DNA Replication & Gene Expression

IB HL Study Guide

Contents

How to Use This Guide

- **DNA Structure** — A2.1 nucleotide components, double helix, base pairing, antiparallel strands
- **DNA Replication** — A2.2 semi-conservative replication, enzymes, leading and lagging strand synthesis
- **Transcription & Translation** — A2.3 gene expression from DNA to protein, codon table, ribosome function
- **Gene Expression & Epigenetics** — A2.4 regulation of gene expression, lac operon model, DNA methylation, histone modification
- **HL / AHL Only** — deeper enzyme detail, Okazaki fragments, lac operon, miRNA regulation
- **MCQ Practice** — styled like real IB Paper 1 questions
- **Exam Alerts** — the traps and mistakes that cost marks

Aligned to IB Biology 2025 syllabus — A2.1 DNA Structure, A2.2 DNA Replication, A2.3 Transcription & Translation, A2.4 Gene Expression

Before diving in, here is the big picture: **DNA is a master blueprint stored in the nucleus**. When a cell needs to copy itself, it duplicates the blueprint precisely (replication). When a cell needs to build something, it reads a section of the blueprint and makes a working copy (transcription), then uses that working copy as instructions to build a protein (translation). Gene expression is the control system that decides which sections of the blueprint get read, when, and in which cells.

Jump to section: DNA Structure (A2.1) · DNA Replication (A2.2) · Transcription & Translation (A2.3) · Gene Expression & Epigenetics (A2.4) · MCQ Practice · Quick Reference

Videos on this page: DNA Replication · Transcription & Translation

Section 1: DNA Structure (A2.1)

DNA contains all the heritable information of a cell, but this only works because its structure is extraordinarily stable and remarkably precise. Understanding the structure explains every other topic in this section — replication depends on base pairing, transcription depends on the antiparallel strands, and gene regulation depends on the accessibility of the double helix.

1.1 The Double Helix

DNA is a **double-stranded helix** — two polynucleotide chains wound around a common axis, first modelled by Watson and Crick in 1953 using X-ray diffraction data from Franklin and Wilkins.

Key structural features:

- Two **antiparallel** strands (one runs $5' \to 3'$; the other runs $3' \to 5'$)
- The **sugar-phosphate backbone** forms the outer rails of the helix
- **Nitrogenous bases** project inward from each strand and form pairs across the helix
- The helix makes one full turn every **10 base pairs**
- The **major groove** and **minor groove** alternate along the outside — transcription factors bind here

MEMORISE THIS

Key structural terms to memorise:

Term	Definition
Nucleotide	Monomer of DNA; consists of a deoxyribose sugar, a phosphate group, and a nitrogenous base
Antiparallel	The two strands of DNA run in opposite directions ($5' \to 3'$ and $3' \to 5'$)
Complementary base pairing	A pairs with T (2 hydrogen bonds); G pairs with C (3 hydrogen bonds)
Phosphodiester bond	Covalent bond linking the $3'$ carbon of one nucleotide to the $5'$ carbon of the next
Hydrogen bond	Weak bond holding the two complementary strands together; easily broken during replication
Double helix	The three-dimensional shape of DNA: two antiparallel strands coiled around each other
Base pair	One complementary pair across the helix (A—T or G—C)

1.2 Nucleotides and Base Pairing

Each **nucleotide** in DNA consists of three covalently bonded components:

1. A **deoxyribose** sugar (5-carbon; lacks the OH group at $2'$ carbon, unlike ribose in RNA)
2. A **phosphate group** (attached to the $5'$ carbon of the sugar)
3. A **nitrogenous base** — one of four types:
    - **Purines** (double ring): Adenine (A), Guanine (G)
    - **Pyrimidines** (single ring): Thymine (T), Cytosine (C)

Chargaff's Rules — from measurements of base composition in many organisms:

- $[\text{A}] = [\text{T}]$ and $[\text{G}] = [\text{C}]$ in any DNA molecule
- Therefore: $[\text{A}] + [\text{G}] = [\text{T}] + [\text{C}]$ (purines = pyrimidines)
- The ratio $\frac{[\text{A}]+[\text{T}]}{[\text{G}]+[\text{C}]}$ varies between species — this is species-specific

Why 3 H-bonds for G—C? Guanine and cytosine have three complementary groups that form hydrogen bonds, whereas adenine and thymine form only two. This makes G

—C pairs stronger. DNA with a higher G—C content has a higher **melting temperature** (T_m).

⚠ EXAM ALERT

Exam Alert: A pairs with T (NOT U — that is RNA). G pairs with C. A—T has **2** hydrogen bonds; G—C has **3** hydrogen bonds. If you are given a percentage for one base, you can calculate all others using Chargaff's rules. For example: if A = 30%, then T = 30%, so G + C = 40%, meaning G = C = 20%.

WORKED EXAMPLE

Worked Example: Chargaff's Rules Calculation

A DNA sample contains 22% guanine. Calculate the percentage of each of the other three bases.

Step 1: By Chargaff's rules, $[\mathrm{G}] = [\mathrm{C}]$, so C = 22%.

Step 2: G + C = 22 + 22 = 44%, so A + T = 100 - 44 = 56%.

Step 3: By Chargaff's rules, $[\mathrm{A}] = [\mathrm{T}]$, so A = T = 56 / 2 = **28%**.

Answer: A = 28%, T = 28%, G = 22%, C = 22%.

Check: 28 + 28 + 22 + 22 = 100%. Correct.

Section 2: DNA Replication (A2.2)

Before any cell can divide, it must copy its entire genome so that each daughter cell receives a complete set of instructions. Replication is extraordinarily accurate — the error rate is approximately 10^{-9} per base pair — and this accuracy relies on a team of coordinated enzymes acting at the replication fork.

2.1 Semi-Conservative Replication

DNA replication is **semi-conservative**: each new double helix consists of one original (parental) strand and one newly synthesised strand. This was proven by the **Meselson-Stahl experiment** (1958), which used heavy (^{15}N) and light (^{14}N) isotopes of nitrogen to track DNA strands across generations.

```
flowchart TD
    A["Original DNA\n(Both strands parental)"] --> B["Heli
    B --> C["Two template strands\nexposed at replication
    C --> D["Primase adds\nRNA primer (5'→3')"]
    D --> E["DNA Pol III extends\nnew strand (5'→3')"]
    E --> F["Leading strand —\ncontinuous synthesis"]
    E --> G["Lagging strand —\ndiscontinuous synthesis\n(O
    F --> H["DNA Pol I replaces\nRNA primers with DNA"]
```

```
    G --> H
    H --> I["DNA Ligase seals\nnicks between fragments"]
    I --> J["Two identical\ndouble-stranded DNA molecules"
```

Each origin of replication creates a **replication bubble**: two replication forks moving in opposite directions away from the origin. Eukaryotes have **multiple origins of replication** per chromosome (essential given chromosome size); prokaryotes typically have one origin (oriC).

2.2 Enzymes of Replication

MEMORISE THIS

Enzymes and proteins of DNA replication:

Enzyme / Protein	Function
Helicase	Unwinds and separates the double helix by breaking hydrogen bonds between base pairs at the replication fork
Single-strand binding proteins (SSBPs)	Stabilise the separated template strands, preventing re-annealing
Primase	Synthesises a short RNA primer (5—10 nucleotides) on the template strand to provide a free $3'$-OH for DNA polymerase
DNA polymerase III HL	Main replicative polymerase; adds nucleotides to the $3'$ end of the growing strand (can only extend, not start); also has $3' \to 5'$ proofreading exonuclease activity
DNA polymerase I	Removes RNA primers (has $5' \to 3'$ exonuclease activity) and replaces them with DNA nucleotides
DNA ligase	Seals the **nick** (phosphodiester bond) between the newly synthesised DNA and the adjacent fragment, joining Okazaki fragments
Topoisomerase	Relieves torsional stress (positive supercoiling) ahead of the replication fork by cutting and re-joining DNA

2.3 Leading and Lagging Strand Synthesis HL

Because DNA polymerase can only synthesise new DNA in the $5' \to 3'$ direction, the two template strands are replicated differently:

Leading strand (template runs $3' \to 5'$):

- DNA Pol III synthesises continuously in the $5' \to 3'$ direction
- Only one primer needed per origin
- Synthesis moves toward the replication fork (same direction as fork movement)

Lagging strand (template runs $5' \to 3'$):

- Must be replicated in short fragments, each starting with a new RNA primer

- These fragments are called **Okazaki fragments** (~100—200 nt in eukaryotes; ~1000—2000 nt in prokaryotes)
- Each Okazaki fragment is synthesised $5' \rightarrow 3'$, but moves away from the replication fork
- After synthesis: Pol I removes primers, fills gaps; Ligase joins fragments

⚠ EXAM ALERT

Exam Alert: DNA polymerase works $5' \rightarrow 3'$ ONLY. The leading strand is synthesised toward the fork; the lagging strand is synthesised AWAY from the fork in Okazaki fragments. Both strands use the SAME fork but have different topologies. A classic exam trap: "synthesis is continuous on both strands" — only the leading strand is continuous.

WORKED EXAMPLE

Worked Example: IB-Style Replication Question

A circular bacterial chromosome is 4.6 million base pairs long. It replicates from a single origin of replication. Identify two enzymes active at the replication fork and state the function of each. (4 marks)

Model answer:

1. **Helicase** — breaks hydrogen bonds between complementary base pairs, unwinding the double helix to expose the template strands. (2 marks)

2. **DNA polymerase III** — adds deoxyribonucleotides to the $3'$-OH end of the growing strand in a $5' \rightarrow 3'$ direction, using the template strand as a guide by complementary base pairing. (2 marks)

Note: You could also name primase (synthesises RNA primer to provide $3'$-OH start point) or DNA ligase (joins Okazaki fragments on lagging strand). Always state both the enzyme name AND its specific function.

IB TIP

IB Tip: For 6-mark questions on replication, examiners want: (1) semi-conservative model named, (2) unwinding by helicase, (3) primer by primase, (4) complementary base pairing with DNA Pol III, (5) leading/lagging strand distinction, (6) Okazaki fragments and ligase. Hitting all six points secures full marks.

▸**Watch: DNA Replication — Leading & Lagging Strands** VIDEO

Section 3: Transcription & Translation (A2.3)

Gene expression converts stored genetic information into functional proteins. This happens in two stages: **transcription** (DNA → RNA, in the nucleus) and **translation** (RNA → protein, at the ribosome). Together, these form the **central dogma of molecular biology**: DNA → RNA → Protein.

3.1 Transcription

Before the mechanics, here is the key idea: only one strand of DNA is used as a template for any given gene, and the RNA copy produced is a complementary, antiparallel copy of that template strand.

Stages of transcription:

1. **Initiation** — RNA polymerase binds to the **promoter** region upstream of the gene. The promoter contains consensus sequences (e.g. TATA box) that are recognised by transcription factors. HL
2. **Elongation** — RNA polymerase unwinds the DNA locally and moves $3' \rightarrow 5'$ along the template strand, synthesising a complementary mRNA strand in the $5' \rightarrow 3'$ direction. Uracil (U) pairs with adenine (A) in the template — RNA contains **no thymine**.
3. **Termination** — RNA polymerase reaches a terminator sequence and dissociates, releasing the pre-mRNA transcript.

Post-transcriptional processing (eukaryotes only): HL

- **5' cap** (methylated guanosine) is added — protects mRNA from degradation and aids ribosome attachment
- **Poly-A tail** (100—250 adenine nucleotides) added at $3'$ end — increases stability and aids export
- **Splicing** — introns (non-coding sequences) are removed by the **spliceosome**; exons (coding sequences) are joined together
- The mature mRNA is exported from the nucleus through nuclear pores to the cytoplasm

> ⚠ EXAM ALERT
>
> **Exam Alert:** Transcription uses RNA polymerase, NOT DNA polymerase. RNA uses Uracil instead of Thymine. The template strand is also called the **antisense strand**. The non-template strand (coding strand) has the same sequence as the mRNA (except T → U). Students frequently mix up which strand is read and in which direction.

3.2 Translation

Translation takes place at the **ribosome** in the cytoplasm (or on the rough endoplasmic reticulum for secreted proteins). It is the process of decoding the mRNA sequence into an amino acid sequence.

Key players:

> MEMORISE THIS
>
> **Translation vocabulary:**

Term	Definition
Codon	A triplet of three consecutive nucleotides on mRNA; codes for one amino acid (or start/stop signal)
Anticodon	The complementary triplet on tRNA that base-pairs with the mRNA codon
tRNA	Transfer RNA; carries a specific amino acid to the ribosome; has an anticodon loop that matches the mRNA codon
rRNA	Ribosomal RNA; structural and catalytic component of the ribosome; catalyses peptide bond formation
Start codon	AUG — codes for methionine; marks the beginning of the coding sequence
Stop codon	UAA, UAG, or UGA — no corresponding tRNA; signals termination of translation
Ribosome	Two subunits (small + large); has three tRNA-binding sites: A (aminoacyl), P (peptidyl), E (exit)
Peptide bond	Covalent bond formed between amino acids during elongation; catalysed by rRNA (peptidyl transferase)
Polypeptide	A chain of amino acids joined by peptide bonds; folds into a functional protein after translation

Ribosome subunits:

- Eukaryotes: 80S ribosome = 60S large subunit + 40S small subunit
- Prokaryotes: 70S ribosome = 50S large subunit + 30S small subunit

Stages of translation:

1. **Initiation** — the small ribosomal subunit binds to the 5′ cap of mRNA and scans for the start codon (AUG). Initiator tRNA (carrying methionine) binds at the P site. Large subunit joins.
2. **Elongation** — a charged tRNA enters the A site; its anticodon base-pairs with the mRNA codon. A peptide bond forms between the growing polypeptide (in P site) and the new amino acid (in A site). The ribosome translocates one codon in the 5′ → 3′ direction; the now-empty tRNA moves to the E site and exits.
3. **Termination** — a stop codon (UAA, UAG, or UGA) enters the A site; no tRNA matches. Release factors cause the ribosome to dissociate and the polypeptide is released.

Polyribosomes (polysomes): Multiple ribosomes can translate the same mRNA simultaneously, increasing the rate of protein production.

WORKED EXAMPLE

Worked Example: Decoding a Codon Table

A segment of the template DNA strand reads: `3'–TAC–CGT–ATT–ACT–5'`

(a) Write the sequence of the corresponding mRNA. (b) Using the codon table below, determine the amino acid sequence produced.

Codon	Amino Acid
AUG	Methionine (Met) — Start
GCA	Alanine (Ala)
UAA	Stop
UGA	Stop
GUG	Valine (Val)
UAU	Tyrosine (Tyr)

Step 1 (a): The template DNA strand is $3' \rightarrow 5'$, so transcription produces mRNA in the $5' \rightarrow 3'$ direction. Replace T with U and write the complementary bases:

Template (DNA): 3′-TAC-CGT-ATT-ACT-5′

mRNA: 5′-AUG-GCA-UAA-UGA-3′

Step 2 (b): Decode each codon:

- AUG = Methionine (start)
- GCA = Alanine
- UAA = Stop

Answer: The polypeptide is **Met-Ala** (translation stops at UAA; UGA is never read).

Key point: Translation always starts at AUG and stops at the first stop codon encountered.

EXAM ALERT

Exam Alert: mRNA carries codons; tRNA carries anticodons. The anticodon is complementary AND antiparallel to the codon. If mRNA codon = 5'-AUG-3', the anticodon = 3'-UAC-5' (written as 5'-CAU-3'). Students routinely forget the antiparallel orientation. Also: rRNA is the structural and catalytic component of the ribosome — it is NOT the RNA that carries the genetic message (that is mRNA) or the amino acid (that is tRNA).

Watch: Transcription & Translation VIDEO

Section 4: Gene Expression & Epigenetics (A2.4)

Not all genes are active in every cell at every time. A liver cell and a neuron contain identical DNA, yet they are structurally and functionally completely different —

because different genes are switched on. Gene expression regulation allows a single genome to produce hundreds of specialised cell types.

4.1 Regulation of Gene Expression

Gene expression can be regulated at multiple levels:

- **Transcriptional regulation** — controlling whether RNA polymerase transcribes a gene (most common and energetically efficient)
- **Post-transcriptional regulation** — controlling mRNA processing, stability, or translation (e.g. by miRNA) HL
- **Post-translational regulation** — modifying the protein after synthesis (e.g. phosphorylation, cleavage)

The lac Operon as a Model HL

The **lac operon** in *Escherichia coli* is the classic model for transcriptional regulation. It controls the expression of genes needed to metabolise lactose, and only switches them on when lactose is present AND glucose is absent.

Operon components:

- **Promoter** — site where RNA polymerase binds
- **Operator** — DNA sequence where the repressor binds (between promoter and structural genes)
- **Structural genes** — *lacZ* (β-galactosidase), *lacY* (permease), *lacA* (transacetylase)
- **Repressor gene** (*lacI*) — always expressed; produces the lac repressor protein

Regulation logic:

Condition	Repressor state	Transcription
No lactose, glucose present	Repressor binds operator	OFF (genes not needed)
Lactose present, no glucose	Allolactose (inducer) binds repressor → repressor changes shape → releases operator	ON (lactose needs to be metabolised)
Both lactose and glucose present	Repressor off operator, but low cAMP → weak transcription	Weak (cell prefers glucose)

cAMP and catabolite repression (when glucose is absent, cAMP is high, activating CAP protein that enhances transcription) ensures the *lac* genes are fully active only when lactose is the only carbon source available. This is **positive regulation** by CAP alongside **negative regulation** by the repressor.

> **IB TIP**
>
> **IB Tip:** The lac operon is a classic 6-mark question. Always describe: (1) what the repressor does by default (blocks transcription), (2) what the inducer is (allolactose,

> not lactose directly), (3) what happens when allolactose binds the repressor (shape change, release of operator), and (4) the result (RNA polymerase can transcribe). Full marks require mechanistic detail, not just “the genes switch on”.

Post-Transcriptional Regulation by miRNA HL

MicroRNA (miRNA) are small non-coding RNA molecules (~22 nucleotides) that regulate gene expression after transcription.

Mechanism:

1. miRNA genes are transcribed to form a long **primary miRNA (pri-miRNA)**
2. The pri-miRNA is cleaved in the nucleus by the enzyme **Drosha** into a ~70 nt hairpin structure called the **pre-miRNA**, which is exported to the cytoplasm
3. The pre-miRNA is processed by the enzyme **Dicer** into a short double-stranded RNA (~22 nt)
4. One strand is loaded into the **RISC** (RNA-induced silencing complex)
5. RISC uses the miRNA as a guide to find complementary sequences in the **3’ UTR of target mRNA**
6. RISC either **cleaves the mRNA** (if match is perfect) or **represses translation** (if match is partial)

Result: the target gene’s protein is NOT produced (or is produced at reduced levels), even though the gene was transcribed. This is called **RNA interference (RNAi)**.

miRNAs are crucial for development and cell differentiation — a single miRNA can regulate hundreds of target genes simultaneously.

4.2 Epigenetics

Epigenetics is the study of heritable changes in gene expression that do NOT involve changes to the DNA sequence itself. The DNA code is unchanged — what changes is how accessible or readable that code is.

The two major epigenetic mechanisms are:

1. DNA Methylation

- Addition of a methyl group ($-CH_3$) to cytosine bases, typically at CpG dinucleotides
- Methylation of a promoter region generally **silences** gene expression by preventing transcription factor binding
- Methylation patterns are heritable through cell division (maintained by methyltransferase enzymes)
- Example: X-chromosome inactivation (Barr body formation in mammalian females) is maintained by methylation

2. Histone Modification

- DNA is wrapped around **histone** proteins to form **nucleosomes** (fundamental unit of chromatin)
- Chemical groups can be added to the histone tails:
  - **Acetylation** (addition of acetyl groups) → loosens chromatin (**euchromatin**) → gene expression **ON**
  - **Deacetylation** → tightens chromatin (**heterochromatin**) → gene expression **OFF**
  - **Methylation of histones** can either activate or repress, depending on which histone residue is modified
- The combination of histone modifications across a region is called the **histone code**

MEMORISE THIS

Epigenetics key terms:

Term	Definition
Epigenetics	Heritable changes in gene expression not due to changes in DNA sequence
DNA methylation	Addition of $-CH_3$ to cytosine; typically silences gene expression
Histone	Protein around which DNA is wrapped; modifications alter chromatin accessibility
Nucleosome	Unit of chromatin: 8 histones + ~147 bp of DNA wound around them
Euchromatin	Loosely packed chromatin; genes are accessible and expressed
Heterochromatin	Tightly packed chromatin; genes are silenced
Acetylation	Addition of acetyl group to histone tail; associated with active gene expression
CpG island	Region of DNA with high CpG dinucleotide density; often found at promoters

EXAM ALERT

Exam Alert: Epigenetic changes are NOT mutations. A mutation changes the DNA sequence permanently (and randomly). An epigenetic change modifies how the DNA is packaged or read — the sequence itself is unchanged. Epigenetic changes are potentially reversible and can be influenced by environment (diet, stress, toxins). This distinction is a frequent MCQ distractor and is commonly worth 2 marks in Paper 2.

WORKED EXAMPLE

Worked Example: Epigenetics Question

Explain why two genetically identical twins may develop different diseases later in life. (3 marks)

Model answer:

Although identical twins have the same DNA sequence (same genotype), their epigenomes can diverge over time due to differences in environment (1). DNA methylation patterns may differ — for example, genes involved in disease susceptibility may be methylated (silenced) in one twin but not the other (1). Histone acetylation patterns may also differ, altering chromatin accessibility and changing which genes are actively transcribed (1).

Key principle: Same genotype, different epigenome → different phenotype (including disease risk).

MCQ Practice

Question 1

A DNA molecule contains 34% guanine. What is the percentage of adenine?

A. 34%

B. 32%

C. 16% ← CORRECT

D. 68%

Why: G = 34%, so by Chargaff's rules C = 34%. Therefore A + T = 100 − 68 = 32%, and since A = T, adenine = **16%**. Option A confuses G with A. Option B gives A + T combined. Option D gives G + C combined.

Question 2

During DNA replication, what is the function of DNA ligase?

A. Unwind the double helix at the replication fork

B. Synthesise new nucleotides in the $5' \rightarrow 3'$ direction

C. Seal the nicks between Okazaki fragments on the lagging strand ← CORRECT

D. Remove RNA primers from the template strand

Why: Ligase joins the $3'$ end of one fragment to the $5'$ end of the next by forming a phosphodiester bond — it seals the nick. Option A is helicase. Option B describes DNA Pol III (although ligase does form a bond, it is not synthesising new nucleotides). Option D is DNA Pol I.

Question 3

Which of the following is a difference between transcription and DNA replication?

A. Transcription uses deoxyribonucleotides; replication uses ribonucleotides

B. Transcription produces a single-stranded product; replication produces a double-stranded product ← CORRECT

C. Transcription requires a primer; replication does not

D. Transcription occurs in the cytoplasm; replication occurs in the nucleus

Why: mRNA is single-stranded; the replicated DNA is double-stranded. Option A has the types reversed — transcription uses ribonucleotides (to make RNA); replication uses deoxyribonucleotides. Option C has it backwards — replication requires a primer; transcription does not. Option D has it backwards for eukaryotes — both occur in the nucleus (transcription in nucleus; replication in nucleus).

Question 4

A ribosome is at a codon that reads 5'-UAA-3'. What happens next?

A. A tRNA carrying tyrosine enters the A site

B. The ribosome translocates one codon in the $3' \rightarrow 5'$ direction

C. Translation terminates and the polypeptide is released ← CORRECT

D. The mRNA is degraded immediately

Why: UAA is one of the three stop codons (UAA, UAG, UGA). No tRNA has a matching anticodon, so a **release factor** binds, triggering release of the polypeptide and dissociation of ribosome subunits. Option A is wrong — UAA codes for no amino acid (tyrosine codons are UAU and UAC). Option B gets the direction wrong (ribosomes move $5' \rightarrow 3'$ along mRNA). Option D is not triggered directly by a stop codon.

Question 5 HL

In the lac operon, what is the role of allolactose?

A. It directly activates RNA polymerase binding to the promoter

B. It acts as an inducer by binding the repressor and preventing it from binding the operator ← CORRECT

C. It methylates the operator region to allow transcription

D. It is the substrate for the enzyme encoded by the *lacZ* gene

Why: Allolactose (a metabolite of lactose) binds to the lac repressor protein, causing a conformational change that reduces the repressor's affinity for the operator. When the repressor is released from the operator, RNA polymerase can transcribe the structural genes. Option A is the role of CAP (with cAMP), not allolactose. Option C describes an epigenetic mechanism, not operon regulation. Option D confuses the regulatory role of allolactose with the metabolic substrate of the lacZ enzyme.

Question 6

Which statement about epigenetic changes is correct?

A. Epigenetic changes always involve a mutation in the DNA base sequence

B. Histone acetylation tightens chromatin and silences gene expression

C. DNA methylation of a promoter region typically reduces gene expression ← CORRECT

D. Epigenetic changes are never passed from parent cell to daughter cell

Why: Methylation of CpG sites in a promoter prevents transcription factor binding and typically silences the gene. Option A is the definition of a mutation — epigenetic changes do NOT alter the DNA sequence. Option B has the effect of acetylation reversed — acetylation loosens chromatin (euchromatin) and activates expression; it is deacetylation that silences. Option D is false — DNA methylation patterns are heritable and are maintained during cell division by methyltransferase.

Question 7

Which row correctly describes the differences between mRNA, tRNA, and rRNA?

	mRNA	tRNA	rRNA
A	Carries anticodon	Carries amino acid	Found in nucleus
B	Carries genetic code	Carries anticodon	Structural component of ribosome
C	Produced in cytoplasm	Carries codons	Catalyses peptide bonds
D	Single-stranded only	Double-stranded	Structural only

B is CORRECT ← CORRECT

Why: mRNA carries the genetic code (codons); tRNA carries the anticodon and the amino acid; rRNA is both structural and catalytic (peptidyl transferase activity). Option A assigns the anticodon to mRNA (wrong — that is tRNA). Option C says mRNA is made in the cytoplasm (wrong — transcription occurs in the nucleus). Option D says tRNA is double-stranded (wrong — tRNA is single-stranded but folds into a cloverleaf shape).

Quick Reference

DNA vs RNA: Key Differences

Feature	DNA	RNA
Sugar	Deoxyribose	Ribose
Bases	A, T, G, C	A, U, G, C
Strands	Double-stranded (usually)	Single-stranded (usually)
Location	Mainly nucleus	Nucleus and cytoplasm
Function	Long-term genetic storage	Gene expression intermediary
Stability	Very stable	Relatively unstable

Replication vs Transcription vs Translation

Feature	Replication	Transcription	Translation
Template	Both DNA strands	One DNA strand	mRNA
Product	DNA	RNA (mRNA/tRNA/rRNA)	Polypeptide
Enzyme	DNA Pol III	RNA polymerase	Ribosome (rRNA)
Location	Nucleus	Nucleus	Cytoplasm/RER
Direction	$5' \rightarrow 3'$	$5' \rightarrow 3'$ (product)	$5' \rightarrow 3'$ along mRNA
Primer needed?	Yes (RNA primer)	No	No

Mixed Practice — Exam Style

> **IB TIP**
>
> **How to use this section:** Unlike topic-specific practice, these questions are interleaved — they mix all topics from this guide in random order. Before answering, identify *which concept or topic area* the question is testing. This is exactly the skill you need on Paper 1 and Paper 2, where you don't know in advance which topic each question covers.

1. **[DNA Replication]** During DNA replication, DNA polymerase III can only add nucleotides in the $5' \rightarrow 3'$ direction. What is the consequence of this for the lagging strand?

   A. The lagging strand is synthesised continuously in one long piece

B. The lagging strand is synthesised in short fragments (Okazaki fragments) in the $5' \rightarrow 3'$ direction, away from the replication fork

C. The lagging strand is synthesised by RNA polymerase, not DNA polymerase

D. The lagging strand cannot be replicated and is lost after each division

2. **[Transcription]** During transcription, which strand of the DNA double helix serves as the template?

   A. Both strands are transcribed simultaneously

   B. The coding (sense) strand, read $3' \rightarrow 5'$, is transcribed

   C. The template (antisense) strand, read $3' \rightarrow 5'$ by RNA polymerase, is used to produce mRNA with the same sequence as the coding strand (with U instead of T)

   D. The choice of strand depends on the organism (prokaryote vs eukaryote)

3. **[Gene Regulation]** The lac operon in *E. coli* is repressed in the absence of lactose. When lactose is present, it acts as an inducer. The inducer molecule:

   A. Directly activates RNA polymerase to transcribe the structural genes

   B. Binds to the repressor protein, changing its shape so it can no longer bind to the operator, allowing transcription to proceed

   C. Degrades the repressor protein permanently

   D. Methylates the promoter region, permanently activating the operon

4. **[Translation — Codons]** The mRNA sequence $5'$-AUG-CGU-UAA-$3'$ codes for:

   A. Two amino acids followed by a stop codon (methionine, arginine, stop)

   B. Three amino acids (methionine, arginine, glutamine)

   C. One amino acid only (methionine is the start and stop codon)

   D. This sequence cannot be translated because UAA is a start codon

5. **[Mutations — Distractor]** A single base insertion (frameshift mutation) at position 3 of a coding sequence is generally more harmful than a single base substitution (point mutation) at the same position. The best explanation is:

   A. Insertions change the molecular mass of the protein more than substitutions

   B. Insertions cause a shift in the reading frame, altering every codon downstream and likely producing a non-functional protein; substitutions change

at most one amino acid, and the genetic code's degeneracy may mean the amino acid is unchanged

C. Substitutions are always silent mutations; insertions are always missense mutations

D. Insertions are more harmful only if they occur in an exon, not an intron

6. **[DNA Replication — Enzymes]** Helicase is to unwinding the double helix as _____ is to sealing nicks between Okazaki fragments.

   A. DNA polymerase I

   B. Primase

   C. DNA ligase

   D. Topoisomerase

7. **[Post-Transcriptional Modification]** In eukaryotes, pre-mRNA is processed before leaving the nucleus. Which of the following is NOT a step in this processing?

   A. Addition of a 5' methyl-guanosine cap

   B. Addition of a poly-A tail to the 3' end

   C. Splicing out of introns by the spliceosome

   D. Methylation of the DNA template to prevent re-transcription

8. **[Translation — Ribosomes]** During translation, the ribosome moves along the mRNA in the $5' \rightarrow 3'$ direction. The A site, P site, and E site have specific functions. In what order does a tRNA carrying an amino acid move through these sites?

   A. E site → P site → A site

   B. A site → P site → E site

   C. P site → A site → E site

   D. A site → E site → P site

9. **[Mutations — Types]** A point mutation changes a codon from GAA (glutamic acid) to GUA (valine). This mutation is most accurately described as:

   A. A silent mutation — the amino acid is unchanged

   B. A nonsense mutation — the codon now codes for a stop

C. A missense mutation — a different amino acid is incorporated, potentially altering protein structure and function

D. A frameshift mutation — the reading frame is disrupted

10. **[Gene Expression — Central Dogma]** A researcher discovers a virus that can copy its RNA genome into DNA inside a host cell. This process requires an enzyme called reverse transcriptase. This finding challenges which aspect of the original central dogma?

    A. That DNA can be transcribed to RNA

    B. That RNA can be translated to protein

    C. That information flows in one direction only — originally stated as DNA → RNA → protein, with no provision for RNA → DNA

    D. That proteins can be converted back to nucleic acids

▶ Show Answers

May 2026 Prediction Questions

⚠ EXAM ALERT

These are NOT official IB questions. These are trend-based practice questions written to reflect the topic areas and question styles most likely to appear on the May 2026 IB Biology HL Paper 2. Based on recent exam patterns (2022-2025), expect heavy weighting on: regulation of gene expression, the process of translation, types of mutations and their consequences, and epigenetics.

WORKED EXAMPLE

Question 1 [Gene Expression Regulation] [~8 marks]

In prokaryotes, the lac operon is a well-studied model of gene regulation.

(a) Outline the structure of the lac operon, identifying the role of the promoter, operator, and structural genes.

(b) Explain how the lac operon is regulated when lactose is absent and when lactose is present.

(c) Suggest why regulation of gene expression is important for a bacterium's survival, with reference to energy efficiency.

▶ Show Solution

WORKED EXAMPLE

Question 2 [Translation] [~7 marks]

Describe the process of translation in eukaryotic cells.

(a) Outline the roles of mRNA, tRNA, and ribosomes during translation.

(b) Describe the events that occur during the elongation phase of translation.

(c) Explain how a single mRNA molecule can be translated by multiple ribosomes simultaneously, and state the term for this structure.

▶ Show Solution

WORKED EXAMPLE

Question 3 [Mutations] [~6 marks]

A gene has the following mRNA sequence for the first six codons:

AUG-GAA-UUU-CGG-UAC-AAG

A point mutation changes the third codon from UUU to UAU.

(a) State the type of point mutation this represents and explain your reasoning.

(b) Using the genetic code, determine the amino acid change caused by this mutation and predict its likely effect on protein function.

(c) Distinguish between a missense mutation and a nonsense mutation, giving an example codon change for each.

▶ Show Solution