Date.....................

Nucleosides

1. Ribose + Adenine — Adenosine
2. Deoxyribose + Adenine — Deoxyadenosine

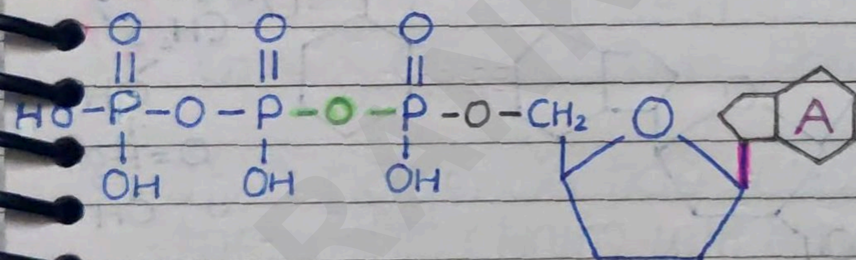
3. Ribose + Guanine — Guanosine
4. Deoxyribose + Guanine — Deoxyguanosine

5. Ribose + Cytosine — Cytidine
6. Deoxyribose + Cytosine — Deoxycytidine

7. Ribose + Uracil — Uridine (RNA only)

8. Deoxyribose + Thymine — Thymidine

Types of bonds in nucleotides



- Phosphoester bond
- Phospho anhydride
- N-glycosidic bond

Adenosine triphosphate (ATP)

[ATP] [GDP] [CMP] [UTP] — Ribose

[dATP] [dGTP] [dTTP] [~~dUMP~~] — Deoxyribose

ATP : Adenylic acid (Adenylate)
GTP : Guanydylic acid
TTP : Thymidylic acid
CTP : Cytidylic acid
UTP : Uridylic acid

Spiral

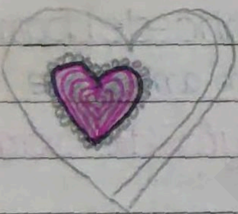
RANKBAAZ.COM

Date 3

Polynucleotide chain

5'

3'



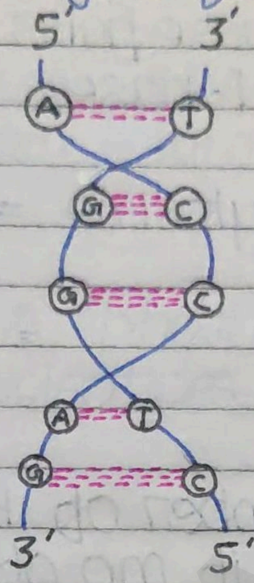
Two nucleotides are linked through 3'-5' Phosphodiester bond to form a dinucleotide chain.

A polynucleotide thus have a free phosphate moiety at 5' end and on the 3' end, it has a free OH group.

Spiral

Date.............

BACKBONE : In a polynucleotide chain is formed by sugar & phosphate. The N-bases linked to the sugar moiety project from the backbone



POLYNUCLEOTIDE CHAIN

- 100 base pairs mean 100-100 bases on both polynucleotide chains.

- A ═══ T (2-H bonds)
  G ═══ C (3-H bonds).

- If a polynucleotide chain has 100 nucleotide then this chain will have 99 3'-5' phosphodiester bond (For linear DNA)

- TYPES OF DNA

Z ⟶ 12 bp / turn — Left handed helix
A ⟶ 11 bp / turn
B ⟶ 10 bp / turn
C ⟶ 9 bp / turn
D ⟶ 8 bp / turn

A, B, C, D — Right handed helix

Date.................

Ques. In a 10 turn of B-DNA we have 20% guanine then total number of cytosine is ?

For 10 turns of B-DNA, base pairs = $10 \times 10 = 100$ bp

N-bases = 200

No of guanine = No of cytosine = $\frac{20}{100} \times 200$

= 40 bases

Ques. Calculate the number of H bond in a 100 base pair B-DNA if no of guanine is 40?

No. of G ≡ C base pairs = 40

No. of A = T base pairs = 60

Total H bonds = $3 \times 40 + 2 \times 60$

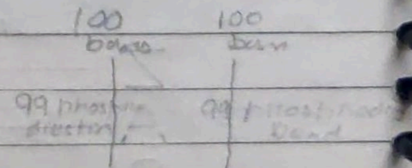
= 240 Hydrogen bonds

Length of DNA = $0.34 \times 10^{-9} \times$ (No of bp)

= $0.34 \times 10^{-9} \times 100$

= 34 nm

For above question



No of phosphodiester bond = 198

No of phosphoester bond = 2 × (No. of phosphodiester bonds)

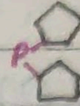
= 2×198

= 396 bonds

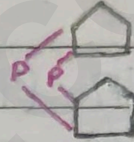
Date....................

No. of glycosidic bond = No. of N-bases
= 200

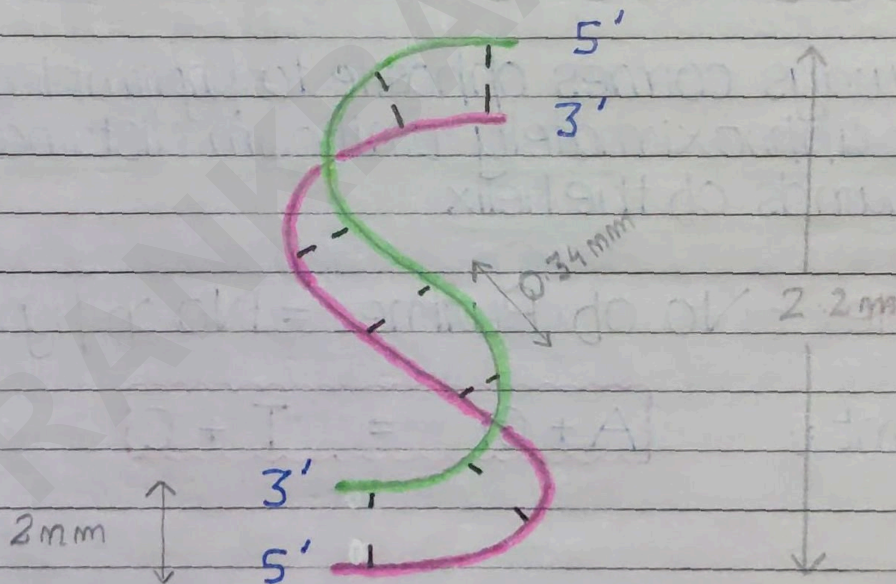
Note: For a linear DNA, phosphodiester bond is always 1 less than total no of nucleotides in that chain



For circular DNA, phosphodiester bonds are equal to the no of nucleotides



SALIENT FEATURES OF B-DNA



Watson & Crick explained the secondary structure of B-DNA based on the X-Ray diffraction data produced by Maurice Wilkins & Rosalind Franklin. One of the hallmarks of this proposition was base pairing b/w 2 strands of polynucleotide chains. However this proposition was also based on observations of Erwin Chargaff.

Spiral
RANKERZ.COM

Validity of chargaff's rule: It is not applicable to single stranded poly nucleotide chain

Eg RNA → messenger RNA, ribosomal RNA, transfer RNA } Not applicable

Date.....................

1. DNA is made up of 2 polynucleotide chains (double helix structure) having anti-parallel polarity.

2. Pitch of helix = 3.4 nm (34 Å)
Rise per base pair = 0.34 nm (3.4 Å)
Diameter (distance b/w 2 chains) = 2 nm

3. Total turn in one pitch = 360°
Turn per base pair = 36°

4. Total base pair (in diploid) = 6.6×10^9 bp
Length of DNA = 0.34×10^{-9} (No of bp)
$= 0.34 \times 10^{-9} \times 6.6 \times 10^9$
$= 2.2$ m

5. Purine always comes opposite to pyrimidine, this generates approximately uniform distance b/w the 2 strands of the helix

Chargaff's Rule: No of Purine = No of pyrimidine

Statement: $A + G = T + C$

ie $\frac{A+G}{T+C} = 1$ But $\frac{A+T}{G+C}$ = constant
base ratio

$(A+T)/(G+C)$ is constant for a given specie, $G \equiv C$ content is more in organisms with stable DNA

1. Human = 1.52
2. E Coli = 0.92

Spiral

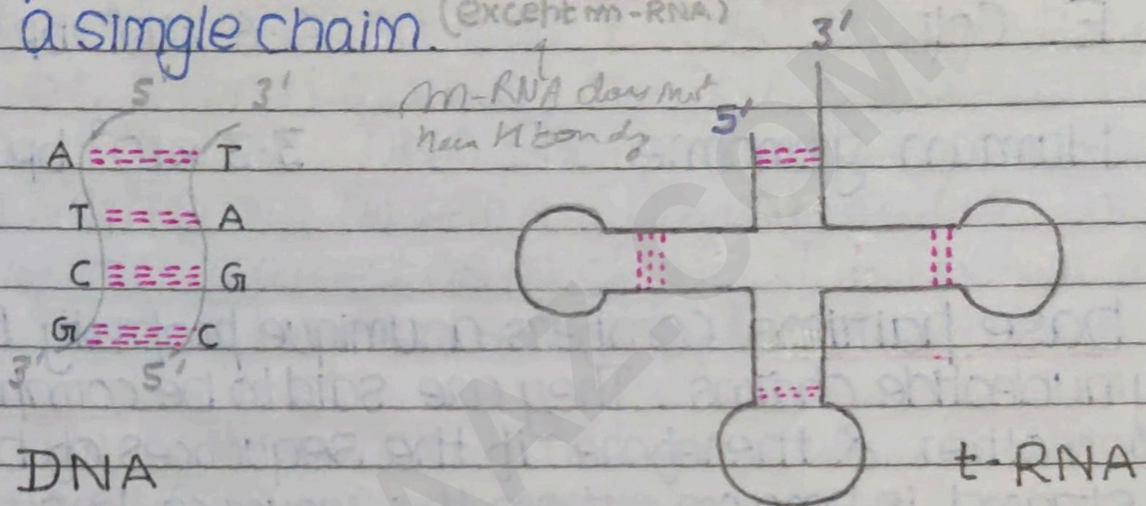
Plane of one base pair stacks over the other in double helix. In addition to H-bonds, confers stability to the helical structure.

Date.....................

NOTE : DNA is an acidic substance present in the nucleus was first identified by Friedrich Meischer in 1869, he named it as NUCLEIN

DNA show inter H bonding between 2 polynucleotide chains, but RNA shows intramolecular H bonding within a single chain. (except m-RNA)

m-RNA does not have H bonds



Functional form of B-DNA = 2° structure

Functional form of m-RNA = 1° structure
Functional form of t-RNA = 2° structure (and 3°)
Functional form of r-RNA = 3° structure
catalyst

Comparison between RNA

Amount : r-RNA > t-RNA > m-RNA
80% 15% 5%

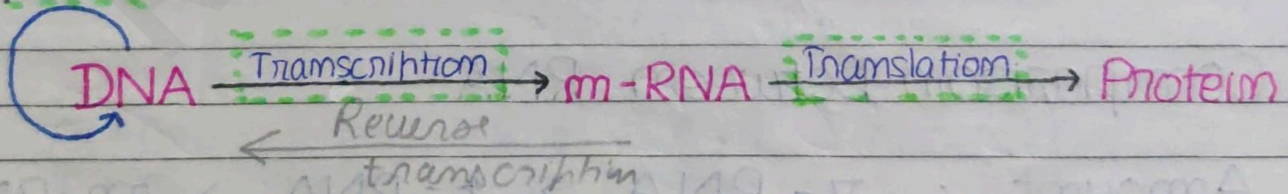
Stability : r-RNA > t-RNA > m-RNA

Size of RNA : m-RNA > r-RNA > t-RNA

Organism		
1. ϕ×174 Bacteriophage ssDNA	5386 nucleotide	
2. Bacteriophage lambda dsDNA	48502 bp	
3. E. Coli dsDNA	$4 \cdot 6 \times 10^6$ bp	
4. Human genome ds DNA	$3 \cdot 3 \times 10^9$ bp	

The base pairing confers a unique property to the polynucleotide chains. They are said to be complimentary to each other & therefore if the sequences of bases in one strand is known, then the sequence in other strand strand can be predicted.

Francis Crick proposed the central dogma of molecular biology which states that genetic info flows from



In some viruses, the flow of information is in reverse direction ie from RNA to DNA, this is called to be REVERSE TRANSCRIPTION using reverse transcriptase which is a DNA polymerase (RNA dependent DNA polymerase)

Date..................

Packaging of DNA helix...

In a typical cell a large size DNA ie 2.2m is accumulated in a small nucleus whose dimension is 10^{-6}m

In prokaryotic cells such as E. Coli, though they do not have a defined nucleus, the DNA is not scattered throughout the cell. DNA (being negatively charged) is held with some proteins (having positive charges) in a region termed as nucleoid. The DNA in nucleoid is organised in large loops held by proteins (Polyamines)

NOTE: In prokaryotic cell histone protein is absent but Packaging of DNA occurs around

1. Non-Histone protein
2. Basic proteins
3. Polyamines ⟶ Positively charged.

In E. Coli, DNA contains 4.6×10^6 bp thus total length is 1.56 mm while the size of cell is about 1 um
↳ Acc to NCERT 1.36 mm

Packaging of DNA in EUKARYOTES

In eukaryotes, there is a set of positively charged basic proteins called histones.

(A protein acquires charge depending upon the abundance of amino acids residues with charged side chains) Histones are rich in basic amino acid residues ARGENINE and LYSINES and are positively charged

Spiral

Date.....................

Histone proteins are of 5 types ie.

- → H_1
- → H_{2A}
- → H_{2B}
- → H_3
- → H_4

2 molecules each of H_{2A}, H_{2B}, H_3, H_4 are joined together to form a unit of 8 molecules histone octamer

The DNA is coiled around the histone octamer / nu body around $1\frac{3}{4}$ times or 1.75 times

The octamer and the DNA coiled above it is known as nucleosome, the DNA present on nucleosome is about 200 bp.

The DNA present between nucleosomes is called linker DNA containing about 80 bp

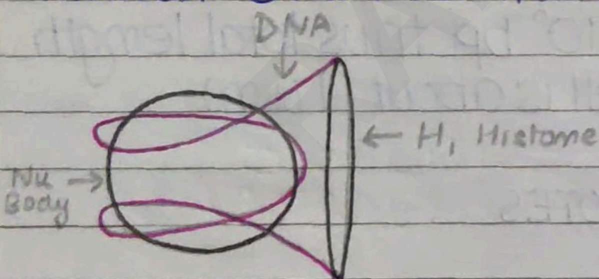


Fig. Nucleosome

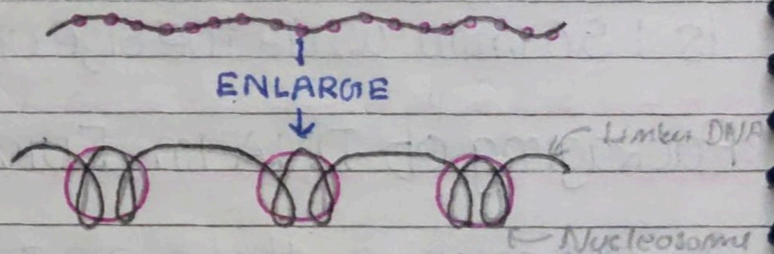


Fig. Beads on string

The coiling of DNA around the histone octamers occurs due to the fact that DNA is negatively while histone is positively charged.

Spiral

Date.................

Further coiling and condensation take place & ultimately forms chromosome. At higher level of packaging, non-histone chromosomal (NHC) protein, is also required.

No of nucleosomes in human somatic DNA strand $= \frac{6.6 \times 10^9}{200}$

$= 3.3 \times 10^7$ Nucleosom

The nucleosomes in chromatin are seen as 'beads on string' structure when viewed under electron microscope [EM]

CHROMATIN

Heterochromatin	Euchromatin
• Dark stained	• Lightly stained
• Genetically inactive (Transcription absent)	• Genetically active (transcription occurs)
• Late replicable	• Early replicable

The beads-on-strings structure in chromatin is packaged to form chromatin fibres that are further coiled and condensed at metaphase of cell cycle form compact mitotic chromosomes.

Spiral

Agar kisi gene ko inactive karna h to cell use heterochromatin me convert kar deta h
Eg: Our skin cells do have genes for insulin formation but in inactive form (Heterochromatin)

Date.....................

THE SEARCH FOR GENETIC MAT.

I. 1928 Griffith : TRANSFORMATION

- Streptococcus pneumoniae (Pneumococcus)

When Streptococcus pneumoniae bacteria are grown on culture plate :

→ Some produce smooth shiny colonies — S strain
→ While others produce rough colonies — R strain

This is becase S - strain bacteria has mucous coat (Polysaccharide), while R strain does not

S strain (S-III) → Smooth
→ Capsulated
→ Pathogenic
→ Virulent

R strain (R-II) → Rough
→ Non-capsulated
→ Non-pathogenic
→ Avirulent

S strain — INJECT → Mice dies
R strain — INJECT → Mice live

When griffith injected a mixture of heat killed S & live R strain, the mice died. Moreover he recovered live S bacteria from the dead mice

S strain (Heat killed) + R strain — INJECT → Mice dies

Spiral

Date.....................

He concluded that the R strain bacteria had been **transformed** by heat killed S strain.

← Basically DNA

Some "**Transforming principle**" must have been transferred from heat killed S strain had enabled the R strain to synthesise a smooth polysaccharide coat & become virulent. This must be due to the transfer of the genetic material.

However the biochemical nature of genetic material was not defined from his experiments.

II. 1933-44 Biochemical characterisation of the "Transforming principle"

Prior to the work of Avery, Macleod, McCarty, the genetic material was thought to be protein

They purified biochemicals (proteins, DNA, RNA, etc) from heat killed S bacteria. They discovered that DNA alone from S bacteria caused R bacteria to become transformed.

Heat Killed S-III + R-II —(→ RNAase/Protease) INJECT → Mice dies

Heat killed S-III + R-II —(→ DNAase) INJECT → Mice lives

Protease & RNAase cannot affect transformation but DNAase inhibited transformation.

Spiral

Date.................

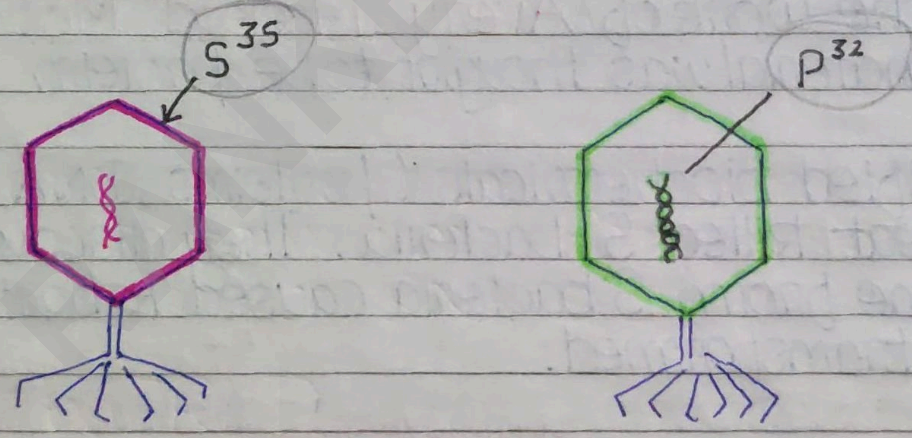
They comcluded that DNA is the heniditany matenial matenial but mot all biologists wene comvimced.

III. 1952 The umequivocal pnoob

The umequivocal pnoob that DNA is the genetic matenial came bnom the expenimemts Albned Hershey and Martha Chase (1952). The wonked om Vinuses that imbect bactenia (bacteniophages)

Hershey & chase obsenved lytic cycle ob bacteniophage. The bactenial cell treats the vinal genetic matenial as ib it was its own and subsequently mamubactunes more vinus particles.

EXPERIMENT



- Radioactive S^{35} labelled pnoteim capsule
- Radioactive P^{32} labelled DNA

→ Radiactive phages wene allowed to attach to E.coli bactenia [INFECTION]

→ The vinal coats wene nemoved bnom the bactenia by agitating them im a blemden [BLENDING]

Spiral

No radioactivity was detected in supernatant when P^{32} virus was allowed to attack

Date....................

→ The virus particles were separated from bacteria by spinning them in a centrifuge [CENTRIFUGATION]

In this experiment, 2 different bacteriophage is labelled, one with radioactive sulphur S^{35} & other with radioactive phosphorous P^{32}

S being a constituent of amino acid is incorporated in protein ∴ This virus possesses viral protein coat

P being a constituent of nucleotide is incorporated in DNA ∴ This virus possesses viral DNA.

After centrifugation it was observed that P^{32} was present inside host cell while S^{35} was present in supernatant fluid indicating that DNA was the genetic material & Not protein

PROPERTIES OF GENETIC MATERIAL

From Hershey chase experiment its an established fact that DNA acts as genetic material. Also in some viruses RNA acts as genetic material

Eg. TMV virus
QB bacteriophage

FEATURE 1: It should be able to generate its replica
FEATURE 2: Should be chemically & structurally stable
FEATURE 3: Should provide scope for slow mutation
FEATURE 4: Should express itself in the form of mendelian characters

Date.....................

DNA is more stable than RNA because :

- 2' OH group makes RNA labile (easily altered) & easily degradable.
- RNA is known to be catalytic (Ribozyme), hence reactive. DNA never acts as catalyst
- Presence of thymine at the place of uracil also confers additional stability to DNA
- Both RNA & DNA are able to mutate, but RNA being unstable mutate faster

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

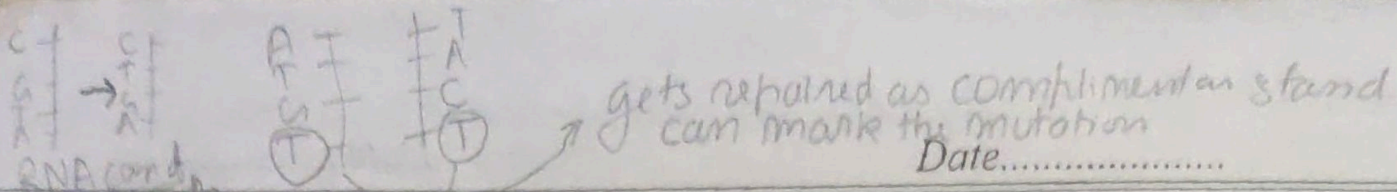
DNA being more stable is preferred for storage of genetic information. For transmission of genetic information RNA is better.

RNA WORLD

RNA was the first genetic material. RNA is furthermore involved in process like

→ Metabolism (Forms enzymes → Proteins)
→ Splicing
→ Translation

Spiral



DNA being double stranded and having complementary strand further resists changes by evolving a repair mechanism

DNA REPLICATION

While proposing the double helical structure, Watson & Crick has immediately proposed a scheme for replication of DNA.

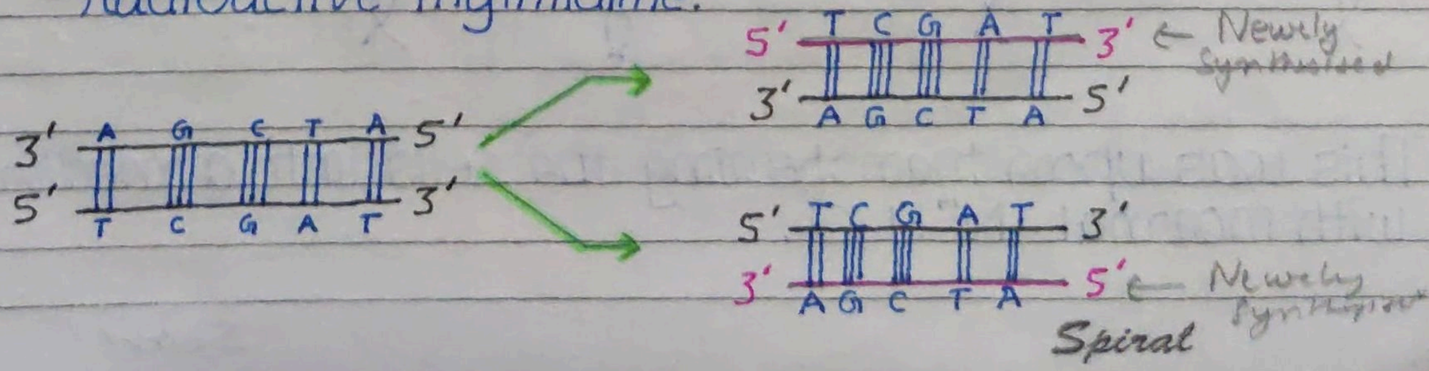
" It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for genetic material."
— Watson

The scheme suggested that the two strands would separate and act as template for synthesis of complimentary strands.

Experimental proof of DNA replication...

Semiconservative mode of DNA replication was predicted by Watson & Crick but was experimentally proven by

1. Matthew Meselson & Franklin Stahl (1958) in E.coli
2. Taylor & His collegues (1958) in Vicia faba using radioactive thymidine.



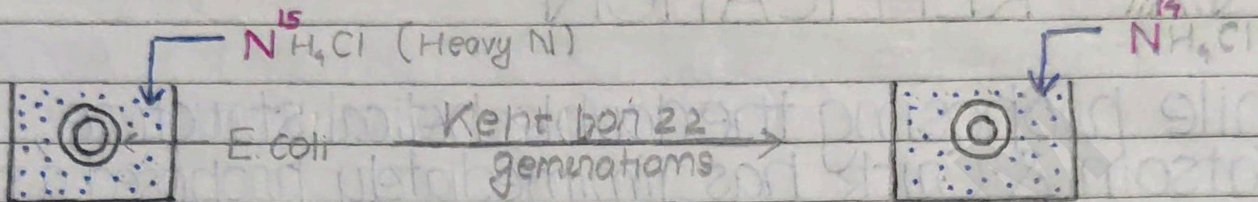
Spiral

N^{15} = Heavy isotope ✓
~~Radioactive isotope~~ ✗

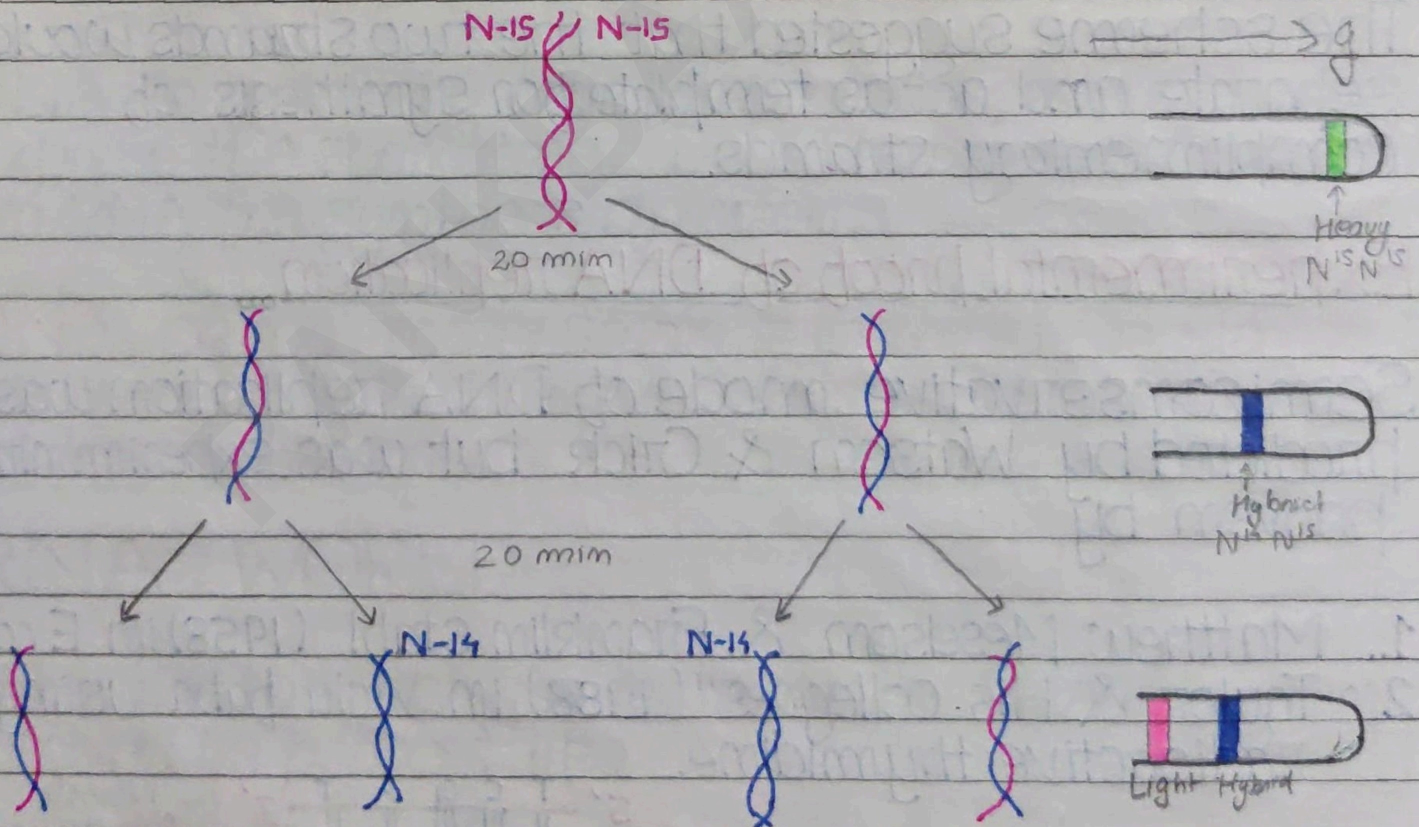
Date....................

Meselsom & stahl experimemt

DNA replicatiom is semi comsenvative is experimemtally proved by Meselsom & stahl im 1958



E. Coli was cultured im a medium comtaiming $N^{15}H_4Cl$ as the omly source of mitrogem for mamy gemenatioms The result was that the mewely symthesised DNA strand has imcorporated N^{15} from the medium.



This was upom tramsferrimg the cells imto a medium with mormal $N^{14}H_4Cl$.

Ecoli connection: (NCERT) Ecoli completes the process of DNA replication in 18 minutes

→ E Coli divides in 20min in which 18 min it gives to DNA replication
38 min thing was UPDATED

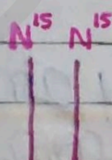
Date.....................

Density of DNA is measured using CsCl density gradient method. In newely formed DNA molecule one strand is parental while other strand is new. Thus it is called semiconservative mode of replication

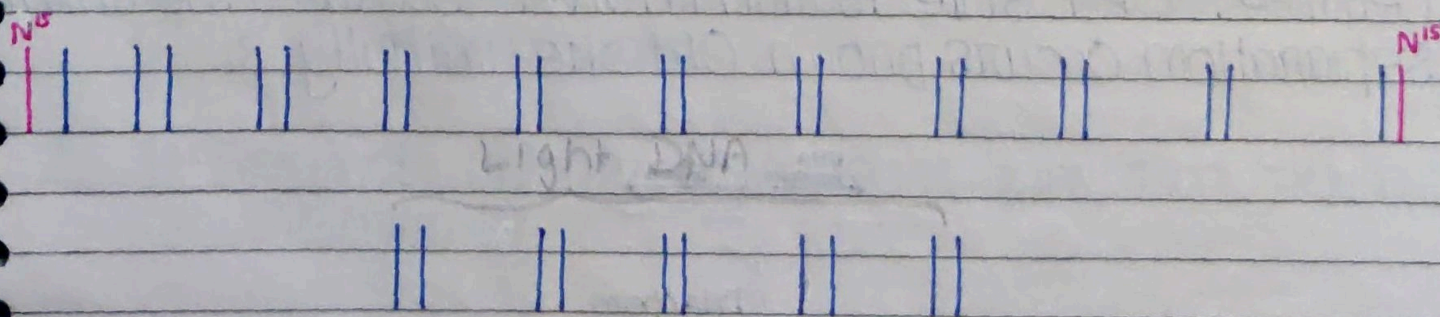
* DNA replication is semi-conservative is also proved by Taylor Et al. with the help of tritiated thymidine on tip of vicia faba (Faba beans)

* DNA replication is semi-conservative was also proven by Cairn with the help of tritiated thymidine on E. Coli bacteria using autoradiography.

Ques (NCERT): If E. coli was allowed to grow for 80min, find ratio of light to hybrid density DNA



After 4 divisions (80min). 16 DNA molecules are synthesised as shown



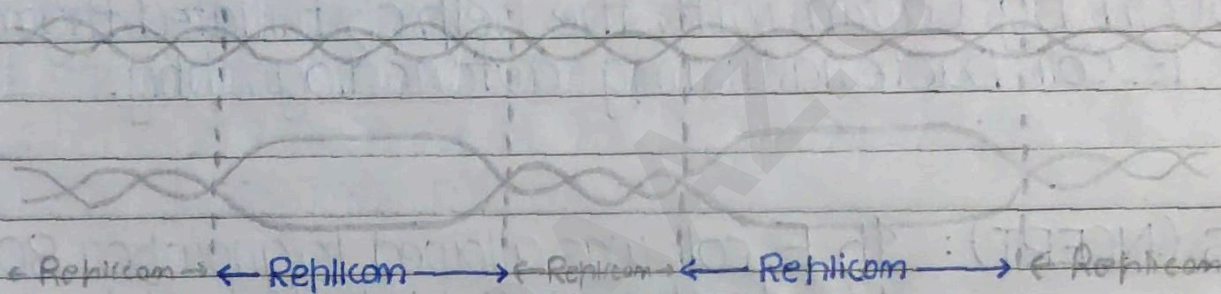
Ratio of Light : Hybrid = 14 : 2

Spiral

THE MACHINERY AND ENZYMES

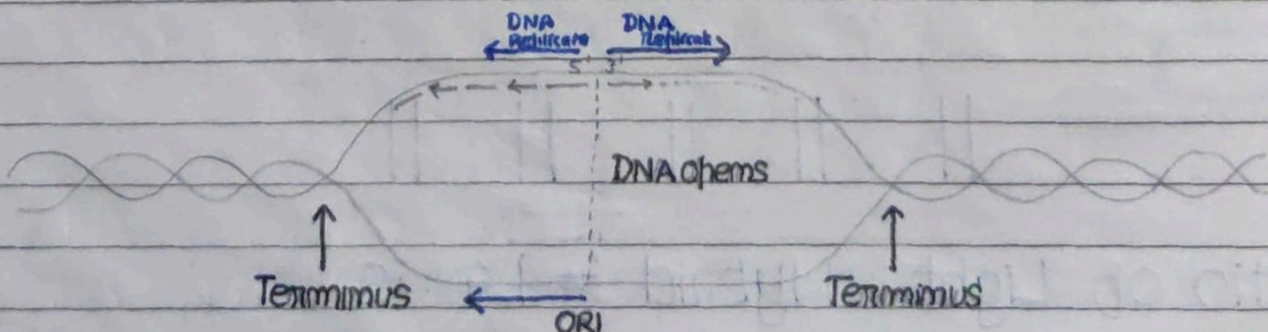
In prokaryotic cell, DNA polymerase is found in cytoplasm while in eukaryotes, it is present in the nucleus as DNA polymerase is required during DNA replication.

In Eukaryotes DNA replication occurs in S-phase while in prokaryotes it occurs prior to fission



Total DNA is structurally and functionally divided into segments called replicon. All replicons do not start DNA replication simultaneously.

Each replicon has 2 terminus and 1 ori at the centre. ORI site is rich in A-T because the strand separation occurs from ORI site initially.



DNA Replication is bidirection

Spiral

NCERT: For long DNA molecules, since the 2 strands cannot be separated in its entire length (due to very high energy requirement, the replication occurs within a small opening of the helix called Replication fork

Date.....................

In eukaryotes, 1000s of replicons (∵ ORI) are present, but in prokaryotes only 1 ori site is present ∴ The whole DNA strand acts as replicon

DNA replication starts in the centre of replicon and in both directions.

DNA replication steps

Helicase: Breaks H bond

DNA replication initiated by separation between the strands, that occurs with the help of DNA Helicase and tension is released by topoisomerase enzyme.

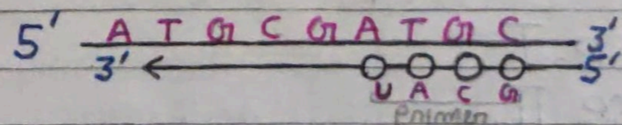
Note: In prokaryotes, in place of topoisomerase & helicase, a single enzyme DNA gyrase is present

After separation, SSBP (single strand binding protein) binds to the DNA strand to prevent self reassociation of the 2 strands due to complimentarity.

DNA dependent DNA polymerase has 2 conditions

1. DNA Polymerase catalyse polymerisation in 5'-3' direction only.

2. It cannot initiate DNA replication, and therefore require primer.

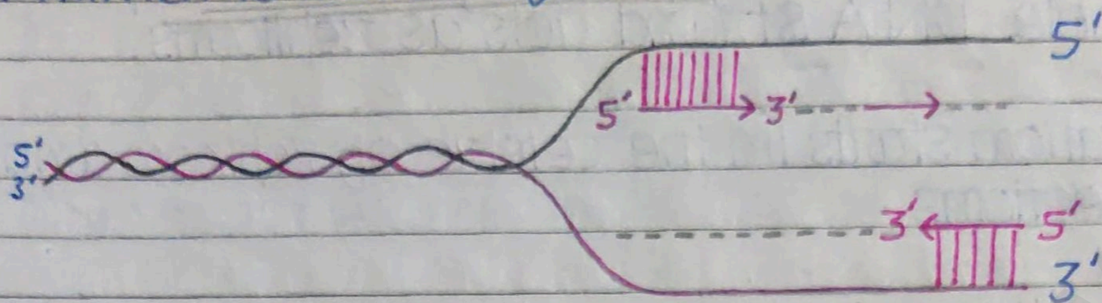


Spiral

we study this

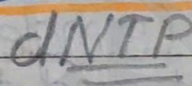
Date.................

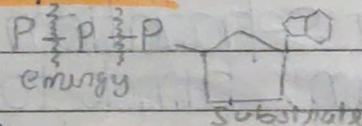
Firstly synthesis of primer occurs by primase enzyme. (RNA Polymerase)
Primer is made of RNA



DNA replication occurs by complimentary base pairing. The replication can start only in 5'-3' direction. Hence synthesis on 3'-5' strand is continous, known as Leading strand and that on 5'-3' end is discontinous, known as Lagging strand (Also called Okazaki fragments)

The deoxyribo nucleoside triphosphate serves duel purpose :





1. Acts as substrate for DNA polymerase
2. Provide energy for polymerisation (due to the 2 terminal phosphate bond) as replication is a very energetically expensive process.

DNA polymerase polymerises 2000 bp per second with high degree of accuracy. In prokaryotes DNA polymerase is of 3 types :

1. DNA Polymerase I
2. DNA Polymerase II
3. DNA Polymerase III

All 3 have polymerase as well as exonuclease activity

Spiral

Primer hamesha RNA ka hi hoga because even ek chhota sa DNA strand banane ke liye bhi DNA polymerase cahiye but DNA polymerase kabhi bhi ZERO se nai start kar sakti, lekin RNA polymerase can start synthesis of RNA strand from ZERO

Date.....................

* All the primers are removed by DNA polymerase I
* At the place of primers, synthesis of DNA occurs by DNA polymerase I

* Proof reading & editing take place by DNA polymerase III

* Major repairing by DNA polymerase I (DNA repair enzyme). Minor repairing by DNA polymerase II.

* DNA replication takes place by DNA Polymerase III (Main DNA polymerase) → Prokaryotes

All the fragments are joined by DNA ligase. DNA replication is therefore semiconservative and semi discontinous.

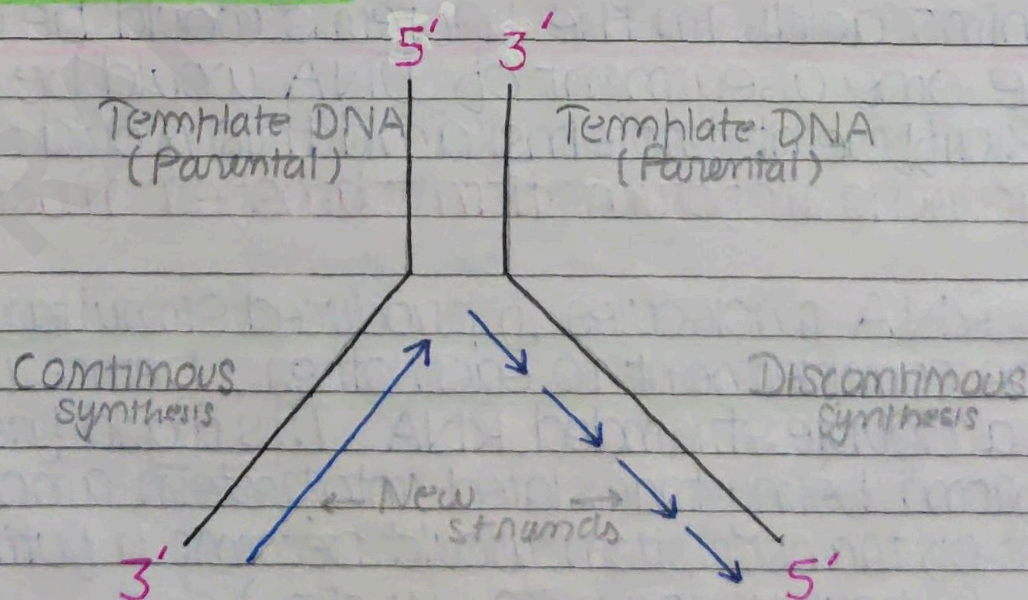


Fig. Replication fork

Spiral

Transcription } Prokaryotes → Cytoplasm
DNA replication } Eukaryotes → Nucleus

Date.....................

TRANSCRIPTION

→ DNA dependent RNA polymerase
↳ Does not require PRIMER

It is the process of copying genetic information from one strand of the DNA into RNA is called transcription

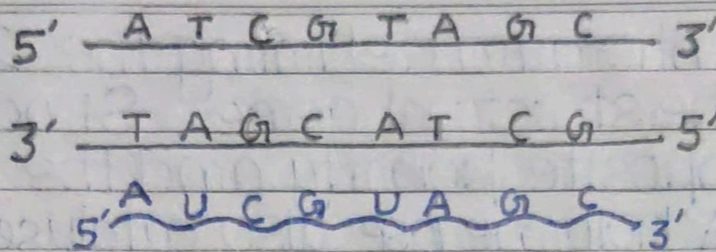
Unlike DNA replication where total 2.2m DNA of an organism gets duplicated, in transcription only a segment and only one of the strands is copied into RNA. Here, only one strand is template and the other one is non-template, in DNA replication both strands were template strand.

Both strands of DNA are not being copied during transcription because:

(1) If both strands act as template, they would code for RNA molecule with 2 different sequences and in turn, if they code for proteins, the sequence of amino acids in the proteins would be different. Hence one segment of DNA would be coding for 2 different proteins and this would complicate the genetic information transfer machinery

(2) The 2 RNA molecules if produced simultaneously would be complement to each other, hence would form a double stranded RNA. This would prevent RNA from being translated into protein and the process of transcription would become a futile one (futile = having no success ; useless)

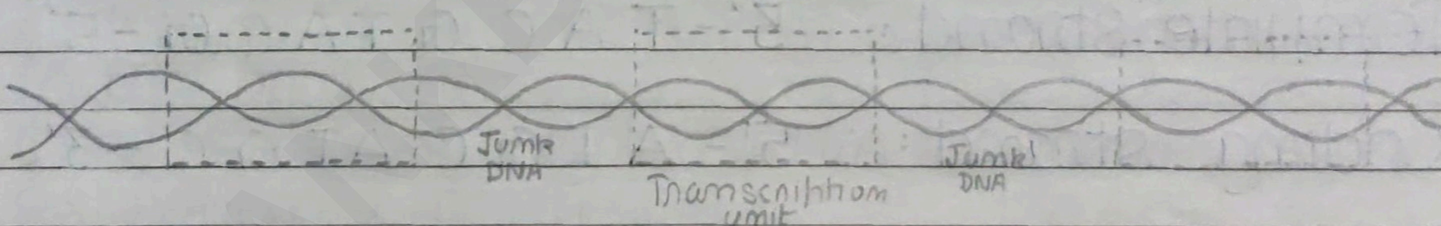
Spiral



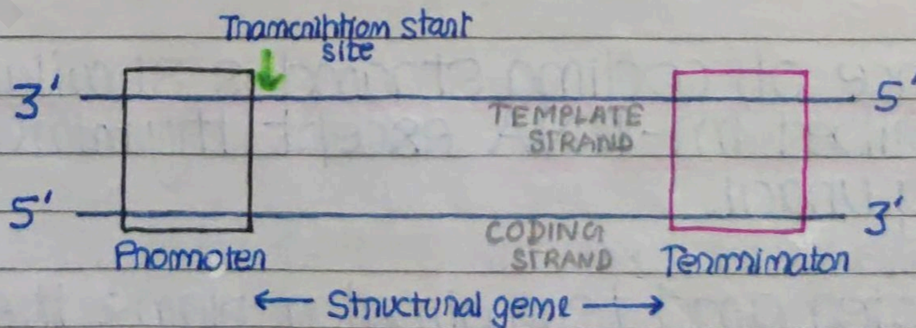
Here also the principle of complementarity governs the process of transcription, except the adenine now forms a complementary uracil instead of thymine in the m-RNA

TRANSCRIPTION UNIT

Since transcription does not occur for entire DNA ∴ it becomes a necessity to define the boundaries that would demarcate the region and the strand of DNA that would be transcribed



Transcription unit is a segment of DNA which shows independent event of transcription.



A transcription unit has 3 regions :

1. Promoter
2. Structural gene
3. Terminator

Spiral

Date......................

There is a convention in defining the 2 strands of the DNA in the structural gene. Since the 2 strands have opposite polarity and the DNA dependent RNA polymerase also catalyse the polymerisation in only one direction ie 5' → 3', the strand that has polarity 3' → 5' acts as a template and is also referred to as template strand.

The other strand which has the polarity 5' → 3' and the sequence similar to RNA is displaced during transcription. Strangely, this strand which does not code for anything is called coding strand. All the references while defining transcription unit are made with the coding strand.

→ Non coding / Antisense strand

Template strand : 3' - T A C G T A C G - 5'

Coding strand : 5' - A T G C A T G C - 3'

→ Non template / Sense strand

Transcribed RNA strand : 5' - AUGCAUGC - 3'

The sequence of coding strand is similar to that of transcribed m-RNA except thymine is replaced by uracil.

The promoter and terminator flank the structural gene (ie present on either sides)

Promoter is said to be located towards 5'-end (upstream) of the structural gene

Spiral

Date...................

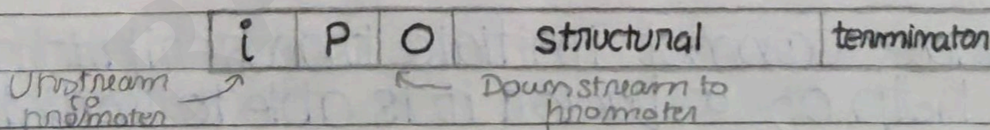
Promoter is a DNA sequence that provides binding site for RNA polymerase. The presence of a promoter in a transcription unit also defines the template & coding strands (as promoter lies with the coding strand)

By switching the position of promoter with the terminator, coding and template strands are reversed as transciption occurs on strand that doesnt have promoter (template strand)

The terminator is present towards 3'-end (downstream) of the coding strand and it usually defines the end of the process of transcription

a gene is anything that codes for any RNA, but regulatory genes do not code for any RNA

NOTE : There are additional regulatory sequences that may be present further upstream or downstream to the promoter. Eg some of the properties of these sequences are involved in regulation of gene expression (Lac operon)



TRANSCIPTION IN PROKARYOTES

In prokaryotic cell, a single DNA dependent RNA polymerase enzyme is able to synthesize all the diff types of RNA ie m-RNA
r-RNA
t-RNA

NCERT

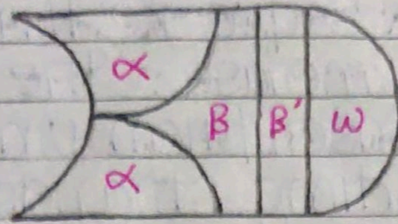
Spiral

Types of RNA

→ r-RNA → Sn-RNA (Small nuclear RNA)
→ t-RNA → Sc-RNA (Small cytoplasmic RNA)
→ m-RNA → Hn-RNA (Heterogenous nuclear RNA)

Date.....................

RNA Polymerase



RNA polymerase is a multisubunit enzyme consisting of $\alpha_2 \beta \beta' \omega$ (Core enzyme)

All components except σ factor is known as core enzyme. If core enzyme binds to σ factor then it is called holoenzyme

Transcription is completed in 3 steps viz.

1. Initiation
2. Elongation
3. Termination

INITIATION

RNA polymerase cannot initiate transcription but with the help of σ factor it is able to initiate transcription

Note: On its own RNA polymerase is only capable of catalysing the process of elongation. It associates transiently with initiation factor (σ) and termination factor (ρ) to initiate and terminate transcription respectively.

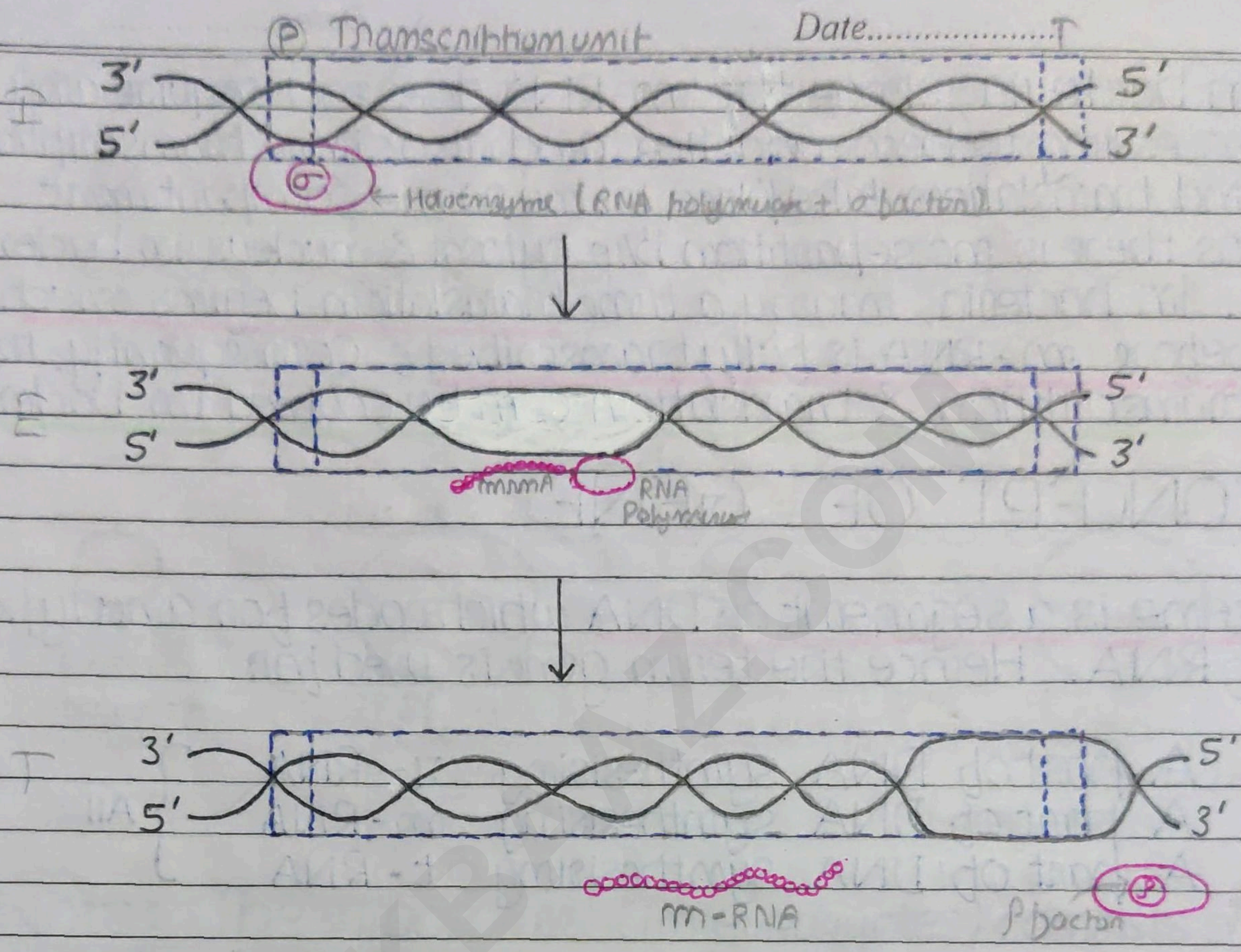


Fig: Process of transcription (Bacterial)

RNA polymerase also uses nucleoside triphosphate as substrate and polymerises in template dependent manner ie following complementarity

RNA polymerase also somehow facilitates opening of the helix and continues elongation (NEET - 2020)

Only a short stretch of RNA remains bound to the enzyme. Once the polymerase reaches the terminator region, the nascent RNA falls off, so also the RNA polymerase resulting in termination of transcription (See Figure above)

Spiral

Pregnancy before birth

Due to absence of split gene arrangement

Date.................

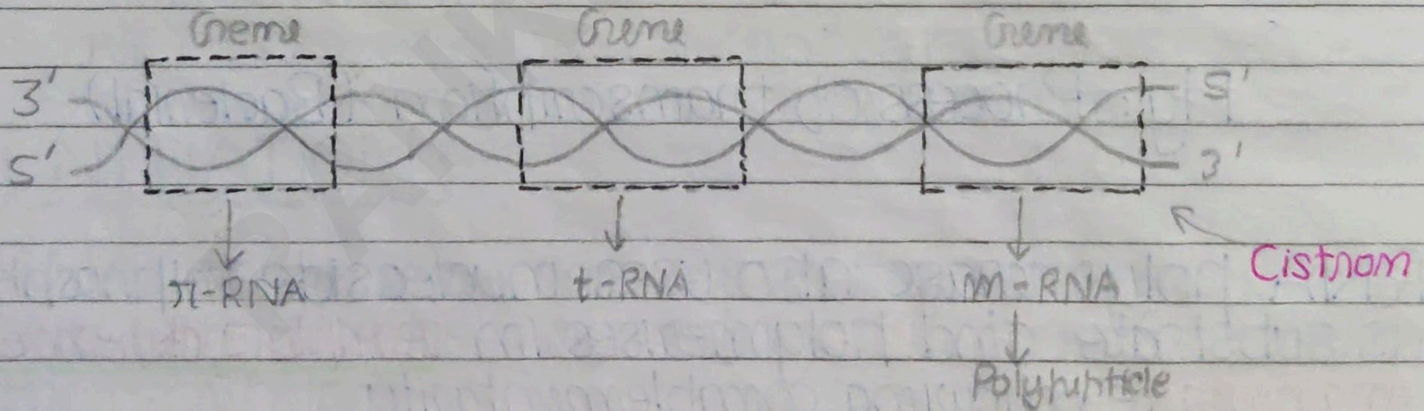
In bacteria, since the m-RNA does not require any processing to become active and also since transcription and translation take place in the same compartment (as there is no separation b/w cytosol & nucleus in bacteria). ∴ In bacteria, many a time translation begins much before m-RNA is fully transcribed. Consequently, the transcription & translation can be coupled in bacteria

CONCEPT OF GENE

Gene is a segment of DNA which codes for any type of RNA. Hence the term gene is used for

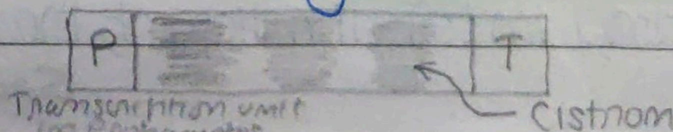
1. A part of DNA synthesising r-RNA
2. A part of DNA synthesising m-RNA
3. A part of DNA synthesising t-RNA

} All



Cistron is the gene that codes for polypeptides (protein). All cistrons are genes but All genes are NOT cistrons

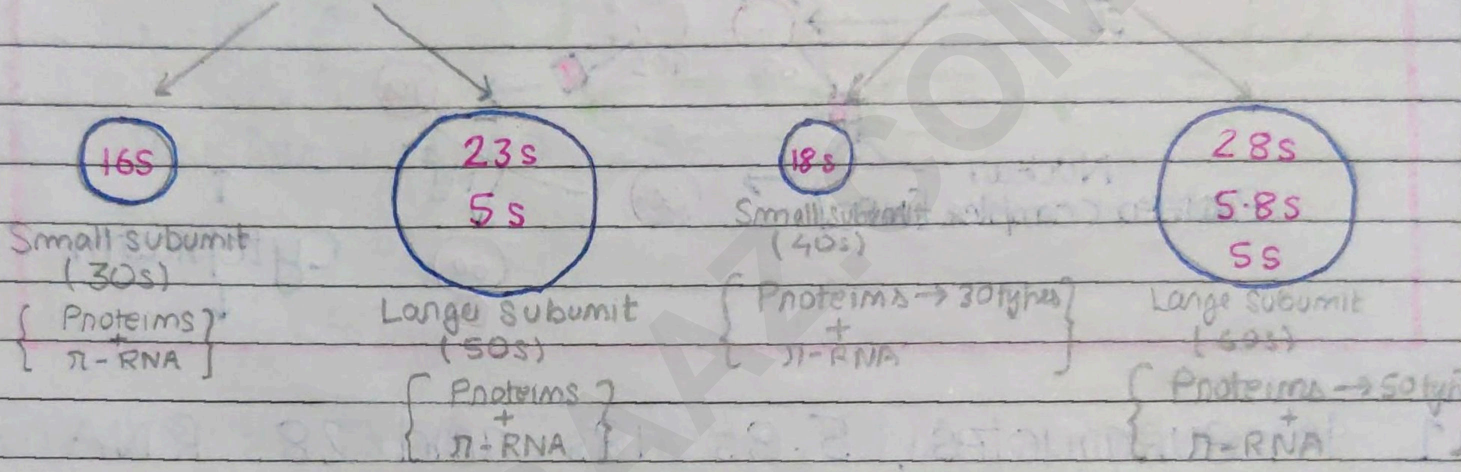
In prokaryotes, a single structural gene has many cistrons (Polycistronic) ∴ more than 1 type of protein is synthsised on same RNA



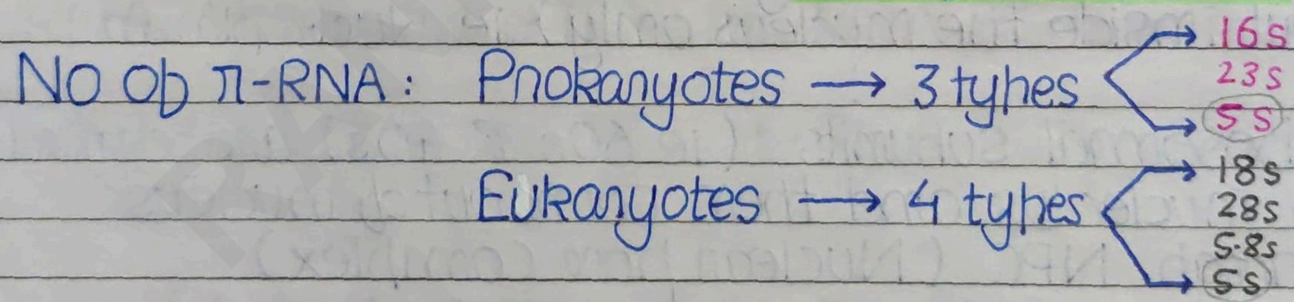
Date....................

In eukaryotes, a single cistron is present in transcription unit ∴ only ONE type of protein is synthesised on m-RNA ie Monocistronic

Assembly of ribosome → Mg^{2+}



Note: The ribosome (in eukaryotes) consists of structural RNA and about 80 different protein (NCERT)



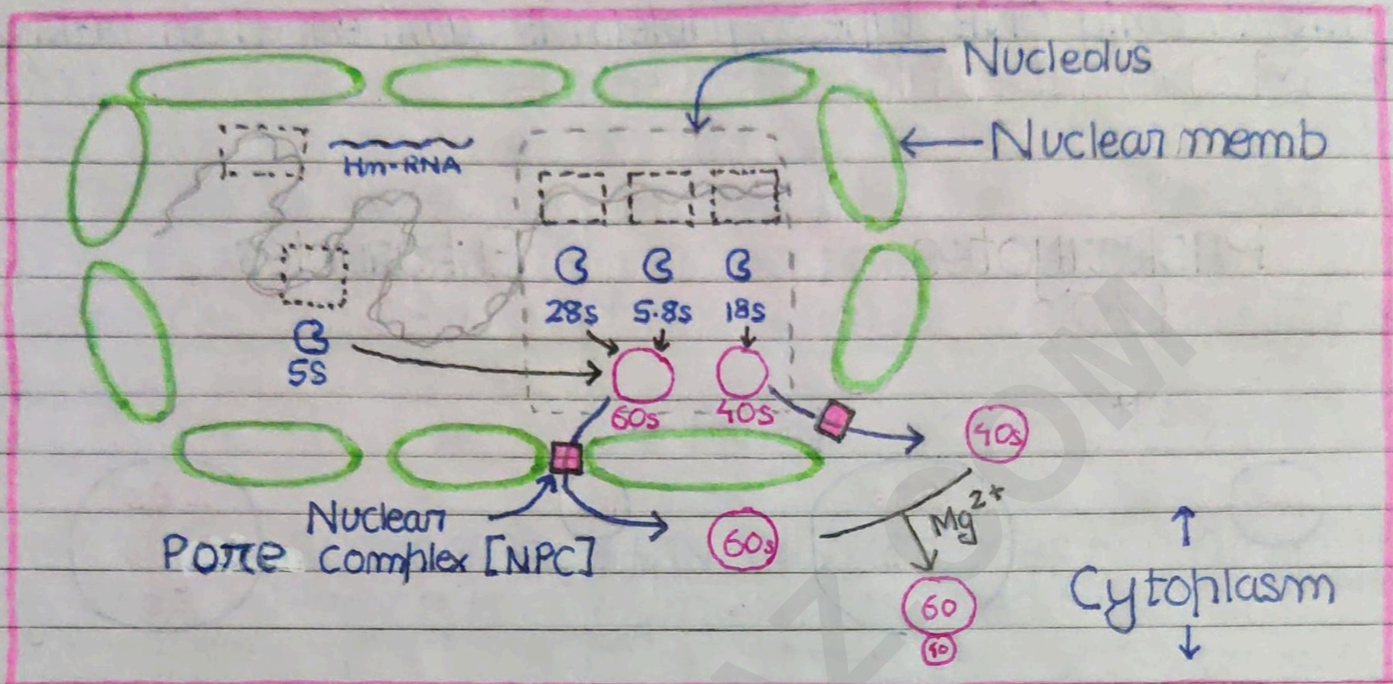
Common type of RNA in both eukaryote and prokaryote is 5s rRNA

16s-RNA : Basis of 6 kingdom classification

Spiral

Date.....................

TRANSCRIPTION IN EUKARYOTES



1. In eukaryotes, 5.8s 18s and 28s RNA are all synthesised in nucleolus

2. The 5s-RNA is synthesised outside nucleolus (but inside the nucleus only) ie Nucleoplasm

3. Ribosomal subunits (ie 60s & 40s) are formed in nucleolus and then come out of nucleus through NPC (Nuclear pore complex)

4. Ribosome subunits combine in the cytoplasm in presence of Mg^{2+} to form dimer

Note: In prokaryotes 23s, 16s, 5s r-RNA are all synthesised in cytoplasm.

Spiral

Sn RNA helps in splicing and it is produced during the transcription

16 s → RNA Polymerase

18 s → RNA Polymerase I

Date.....................

Types of RNA polymerases

(I) In Prokaryotes

Prokaryotes have a SINGLE RNA Polymerase is present which is able to synthesise

1. m-RNA
2. t-RNA
3. r-RNA (16s, 23s, 5s)

(II) In Eukaryotes

Eukaryotes have 3 types of RNA Polymerase

RNA Polymerase I : 5.8s 18s 28s

RNA Polymerase II : hn RNA —Processing→ m-RNA —comes out of nucleus→

RNA Polymerase III : t-RNA , 5s , Sc RNA , Sn

Hn-RNA is processed to m-RNA inside nucleus, after that, the processed m-RNA comes out of the nucleus.

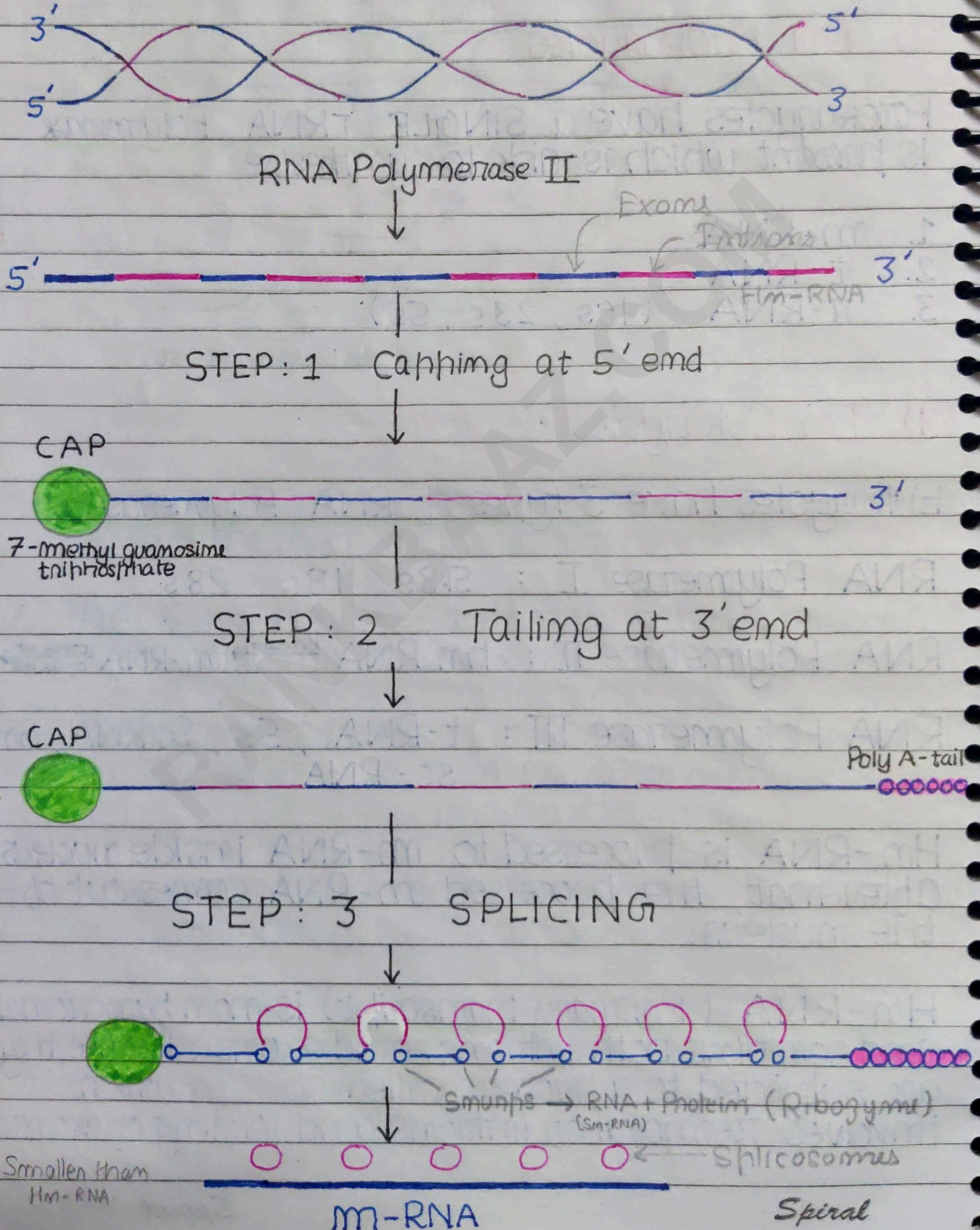
Hn-RNA (Primary transcript) is non functional and contains both introns and exons. Hence they are subjected to a process called splicing which involves removal of introns and joining of exons.

Spiral

NCERT : Presence of introns is reminiscent of antiquity and represents the dominance of RNA-world.
↑ Antique

Date.....................

Post transcriptional modification of Hn-RNA



Spiral

F. Crick → B-DNA model
Central dogma
t-RNA

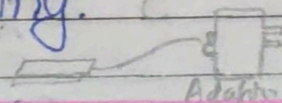
Splicing occurs in nucleus

Date......................

Capping: An unusual nucleotide (7-methyl guanosine triphosphate) is attached to 5'-end.

Tailing: Adenylate (ATP / adenylic acid) residues about 200-300 are added at 3'-end in a template independent manner

Splicing: Exons are coding sequences present in process m-RNA. Introns are non-coding sequences absent in processed m-RNA. Snups (small nucleoprotein) help in splicing.



t-RNA (transfer RNA / Soluble / Adapter RNA)

According to Francis Crick amino acid lack capability to read code on m-RNA

He postulated the presence of an adapter molecule that would on one hand read the code on m-RNA and with other hand bind to a specific amino acid

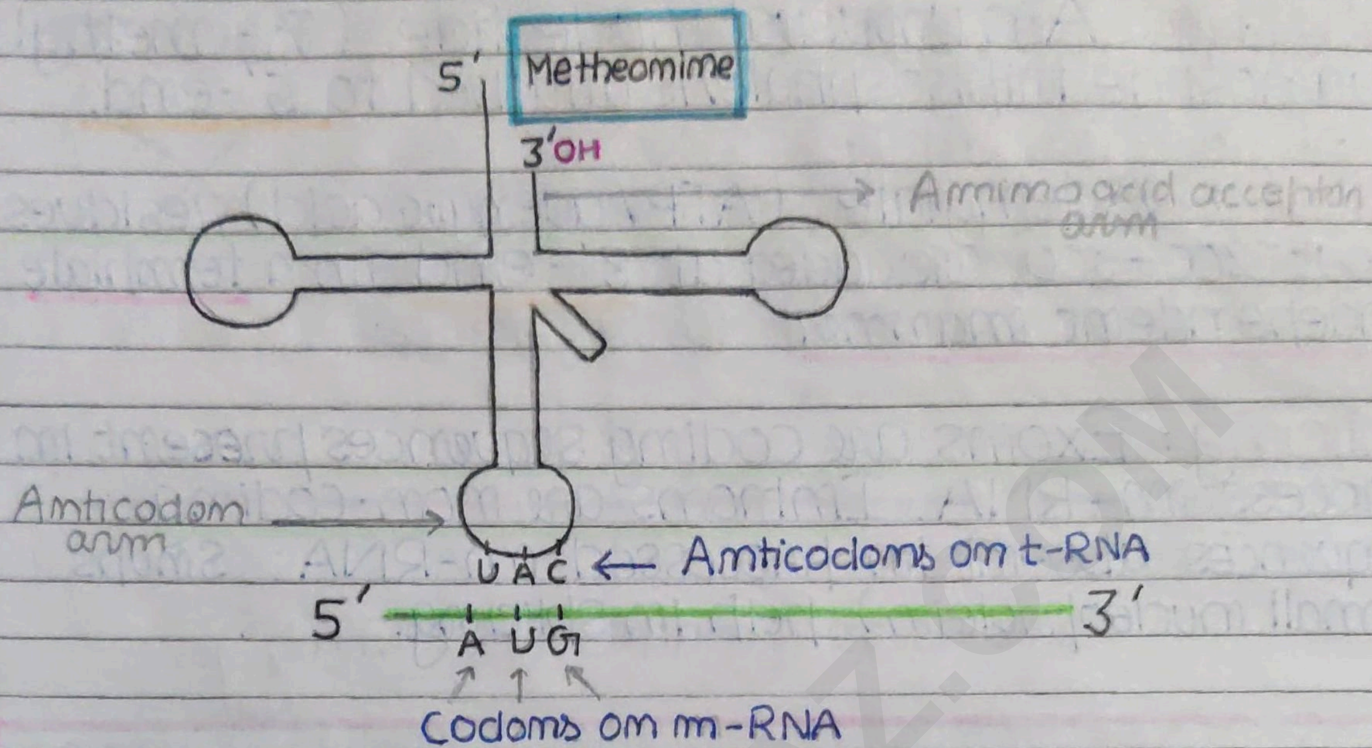
The t-RNA (then called S-RNA [Soluble]) was known before the genetic code was postulated. However its role as an adapter was assigned much later

The secondary structure of t-RNA has been depicted that looks like a Clover leaf shape

In actual structure, t-RNA is a compact molecule which looks like [illegible] L shape
invented

Spiral

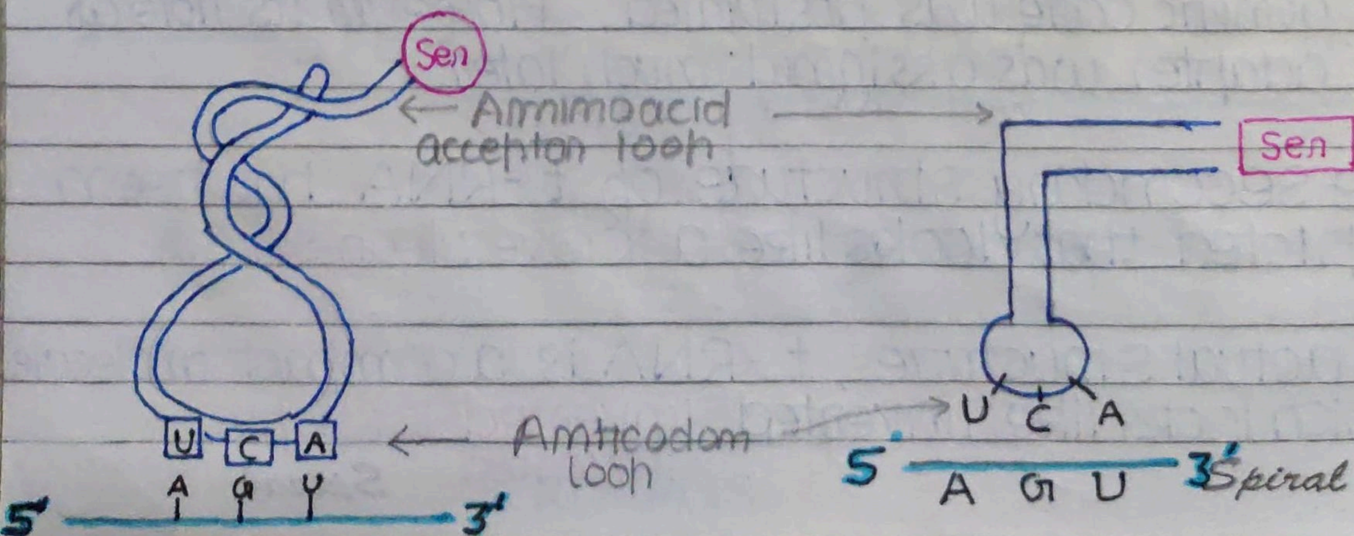
Date.....................



Amticodom arm : Recogmise code om m-RNA (has bases complimentary to codoms on m-RNA)
Ammimo acid accepton arm : Bimds to ammimo acid

t-RNA are specific fon each ammimo acid
Fon imitiatiom, there is a t-RNA specific to metheomime, knowm as imitiaton t-RNA having its amticodom bases complementary to the codom bases om m-RNA

NOTE : There are NO t-RNA fon stop codoms



Spiral

Concept of C-DNA

hn-RNA undergoes post-transcriptional modification to form m-RNA. When this m-RNA undergoes reverse transcription, it forms C-DNA which is smaller than original DNA due to absence of non coding region (introns)

In prokaryotes, it is difficult to express eukaryotic gene because the primary transcript requires modification which cannot happen in prokaryotes.

Hence we synthesise C-DNA from eukaryotic gene and then express this C-DNA into prokaryotes. It can form m-RNA directly.

GENETIC CODE

Both DNA replication and transcription were governed by rule of complementarity, hence it is easy to conceptualise & predict the process

The process of translation requires transfer of genetic information from a polymer of nucleotide (m-RNA) to a polymer of amino acids, neither does any complementarity exist between nucleotides and aa

This led to proposition of a genetic code that could direct the sequence of amino acids during synthesis of proteins.

Spiral

No of codons (64) >> No of amino acids (20)

Template independent → Poly A tail
Severo Ochoa

Date....................

It was George Gamow, a physicist who argued that since there are only 4 bases and if they have to code for 20 amino acids, the code should be made up of three nucleotides. This was a very bold proposition because a permutation combination of 4^3 ($4 \times 4 \times 4$) would generate 64 codons; many more than required

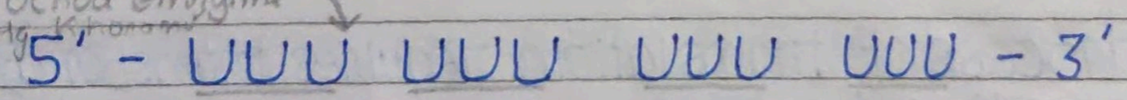
Other contributions

1. Har Gobind Khorana was able to synthesise homopolymers and copolymers
   5'– UUU UUU UUU – 3' Homopolymer
   5 – GUG UGU GUG – 3' Copolymer

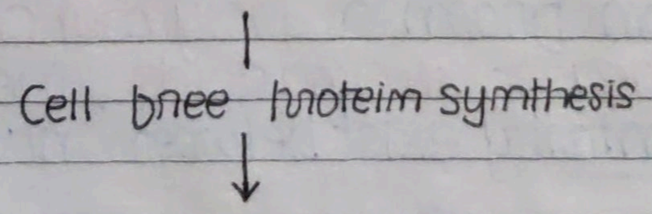
2. Marshall Nirenberg's cell free system for protein synthesis finally helped the code to be deciphered (bina cell ke protein!!)

3. Severo Ochoa gave severo ochoa enzyme (Polynucleotide phosphorylase) was also helpful in polymerising RNA with defined sequences template free (Template free RNA synthesis)
   ∴ It occurs in a template independent manner due to enzymatic synthesis of RNA.

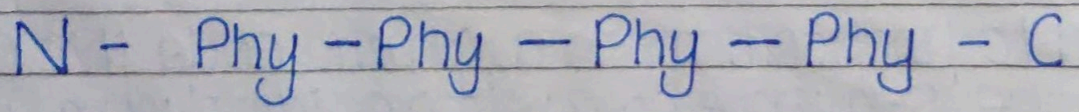
Sequence synthesized using severo ochoa enzyme by Hg. Khorana



5' – UUU UUU UUU UUU – 3'



(Agar bacterial cell ke andar karte to kaise pata chalta ki ye protein humne h ki cell ka h)



N – Phy – Phy – Phy – Phy – C

Hence, using this homopolymer, they could

Date.....................

decipher that UUU codon codes for Phenylalanine

Genetic code - information

1. UUU : First codon to be discovered
   Codes for phenyl alanine

2. AUG : Initiation codon (Non-degenerate)
   Codes for methionine

3. UGG : Non degenerate codon
   Codes for tryptophan

4. UGA UAG UAA : Stop codon
   Dont code for anything

Salient features of genetic code

There are 6 salient features of a genetic code

TRIPLET : 61 codons code for amino acids while remaining 3 dont code

NO PUCTUATIONS: The codon is read in m-RNA in a contigous fashion ie without any comma (s) or fullstops.

NEARLY UNIVERSAL : A codon would code for the same amino acid in all organism However some exceptions to this rule have been found in mitochondrial codons & in some protozoa

Spiral

Date..................

Genetic engineering is possible only due to the universal nature of the genetic code.

4. Degenerate : Some amino acids are coded by more than one codons. However AUG and UGG are non-degenerate

5. Non ambigous : One codon can code for only one amino acid (specificity) GUG (Valine) is ambigous codon

6. Non overlapping : Single nucleotide is not involved in formation of more than one codon

AUG CAU — Right

AUG CAU — Wrong

MUTATIONS AND GENETIC CODE

5'- GUG | AUG UGG AGG UAA | GAG - 3'
(UTR on both ends)

Replaced by U

5'- GUG | AUG UGU AGG UAA | GAG - 3'
(UTR on both ends)

In substitution type of gene mutation, only one amino acid is changed in the peptide chain. It can also produce silent mutation if codon is converted to another codon that codes for same amino acid.

Spiral

Sickle cell anemia (substitution mutation) Date.........

A classic example of point mutation is a change of single base pair in the gene for Beta-globin chain of haemoglobin that results in change of glutamic acid to valine (ie it affected only one amino acid in whole peptide chain)

Conclusion

1. Substitution of one nucleotide affects 1 codon and ∴ only one amino acid

2. Insertion or deletion of one or two bases change the reading frame from the point of insertion or deletion. However such mutations are called frameshift insertion or deletion mutations respectively. → Thallassemia

3. Insertion or deletion of 3 base adds or deletes one codon and ∴ an amino acid. The reading frame remains unaltered from that point onwards

TRANSLATION

Translation refers to process of polymerisation of amino acid to form a polypeptide.

Translation location
1. Prokaryotes, Eukaryotes } Cytoplasm
2. Organelles → Chloroplast, → Mitochondria (own DNA & 70s ribosome)

Spiral

All 3 RNA help in translation

m-RNA	t-RNA	r-RNA
→ Have information to give primary structure of Polypeptide	→ Reads code on m-RNA and brings aminoacid to ribosome	→ Helps in making Peptide bond

STEP I : Activation of amino acid

Amino acids are activated in the presence of ATP and linked to their cognate t-RNA to form charged t-RNA or aminoacyl t-RNA.

3'-OH HOOC – C – NH₂ (with H above and R below the C)

Ester bond

At 3'-end of t-RNA, COOH of amino acid is linked through ester bond.

Charging of t-RNA occurs in 2 steps;

1. $AA + ATP + E_1 \longrightarrow AA\text{-}AMP\text{-}E_1 + 2P_i$
   (AA-AMP-E₁ = complex)

   E_1 (enzyme) = Aminoacyl t-RNA synthatase

2. Activated AA binds on specific t-RNA then enzyme and AMP is released forming charged t-RNA

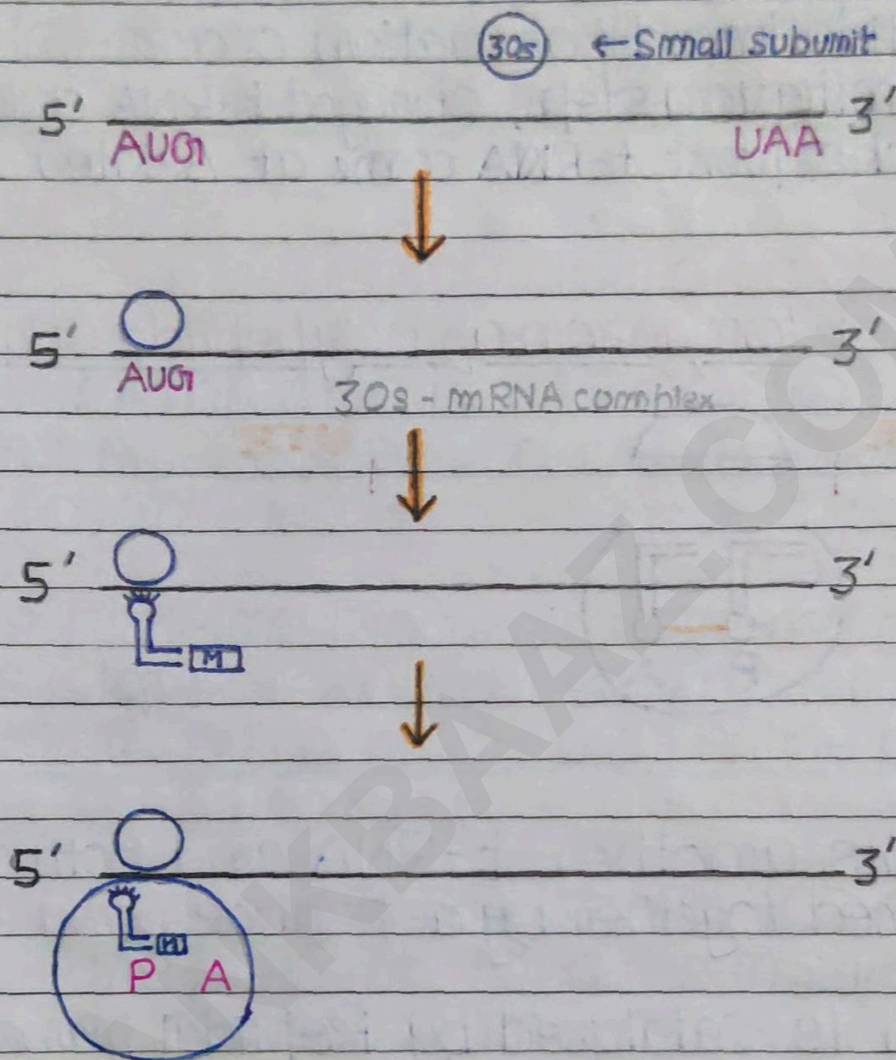
   $AA\text{-}AMP\text{-}E_1 + t\text{-}RNA \longrightarrow AA\text{-}tRNA + AMP + E_1$
   (AA-AMP-E₁ = complex; AA-tRNA = Charged t-RNA)

A large no of charged t-RNA are always present inside cytoplasm.

Spiral

Date.....................

STEP 2 : Imitiation ob polyhehtide chaim



Abter charging ob t-RNA, the Smaller subumit ob ribosome ie 30s (im Prokaryotes) bimds om m-RNA recogmised by Shimedelgnamo sequemce

Them 1st tRNA with imitiatiom ammimo acid (Metheomime im eukaryotes amd bormyl metheomime im Prokaryote) bimds om 30s m-RNA comhlex.

50s subumit ob RNA bimds to the above comhlex amd it has 2 sites
1. Psite
2. Asite

First t-RNA is limked to P site

Spiral

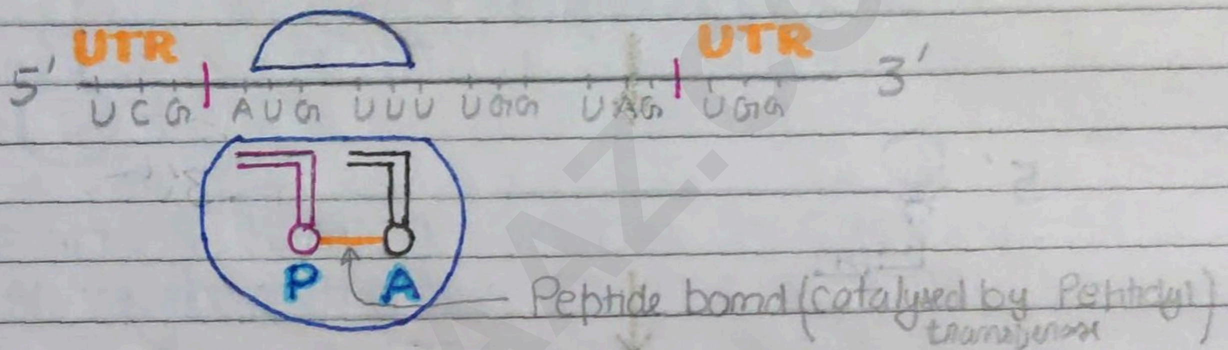
If 2 charged t-RNA are brought close enough, the formation of peptide bond between them would be favoured energetically !! The presence of catalyst (Peptidyl transferase) enhances the rate of peptide bond formation

Date.....................

During initiation of peptide chain

1. NO peptide bond formation occurs
2. Only in initiation step, charged t-RNA comes at P site. (subsequent t-RNA come at A-site)

STEP 3 : Elongation of polypeptide chain



A site accepts another t-RNA and both amino acids are joined together by a peptide bond

This reaction is catalysed by Peptidyl transferase (23s r-RNA present on 50s subunit). In eukaryotes it is 28s r-RNA present on 60s subunit

The UTRs (Untranslated regions) are present at both 5'end before start codon and at 3'-end after stop codon. They are required for efficient translation process

After formation of 1st peptide bond, ribosomes moves towards 5'→3' known as translocation and it occurs with the expenditure of GTP

Spiral

Date.................

NCERT : The ribosomes move from codon to codon along the m-RNA. Amino acids are added one by one, translated into protein (Polypeptide) sequences, dictated by DNA and represented by m-RNA

STEP 4 : Terminalisation of polypeptide chain

For termination to take place, stop codon must be present on A-site, no t-RNA would come for stop codon

RF

Some proteins known as release factors binds to the stop codon, terminating translation and releasing the complete polypeptide from the ribosome

Ques. A Polypeptide chain 100 amino acids are present, find the minimum number of nucleotide in m-RNA?

Ans. Amino acids = 100

1 amino acid is coded by 3 nucleotides
$\therefore$ total : $3 \times 100 = 300$ nucleotide

In addition to these 300 nucleotides, 3 more nucleotide (Stop codons) are also present which do not code for amino acid.

Total Nucleotides = 300 + 3
= 303 Nucleotide

Spiral

Regulation matlab protein banne dena ya nahi banne dena !!

Date.....................

REGULATION OF GENE EXP

Regulation of gene expression refers to a very broad term that may occur at various levels. In eukaryotes the regulation can be exerted at (NCERT)

(i) Transcriptional level (Formation of 1° transcript)
(ii) Processing level (Regulation of splicing) — Processing = Post transcriptional modification
(iii) Transport of mRNA from nucleus to cytoplasm.
(iv) Translational level.

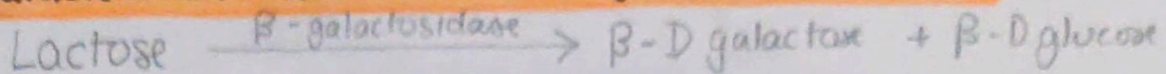
NOTE : In prokaryotes only regulation at transcriptional and translational level is possible

NCERT : Control of the rate of transcriptional initiation is the predominant site for control of gene expression.

Types of gene

1. Constitutive gene : Those genes which are constantly expressing themselves in a cell bcoz their products are required for normal cellular activities. Eg : genes for glycolysis (Hexokinase), ATPase gene, i gene in lac operon

2. Non constitutive (luxury) gene : These genes are not always expressing themselves in a cell. They are switched on or off according to the requirement of cellular activities. Eg : Z, Y and A gene in the lac operon.



Lactose $\xrightarrow{\beta\text{-galactosidase}}$ β-D galactose + β-D glucose

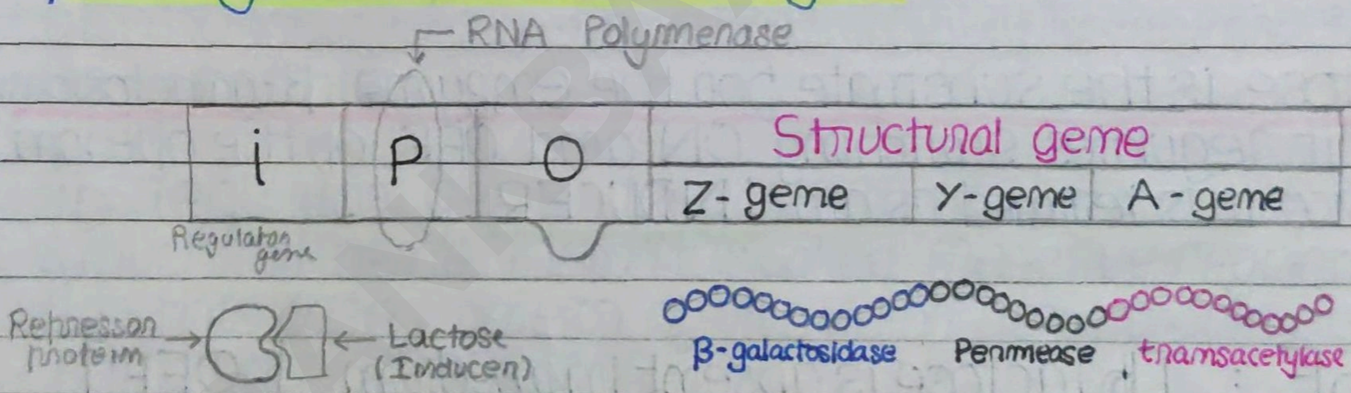
Date.....................

THE LAC OPERON MODEL

The elucidation of the lac operon was also a result of close association between a GENETICIST (Francois Jacob) and a BIOCHEMIST (Jacque Monod)

DEFINITION: When more than one structural genes (Polycistronic) have a common promoter and regulator gene due to which all structural genes express together. This arrangement is called an operon

Note: Since polycistronic condition is present only in prokaryotes ∴ Operons are studied only in prokaryotes and not eukaryotes.



The lac operon consists of one regulator gene ie i-gene (Here i does not refer to inducer, rather it is derived from the word inhibitor) i-gene codes for repressor of Lac operon

Also lac operon consists of 3 structural genes

1. Z gene : β-galactosidase hydrolyses lactose
2. Y gene : Permease increases permeability of β-gal to cell
3. A gene : Transacetylase helps in acetylation rxn

Spiral

Lac operon is inducible operon

Date....................

All 3 gene products in lac operon are required for metabolism of lactose, ie in the same metabolic pathway

NCERT : In other operons as well, the genes present in the operon are needed together to function in the same or related metabolic pathway.

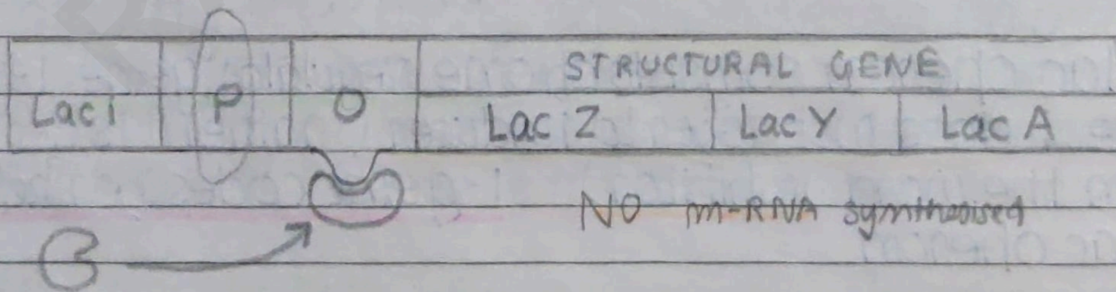
[NCERT]

In the absence of a PREFFERED carbon source ie GLUCOSE , if lactose is present in medium of the bacteria then lactose is transported into cells by the activity of permease (Remember, a low level of expression of lac operon has to be present in the cell at all time otherwise lactose cannot enter the cells)

↳ Agar lactose nai hoga to Permease nai banega, aur Permease nai banega to lactose andar enter kaise karega ? Therefore, small level operons always under expression!!

Lactose is the substrate for the enzyme β-galactosidase and it regulates switching ON and OFF of the operon. Hence it is termed as an INDUCER

CASE : If lactose is absent in medium [OFF]

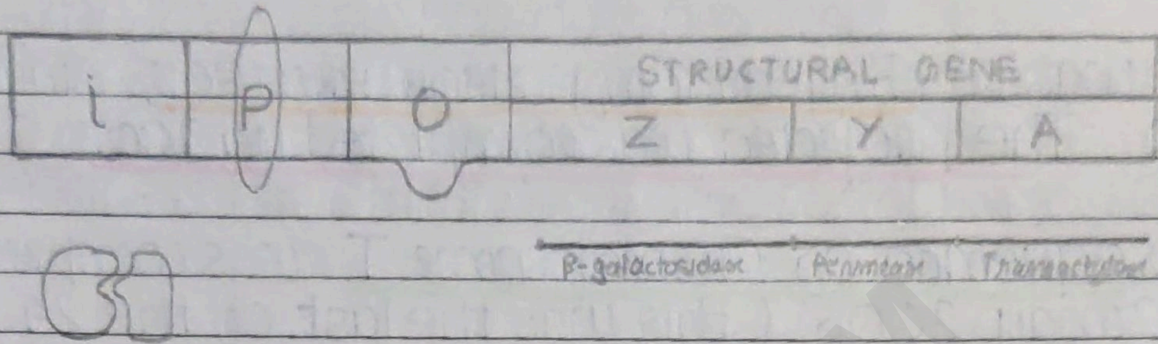


The repressor of the operon is synthesised constitutively from i gene . The repressor protein binds to the operator region of the operon and prevents RNA Polymerase from transcribing the operon

RNA Polymerase

Date.....................

CASE : If lactose is present in medium [ON]



In the presence of an inducer such as LACTOSE or ALLOLACTOSE, the repressor is inactivated by interaction with inducer. This allows the RNA polymerase to access to the promoter and transcription proceeds

NCERT : Essentially, regulation of lac operon can also be visualised as regulation of enzyme synthesis by its substrate.

If regulator gene decreases the expression of the structural gene then it is called negative regulation. In this case protein of regulator gene would be called as REPRESSOR PROTEIN

Regulation of lac operon by repressor is referred to as NEGATIVE Regulation

Lac operon is under control of POSITIVE regulation as well, but it is beyond the scope of discussion at this level.

Bioinformatics: Branch of biology that deals with biological data storage, retrieval & use

Human genome (haploid set) was sequenced

Date.....................

HUMAN GENOME PROJECT

HGP was a 13 year long MEGA PROJECT, started in 1990. The projected was completed in 2003

The sequence of Chromosome I was completed only in may 2006 (this was the last of the 24 human chromosomes to be sequenced)

The project was coordinated by

1. U.S. Department of Energy
2. National institute of health, USA
3. Wellcome trust, UK
4. Additional contributions from → Japan, France, Germany, China

Human genome is said to have 3×10^9 bp and if the cost of sequencing required is US \$ 3 / bp, the total estimated cost of the project would be approx 9 Billion US dollars

HGP was closely associated with the rapid development in the area of BIOINFORMATICS due to enormous amount of data expected to be generated.

Goals of HGP

1. Identify all the approximately 20000 - 25000 genes in human DNA
2. To determine the sequences of 3 billion chemical base pairs that make up human DNA

Spiral

Date..................

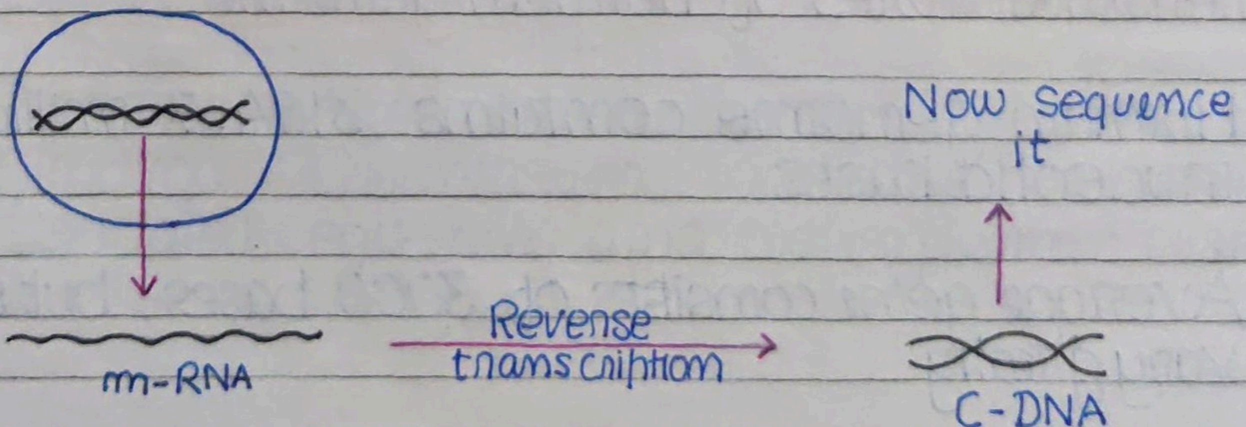
(3) Store this information in database
(4) Improve tools for data analysis
(5) Transfer related technologies to related sectors, such as industries
(6) Address the ethical, legal and social issues (ELSI)

Many non-human model organisms are also sequenced
→ Bacteria
→ Yeast
→ Caenorhabditis elegans (Free living Non Pathogenic nematode)
→ Drosophila
→ Plants : Rice & Arabdopsis

Methodologies: The method involved in HGP involved 2 major approaches :

1. Expressed sequence tags (ESTs)

It focuses on identifying all the genes that are expressed as RNA

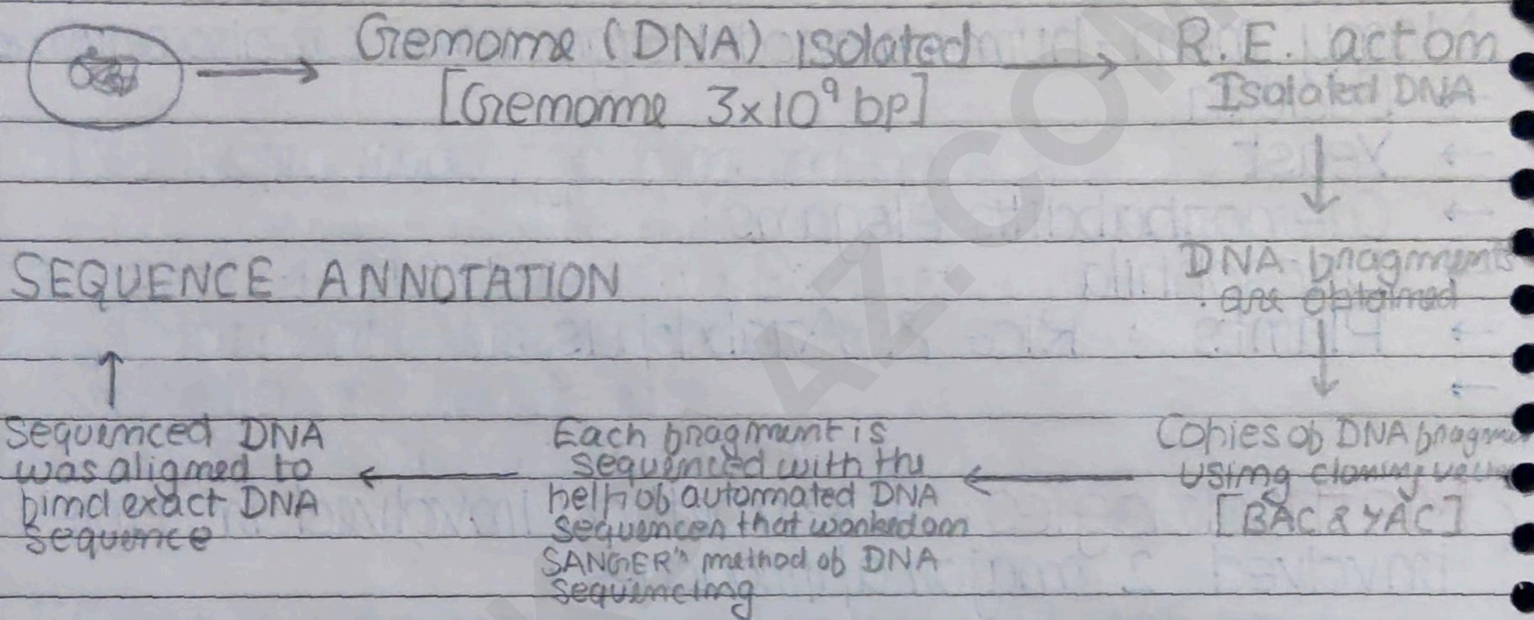


Spiral

Date......................

2. Sequence annotation

Its a blind approach of simply sequencing the whole set of genome that contained all the coding and non-coding sequence, and later assigning different regions in the sequence with functions.



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

Salient features of human genome

a) Human genome contains 3164.7 million nucleotide bases

b) Average gene consists of 3000 bases; but sizes vary greatly

Largest gene : Dystrophin ⟶ 2.4 million bases
Smallest gene : TDF ⟶ 14 bases

Spiral

Date....................

c) Total number of genes are estimated to be 30,000 much lower than previous estimates of 80,000 - 1,40,000 (Almost 99.9% N. Base are exactly same in all people)

d) The functions are unknown for over 50% of the discovered genes

e) Less than 2% of the genome codes for protein

f) Repeated sequences make up very large portion of the human genome

5' — ATT ATT ATT ATT ...... 3'

g) Repeated sequences are stretches of DNA sequences that are repeated many times, sometimes hundred to thousand times. They are said to have no direct coding function

h) Chromosome I has maximum genes (2968) and Chromosome Y has fewest (231)

i) Scientists have identified about 1.4 M locations where single base DNA differs (Single nucleotide polymorphism - SNPs) in humans.

j) These sequences revolutionarise the process of finding 'Chromosomal locations' for disease associated sequences and tracing human history.

Spiral

DNA fingerprinting was discovered by Alec Jeffreys

Polymorphism in coding sequence Polymorphism in Non-coding sequence

Date.....................

DNA FINGERPRINTING

As already stated 99.9 % of base sequence among humans is same ∴ No of bases that differ in humans are 0.1 % of 3×10^9 bp

No of bases that differ $= \frac{0.1}{100} \times 3 \times 10^9 = 3 \times 10^6$ bp

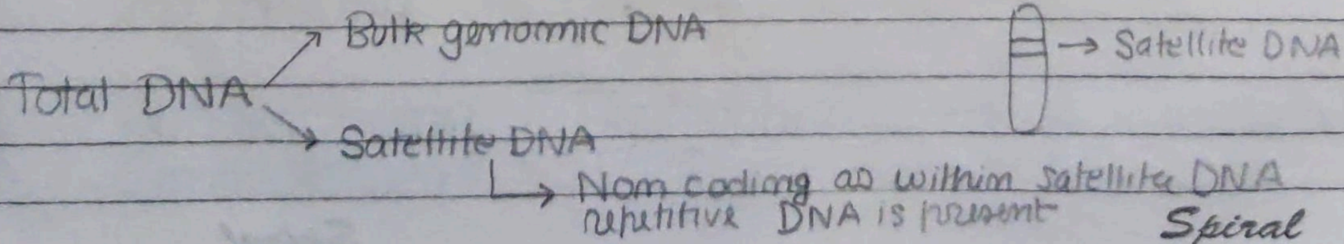
In DNA fingerprinting, no genome sequencing is done as cost of doing so would be too high @ 3 USD/bp to sequence 3×10^6 bp. Basis of DNA fingerprinting is Polymorphism (Variation at genetic level) which arises due to mutation.

Satellite DNA

Satellite DNA are mainly repetitive sequence hota hai

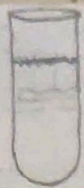
These are specific non-coding region on chromosome and are essentially repetitive DNA showing high degree of polymorphism and forms the basis of DNA fingerprinting.

Since DNA from every tissue shows the same degree of polymorphism, they become very useful identification tool in forensic applications. Further as the polymorphisms (DNA) are inheritable from parents to children, DNA fingerprinting is the basis of paternity testing, in case of disputes.



Spiral

CsCl density gradient method



← Bulk genomic DNA (Higher Peak)
] Smaller Peaks → Satellite DNA

Date.....................

These repetitive DNA are separated from bulk genomic DNA as different peaks during CsCl density gradient

The bulk DNA forms a major peak and the other smaller peaks are referred to as satellite DNA

Depending upon
1. Base composition (AT rich or GC rich)
2. Length of segment
3. Number of repetitive units

The satellite is classified into many categories

1. Minisatellite : Used in DNA finger Printing VNTR
2. Microsatellite : SSR (Single sequence repeats) or STR (Short tandem repeats)

(Mutation in coding region)

NCERT: Allelic sequence variations (Variation in coding regions) has traditionally been recognised as DNA polymorphism if its frequency is greater than 0.01. In simple terms, if an "Inheritable mutation" is observed in a population at high frequency it is referred to as DNA Polymorphism.

NCERT: The probability of such variations to be observed in non-coding DNA would be higher as mutation in these sequences may NOT have any immediate effect in an individuals reproductive capacity. These mutations keep on accumulation generation after generation, and form one of the Basis of variability / Polymorphism

VNTR → 0.1 - 20 Kb
SSR → 1 to 6 bp

The technique used in DNA fingerprinting is RFLP
(Restriction fragment length polymorphism)

Date.....................

The technique of DNA fingerprinting was initially developed by Alec Jeffreys. He used a satellite DNA as probe that shows very high degree of polymorphism. It was called VARIABLE NUMBER OF TANDEM REPEATS ie VNTR

The technique involved Southern blotting hybridisation using radiolabelled VNTR as a probe.

VNTR belongs to a class of satellite DNA referred to as mini-satellite. A small DNA sequence is arranged tandemly in many copy numbers. The number of repeats show very high degree of polymorphism. As a result the size of VNTR varies in size from 0.1 to 20 Kb. Consequently, after hybridisation with VNTR probe the autoradiogram give many bands of different sizes. These bands give characteristic pattern for an individual DNA. It differs from individual to individual except Monozygotic twin

Sensitivity of this technique has been increased by use of PCR.

Spiral