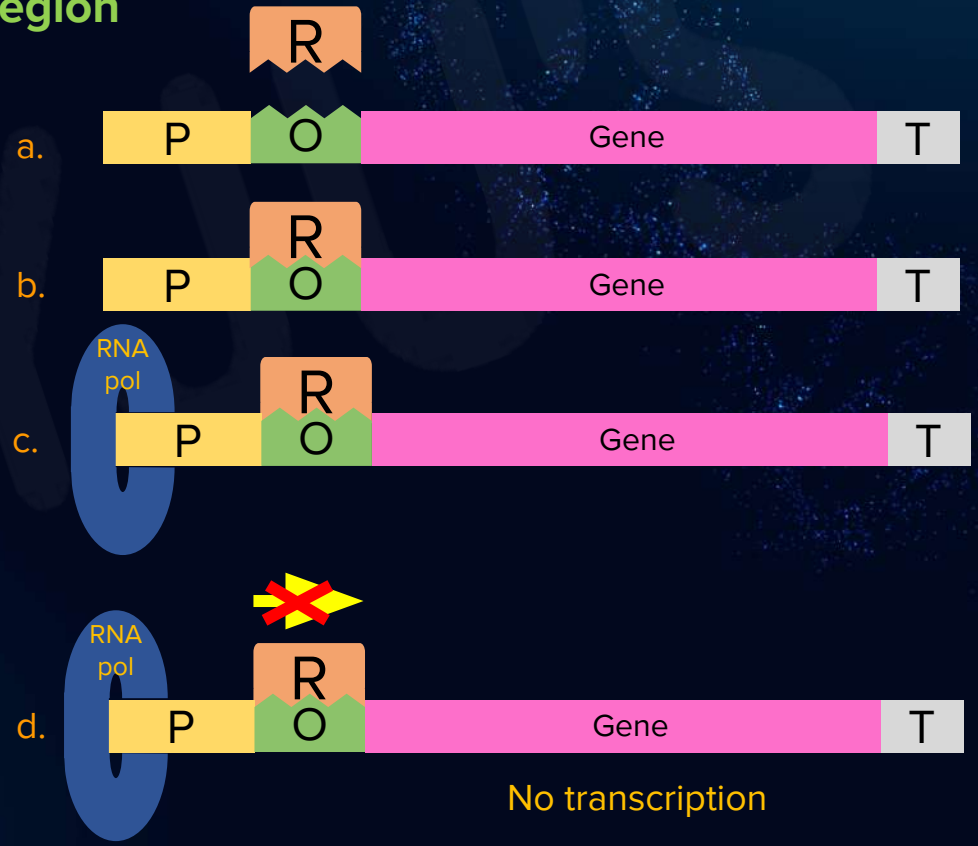
- RNA polymerase binds promoter region, it moves through the promoter and the operator region.
- Once it crosses operator region, transcription occurs as usual and the rate of transcription is normal.



Prokaryotic Gene Regulation

Case 2 : Repressor bound to operator region

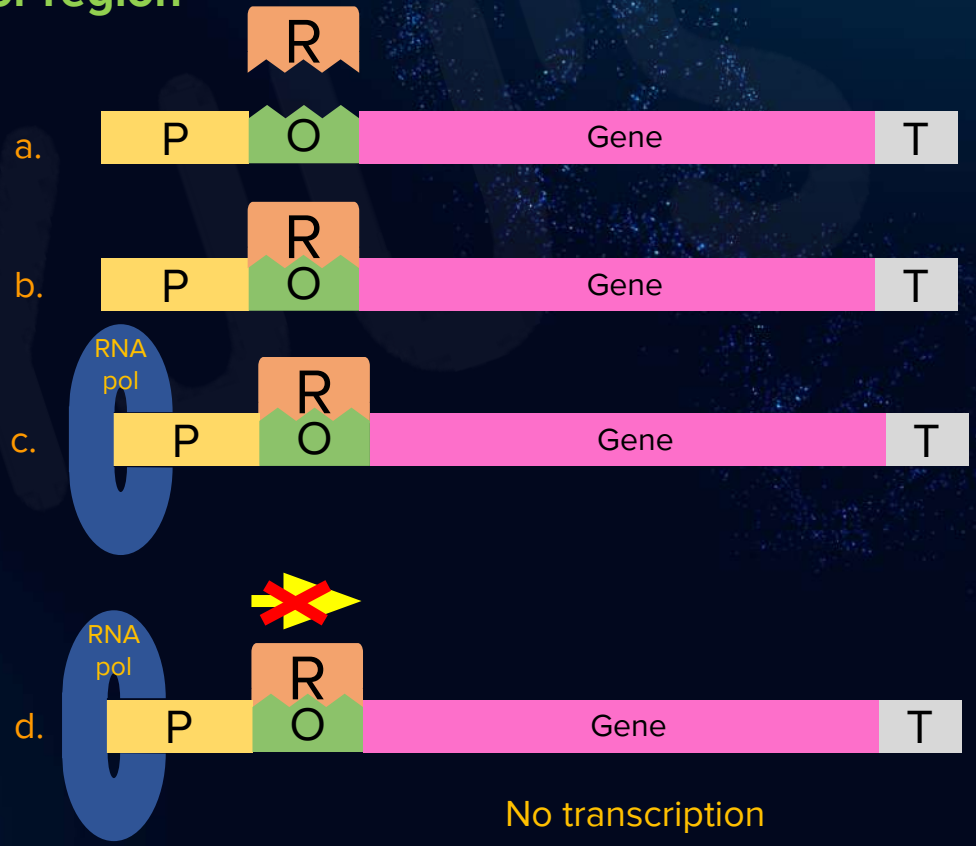
- Repressor protein binds at the operator region (fig. b).
- RNA polymerase binds to promoter and moves further (fig. c).
- Since repressor is present on the operator, RNA polymerase cannot move further (fig. d).
- Hence RNA polymerase falls off.
- So, no transcription.



Prokaryotic Gene Regulation

Case 2 : Repressor bound to operator region

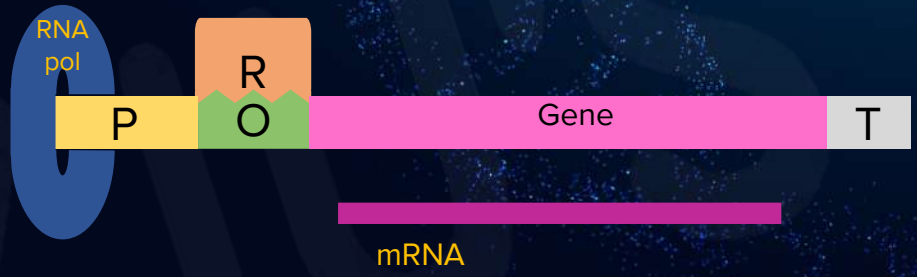
- Once bound to the operator region, the repressor does not get fixed over there permanently.
- Every protein has a fixed life span.
- So, repressor protein bound to operator region is degraded after some time and is replaced by another repressor protein.
- Thus, repressor protein should be constitutively produced.



Prokaryotic Gene Regulation

Case 3 : RNA polymerase sneaks before repressor binds to the operator region

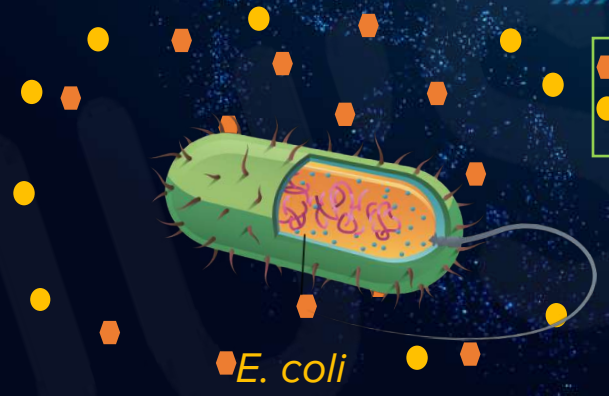
- Sometimes RNA polymerase sneaks in and crosses operator region before the new repressor binds to the operator.
- Therefore, transcription occurs.
- However, this happens very rarely.
- Because of this rare event, low level of transcription keeps occurring even in the presence of repressor.
- The transcribed mRNA is translated into the respective protein.
- Hence, this helps in maintaining basal level of proteins in the cell at any given time.



Negative regulation

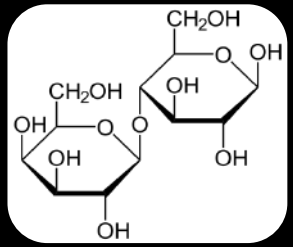
Presence of repressor affects the functioning of RNA polymerase negatively hence decreasing the rate of transcriptional initiation.

Lac Operon



- *E. coli* shows similar kind of negative regulation for its lactose metabolism.
- *E. coli* prefers glucose over lactose as energy source.
- However, sometimes in absence of glucose, lactose has to be utilized by *E. coli* as a substitute of energy.
- Lactose/ β galactoside is a dimeric sugar (Disaccharide) consisting of glucose and galactose.

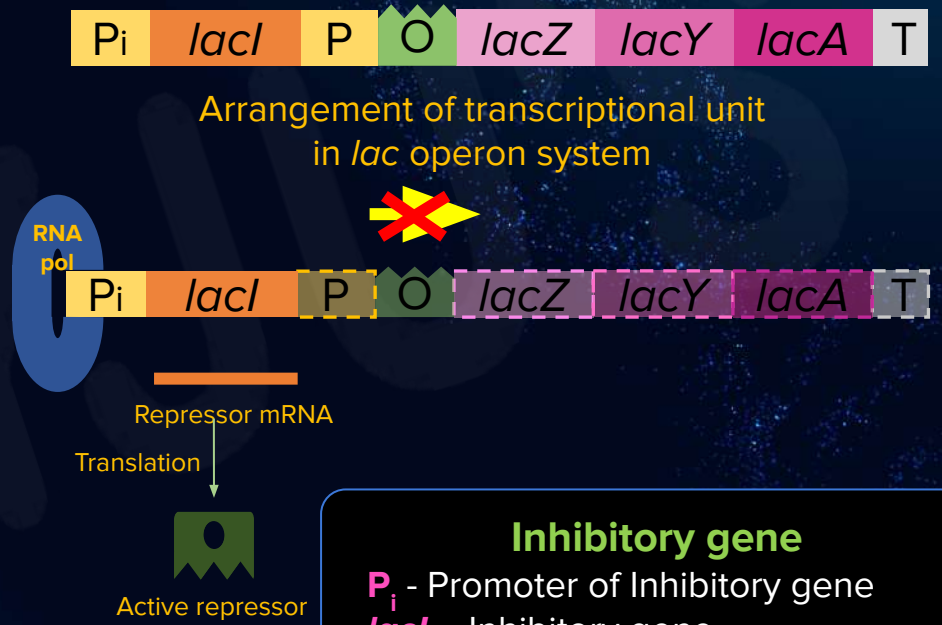
Lactose



- β – galactoside
- Contains **glucose** and **galactose**

Lac Operon

- RNA polymerase bind to the promoter of inhibitory gene and transcribes repressor mRNA .
- Repressor mRNA is translated into active repressor protein.
- Active repressor protein has affinity for operator region, so it goes and binds to the operator region.
- Meanwhile another repressor protein is already translated and formed.
- Since repressor is present at operator, RNA polymerase cannot move forward and hence it falls off.



Inhibitory gene

P_i - Promoter of Inhibitory gene
lacI – Inhibitory gene

- Regulatory gene for *lac* operon
- Expressed constitutively
- **C**odes for **repressor protein**

Lac Operon

Promoter and operator

P - Promoter of *lac* gene
O - Operator of *lac* gene where the repressor protein binds

lac gene

- Structural gene which codes for a polycistronic *lac* mRNA and lactose metabolizing enzymes
- A non – constitutive gene of *lac* operon

Lac Operon

lacZ gene

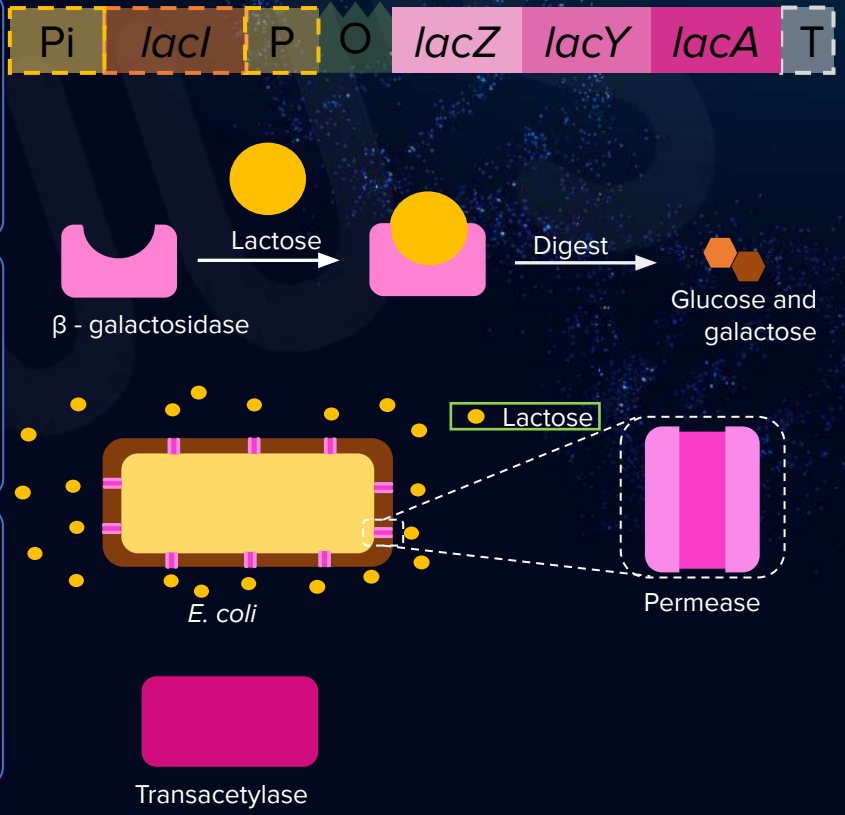
- Lac Z gene codes for β – galactosidase enzyme.
- Lactose binds to the active site of β – galactosidase.
- Lactose gets digested here into glucose and galactose.

lacY gene

- Lac Y gene codes for permease enzyme which is a cell membrane bound enzyme.
- It make the cell membrane of *E.coli* permeable to lactose.

lacA gene

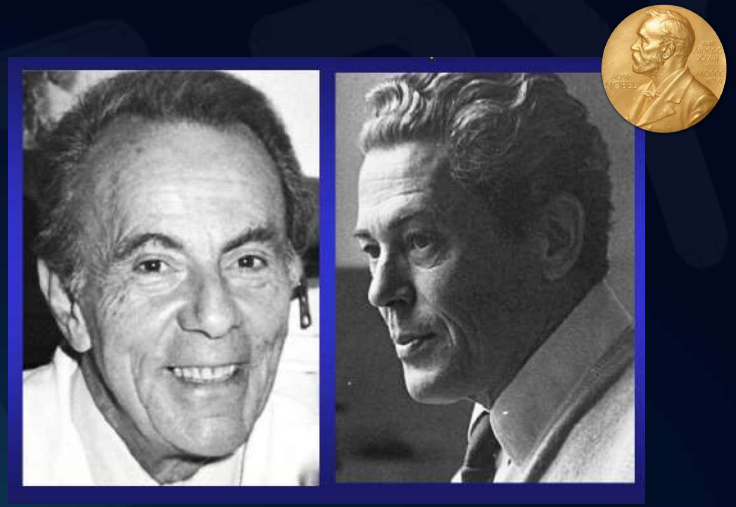
- Lac A gene codes for transacetylase enzyme.
- It helps in trans - acetylation reaction.
- Other functions of transacetylase is not known in great detail.



Lac Operon

Operon

Operon is defined as a system where polycistronic structural gene is regulated by a common promoter and regulatory protein.
E.g. *lac* operon, *trp* operon, *ara* operon, *his* operon, *val* operon etc.



Francois Jacob (Geneticist)

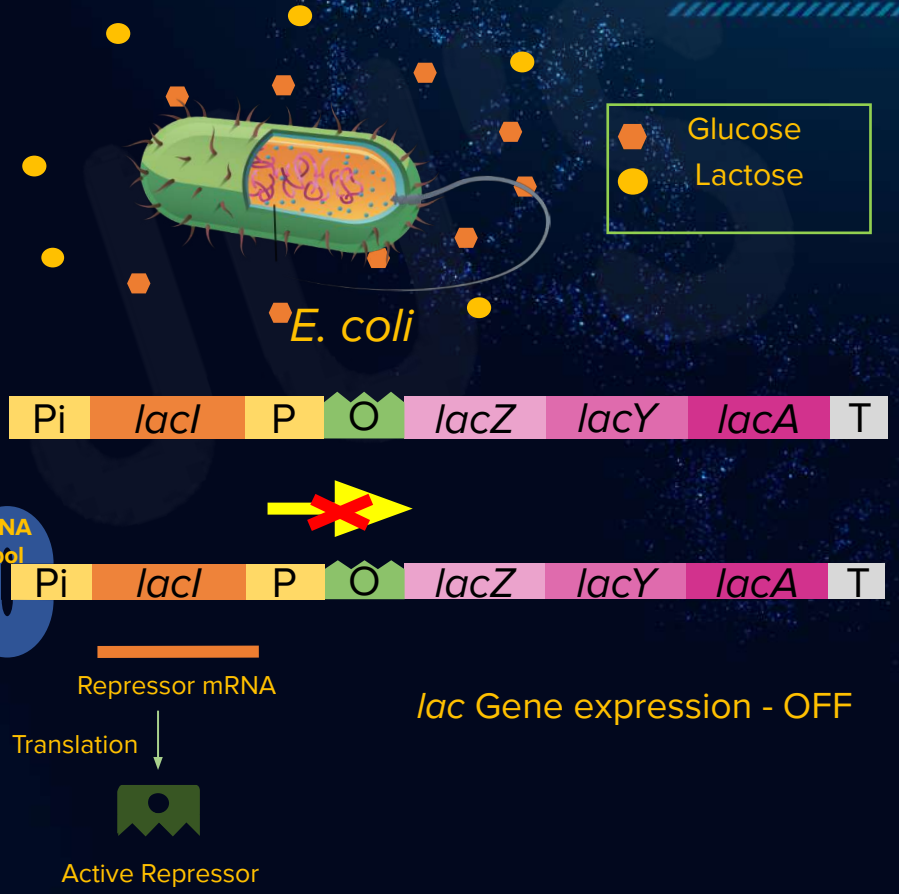
Jacque Monod (Biochemist)

- Lac Operon was discovered by Francois Jacob and Jacque Monod.
- They shared Nobel prize for the discovery of Lac operon.

Lac Operon

Scenario 1 : *E. coli* does not feed on lactose

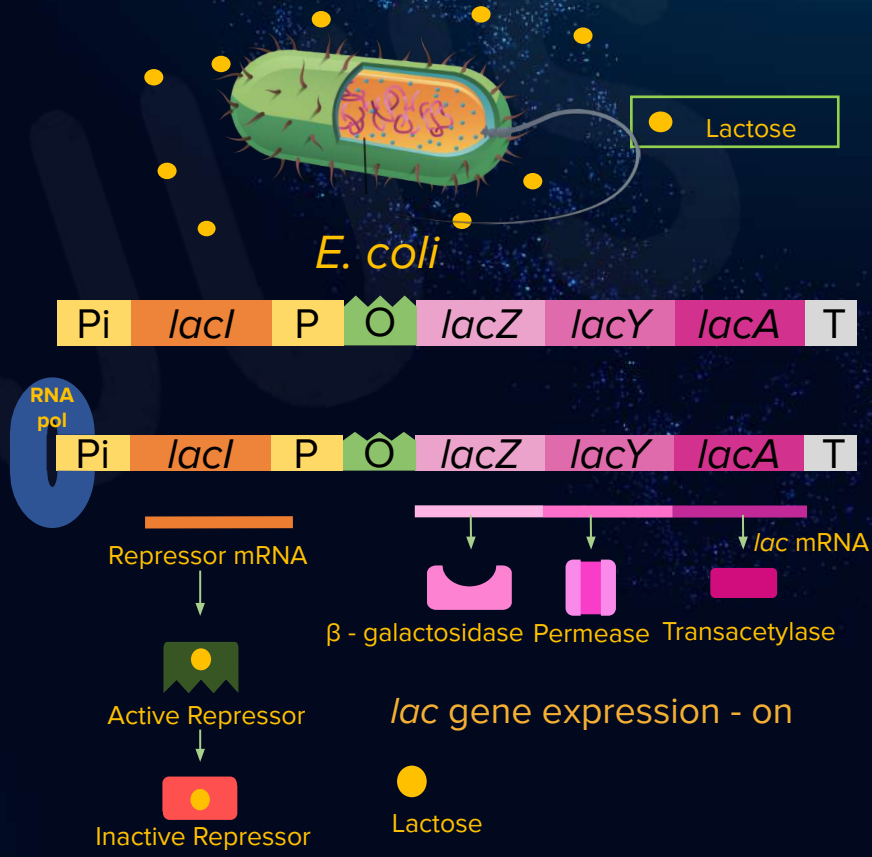
- *E. coli* does not feed on lactose normally because *E. coli* prefers glucose over lactose.
- RNA polymerase binds to the promoter of inhibitory gene, transcribes repressor mRNA which forms active repressor. which binds to the operator region and RNA polymerase cannot move further.
- So, no transcription of lac genes occurs.
- Hence lac gene expression is off.



Lac Operon

Scenario 2 : *E. coli* feeds on lactose

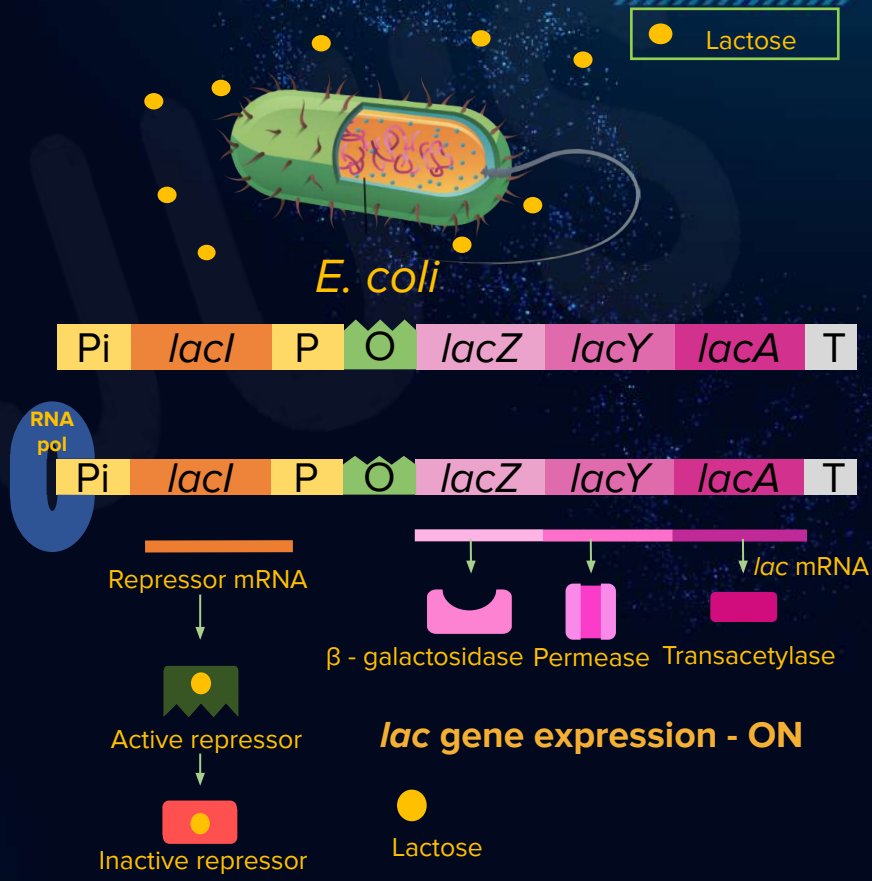
- When there is no glucose present in the environment, *E. coli* feeds on lactose as an alternative source of food for obtaining energy.
- RNA polymerase binds to the promoter of inhibitory gene, and transcribes repressor mRNA which forms active repressor.
- Lactose binds to the active repressor and makes it inactive.



Lac Operon

Scenario 2 : *E. coli* feeds on lactose

- The inactive repressor cannot bind to the operator now because it has lost its affinity.
- Hence operator region is free and RNA pol can move further to express lac gene.
- The lac mRNA is transcribed.
- Thus, lac gene expression is switched ON.
- This completes the negative regulation of gene expression in prokaryotes.



Lac Operon

- In lac operon, lactose is the inducer molecule because only when lactose binds to the active repressor it turns inactive.
- Inactive repressor cannot bind to the operator region and hence cannot block the path of RNA polymerase.
- RNA polymerase moves freely and transcribes the lac gene.
- Hence, this induces the expression of lac gene.

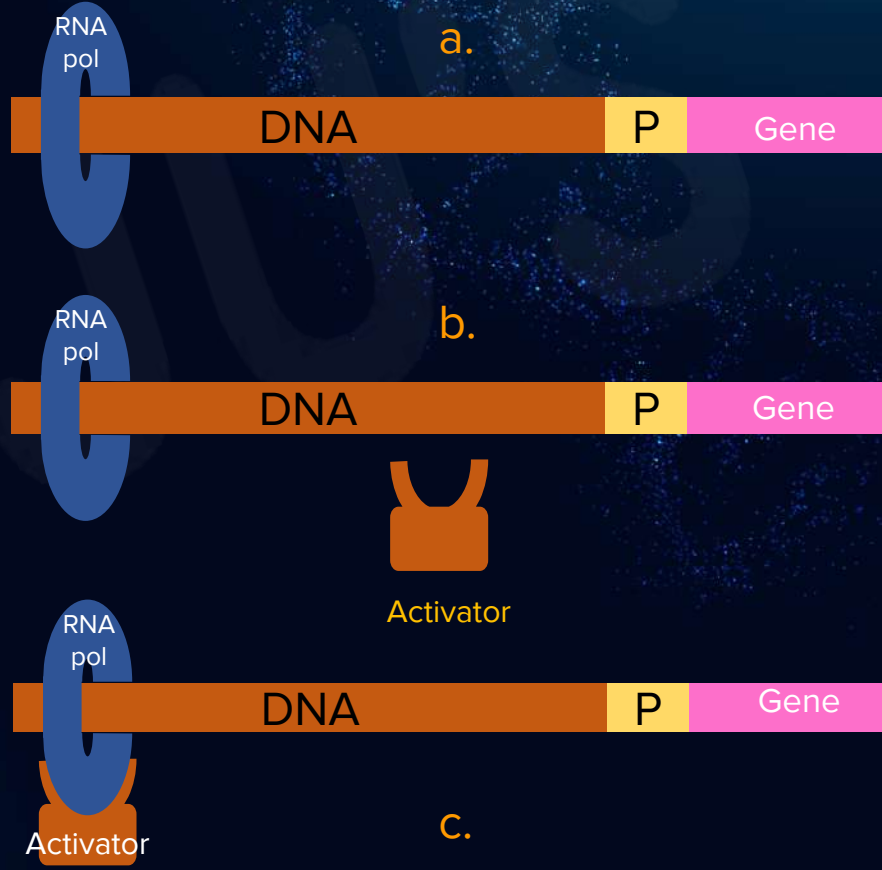
Scenario	Lactose in *E. coli*	Regulatory protein (repressor)	Regulation of *lac* gene	Gene expression
1	Absent	Active	Negative	Switch OFF
2	Present	Inactive	No	Switch ON

Did You Know?

Lactose has an isomeric form called allolactose which also acts as inducer and inactivates the repressor on binding.

Prokaryotic Gene Regulation

- Normally RNA polymerase binds to the DNA and scans for the promoter region in the entire DNA.
- In positive regulation, a regulatory protein called activator is present in the cell.
- The activator protein binds to the RNA polymerase and helps it to reach the promoter region faster.
- Thus, the rate of transcriptional initiation increases.
- This completes the positive regulation of gene expression in prokaryotes.



Prokaryotic Gene Regulation

Case	Type of regulation on RNA polymerase	Regulatory protein	Rate of transcriptional initiation
1	No regulation	-	Normal
2	Positive regulation	Activator	Faster than normal
3	Negative regulation	Repressor	Slower than normal

Which enzyme will be produced in a cell if there is a nonsense mutation in the *lacY* gene?

a) Transacetylase

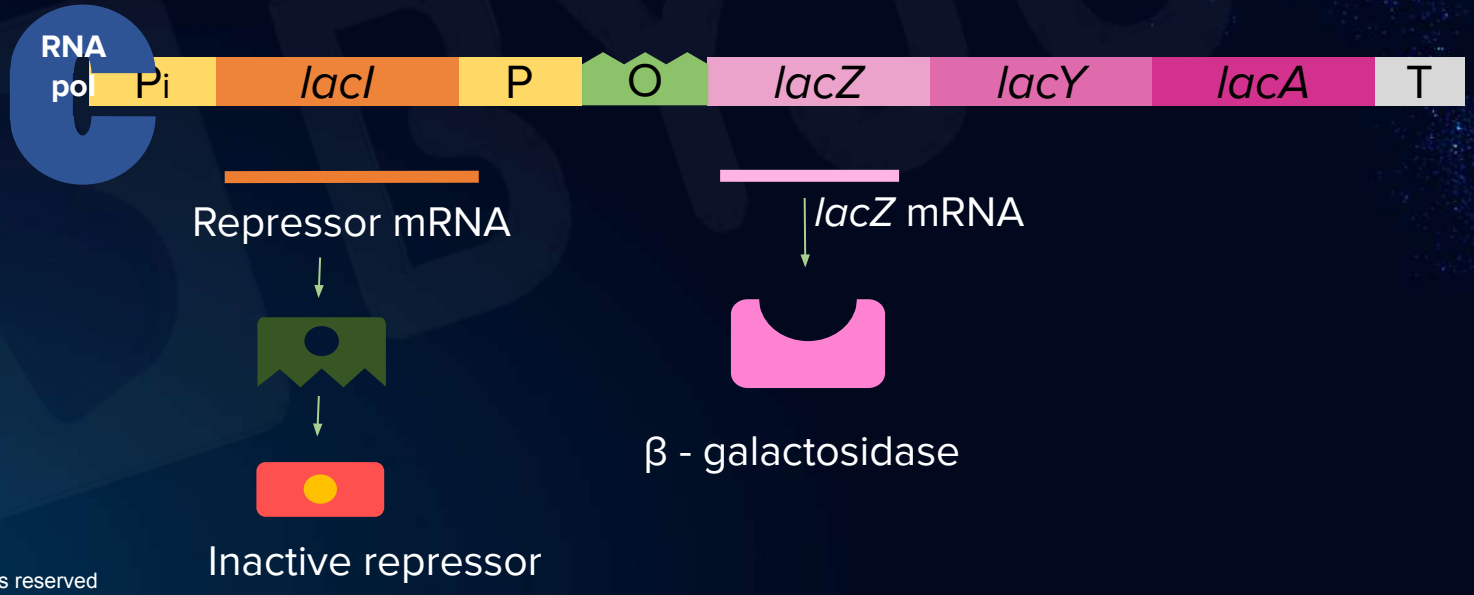
b) Lactose permease and transacetylase

c) Lactose Permease

d) β – Galactosidase

Discussion

- Nonsense mutation is a mutation in which a stop codon arises in a gene.
- In question, nonsense mutation has occurred in lacY gene.
- Therefore, RNA polymerase stops when it encounters stop codon in the lacY gene.
- The lacZ mRNA formed is translated into protein.
- Hence only β – galactosidase enzyme is produced in the cell.



Which enzyme will be produced in a cell if there is a nonsense mutation in the *lacY* gene?

a) Transacetylase

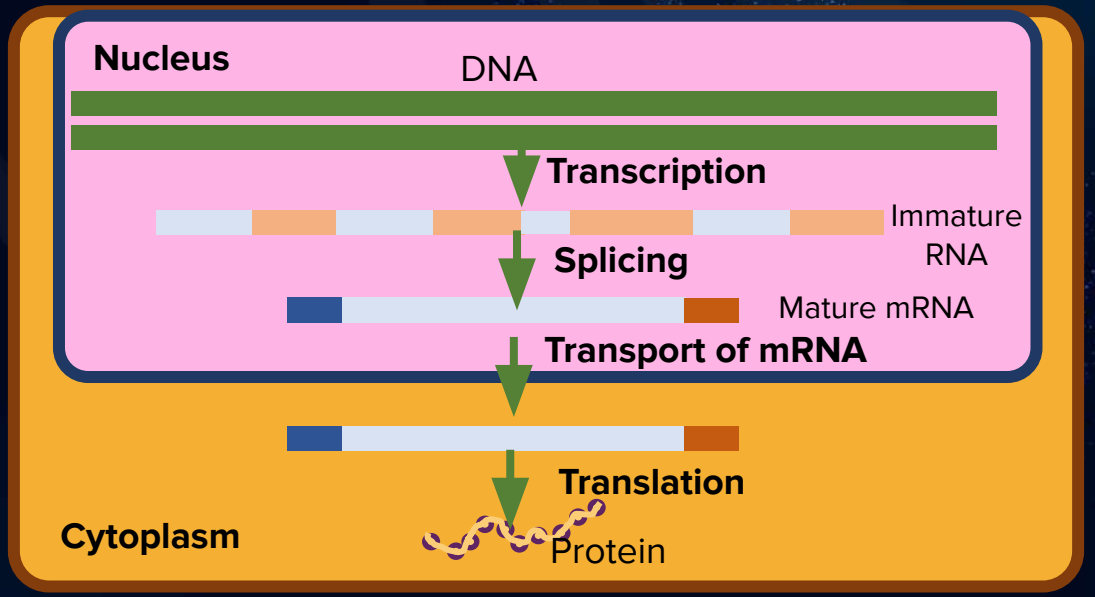
b) Lactose permease and transacetylase

c) Lactose Permease

d) β – Galactosidase

Eukaryotic Gene Regulation

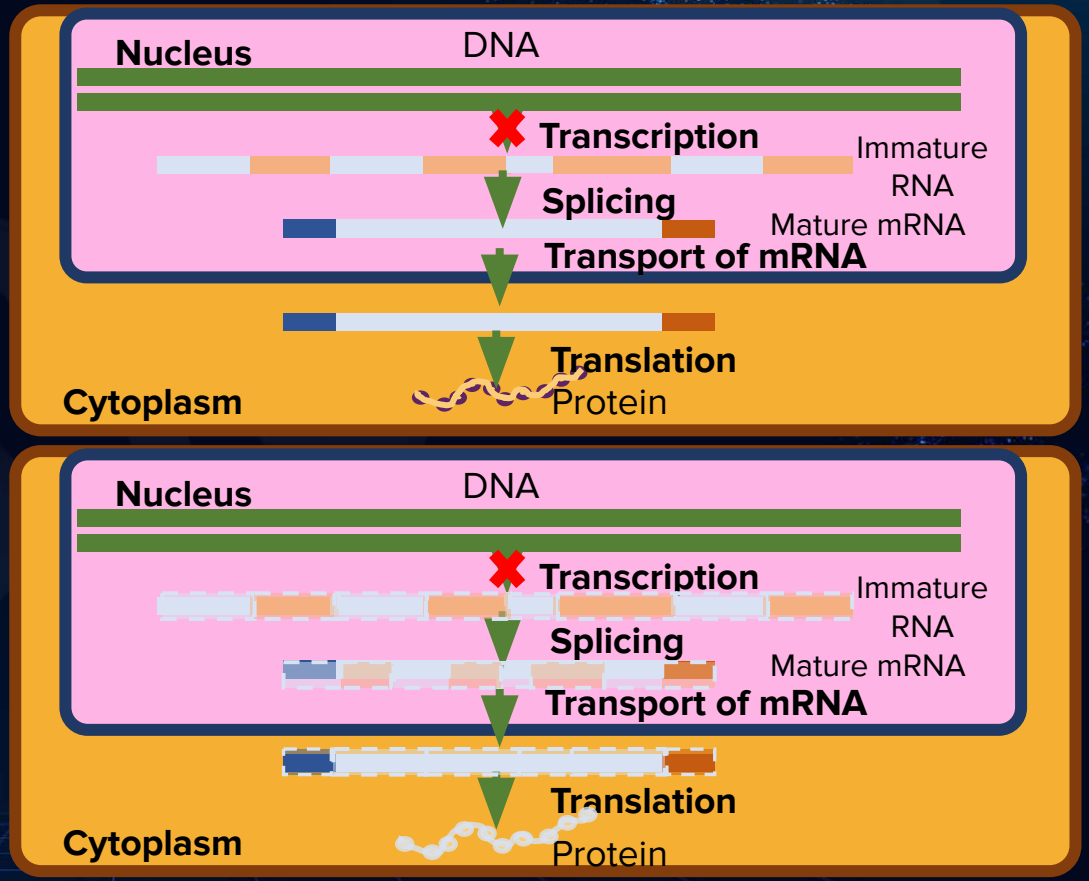
- DNA in the nucleus undergoes transcription to produce healthy primary transcript called immature RNA which undergoes process of splicing, capping and tailing to give rise to mRNA.
- This processing is must for transporting mRNA into cytoplasm.
- So, the mRNA formed is transported from the nucleus into the cytoplasm because translation can occur only in cytoplasm.
- In cytoplasm, mRNA is translated into protein.



Eukaryotic Gene Regulation

Stage 1

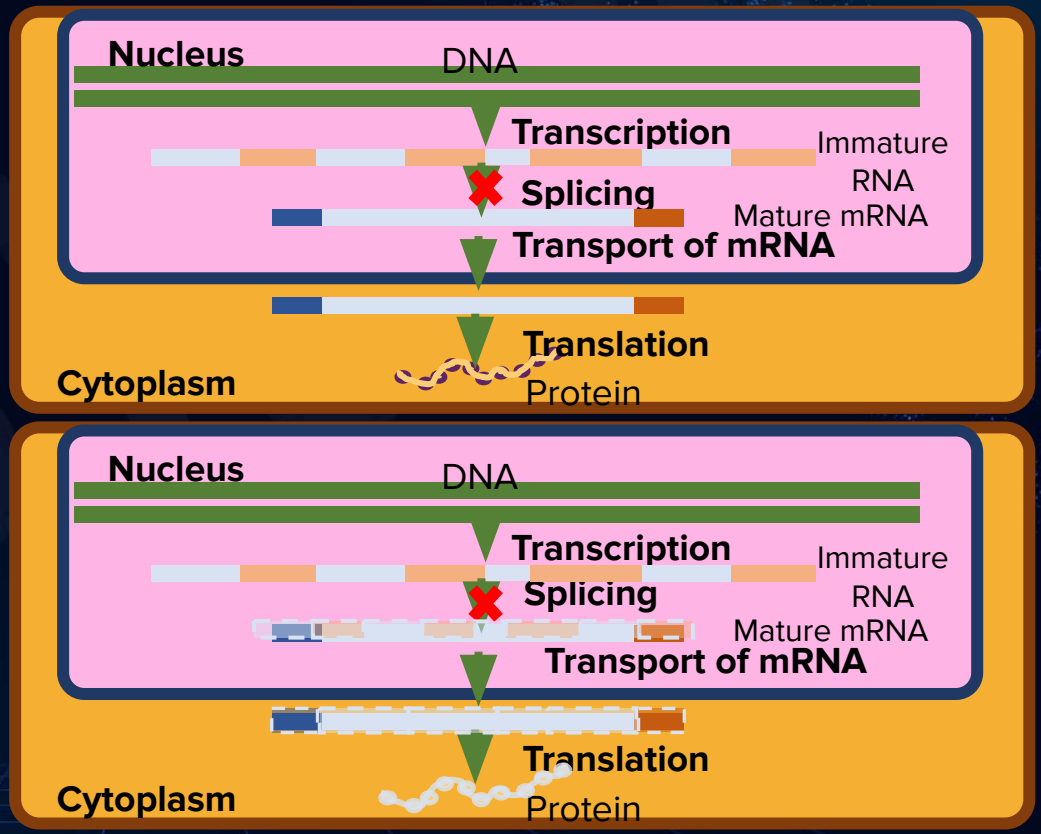
- If the process of transcription is blocked somehow, then no transcription will occur.
- Hence, immature mRNA will be formed.
- Hence no further processes.
- Thus, no protein formed.



Eukaryotic Gene Regulation

Stage 2

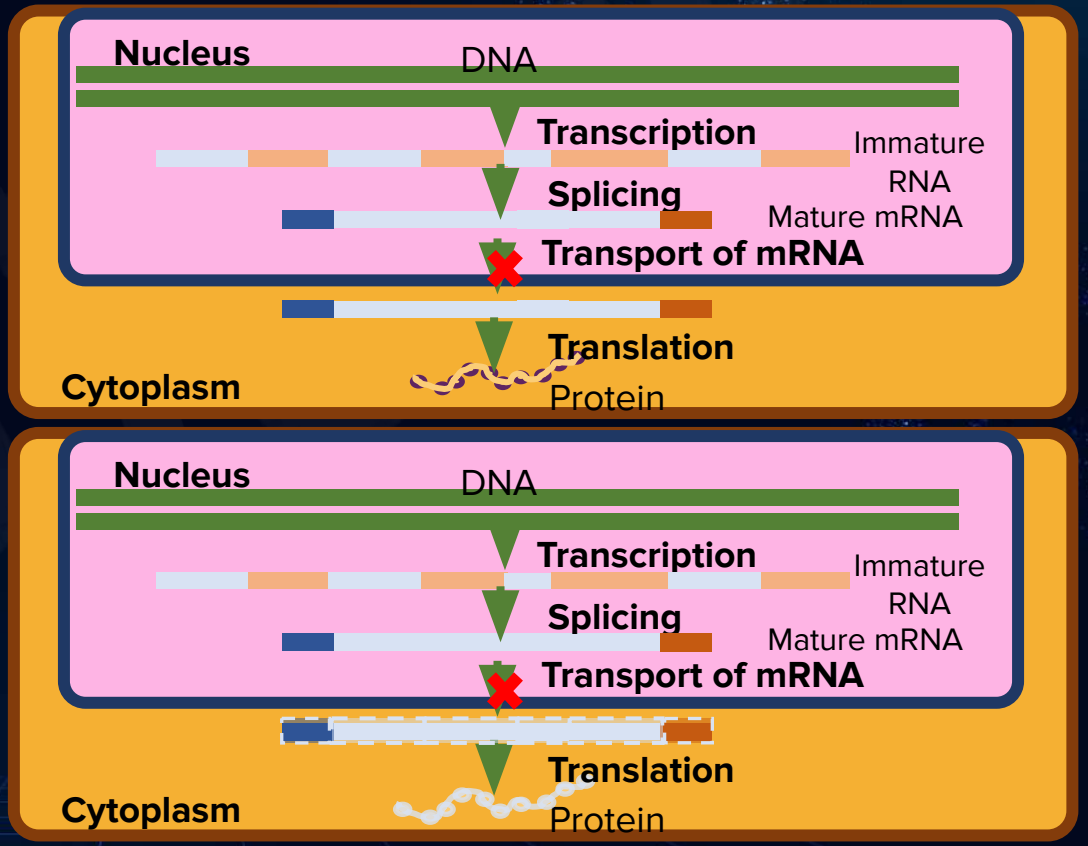
- Suppose the immature mRNA is formed, even after that eukaryotes have chance to regulate gene expression.
- It can be stopped by blocking the splicing process.
- Therefore, no mRNA will be formed.
- No further processes will occur.
- Hence, no protein will be formed.



Eukaryotic Gene Regulation

Stage 3

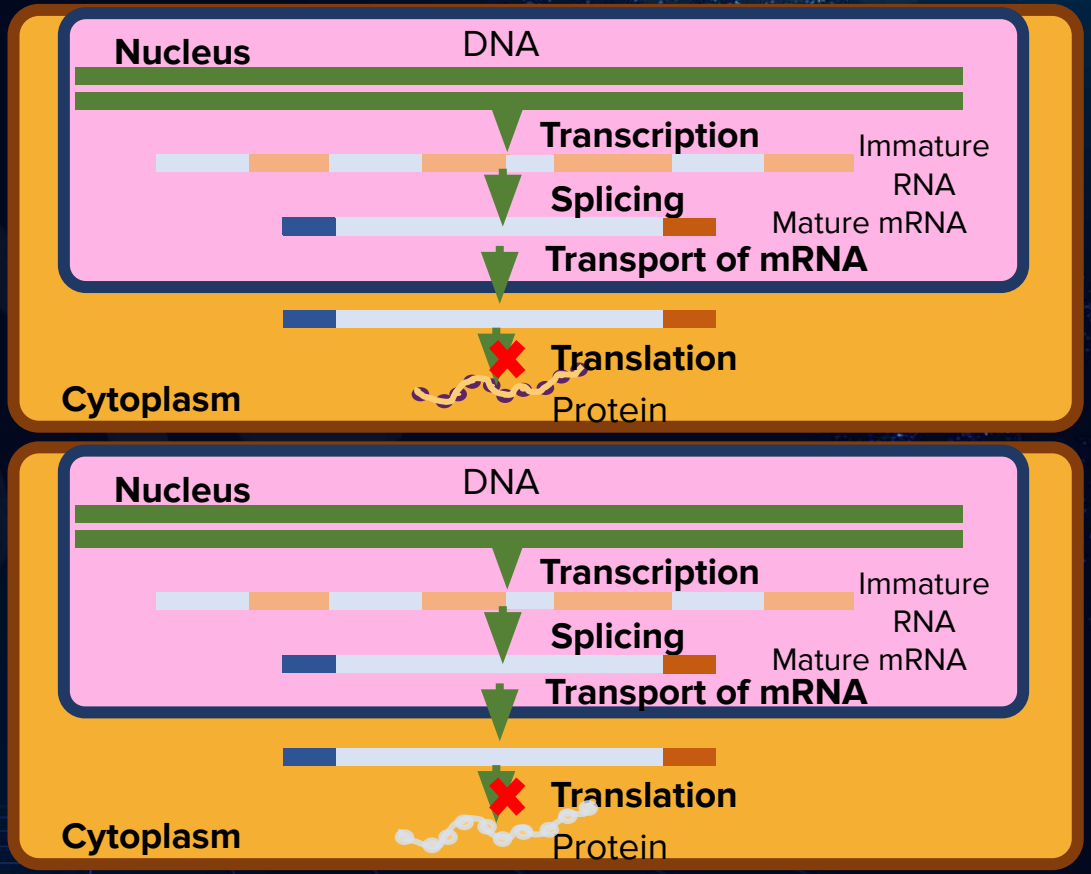
- Mature mRNA formed in nucleus is transported to cytoplasm where translation occurs.
- In absence of successful transportation of mRNA to cytoplasm, no translation occurs and hence no protein is formed.



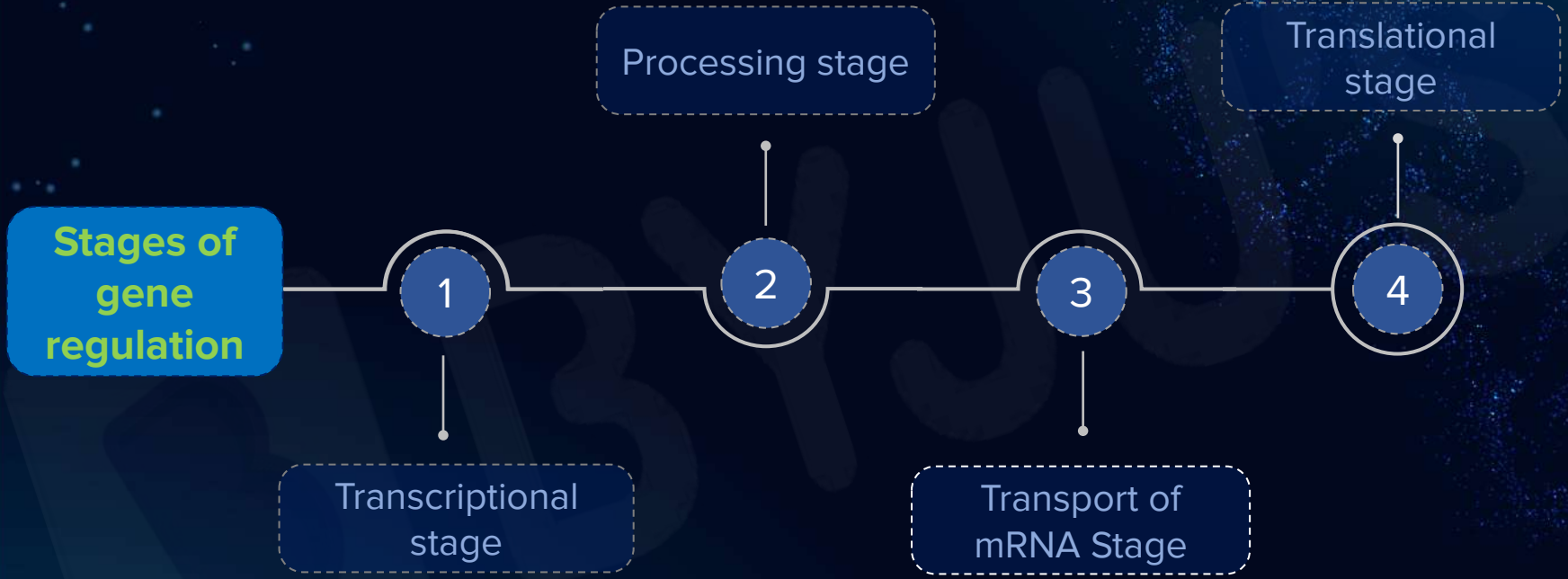
Eukaryotic Gene Regulation

Stage 4

- Suppose mature mRNA is transported to the cytoplasm.
- Still gene expression can be controlled by degrading mRNA in cytoplasm or not making mRNA available for translation.
- Thus, no translation will occur and hence no protein will be formed.

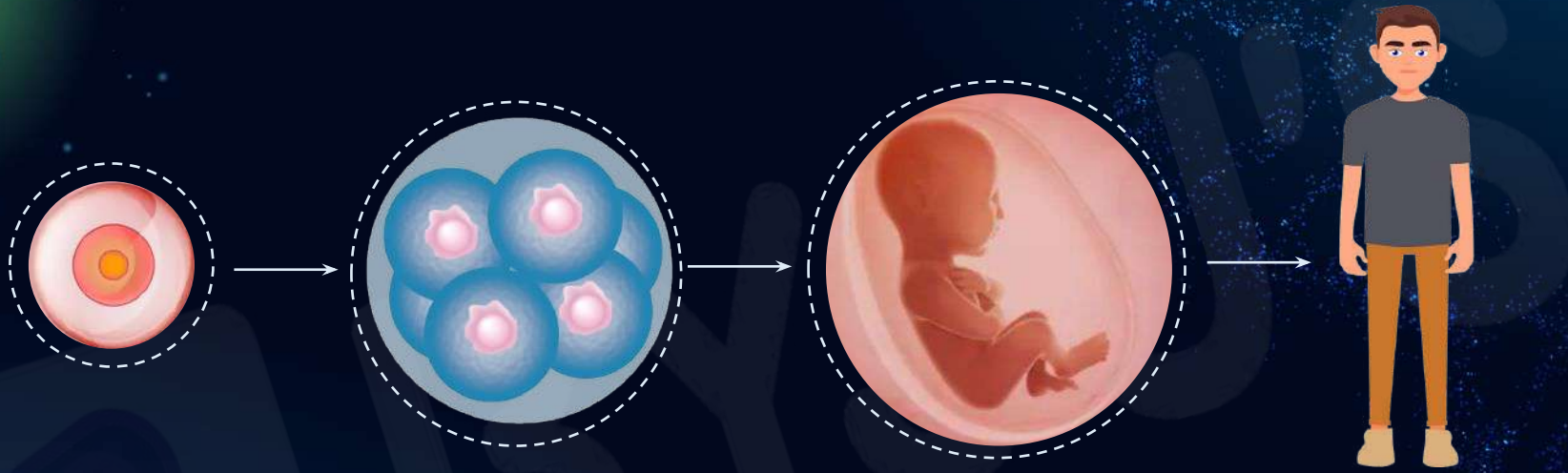


Eukaryotic Gene Regulation



- Eukaryotes regulate their gene expression very intricately and precisely at various stages.
- This way they conserve their energy by producing only what is needed.

Did You Know ?



Controlled and coordinated regulation of expression of several set of genes help embryo to develop into an adult.

Summary

- Process by which genetic information stored in the DNA is converted into protein within the cell is called gene expression.
- Regulation of gene expression is the process of turning gene expression on or off.
- There are two types of gene based on gene expression constitutive genes and non – constitutive genes.
- Constitutive genes are expressed continuously and non – constitutive genes are not expressed continuously.
- Regulation of non – constitutive gene expression can be positive or negative.

Summary

Prokaryotic gene regulation

- **Case 1 :** Repressor not bound to operator region - transcription will occur
- **Case 2 :** Repressor bound to operator region - since repressor is present on the operator, RNA polymerase cannot move further and there is no transcription.
- **Case 3 :** RNA polymerase sneaks before repressor binds to the operator region - If RNA polymerase sneaks in and crosses operator region before the new repressor binds to the operator then transcription occurs.

Lac operon

- Lac operon was discovered by Francois Jacob and Jacque Monod.
- Operon is defined as a system where polycistronic structural gene is regulated by a common promoter and regulatory protein.

Summary

Lac operon

Scenario 1 : *E. coli* does not feed on lactose

- *E. coli* prefers glucose over lactose.
- So, no transcription of lac genes occursand the lac gene expression is OFF.

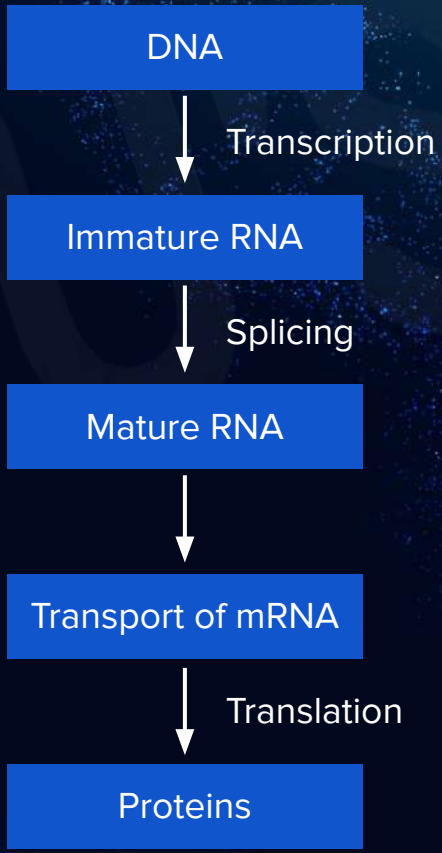
Scenario 2 : *E. coli* feeds on lactose

- When there is no glucose present in the environment, then *E. coli* feeds on lactose.
- The lac mRNA is transcribed and the lac gene expression is switched ON.
- This completes the negative regulation of gene expression in prokaryotes.
- In positive regulation, a regulatory protein called activator is present in the cell.
- The activator protein binds to the RNA polymerase and helps it to reach the promoter region faster, thus, increasing the rate of transcriptional initiation.

Summary

Eukaryotic gene regulation

- The process of gene expression in eukaryotes has multiple steps hence there are multiple stages at which gene expression can be regulated.
- It can be blocked during transcription process or splicing process or transportation process or translation process.
- Blockage of any of the steps will stop the production of proteins.



BYJU'S Classes Notes

Molecular Basis of Inheritance

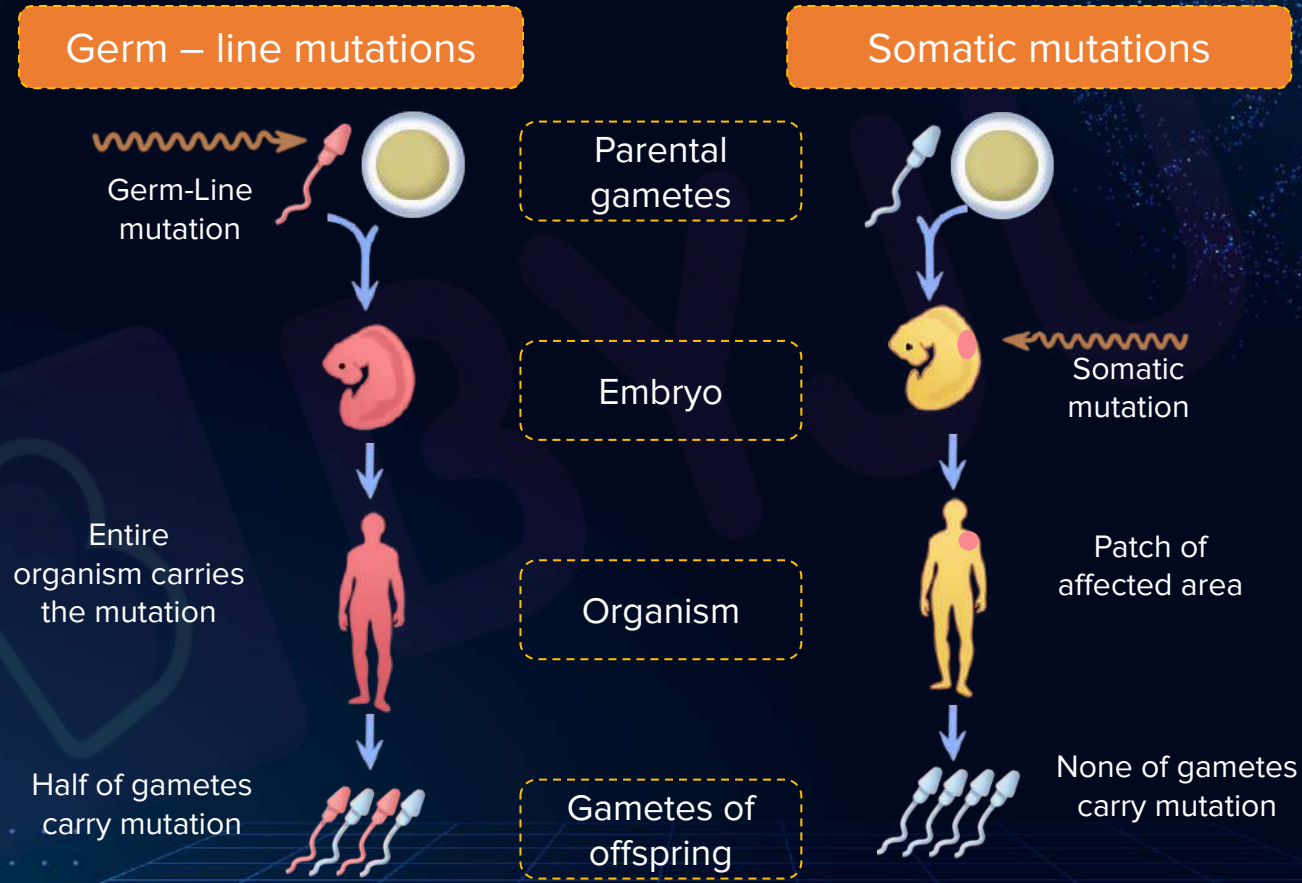
DNA Fingerprinting: Satellite DNA, VNTR, DNA Polymorphism, Steps in DNA Fingerprinting and Applications

Key Takeaways

1. Dna fingerprinting
2. Satellite DNA
3. DNA polymorphism
4. Steps in DNA fingerprinting
5. Applications

Summary

Recall ! Types of Mutation



Recall ! Intron and Exon

- Introns: Non-coding or intervening sequence
- Exons: Coding or functional sequence

- Primary transcript in eukaryotes has both introns and exons.

DNA Fingerprinting

Alec Jeffreys

- DNA fingerprinting is a technique used to determine the characteristic of an individual's DNA.
- It is used to compare DNA of two individuals.
- This method was discovered by Sir Alec Jeffreys.

DNA Fingerprinting

- Analyzing DNA of two different individuals:
  - 99.9% genome is similar
  - Differ by 0.1%
- 0.1% difference is used for DNA fingerprinting
- World population - More than 7.9 billion

Repetitive Sequences

- Repetitive sequences are repeats or repetitive elements that occur multiple times in the nucleic acid sequences(DNA/RNA).
- In introns, the sequences can be both repetitive and non-repetitive.
- The number of these repeats are different in different individuals.
- Thus, it is used in the technique of DNA fingerprinting.

5’ CTCATGATGATGATGATGTCATCCCGAAATCGTAGCTA 3’

Repetitive sequence

5’ CTTAGGATTCAATCCGATTCATCCCGAAATCGT 3’

Non-repetitive sequence

Repetitive Sequences

Person 1 (5 Repeats)

5’ CTCATGATGATGATGATGTCATCCCGAAATCGTAGCTA 3’

Person 2 (7 Repeats)

5’ CTCATGATGATGATGATGATGATGCGAAATCGTAGCTA 3’

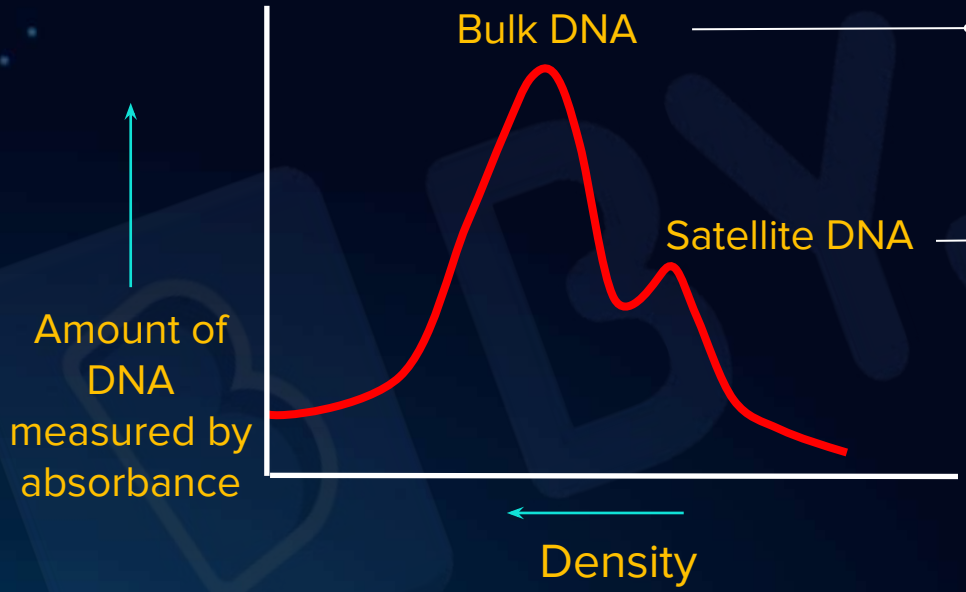
- The number of copies (repeats) in different chromosomes in an individual is also different, but in the overall genome will be constant for all cells in the body.

Repetitive Sequences

- Genome was taken and density centrifugation was performed.

- We obtain two bands based on density :
  - Lighter band - Less dense DNA-non coding sequence
  - Heavier band - bottom layer-coding sequence

Repetitive Sequences



Graphical representation

The main band is formed by the denser DNA (Bulk DNA) sequence which forms the larger peak.

Satellite DNA forms a lighter bands which is followed by the larger peak.

Satellite DNA

Definition: Highly repetitive DNA sequence that does not code for proteins used for DNA fingerprinting.

Classification criteria:

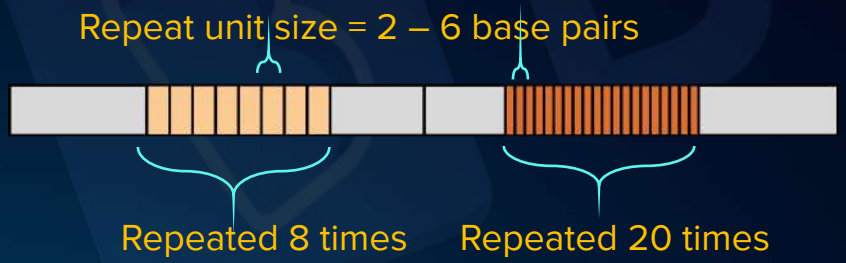
1. Length of sequence
2. Number of repetitive units
3. Base composition

Satellite DNA - Types

Micro satellite

2-6 base pairs repeating units in tandem repeats

Short Tandem Repeats

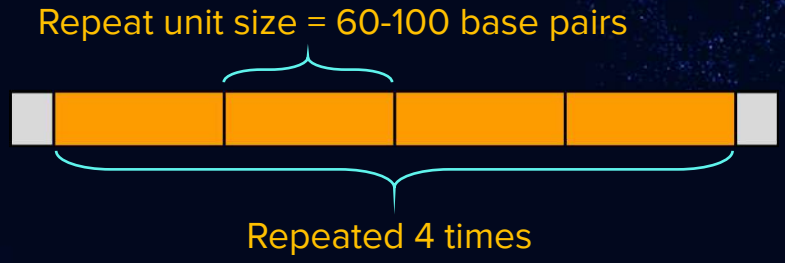


Short Tandem Repeats (STR)

Mini satellite

10-100 base pairs repeating units in tandem repeats

Variable Number Tandem Repeats



Variable Number Tandem Repeats (VNTR)

Satellite DNA - Types

Variable Number Tandem Repeats (VNTR)

- The number of repeats in VNTR show variation.
- Hence, they are commonly used in DNA fingerprinting.

Tandem Repeats

- Tandem means one after the other, if there is a break then it is not a tandem repeat.
- If there is a mutation, the number of repeats changes and hence the sequence ends up having different number of tandem repeats.

Satellite DNA

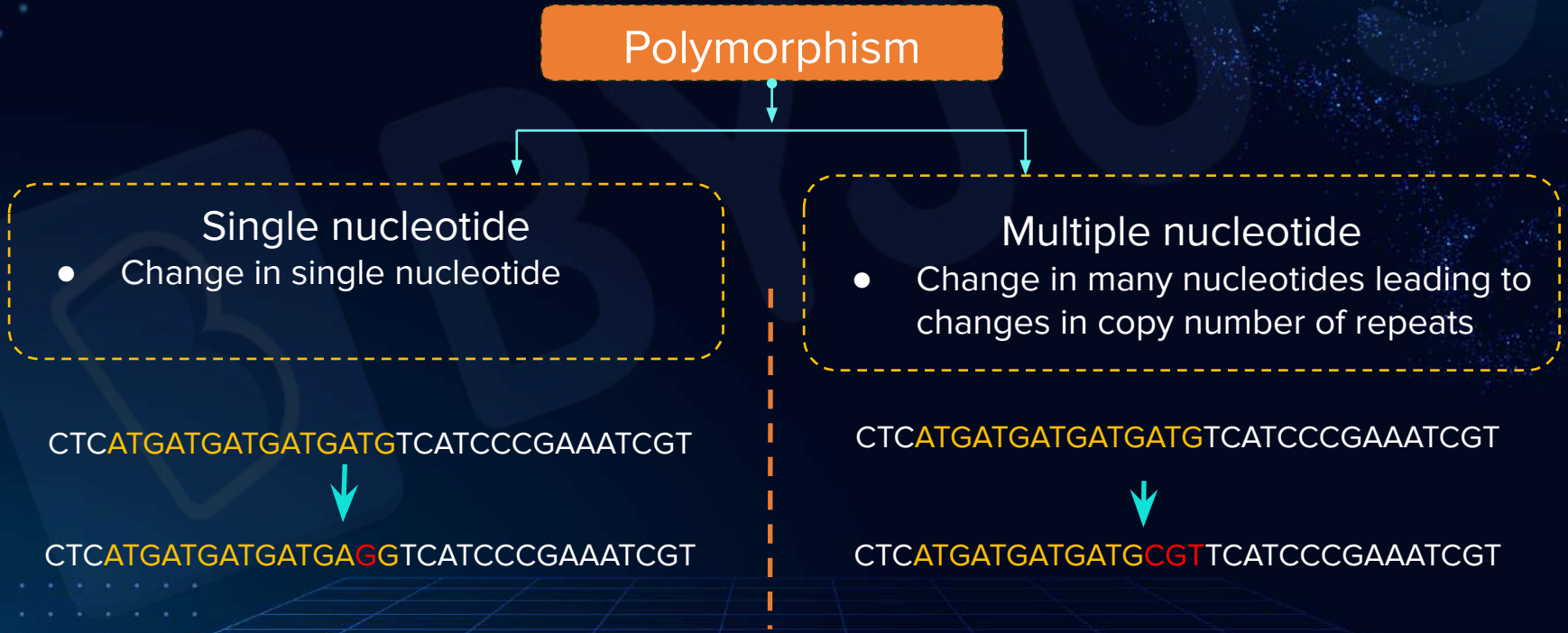
- Satellite DNA does not code for protein

- Mutation in satellite DNA does not lead to genetic disorder.

- Thus, changes caused by mutations in satellite DNA accumulate in germ cell generation after generation.

Polymorphism

- It is the inheritable mutation observed in a population at high frequency (Frequency > 0.01).
- It plays a major role in evolution.



Steps of DNA Fingerprinting

DNA isolation: DNA isolation is performed using biological sample.

1

The biological sample can be any cell/tissue from the person's body

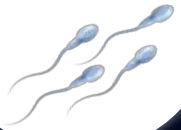
Biological sample

Hair root

Saliva/ cheek cells

Blood

Sperm cells



Steps of DNA Fingerprinting

DNA isolation: Sample is processed and DNA is isolated.

Hair root from crime scene

Crime scene DNA

Hair root from suspect 1

Man 1 DNA

Man 2 DNA

Hair root from suspect 2

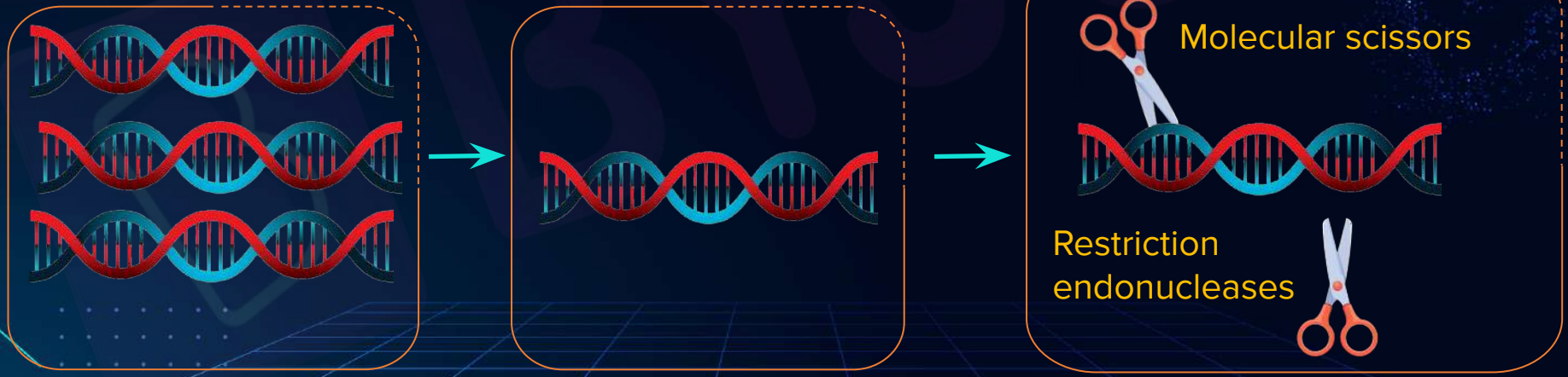
Steps of DNA Fingerprinting

Restriction digestion: Isolated DNA is then digested with restriction enzymes viz. restriction endonuclease, to convert it into small fragments.

DNA isolation

2

The restriction enzymes cut DNA into multiple fragments, hence this process is called restriction digestion.



Steps of DNA Fingerprinting

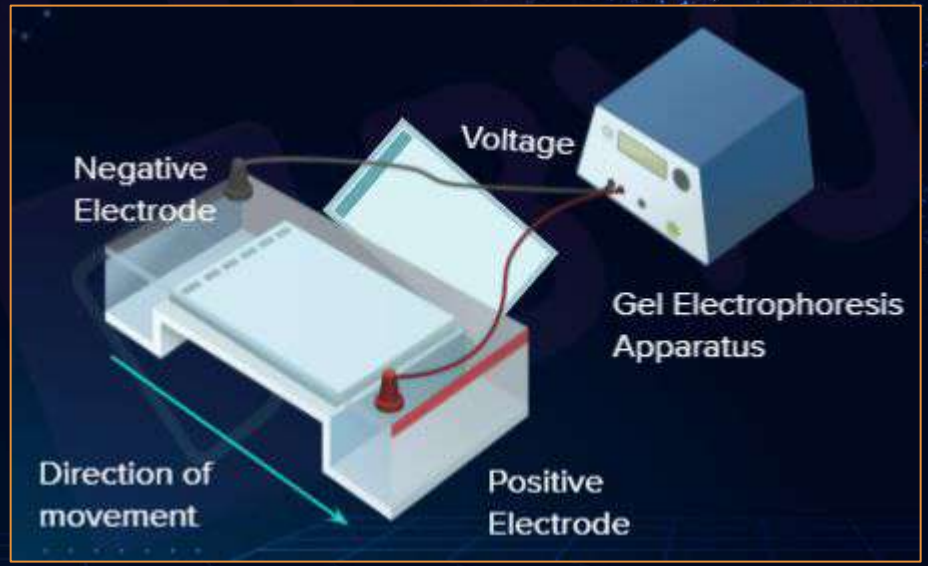
Electrophoresis: It is a process of separating DNA fragments on the basis of their sizes and charges.

DNA isolation

Restriction digestion

3

- Digested samples are loaded into electrophoresis gel.
- The samples move under the influence of electric charge and slowly the fragments separate.

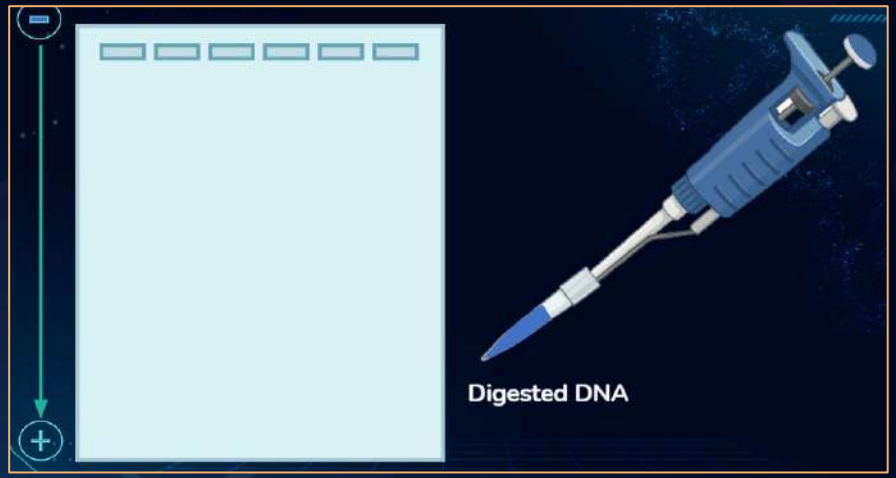


Steps of DNA Fingerprinting

Electrophoresis:

- Restriction digested DNA samples are added to electrophoresis chamber.

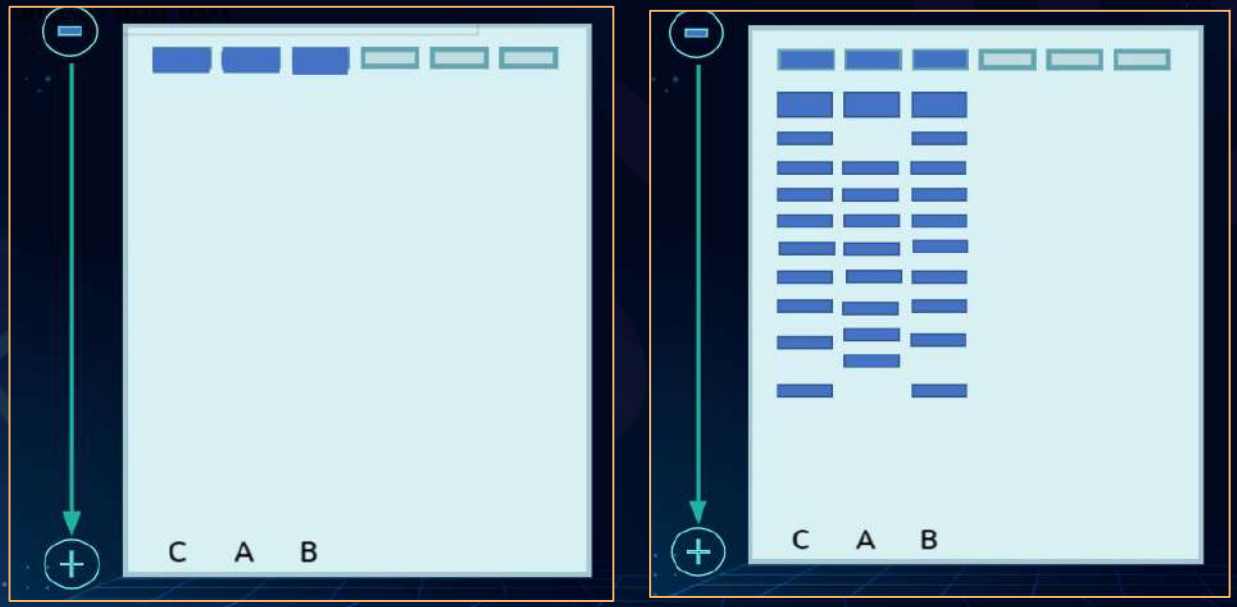
C- Crime scene DNA
A- Suspect 1
B- Suspect 2



Steps of DNA Fingerprinting

Electrophoresis:

- DNA is negatively charged.
- The samples are hence loaded towards negative electrode.
- As the circuit is complete, the samples run towards positive electrode.



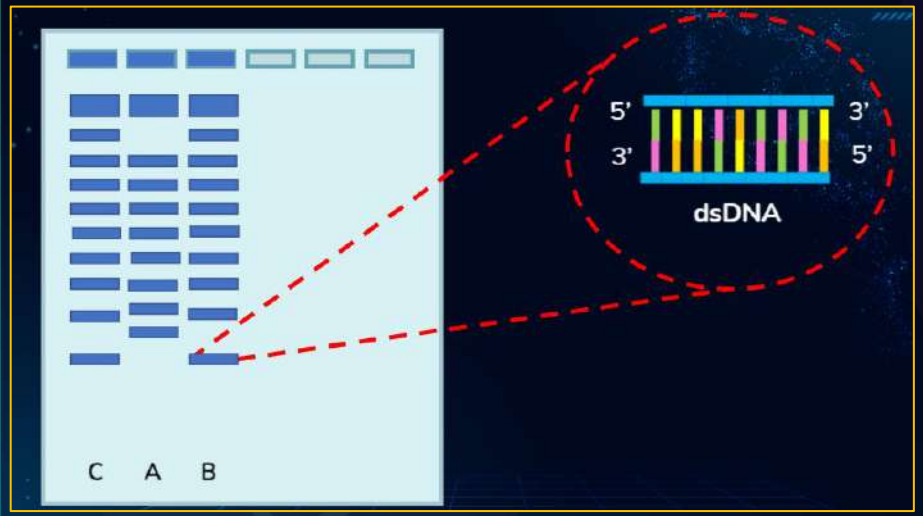
C- Crime scene DNA
A- Suspect 1
B- Suspect 2

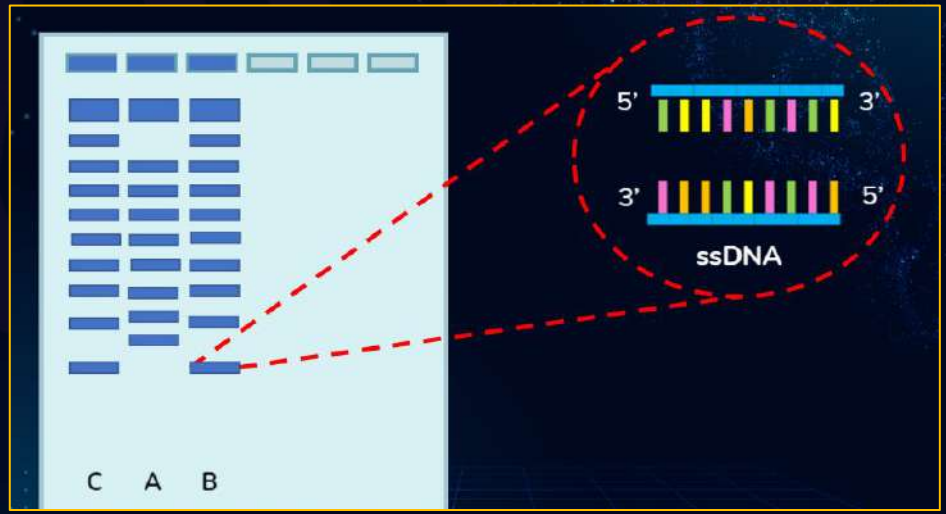
Electrophoresis gel

Steps of DNA Fingerprinting

Post electrophoresis:

- Separated DNA in the gel is double stranded DNA (dsDNA).
- Separated DNA in the gel is treated to make single stranded DNA (ssDNA).





C- Crime scene DNA

A- Suspect 1

B- Suspect 2

Steps of DNA Fingerprinting

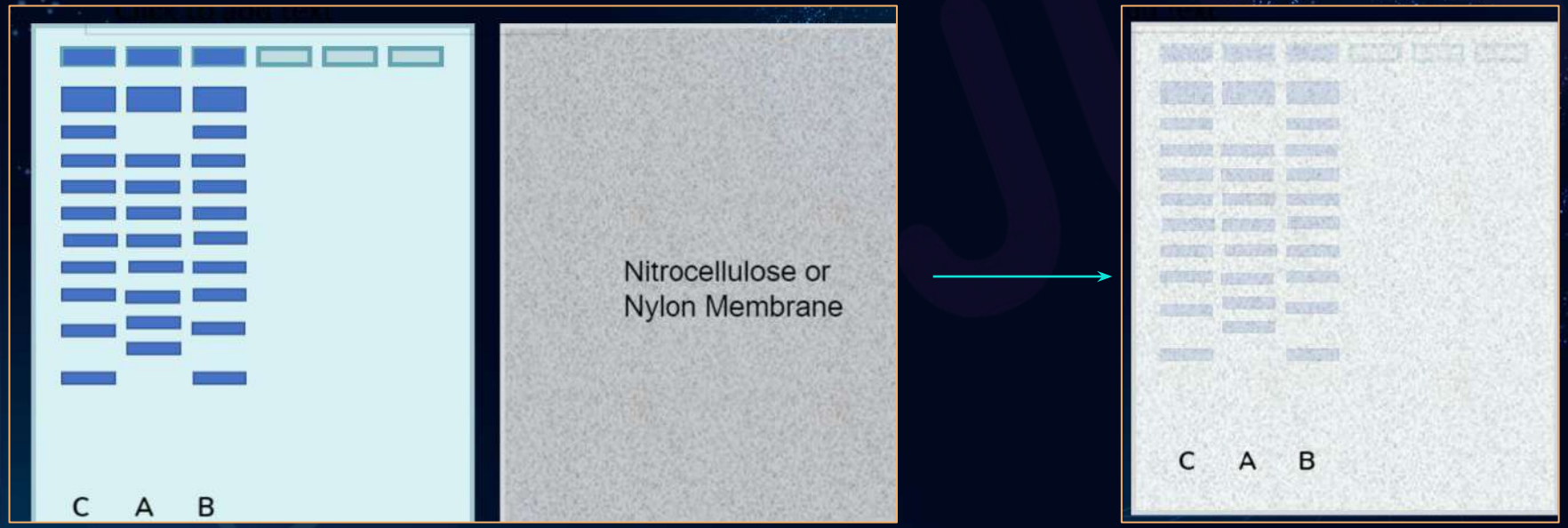
Southern Blot: It is the process of detecting a specific DNA sequence in DNA samples.

- After electrophoresis, the separated DNA samples are transferred to a nylon membrane. This process is called southern blotting.

Steps of DNA Fingerprinting

Southern blot

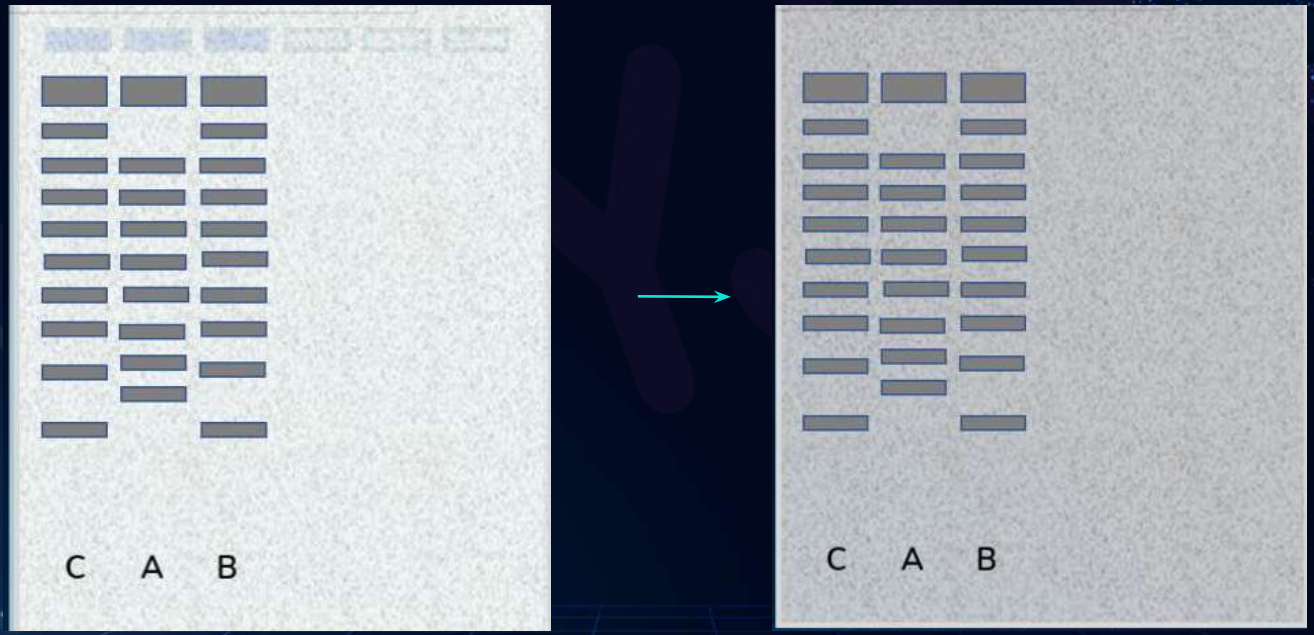
- After treatment is done, the nylon membrane is placed on the gel.



Steps of DNA Fingerprinting

Southern blot

- After bands are transferred to the nylon membrane, the gel is removed .



Steps of DNA Fingerprinting

Hybridization: It is the process where sequence specific binding of two complementary DNA sequences occurs.

DNA isolation

Restriction digestion

Electrophoresis

Southern blot

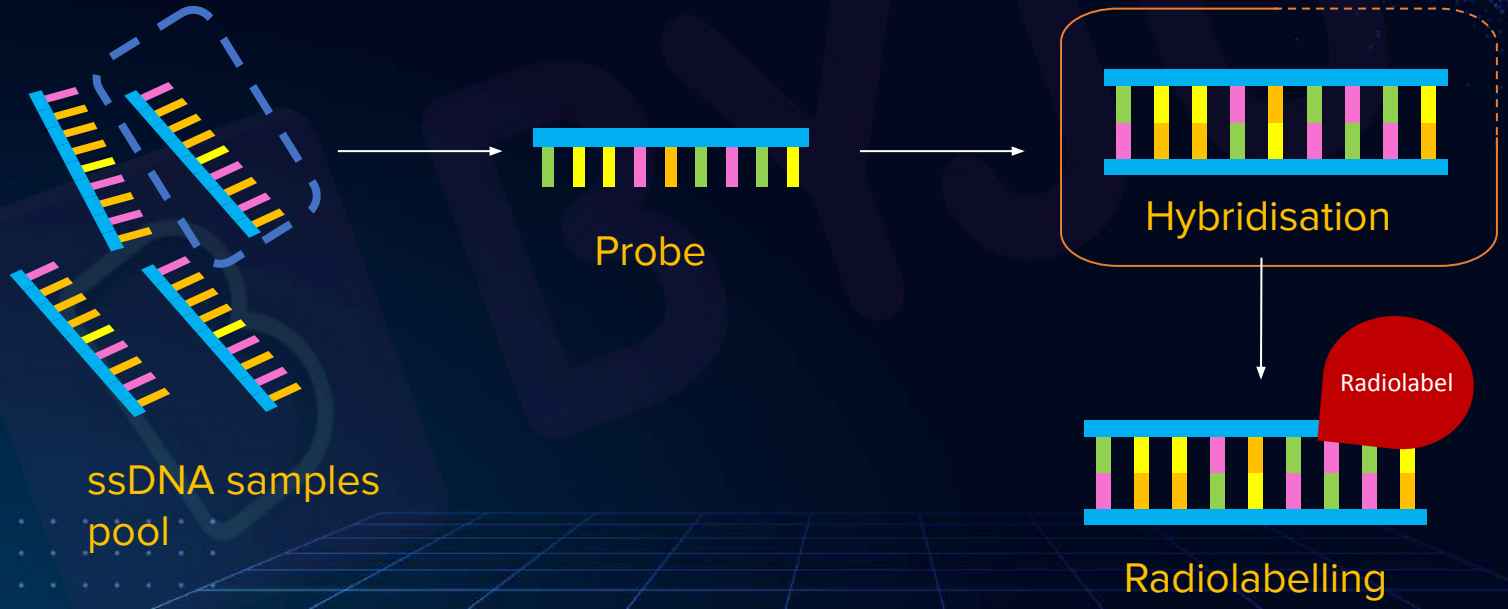
5

- **Probe:** It is an artificially designed sequence that is complementary to the DNA sequence to be selected.
- Probe can only bind to our desired sequence.

Steps of DNA Fingerprinting

Hybridisation

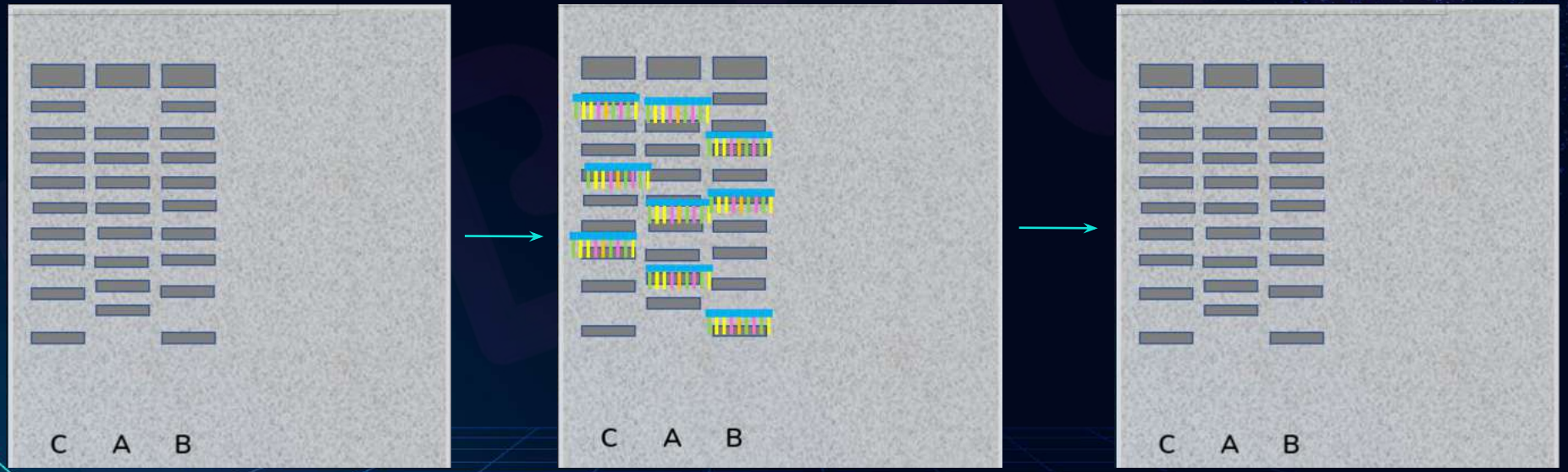
- In a DNA sample, there are a lot of DNA fragments.
- Probe is labelled with radioactive substance.
- As probe binds to the complementary DNA, they send out signal (radiolabels).



Steps of DNA Fingerprinting

Hybridisation

- The bands transferred to nylon membrane have pool of ssDNA fragments.
- VNTR specific probes hybridise to the complementary VNTR sequence present on the membrane.
- Unattached probes are washed away.



Steps of DNA Fingerprinting

Autoradiography: Hybridised DNA with probe is subjected to X-Ray, label emits signal.

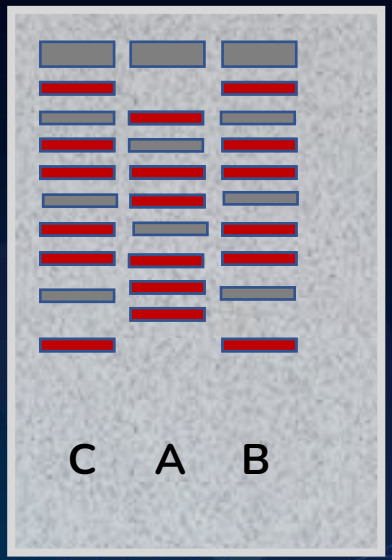
DNA isolation | Restriction digestion | Electrophoresis | Southern blot | Hybridisation | 5

- This is performed to confirm if hybridisation had occurred or not.
- The probes not attached, are washed away.
- Hence, signal will be passed for the sequences which have attached to the radiolabeled probes.
- The banding pattern obtained after exposure to x-ray is analyzed.

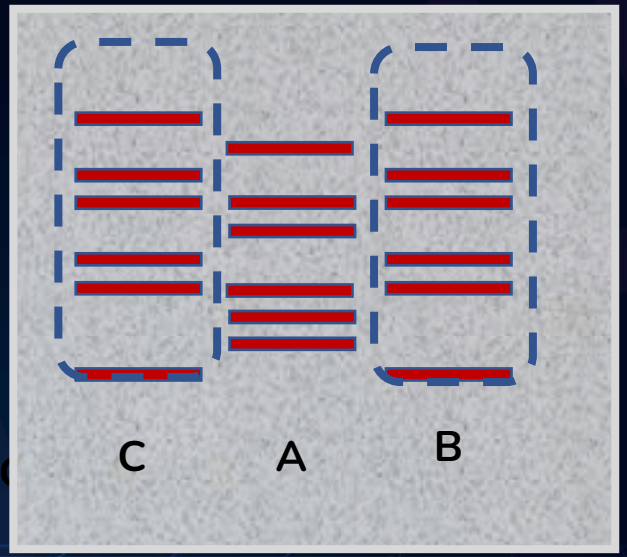
Steps of DNA Fingerprinting

Autoradiography

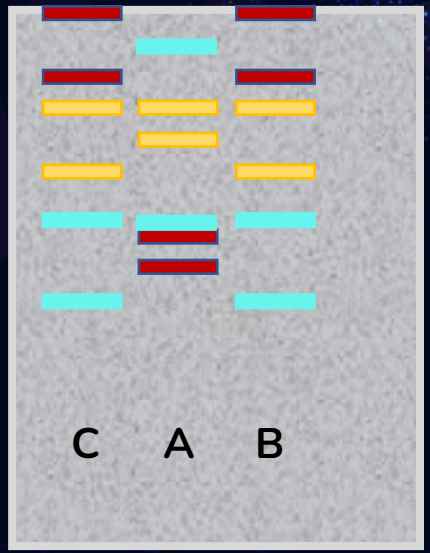
- Given a sample DNA C, it should match with one of the two person's DNA A or B.
- On performing the above steps, it was found that sample DNA is matching with the DNA of person B.



Autoradiograph



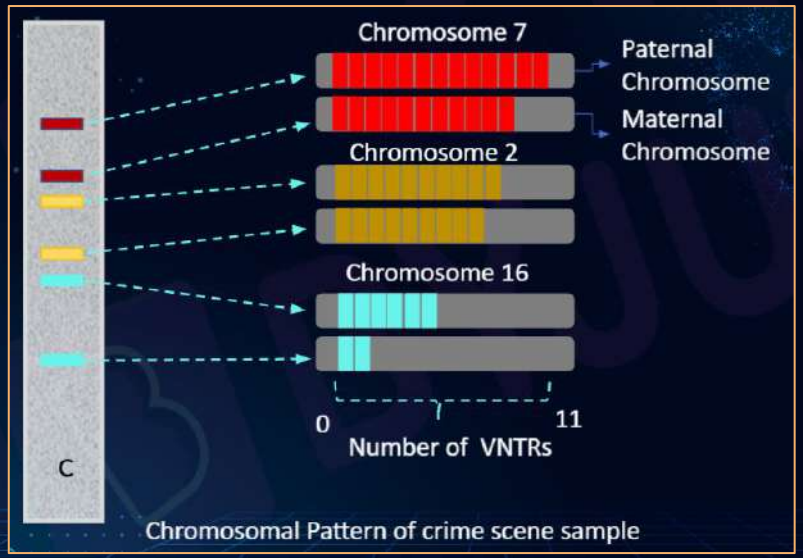
Matching the banding patterns



C matching with B

Chromosomal Representation

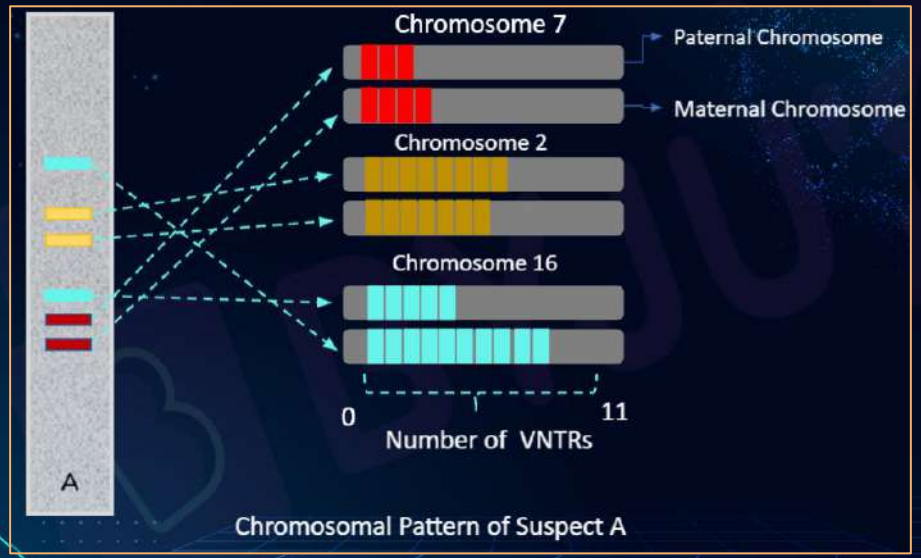
- DNA bands in the gel represent the number of short tandem repeats.
- For eg. Bands of chromosome 7 moved to very less distance in electrophoresis, hence, contain more repeats.
- In case of chromosome 16, it has travelled more distance. Hence, it can be concluded that it has less number of repeats.

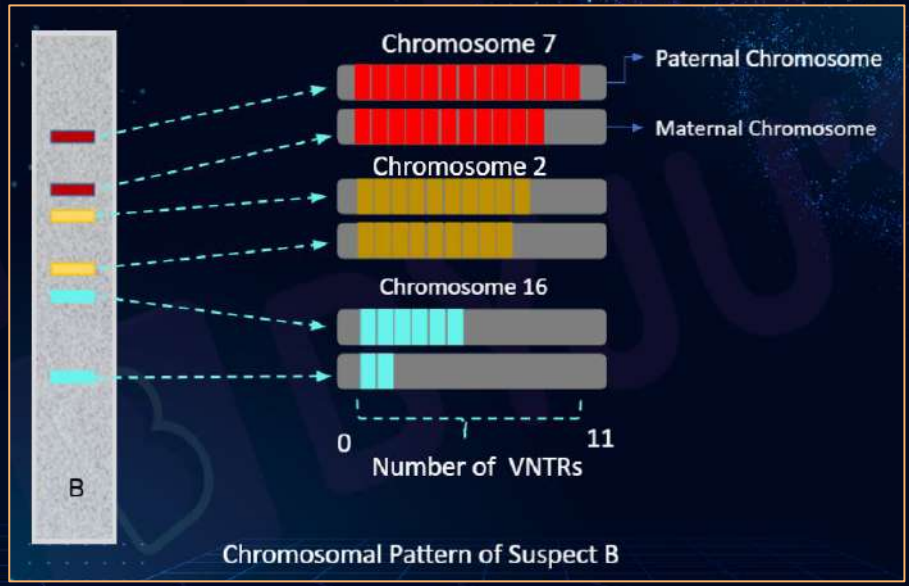


Chromosomal Pattern of crime scene sample

Chromosomal Representation

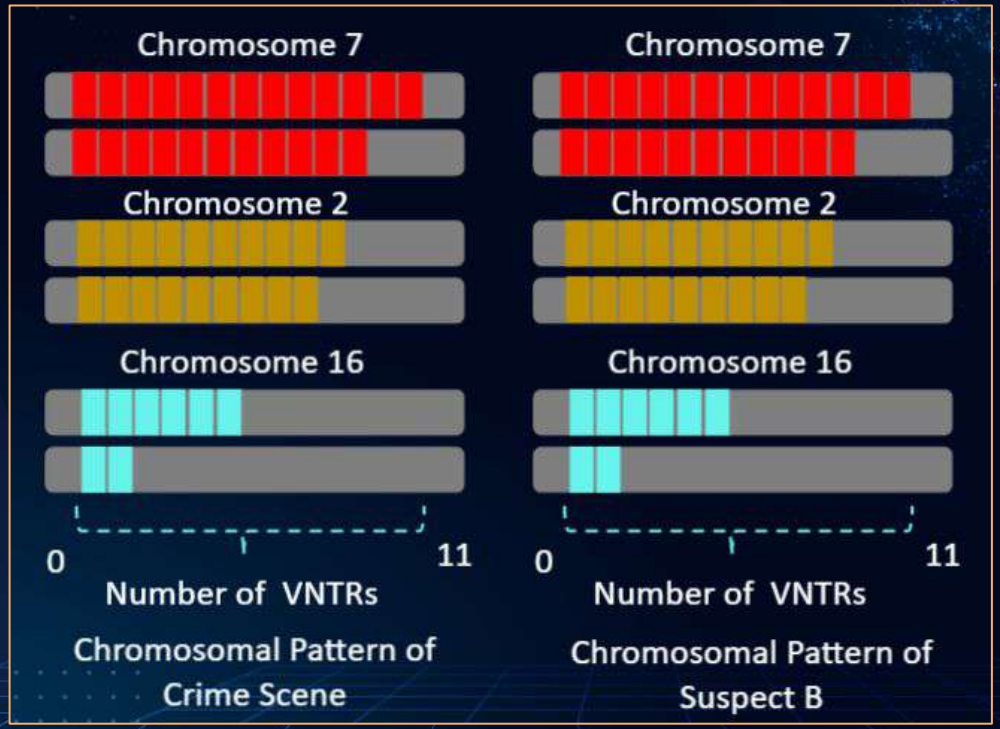
- In the gel images shown, the color differentiation is shown for our understanding.
- The pattern of the tandem repeats of the sample obtained from suspect A and the suspect B is different, indicating unique vntr patterns of different individuals.



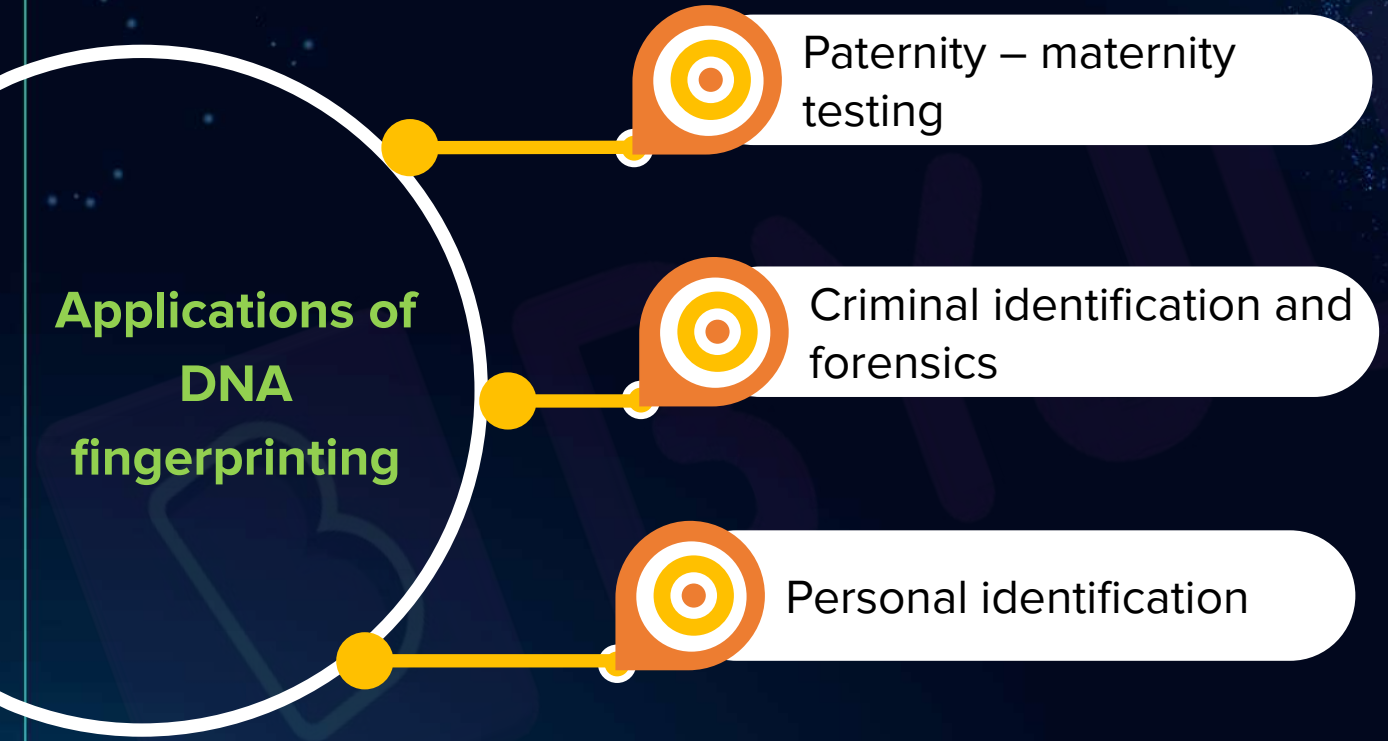


Chromosomal Representation

- The pattern of the tandem repeats of the sample obtained from crime scene and that of suspect B is same.
- Hence, it can be concluded that suspect B is the criminal.



Applications of DNA Fingerprinting



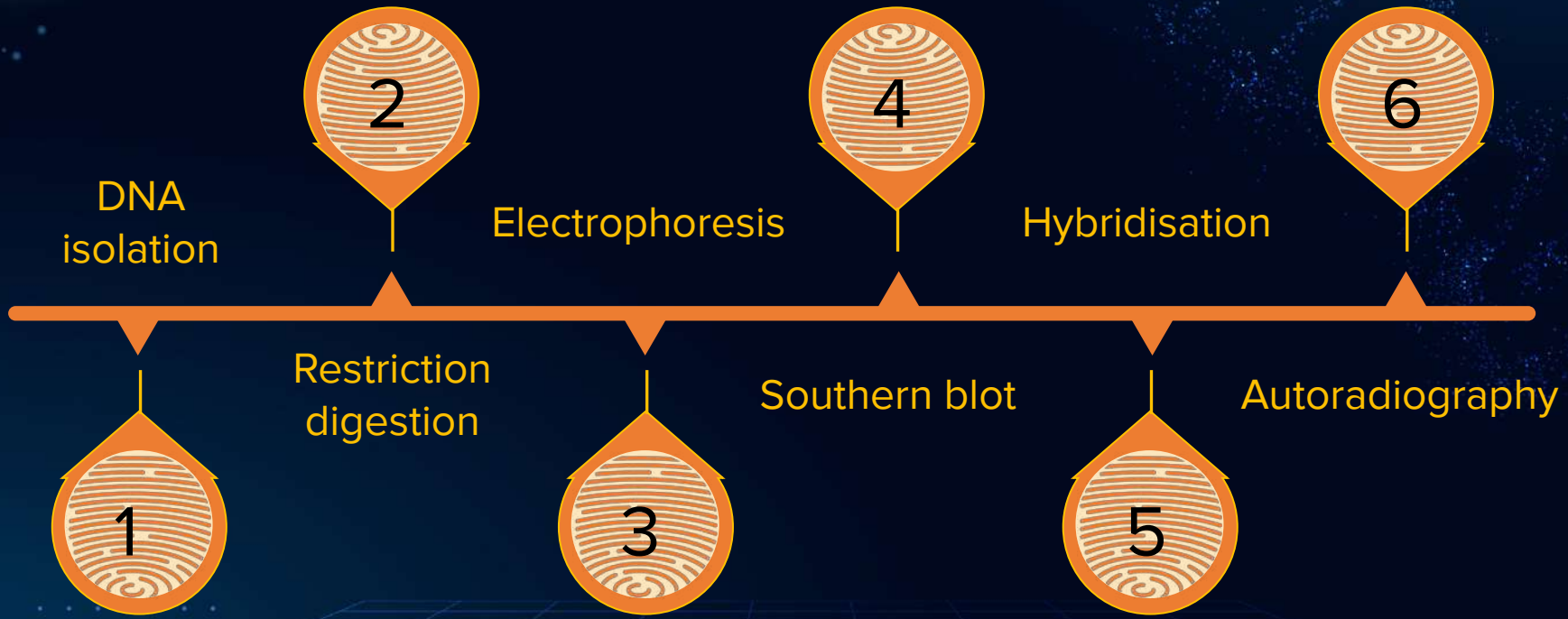
Summary

Variable number tandem repeats (VNTR)

- Show high polymorphism
- Many bands are seen (Size 0.1-20 kb)

Summary

Steps of DNA fingerprinting



BYJU'S Classes Notes

Molecular Basis of Inheritance

Human Genome Project : Goals, Methods, Features and Applications

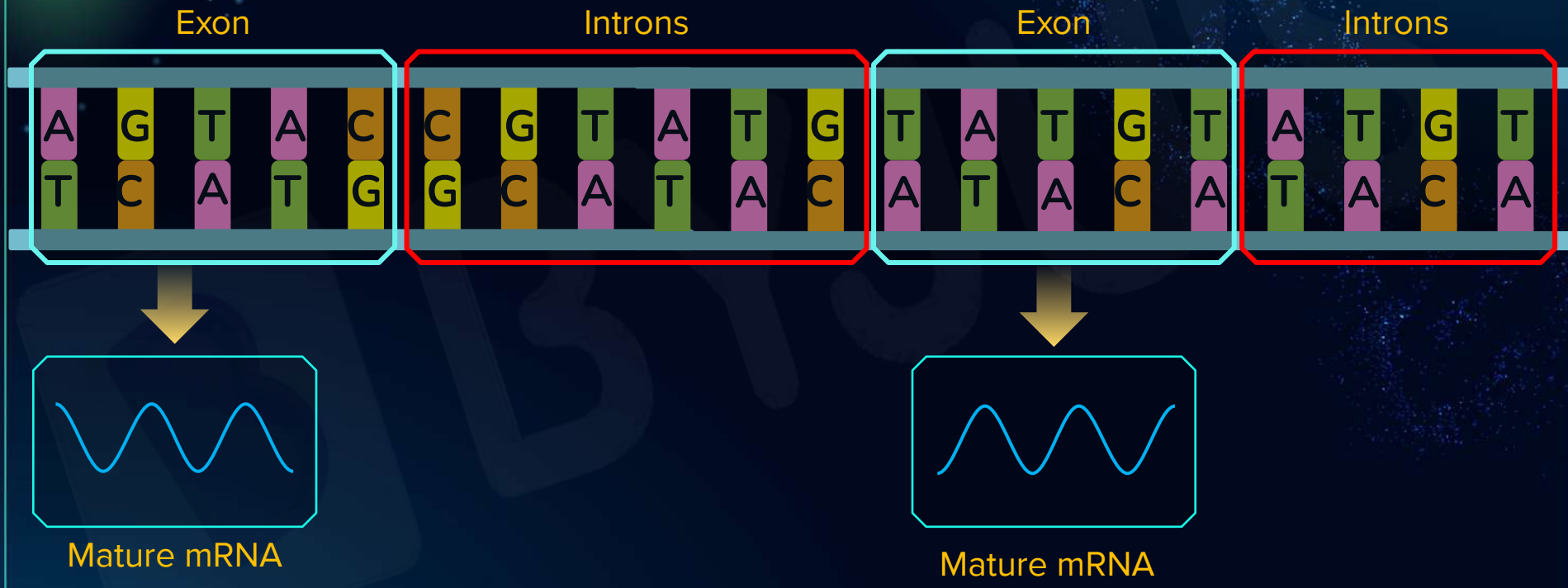
Key Takeaways

Human genome project

- Goals
- Methodology
- Features
- Applications

Summary

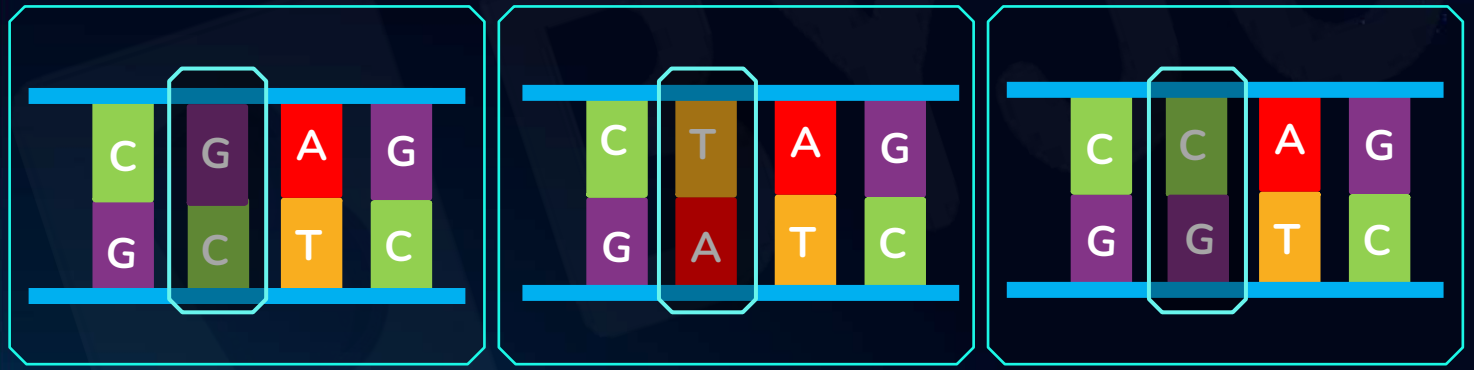
Recall! Exons and Introns



- Exons code for proteins, whereas introns are non-coding regions.

Recall! SNPs

The change in DNA sequence at single base pair is called single nucleotide polymorphism (SNPs).



Sample 1 Sample 2 Sample 3

Similarity Vs Difference

Assumption : To have morphological changes between two individuals, their DNA sequences must differ at least at certain places.
This basic assumption led to the quest of finding the complete human genome sequence.

Human Genome Project

- A thirteen years long project (1990-2003)
- Aim was to sequence the complete human genome
- Also known as ‘mega project’

Coordinated by:

- U.S Department of Energy
- National Institute of Health

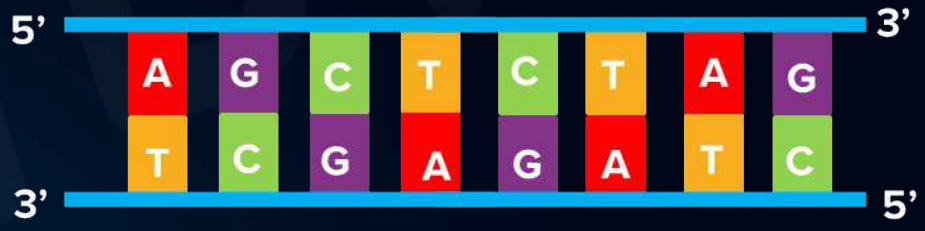
Partners:

- Wellcome Trust (U.K.)
- Japan
- France
- Germany
- China

DNA Sequencing

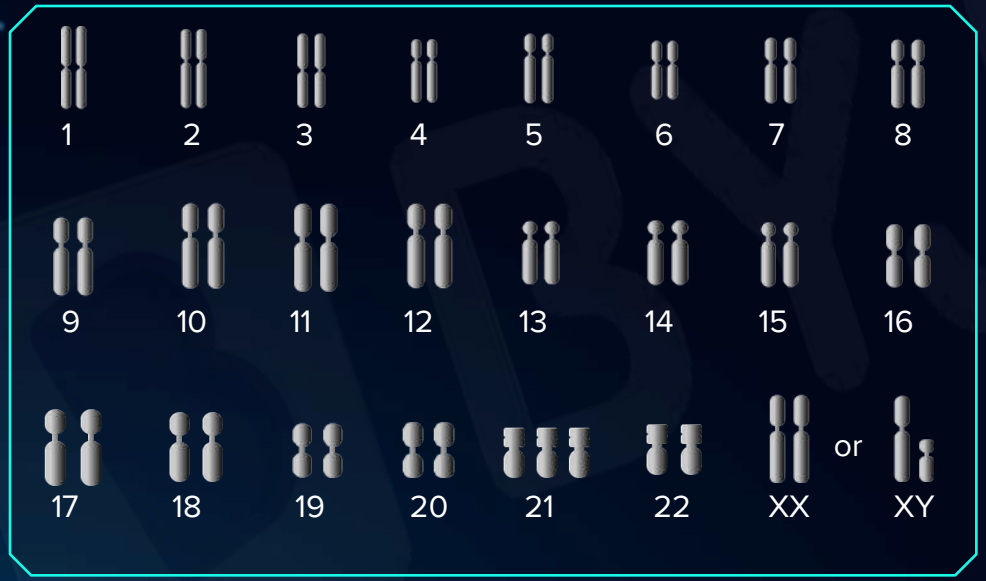
DNA sequencing – It is a process of identifying the exact sequence of nitrogenous bases in the DNA.

For example, DNA sequence of upper strand is AGCTCTAG and of lower strand is TCGAGATC

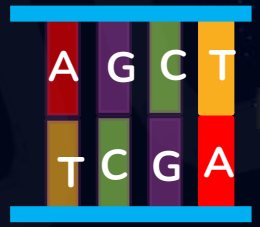


Human Genome Project

The number of base pairs of the entire genome is approx 3×10^9



Genome



Approx. $ 3 US dollar is the cost for the sequencing of 1 base pair.

$= \$3 \text{ USD} \times 3 \times 10^9$ base pairs

= $ 9 billion US dollar (approx.)

$ 9 billion US dollar = 900 crores INR

Human Genome Project

If the obtained sequences were in the form of books, 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.

1000 letters in a page

Each book has 1000 pages

3300 Books

The obtained sequences made storage, retrieval and analysis of data very difficult. Hence, there was an urgent need to find out a user friendly way to store, retrieve and analyse the data.

Human Genome Project

Bioinformatics : Hybrid field that deals with biological data and uses computer science to store, retrieve and analyse them

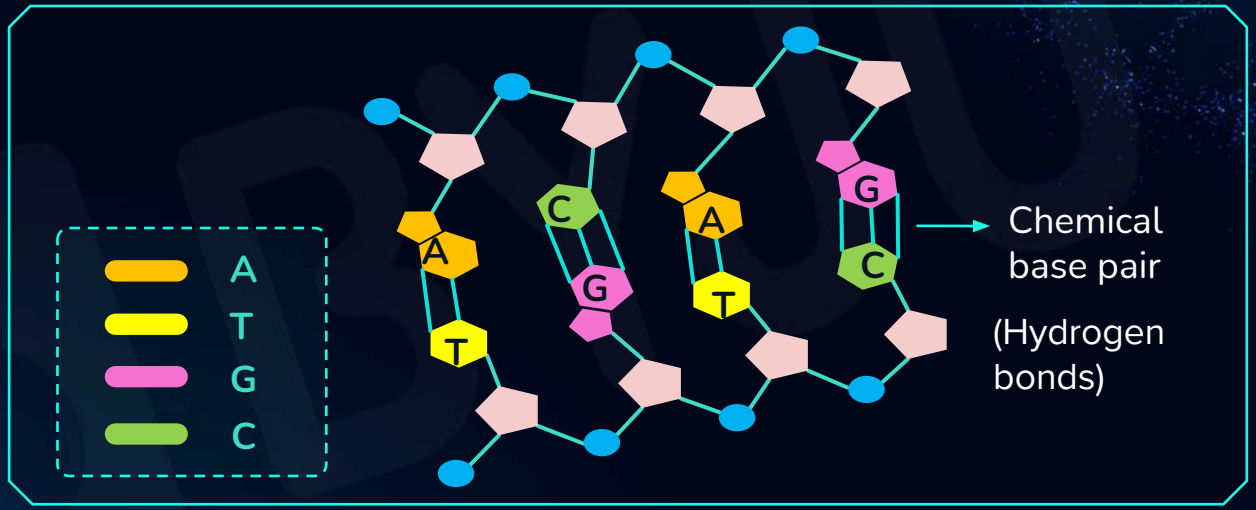
- The storage, retrieval and analysis of the data became very easy with the help of computers.
- To analyze the biological data such as DNA sequences, usage of computer and mathematics was increased.

Human Genome Project : Goals

Identification of approx. 20,000-25,000 genes in human DNA

Human Genome Project : Goals

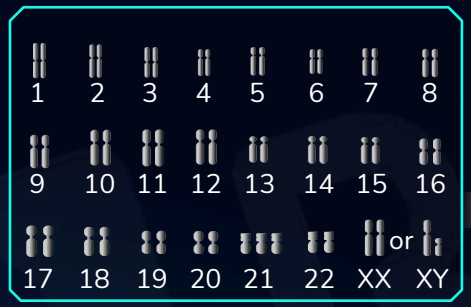
Determination of 3 billion chemical base pairs of human DNA



- Within the genes, bases are combined with the help of hydrogen bonding.
- The combination of purines (A & G) and pyrimidine (T & C) by hydrogen bonding is called chemical base pairing.

Human Genome Project : Goals

Store the information in databases



Genome

DNA sequence information

Computer

- Whenever a large data in computer has to be stored, it requires a database where data is stored.
- So, to store the DNA sequences logically in the computer, a database to store it, is required.
- Example – GSDB- (Genome sequencing database) is controlled by NCGR (National Center for Genome Resources).

Human Genome Project : Goals

Improvement of tools for data analysis

Data analysis

- Improvement of tools indicates using fast computational techniques to retrieve and analyze the genetic sequence data quickly.
- Now, computational techniques include those software which help scientists to manage the data efficiently (Eg: BLAST and FASTA).
- Moreover, using tools and techniques to make data accessible to everyone is always beneficial to mankind.

Human Genome Project : Goals

Transfer related technologies to other industries

Agriculture

Hardier, more nutritious, and healthier crops and animals by sequencing their genomes and modifying them for improvements.

Medical

Identification of humans, other animals, plants, and microbes; evolutionary and human anthropological studies made easy

Pharmaceutical

Personalized drug (based upon individual's genetic sequence) and other therapies

Human Genome Project : Goals

Address ethical, legal and social issues that may arise from HGP

Issues

Ethical

Killing female foetus if genetic sequence is known before her birth (female foeticide)

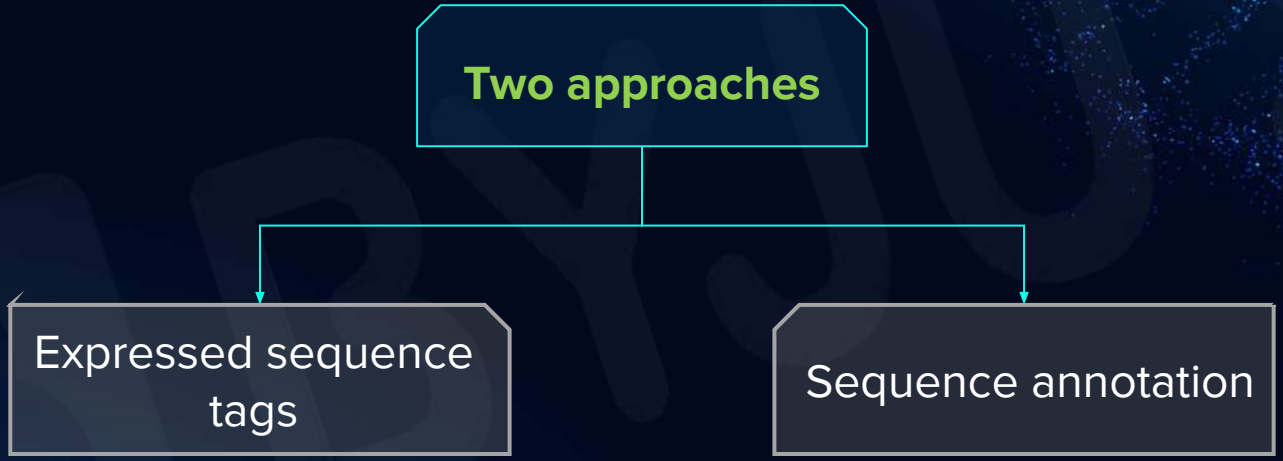
Social

Discrimination on the bases of race, by the society

Legal

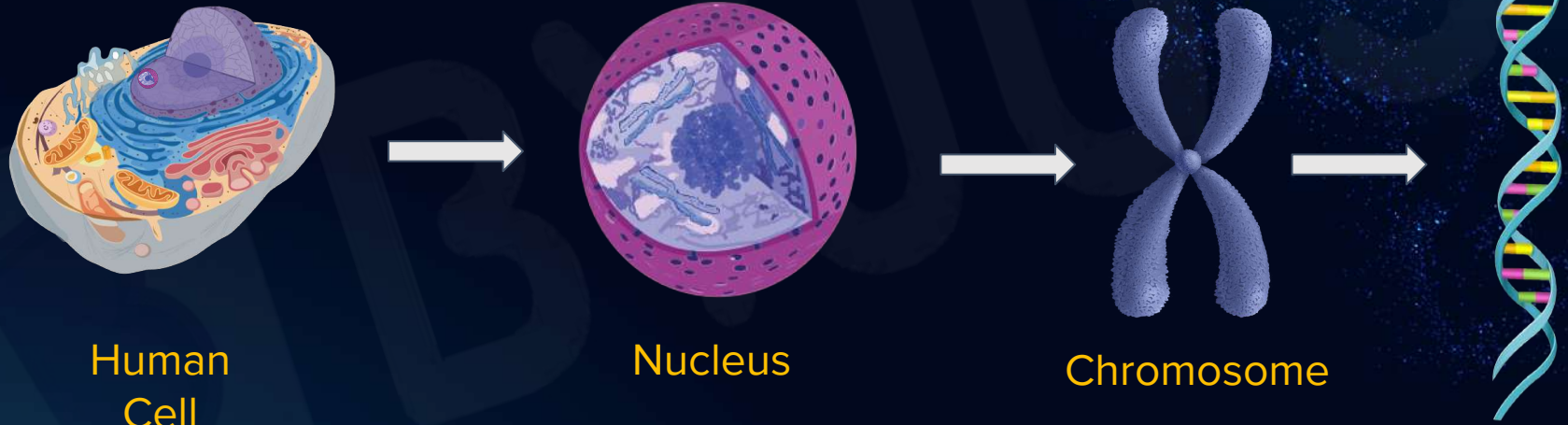
Denying insurance if the individual has an altered DNA sequence because of higher risk of death

Human Genome Project : Methodology



Human Genome Project : Methodology

Expressed sequence tags



Human Cell

Nucleus

Chromosome

DNA

- DNA isolation is the prime and first step, wherein DNA is isolated from the nucleus of human cell.

Human Genome Project : Methodology

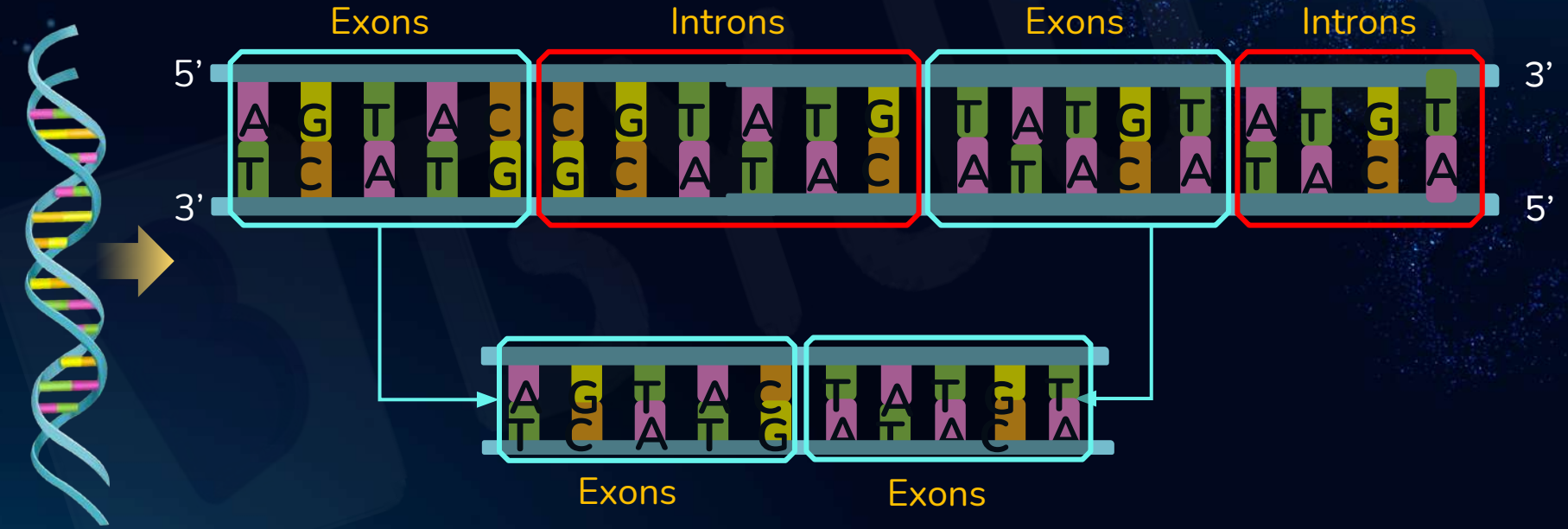
Expressed sequence tags

- In this approach, all the coding genes are selected and sequenced.
- Coding genes means genes which can be expressed as mRNA.
- Since, introns which are present between two exons are removed during mRNA synthesis they won’t be sequenced.
- Therefore, no introns are sequenced in this approach.

Genomic DNA

Human Genome Project : Methodology

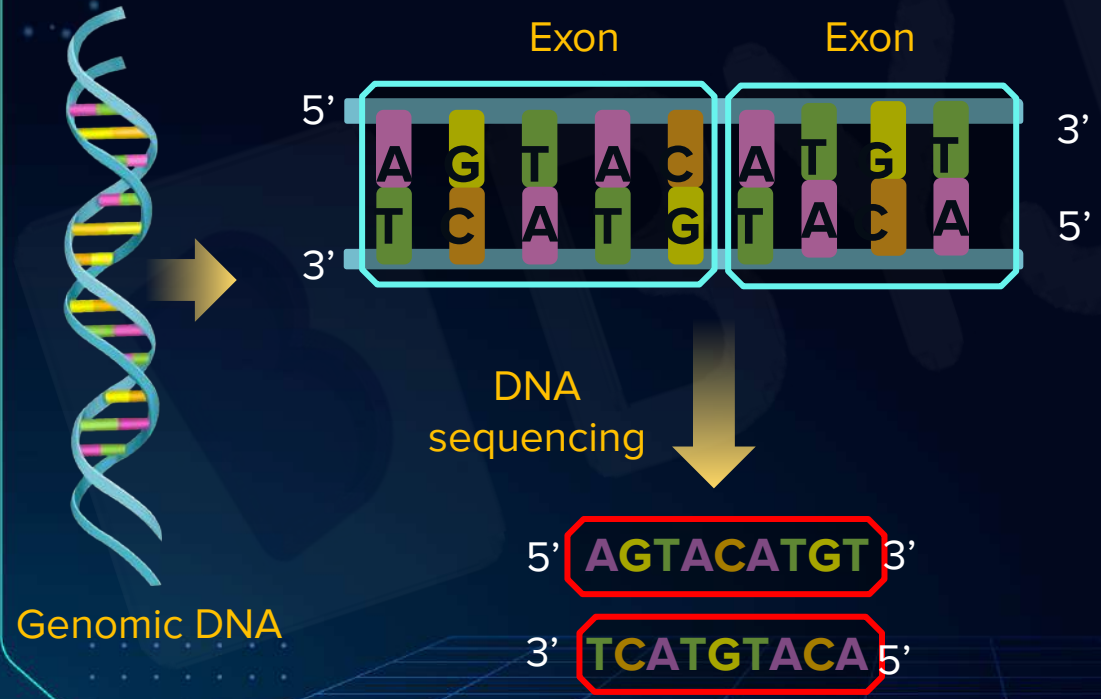
Expressed sequence tags



- Hence, exons are sequenced and analysed as introns, which are non-coding, are removed.

Human Genome Project : Methodology

Expressed sequence tags



- Exons are sequenced and analysed.
- Therefore, only expressed genes are sequenced.
- This process does not give the complete genome but at least the sequence of the expressed regions can be determined.

Human Genome Project : Methodology

Sequence annotation

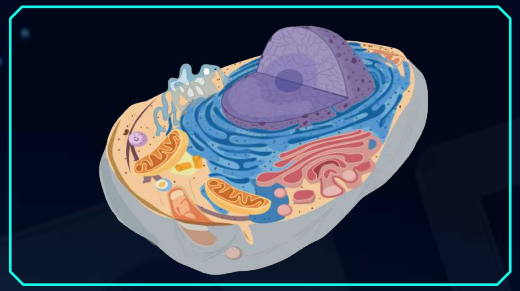
- To sequence DNA
- To mark

Steps in sequence annotation

1. Isolation of DNA
2. Amplification of DNA or creation of copies
3. Sequencing of amplified fragmented DNA
4. Annotation and assigning of DNA

Sequence Annotation

Isolation of DNA



Human Cell

Genomic DNA

- Isolation of the human genomic DNA from the nucleus of the cell is the first step.
- The genomic DNA isolated from the cell is cut into fragments using enzymes.
- To make multiple copies of the cut strand, either a bacteria or yeast is utilised.
- The fragment is inserted initially into them.

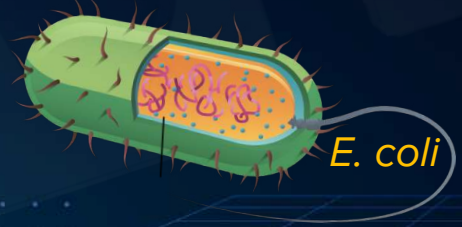
Sequence Annotation

Vectors (Vehicle)

Vectors are DNA molecules that are used as a vehicle to carry foreign DNA fragments into other cells.

BAC

- Bacterial Artificial Chromosome Vector
- Used to transfer DNA to bacteria



E. coli

YAC

- Yeast Artificial Chromosome Vector
- Used to transfer DNA to yeast

Yeast

Sequence Annotation

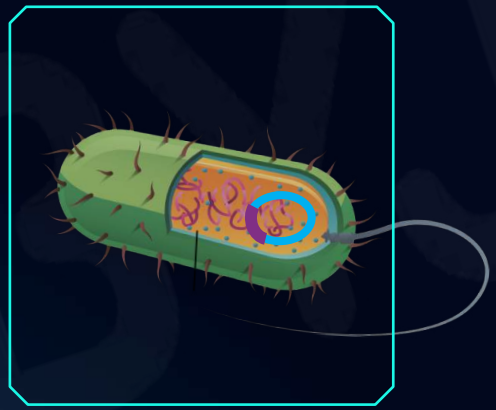
Amplification

- In order to increase the number of DNA fragments, the vector (eg BAC) is used and cut open using enzymes.
- Vector is then cut and the DNA fragment is inserted and joined with the help of DNA ligase.
- Vector + DNA fragment = Recombinant DNA molecules

Sequence Annotation

Amplification

Recombinant DNA (rDNA)

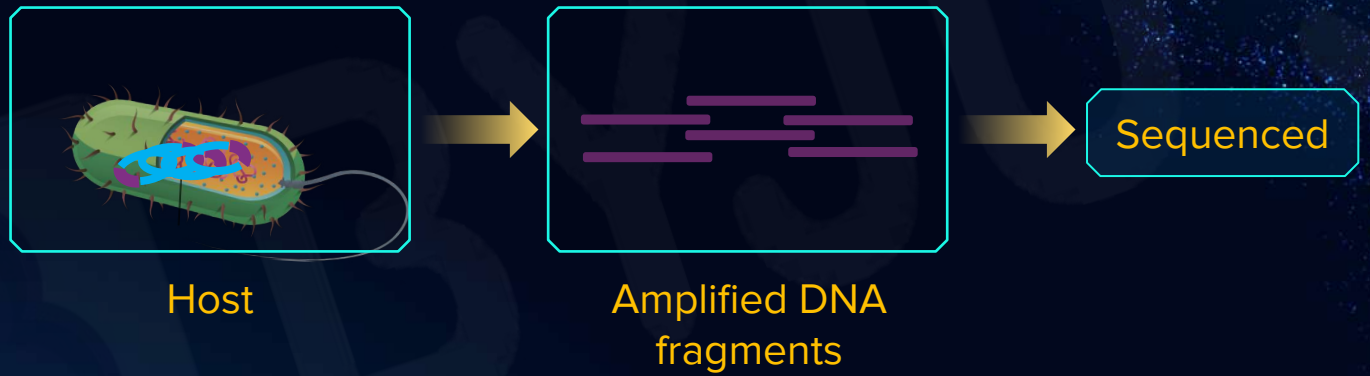


Bacteria

- The rDNA is then inserted into the bacterial cell/yeast cell (depending upon what vector was used), so that the fragmented DNA can amplify (increase in number, making more copies).
- The amplification of the fragmented DNA would ease the DNA sequencing process.

Sequence Annotation

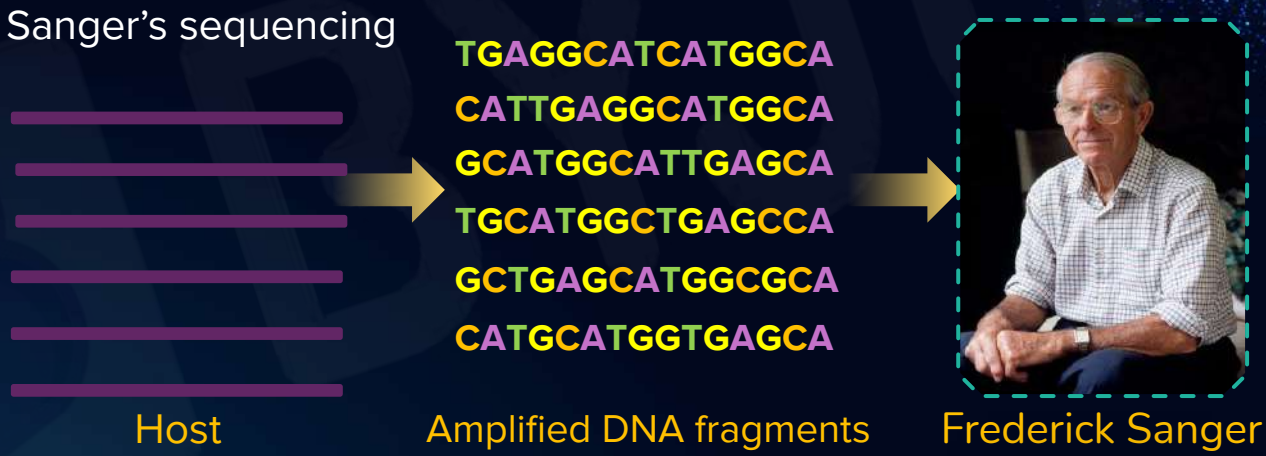
Sequencing of amplified fragmented DNA



- Vector then replicates within the bacteria and makes several copies of itself.
- Along with it, DNA fragment will also replicate and make several copies of the same.
- The amplified fragmented DNA is then isolated and sequenced.

Sequence Annotation

Sequencing of amplified fragmented DNA



- Frederick sanger is known for devising a method for the sequencing of DNA nucleotides.

Sequence Annotation

Sequencing of amplified fragmented DNA

TGAGGCATCATGGCA

CATTGAGGCATGGCA

GCATGGCATTGAGCA

TGCATGGCTGAGCCA

GCTGAGCATGGCGCA

CATGCATGGTGAGCA

- The sequenced DNA are then arranged using specialized computer based programs.

Sequence Annotation

Sequencing of amplified fragmented DNA

Overlapping DNA sequence

Complete Genome Sequence

- The overlapping sequences helps in finding the common regions in different DNA fragments and thus a common genomic sequence.

Sequence Annotation

Annotating and assigning

- Once the genome is sequenced, the different regions of the genome are annotated or marked, for example, which gene is coding for proteins and which are non-coding proteins etc.

Sequence Annotation

Annotating and assigning

TGCCATGATGCCA TGACCATGATGCCA TGAGGCATCATGGCA TGACCATGATGCCATGA TGACCATGATG

Annotated DNA sequence

Assigned to chromosomes

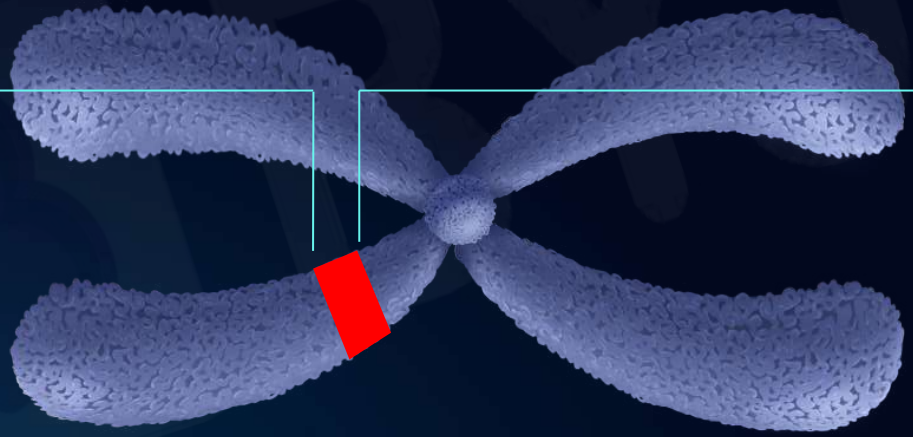
- The complete DNA sequence is then assigned to respective chromosomes.

Sequence Annotation

Annotating and assigning

Gene mapping

TGCCATGATGCCA TGACCATGATGCCA TGAGGCATCATGGCA TGACCATGATGCCATGATGACCATGATG

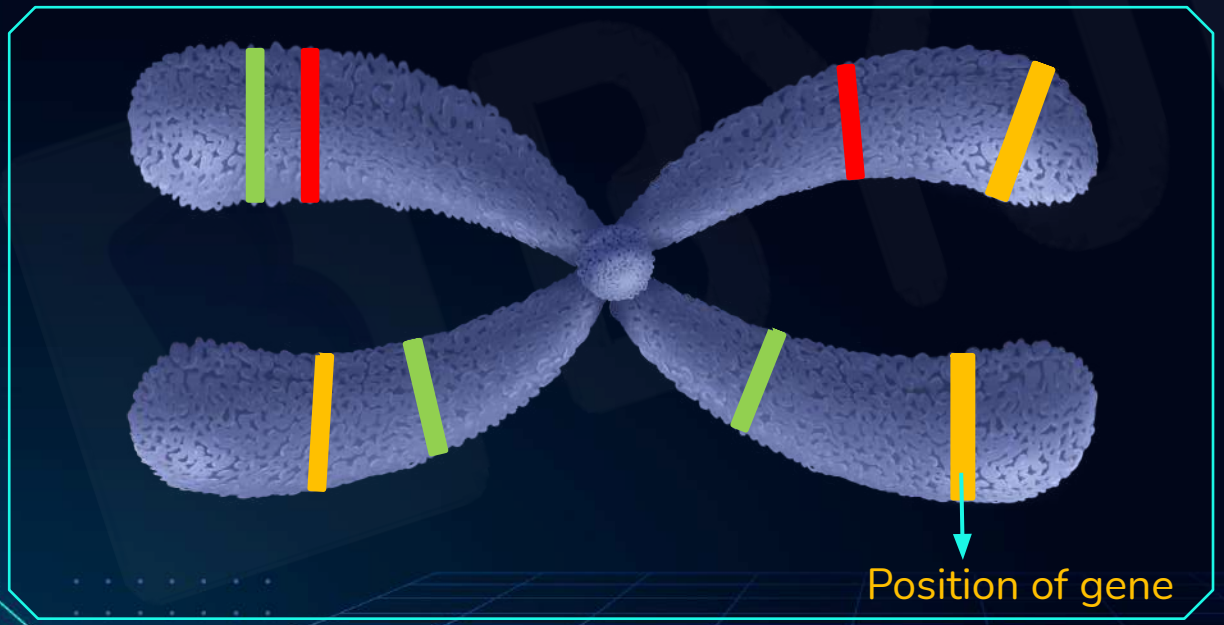


- The process of assigning the position of genes on to the chromosomes is called genetic mapping.

Sequence Annotation

Annotating and assigning

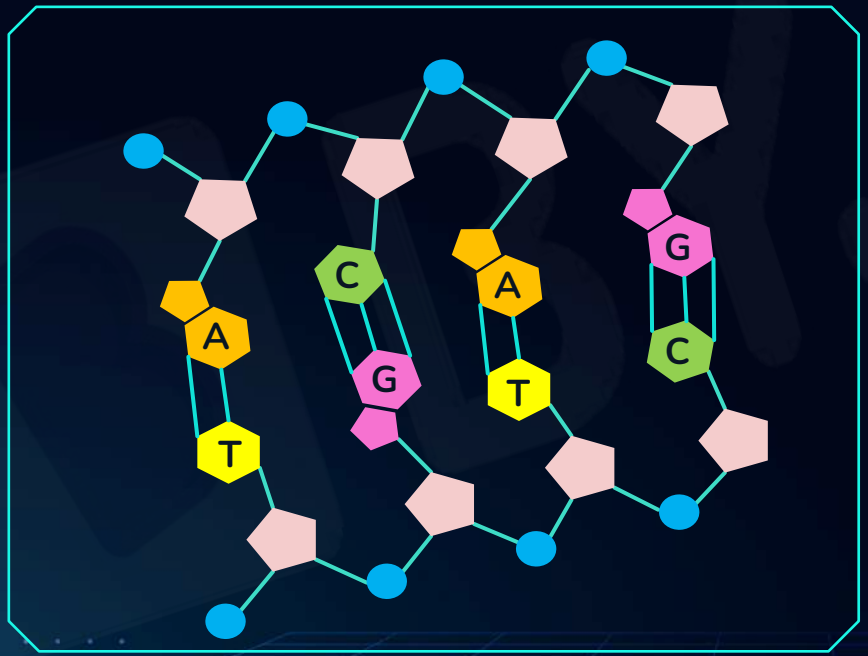
Gene mapping



- Position of different genes can be found out on the chromosome.

Human Genome Project : Features

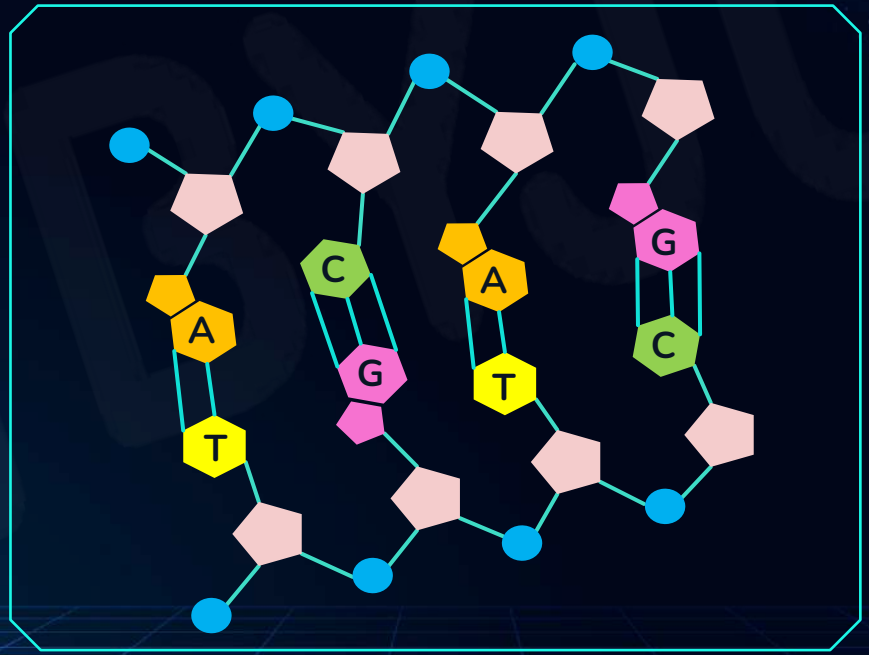
Human genome contains approx. 3164.7 million bp



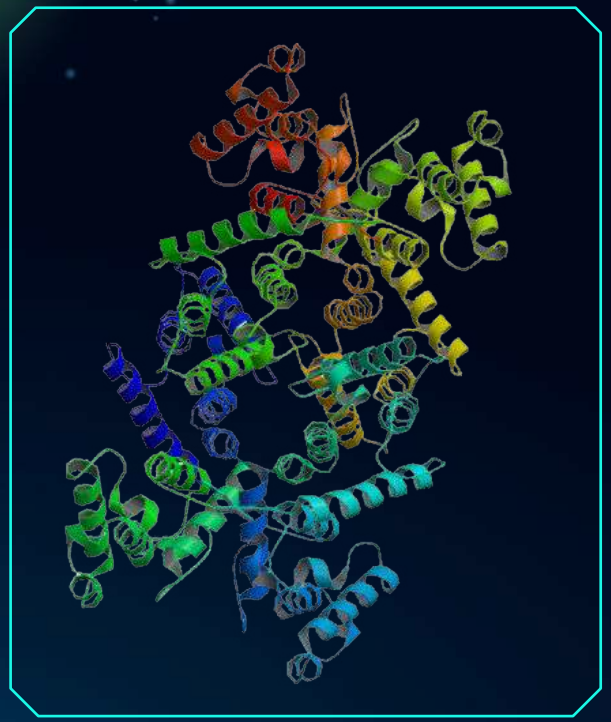
- Nucleotides have adenine, thymine, cytosine and guanine as the 4 different nitrogen bases in DNA.

Human Genome Project : Features

Average size of human gene is 3000 bases, it means around 3000 nucleotide bases make up one gene.



Did You Know?



Dystrophin

Largest human gene

Dystrophin
Largest known human gene with 2.4 million bases, this gene codes for a protein that play role in the strengthening of the muscles fibers and protect them from injury.

Human Genome Project : Features

Total number of genes is 30,000

Estimation – 80,000 to 1,40,000

HGP

Actual number – 30,000

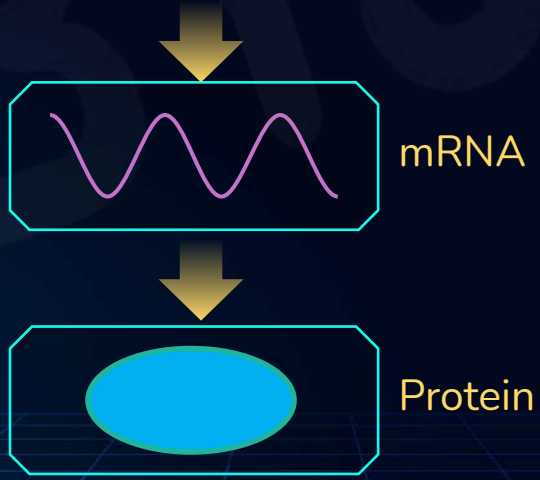
- Initially, it was estimated that the total number of genes could be 80-140 thousand but after HGP was done, it was found that the total number of genes is 30 thousand only.

Human Genome Project : Features

The function of over 50% of discovered genes is unknown

Human Genome Project : Features

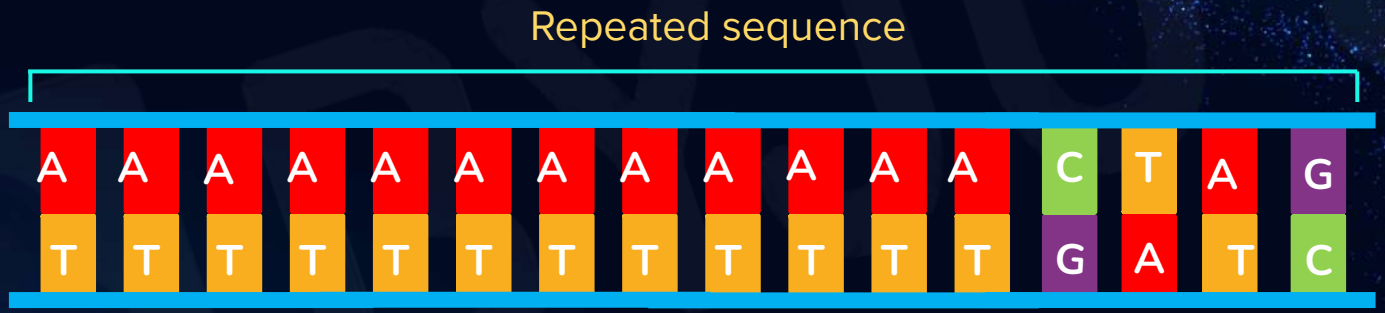
Less than 2% of the genome codes for protein, rest of the genes have unknown function.



- Total number of genes coding for protein in human genome is 30,000 which is less than 2% of the total genome and almost 50% of the discovered genes have known functions.

Human Genome Project : Features

Repeated sequences make up a large portion of human genome

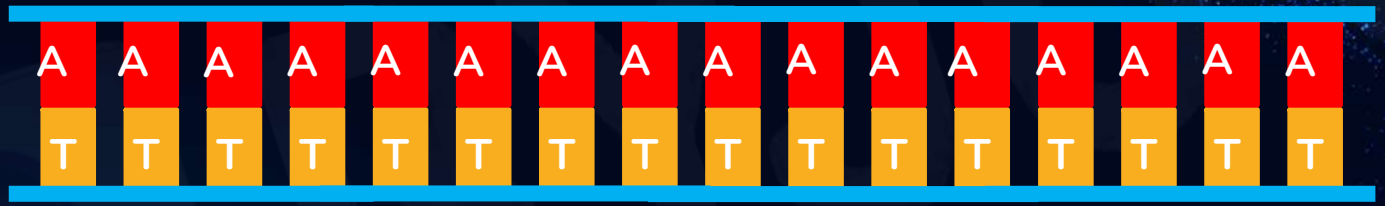


- Most of the genome contains repetitive sequences, like AAAAA or TTTTT or GGGG or CCCCC, they repeat for very long distances like thousands of bases are repetitive only.
- Variable number of tandem repeat (VNTR) is used in genome mapping or in parental identification.

Human Genome Project : Features

Repeated sequences may stretch to sometimes hundred to thousand times

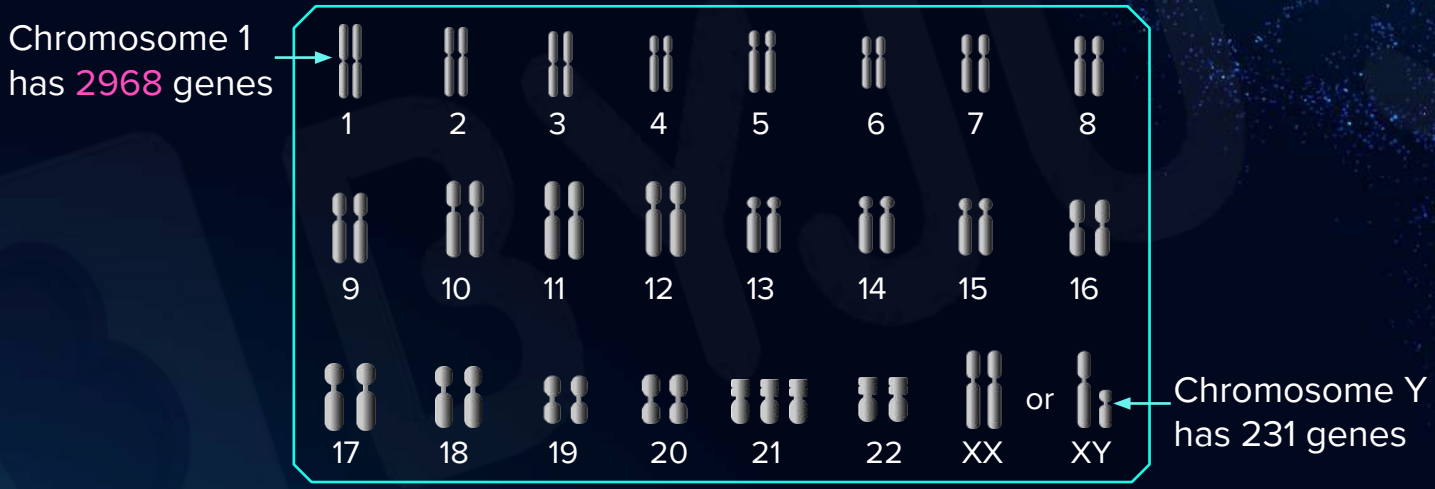
No direct coding function



- Repetitive sequences are nucleotide sequences that are repeated many times, sometimes hundred to thousand of times.
- For eg. AT rich region will have less strength, because of only 2 hydrogen bonds, than GC bond which has three hydrogen bonds.
- So, the function of this region could be to make replication bubble during DNA replication.
- Mostly long DNA repeated sequences have no function but to provide structure to chromosome.

Human Genome Project : Features

Chromosome 1 has 2968 genes while Y-chromosome has 231 genes

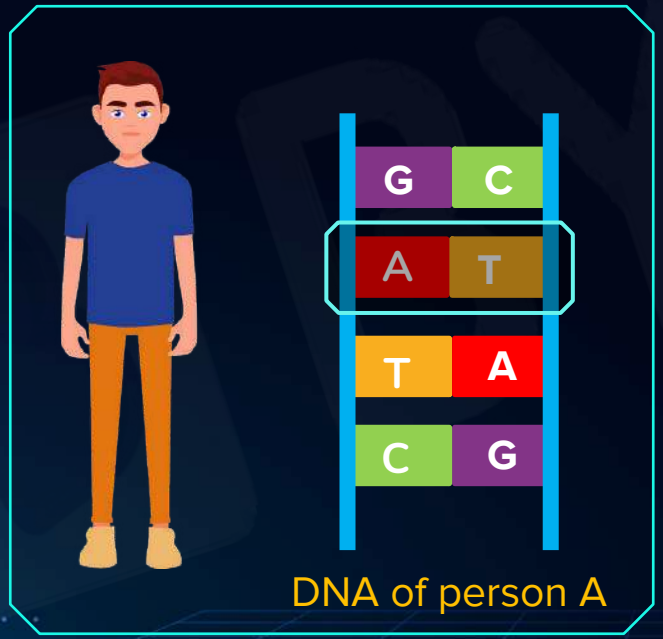


Genome

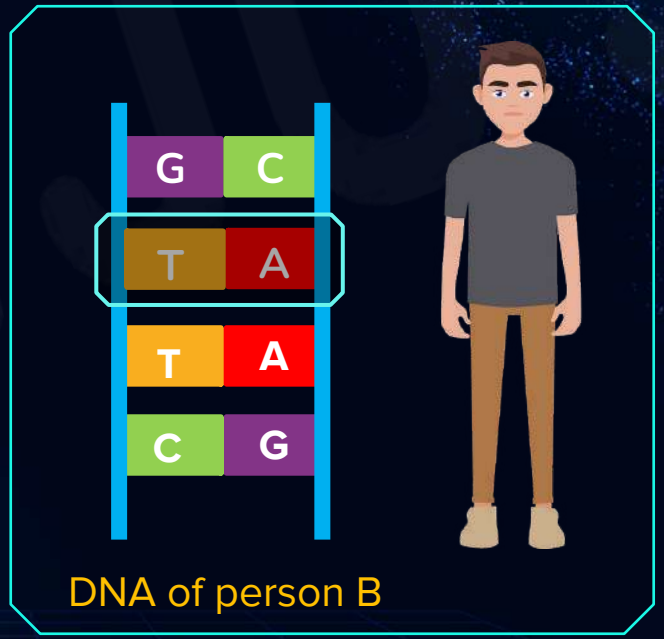
- Chromosome 1 was the last chromosome of the 24 human chromosomes to be sequenced.

Human Genome Project : Features

Approx. 1.4 million location on DNA where single base pair difference occur in humans



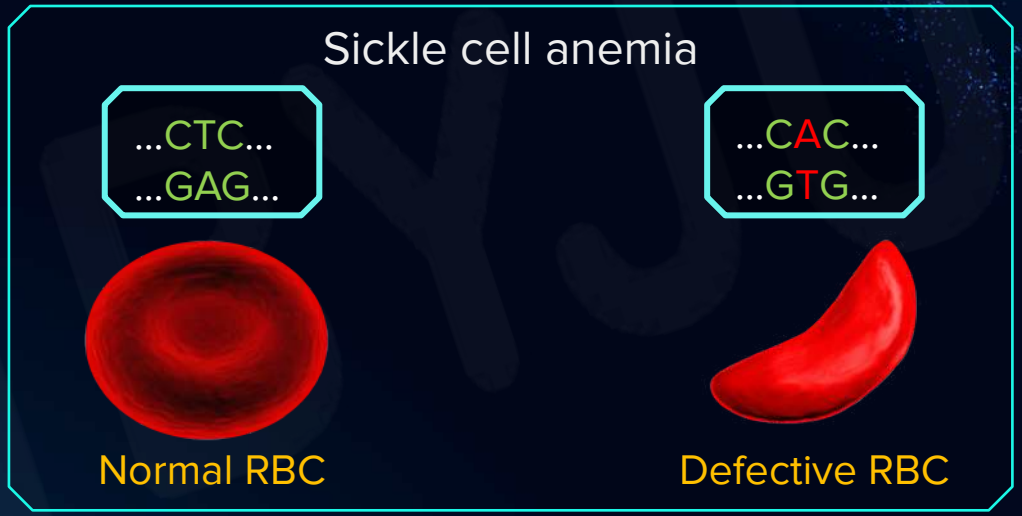
DNA of person A



DNA of person B

Human Genome Project : Applications

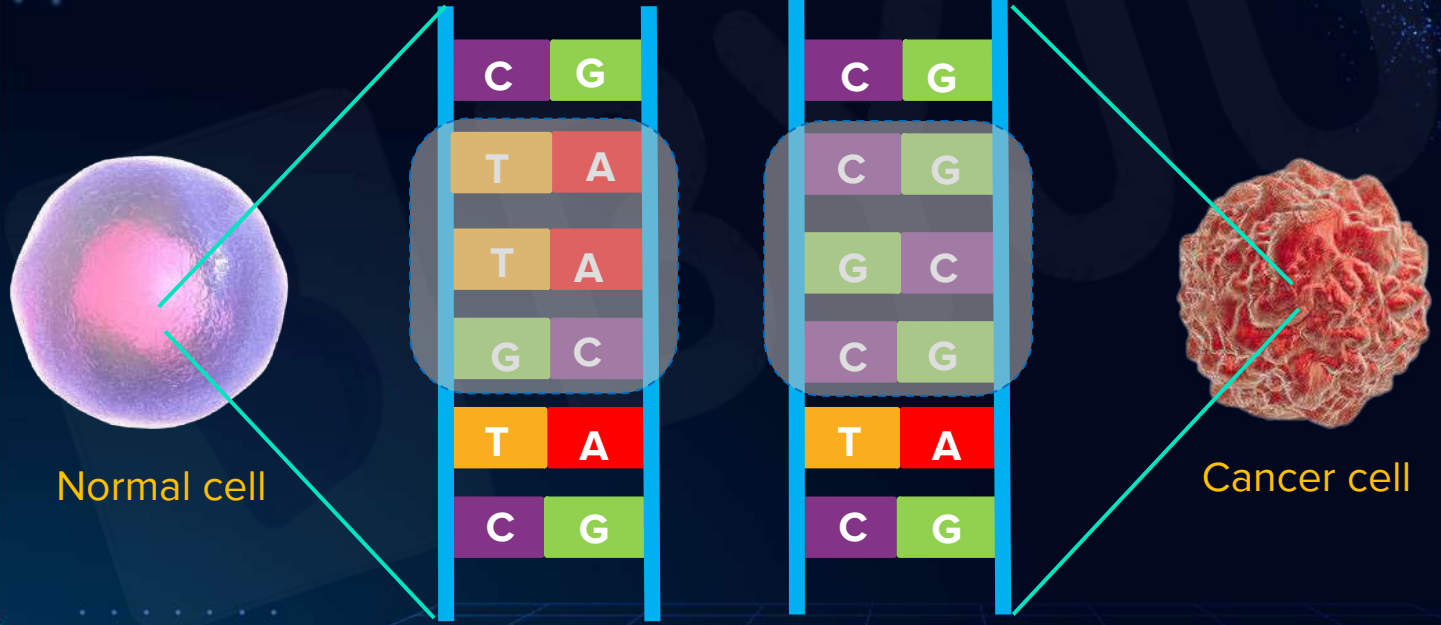
Disease detection



- With the help of genomic sequence, the analysis of changed base pair allow doctors to identify the disease.

Human Genome Project : Applications

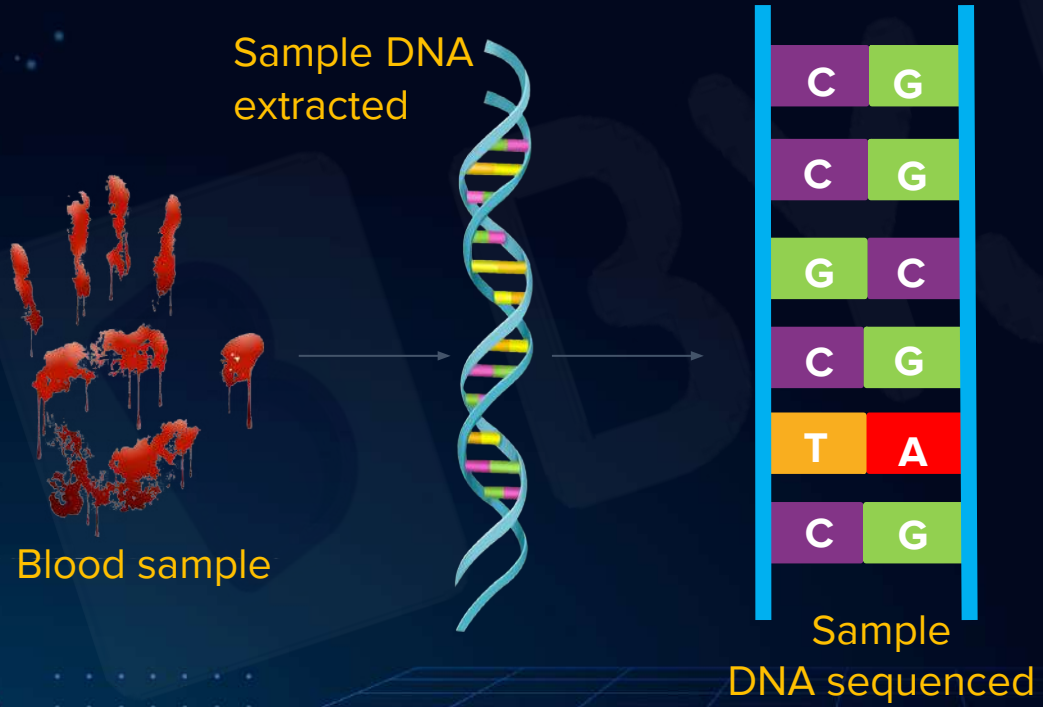
Study of cancerous cell



- We can identify the genetic sequences responsible for the cancerous growth of the normal cell.
- Early diagnosis of cancer can be done.

Human Genome Project : Applications

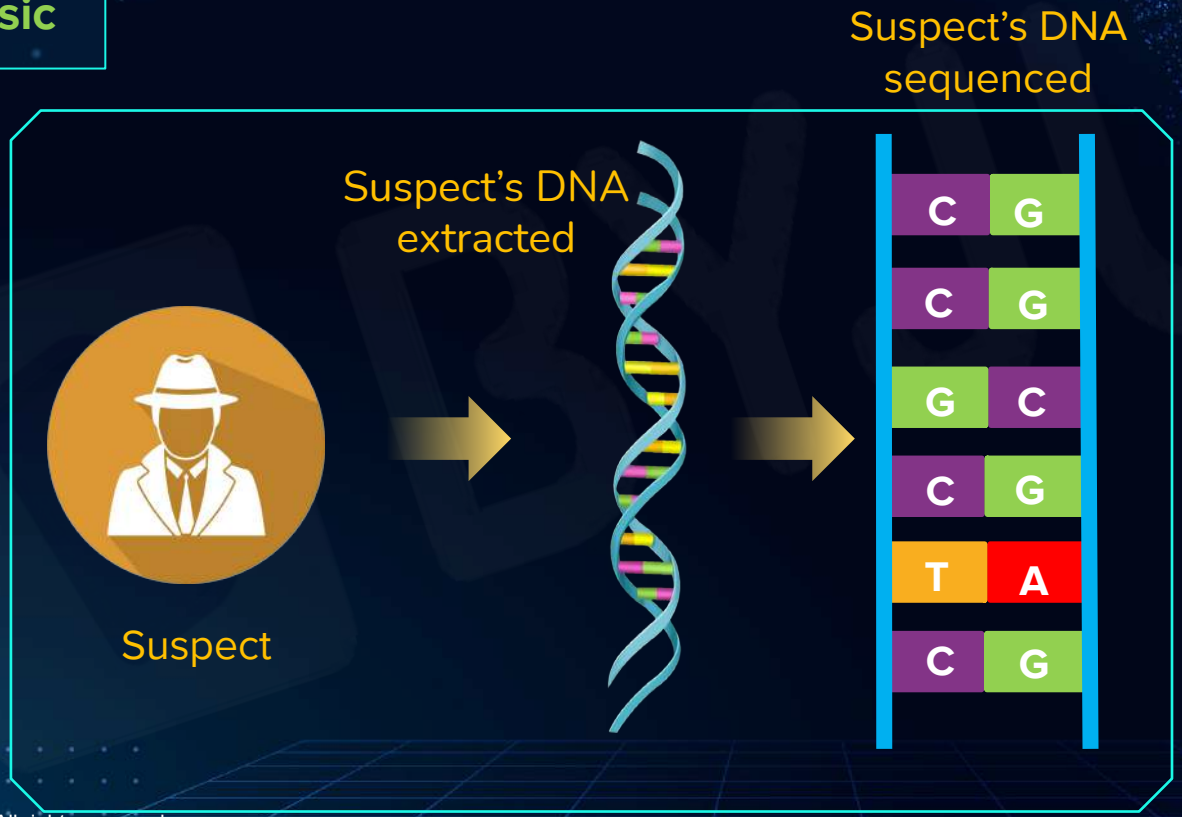
Forensic



- In forensic studies, blood samples are taken to the laboratory, where they isolate the DNA from blood sample obtained at crime scene .

Human Genome Project : Applications

Forensic

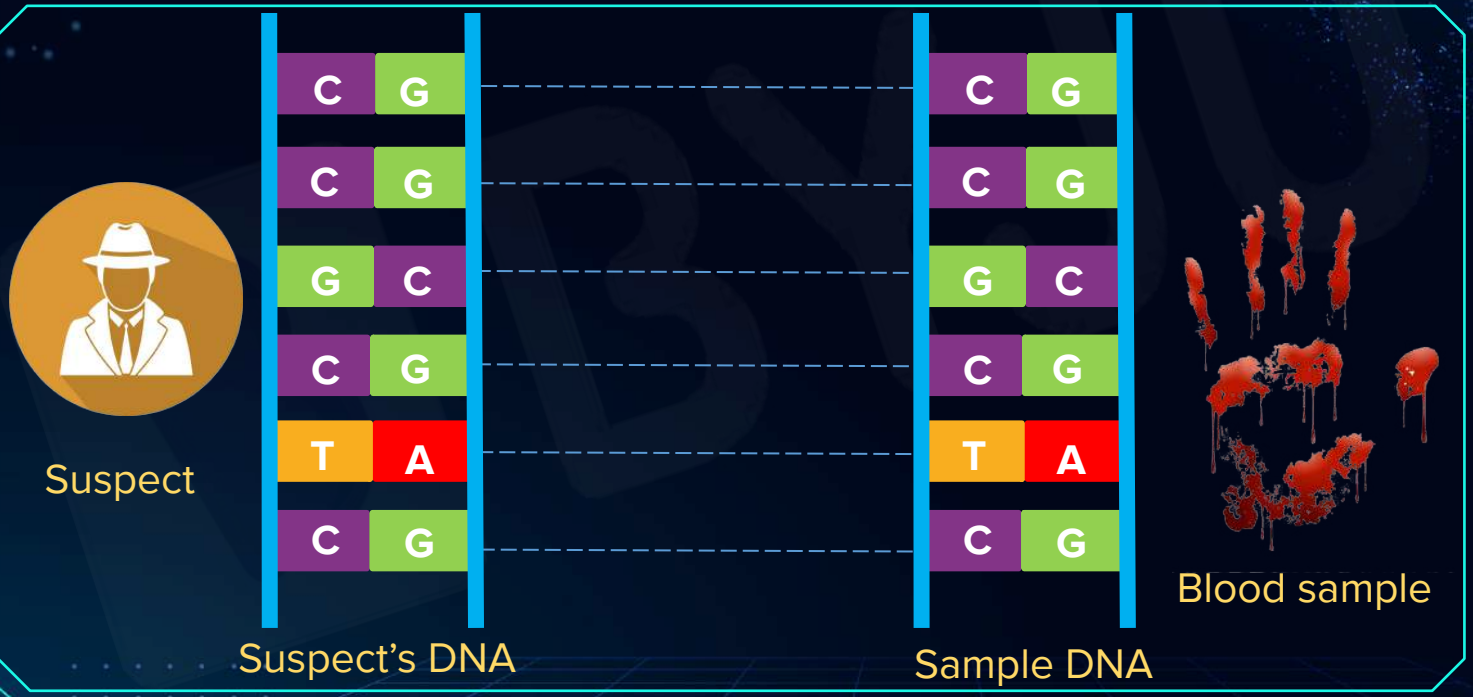


- The blood sample or any other sample like hair, nail etc., is taken to extract DNA of suspect.

Human Genome Project : Applications

Forensic

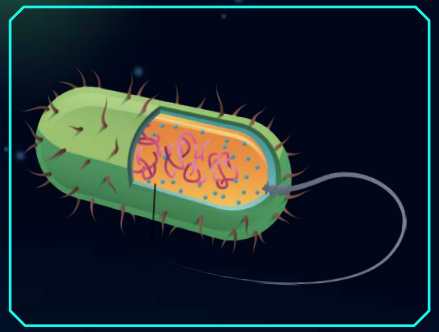
DNA matching



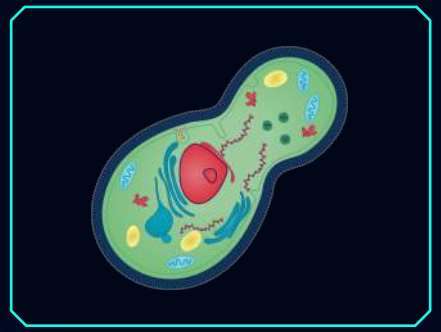
- Upon matching the DNA from sample and from suspect, guilty is caught.

Did You Know ?

Non-human model sequenced genomes



E. coli



Yeast

Rice

Caenorhabditis elegans

Drosophila

Arabidopsis

It helps in the analysis of crops and also helps in increasing the quality of crops.

Which of the following was not used in HGP?

a) DNA sequencers

b) BAC and YAC vectors

c) Plant cells

d) None of the above

Which of the following was not used in HGP?

a) DNA sequencers

b) BAC and YAC vectors

c) Plant cells

d) None of the above

Summary

HGP-Goals

- Identification of approx. 20,000-25,000 genes in human DNA
- Determination of 3 billion chemical base pairs of human DNA
- Storing the information in databases
- Improvement of tools for data analysis
- Transfer related technologies to other industries
- Address ethical, legal and social issues that may arise from HGP

Summary

HGP-Methodology

Two approaches

Expressed sequence tags

i) DNA is isolated from the cell.
ii) mRNA is obtained from this DNA.
iii) Since introns, which are present between two exons, are removed during mRNA synthesis they won't be sequenced.
iv) This way, all the coding genes are isolated and sequenced.

Sequence annotation

i) DNA is isolated from the cell.
ii) This DNA is cut into fragments using enzymes.
iii) DNA fragments are then inserted into vectors/carriers such as bacteria or yeast.
iv) Amplification of DNA.
v) Sequencing, annotating and assigning of DNA.

Summary

Features of HGP

- Human Genome contains approx. 3164.7 million bp
- Average size of gene is 3000 bases
- Human genome has 30000 genes
- Function of over 50% genes is known.
- Only 2% of genome codes for proteins.
- Most of the genome contains repetitive sequences, like AAAAA or TTTTT or GGGG or CCCCC.They are also known as VNTR and are used in genome mapping or in parental identification.
- Chromosome I has 2968 genes while Y-chromosome has 231 genes
- Regions in DNA having one base pair difference are called Single Nucleotide Polymorphisms (SNPs). SNPs can be used in disease detection.

Summary

Applications of HGP

- Early diagnosis of cancer cells
- In forensic medicine to match DNA samples of suspects to reach the criminals
- With genomic sequence, disease like sickle cell anemia can be detected