collegedunia

Collegedunia NCERT Solutions

Molecular basis of inheritance ncert solutions: step-by-step coloured PDF for the 2026-27 NCERT (Latest Edition), Class 12th Biology Chapter 5

Chapter 5: Molecular Basis of Inheritance

About this Chapter

This chapter traces how scientists figured out that **DNA** is the genetic material and how the information it carries is read, copied and expressed inside a cell. You will study the **Watson and Crick double helix**, semi-conservative DNA replication, the central dogma flowing from DNA to RNA to protein, transcription in bacteria and eukaryotes, the genetic code, the **lac operon** and the Human Genome Project. By the end you can solve every NCERT exercise question for Class 12th Biology Chapter 5 with confidence.

Topics covered: DNA structure • Search for genetic material • Replication • Transcription • Genetic code • Translation • Regulation (lac operon) • Human Genome Project • DNA fingerprinting

Quick Formula Sheet

Chargaff's base-pairing rule:
$\%\text{A} = \%\text{T}, \quad \%\text{G} = \%\text{C}$
$\%\text{A} + \%\text{T} + \%\text{G} + \%\text{C} = 100\%$

Central dogma:
$\text{DNA} \xrightarrow{\text{transcription}} \text{RNA} \xrightarrow{\text{translation}} \text{Protein}$

DNA length (B-form):
$L = N \times 0.34\,\text{nm}$

Complementary strands run anti-parallel:
$5' \to 3'$ pairs with $3' \to 5'$

mRNA from coding strand:
same sequence, T $\Rightarrow$ U

Also see for this chapter: Revision Notes | Formula Sheet | Exemplar Solutions

Exercises

Q 5.1 **Group the following as nitrogenous bases and nucleosides: Adenine, Cytidine, Thymine, Guanosine, Uracil and Cytosine.**

SOLUTION

Concept used. A **nitrogenous base** is a single nitrogen-containing ring molecule that belongs to one of two families: the double-ring *purines* (adenine, guanine) and the single-ring *pyrimidines* (cytosine, thymine, uracil). When a nitrogenous base is attached through an N-glycosidic bond to a five-carbon sugar (*ribose* in RNA, *deoxyribose* in DNA), the base–sugar unit is called a **nucleoside**. If a phosphate group is further joined to the 5′ carbon of the sugar, the unit becomes a **nucleotide**.

☞ **Naming rule**

Nucleoside names end in *-osine* (for purines: adenosine, guanosine) or *-idine* (for pyrimidines: cytidine, thymidine, uridine). Plain base names (adenine, guanine, cytosine, thymine, uracil) end in *-ine*. The suffix tells you whether the sugar is already attached.

Step 1. **Step 1: classify each item by its ending.**

- Adenine (*-ine*) ⇒ free base.
- Cytidine (*-idine*) ⇒ base + ribose ⇒ nucleoside (cytosine + ribose).
- Thymine (*-ine*) ⇒ free base.
- Guanosine (*-osine*) ⇒ base + ribose ⇒ nucleoside (guanine + ribose).
- Uracil (*-il/-ine*) ⇒ free base (RNA pyrimidine).
- Cytosine (*-ine*) ⇒ free base.

Step 2. **Step 2: assemble the two groups.**

- **Nitrogenous bases:** Adenine, Thymine, Uracil, Cytosine.
- **Nucleosides:** Cytidine, Guanosine.

Step 3. **Step 3: sanity check.** Free bases must have no sugar attached. Adenine, Thymine, Uracil and Cytosine are all just ring molecules. Cytidine = cytosine + ribose; Guanosine = guanine + ribose. Both have a sugar attached, so they are nucleosides. Check passes.

Final Answer: Nitrogenous bases: **Adenine, Thymine, Uracil, Cytosine**. Nucleosides: **Cytidine, Guanosine**.

☞ **Exam Tip**

A one-mark trap on NEET: students see *cytidine* and *cytosine* together and tick both as bases. Remember the rule "*-idine/-osine* = sugar attached = nucleoside". A nucleotide additionally has a phosphate; you would recognise it by the prefix *adenosine mono/di/triphosphate* etc.

EXPERT'S SOLUTION *: Aanya Iyer, M.Sc Biotechnology, AIIMS Delhi*

Structural angle. Sort each name into a 3-tier ladder "*base* → *nucleoside* → *nucleotide*" and the answer falls out without memorisation.

Step 1. **Step 1: write the ladder.** Each level adds one chemical group:

$$\underset{\text{(ring only)}}{\textbf{base}} \xrightarrow{\textit{add ribose}} \underset{\text{(base + sugar)}}{\textbf{nucleoside}} \xrightarrow{\textit{add phosphate}} \underset{\text{(base + sugar + PO}_4\text{)}}{\textbf{nucleotide}}$$

Step 2. **Step 2: place every given name on the ladder.** Adenine, Thymine, Uracil and Cytosine carry no sugar and no phosphate, so they sit on the lowest rung (bases). Cytidine and Guanosine carry the sugar (their suffix gives them away), so they sit one rung up (nucleosides). None of the six has a phosphate, so the top rung stays empty here.

Step 3. **Step 3: read off the groups.** Bases (lowest rung): Adenine, Thymine, Uracil, Cytosine. Nucleosides (middle rung): Cytidine, Guanosine.

Why this matters. The same ladder lets you instantly classify ATP, GTP, AMP, dGMP and similar names that appear in the biotechnology chapters: count the chemical groups, place on the ladder, and you have the category.

> **Final Answer:** Bases: **Adenine, Thymine, Uracil, Cytosine**; Nucleosides: **Cytidine, Guanosine**.

Q 5.2 If a double stranded DNA has 20 per cent of cytosine, calculate the per cent of adenine in the DNA.

SOLUTION

Concept used. **Chargaff's rule of base pairing** states that in any double-stranded DNA molecule, adenine pairs only with thymine and guanine pairs only with cytosine. Because every A on one strand faces a T on the other, and every G faces a C,

$$\%\,\mathrm{A} = \%\,\mathrm{T} \quad \text{and} \quad \%\,\mathrm{G} = \%\,\mathrm{C}.$$

All four bases together account for every nucleotide in the molecule, so

$$\%\,\mathrm{A} + \%\,\mathrm{T} + \%\,\mathrm{G} + \%\,\mathrm{C} = 100\%.$$

This is the same equivalence rule that confirmed the double-helix hypothesis when Watson and Crick built their model in 1953.

☞ **Why A=T and G=C**

A and T form two hydrogen bonds; G and C form three. Two-bond pairs fit between two strands of a helix at the same width as three-bond pairs because A (purine) is paired with T (pyrimidine), and G (purine) with C (pyrimidine). So the helix has uniform width and the base counts must match in pairs.

Step 1. **Step 1: write down what is given.**

$$\%\,\mathrm{C} = 20\%.$$

We are asked to find $\%\,\mathrm{A}$.

Step 2. **Step 2: use $\%\,\mathrm{G} = \%\,\mathrm{C}$.** From Chargaff's rule,

$$\%\,\mathrm{G} = \%\,\mathrm{C} = 20\%.$$

Step 3. **Step 3: add up the GC content.**

$$\%\,\mathrm{G} + \%\,\mathrm{C} = 20\% + 20\% = 40\%.$$

So the remaining bases (A and T together) must make up

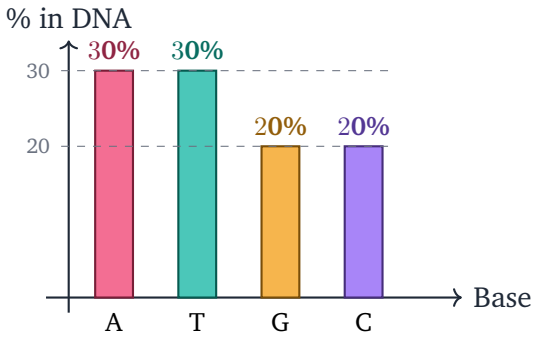
$$100\% - 40\% = 60\%.$$

Step 4. **Step 4: split the AT content equally.** Because $\%\,\mathrm{A} = \%\,\mathrm{T}$,

$$\%\,\mathrm{A} + \%\,\mathrm{T} = 60\% \;\Rightarrow\; 2 \times \%\,\mathrm{A} = 60\% \;\Rightarrow\; \%\,\mathrm{A} = \frac{60\%}{2} = 30\%.$$

Step 5. **Step 5: sanity check.**

$\%\,\mathrm{A} + \%\,\mathrm{T} + \%\,\mathrm{G} + \%\,\mathrm{C} = 30\% + 30\% + 20\% + 20\% = 100\%$. ✓



Base composition: A = T = 30%, G = C = 20%.

Final Answer: The percentage of adenine in the DNA is **30%**.

✗ Common Mistake

A frequent slip is to assume $\%\,\mathrm{A} = \%\,\mathrm{C}$ or to split the entire 100% four-way without first removing the GC fraction. Always: (i) double the given base, (ii) subtract from 100, (iii) divide the remainder by 2.

EXPERT'S SOLUTION *: Pranav Sharma, M.Sc Molecular Biology, NCBS Bangalore*

Quick reading. Chargaff's rule is really one equation in one unknown if you write it cleanly. Set $a = \%\,\mathrm{A}$ and use the constraint to solve directly.

Step 1. **Step 1: define the variable.** Let $a = \%\,\mathrm{A}$. Then by Chargaff's rule, $\%\,\mathrm{T} = a$, $\%\,\mathrm{C} = 20\%$ (given), and $\%\,\mathrm{G} = \%\,\mathrm{C} = 20\%$.

Step 2. **Step 2: write the conservation equation.** All bases together account for 100% of the molecule:

$$\%\,\mathrm{A} + \%\,\mathrm{T} + \%\,\mathrm{G} + \%\,\mathrm{C} = 100\%.$$

Substitute:

$$a + a + 20\% + 20\% = 100\%.$$

Step 3. **Step 3: solve.**

$$2a + 40\% = 100\% \;\Rightarrow\; 2a = 60\% \;\Rightarrow\; a = 30\%.$$

Step 4. **Step 4: verify against Chargaff.** $\%\,\mathrm{A}/\%\,\mathrm{T} = 30/30 = 1.00$ and $\%\,\mathrm{G}/\%\,\mathrm{C} = 20/20 = 1.00$. Both ratios equal unity, confirming the molecule is true double-stranded DNA.

Why this matters. The same one-variable approach handles every NCERT and entrance question of the form "given one base %, find another". Tie the variable to A or G, write the 100% constraint, solve.

Final Answer: $\%\,\mathrm{A} = \mathbf{30\%}$.

Q 5.3 **If the sequence of one strand of DNA is written as follows:**
5′ -ATGCATGCATGCATGCATGCATGCATGC-3′
Write down the sequence of complementary strand in $5' \to 3'$ direction.

SOLUTION

Concept used. A DNA double helix is built from two **antiparallel** strands. "Antiparallel" means that if one strand runs $5' \to 3'$ from left to right, the other strand runs $3' \to 5'$ over the same region. The bases on the two strands are joined by hydrogen bonds following

complementary base pairing:

$$A \leftrightarrow T \qquad \text{and} \qquad G \leftrightarrow C.$$

To write the complementary strand of any given sequence:

1. Replace each base in the given strand with its partner (A→T, T→A, G→C, C→G). This gives the partner sequence in the *antiparallel* direction.
2. Reverse the resulting sequence so that we read it in $5' \to 3'$ direction (because by convention every DNA sequence is written 5′ end first).

> ☞ **Antiparallel orientation**
>
> The 5′ end carries a free phosphate on the 5′ carbon of the sugar; the 3′ end carries a free –OH on the 3′ carbon. Two strands in a helix cannot both end the same way, so they must run opposite ways. This is why the partner sequence is reversed.

Step 1. **Step 1: write the given strand with the 5′ end clear.**

5′-A T G C A T G C A T G C A T G C A T G C A T G C A T G C-3′

That is 28 bases long (4 bases × 7 repeats of ATGC).

Step 2. **Step 2: write the partner base under each position.** Using A→T, T→A, G→C, C→G:

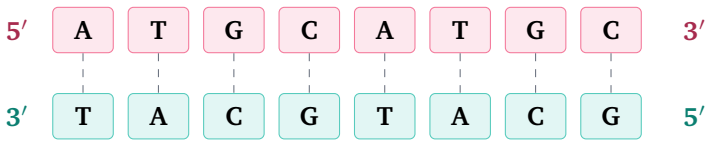
3′-T A C G T A C G T A C G T A C G T A C G T A C G T A C G-5′

Notice the partner strand is written $3' \to 5'$ because the leftmost base of the original was at 5′, so its partner is at 3′.

Step 3. **Step 3: reverse the partner strand to put it in $5' \to 3'$ direction.** Reading the line in Step 2 right-to-left, and writing it left-to-right gives:

5′-G C A T G C A T G C A T G C A T G C A T G C A T G C A T-3′

Step 4. **Step 4: sanity check.** Length: 28 bases, matches the input. ✓ First base of new strand = G, last base = T. The first base (G) pairs with the last base of the original (C, at the 3′ end), which is correct for antiparallel pairing. ✓



(only the first 8 of 28 bases shown; pattern repeats)

Final Answer: Complementary strand in $5' \to 3'$ direction:
5′ -GCATGCATGCATGCATGCATGCATGCAT-3′

✗ Common Mistake

A very common error is to swap each base (A→T etc.) but forget to *reverse* the sequence. That gives the partner read 3′→5′, which is the wrong convention. Always reverse before reporting.

EXPERT'S SOLUTION *: Rohit Bhat, Ph.D Molecular Genetics, NCBS Bangalore*

Picture-first. Imagine the two strands as a zip with the 5′ end of strand 1 lined up with the 3′ end of strand 2. Pair every tooth (base), then read strand 2 the other way around.

Step 1. **Step 1: chunk the input into the repeating motif.** The given strand is the 4-base motif ATGC repeated 7 times. Whatever rule we derive for one ATGC unit applies to all seven.

Step 2. **Step 2: complement and reverse one motif.** ATGC complemented is TACG (still 5′→3′ relative to partner-aligned position, i.e. written 3′→5′ here). Reverse: GCAT. So the complementary 4-base unit, in 5′→3′, is **GCAT**.

Step 3. **Step 3: stitch 7 copies of GCAT together.**

$$5'\text{-GCAT GCAT GCAT GCAT GCAT GCAT GCAT-}3'$$

Removing the spaces:

$$5'\text{-GCATGCATGCATGCATGCATGCATGCAT-}3'$$

Step 4. **Step 4: cross-check by base counts.** Original: 7 A, 7 T, 7 G, 7 C. Complement should also have 7 A, 7 T, 7 G, 7 C (because A on one strand becomes T on the other, etc.). Count in our answer: 7 G, 7 C, 7 A, 7 T. ✓

Why this matters. Spotting a repeating motif converts what looks like a 28-base problem into a 4-base problem. This trick saves real time in entrance exams when sequences are long.

Final Answer: **5′ -GCATGCATGCATGCATGCATGCATGCAT-3′.**

Q 5.4 **If the sequence of the coding strand in a transcription unit is written as follows:**

5′ -ATGCATGCATGCATGCATGCATGCATGC-3′

Write down the sequence of mRNA.

SOLUTION

Concept used. In a transcription unit, the two DNA strands play different roles. The **template strand** (running $3' \to 5'$) is actually read by RNA polymerase to build the mRNA. The other strand, running $5' \to 3'$ with a sequence identical to the mRNA (apart from T $\to$ U), is called the **coding strand** (or "sense" strand). Therefore the mRNA can be written directly from the coding strand using one rule only:

copy the coding strand as-is, but replace every T with U.

The $5' \to 3'$ direction is preserved because mRNA grows in the $5' \to 3'$ direction, the same direction in which the coding strand is conventionally written.

> ☞ **Why coding strand ≡ mRNA up to T/U**
>
> RNA polymerase reads the template strand $3' \to 5'$, adding ribonucleotides $5' \to 3'$. The new mRNA base is complementary to the template base, which is itself complementary to the coding-strand base. Two complements bring us back to the original, with the change that RNA uses uracil instead of thymine.

Step 1. **Step 1: identify what we already have.**

coding strand $= 5'$-ATGCATGCATGCATGCATGCATGCATGC-$3'$.

This is 28 nucleotides (4 bases $\times$ 7 repeats).

Step 2. **Step 2: apply the T $\to$ U substitution while keeping every other base unchanged.** Going base by base: A stays A, T $\to$ U, G stays G, C stays C. So ATGC becomes AUGC, and the whole strand becomes:

$5'$-AUGCAUGCAUGCAUGCAUGCAUGCAUGC-$3'$.

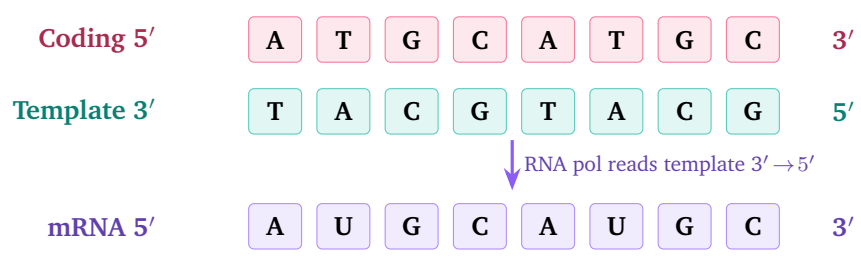
Step 3. **Step 3: write the answer in standard mRNA direction.** mRNA is always reported $5' \to 3'$ (the direction in which a ribosome reads it). Our sequence is already in that direction, so no reversal is needed.

Step 4. **Step 4: optional cross-check via the template strand.** Template strand = complement of coding strand, written antiparallel:

$3'$-TACGTACGTACGTACGTACGTACGTACG-$5'$.

Reading this template $3' \to 5'$ and writing the complementary RNA base for each: T $\to$ A, A $\to$ U, C $\to$ G, G $\to$ C, ... gives exactly

$5'$-AUGCAUGCAUGCAUGCAUGCAUGCAUGC-$3'$. ✓



Coding ↔ template ↔ mRNA: the mRNA copies the coding strand and replaces every T with U.

Final Answer: mRNA sequence: **5′ -AUGCAUGCAUGCAUGCAUGCAUGC-3′**

EXPERT'S SOLUTION *: Sneha Reddy, M.Sc Biotechnology, AIIMS Delhi*

Strategic angle. The shortest route from coding-strand to mRNA is a single character substitution: T → U. No reversing, no complementing. Everything else is a check.

Step 1. **Step 1: state the rule cleanly.** mRNA$(5' \to 3')$ = coding$(5' \to 3')$ with the single substitution T → U. This rule comes from the two complementations (coding ↔ template, then template → mRNA) cancelling out, except for the change in pyrimidine.

Step 2. **Step 2: do the substitution motif by motif.** The coding strand has 7 copies of ATGC. Each ATGC becomes AUGC, so the mRNA has 7 copies of AUGC:

$$5'\text{-AUGC AUGC AUGC AUGC AUGC AUGC AUGC-}3'.$$

Removing the spaces gives 28 nucleotides.

Step 3. **Step 3: confirm the reading frame.** Read in triplets from the 5′ end: `AUG | CAU | GCA | UGC | AUG | CAU | GCA | UGC | AU`. The first codon is the start codon AUG (methionine), as expected for a transcription unit whose coding strand starts with ATG.

Why this matters. In entrance papers, recognising that "mRNA = coding strand with T→U" converts almost every transcription question into a one-substitution exercise. Reach for the template strand only when the question explicitly hands you one.

Final Answer: mRNA = **5′ -AUGCAUGCAUGCAUGCAUGCAUGC-3′**.

Read the Full Class 12 Biology Chapter 5 Revision Notes →

Q 5.5 **Which property of DNA double helix led Watson and Crick to hypothesise semi-conservative mode of DNA replication? Explain.**

SOLUTION

Concept used. **Semi-conservative replication** is the model in which each daughter DNA molecule consists of one *parental* strand and one *newly synthesised* strand. The clue that pointed Watson and Crick towards this model lies in the geometry of the helix itself: the two strands of DNA are **complementary** (every A pairs with T, every G with C) and

they are joined only by weak hydrogen bonds. The complementarity means that if the two strands are gently pulled apart, each strand already carries the full information needed to rebuild its partner. The weak hydrogen bonds (in contrast to covalent backbone bonds) allow the strands to come apart without the molecule being destroyed.

Step 1. **Step 1: the key property – complementary base pairing.** Within the helix, the order of bases on one strand *determines* the order of bases on the other. There is no independent information on the two strands; they are mirror images of each other through the A–T and G–C rule.

Step 2. **Step 2: physical separability via hydrogen bonds.** The two strands are held together by 2 hydrogen bonds (A–T) or 3 (G–C). These are weaker than the covalent phosphodiester bonds along each backbone, so the strands can unwind without breaking either backbone. This makes strand separation a feasible step in replication.

Step 3. **Step 3: the implication Watson and Crick drew.** Because each strand carries the complete information for its partner, the molecule can be copied by: (i) unwinding the helix, (ii) using each parental strand as a template, (iii) laying down a new strand whose bases are complementary to the parental template. After one round of replication, each daughter molecule has one old strand and one new strand – hence "semi-conservative".

Step 4. **Step 4: experimental confirmation.** This hypothesis was later proved in 1958 by **Meselson and Stahl** using density-gradient centrifugation of $E.\ coli$ DNA labelled with ^{15}N, confirming the prediction beyond doubt.

Each daughter inherits one parental strand; the partner is built fresh – the essence of semi-conservative replication.

Final Answer: The **complementarity of the two strands**, joined by weak hydrogen bonds, allowed each strand to act as a template for a new partner, leading directly to the semi-conservative model.

♥ The bigger picture

Watson and Crick's two-paragraph 1953 paper hinted: "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying

mechanism for the genetic material." That single sentence launched molecular biology.

EXPERT'S SOLUTION *: Vivaan Joshi, Ph.D Molecular Biology, NCBS Bangalore*

Strategic angle. The question is really asking: "*what about the helix made copying it obvious?*" Frame the answer around two ideas, information mirroring and physical separability, and finish with the experiment that confirmed it.

Step 1. **Step 1: information mirroring.** Watson and Crick noticed that A pairs only with T and G only with C. This is a one-to-one mapping, so each strand uniquely determines the other. A copy of strand-1, paired with strand-2, is exactly the same DNA molecule. So the molecule *contains its own template*.

Step 2. **Step 2: physical separability.** The forces holding the two strands together are hydrogen bonds (2 in A–T, 3 in G–C). These are far weaker than the covalent backbone bonds, so a mild input of energy, provided by enzymes like helicase, can pry the strands apart without damaging them. A model that requires breaking the covalent backbone would not be biologically sensible.

Step 3. **Step 3: how the two ideas combine.** Put together: separate the two strands; use each as the template for a new strand; you get two daughter helices, each with one old and one new strand. This is what we now call semi-conservative replication.

Step 4. **Step 4: the experimental seal.** Meselson and Stahl (1958) grew *E. coli* on heavy nitrogen (^{15}N) so all DNA was heavy, then shifted them to light nitrogen (^{14}N). After one round of replication every molecule was of intermediate density (one heavy strand + one light strand), exactly as semi-conservative predicts. The conservative model would have given two density bands; the dispersive model would have given a slowly lightening single band over many generations. Both were ruled out.

Why this matters. The leap from helix *structure* to helix *function* (replication) is the textbook example of how a good structural model immediately suggests a mechanism. The same logic later guided the discovery of transcription and translation.

> **Final Answer:** The **complementary base-pairing of the two antiparallel strands**, held by weak H-bonds, made the semi-conservative model inevitable.

Q 5.6 **Depending upon the chemical nature of the template (DNA or RNA) and the nature of nucleic acids synthesised from it (DNA or RNA), list the types of nucleic acid polymerases.**

SOLUTION

Concept used. A **nucleic acid polymerase** is an enzyme that synthesises a nucleic-acid strand by adding nucleotides one at a time, using a pre-existing nucleic acid as a template. Two choices define the family of the enzyme:

1. **What kind of template** does the enzyme read – DNA or RNA?
2. **What kind of product** does the enzyme build – DNA or RNA?

With two choices per axis, we get $2 \times 2 = 4$ possible types, each with a distinct biological role.

Step 1. **Step 1: list all four template-product combinations.**

Template	Product	Enzyme		Biological role
DNA	DNA	DNA-dependent polymerase	DNA	DNA replication
DNA	RNA	DNA-dependent polymerase	RNA	Transcription
RNA	DNA	RNA-dependent polymerase	DNA	Reverse transcription (retroviruses)
RNA	RNA	RNA-dependent polymerase	RNA	Replication of RNA viruses

Step 2. **Step 2: describe each enzyme briefly.**

- **DNA-dependent DNA polymerase**: copies DNA into DNA during cell division so each daughter cell gets a full genome.
- **DNA-dependent RNA polymerase**: copies a DNA template into mRNA, rRNA or tRNA during transcription. In eukaryotes there are three sub-types (RNA pol I, II and III) for different RNA classes.
- **RNA-dependent DNA polymerase** (**reverse transcriptase**): used by retroviruses such as HIV to write their RNA genome back into DNA so it can integrate into the host genome.
- **RNA-dependent RNA polymerase** (RdRp): used by RNA viruses such as influenza, polio and SARS-CoV-2 to copy their RNA genome into more RNA.

Step 3. **Step 3: sanity check.** The central dogma adds the reverse-transcriptase arrow only after Temin and Baltimore's 1970 discovery; the RdRp arrow was added for RNA viruses. Together the four enzymes cover every information-transfer route between DNA and RNA actually seen in nature.

Final Answer: Four types: **DNA-dependent DNA polymerase, DNA-dependent RNA polymerase, RNA-dependent DNA polymerase (reverse transcriptase), and RNA-dependent RNA polymerase**.

☞ Exam Tip

Reverse transcriptase is the only enzyme on the list that goes "backwards" (RNA → DNA). Examiners love asking which enzyme breaks the original central dogma; it is this one.

EXPERT'S SOLUTION *: Aditi Nair, Ph.D Virology, NCBS Bangalore*

Picture-first. Draw the 2×2 table of (template, product) combinations. Each cell is one enzyme family with a specific job.

Step 1. **Step 1: build the 2×2 grid.** Rows = template (DNA, RNA). Columns = product (DNA, RNA). Every cell is a real enzyme found in nature.

Step 2. **Step 2: fill in the cells.**

- DNA → DNA: **DNA-dependent DNA polymerase** (replication, S-phase of cell cycle).
- DNA → RNA: **DNA-dependent RNA polymerase** (transcription, every gene expression event).
- RNA → DNA: **Reverse transcriptase** (retroviruses; also the basis of cDNA cloning).
- RNA → RNA: **RNA replicase / RdRp** (RNA viruses; not present in healthy human cells).

Step 3. **Step 3: connect to the central dogma.** The first two cells are the original Crick-1958 dogma. Reverse transcriptase added the RNA-to-DNA arrow in 1970. RdRp completes the picture by allowing RNA to copy itself, which is essential for the life cycle of every RNA virus.

Why this matters. Drugs that block reverse transcriptase (zidovudine, lamivudine) stop HIV from establishing infection. Drugs that block RdRp (remdesivir, molnupiravir) stop the replication of SARS-CoV-2. The taxonomy of polymerases drives modern antiviral drug design.

Final Answer: Four polymerases: **DNA-dep. DNA pol, DNA-dep. RNA pol, RNA-dep. DNA pol (reverse transcriptase), RNA-dep. RNA pol.**

Q 5.7 How did Hershey and Chase differentiate between DNA and protein in their experiment while proving that DNA is the genetic material?

SOLUTION

Concept used. The Hershey–Chase experiment (1952) used **radioactive isotope labelling** to tell apart the two main components of a virus – DNA and protein – as they entered a host bacterium. The trick relies on the chemistry of these two molecules:

- **DNA contains phosphorus but no sulphur** (phosphate is part of the backbone, PO_4^{3-}).
- **Protein contains sulphur but no phosphorus** (sulphur sits inside the amino acids cysteine and methionine).

So if you grow viruses in a medium containing radioactive sulphur (^{35}S), only their proteins become radioactive. If you grow viruses in a medium containing radioactive phosphorus (^{32}P), only their DNA becomes radioactive. By following the radioactivity, you follow either the protein or the DNA, but never both at once.

Step 1. **Step 1: prepare two separate batches of bacteriophage T2.**

- **Batch A:** grow viruses in medium with ^{35}S. Outcome: virus protein coat is labelled, DNA is not.
- **Batch B:** grow viruses in medium with ^{32}P. Outcome: virus DNA is labelled, protein coat is not.

Step 2. **Step 2: infect *E. coli* with each batch separately.** Allow viruses time to attach to the bacterial wall and inject their genetic material.

Step 3. **Step 3: shear the empty viral coats off the bacteria using a blender.** Whatever is inside the cell stays inside; whatever is outside (the empty coats) gets stripped from the bacterial surface.

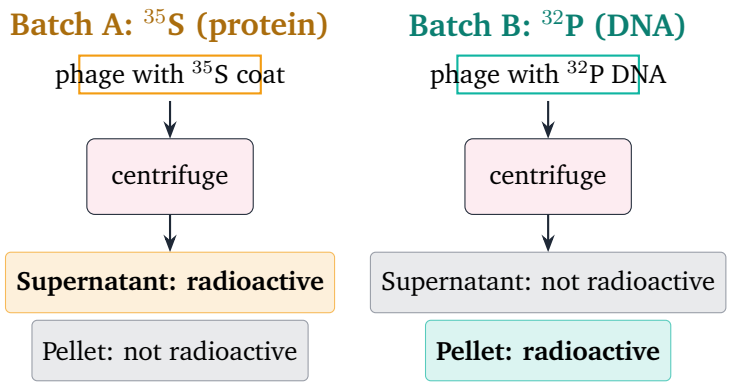
Step 4. **Step 4: centrifuge to separate.**

- Heavy bacteria sink to the pellet.
- Lighter, empty viral coats stay in the supernatant.

Step 5. **Step 5: measure radioactivity in pellet vs. supernatant.**

- Batch A (^{35}S, protein): radioactivity stays in the *supernatant* (the empty coats). Protein never entered the bacterium.
- Batch B (^{32}P, DNA): radioactivity ends up in the *pellet* (inside the bacteria). DNA did enter the bacterium.

Step 6. **Step 6: interpret.** Only the material that enters the host can be the genetic material (because the new viruses inside the bacteria came from that material). Since only DNA entered, **DNA is the genetic material**.



Only ^{32}P (DNA) gets into the bacteria. DNA is the genetic material.

Final Answer: Hershey and Chase used **^{35}S to tag protein** and **^{32}P to tag DNA**; only the ^{32}P entered the bacteria, proving **DNA, not protein, is the genetic material**.

EXPERT'S SOLUTION *: Diya Banerjee, Ph.D Molecular Biology, NCBS Bangalore*

Structural observation. The experiment is a clean "two controls in one" design: one isotope per molecule, and each isotope is chemically present in exactly one component, never both.

Step 1. **Step 1: pick chemically clean labels.** Sulphur appears in cysteine and methionine – both amino acids – but *nowhere* in DNA. Phosphorus appears in the DNA backbone phosphate but *nowhere* in the 20 standard amino acids. So ^{35}S unambiguously means "protein" and ^{32}P unambiguously means "DNA". No experimentally relevant overlap.

Step 2. **Step 2: grow the phage so the label is incorporated.** Bacteriophages cannot reproduce on their own, so you grow them inside bacteria cultured in ^{35}S-medium (Batch A) or ^{32}P-medium (Batch B). The new phages assemble using whatever atoms are in the medium, so they emerge with one or the other isotope built in.

Step 3. **Step 3: infect and shear.** Add the labelled phages to fresh, unlabelled *E. coli*. Allow attachment and injection. Then run the bacterium–phage mix in a kitchen blender for a few minutes – enough shear to knock the empty viral coats off the bacterial surface but not enough to lyse the bacteria.

Step 4. **Step 4: separate by centrifugation, count radioactivity.** Bacteria are heavy; they form a pellet. Free phage coats are light; they stay in the supernatant. Then run the pellet and the supernatant in a scintillation counter:

- Batch A radioactivity sits in the supernatant (protein stayed outside).
- Batch B radioactivity sits in the pellet (DNA went inside).

Step 5. **Step 5: conclude.** Whatever enters the host is what carries the genetic information, because the next generation of phages must be templated on it. Only DNA entered. So DNA is the genetic material.

Why this matters. Hershey–Chase finally closed the debate that started with Griffith (1928) and Avery–MacLeod–McCarty (1944). By using radioactivity instead of biochemistry, it produced an answer that even sceptics of the earlier experiments could not refute.

Final Answer: Different isotopes tag different molecules: ^{35}**S** $\rightarrow$ **protein,** ^{32}**P** $\rightarrow$ **DNA**; only the DNA isotope entered the bacteria.

Q 5.8 **Differentiate between the followings:**

(a) Repetitive DNA and Satellite DNA

(b) mRNA and tRNA

(c) Template strand and Coding strand

SOLUTION

Concept used. Each pair contrasts two related-but-distinct molecules or sequences in molecular biology. To differentiate them cleanly we list, for each pair, the property that is unique to each member and the property they share, in a side-by-side table.

(a) Repetitive DNA vs. Satellite DNA.

Feature	Repetitive DNA	Satellite DNA
Definition	Any DNA sequence that occurs in multiple copies in the genome.	A *subset* of repetitive DNA in which the repeats are short and highly tandem, so they form a band of different density from bulk DNA.
Buoyant density	Same as bulk genomic DNA.	*Different* from bulk DNA – forms a separate "satellite" peak on a CsCl density gradient.
Function	Some repeats are functional (rRNA genes); many are non-coding.	Largely non-coding; serves structural roles (centromeres, telomeres).
Use in genetics	Genome organisation studies.	Forms the basis of **DNA fingerprinting** (VNTRs, minisatellites).

(b) mRNA vs. tRNA.

Feature	mRNA (messenger RNA)	tRNA (transfer RNA)
Role	Carries the coded message (codons) from DNA to the ribosome.	Brings the correct amino acid to the ribosome and reads the mRNA codon.
Size	Long; varies with the length of the gene (typically 500–5000 nt).	Short; about 73–90 nucleotides.
Shape	Linear, unfolded.	Folded into a **clover-leaf** secondary structure (L-shaped in 3D).
Key sites	Contains codons; start codon AUG, stop codons UAA, UAG, UGA.	Contains an **anticodon** loop (reads codon) and a 3′ CCA end that binds the amino acid.
Fate after translation	Degraded after a few rounds.	Re-used many times.

(c) Template strand vs. Coding strand.

Feature	Template strand	Coding strand
Polarity	$3' \to 5'$ (in the direction of transcription).	$5' \to 3'$.
Role in transcription	Read by RNA polymerase to build mRNA.	*Not* read by RNA polymerase; sits aside.
Sequence vs. mRNA	Complementary and antiparallel to mRNA.	Same sequence as mRNA, except T in DNA replaces U in RNA.
Other names	Antisense strand, non-coding strand.	Sense strand.

Step 1. **Step 1: spot the relationship in each pair.** (a) Satellite DNA is a *type* of repetitive DNA. (b) mRNA and tRNA are two separate RNA classes with non-overlapping roles in translation. (c) Template and coding are the two strands of one transcription unit, only one of which is actually read.

Step 2. **Step 2: highlight the discriminator.** For each pair, the single most useful discriminator is: density-gradient behaviour (a), structural shape (b), and polarity/role in transcription (c).

Final Answer: See tables above. Key discriminators:
(a) satellite DNA forms a separate density band; repetitive DNA is the broader category.
(b) mRNA carries codons; tRNA carries an anticodon and an amino acid.
(c) template is read $3' \to 5'$ to make mRNA; coding strand is the same as mRNA except T/U.

♥ DNA fingerprinting begins here

The reason satellite DNA matters beyond textbooks: short tandem repeats (a satellite subset) vary in copy number between individuals. That variation is what makes a DNA fingerprint unique – and is the foundation of forensic identification.

EXPERT'S SOLUTION *: Karan Verma, M.Sc Molecular Biology, JNU*

Quick reading. Instead of memorising long lists, anchor each pair on a single distinguishing experiment or property; everything else flows from there.

Step 1. **Step 1: anchor (a) on the CsCl gradient.** Spin total genomic DNA in caesium chloride. Most DNA settles in one band; satellite DNA settles in a clearly separate band because its base composition (often AT-rich) gives it a different density. "Satellite" literally refers to that offset band. So satellite DNA $\subset$ repetitive DNA, but the special density signature is the distinguishing test.

Step 2. **Step 2: anchor (b) on shape and job.** mRNA is a long messenger; tRNA is a small adapter with a clover-leaf fold. mRNA *is read*; tRNA *reads*. mRNA is one-time-use; tRNA is recycled. Pin every other difference on the message-vs-adapter dichotomy.

Step 3. **Step 3: anchor (c) on which strand RNA polymerase actually touches.** RNA polymerase binds the template ($3' \to 5'$) and writes mRNA ($5' \to 3'$) complementary to it. The coding strand sits idle on the side; its job is to be the *record* of what the mRNA will say (just with T instead of U). Knowing that "coding = mRNA except T/U" is enough to answer almost every exam question on this pair.

Why this matters. Each pair maps to a real experimental technique you will meet in the biotechnology chapters: density-gradient centrifugation for satellite DNA, ribosome-profiling for mRNA/tRNA function, and primer design for the coding strand.

Final Answer: Anchor each pair on one diagnostic property: density band (a), shape and recycling (b), polarity and reading by RNA pol (c).

Q 5.9 List two essential roles of ribosome during translation.

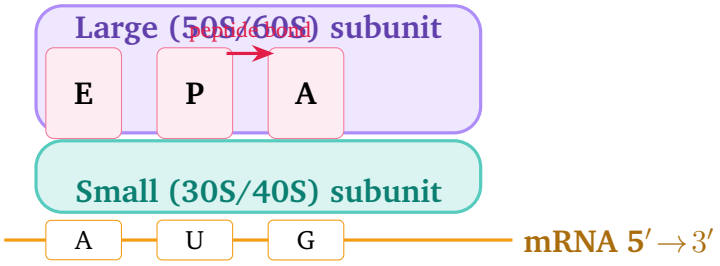
SOLUTION

Concept used. The ribosome is the molecular machine that builds a polypeptide chain from an mRNA template. It is made of rRNA (ribosomal RNA) and ribosomal proteins, and it has two subunits – a small subunit (30S in bacteria, 40S in eukaryotes) and a large subunit (50S/60S). Inside the ribosome there are three tRNA-binding pockets called the A, P and E sites. During translation the ribosome performs two distinct, essential tasks.

Step 1. **Role 1: holds the mRNA, the codon, and the matching tRNAs together in space.** The small subunit binds the mRNA and slides along it codon by codon. The A site exposes the next codon to be read. Incoming aminoacyl-tRNAs base-pair their anticodon with that codon inside the ribosome's reading frame. The P site holds the tRNA carrying the growing peptide chain.

Without the ribosome, the mRNA codon and the matching tRNA anticodon would diffuse in solution and would not stay paired long enough for chemistry to happen.

Step 2. **Role 2: catalyses the formation of peptide bonds between adjacent amino acids (the peptidyl transferase activity).** The peptide bond between the amino acid on the P-site tRNA and the amino acid on the A-site tRNA is formed by a catalytic site in the large subunit of the ribosome. That catalytic site is made entirely of rRNA, so the ribosome is a true ribozyme – an RNA enzyme. After the bond is made, the ribosome translocates to expose the next codon, repeating the cycle until a stop codon is reached.



The ribosome holds mRNA and tRNAs in place (role 1) and catalyses peptide-bond formation in the large subunit (role 2).

Final Answer: **Role 1:** hold the mRNA and two adjacent tRNAs together so that codon–anticodon pairing happens precisely.
Role 2: catalyse peptide-bond formation between adjacent amino acids (peptidyl transferase activity of the large subunit's rRNA).

EXPERT'S SOLUTION *: Meera Pillai, M.Sc Biochemistry, AIIMS Delhi*

Strategic angle. The two roles map perfectly onto the two ribosomal subunits: the *small* subunit handles reading, the *large* subunit handles peptide-bond chemistry. Frame the answer around that split.

Step 1. **Step 1: assign the reading role to the small subunit.** The small subunit binds the mRNA. It positions one codon at a time over the A-site, where the matching aminoacyl-tRNA docks. Codon–anticodon base pairing (three H-bond contacts) is far too weak to hold tRNA on its own in the cytosol; the ribosome provides the structural scaffold that makes this transient interaction productive.

Step 2. **Step 2: assign the chemistry role to the large subunit.** Inside the large subunit there is a catalytic cleft at the "peptidyl transferase centre". Here the amine group of the A-site amino acid attacks the carbonyl of the P-site amino acid and a new peptide bond forms, lengthening the chain by one residue. Crystallography showed this active site is built entirely from rRNA, so the ribosome is a **ribozyme**.

Step 3. **Step 3: tie the two roles to the elongation cycle.** Each elongation cycle is: (i) tRNA reading (small subunit), (ii) peptide-bond formation (large subunit), (iii) translocation. Lose either role and translation halts.

Why this matters. Many antibiotics work by attacking exactly one of these two roles. Streptomycin blocks the small subunit's reading function; chloramphenicol blocks the large subunit's peptidyl transferase. Knowing the two roles is therefore not just biology – it is the foundation of antibacterial drug design.

> **Final Answer:** **(1)** Positioning mRNA and tRNAs for accurate codon–anticodon pairing (small subunit). **(2)** Catalysing peptide-bond formation (large subunit, peptidyl transferase activity).

Q 5.10 **In the medium where *E. coli* was growing, lactose was added, which induced the *lac* operon. Then, why does *lac* operon shut down some time after addition of lactose in the medium?**

SOLUTION

Concept used. The **lac operon** is a cluster of three structural genes (z, y, a) controlled together by a single promoter and operator, and a regulatory gene i that codes for the **repressor protein**. When lactose is absent, the repressor sits on the operator and blocks RNA polymerase from transcribing the operon. When *lactose* is present, a small amount of lactose is converted by basal β-galactosidase to **allolactose**, which acts as the

inducer. Allolactose binds the repressor, changes its shape, and pulls it off the operator – so transcription begins. This explains why lactose "switches on" the operon. The question is the opposite: once switched on, why does the operon *switch off* again after some time?

Step 1. **Step 1: identify what changes as the operon runs.** With the operon expressed, the structural genes are translated into three enzymes:

- $z \rightarrow$ **β-galactosidase**: cleaves lactose into glucose + galactose.
- $y \rightarrow$ **permease**: pumps more lactose into the cell.
- $a \rightarrow$ **transacetylase** (a transferase).

So as the operon runs, the lactose pool in the cell is rapidly being broken down by β-galactosidase.

Step 2. **Step 2: trace the consequence for the inducer.** Allolactose is itself derived from lactose (a side reaction of β-galactosidase). When lactose is consumed, the amount of allolactose also falls.

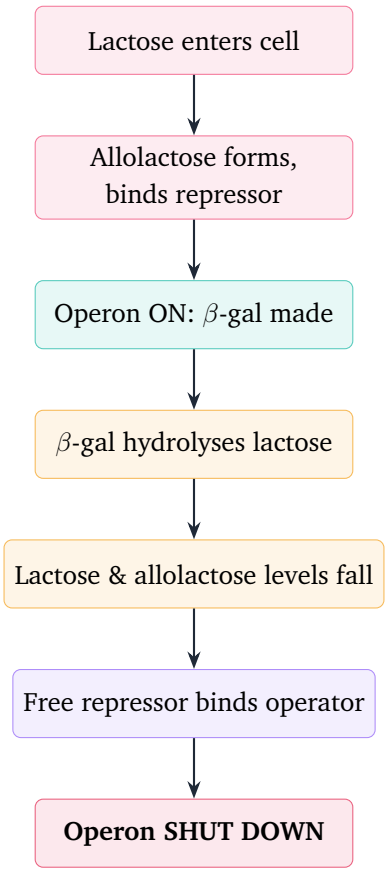
Step 3. **Step 3: link to the repressor.** With less allolactose around, the equilibrium

$$\text{repressor} + \text{allolactose} \rightleftharpoons \text{repressor-allolactose complex}$$

shifts to the left: free repressor reappears.

Step 4. **Step 4: shut-down step.** The free repressor binds back to the operator, physically blocks RNA polymerase, and transcription of the structural genes stops. Existing mRNAs are degraded quickly (mRNA half-life in bacteria is ~ 2 minutes), so within a few minutes the cell stops making the lac enzymes and the operon is back to its "off" state.

Step 5. **Step 5: name the kind of regulation.** Because lactose / allolactose is both the substrate *and* the inducer, the operon is autoregulated by negative feedback: the operon's output (enzymes) destroys the input (inducer), which turns the operon off. This is the classic example of **end-product (substrate) regulation**.



Final Answer: Once β-galactosidase digests away the lactose, the inducer (allolactose) is no longer present in enough quantity to keep the repressor off the operator. The repressor rebinds, blocks transcription, and the operon switches off, a negative-feedback, substrate-depletion shut-down.

✗ Common Mistake

A common error is to write that "glucose builds up and shuts down the operon" or "cAMP rises". Those describe catabolite repression by glucose, a separate mechanism. The direct reason the operon switches off after induction is simply that the inducer (lactose / allolactose) is consumed.

EXPERT'S SOLUTION *: Tara Desai, Ph.D Molecular Genetics, NCBS Bangalore*

Strategic angle. The lac operon is a feedback loop. As long as inducer is present, the loop runs. Remove the inducer and the loop collapses back to the resting (repressed) state. Tracking inducer concentration is therefore enough to answer the question.

Step 1. **Step 1: write the loop in three lines.** (a) Lactose $\rightarrow$ allolactose binds repressor $\rightarrow$ operon ON. (b) ON $\Rightarrow$ β-galactosidase made $\Rightarrow$ lactose broken down. (c) When lactose is depleted, allolactose drops, repressor is freed and

binds the operator ⇒ operon OFF.

Step 2. **Step 2: identify why this is inevitable, not accidental.** The cell makes the very enzyme that destroys its own inducer. So as soon as the operon's output rises (more β-galactosidase), the rate of inducer destruction rises too. Eventually inducer concentration drops below the binding threshold of the repressor and the operon switches off.

Step 3. **Step 3: estimate the timescale.** Bacterial mRNAs are degraded in 1–3 minutes. Once the repressor rebinds, new transcription stops within seconds; the existing mRNAs are gone within a few minutes; the existing enzymes degrade more slowly but are no longer replaced. So the operon can shut down in less than 10 minutes after lactose runs out.

Step 4. **Step 4: name the principle.** This is **negative feedback** via substrate consumption, the same principle used in many other inducible operons. It ensures the cell does not waste energy building enzymes for a nutrient that is no longer available.

Why this matters. The lac operon is the textbook example of how cells couple resource availability to gene expression – a strategy used throughout the bacterial world for sugars, amino acids and antibiotics.

> **Final Answer:** Because β-galactosidase consumes lactose (and hence allolactose), the inducer disappears, the repressor rebinds the operator, and the operon switches off.

Q 5.11 **Explain (in one or two lines) the function of the followings:**
(a) Promoter
(b) tRNA
(c) Exons

SOLUTION

Concept used. Each of these three terms names a different piece of the transcription–translation pipeline. We give a precise one-or-two-line function for each, and follow up with a one-line expansion of *why* that function matters.

Step 1. **(a) Promoter.** A **promoter** is a short DNA sequence located *upstream* of a gene (with consensus elements at positions -10, the Pribnow box, and -35 relative to the transcription start site in bacteria) that serves as the binding site for RNA polymerase. By providing this binding site, the promoter defines exactly where transcription will start and which of the two DNA strands will be used as the

template.

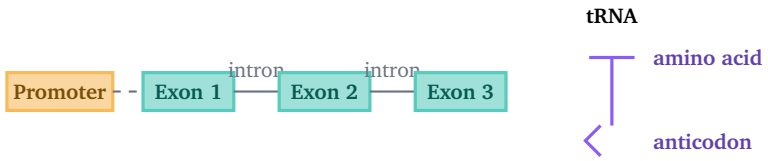
Why this matters. Different promoters have different strengths, so they control how often the gene downstream of them is transcribed. Mutations in a promoter can therefore increase or decrease gene expression without altering the protein-coding sequence at all.

Step 2. **(b) tRNA (transfer RNA).** A **tRNA** is a small ($\sim$ 73–90 nt) clover-leaf RNA that brings a specific amino acid to the ribosome and reads the mRNA codon via its **anticodon** loop, so that the correct amino acid is added next in the growing polypeptide.

Why this matters. Each of the 20 amino acids has at least one dedicated tRNA. Without tRNAs the codon table on the mRNA would have no way of being translated into a protein sequence – tRNAs are the physical embodiment of the genetic code.

Step 3. **(c) Exons.** **Exons** are the segments of a eukaryotic protein-coding gene whose sequences are kept in the mature mRNA after splicing. They contain the codons that ultimately get translated into the protein sequence.

Why this matters. Splicing alternative combinations of exons (alternative splicing) lets one gene produce multiple distinct proteins, which is one major reason why humans have only $\sim$ 20,000 genes but hundreds of thousands of distinct proteins.



Promoter (orange) starts transcription. Exons (green) survive splicing. tRNA (right) reads codons and carries the amino acid.

Final Answer: **(a)** Promoter: DNA site upstream of a gene where RNA polymerase binds; sets the start point of transcription.
(b) tRNA: clover-leaf RNA that carries an amino acid to the ribosome and reads the mRNA codon via its anticodon.
(c) Exons: the coding segments of a eukaryotic gene that are retained in the mature mRNA and translated into protein.

EXPERT'S SOLUTION *: Aditya Kumar, M.Sc Molecular Biology, JNU*

Strategic angle. Treat each term as one "component in the gene expression assembly line". Identify the step where it acts and the consequence of removing it.

Step 1. **(a) Promoter.** Position: just upstream of every gene. Acts at: *start* of

transcription. Function: gives RNA polymerase a place to bind. Remove it: no transcription, gene silent. Strong vs. weak promoters tune how often a gene is read.

Step 2. **(b) tRNA.** Position: cytoplasm, ribosome A/P/E sites. Acts at: *translation*. Function: brings amino acids matching each codon. Remove it: codons unreadable, translation halts. There are at least 31 distinct tRNAs covering all 61 sense codons in humans.

Step 3. **(c) Exons.** Position: within a eukaryotic gene, interspersed with introns. Acts at: *mRNA maturation* (splicing). Function: the portions kept in the mature mRNA. Remove all: no protein coding info. Alternative exon usage produces protein **isoforms** from one gene.

Why this matters. Promoter → tRNA → exon traces the entire path from DNA to protein in a eukaryote: deciding when to transcribe (promoter), maturing the mRNA (exons after splicing), and finally building the protein (tRNAs reading codons).

> **Final Answer:** Promoter = transcription start site; tRNA = codon reader and amino acid carrier; exons = coding pieces kept in the mature mRNA.

Q 5.12 Why is the Human Genome project called a mega project?

SOLUTION

Concept used. The **Human Genome Project (HGP)** was an international research programme (1990–2003) that aimed to determine the complete nucleotide sequence of the $\sim 3 \times 10^9$ base pairs of the human genome and to identify every gene in it. It is called a "mega project" not because of one feature but because of the scale, cost, duration, organisational scope, technological demands and societal impact – all of which were unprecedented for any single biology project up to that point.

Step 1. **Step 1: scale of data.** The human genome contains about 3.1×10^9 base pairs. Sequencing each pair, storing it, checking it and annotating it meant handling a dataset that, in 1990 computing terms, was enormous. If printed at 1000 letters per page in standard books, the data would fill about 3300 books of 1000 pages each.

Step 2. **Step 2: cost.** The total budget was approximately **US $ 9 billion**. Such a sum, contributed by multiple governments (mainly the US Department of Energy and

NIH, with UK, France, Germany, Japan and China), is in the same league as a space exploration programme.

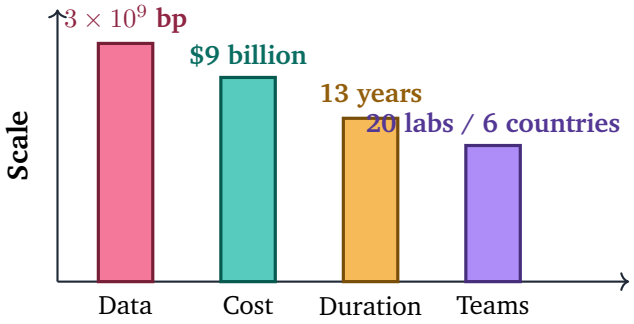
Step 3. **Step 3: duration.** The project ran for about **13 years**, with a working draft published in 2001 and the substantially complete sequence announced in 2003. A decade-plus commitment is unusual for any biology project; most biology grants run 3–5 years.

Step 4. **Step 4: technology demands.** HGP drove the development of **automated DNA sequencers** (capillary-based Sanger machines), bacterial artificial chromosome (BAC) cloning libraries, expressed sequence tag (EST) databases, and bioinformatics tools. The cost-per-base fell by orders of magnitude during the project.

Step 5. **Step 5: international coordination.** Twenty laboratories across six countries (US, UK, France, Germany, Japan, China) sequenced different chromosomes in parallel, with regular data deposition into the public GenBank database. Coordinating that effort – with shared standards, shared software and free open access to data – was itself a major achievement.

Step 6. **Step 6: societal and ethical impact.** For the first time, biological research had to be paired with formal ELSI (**Ethical, Legal and Social Implications**) programmes covering privacy, informed consent, gene patents and insurance discrimination. About 5% of the HGP budget was set aside for ELSI work.

Step 7. **Step 7: legacy.** The HGP made personalised medicine, comparative genomics, rapid cancer-genome sequencing, and the COVID-19 vaccine pipelines (which depend on fast viral-genome sequencing) possible. Its return-on-investment, conservatively estimated, runs into hundreds of billions of dollars.



HGP at a glance: vast data, very high cost, long duration and global teamwork.

Final Answer: Because the HGP combined an **unprecedented scale of data (3×10^9 bp), cost (US $9 billion), duration ($\sim$ 13 years), international coordination, technology development and societal/ethical impact**, no single feature but the combination of all of them earned it the "mega" label.

♥ The post-HGP era

After 2003, the cost of sequencing one human genome fell from $ 2.7 billion to under $ 1000 – a faster-than-Moore's-law collapse that turned HGP from a one-off achievement into routine clinical practice. Every cancer-genome panel and every COVID-variant surveillance report sits on top of HGP foundations.

EXPERT'S SOLUTION *: Yash Mehta, M.Sc Bioinformatics, IIT Madras*

Quick reading. "Mega" = many dimensions of magnitude. The right answer is a small bulleted scorecard, one bullet per dimension.

Step 1. **Step 1: list six dimensions of magnitude.** Data (~ 3.1 Gb); cost ($\sim$ \$9 billion); time (~ 13 years); manpower (thousands of scientists across six countries); technology (drove next-generation sequencing platforms); societal stakes (ELSI programme covering ethics and policy).

Step 2. **Step 2: cross-reference against earlier biology.** Before HGP, individual labs typically sequenced single genes of ~ 1000 bp. HGP scaled this up by a factor of $\sim 3 \times 10^6$. No previous biological project was even a tenth the size.

Step 3. **Step 3: cross-reference against engineering megaprojects.** At \$9 billion and over a decade, HGP sits at the same scale as the Hubble Space Telescope, the Channel Tunnel or the Large Hadron Collider – and it produced an open scientific deliverable used worldwide every day.

Step 4. **Step 4: summarise.** Because the project was unprecedented on six independent magnitudes at once, "mega project" is more than a label – it is an accurate technical description.

Why this matters. The HGP is the model that later international "mega" biology projects (ENCODE, the Cancer Genome Atlas, the Human Cell Atlas, the Earth BioGenome) deliberately copied.

> **Final Answer:** Mega in **data, cost, duration, manpower, technology and societal impact** – six independent magnitudes, all unprecedented.

Q 5.13 What is DNA fingerprinting? Mention its application.

SOLUTION

Concept used. DNA fingerprinting is a laboratory technique that identifies an individual from a tiny amount of biological sample by comparing highly variable, person-specific regions of the genome. The variability comes from short DNA sequences

called **VNTRs** (Variable Number Tandem Repeats) or **minisatellites**, which are part of the satellite DNA component of the genome. Each person has a unique combination of copy numbers at many VNTR loci – so unique that the chance of two unrelated people sharing the same pattern is vanishingly small. The technique was developed by **Alec Jeffreys** in 1984.

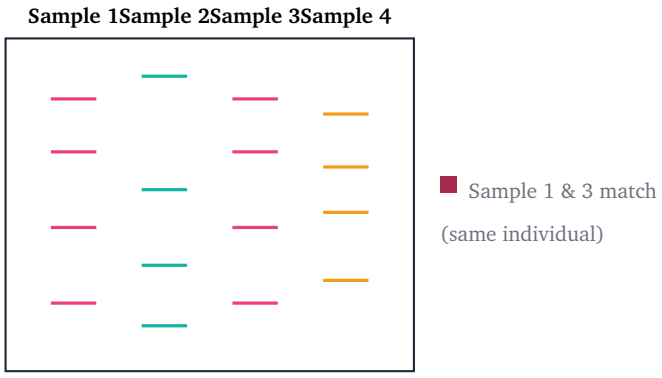
Step 1. **Step 1: isolate DNA from the sample.** Sources can be blood, semen, skin, hair root, saliva, bone – any tissue with nucleated cells.

Step 2. **Step 2: amplify and cut.** Either amplify the VNTR loci with the polymerase chain reaction (**PCR**), or digest the genomic DNA with a restriction enzyme that cuts *outside* the VNTR region. The fragment length then reflects the number of repeats in that region.

Step 3. **Step 3: separate by size.** Run the fragments through an agarose gel by electrophoresis; shorter fragments migrate further.

Step 4. **Step 4: detect with a probe.** Transfer the bands to a membrane (Southern blot) and probe with a radioactively labelled VNTR sequence to make the bands visible. The result is a ladder of bands – the *fingerprint*.

Step 5. **Step 5: compare.** Compare the fingerprints of the unknown sample and one or more reference samples. Identical fingerprints across many loci means the samples are from the same individual (or identical twins). Sharing half the bands on average means a parent–child relationship.



Representative DNA fingerprint gel: identical band patterns identify the same individual.

Applications:

- **Forensic science** – matching biological evidence at a crime scene (blood, semen, hair) to a suspect.
- **Paternity / maternity testing** – a child shares half the bands with each biological parent.
- **Identification of victims** of mass disasters, plane crashes, fires.
- **Pedigree analysis** – tracing family trees, identifying smuggled wildlife, etc.
- **Population and evolutionary studies** – estimating genetic distance between populations, conservation genetics of endangered species.

> **Final Answer:** **DNA fingerprinting** is the identification of an individual by analysing person-specific variable regions of DNA (VNTRs) using PCR/restriction digestion + electrophoresis + DNA hybridisation. **Applications:** forensic identification, paternity testing, victim identification, pedigree analysis, population genetics.

EXPERT'S SOLUTION *: Ananya Chatterjee, M.Sc Forensic Science, AIIMS Delhi*

Strategic angle. The answer has two halves: a one-line definition ("what is it"), and a 3–5 item list ("what is it used for"). Make the definition mention VNTRs, electrophoresis and the uniqueness claim; make the application list concrete.

Step 1. **Step 1: define.** DNA fingerprinting is the identification of an individual from a small biological sample by analysing their unique pattern of VNTR (minisatellite) bands on an electrophoresis gel.

Step 2. **Step 2: justify uniqueness.** With ~ 10 VNTR loci, each having (say) 20 possible copy numbers in the population, the number of distinct fingerprints is roughly $20^{10} = 1.024 \times 10^{13}$ – about 10^4 times the human population. So mismatches are practically impossible by chance.

Step 3. **Step 3: list applications.**

1. Forensics – linking a suspect to a crime scene with the kind of certainty courts can rely on.
2. Paternity / maternity disputes.
3. Identification of disaster victims when conventional methods (dental records) fail.
4. Wildlife conservation – catching poachers, verifying pedigree of zoo animals.
5. Anthropology / population genetics – studying migration patterns and ancestry.

Why this matters. DNA fingerprinting has shifted criminal justice towards evidence-based prosecution and has exonerated many wrongly convicted prisoners (the Innocence Project in the USA has used it to overturn more than 300 wrongful convictions).

> **Final Answer:** Identification of an individual via unique VNTR band patterns; used in forensics, paternity, disaster victim ID, wildlife conservation and population genetics.

Q 5.14 Briefly describe the following:

(a) Transcription
(b) Polymorphism
(c) Translation
(d) Bioinformatics

SOLUTION

Concept used. Each of these four terms names one of the foundational ideas of molecular biology. We define each precisely, then briefly describe what happens in the process or why the phenomenon matters.

(a) Transcription. Transcription is the process by which a single-stranded mRNA molecule is synthesised on a DNA template, by the enzyme RNA polymerase. It has three phases:

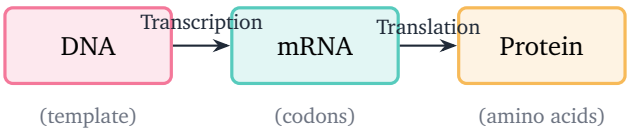
- **Initiation**: RNA polymerase binds the *promoter* of a gene and unwinds the local DNA.
- **Elongation**: the polymerase reads the template strand $3' \to 5'$ and builds a complementary mRNA $5' \to 3'$, using ribonucleoside triphosphates as substrates and following the rule A→U, T→A, G→C, C→G.
- **Termination**: at a terminator sequence the polymerase, DNA and nascent mRNA dissociate.

In eukaryotes the mRNA is further processed (5′ cap, 3′ poly-A tail, splicing of introns) before leaving the nucleus.

(b) Polymorphism. Polymorphism is the occurrence of two or more variants (alleles or sequence forms) of a DNA sequence in a population at a frequency $> 1\%$ for the rarer variant. Most polymorphisms are single-nucleotide polymorphisms (SNPs): a single base position where different individuals carry different bases (e.g. A in some, G in others). VNTRs (used in DNA fingerprinting) are a length polymorphism. Polymorphism is the raw material of genetic diversity within a species and is the basis for studying disease susceptibility, ancestry and evolution.

(c) Translation. Translation is the process by which the codon sequence of mRNA is decoded into the amino acid sequence of a polypeptide, on the ribosome. Each *tRNA* brings one amino acid that matches the codon being read at the ribosome's A-site; the ribosome's large subunit catalyses peptide-bond formation; the ribosome then translocates one codon along the mRNA. The chain starts at the AUG start codon and ends at a stop codon (UAA, UAG, or UGA). The new polypeptide then folds (often with the help of chaperones) into a functional protein.

(d) Bioinformatics. Bioinformatics is the application of computational tools – algorithms, databases and statistical models – to store, search and analyse biological data, especially DNA and protein sequences. Typical tasks include genome assembly, sequence alignment, gene prediction, protein-structure prediction, phylogenetic-tree construction and gene-expression analysis. Bioinformatics became indispensable during the Human Genome Project, and today drives drug discovery, COVID variant tracking and personalised medicine.



The central dogma: transcription makes mRNA from DNA; translation builds protein from mRNA.

Final Answer: **(a)** Transcription = DNA → mRNA by RNA polymerase (initiation → elongation → termination).
(b) Polymorphism = $> 1\%$ frequency of two or more sequence variants (SNPs, VNTRs) at a locus in a population.
(c) Translation = mRNA codons decoded into amino acids on the ribosome with tRNA adapters.
(d) Bioinformatics = computational analysis of biological sequence and structural data.

EXPERT'S SOLUTION *: Ishaan Kapoor, Ph.D Computational Biology, IISc Bangalore*

Picture-first. Place each term on the molecular-biology pipeline. (a) and (c) are two of the three core arrows of the central dogma; (b) describes the raw material those arrows act on; (d) is the analytical layer that sits on top of all the sequence data the other three generate.

Step 1. **(a) Transcription.** Think of it as "the gene being read aloud" from DNA into mRNA. RNA polymerase is the reader. Promoter = where to start reading. Template strand = the side actually read. Result = a single-stranded mRNA in $5' \to 3'$ direction. In eukaryotes the raw transcript is then capped, tailed and spliced before export to the cytoplasm.

Step 2. **(b) Polymorphism.** Think of it as "natural genetic variation that is common enough to matter". SNPs occur once every ~ 1000 bases on average in humans – that is roughly 10 million SNPs per individual. This variation underlies blood groups, disease susceptibility (diabetes, heart disease), drug response (pharmacogenomics) and ancestry tracing.

Step 3. **(c) Translation.** Think of it as "the mRNA being executed by the ribosome to build the corresponding protein". The ribosome reads three bases at a time (codon), an aminoacyl-tRNA matching that codon donates its amino acid, the new peptide bond forms, and the ribosome shifts to the next codon. Continues until a stop codon is hit.

Step 4. **(d) Bioinformatics.** Think of it as "biology done with a computer". Without it, the Human Genome Project's raw data would be unusable. With it, we can search GenBank for a homologous gene in seconds, predict a protein structure with AlphaFold, or assemble a viral genome from a nasal swab in hours.

Why this matters. Together these four ideas form the everyday toolkit of every working biologist today: you transcribe (a), you translate (c), you study natural variation in your population (b), and you let the computer crunch the data (d).

> **Final Answer:** Transcription = DNA → mRNA; Polymorphism = common DNA variation in a population; Translation = mRNA → protein; Bioinformatics = computational analysis of biological data.

Related Collegedunia Resources

Same chapter — other resources:

- Revision Notes
- Formula Sheet
- NCERT Book PDF
- Exemplar Book PDF
- Exemplar Solutions
- Handwritten Notes

Continue learning:

- Ch 10: Microbes in Human Welfare
- Ch 12: Ecosystem
- Class 12 Biology — All Chapters

Key Takeaways

- **DNA structure.** Double helix of two antiparallel strands held by hydrogen bonds; A–T (2 H-bonds) and G–C (3 H-bonds); Chargaff's rule: %A = %T and %G = %C.
- **DNA is the genetic material.** Hershey–Chase (1952) used ^{35}S (protein) and ^{32}P (DNA) labels to show that only DNA enters the bacterium during phage infection.
- **Replication is semi-conservative.** Each daughter DNA contains one parental strand and one new strand; proved by Meselson–Stahl using ^{15}N density gradients.
- **Central dogma.** DNA → RNA (transcription, by RNA polymerase) → Protein (translation, on the ribosome with tRNA adapters). Reverse transcription (RNA → DNA) is the exception used by retroviruses.
- **Coding vs. template strand.** Coding strand = mRNA sequence with T in place of U; template strand = actual strand read by RNA polymerase $3' \to 5'$.
- **Ribosome.** Two-subunit ribozyme: small subunit reads the mRNA codon; large subunit catalyses peptide-bond formation (peptidyl transferase activity of the rRNA).
- **lac operon.** Lactose → allolactose → binds repressor → operon ON. β-galactosidase consumes lactose → allolactose falls → repressor rebinds → operon OFF. Classic negative-feedback induction.
- **Human Genome Project (1990–2003).** $\sim 3 \times 10^9$ bp, \$ 9 billion, 13 years, 6 countries – a mega project in every dimension; foundation of modern bioinformatics and personalised medicine.
- **DNA fingerprinting.** VNTR (minisatellite) band patterns are unique to each individ-

ual; used in forensics, paternity testing, disaster victim ID, wildlife conservation and population genetics.

End of Exercises. For more Class 12 Biology resources, visit **collegedunia.com/exams/ncert**.