Collegedunia NCERT Notes

The Ultimate NCERT Revision Guide for Class 12 Biology

Full-colour diagrams · New NCERT 2026-27 syllabus

Chapter 5: Molecular Basis of Inheritance

Class 12 Biology — 12th NCERT (2026-27)

Also see for this chapter: NCERT Solutions | Formula Sheet | Exemplar Solutions

How to use these notes

This guide covers the full NCERT Chapter 5 — DNA structure, the search for genetic material, replication, transcription, the genetic code, translation, gene regulation (*lac* operon), the Human Genome Project, and DNA fingerprinting — plus the extra depth NEET aspirants need. Every process is supported by a labelled diagram or the original NCERT figure, key facts sit in coloured boxes, and a Quick Reference section at the end consolidates every value you must memorise. This is one of the highest-yield chapters in NEET Biology (typically 4–6 questions).

Contents

1 The DNA: Structure and Packaging **2**

1.1 Nucleotide: the Repeating Unit 3

1.2 The Double Helix — the Watson–Crick Model 4

1.3 Packaging of the DNA Helix 5

2 The Search for Genetic Material **7**

2.1 Griffith's Transforming Principle (1928) 7

2.2 Avery, MacLeod and McCarty (1933–44) 7

2.3 The Hershey–Chase Experiment (1952) 8

2.4 Properties of a Good Genetic Material: DNA vs RNA 9
2.5 The RNA World 10

3 DNA Replication **10**
3.1 Semiconservative Replication 10
3.2 The Meselson–Stahl Experiment (1958) 11
3.3 The Machinery and the Enzymes 12

4 Transcription **13**
4.1 The Transcription Unit 13
4.2 Transcription Unit and the Gene 14
4.3 Types of RNA and the Process of Transcription 14

5 Genetic Code and Translation **15**
5.1 The Genetic Code 15
5.2 Mutations and the Genetic Code 16
5.3 tRNA — the Adapter Molecule 16
5.4 Translation 17

6 Regulation of Gene Expression: The *Lac* Operon **18**
6.1 The Operon Concept 18
6.2 Switching the Operon On and Off 18

7 Human Genome Project and DNA Fingerprinting **19**
7.1 The Human Genome Project (HGP) 19
7.2 DNA Fingerprinting 20

8 Quick Reference Summary **22**
8.1 Key Numbers to Memorise 22
8.2 Base Pairing and the Code 22
8.3 People and their Contributions 23

1 The DNA: Structure and Packaging

The previous chapter established that genes are the units of inheritance, but it did not say what genes are made of *chemically*. This chapter answers that question. Deoxyribonucleic acid (DNA) is the genetic material in the overwhelming majority of organisms; ribonucleic acid (RNA) is the genetic material in some viruses and

otherwise functions as a messenger, adapter, structural, and catalytic molecule. We begin with the structure of DNA, because every later process — replication, transcription, translation — is a direct consequence of that structure.

1.1 Nucleotide: the Repeating Unit

DNA and RNA are long polymers of **nucleotides**. Each nucleotide is built from three parts joined together: a nitrogenous base, a pentose sugar, and a phosphate group.

- **Nitrogenous base.** Two chemical classes: **purines** (Adenine, Guanine — double-ring) and **pyrimidines** (Cytosine, Thymine, Uracil — single-ring). Cytosine occurs in both DNA and RNA. Thymine is unique to DNA; Uracil replaces thymine in RNA.
- **Pentose sugar. Deoxyribose** in DNA, **ribose** in RNA. Ribose carries an extra $-$OH at the 2′ carbon; deoxyribose has $-$H there.
- **Phosphate group.** Links nucleotides together and gives the molecule its acidic, negatively charged character.

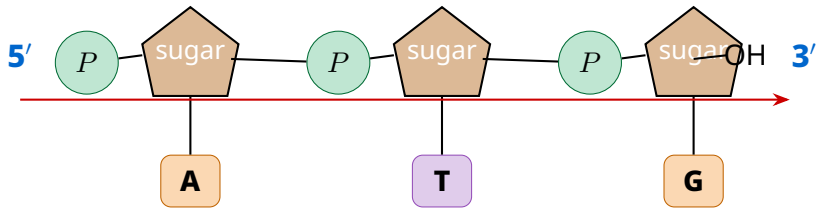
A base joined to a sugar (through an N-glycosidic linkage at the 1′ carbon) is a **nucleoside** (e.g. adenosine, deoxyadenosine). When a phosphate is added to the 5′ carbon of the nucleoside, the unit becomes a **nucleotide**.

Building up the polymer

$$\begin{aligned} \text{Base} + \text{Sugar} &\rightarrow \text{Nucleoside} && \text{(N-glycosidic bond)} \\ \text{Nucleoside} + \text{Phosphate} &\rightarrow \text{Nucleotide} && \text{(phosphoester bond)} \\ \text{Nucleotide} + \text{Nucleotide} &\rightarrow \text{Dinucleotide} && \text{(3′–5′ phosphodiester)} \end{aligned}$$

$$\text{Dinucleotide} \longrightarrow \text{Polynucleotide chain}$$

Two nucleotides are joined by a **3′–5′ phosphodiester linkage**; many such links build the polynucleotide chain.



Chain has a definite polarity: free 5′-phosphate at one end, free 3′-OH at the other

The diagram above shows the sugar–phosphate backbone with the nitrogenous bases projecting outward. The free phosphate at one terminus defines the **5′-end**; the free $-$OH at the other defines the **3′-end**. This directionality is fundamental: enzymes read and build nucleic acids only in the $5' \rightarrow 3'$ direction.

DNA vs RNA — the two chemical differences

(1) Sugar: deoxyribose (DNA) carries $-$H at 2′; ribose (RNA) carries $-$OH at 2′. (2) Base: DNA uses **thymine** (5-methyl uracil); RNA uses **uracil**. The extra 2′-OH and the absence of thymine make RNA chemically more reactive and less stable than DNA.

1.2 The Double Helix — the Watson–Crick Model

DNA was first isolated by Friedrich Meischer in 1869, who called it 'Nuclein'. Its three-dimensional structure remained unsolved until 1953, when James Watson and Francis Crick proposed the **double-helix model**, building on the X-ray diffraction images produced by Maurice Wilkins and Rosalind Franklin and on Erwin Chargaff's rule.

Chargaff's rule: in any double-stranded DNA, the amount of Adenine equals Thymine and the amount of Guanine equals Cytosine, so $\frac{A}{T} = 1$ and $\frac{G}{C} = 1$. This was the experimental clue that bases pair specifically.

Salient features of the B-DNA double helix

1. Two anti-parallel polynucleotide chains: one runs $5' \to 3'$, the other $3' \to 5'$. The sugar–phosphate backbone is outside; bases point inward.
2. Base pairing by hydrogen bonds: **A=T (2 H-bonds)**, **G≡C (3 H-bonds)**. A purine always pairs with a pyrimidine, keeping the helix diameter uniform (≈ 2 nm).
3. Right-handed coil. Pitch (one full turn) $= 3.4$ nm, with ≈ 10 base pairs per turn, so the rise per base pair $= 0.34$ nm.
4. Base stacking (hydrophobic interaction between stacked pairs) adds stability, additional to the H-bonds.

The original NCERT figure below shows the two anti-parallel strands held together by hydrogen bonds — note A pairs with T by two dashed bonds and G pairs with C by three.

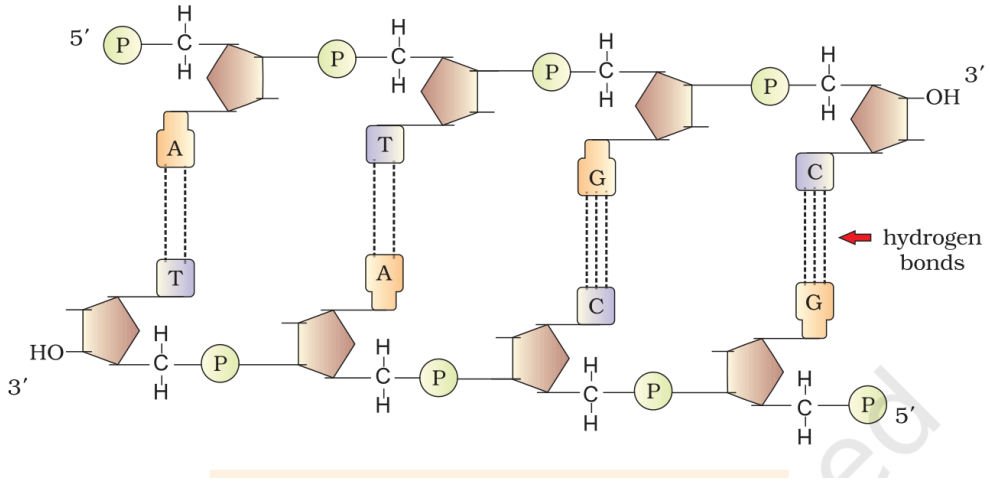
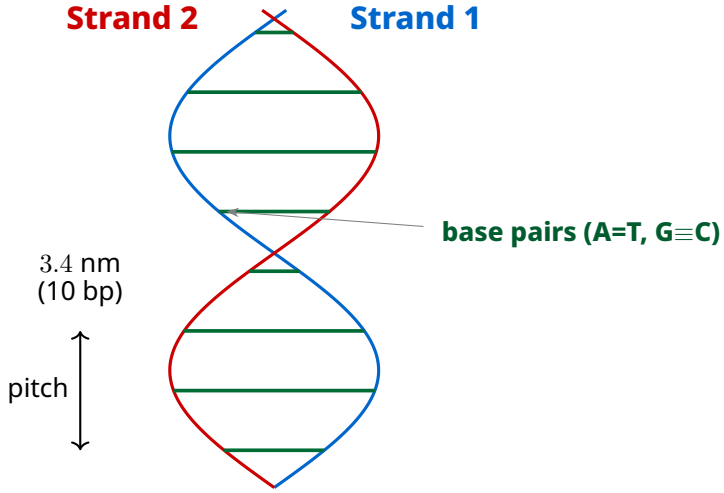


Fig 5.2 — Double-stranded polynucleotide chain. The two strands are anti-parallel ($5' \to 3'$ on top, $3' \to 5'$ below); A=T joined by 2 H-bonds, G≡C by 3.



The right-handed double helix above: the two backbones spiral around a common axis, base pairs lie flat like the rungs of a twisted ladder.

Central Dogma

Right after the double helix, Crick stated the **Central Dogma**: genetic information flows DNA $\to$ RNA $\to$ Protein (replication keeps DNA copies; transcription makes RNA; translation makes protein). In some viruses (retroviruses like HIV) the flow reverses, RNA $\to$ DNA, by **reverse transcription** — the basis of the term 'Teminism' / reverse transcriptase.

1.3 Packaging of the DNA Helix

The length of a DNA molecule is enormous compared with the cell. In a typical human diploid cell there are 6.6×10^9 bp; multiplying by the rise per base pair gives the length.

Length of DNA in a cell

$$\text{Length} = (\text{number of bp}) \times 0.34 \times 10^{-9} \text{ m}$$

For a human diploid cell: $6.6 \times 10^{9} \times 0.34 \times 10^{-9}$ m ≈ 2.2 m.
Yet the nucleus is only $\approx 10^{-6}$ m across — a packaging ratio of roughly a million.

Numerical you must be able to do

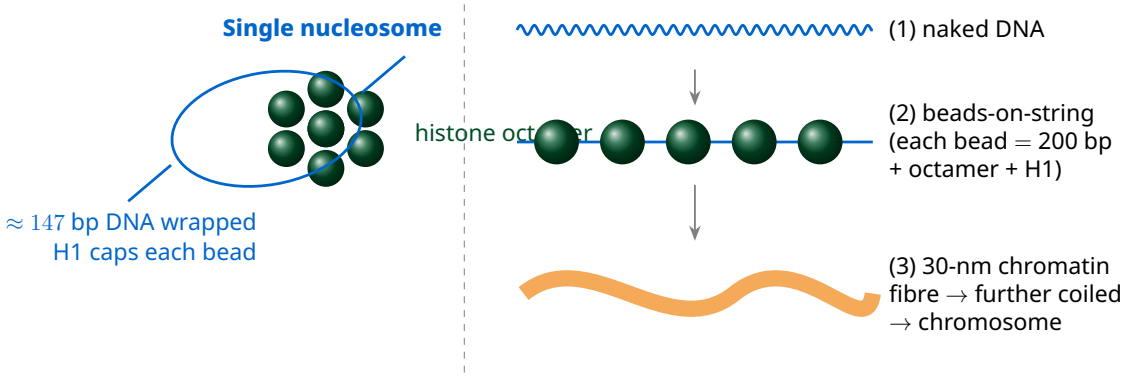
"If *E. coli* DNA is 1.36 mm long, find the number of base pairs."

$$\text{bp} = \frac{1.36 \times 10^{-3} \text{ m}}{0.34 \times 10^{-9} \text{ m/bp}} = 4.0 \times 10^{6} \text{ bp}.$$

Always convert mm to m first, then divide by 0.34 nm.

In prokaryotes (e.g. *E. coli*), there is no nucleus. The negatively charged DNA is held by positively charged proteins in a region called the **nucleoid**, organised into large supercoiled loops.

In eukaryotes, packaging is hierarchical and uses basic, positively charged proteins called **histones** (rich in the basic amino acids **lysine** and **arginine**). Eight histone molecules form a **histone octamer**; the negatively charged DNA wraps around it (≈ 200 bp) to form a **nucleosome**.



Nucleosomes are the repeating unit of **chromatin**. Under the electron microscope, chromatin looks like 'beads-on-string'. The string is condensed further into chromatin fibres and, at metaphase, into chromosomes, with the help of **Non-histone Chromosomal (NHC) proteins**.

Euchromatin vs Heterochromatin

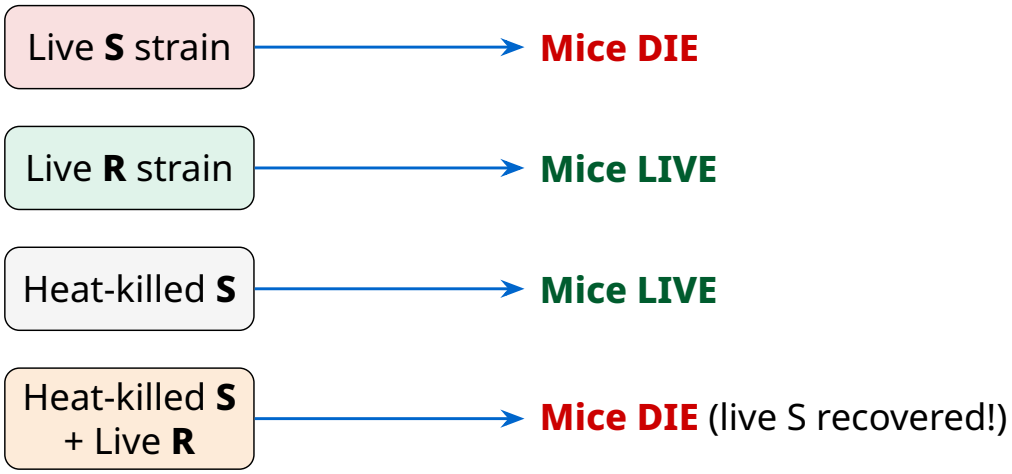
Euchromatin — loosely packed, stains light, **transcriptionally active**.
Heterochromatin — densely packed, stains dark, **transcriptionally inactive**.

2 The Search for Genetic Material

It took nearly a century after Meischer's nuclein and Mendel's factors before DNA was proven to be the genetic material. The proof came in three classic stages.

2.1 Griffith's Transforming Principle (1928)

Frederick Griffith worked with *Streptococcus pneumoniae*. The **S strain** has a smooth polysaccharide (mucous) coat and is virulent; the **R strain** is rough, has no coat, and is non-virulent.



The striking result is the last row: heat-killed S mixed with live R killed the mice, and **living S** bacteria were recovered. Griffith concluded that some 'transforming principle' had passed from the dead S cells and converted live R into virulent S. He did not know its chemical nature.

Transformation

A non-virulent strain acquired a new, heritable characteristic (the polysaccharide coat) by taking up material from dead cells. The conclusion: a chemical substance can carry genetic information from one cell to another.

2.2 Avery, MacLeod and McCarty (1933–44)

Oswald Avery, Colin MacLeod and Maclyn McCarty purified the biochemicals (protein, DNA, RNA) from heat-killed S cells and tested each for transforming ability.

Treatment of heat-killed S extract	Did transformation of R cells occur?
Add protease (digests protein)	Yes — so protein is not the genetic material
Add RNase (digests RNA)	Yes — so RNA is not the genetic material
Add DNase (digests DNA)	**No** — transformation blocked

Only digestion with **DNase** stopped transformation. They concluded that **DNA is the transforming principle** — the genetic material. Many biologists were still not fully convinced (DNA was thought too simple a molecule).

2.3 The Hershey–Chase Experiment (1952)

Alfred Hershey and Martha Chase gave the unequivocal proof using **bacteriophages** (viruses that infect bacteria). A phage injects its genetic material into the bacterium; the question was whether the injected material is DNA or protein. They exploited a chemical difference: **DNA contains phosphorus but no sulfur; protein contains sulfur but no phosphorus**.

- One batch of phages grown with radioactive $^{32}P \rightarrow$ **labelled DNA**.
- Another batch grown with radioactive $^{35}S \rightarrow$ **labelled protein coat**.
- Phages were allowed to infect *E. coli*; the empty coats were sheared off in a **blender** and separated from cells by **centrifugation**.

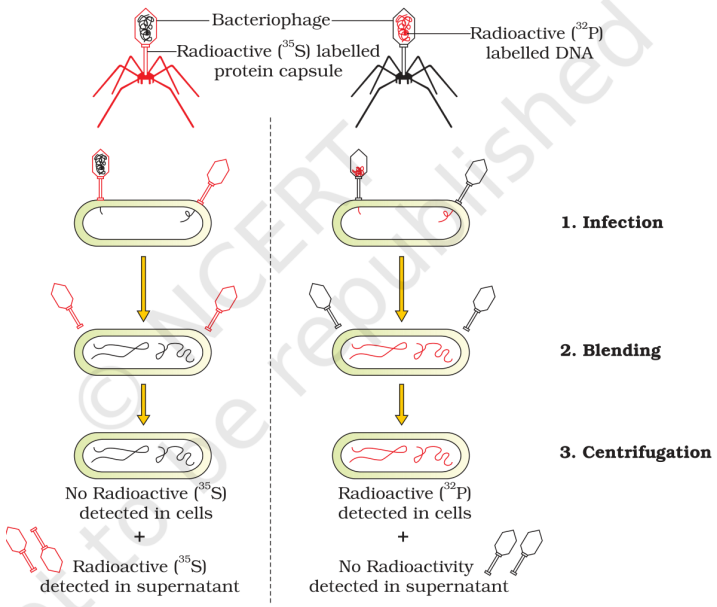


Figure 5.5 The Hershey-Chase experiment

Fig 5.5 — The Hershey–Chase experiment. ^{32}P (DNA) entered the bacteria; ^{35}S (protein) stayed outside in the supernatant.

Result and conclusion

Bacteria infected with ^{32}P phages **became radioactive** — DNA had entered.
Bacteria infected with ^{35}S phages were **not radioactive** — protein stayed outside.
Therefore **DNA, not protein, is the genetic material** that passes from virus to bacterium.

Common Mistake

Do not say Hershey–Chase used radioactive *nitrogen* or *carbon*. Those elements occur in *both* DNA and protein and cannot distinguish them. The whole design rests on ^{32}P being unique to DNA and ^{35}S unique to protein.

2.4 Properties of a Good Genetic Material: DNA vs RNA

Hershey–Chase settled the protein-vs-DNA debate, but RNA is the genetic material in some viruses (Tobacco Mosaic Virus, Qβ bacteriophage). Why is DNA the predominant choice? A molecule must satisfy four criteria.

Criteria for genetic material

1. It should be able to **replicate** (generate its own copy).
2. It should be chemically and structurally **stable**.
3. It should allow **slow mutation** (variation for evolution).
4. It should be able to **express** itself as 'Mendelian characters'.

Both DNA and RNA can replicate (base pairing/complementarity) and both can mutate, but:

Property	DNA	RNA
2′ group on sugar	$-$H (deoxyribose)	$-$OH (ribose) — reactive, labile
Strands	Double-stranded	Mostly single-stranded
Base	Thymine (extra stability)	Uracil
Chemical stability	High (better for storage)	Low (mutates faster)
Catalytic activity	Generally none	Can be catalytic (ribozyme)
Direct protein coding	No (needs RNA)	Yes

Why DNA stores, RNA transmits

The 2′-OH and single-stranded nature make RNA reactive and unstable — good for a short-lived, active worker, bad for long-term storage. DNA's double strand allows repair using the complementary strand, and thymine adds extra stability. So **DNA is the better store of information; RNA is better for the dynamic transfer and expression of it**.

2.5 The RNA World

Which came first? Evidence suggests **RNA was the first genetic material**. Essential life processes (metabolism, translation, splicing) evolved around RNA, and RNA can act both as information carrier *and* as a catalyst (ribozyme). Being reactive, RNA was unstable; DNA then evolved from RNA with chemical modifications (loss of 2′-OH, thymine for uracil, double strand) that make it more stable.

Order of discovery / proof

G–A–H: **G**riffith (transformation, 1928) → **A**very, MacLeod & McCarty (DNA is the principle, 1944) → **H**ershey–Chase (unequivocal proof, 1952). "**G**reat **A**ll-rounders **H**elped."

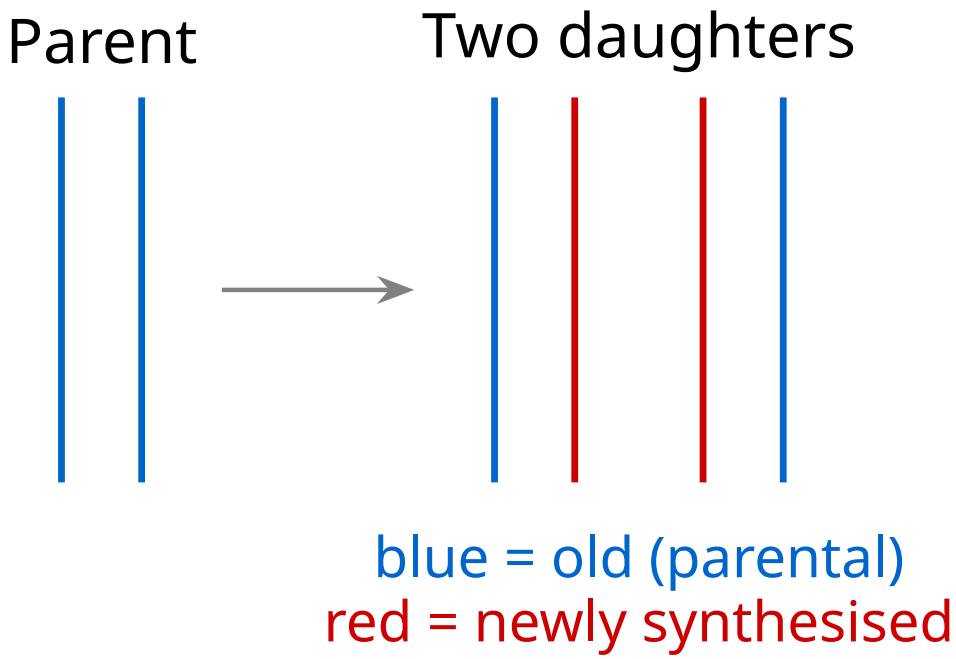
Solve the NCERT Exercises for this Chapter

3 DNA Replication

Watson and Crick noted that the specific base pairing immediately suggested a copying mechanism: each strand is the template for a new complementary strand.

3.1 Semiconservative Replication

The scheme they proposed is **semiconservative**: the two parental strands separate, each acts as a template, and each daughter DNA molecule ends up with **one parental (old) strand and one newly synthesised strand**.



Semiconservative model

Each new DNA = $\frac{1}{2}$ old + $\frac{1}{2}$ new. This conserves one original strand in every daughter molecule — hence "semi"-conservative.

3.2 The Meselson–Stahl Experiment (1958)

Matthew Meselson and Franklin Stahl proved semiconservative replication in *E. coli*.

1. Grew *E. coli* for many generations in $^{15}NH_4Cl$ (^{15}N = heavy, *non-radioactive* isotope). All DNA became heavy ($^{15}N/^{15}N$), distinguishable by **CsCl density-gradient centrifugation**.
2. Transferred cells to normal $^{14}NH_4Cl$ medium; sampled DNA at intervals (*E. coli* divides every 20 minutes).
3. After **one generation (20 min)**: all DNA was of **hybrid** ($^{15}N/^{14}N$) intermediate density.
4. After **two generations (40 min)**: equal amounts of **hybrid** and **light** ($^{14}N/^{14}N$) DNA.

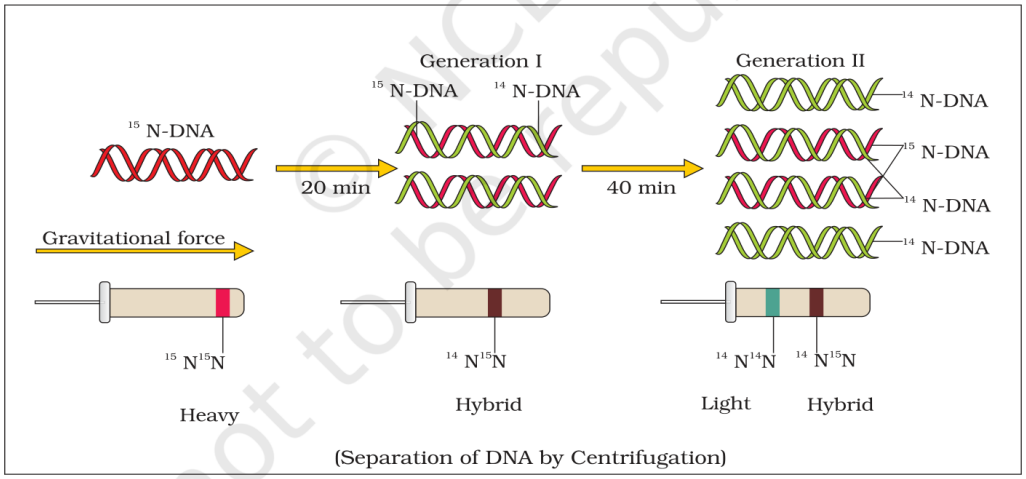


Fig 5.7 — Meselson and Stahl's experiment. One generation gives only hybrid-density DNA; the second gives a 1:1 mix of hybrid and light — exactly the prediction of the semiconservative model.

Predicting the bands

After n generations in ^{14}N: number of **hybrid** molecules stays $= 2$ (the two original ^{15}N strands), and the rest are **light**. So after 80 min (4 gen, 16 molecules): 2 hybrid : 14 light $= 1:7$.

A similar experiment by Taylor and colleagues (1958) on *Vicia faba* (faba bean) using radioactive thymidine proved that chromosomes in higher organisms also replicate semiconservatively.

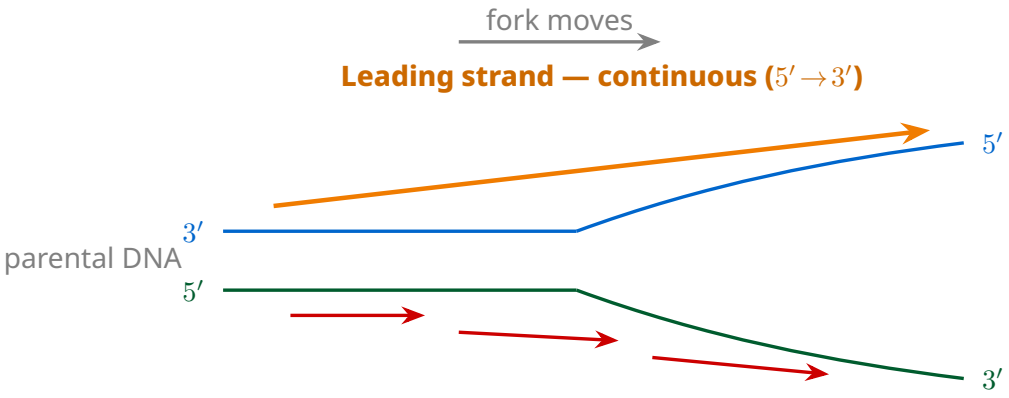
3.3 The Machinery and the Enzymes

The chief enzyme is **DNA-dependent DNA polymerase** (it uses a DNA template to polymerise deoxynucleotides). It is extremely fast and accurate — in *E. coli* the average rate is ≈ 2000 bp per second, completing 4.6×10^6 bp in about 18 minutes.

Energetics and direction

Substrates are **deoxynucleoside triphosphates (dNTPs)**; they serve a **dual purpose** — substrate *and* energy source (the two terminal phosphates are high-energy bonds, as in ATP).
DNA polymerase synthesises strictly in the **5′ → 3′ direction**.

Because polymerase works only 5′ → 3′, the two template strands cannot be copied identically at the replication fork:



Continuous vs discontinuous synthesis

On the template with polarity 3′ → 5′, the new strand grows **continuously** (leading strand). On the template with polarity 5′ → 3′, synthesis is **discontinuous** (lagging strand) — short fragments (Okazaki fragments) later sealed by **DNA ligase**.

Replication begins at a definite **origin of replication** (*ori*); polymerases cannot start on their own. In recombinant DNA technology, the *ori* is the sequence a vector must provide so that the foreign DNA can replicate. In eukaryotes, replication occurs in the **S phase** of the cell cycle and must be tightly coordinated with cell division (failure causes polyploidy).

Common Mistake

^{15}N is **not radioactive** — it is a heavy stable isotope, separated only by *density* on a CsCl gradient. Confusing it with a radioactive label is a frequent NEET trap.

4 Transcription

Transcription is the copying of genetic information from one strand of DNA into RNA. The principle of complementarity again governs it, except that adenine on DNA now pairs with **uracil** (not thymine) in the RNA. Unlike replication, only a **segment** of DNA and only **one strand** is copied.

Why only one strand, only a segment

If both strands were transcribed: (a) two RNAs of different sequence would code for two different proteins from one gene, complicating information transfer; (b) the two RNAs would be complementary, form double-stranded RNA, and could not be translated. So copying one strand of a defined segment is essential.

4.1 The Transcription Unit

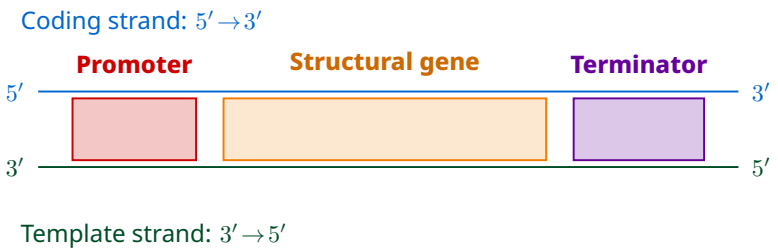
A transcription unit has three regions: a **Promoter**, the **Structural gene**, and a **Terminator**.

The two strands are named with respect to the RNA polymerase, which also moves $5' \rightarrow 3'$:

- **Template strand** — polarity $3' \rightarrow 5'$; the strand actually read by RNA polymerase.
- **Coding strand** — polarity $5' \rightarrow 3'$; same sequence as the RNA (except T for U); does not code for anything but is the reference for naming.

Worked example — writing the RNA

Template strand: $3'$-A T G C A T G C A T G C-$5'$
Coding strand: $5'$-T A C G T A C G T A C G-$3'$
RNA (complementary to template, U replaces T): $5'$-U A C G U A C G U A C G-$3'$
The RNA has the same sequence as the coding strand, with U in place of T.



Promoter is *upstream* (towards 5′ of coding strand); terminator is *downstream* (towards 3′).

The **promoter** provides the binding site for RNA polymerase and, by its position, defines which strand is template and which is coding. The **terminator** defines the end of transcription.

4.2 Transcription Unit and the Gene

A **gene** is the functional unit of inheritance. Defined as the DNA segment coding for a polypeptide, the structural gene is a **cistron**. Structural genes can be:

- **Monocistronic** — one gene per mRNA; mostly **eukaryotes**.
- **Polycistronic** — several genes on one mRNA; mostly **bacteria**.

In eukaryotes, genes are **split**: the coding/expressed sequences are **exons**; the intervening sequences that are removed are **introns**. Inheritance is also influenced by promoter and regulatory sequences (loosely called regulatory genes), even though these do not code for any RNA or protein.

Exon vs Intron

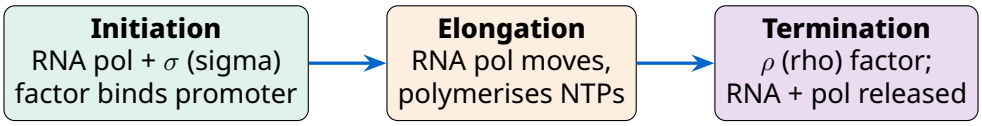
EXons are **EX**pressed (stay in mature RNA). **IN**trons are **IN** the way and are removed ("**IN**-the-bin").

4.3 Types of RNA and the Process of Transcription

In bacteria there are three RNAs, all made by a **single DNA-dependent RNA polymerase**:

- **mRNA** (messenger) — provides the template for protein synthesis.
- **tRNA** (transfer) — brings amino acids and reads the genetic code.
- **rRNA** (ribosomal) — structural and catalytic role in translation.

Transcription has three steps. The RNA polymerase alone can only **elongate**; it needs accessory factors to start and stop:



Initiation and termination factors

σ **(sigma) factor** — initiation factor; alters RNA polymerase so it recognises the promoter and starts.
ρ **(rho) factor** — termination factor; ends transcription and releases the RNA.

In **bacteria** there is no nucleus, so transcription and translation are **coupled** — translation of the mRNA can begin before transcription is complete.

In **eukaryotes** there are two extra complexities:

1. **Three RNA polymerases** in the nucleus, with a clear division of labour:
   - RNA Polymerase I $\rightarrow$ rRNAs (28S, 18S, 5.8S)
   - RNA Polymerase II $\rightarrow$ precursor of mRNA (**hnRNA**)

- RNA Polymerase III $\rightarrow$ tRNA, 5sRNA, snRNAs

2. The primary transcript (**hnRNA**) contains both exons and introns and is non-functional. It is processed by:
   - **Splicing** — introns removed, exons joined in order.
   - **Capping** — an unusual nucleotide (methyl guanosine triphosphate) added to the 5′-end.
   - **Tailing** — 200–300 adenylate residues (poly-A tail) added to the 3′-end.

   The fully processed hnRNA, now **mRNA**, is exported for translation.

Why split genes matter

The presence of introns is considered an ancient feature of the genome (a remnant of the RNA world). Alternative splicing of the same hnRNA can yield several different proteins from one gene — one reason humans have far more proteins than genes.

5 Genetic Code and Translation

5.1 The Genetic Code

No chemical complementarity exists between nucleotides and amino acids, so a **code** must relate them. George Gamow argued that with only 4 bases, a code for 20 amino acids needs a combination of bases: a triplet ($4^3 = 64$ codons) is more than enough.

Salient features of the genetic code

1. The codon is a **triplet**. 61 codons code for amino acids; 3 are **stop** codons.
2. It is **degenerate** — one amino acid may have more than one codon.
3. It is read **contiguously**, with **no punctuation** (no commas).
4. It is nearly **universal** (UUU = phenylalanine from bacteria to humans). A few exceptions in mitochondria/protozoa.
5. **AUG** is dual — it codes for Methionine *and* acts as the **initiator** codon.
6. **UAA, UAG, UGA** are **stop (terminator)** codons.

Har Gobind Khorana (synthetic RNA with defined sequences), Marshall Nirenberg (cell-free system) and Severo Ochoa (polynucleotide phosphorylase) together deciphered the code.

Stop codons

U Are Annoying, U Are Gone, U Go Away — UAA, UAG, UGA. Mnemonic "**U**AA **U**AG **U**GA = stop".

5.2 Mutations and the Genetic Code

Mutation studies link genes to function. **Point mutation:** a single base-pair change. The classic example is **sickle-cell anaemia** — one base change in the β-globin gene replaces glutamate with valine.

Frameshift mutation: insertion or deletion of one or two bases shifts the entire reading frame from the point of change onward. Inserting/deleting bases in multiples of **three** adds or removes whole codons but leaves the downstream frame unchanged.

```
Original:  RAM HAS RED CAP
Insert B:  RAM HAS BRE DCA P     (frameshift)
Insert BIG: RAM HAS BIG RED CAP  (frame restored)
Delete R,E,D: RAM HAS CAP        (3 deleted ⇒ frame intact)
```

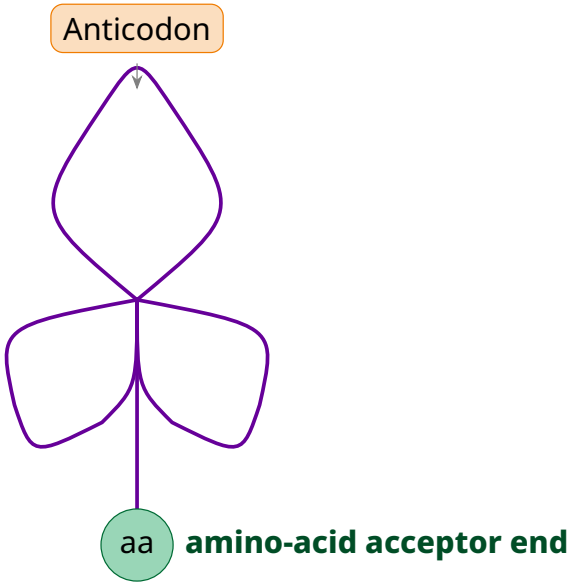
Frameshift logic

The genetic code is read three bases at a time with no punctuation. Add or remove a number *not* divisible by 3 and every codon after the change is misread. This is why insertions/deletions are usually far more damaging than substitutions.

5.3 tRNA — the Adapter Molecule

Crick predicted an **adapter molecule** that reads the code at one end and carries the matching amino acid at the other. That molecule is **tRNA**.

- An **anticodon loop** with bases complementary to the mRNA codon.
- An **amino-acid acceptor end** that binds a specific amino acid.
- A specific **initiator tRNA** for AUG; there are **no tRNAs for stop codons**.



The secondary structure of tRNA is the familiar **clover-leaf**; the actual 3-D shape is a compact **inverted-L**.

5.4 Translation

Translation is the polymerisation of amino acids into a polypeptide, in an order dictated by the mRNA. Amino acids are joined by **peptide bonds**.

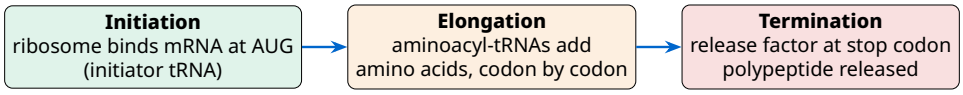
Charging of tRNA (aminoacylation)

$$\text{Amino acid} + \text{ATP} + \text{tRNA} \rightarrow \text{aminoacyl-tRNA} + \text{AMP} + \text{PP}_i$$

(enzyme: aminoacyl-tRNA synthetase)

Activating the amino acid first ("charging" the tRNA) makes the later peptide-bond formation energetically favourable.

The **ribosome** is the cellular factory — structural rRNAs $+ \approx 80$ proteins, in a **large** and a **small** subunit. The large subunit's **23S rRNA** (in bacteria) is the catalyst (a **ribozyme**) for peptide-bond formation.



UTRs

mRNA has **untranslated regions (UTRs)** before the start codon (5′-UTR) and after the stop codon (3′-UTR). They are not translated but are required for efficient translation.

Two roles of the ribosome — exam favourite

(1) It provides the **platform/site** that brings mRNA and aminoacyl-tRNAs together (two sites in the large subunit). (2) Its rRNA is the **catalyst (ribozyme)** for peptide-bond formation.

6 Regulation of Gene Expression: The *Lac* Operon

A cell does not make every protein all the time. **Gene expression is regulated** so that proteins are made only when needed — in *E. coli*, the lactose-digesting enzyme β-galactosidase is made only when lactose is present. In eukaryotes regulation can act at the transcriptional, processing, transport, or translational level; in prokaryotes the chief control point is **transcription initiation**.

6.1 The Operon Concept

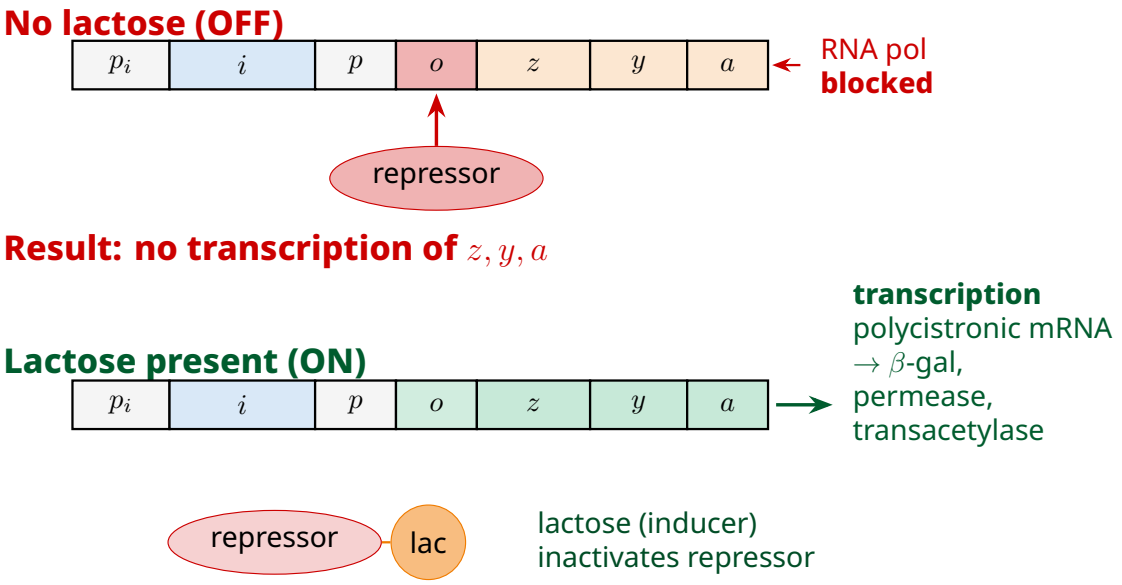
In bacteria, several related genes share a common promoter and regulatory genes — this unit is an **operon** (e.g. *lac*, *trp*, *ara*, *his* operons). The **operator** lies adjacent to the promoter; a **repressor** protein binds the operator to switch the operon off.

Components of the *lac* operon

Regulatory gene: *i* gene $\rightarrow$ codes the **repressor** (i = inhibitor, *not* inducer).
Structural genes:
z $\rightarrow$ β-galactosidase (hydrolyses lactose $\rightarrow$ galactose + glucose)
y $\rightarrow$ permease (increases cell permeability to β-galactosides)
a $\rightarrow$ transacetylase
All three are needed for lactose metabolism — a polycistronic mRNA.

6.2 Switching the Operon On and Off

The repressor is made constitutively (all the time) from the *i* gene. **Lactose** acts as the **inducer**.



Negative regulation of the *lac* operon

No lactose: repressor binds the operator → RNA polymerase blocked → operon **OFF**.
Lactose present: lactose (allolactose) binds and **inactivates** the repressor → operator free → RNA polymerase transcribes → operon **ON**.
Control by a repressor is **negative regulation**. (The *lac* operon also has positive regulation, beyond this level.)

Common Mistake

In "*lac* operon" the *i* gene stands for **inhibitor** (it codes the repressor), *not* inducer. Also, glucose or galactose **cannot** act as the inducer — only lactose/allolactose can. These are classic NEET distractors.

Substrate-regulated enzyme synthesis

The *lac* operon is regulated by the amount of its substrate (lactose). Seen this way, it is an elegant example of an enzyme system that is switched on only by the very molecule it acts upon — the cell never wastes energy making β-galactosidase when there is no lactose to digest.

7 Human Genome Project and DNA Fingerprinting

7.1 The Human Genome Project (HGP)

Launched in **1990**, the HGP aimed to sequence the entire human genome ($\approx 3\times10^9$ bp). At an early cost of about US$3 per bp the total estimate was ≈ US$9 billion. It was a 13-year mega project (coordinated by the US Dept. of Energy and the NIH,

with the Wellcome Trust and others), **completed in 2003**, and drove the rise of **Bioinformatics**.

Goals of HGP

- Identify all $\approx$ 20,000–25,000 genes in human DNA.
- Determine the sequence of the ≈ 3 billion base pairs.
- Store the information in databases; improve analysis tools.
- Transfer related technologies to industry.
- Address Ethical, Legal and Social Issues (**ELSI**).

Methodologies: (1) **Expressed Sequence Tags (ESTs)** — focus only on the expressed genes (mRNA); (2) **Sequence Annotation** — blindly sequence the whole genome (coding + non-coding) then assign function. DNA was fragmented, cloned in **BAC** (bacterial artificial chromosomes) and **YAC** (yeast artificial chromosomes), sequenced by **Sanger**'s method, and assembled by computer using overlapping regions.

Salient features of the human genome

- ≈ 3164.7 million bp; average gene ≈ 3000 bases (largest: dystrophin, 2.4 Mb).
- $\approx 30{,}000$ genes (far fewer than the earlier 80,000–140,000 estimate). 99.9% of bases identical in all humans.
- Functions unknown for $>$50% of genes; $<$2% of the genome codes for proteins.
- Repetitive sequences make up a very large part of the genome.
- Chromosome 1 has the most genes (2968); the Y chromosome the fewest (231).
- ≈ 1.4 million sites of single-base differences — **SNPs** (single-nucleotide polymorphisms).

7.2 DNA Fingerprinting

Since 99.9% of human DNA is identical, individuals differ at the remaining $\approx$ 0.1%. **DNA fingerprinting** is a quick way to compare those differences without sequencing the whole genome. It targets **repetitive DNA** called **satellite DNA** — short sequences repeated many times that show a very high degree of **polymorphism**.

VNTR and the technique

Alec Jeffreys developed DNA fingerprinting using a satellite DNA probe called **VNTR** (Variable Number of Tandem Repeats), a class of **mini-satellite**. Copy number varies greatly between individuals (size 0.1–20 kb), giving each person a characteristic band pattern. Steps (Southern blot method):

1. Isolation of DNA

2. Digestion by restriction endonucleases
3. Separation of fragments by electrophoresis
4. Blotting (transfer) to a nylon / nitrocellulose membrane
5. Hybridisation with a labelled VNTR probe
6. Detection of hybridised fragments by autoradiography

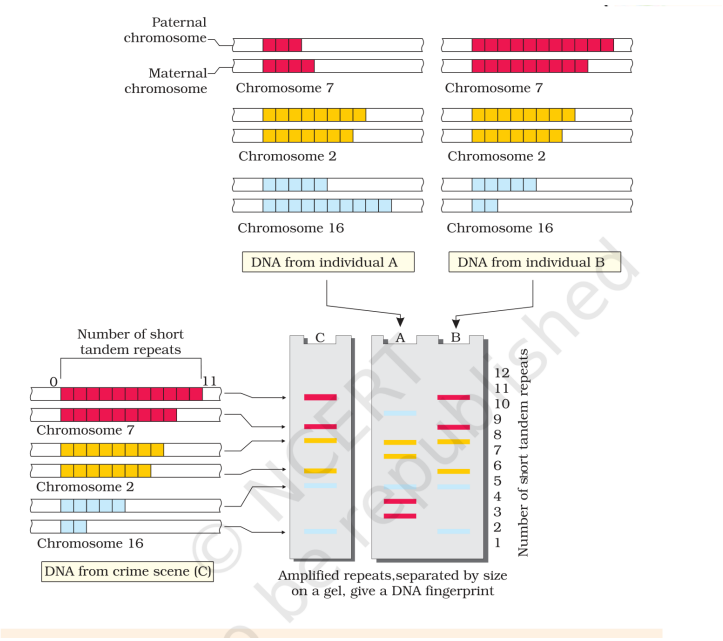


Fig 5.16 — Schematic of DNA fingerprinting. The banding pattern of crime-scene DNA (C) matches individual B, not A.

Why DNA fingerprinting works

Polymorphism (variation at genetic level) arises from mutations and is high in non-coding satellite DNA (mutations there have little effect on fitness, so they accumulate). VNTR copy number is inherited from parents, so the technique works for **paternity testing**, **forensics**, and population/genetic-diversity studies. The pattern differs between any two individuals *except identical (monozygotic) twins*. PCR (Chapter 9) makes it possible from DNA in a single cell.

Where you meet DNA fingerprinting

Crime-scene identification, settling disputed paternity/maternity, identifying disaster victims, verifying pedigree in conservation breeding, and tracing population ancestry — all rest on the same VNTR band-pattern logic.

Related Collegedunia Resources

Same chapter — other resources:
- NCERT Solutions
- Formula Sheet
- NCERT Book PDF
- Exemplar Book PDF
- Exemplar Solutions
- Handwritten Notes

Continue learning:
- Ch 2: Human Reproduction
- Ch 4: Principles of Inheritance and Variation
- Class 12 Biology — All Chapters

8 Quick Reference Summary

8.1 Key Numbers to Memorise

Quantity	Value
Distance between two consecutive base pairs	0.34 nm (0.34×10^{-9} m)
Pitch of the helix (one full turn)	3.4 nm
Base pairs per turn	≈ 10
Helix diameter	≈ 2 nm
Length of DNA in a human diploid cell	≈ 2.2 m (6.6×10^9 bp)
E. coli genome	4.6×10^6 bp
Human haploid genome	3.3×10^9 bp (≈ 3164.7 Mb total)
Polymerisation rate (*E. coli*)	≈ 2000 bp s^{-1}
bp in one nucleosome	≈ 200 bp
Number of human genes	$\approx 30{,}000$ (HGP)
Poly-A tail length	200–300 adenylates
HGP duration / cost	13 years, $\approx$ US$9 billion

8.2 Base Pairing and the Code

Must-know rules

Chargaff: $A = T$, $G = C$; $A + G = T + C$ (purines = pyrimidines).
H-bonds: A=T (2 bonds), G≡C (3 bonds).
Central Dogma: DNA $\rightarrow$ DNA (replication); DNA $\rightarrow$ RNA (transcription) $\rightarrow$ Protein (translation); reverse transcription: RNA $\rightarrow$ DNA.
Code: 64 codons; 61 sense + 3 stop (UAA, UAG, UGA); AUG = start = Met; triplet, degenerate, non-overlapping, comma-less, nearly universal.

Chargaff numericals

If cytosine = 20%, then $G = 20\%$, so $G + C = 40\%$, and $A + T = 60\%$, giving $A = T = 30\%$. Always: %A = %T and %G = %C; the four add to 100.

8.3 People and their Contributions

Scientist(s)	Contribution
Friedrich Meischer (1869)	First isolated DNA — called it 'Nuclein'
Watson & Crick (1953)	Double-helix model of DNA
Wilkins & Franklin	X-ray diffraction of DNA
Erwin Chargaff	Base-equivalence rule ($A = T$, $G = C$)
Griffith (1928)	Transforming principle
Avery, MacLeod, McCarty	DNA is the transforming principle
Hershey & Chase (1952)	DNA is the genetic material (phage expt)
Meselson & Stahl (1958)	Semiconservative replication
Taylor et al. (1958)	Semiconservative replication in *Vicia faba*
Gamow	Proposed the triplet code
Khorana, Nirenberg, Ochoa	Deciphered the genetic code
Alec Jeffreys	DNA fingerprinting (VNTR)
Sanger	DNA (and protein) sequencing method
Jacob & Monod	*Lac* operon model

One-line revision of the whole chapter

DNA is an anti-parallel, complementary double helix that **replicates** semiconservatively, is **transcribed** (one strand, a segment) into RNA, whose triplet **code** is **translated** by tRNA adapters on ribosomes into protein; expression is **regulated** (the *lac* operon as the model); the **HGP** sequenced the human genome and **DNA fingerprinting** exploits VNTR polymorphism for identity.

End of Chapter 5 — Molecular Basis of Inheritance. Revise the Quick Reference table the night before the exam.