

Table of Contents

1. Introduction
2. The DNA
3. Structure of Polynucleotide Chain
4. Structure of DNA
5. Central Dogma
6. Packaging of DNA Helix
7. DNA Packaging in Prokaryotes
8. DNA Packaging in Eukaryotes
9. Transforming Principle
10. Biochemical Characterisation of Transforming Principle (Avery Experiment)
11. DNA is the Genetic Material (Hershey Chase Experiment)
12. Properties of Genetic Material (DNA versus RNA)
13. RNA World
14. Replication
15. The Experimental Proof (Meselson and Stahl Experiment)
16. The Machinery and the Enzymes
17. Transcription
18. Types of RNA and Process of Transcription
19. Transcription in Eukaryotes
20. Genetic Code
21. Mutations and Genetic Code
22. tRNA – the Adapter Molecule
23. Translation
24. Protein Biosynthesis
25. Polyribosomes
26. Regulation of Gene Expression
27. The Lac Operon
28. Human Genome Project
29. Methodologies of HGP
30. Salient Features of HGP
31. Applications of HGP
32. DNA Fingerprinting
33. Southern Blotting
34. DNA Sequencing by Sanger's Method

Introduction

Overview

- At Mendel's time, inheritance factors were unknown.
- **DNA** and **RNA** are nucleic acids made of nucleotides.
- DNA is genetic material in most organisms.
- RNA acts as genetic material in some viruses and functions mainly as a **messenger**, adapter, structural, or **catalytic molecule**.

Memory Aids

- DNA = Genetic blueprint in most life forms.
- RNA = Messenger and catalyst in cells.

Quick Checkpoints

- What are the two main nucleic acids?
- Which nucleic acid acts as genetic material in most organisms?

The DNA

Structure and Composition

- DNA is a long polymer of deoxyribonucleotides forming two **complementary strands** running antiparallel.
- Strands held by hydrogen bonds between opposite nitrogen bases.
- Four types of nucleotides form DNA.
- DNA length is measured by number of nucleotides or base pairs (bp).
- Examples: Φ x174 (5386 nucleotides), lambda phage (48502 bp), *Escherichia coli* (4.6 million bp), human haploid DNA (3.3 billion bp).

Memory Aids

- DNA strands run antiparallel and complement each other.
- Base pairs define DNA length and organism specificity.

Quick Checkpoints

- What is the orientation of DNA strands?
- How is DNA length defined?

Structure of Polynucleotide Chain

Components and Linkages

- Nucleotide = nitrogenous base + pentose sugar + phosphate group.
- Purines: adenine, guanine; Pyrimidines: cytosine, uracil (RNA), thymine (DNA).
- Cytosine common to DNA and RNA; uracil replaces thymine in RNA.
- Base linked to 1'C sugar via **N-glycosidic linkage** forming nucleoside.
- Phosphate linked to 5'-OH of nucleoside via phosphodiester bond forms nucleotide.
- Nucleotides linked 3'-5' phosphodiester form dinucleotides and polynucleotide chains.
- RNA nucleotides have extra -OH at 2' position; DNA identified by Friedrich Miescher in 1869 as 'Nuclein'.

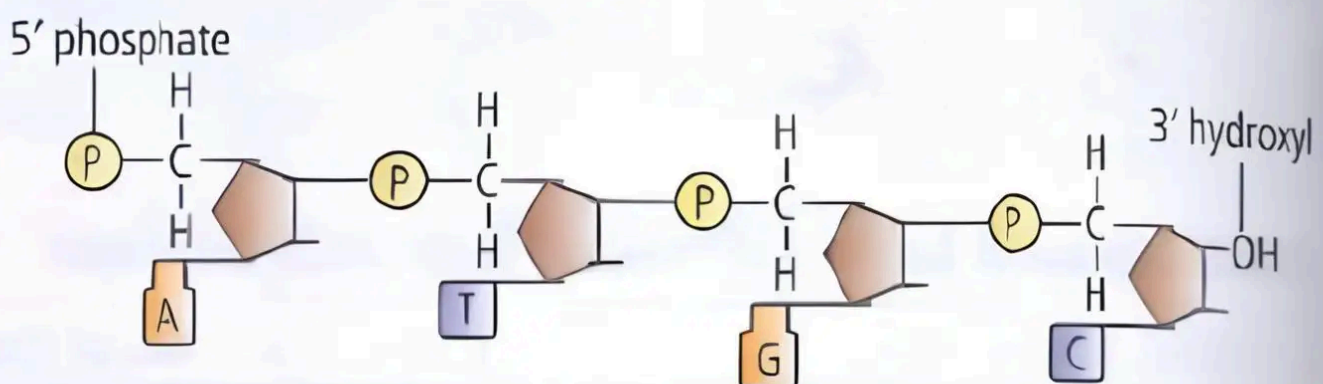


Fig.: A polynucleotide chain

Memory Aids

- Nucleotide = Base + Sugar + Phosphate.
- RNA has uracil; DNA has thymine.
- Phosphodiester bonds link nucleotides 3' to 5'.

Quick Checkpoints

- What are the three parts of a nucleotide?
- Which base replaces thymine in RNA?
- What type of bond links nucleotides?

Structure of DNA

Double Helix Model

- Proposed by James Watson and Francis Crick in 1953, based on X-ray data by Wilkins and Franklin.
- Chargaff's rules: purines = pyrimidines ($A + G = T + C$), $A = T$, $G = C$.
- Base ratio $A + T / G + C$ varies by species but constant within species.
- Each DNA strand acts as template for identical daughter DNA.
- Features:
 - Two polynucleotide chains with sugar-phosphate backbone; bases inside.
 - Chains are antiparallel ($5' \rightarrow 3'$ and $3' \rightarrow 5'$).
 - Bases paired by hydrogen bonds: A-T (2 bonds), G-C (3 bonds); purine pairs with pyrimidine.
 - Right-handed double helix with 3.4 nm pitch and ~10 bp per turn.
 - Base pairs stack for stability.

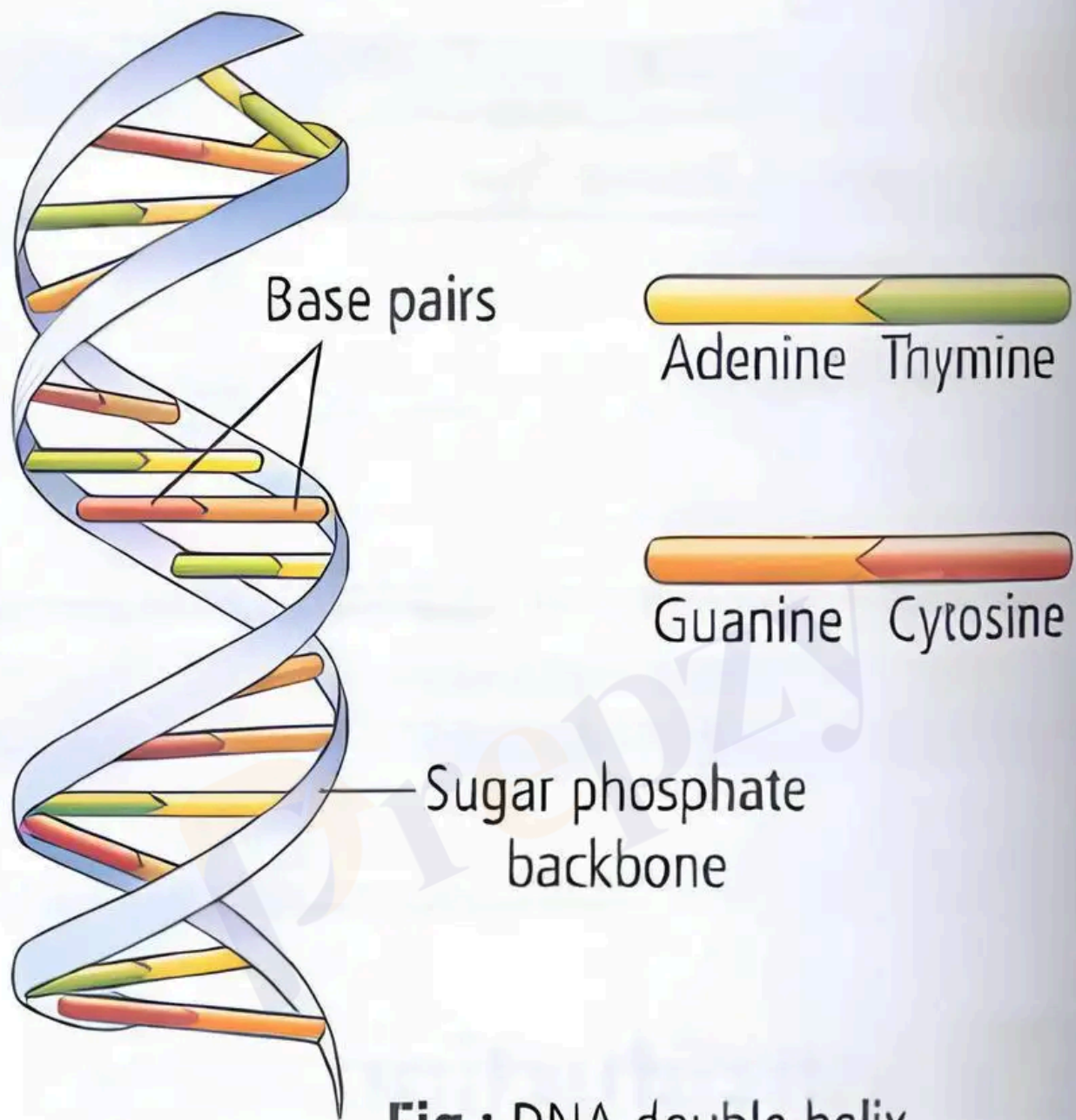


Fig.: DNA double helix

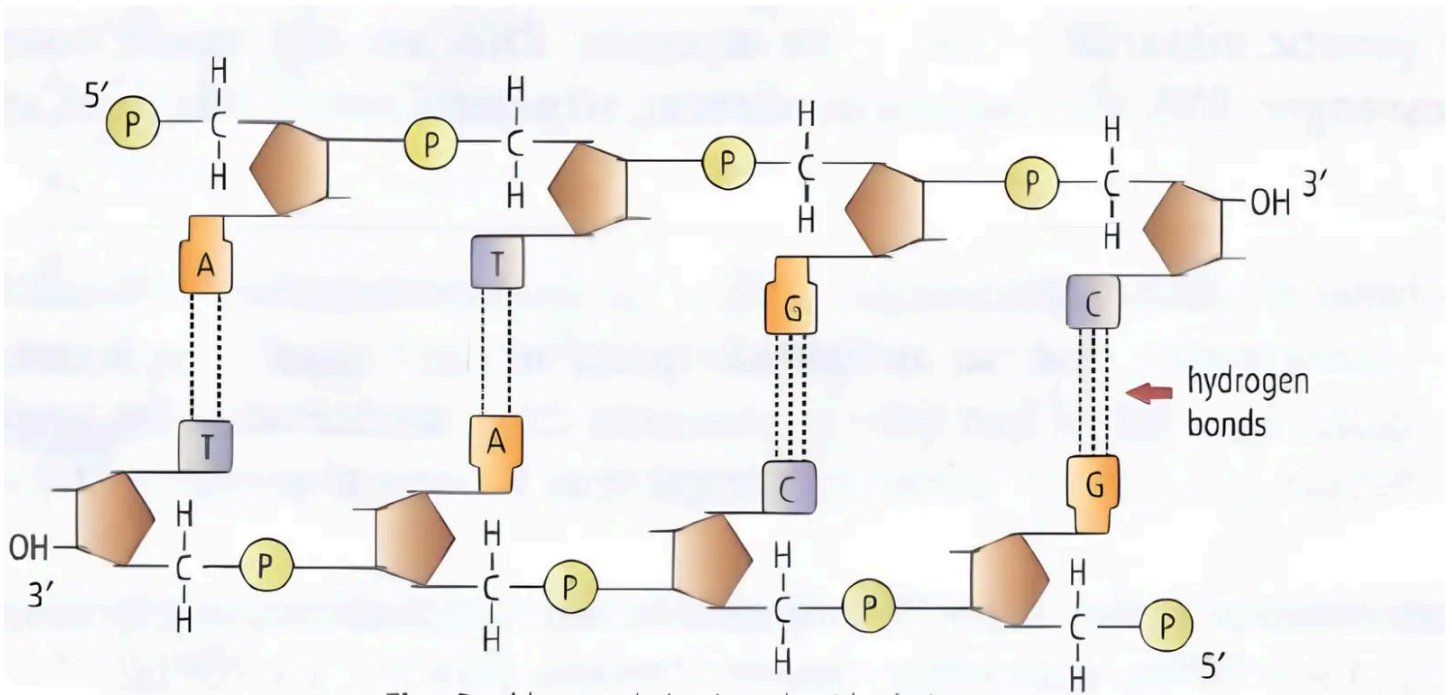


Fig.: Double stranded polynucleotide chain

Memory Aids

- DNA strands are antiparallel and complementary.
- A pairs with T; G pairs with C.
- Double helix is right-handed with 10 bp per turn.

Quick Checkpoints

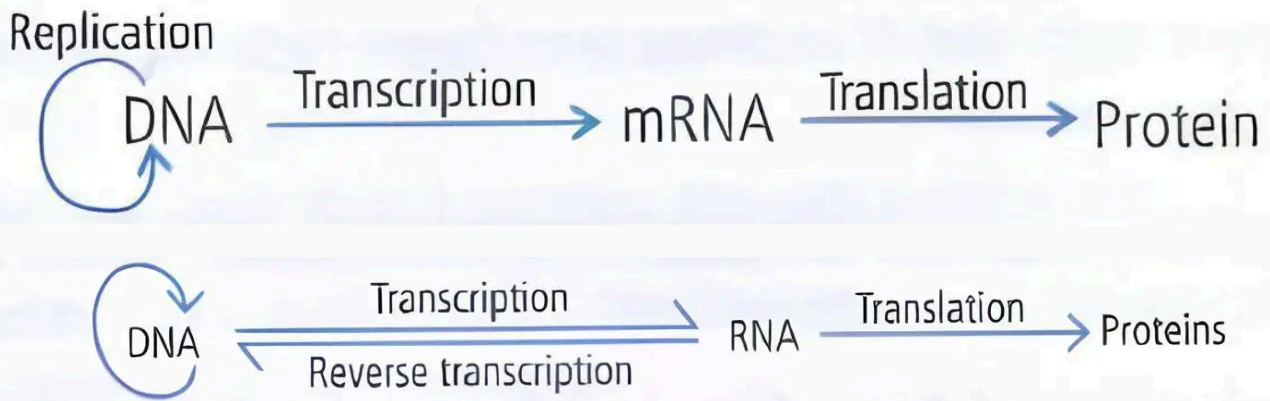
- Who proposed the DNA double helix model?
- What are Chargaff's base pairing rules?
- Describe the polarity of DNA strands.

Central Dogma

Genetic Information Flow

- Proposed by Francis Crick.
- Genetic information flows: DNA → RNA → Protein.
- Steps: DNA replication, transcription (DNA to mRNA), translation (mRNA to protein).

- In some viruses, reverse flow occurs: RNA → DNA (reverse transcription).



Memory Aids

- DNA → RNA → Protein is the central dogma.
- Reverse transcription occurs in some viruses.

Quick Checkpoints

- What is the direction of genetic information flow?
- What is reverse transcription?

Packaging of DNA Helix

DNA Length and Compaction

- DNA length = number of base pairs × distance between base pairs.
- Typical mammalian cell DNA length ≈ 2.2 meters.
- DNA is compacted to fit in small nucleus.

Memory Aids

- DNA is very long but tightly packed.

Quick Checkpoints

- How is DNA length calculated?
- Why is DNA packaging necessary?

DNA Packaging in Prokaryotes

Supercoiling and Nucleoid

- DNA is supercoiled with help of RNAs and non-histone proteins like polyamines.
- Compacted DNA mass is called nucleoid.

Memory Aids

- Prokaryotic DNA is supercoiled and forms nucleoid.

Quick Checkpoints

- What is nucleoid?
- How is prokaryotic DNA compacted?

DNA Packaging in Eukaryotes

Histones and Nucleosomes

- DNA wraps around histones (lysine and arginine rich proteins).
- Histones form octamer (eight molecules) called histone octamer.
- DNA + histone octamer = nucleosome.
- Five histone types: H₁, H₂A, H₂B, H₃, H₄; H₂A, H₂B, H₃, H₄ form octamer.
- Typical nucleosome contains ~200 bp DNA.
- Nucleosomes form chromatin, visible as thread-like bodies in nucleus.
- Chromatin appears as 'beads-on-string' under electron microscope.

- Chromatin fibers coil and condense into chromosomes during metaphase.
- Non-histone chromosomal proteins (NHC) assist higher-level packaging.
- Chromatin types: euchromatin (loosely packed, active), heterochromatin (densely packed, inactive).
- DNA between nucleosomes is linker DNA or interbead DNA, bearing H₁ histone (plugging and marker protein).
- Nucleosome + linker DNA = chromatosome.

Prepzy

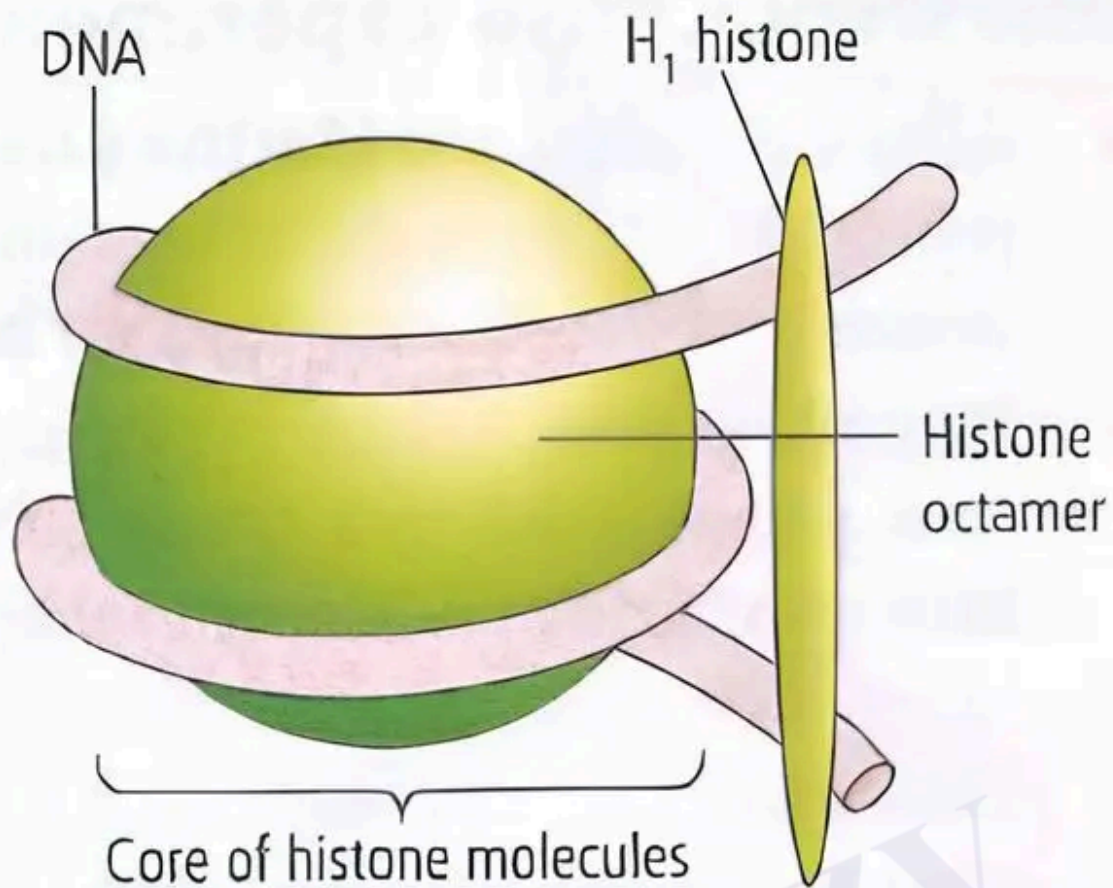


Fig.: Nucleosome



Fig.: EM picture - 'Beads-on-String'

Memory Aids

- Histones package DNA into nucleosomes.
- Chromatin = nucleosomes + linker DNA.
- Euchromatin = active; heterochromatin = inactive.

Quick Checkpoints

- What proteins help package eukaryotic DNA?
- Define nucleosome and chromosome.
- Difference between euchromatin and heterochromatin?

Transforming Principle

Key Findings

- 1928: Frederick Griffith studied *Streptococcus pneumoniae* with two strains: S-type (virulent, smooth colonies) and R-type (non-virulent, rough colonies).
- Heat-killed S strain alone did not kill mice; live R strain alone did not kill mice.
- Mix of heat-killed S strain and live R strain killed mice, showing transformation of R to virulent S type.
- Conclusion: A 'transforming principle' from dead S strain transferred genetic material to R strain, enabling virulence.

Memory Aids

- Griffith's experiment: Dead S + Live R = Dead mice (transformation).
- Transforming principle = genetic material transfer.

Quick Checkpoints

- What are the two strains of *Streptococcus pneumoniae*?

- What did the heat-killed S strain demonstrate?
- What is the significance of the transforming principle?

Biochemical Characterisation of Transforming Principle (Avery Experiment)

Key Findings

- Oswald Avery, Colin MacLeod, and Maclyn McCarty (1933-44) purified proteins, DNA, and RNA from heat-killed S cells.
- Only DNA caused transformation of live R cells into S cells.
- Protease and RNase treatments did not stop transformation; DNase treatment did.
- Conclusion: DNA is the hereditary material, not protein or RNA.

Memory Aids

- DNase stops transformation → DNA is genetic material.
- Protease and RNase do not affect transformation.

Quick Checkpoints

- Which enzyme inhibited transformation?
- What was the role of protease and RNase?
- What did Avery's experiment prove about genetic material?

DNA is the Genetic Material (Hershey Chase Experiment)

Key Findings

- Alfred Hershey and Martha Chase (1952) used bacteriophages to prove DNA is genetic material.
- Phages labeled with radioactive phosphorus (^{32}P) in DNA and radioactive sulfur (^{35}S) in protein.

- Only radioactive DNA entered *Escherichia coli* cells, not protein.
- Conclusion: DNA carries genetic information, not protein.

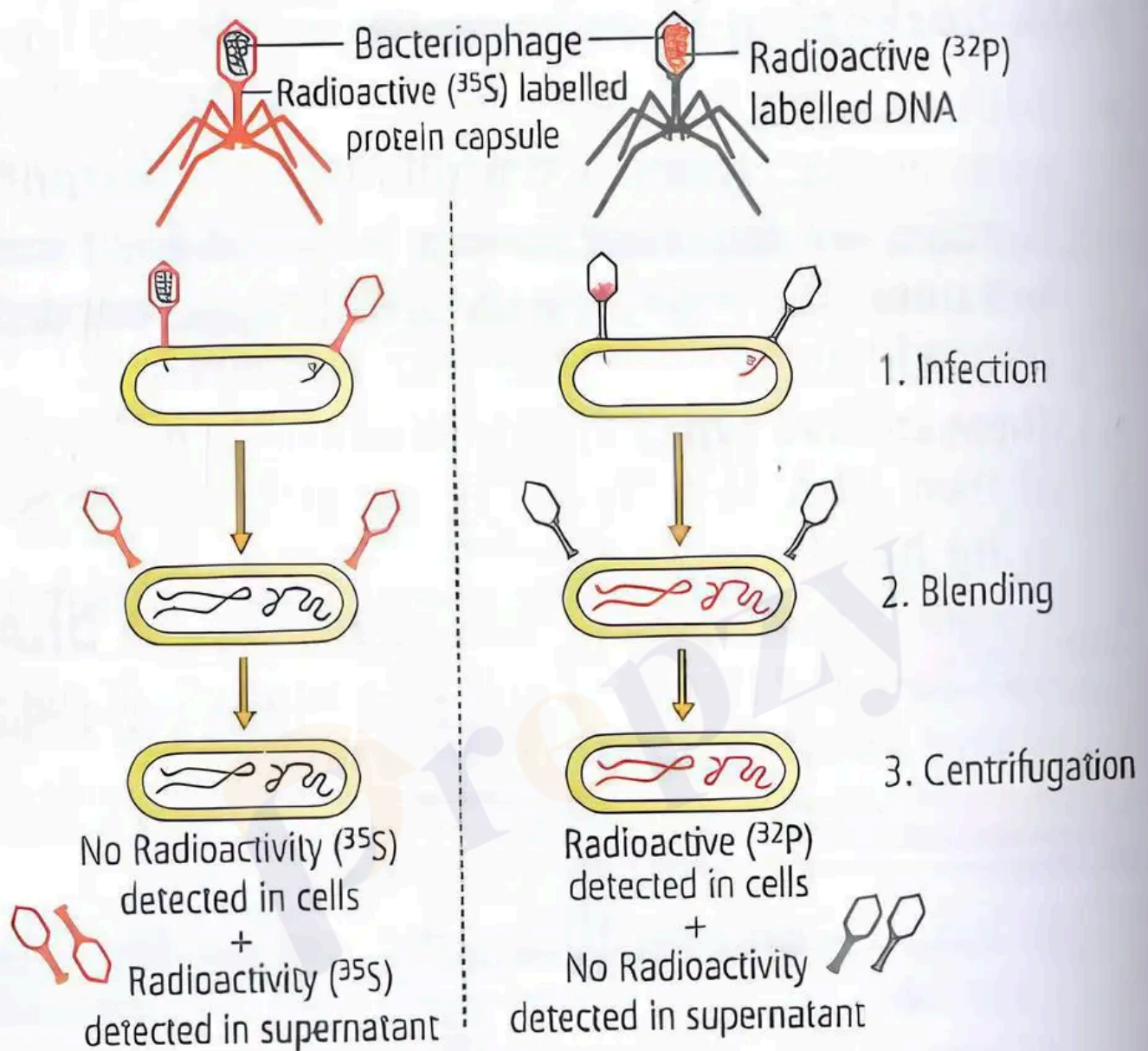


Fig.: The Hershey-Chase experiment

Memory Aids

- ³²P labels DNA; ³⁵S labels protein.
- DNA enters bacteria; protein does not.
- DNA = genetic material.

Quick Checkpoints

- What isotopes were used to label DNA and protein?
- Which component entered bacterial cells?
- What was the conclusion of the Hershey-Chase experiment?

Properties of Genetic Material (DNA versus RNA)

Key Properties

- Some viruses use RNA as genetic material (e.g., Tobacco Mosaic virus, QB bacteriophage).
- Genetic material must: replicate, be chemically and structurally stable, allow mutation, and express Mendelian characters.
- DNA and RNA can replicate due to base pairing; proteins cannot.
- DNA is chemically more stable than RNA due to absence of reactive 2'-OH group and presence of thymine instead of uracil.
- RNA is reactive, catalytic, and degrades easily.
- RNA mutates faster; RNA viruses evolve faster.
- RNA directly codes proteins; DNA depends on RNA for protein synthesis.
- DNA is preferred for genetic information storage; RNA for transmission.

Memory Aids

- DNA = stable, stores info; RNA = reactive, transmits info.
- Thymine in DNA adds stability; uracil in RNA less stable.
- RNA mutates faster → faster evolution.

Quick Checkpoints

- What are the criteria for genetic material?
- Why is DNA more stable than RNA?
- How do RNA and DNA differ in mutation rates and roles?

RNA World

RNA as First Genetic Material

- RNA was the first genetic material supporting essential life processes like metabolism, translation, and splicing.
- RNA acted both as genetic material and catalyst but was chemically reactive and unstable.
- **DNA evolved from RNA** with chemical changes making it more stable.
- DNA is double-stranded with complementary strands and repair mechanisms to resist changes.

Types of Non-Genetic RNA

- Based on size and function, three main non-genetic RNAs: mRNA (~3.5%), tRNA (~15%), rRNA (~80%).
- *mRNA*: Transcribed from genes, serves as template for protein synthesis by ribosomes.
- *tRNA*: Also called soluble, acceptor, or adapter RNA; has amino acid binding site opposite anticodon.
- *rRNA*: Structural component of ribosomes, synthesized in nucleolus.

Memory Aids

- RNA = First genetic material and catalyst.
- DNA = Stable, double-stranded, repair-capable.
- mRNA = Messenger, carries code.
- tRNA = Transfer, carries amino acids.
- rRNA = Ribosome, protein factory.

Quick Checkpoints

- What roles did RNA play in early life?
- How is DNA chemically more stable than RNA?
- What are the main types of non-genetic RNA and their functions?
- Where is rRNA synthesized?

Replication

Process and Mechanism

- DNA replication is the process of making an identical copy of double-stranded DNA using the original DNA as a template.
- Watson and Crick proposed semiconservative replication: each new DNA molecule has one parental and one newly synthesized strand.
- Replication involves unwinding the double helix and synthesizing complementary strands.

Prepzy

Memory Aids

- "Semi" means half old, half new strands in each DNA molecule.
- Complementary base pairing: A pairs with T, G pairs with C.

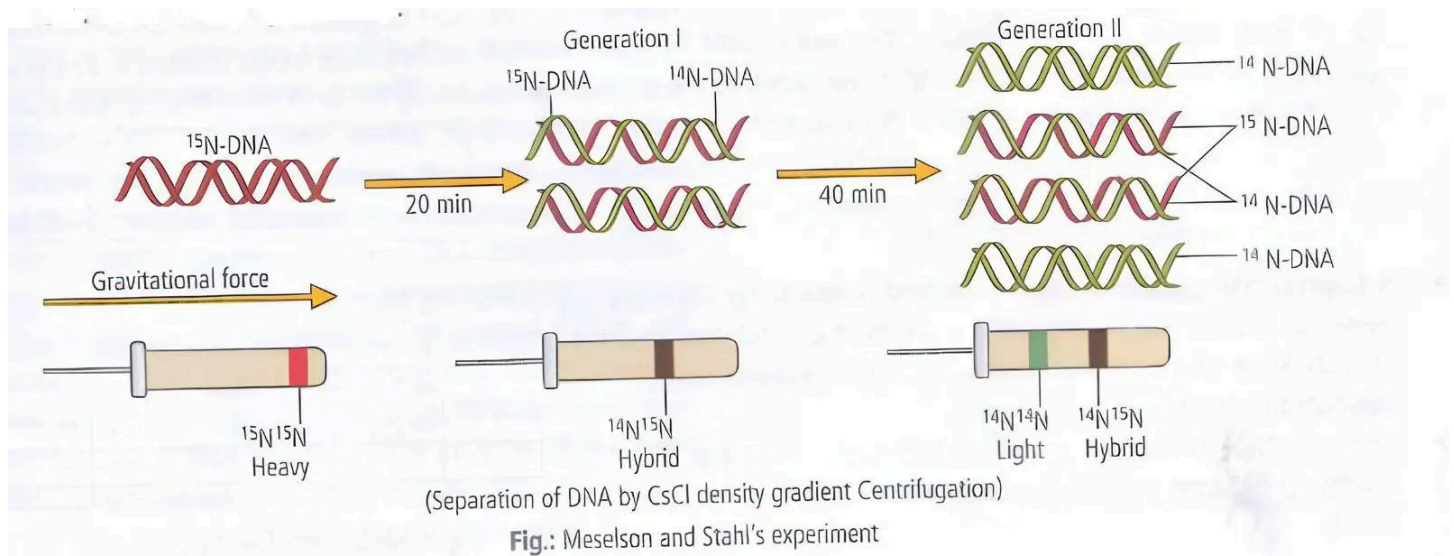
Quick Checkpoints

- What is semiconservative replication?
- Which strands serve as templates during replication?
- How does base pairing ensure accurate copying?

The Experimental Proof (Meselson and Stahl Experiment)

Experiment and Findings

- Meselson and Stahl (1958) proved DNA replicates semi-conservatively using *Escherichia coli* and nitrogen isotopes.
- DNA labeled with heavy nitrogen (^{15}N) was grown in light nitrogen (^{14}N) medium.
- After one generation, DNA density was intermediate, showing hybrid strands with one old and one new strand.
- Subsequent generations showed a mix of hybrid and light DNA, confirming semiconservative replication.
- Taylor et al. confirmed semi-conservative replication in chromosomes of *Vicia faba* using radioactive thymidine.



Memory Aids

- Heavy nitrogen (^{15}N) labels old DNA strands.
- Hybrid DNA has one heavy and one light strand.
- Density gradient centrifugation separates DNA by weight.

Quick Checkpoints

- What isotope was used to label DNA in the experiment?
- How does DNA density change after replication in light nitrogen?
- What does the presence of hybrid DNA indicate?

The Machinery and the Enzymes

Enzymes and Process

- DNA replication is energy-intensive and requires enzymes called catalysts.
- Main enzyme: DNA-dependent DNA polymerase synthesizes new DNA strands using a DNA template.
- Replication speed in *E. coli* is about 2000 base pairs per second.
- Energy comes from breakdown of deoxyribonucleotide triphosphates.
- Helicase unwinds DNA at origin of replication (Ori site).
- Single-stranded binding proteins stabilize separated strands.

- Topoisomerases relieve supercoiling by nicking and resealing DNA strands.
- Replication fork is Y-shaped where DNA strands separate and replication occurs.
- RNA primer initiates new DNA strand synthesis; DNA polymerase adds nucleotides 5' to 3'.
- Leading strand replicates continuously (template 3' to 5').
- Lagging strand replicates discontinuously in Okazaki fragments (template 5' to 3').
- DNA ligase joins Okazaki fragments.
- Replication starts at specific origin sites, not randomly.
- In eukaryotes, replication occurs during S-phase of the cell cycle.

Prepzy

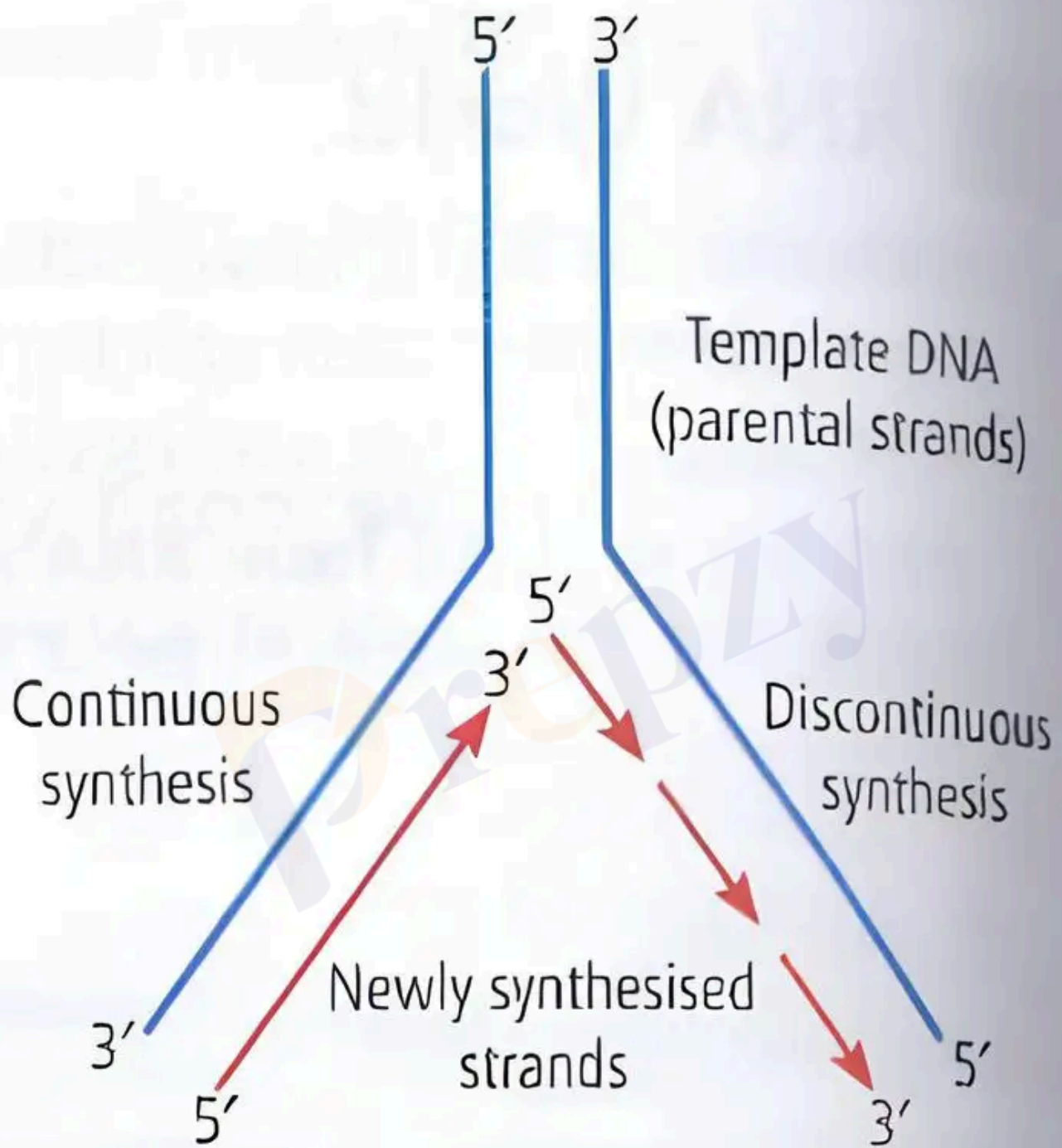


Fig.: Replicating fork

Memory Aids

- Helicase = Unwinds DNA.
- SSB proteins = Stabilize single strands.
- Topoisomerase = Relieves supercoiling.
- DNA polymerase = Synthesizes DNA 5' to 3'.
- RNA primer = Starts DNA synthesis.
- Ligase = Joins DNA fragments.
- Leading strand = Continuous synthesis.
- Lagging strand = Discontinuous synthesis.

Quick Checkpoints

- Which enzyme unwinds the DNA helix?
- What stabilizes single DNA strands during replication?
- Why is RNA primer needed?
- How does replication differ on leading and lagging strands?
- When does DNA replication occur in eukaryotic cells?

Transcription

Process and Mechanism

- Transcription is copying genetic information from one DNA strand into RNA.
- Complementarity governs transcription; adenine pairs with uracil (not thymine).
- Only one DNA strand and a segment are transcribed, unlike replication which duplicates entire DNA.
- Both DNA strands are not copied because:
 - If both strands were templates, they would produce different RNA sequences, complicating protein synthesis.
 - Simultaneous RNA molecules would be complementary, forming double-stranded RNA, blocking translation.
- A transcription unit has three DNA regions: promoter, structural gene, terminator.
- Structural gene can be monocistronic (eukaryotes) or polycistronic (prokaryotes).
- Coding strand sequence matches RNA except thymine is replaced by uracil.

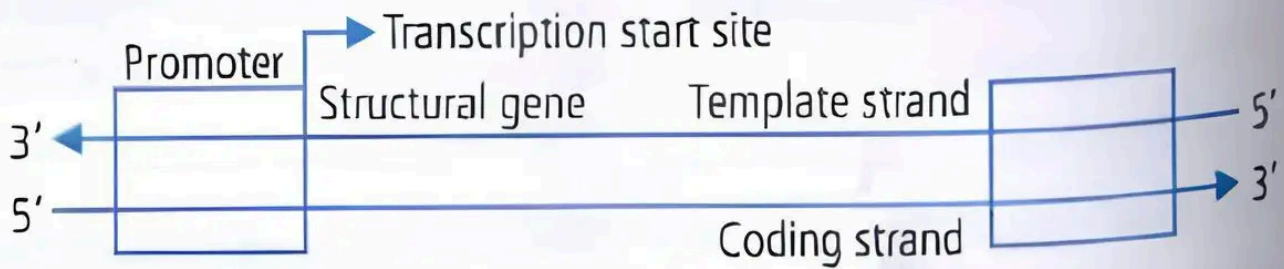


Fig.: Schematic structure of a transcription unit

Memory Aids

- Transcription = DNA to RNA copying.
- Only one strand used to avoid conflicting proteins.
- Promoter starts, terminator ends transcription.
- RNA matches coding strand except T → U.

Quick Checkpoints

- Why is only one DNA strand transcribed?
- What are the three regions of a transcription unit?
- How does RNA sequence relate to the coding strand?

Types of RNA and Process of Transcription

RNA Types and Transcription Steps

- Bacteria have three RNA types: *mRNA* (template), *tRNA* (amino acid carrier), *rRNA* (structural/catalytic).
- Transcription enzyme: DNA-dependent RNA polymerase holoenzyme = core enzyme + sigma factor.
- Sigma factor directs RNA polymerase to promoter, enhancing specificity.
- Transcription stages:
 - **Initiation:** RNA polymerase binds promoter (-35 and -10 regions), DNA locally untwists.

- **Elongation:** Sigma factor leaves after 8-9 nucleotides; RNA polymerase synthesizes RNA 5'→3', unwinding DNA ahead.
- **Termination:** Rho factor binds RNA, catches polymerase at pause site, unwinds DNA-RNA hybrid, releasing polymerase.
- In bacteria, transcription and translation occur in cytoplasm; mRNA requires no processing.

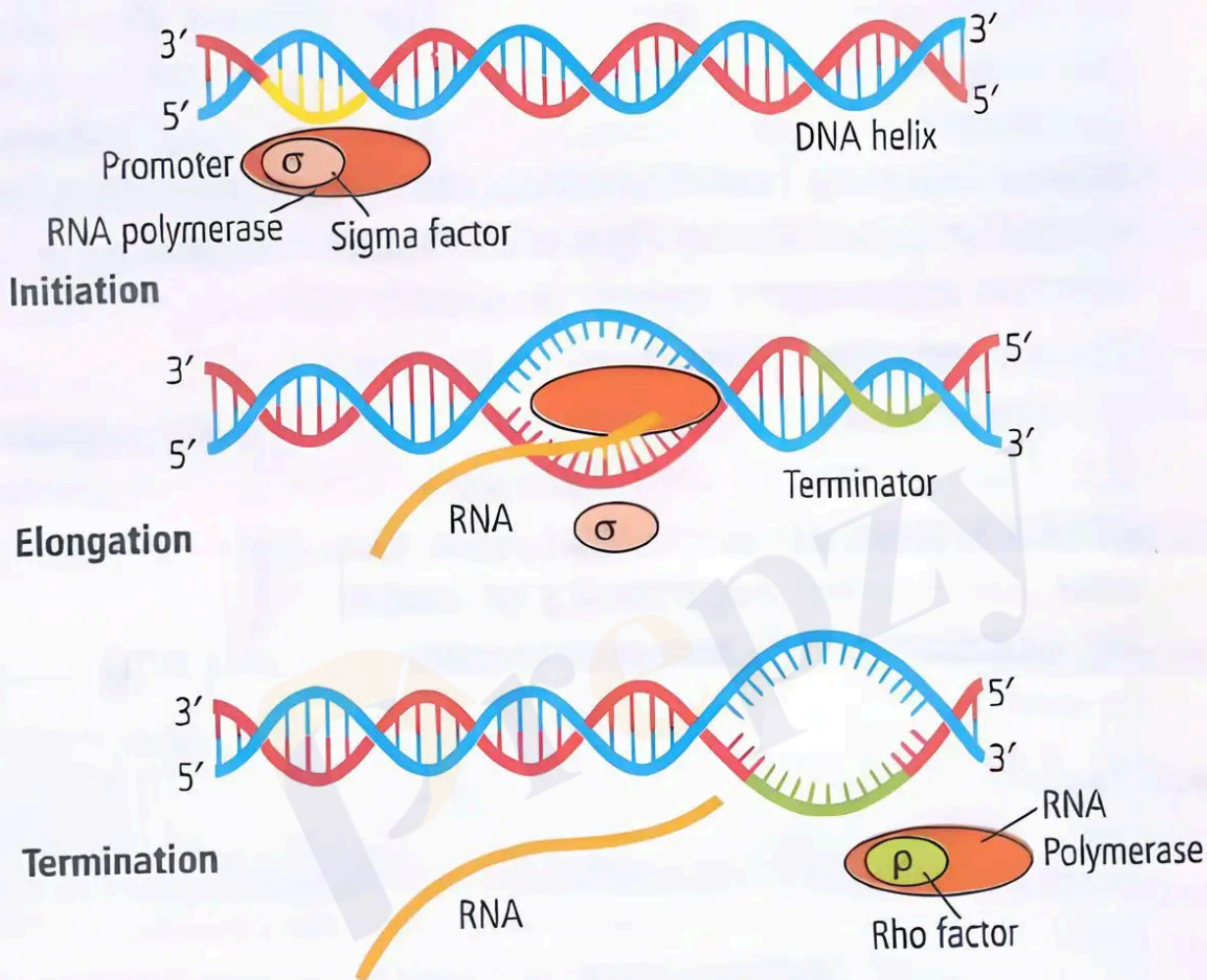


Fig.: Process of transcription in bacteria

Memory Aids

- 3 RNA types: mRNA (message), tRNA (transfer), rRNA (ribosome).
- Holoenzyme = core + sigma factor.
- Stages: Initiation (-35/-10), Elongation (5'→3'), Termination (rho factor).
- Bacterial mRNA is ready for translation immediately.

Quick Checkpoints

- What roles do the three RNA types play?
- What is the function of sigma factor?
- Describe the three transcription stages.
- Where does bacterial transcription occur?

Transcription in Eukaryotes

RNA Polymerases and mRNA Processing

- Three RNA polymerases in nucleus:
 - RNA polymerase I: transcribes rRNAs (28S, 18S, 5.8S).
 - RNA polymerase II: transcribes precursor *mRNA* (hnRNA).
 - RNA polymerase III: transcribes *tRNA*, 5S rRNA, snRNA.
- Transcription occurs in nucleus; *mRNA* moves to cytoplasm for translation.
- *mRNA* maturation includes:
 - 5' cap addition (7-methyl guanosine).
 - 3' poly-A tail addition (200-300 adenines).
 - Removal of introns by RNA splicing.
 - Joining of exons by ligase to form mature *mRNA*.
- Primary transcript called heterogeneous nuclear RNA (hnRNA) or pre-*mRNA*.
- Splicing shows RNA world dominance.

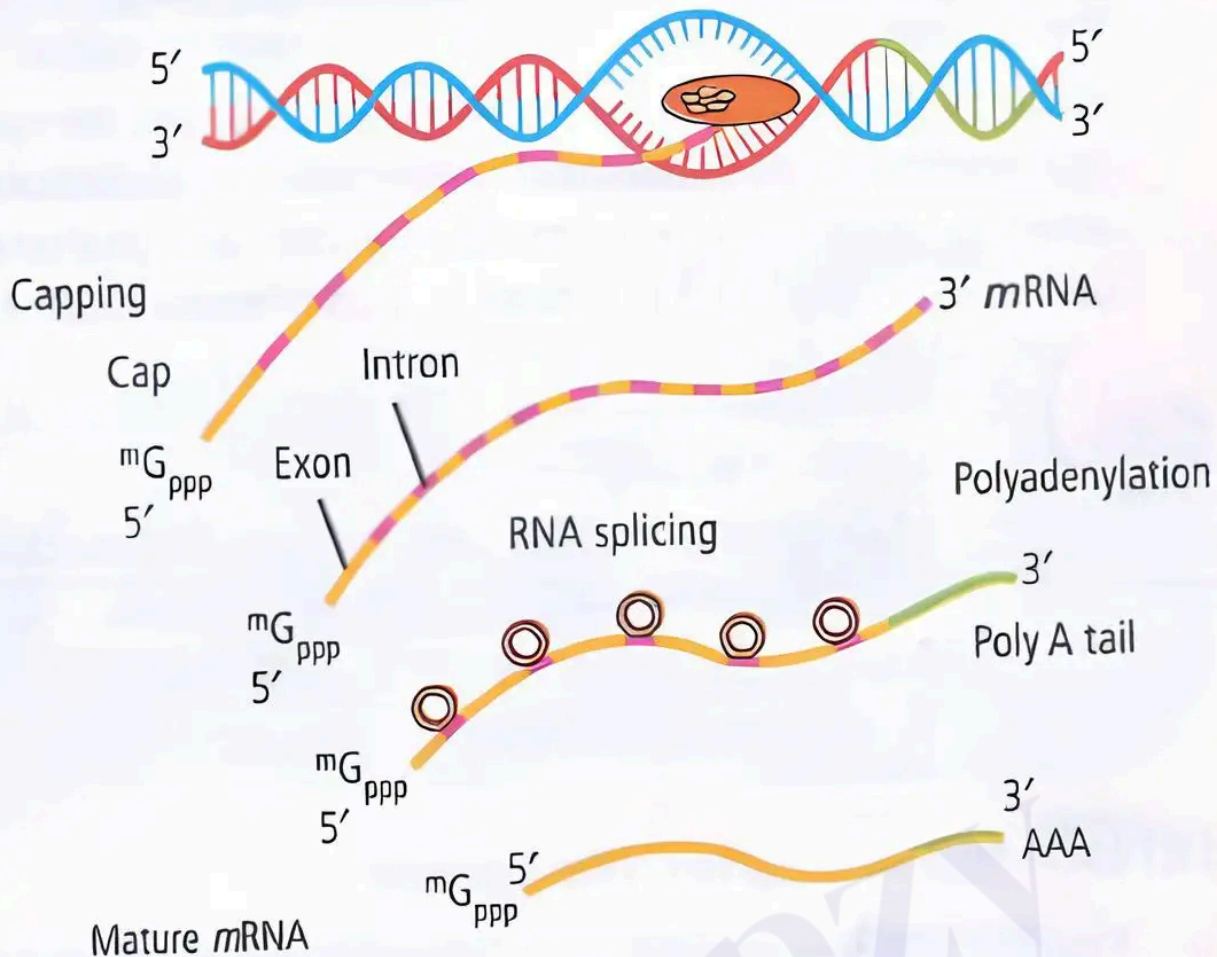


Fig.: Process of transcription in eukaryotes

Memory Aids

- RNA polymerase I = rRNA, II = pre-mRNA, III = tRNA + 5S rRNA + snRNA.
- mRNA processing: cap, tail, splicing.
- hnRNA = pre-mRNA with introns and exons.
- Splicing removes introns, joins exons.

Quick Checkpoints

- Which RNA polymerase transcribes pre-mRNA?
- What modifications occur during mRNA maturation?
- What is hnRNA?
- Why is RNA splicing important?

Genetic Code

Characteristics

- Genetic code is the codon representation decoding RNA into polypeptides.
- Code consists of three nucleotides (triplet code) producing 64 codons for 20 amino acids.
- Universal: Same codon codes same amino acid in all organisms (e.g., UUU codes phenylalanine).
- Degeneracy: Multiple codons can code for one amino acid.
- Non-overlapping: Each base is part of only one codon.
- Non-ambiguous: One codon codes for one amino acid, except AUG which can code methionine (start) or valine (internal).
- Initiator codon: Usually AUG, starts protein synthesis coding methionine.
- Non-sense codons: UAG, UAA, UGA signal termination of protein synthesis.
- Wobble hypothesis: Third base of codon binds loosely allowing one tRNA to recognize multiple codons differing at third base.
- First codon discovered: UUU; Nobel prize awarded to Holley, Nirenberg, Khorana in 1968.

Table: The codons for the various amino acids

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

Memory Aids

- Triplet code = 3 bases per codon.
- Universal code = same in all organisms.
- Degeneracy = multiple codons for one amino acid.

- Wobble = flexibility at third base.
- Start codon = AUG (Methionine).
- Stop codons = UAG, UAA, UGA.

Quick Checkpoints

- What is the length of a codon?
- Which codon initiates protein synthesis?
- Define degeneracy of genetic code.
- What are non-sense codons?
- Explain wobble hypothesis briefly.

Mutations and Genetic Code

Types and Effects

- Point mutation: Single base change causing amino acid substitution (e.g., glutamate to valine causes sickle cell anemia).
- Frameshift mutation: Insertion/deletion of 1 or 2 bases shifts reading frame, altering protein.
- Insertion/deletion of multiples of 3 bases does not shift frame, proving codon is triplet and read continuously.
- Non-sense mutation: Mutation creates stop codon, halting protein synthesis (e.g., UUA to UAA).
- Mis-sense mutation: Codon change leads to different amino acid (e.g., UUA Leucine to UUU Phenylalanine).
- Same sense mutation: Silent mutation; codon changes but amino acid remains same due to degeneracy (e.g., UUA to UUG).

Memory Aids

- Point mutation = single base change.
- Frameshift = insertion/deletion not multiple of 3.
- Non-sense = stop codon formed.
- Mis-sense = amino acid changes.

- Silent mutation = no amino acid change.

Quick Checkpoints

- What causes sickle cell anemia at genetic level?
- Define frameshift mutation.
- Difference between non-sense and mis-sense mutation?
- What is silent mutation?

tRNA – the Adapter Molecule

Structure and Function

- tRNA reads genetic code and binds specific amino acids; also called soluble RNA (sRNA).
- Secondary structure resembles clover-leaf; actual shape is compact inverted L.
- Has anticodon loop with bases complementary to mRNA codon.
- Has amino acid acceptor end for binding amino acids.
- Each tRNA is specific for one amino acid.
- Initiator tRNA starts translation; no tRNA for stop codons.

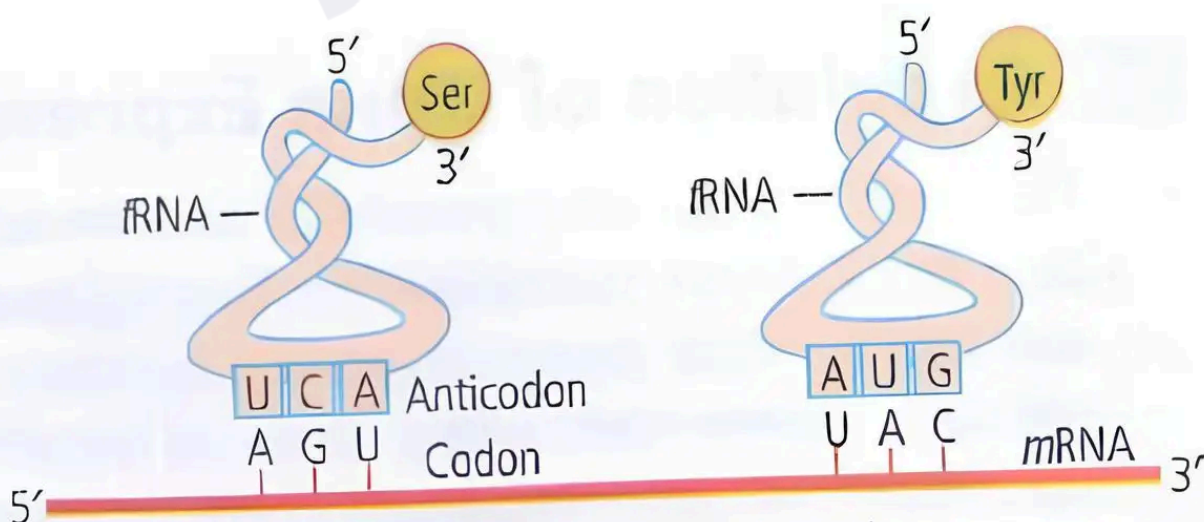


Fig.: tRNA - the adapter molecule

Memory Aids

- tRNA = adapter between mRNA codon and amino acid.
- Anticodon pairs with mRNA codon.
- Initiator tRNA starts protein synthesis.
- No tRNA for stop codons.

Quick Checkpoints

- What is the role of tRNA in protein synthesis?
- Describe the structure of tRNA.
- What is an anticodon?
- Which tRNA initiates translation?
- Are there tRNAs for stop codons?

Translation

Process Overview

- Translation is the polymerisation of amino acids into a polypeptide.
- Amino acid sequence is determined by mRNA base sequence.
- Amino acids are joined by peptide bonds, requiring energy.
- Amino acids are activated with ATP and linked to their specific tRNA (charging/aminoacylation).
- Ribosome, made of rRNAs and ~80 proteins, synthesises proteins.
- Ribosome has two subunits: large and small; translation starts when small subunit binds mRNA.
- Large subunit has two sites for amino acid binding; ribosome acts as catalyst (ribozyme) for peptide bond formation.
- mRNA has untranslated regions (UTRs) at 5' and 3' ends for efficient translation.
- Translation starts at AUG codon recognised by initiator tRNA.
- Elongation: aminoacyl-tRNAs bind codons on mRNA; ribosome moves codon to codon adding amino acids.
- Termination: release factor binds stop codon, releasing polypeptide.

Memory Aids

- "Charging tRNA" = amino acid + tRNA + ATP.
- Ribosome = protein factory with two subunits.
- Start codon AUG = initiator tRNA binding site.
- Peptide bond formation catalysed by ribosomal RNA (ribozyme).
- UTRs ensure translation efficiency.

Quick Checkpoints

- What defines amino acid sequence in translation?
- What is the role of tRNA charging?
- How does ribosome catalyse peptide bond formation?
- What are untranslated regions in mRNA?
- How does translation terminate?

Protein Biosynthesis

Post-Translational Modifications

- Proteins made by ribosomes undergo modifications to become mature.
- Modifications are enzyme-catalysed and target specific protein sequences.

Memory Aids

- PTMs = protein maturation steps after synthesis.
- Enzymes recognise specific sequences for modification.

Quick Checkpoints

- What happens to proteins after synthesis?
- What catalyses post-translational modifications?

Polyribosomes

Simultaneous Translation

- Multiple ribosomes translate a single mRNA simultaneously.
- First ribosome moves past start codon, allowing next ribosome to bind.
- Cluster of ribosomes on mRNA is called polyribosome.
- Found in both prokaryotes and eukaryotes.

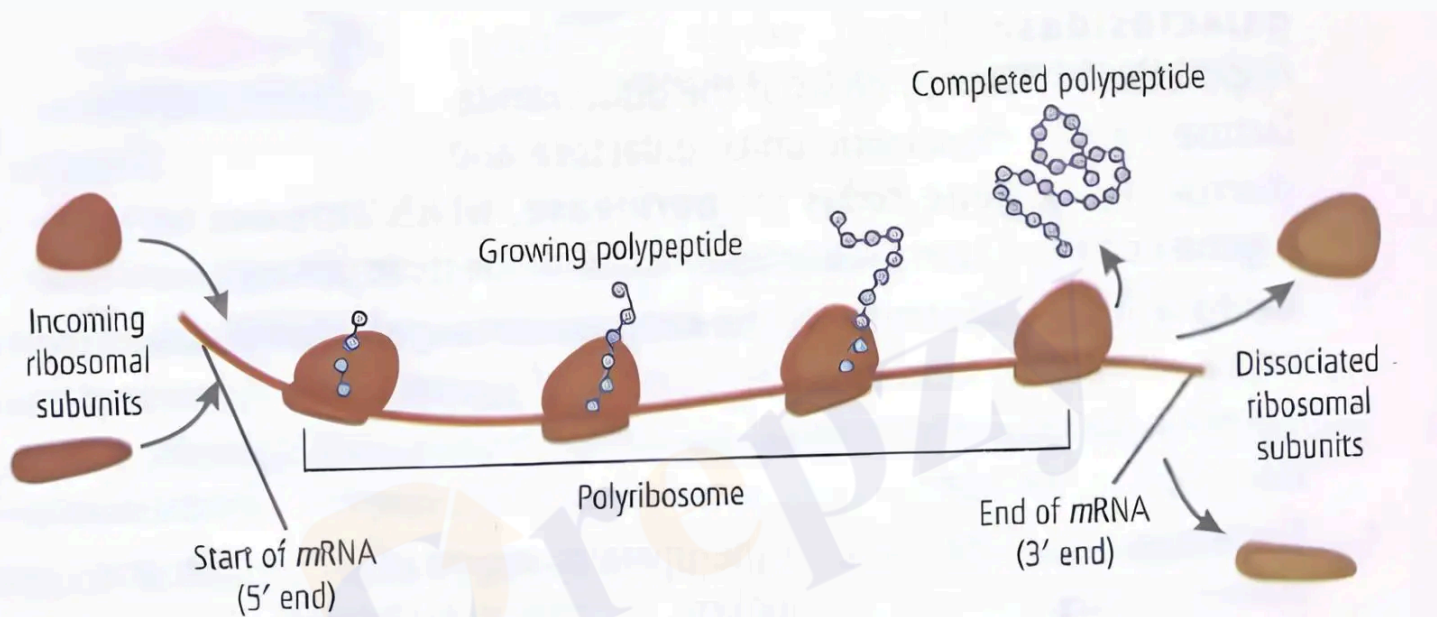


Fig.: Several ribosome bind to an mRNA during translation forming polyribosome

Memory Aids

- Polyribosome = many ribosomes on one mRNA.
- Increases protein production speed.

Quick Checkpoints

- What is a polyribosome?
- How does polyribosome formation increase protein synthesis?

Regulation of Gene Expression

Gene Expression and Regulation

- Genes express to perform specific functions, e.g., beta-galactosidase in *E. coli* hydrolyses lactose into glucose and galactose.
- Gene expression depends on metabolic, physiological, or environmental conditions.
- Two types of gene regulation: **positive** (inducible) and **negative** (repressible).
- Negative regulation: genes express until suppressed by a regulatory product.
- Positive regulation: genes remain off until induced.
- Eukaryotic regulation occurs at transcription, RNA processing, mRNA transport, and translation levels.
- In prokaryotes, gene expression is mainly controlled at transcription initiation.
- RNA polymerase activity at promoters is regulated by proteins binding to operators adjacent to promoters.
- Each operon has a specific operator and repressor, e.g., lac operator binds only lac repressor.

Memory Aids

- Negative = Repressible = Always ON until turned OFF.
- Positive = Inducible = Always OFF until turned ON.
- Prokaryotes control gene expression mainly at transcription initiation.
- Operators are specific to their repressors.

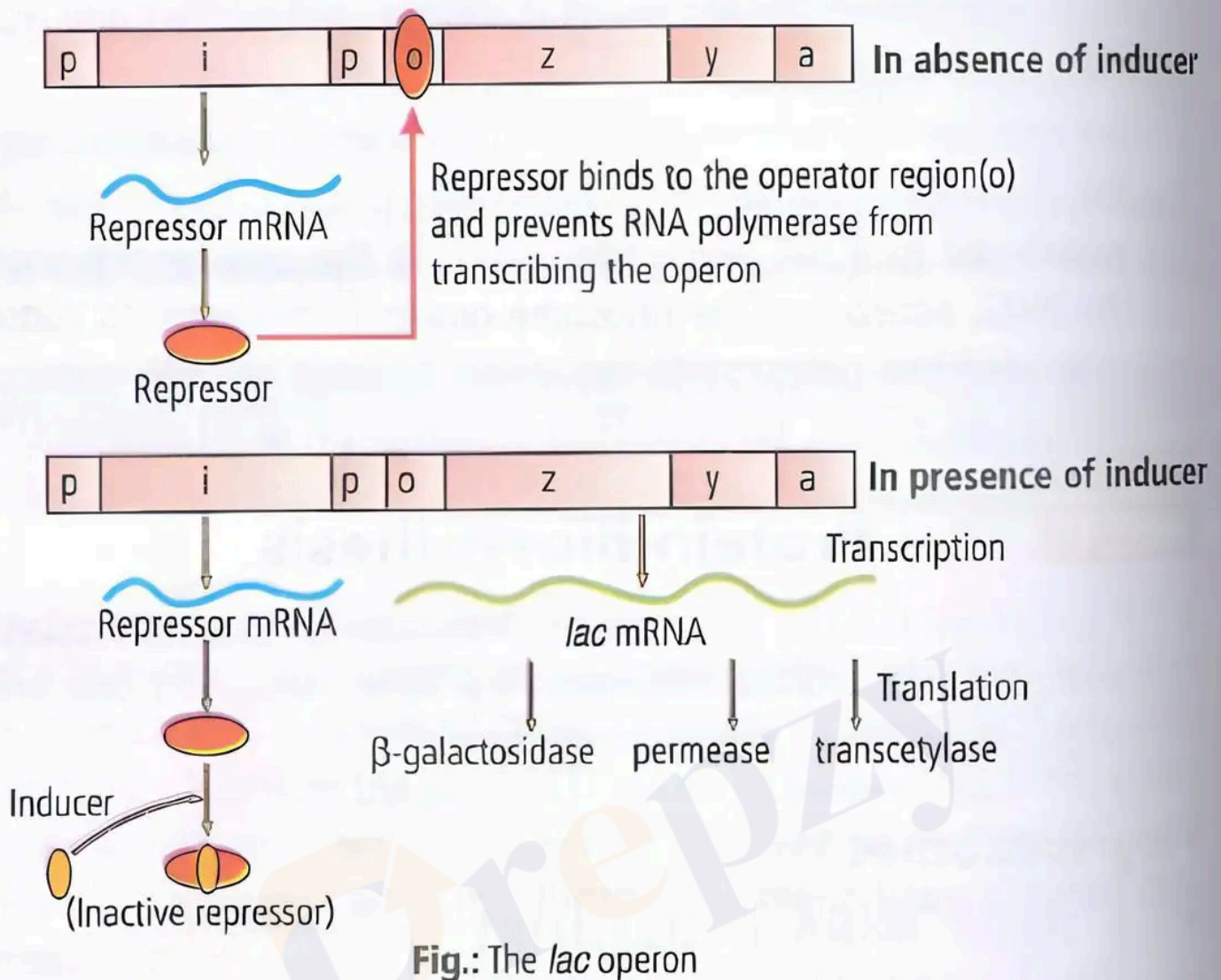
Quick Checkpoints

- What triggers gene expression in positive regulation?
- Where does eukaryotic gene regulation occur?
- What is the role of operator in prokaryotic gene regulation?
- How does negative regulation differ from positive regulation?

The Lac Operon

Structure and Function

- Discovered by Francois Jacob and Jacques Monod.
- Operon: cluster of genes under a single promoter, common in bacteria.
- Lac operon has one regulatory gene (i) and three structural genes (z, y, a).
- i gene codes for lac repressor protein.
- z gene codes for beta-galactosidase, hydrolyzing lactose into glucose and galactose.
- y gene codes for permease, increasing cell permeability to lactose.
- a gene codes for transacetylase, involved in lactose metabolism.
- Lactose acts as an inducer by binding to the repressor and inactivating it.
- In absence of glucose, lactose enters cell via permease and induces operon expression.
- Repressor binds operator to block RNA polymerase and prevent transcription.
- Inducer binding inactivates repressor, allowing transcription of z, y, a genes.
- Regulation by repressor is negative regulation.



Memory Aids

- i gene = inhibitor (repressor).
- z = zero (beta-galactosidase breaks lactose).
- y = yes (permease helps lactose entry).
- a = acetylase (transacetylase enzyme).
- Lactose = inducer that inactivates repressor.
- Negative regulation = repressor blocks transcription until inducer present.

Quick Checkpoints

- What genes are present in lac operon and their functions?

- How does lactose regulate lac operon?
- What is the role of the repressor protein?
- Define negative regulation in lac operon context.

Human Genome Project

Overview

- Launched in 1990 as a mega project called International Human Genome Sequencing Consortium.
- Aimed to identify all human genes, determine their functions, and understand organism formation.
- Coordinated by U.S. Department of Energy and National Institute of Health; completed in 2003.
- Closely linked to bioinformatics development.

Goals

- Identify 20,000-25,000 human genes.
- Sequence 3 billion DNA base pairs.
- Store data in databases.
- Improve data analysis tools.
- Transfer technology to industries.
- Address ethical, legal, social issues (ELSI).
- Develop genetic linkage map using genetic markers.
- Obtain physical map by cloning DNA into BACs and YACs.

Enabling Factors

- Genetic engineering for DNA isolation and cloning.
- Fast DNA sequencing techniques.
- Sequencing of model organisms like bacteria, yeast, *Caenorhabditis elegans*, *Drosophila*, rice, and *Arabidopsis*.

Memory Aids

- HGP = Human genes + sequencing + bioinformatics + ethics.
- Goals: Identify, Sequence, Store, Analyze, Transfer, Ethics, Maps.

Quick Checkpoints

- What is the main aim of HGP?
- Which organizations coordinated HGP?
- What are BACs and YACs used for?

Methodologies of HGP

Techniques Used

- Expressed Sequence Tags (ESTs) to identify genes expressed as RNA.
- Sequence Annotation to assign functions to coding and non-coding regions.
- Automated DNA sequencing based on Sanger's method.
- Generation of overlapping DNA fragments for sequencing.
- Computer programs for sequence alignment and annotation.
- Chromosome 1 sequencing completed in 2006 (last human chromosome sequenced).

Memory Aids

- ESTs focus on expressed genes.
- Sanger method for sequencing.
- Computers essential for sequence assembly.

Quick Checkpoints

- What is the role of ESTs?
- Why are computers needed in HGP?
- When was chromosome 1 sequencing completed?

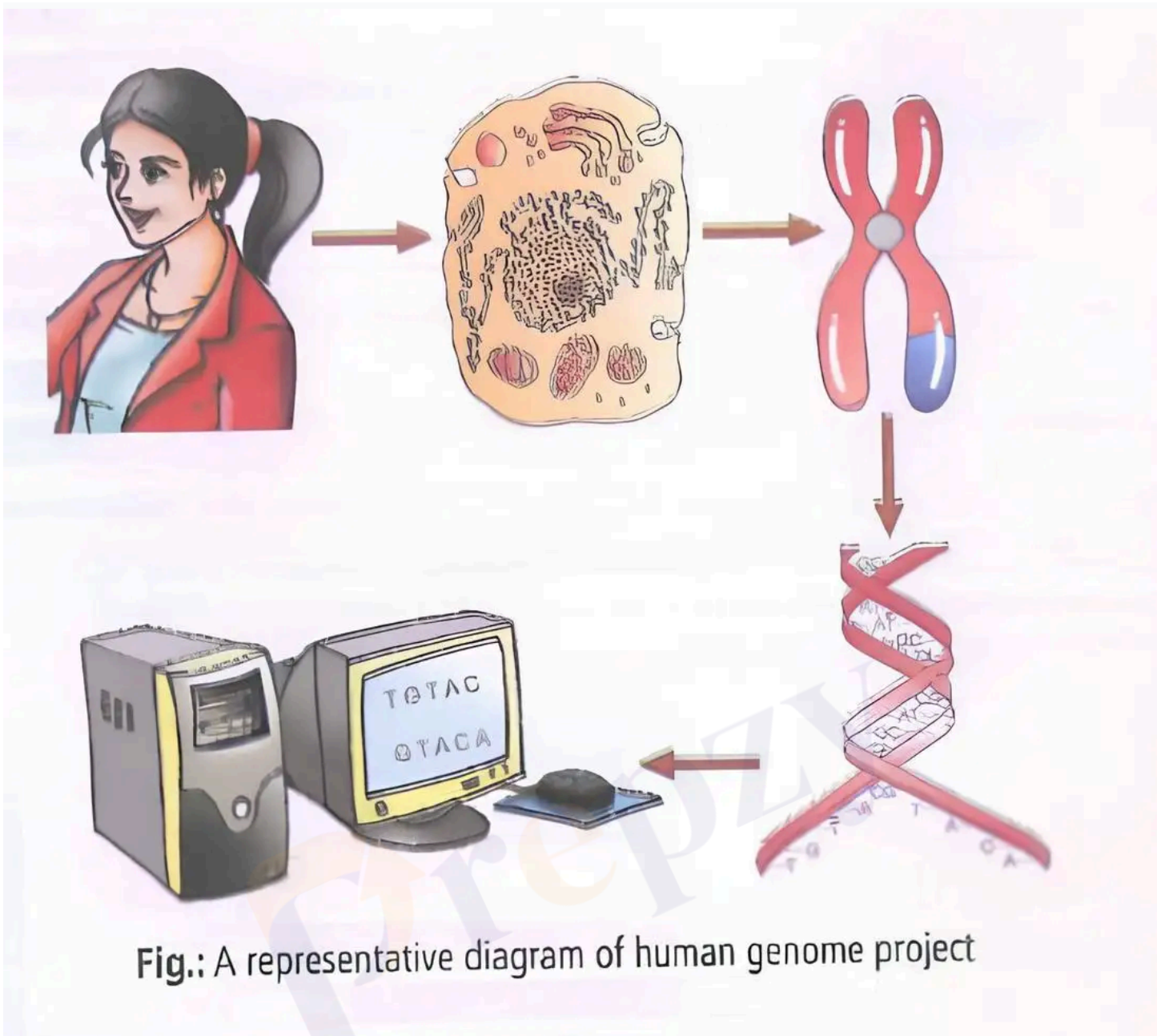


Fig.: A representative diagram of human genome project

Salient Features of HGP

Key Observations

- Human genome has 3,164.7 million nucleotide base pairs.
- Average gene size ~3000 bases; largest gene dystrophin (2.4 million bases).
- Estimated 30,000 genes; 99.9% nucleotide bases identical in all humans.
- Functions of over 50% genes unknown.
- Less than 2% genome codes for proteins.
- Repeated DNA sequences form large genome portion.
- Chromosome 1 has most genes (2968); Y chromosome has fewest (231).
- About 1.4 million single nucleotide polymorphisms (SNPs) identified.

Memory Aids

- Genome size ~3.2 billion bases.
- Most genes on chromosome 1, least on Y.
- SNPs = single base differences in DNA.

Quick Checkpoints

- How many base pairs are in the human genome?
- What percentage of genome codes for proteins?
- What are SNPs?

Applications of HGP

Uses

- Enables systematic, large-scale biological research.
- Allows study of all genes or transcripts in tissues/organs/tumors.
- Helps understand gene and protein networks in life chemistry.

Memory Aids

- HGP = Research scale-up + gene study + network understanding.

Quick Checkpoints

- How does HGP change biological research?
- What can be studied using HGP data?
- Why is understanding gene networks important?

DNA Fingerprinting

Concept and Process

- DNA fingerprinting identifies differences in repetitive DNA regions called Variable Number of Tandem Repeats (VNTRs).
- VNTR size ranges from 0.1 to 20 kb; repetitive DNA forms satellite DNA peaks separate from bulk genomic DNA.
- These sequences show high polymorphism, unique to individuals except identical twins.
- DNA from any tissue (blood, hair, skin, saliva, sperm) shows the same polymorphism pattern.
- Useful in forensic identification and genetic diversity studies.
- Polymorphism arises from mutations; multiple alleles exist at a locus with frequency >0.01 .
- Developed by Alec Jeffreys using satellite DNA probes and Southern blot hybridization.
- Steps: DNA extraction, PCR amplification, restriction enzyme digestion, electrophoresis, denaturation, transfer to membrane (Southern blotting), probe hybridization, autoradiography.

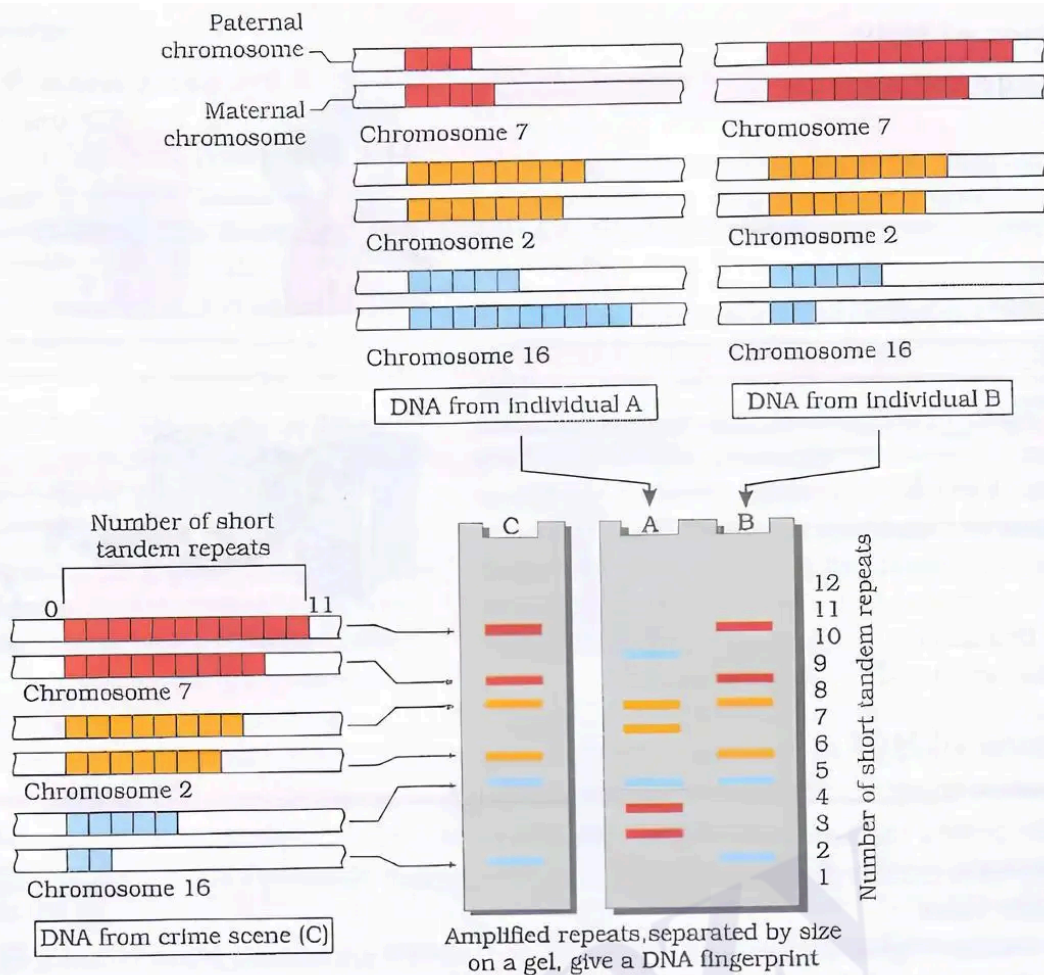


Fig.: Schematic representation of DNA fingerprinting: Few representative chromosomes have been shown to contain different copy number of VNTR. For the sake of understanding different colour schemes have been used to trace the origin of each band in the gel. The two alleles (paternal and maternal) of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

Memory Aids

- VNTR = Variable Number of Tandem Repeats, key to uniqueness.
- Polymorphism = multiple alleles due to mutations.
- Southern blotting = transfer DNA to membrane for probe detection.
- Alec Jeffreys = pioneer of DNA fingerprinting.

Quick Checkpoints

- What is VNTR and its role in DNA fingerprinting?
- How does polymorphism contribute to individual identification?
- What are the main steps in DNA fingerprinting?
- Who developed DNA fingerprinting and what probe was used?

Southern Blotting

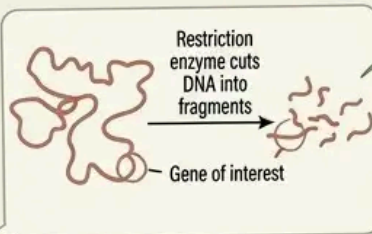
Technique and Steps

- Transfers DNA fragments from gel to membrane, preserving banding pattern.
- DNA immobilized on membrane for hybridization with labeled probes.
- Named after Edwin Southern (1975).
- Steps: DNA digestion by restriction enzymes, gel electrophoresis, transfer to nylon membrane, prehybridization with buffer, hybridization with specific probe, detection by X-ray or other methods.
- Used to detect specific DNA sequences within a mixture.

Prepzy

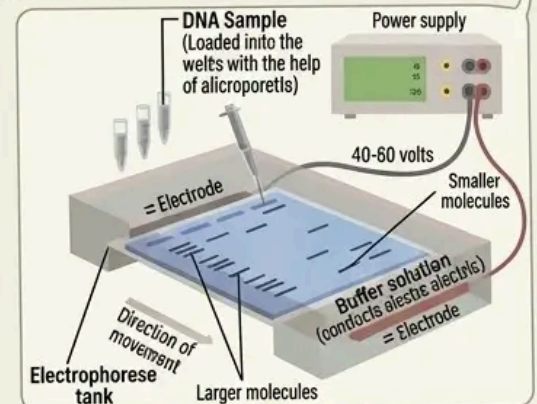
1 DNA Digestion

DNA containing gene of interest is extracted from cells with the help of enzymes and subjected to restriction digestion as to obtain DNA fragments.



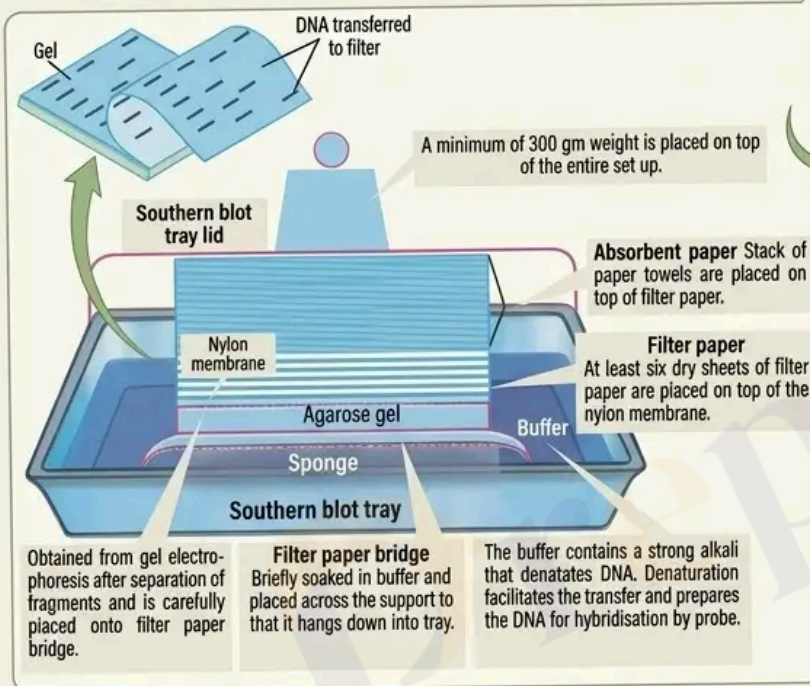
2 Separation of Fragments by Gel Electrophoresis

Fragmented DNA is typically electrophoresed on an agarose gel to separate them on the basis of their molecular weight. DNA size is determined by running the DNA ladders (or molecular weight markers) in the lane adjacent to the sample. By matching the distance covered by fragment and marker the molecular weight and size of DNA fragment can be estimated.



3 Blotting

After electrophoresis, the DNA fragments in the gel are denatured for which it is soaked in about 0.5 M NaOH which separates dsDNA into ssDNA and then transferred onto positively charged nylon or nitrocellulose membrane. The traditional transfer of DNA is done overnight using an upward capillary transfer of DNA from an agarose gel onto a membrane and its subsequent immobilization by UV irradiation (for nylon) or baking for nitrocelluloses.

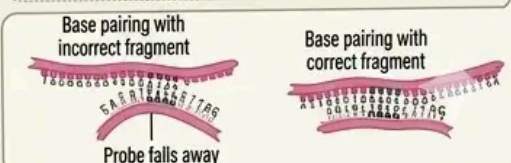


4 Prehybridisation of Southern blot

The nylon membrane is carefully removed from the setup and is incubated in a buffer for 15 minutes containing radioactively labelled probe. Labelled probe is a nucleic acid probe with sequence homologous to the target sequence under study, labelled with radioactivity, fluorescent dye or an enzyme that can generate a chemiluminescent signal when incubated with the appropriate substrate.

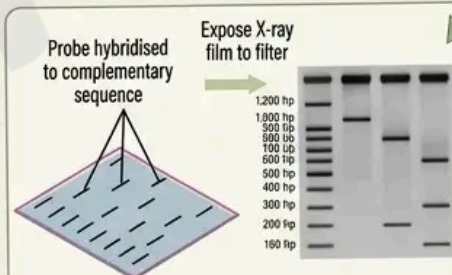
5 Hybridisation

Hybridisation is promoted when the labelled probe is incubated with the DNA fragments immobilised on the blot in a tube that is constantly ressed or in a sealed plastic bag that is placed on a shaker. Random collisions bring small regions of complementary sequences together to start the renaturation. Base pairing proceeds in a super-like fashion as flanking sequences are complementary, probe sticks even after washing. The unhybridised probe is removed by washing several times in buffer. Only the fully hybridized labelled probe molecules with sequences complementary to gene of interest remain bound.



6 Detection

The labelled probe is detected using the method required for the type of label used. For example, radiolabelled probes may be detected using X-ray film or a phosphorimaging instrument.



Memory Aids

- Southern blot = DNA transfer and detection.
- Restriction enzymes cut DNA into fragments.
- Probe binds only to complementary DNA sequences.
- Detection shows dark bands where probe hybridizes.

Quick Checkpoints

- What is the purpose of Southern blotting?
- How are DNA fragments separated before transfer?
- What role does the probe play in Southern blotting?
- How is the hybridized DNA detected?

DNA Sequencing by Sanger's Method

Process and Significance

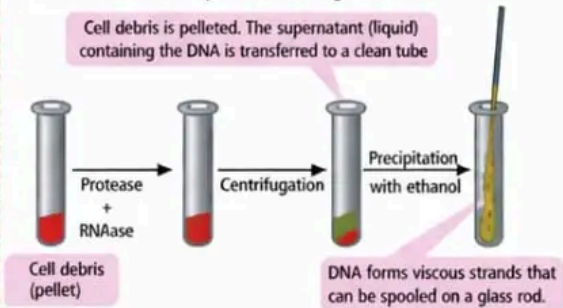
- Determines exact nucleotide order (A, T, C, G) in DNA.
- Developed by F. Sanger in 1977 using enzymatic chain termination.
- Steps: DNA isolation, fragmentation, PCR amplification, denaturation, primer binding, DNA polymerase elongation with normal nucleotides and chain-terminating ddNTPs.
- ddNTPs lack 3' OH group, causing chain termination at random points.
- Fragments separated by gel electrophoresis; sequence deduced from fragment sizes.
- Essential for genome analysis and understanding biological processes.

DNA Sequencing by Sanger's Method

DNA sequencing is the process of determining the precise order of nucleotides *i.e.*, adenine (A), thymine (T), cytosine (C) and guanine (G) along a DNA strand. It helps in genome analysis and hence understanding the biological processes. Sanger's classical method is an enzymatic chain termination method of sequencing developed by **F. Sanger et al.** in 1977. This method is performed in various steps.

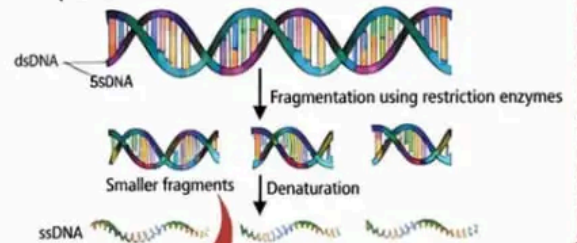
1 DNA Isolation

Genomic DNA to be sequenced can be isolated from blood, saliva or tissue samples of an organism.



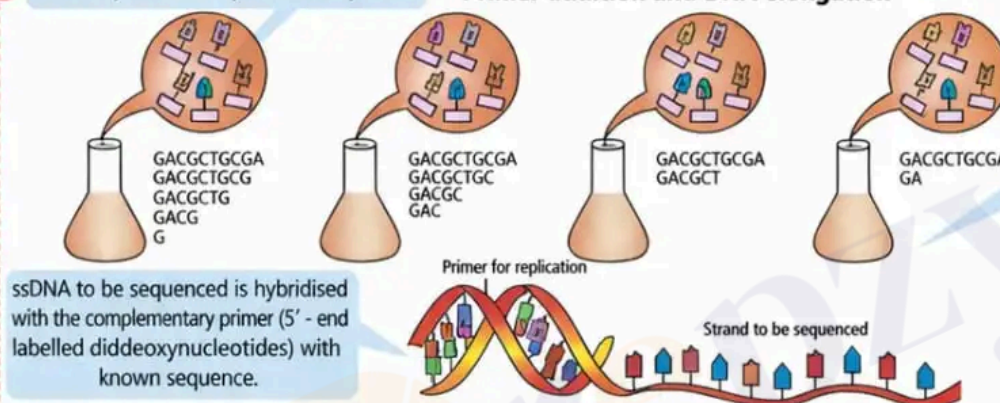
2 DNA Fragmentation, Amplification and Denaturation

The entire genomic DNA extracted is broken down into short fragments and amplified using PCR or cloning vector. Fragmented DNA is denatured to obtain single stranded DNA (ssDNA), which is used as template.



3 ddNTPs may be radioactively or fluorescently labelled **Primer addition and DNA elongation**

ddNTPs may be radioactively or fluorescently labelled

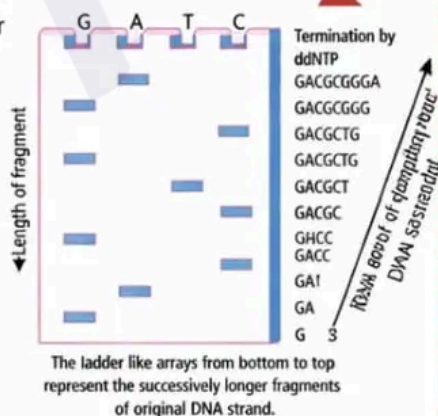


The fragmented sample DNA is divided into 4 separate reaction mixtures, each comprising of 4 normal deoxynucleotides (dNTPs) : dATP, dCTP, dGTP, dTTP, DNA polymerase enzyme, primer, single stranded template DNA and any one of the four dideoxynucleotides (ddNTPs) in each of the flask.

Enzyme starts synthesis of the complementary strand using normal dNTPs but the chain elongation ceases when a ddNTP gets incorporated into the synthesising strand which prevents addition of next nucleotide. Thus, mixture of terminated fragments having different lengths are produced.

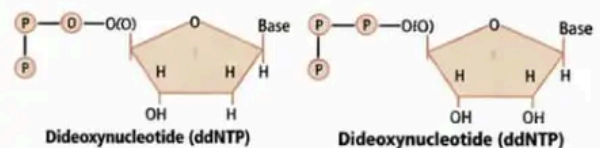
4 Each of the four reaction

Each of the four reaction mixtures are run together on a polyacrylamide gel for electrophoresis. Separated fragments are visualised by autoradiography. From the position of the bands, the sequence of the original DNA template strand is read using sequencing softwares/programmes on computer.



Significance of ddNTPs

Dideoxynucleotides also known as 2', 3' dideoxynucleotides are analogous to the normal dNTPs. DNA polymerase cannot distinguish between them while catalysing polymerisation.



The absence of 3' OH in ddNTPs inhibits the chain elongation as the phosphodiester bond could not be formed with next nucleotide. Hence, terminates the DNA elongation.

Memory Aids

- ddNTPs stop DNA chain growth, enabling sequence determination.
- Chain termination method = Sanger sequencing.

- Gel electrophoresis separates fragments by size.
- Primer initiates DNA synthesis.

Quick Checkpoints

- What causes chain termination in Sanger sequencing?
- Why is gel electrophoresis used in sequencing?
- What is the role of primers in DNA sequencing?
- How does Sanger sequencing help in genome analysis?

Prepzy