MEGA LECTURE

IB DP BIOLOGY

Theme D: Continuity and Change

D2.2 Gene Expression

Revision Notes · Higher Level only

Fahad H. Ahmad

+92 323 509 4443 | Megalecture.com

Original notes prepared for the IB Diploma Programme Biology course (first assessment 2025)

What the syllabus requires

D2.2 is examined at **Higher Level only**. Everything in these notes is HL content. By the end of the topic you should be able to do each of the following. Use the list as a final checklist before the exam.

Understanding (HL)	You should be able to...
Gene expression	Explain that gene expression is a gene being used to make a functional product, and that not every gene is expressed in every cell.
Differential expression	Explain how switching different genes on and off (differential gene expression) produces specialised cell types from one genome.
Regulation of transcription	Describe how promoters, transcription factors, enhancers and silencers control whether a gene is transcribed.
The *lac* operon	Use the prokaryotic *lac* operon of *E. coli* as a model of gene regulation and predict its state with and without lactose.
Epigenetics	Explain heritable changes in expression that do not change the DNA base sequence, via DNA methylation and histone modification.
Environment and expression	Discuss how environmental factors influence gene expression, using examples such as diet, temperature and identical twins.
Cancer (conceptual)	Outline how mutations and epigenetic changes to gene regulation can contribute to cancer.

> **HL only:** The whole of D2.2 is Higher Level. It links closely to D2.1 (cell specialisation), D1.1 - D1.3 (DNA, replication, transcription and translation) and D3.2 (inheritance), and is often examined in combination with them.

1. What gene expression means (HL)

Gene expression is the process by which the information in a gene is used to make a functional product - usually a protein, but sometimes a functional RNA molecule. A gene is *expressed* when it is being transcribed and (for a protein-coding gene) translated; it is *silent* when it is not.

The central idea of D2.2 is that **not all genes are expressed in all cells, or at all times**. Every nucleated body cell in an organism contains the same complete genome, yet a neuron, a muscle cell and a pancreatic cell each use only a subset of those genes. The pattern of genes switched on defines what the cell is and does.

Differential gene expression is the term for different cells expressing different combinations of genes. It is the basis of **differentiation** - the process by which an unspecialised cell becomes a specialised one (D2.1). The genome is a full recipe book; each cell type reads only certain pages.

Cell type	A gene switched ON (expressed)	A gene kept OFF (silent)
Pancreatic beta cell	Insulin gene	Haemoglobin genes
Red blood cell precursor	Haemoglobin (globin) genes	Insulin gene
Rod cell (retina)	Rhodopsin (opsin) gene	Insulin, globin genes

Cell type	A gene switched ON (expressed)	A gene kept OFF (silent)
Skin (epidermal) cell	Keratin genes	Rhodopsin gene

> **Key distinction:** Every cell has the same *genes* (the same DNA); cells differ in which genes are *expressed*. Do not write that a muscle cell 'has different genes' from a nerve cell - it has the same genes, switched on differently.

2. Where expression is regulated (HL)

Gene expression can in principle be controlled at several stages between DNA and a finished protein, but the **most important control point is transcription** - deciding whether an mRNA is made at all. Regulating the first step is efficient, because it avoids wasting resources building products the cell then has to break down.

Control point	What is regulated
Transcriptional (main)	Whether, and how often, RNA polymerase transcribes the gene into mRNA.
Post-transcriptional	RNA processing and splicing, and mRNA stability / degradation.
Translational	Whether ribosomes translate the mRNA into polypeptide.
Post-translational	Modification, folding and activation of the finished protein.

Key players in transcriptional control

- **Promoter** - a DNA sequence just before (upstream of) a gene where RNA polymerase and other proteins bind to begin transcription.
- **Transcription factors** - proteins that bind to DNA and control the rate of transcription. **Activators** increase transcription; **repressors** decrease or block it.
- **Enhancers** - regulatory DNA sequences (sometimes far from the gene) that, when bound by activators, increase transcription of that gene.
- **Silencers** - regulatory DNA sequences that, when bound by repressors, decrease or switch off transcription.

Whether a gene is transcribed in a given cell therefore depends on which transcription factors that cell is making, and on whether the relevant DNA is physically accessible (see epigenetics, section 4).

> **Exam note:** Enhancers and silencers are stretches of *DNA*; activators and repressors are the *proteins* (transcription factors) that bind them. Mixing these up is a common slip.

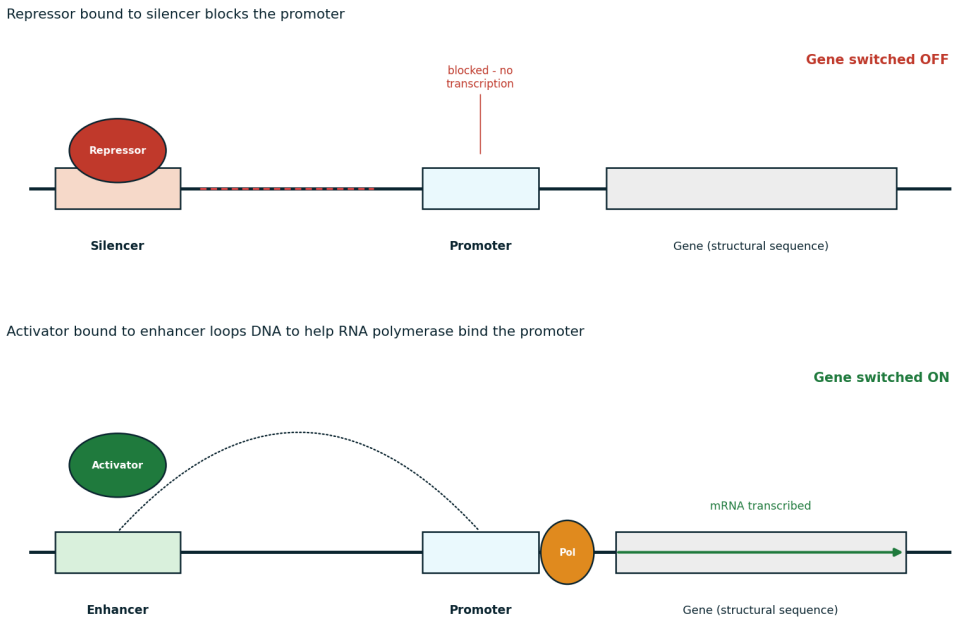

Figure 1. Transcriptional control: a repressor bound to a silencer blocks RNA polymerase at the promoter (gene OFF, top); an activator bound to an enhancer helps RNA polymerase bind the promoter, so the gene is transcribed (gene ON, bottom).

3. A prokaryotic model - the *lac* operon (HL)

Bacteria regulate genes efficiently by grouping related genes into an **operon**: a cluster of genes transcribed together from a single promoter under shared control. The *lac* operon of *Escherichia coli* is the classic worked example of how a cell switches genes on only when they are needed.

Components of the operon

Component	Role
Structural genes (*lacZ, lacY, lacA*)	Code for the enzymes that import and break down lactose (e.g. beta-galactosidase, which splits lactose into glucose and galactose).
Promoter	DNA sequence where RNA polymerase binds to start transcribing the structural genes.
Operator	DNA sequence between promoter and structural genes; the switch to which the repressor binds.
Regulatory gene (*lacI*)	A separate gene that codes for the *lac* **repressor** protein.

How the switch works

The *lac* operon controls enzymes for using lactose. Making those enzymes when no lactose is present would waste energy, so the operon is normally **off** and is switched on only when lactose is available - a form of **induction**.

- **Lactose absent:** the repressor protein binds the operator. This physically blocks RNA polymerase from transcribing the structural genes, so little or no enzyme is made. The operon is **OFF**.
- **Lactose present:** lactose (as its isomer allolactose) binds the repressor and changes its shape, so the repressor can no longer bind the operator. RNA polymerase is now free to transcribe the structural genes, the enzymes are made, and lactose is broken down. The operon is **ON**.

Lactose therefore acts as an **inducer**: its presence induces expression of the very enzymes that digest it. When the lactose runs out, the repressor binds the operator again and the operon switches back off - an economical, self-regulating feedback system.

> **Why it matters:** The *lac* operon shows the general principle that a small molecule from the environment (lactose) can control which genes a cell expresses. The same logic - a signal changing which genes are on - underlies regulation in eukaryotes, though the machinery is more elaborate.

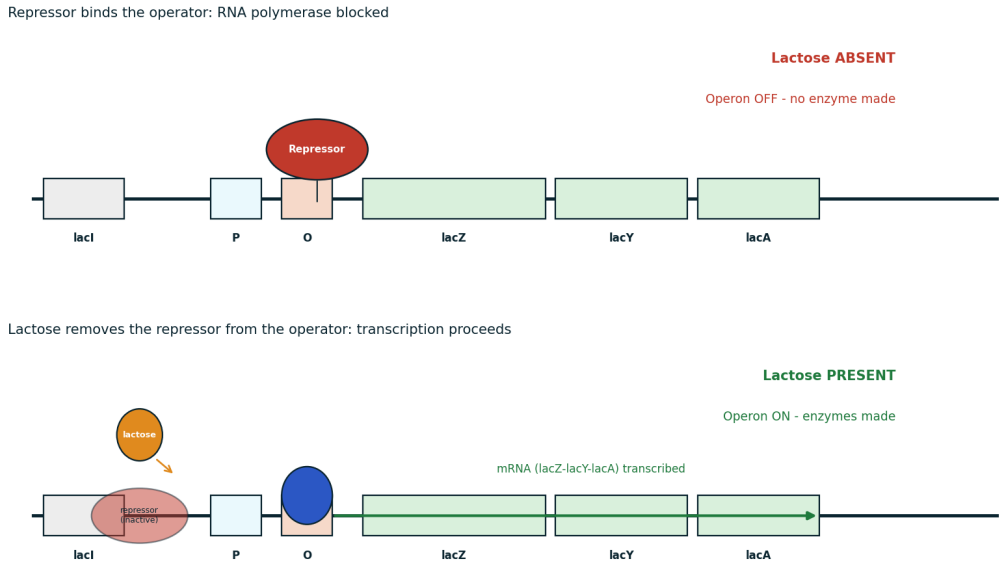

Figure 2. The lac operon. Lactose absent: the repressor binds the operator and blocks RNA polymerase (top, OFF). Lactose present: the repressor leaves the operator and RNA polymerase transcribes lacZ, lacY, lacA (bottom, ON).

4. Epigenetics (HL)

Epigenetics is the study of **heritable changes in gene expression that do NOT involve a change in the DNA base sequence**. Epigenetic changes switch genes on or off, or turn them up or down, while leaving the underlying genetic code unchanged. They can be copied to daughter cells during mitosis, and in some cases passed between generations.

The set of chemical tags on DNA and its associated proteins is called the **epigenome**. Unlike the genome, the epigenome is dynamic: it changes with cell type, developmental stage and environment. Two main mechanisms are studied at HL.

(a) DNA methylation

Methyl groups ($-CH_3$) are added to cytosine bases in the DNA, typically where a cytosine sits next to a guanine. Heavy methylation of a gene's promoter region generally **silences** the gene - it becomes harder for transcription factors and RNA polymerase to bind, so the gene is switched off. Broadly: **more methylation → less expression**.

(b) Histone modification

In eukaryotes DNA is wound around **histone** proteins to form chromatin. Chemical tags on the histone tails change how tightly the DNA is packed and therefore how accessible it is:

- **Histone acetylation** (adding acetyl groups) loosens the packing, opening up the chromatin so genes can be transcribed - it tends to **activate** expression.
- Removing acetyl groups (deacetylation), and some methylation of histones, tightens the packing and tends to **switch genes off**.

Key properties of epigenetic changes

- **Reversible** - tags can be added and removed, so expression can be switched back; this is different from a permanent mutation.
- **Heritable through cell division** - the pattern of tags is copied so daughter cells 'remember' which genes were on, keeping a cell type stable.
- **Influenced by the environment** - diet, stress, toxins and other factors can add or remove tags and so change expression.
- **Do NOT change the base sequence** - the genetic code is untouched; only its use is changed.

> **Analogy:** If the genome is the text of a book, the epigenome is the highlighting and sticky notes: it does not rewrite a single word, but it decides which passages get read. Methylation is like a 'do not read' tag over a gene.

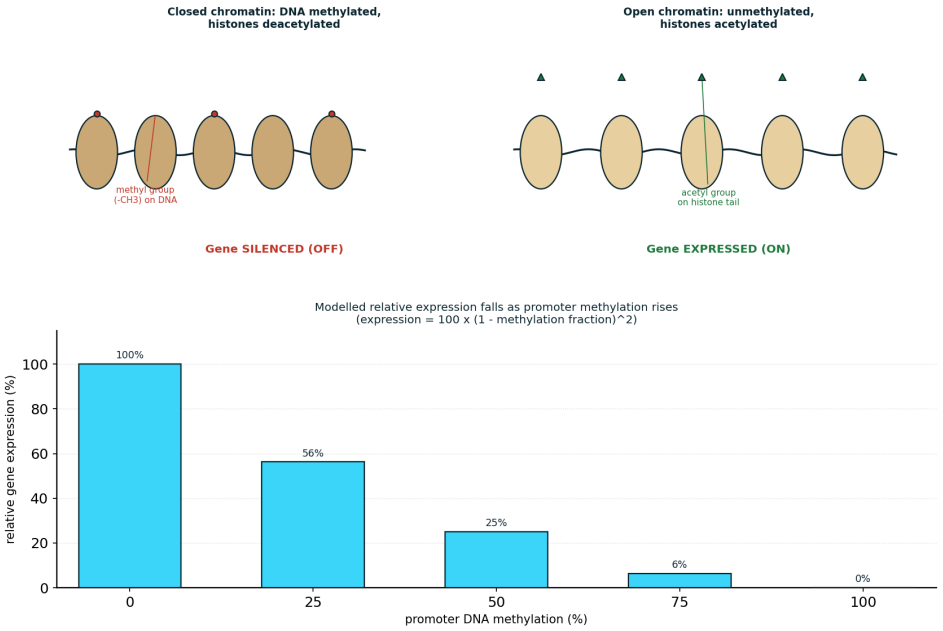

Figure 3. DNA methylation and histone deacetylation pack chromatin tightly and silence a gene (left); an unmethylated, acetylated region opens chromatin so the gene is expressed (right). Bottom: a modelled dose-response showing relative expression falling from 100% to 0% as promoter methylation rises from 0% to 100%.

5. The environment and gene expression (HL)

Because epigenetic tags respond to conditions, an organism's **phenotype is shaped by both its genotype and its environment**. The same genome can produce different outcomes depending on the signals a cell receives. Several examples illustrate the point.

Example	How the environment affects expression
Diet	Nutrients supply the methyl groups used in DNA methylation, so diet can alter methylation patterns and hence which genes are expressed.
Temperature	In some organisms temperature changes expression: for example, coat-colour genes in certain animals are only expressed (dark pigment) in cooler body regions such as the extremities.
Identical twins	Twins share a genome but their epigenomes drift apart over life as they experience different diets, habits and environments - so their gene expression, and some traits and disease risks, diverge with age.
Stress / toxins	Exposure to stress hormones or chemicals can add or remove epigenetic tags, changing expression of the affected genes.

These examples show that having a gene is not the same as expressing it. The environment helps decide which of the genes a cell carries are actually used.

> **Nature and nurture:** Epigenetics gives a molecular explanation for how 'nurture' (environment) acts on 'nature' (genes): environmental signals leave epigenetic marks that adjust expression without altering the DNA sequence.

6. One genome, many cell types (HL)

It is worth stating clearly how differential expression produces specialised cells even though every cell shares the same DNA. The reasoning ties together the whole topic and is a favourite extended-response question.

- All body cells arise by mitosis from one zygote, so all carry the same complete genome (the same genes).
- During development, cells receive different chemical signals depending on their position, which switch on different sets of transcription factors.
- These transcription factors, together with epigenetic marks (methylation and histone modification), switch particular genes on and keep others off - **differential gene expression**.
- Each cell therefore makes only its own characteristic set of proteins, giving it a distinctive structure and function (differentiation).
- Because epigenetic marks are copied at mitosis, a differentiated cell's daughter cells inherit the same expression pattern, so a liver cell divides to make more liver cells - the identity is stable.

> **In one sentence:** Same genes + different genes expressed = different cells. Differentiation is not a change in the DNA but a change in which parts of it are switched on.

7. Gene regulation and cancer (HL, conceptual)

Cancer is a disease of uncontrolled cell division, and at its root it is a disease of **faulty gene regulation**. The genes that normally control the cell cycle - encouraging division when appropriate, and restraining or halting it otherwise - can be disrupted in two broad ways.

- **Mutations** that change the DNA base sequence of regulatory genes - for example a mutation that leaves a growth-promoting gene permanently switched on, or that disables a gene which normally halts division.
- **Epigenetic changes** that alter expression without changing the sequence - for example abnormal methylation that silences a protective gene, or loss of methylation that switches on a growth-promoting gene.

Either route can push a cell into dividing when it should not. The epigenetic route is a striking illustration that **how a gene is expressed** - not just its sequence - matters for health. Because epigenetic changes are reversible in principle, they are an active area of research for possible treatments.

> **Conceptual only:** You are not expected to memorise named oncogenes for D2.2. The point to grasp is that both mutation and epigenetic change can disrupt the regulation of genes controlling the cell cycle, contributing to cancer.

8. Worked examples and skills (HL)

Worked example 1 - state of the lac operon

Predict whether the *lac* operon structural genes are transcribed (a) when lactose is absent, and (b) when lactose is present. Explain each case.

(a) **Lactose absent → operon OFF.** The repressor protein is free to bind the operator, which blocks RNA polymerase from transcribing the structural genes, so essentially no lactose-digesting enzyme is made.

(b) **Lactose present → operon ON.** Lactose (allolactose) binds the repressor and changes its shape so it can no longer bind the operator. RNA polymerase transcribes the structural genes and the enzymes are made.

Principle: the enzymes are produced only when their substrate is available - an economical, inducible system.

Worked example 2 - effect of methylation

A tumour-suppressor gene that normally slows cell division becomes heavily methylated in its promoter region. Predict the effect on the gene's expression and on the cell.

Heavy promoter methylation makes it harder for transcription factors and RNA polymerase to bind, so the gene is **silenced** (switched off). The cell loses that brake on division, so it may divide more than it should - a step towards cancer. Note the DNA base sequence is unchanged; this is an **epigenetic** effect.

Worked example 3 - two cells, same DNA

A liver cell and a neuron from the same person contain identical DNA yet look and behave completely differently. Explain how this is possible.

Both cells carry the same complete genome, but they express **different sets of genes** (differential gene expression). Different transcription factors and different epigenetic marks (DNA methylation and histone modification) switch on the genes needed for liver function in one cell and neuron function in the other, while keeping the rest silent. Each therefore makes a different set of proteins, giving different structure and function.

Worked example 4 - interpreting an epigenetics scenario

Two genetically identical mice are raised on different diets. As adults they differ in coat colour and weight, and analysis shows different DNA methylation patterns at certain genes. Explain what has happened, and state whether their genes have changed.

Diet supplied different amounts of the raw materials used for DNA methylation, so the two mice ended up with different methylation patterns. This changed which genes were expressed, producing the different coat colour and weight - an **environmental effect on gene expression through epigenetics**. Their **genes (DNA base sequence) have NOT changed**; only the epigenetic marks, and therefore expression, differ.

Worked example 5 - identical twins diverging

Identical twins are very alike as children but show increasingly different disease risks as they age. Suggest an epigenetic explanation.

The twins share the same genome, but over the years they experience different environments (diet, stress, smoking, activity). These add and remove epigenetic marks differently in each twin, so their **epigenomes drift apart** and different genes become switched on or off. The resulting differences in gene expression lead to different traits and disease susceptibilities, despite identical DNA.

9. Common pitfalls (HL)

- Writing that different cells 'have different genes'. They have the **same** genes; they **express** different ones.
- Claiming epigenetic changes alter the DNA base sequence. They do **not** - that is the defining feature of epigenetics.
- Getting methylation backwards. DNA methylation of a promoter generally **silences** a gene; it does not switch it on.
- Confusing enhancers/silencers (DNA) with activators/repressors (the proteins that bind them).
- Saying the *lac* repressor binds the promoter. It binds the **operator**; RNA polymerase binds the promoter.

- Treating epigenetic marks as permanent. They are **reversible**, though they can be stable enough to be inherited by daughter cells.

10. Quick reference (HL)

Term	One-line meaning
Gene expression	A gene being used to make a functional product (protein or RNA).
Differential expression	Different cells expressing different combinations of genes.
Promoter	DNA site where RNA polymerase binds to start transcription.
Transcription factor	Protein that binds DNA to raise (activator) or lower (repressor) transcription.
Enhancer / silencer	DNA sequences that, when bound by proteins, increase / decrease transcription.
Operon	Cluster of genes transcribed together under shared control (prokaryotes).
lac operon (no lactose)	Repressor on operator → genes OFF.
lac operon (lactose)	Lactose removes repressor → genes ON (induction).
Epigenetics	Heritable change in expression with NO change to DNA base sequence.
DNA methylation	Methyl tags (usually on promoters) that silence genes.
Histone acetylation	Loosens chromatin → activates expression.
Epigenome	The reversible, environment-sensitive set of tags on DNA and histones.

11. Test yourself (HL)

Attempt these without notes; full worked answers follow.

1. Define gene expression, and explain why not all genes are expressed in all cells.

2. Name the main stage at which gene expression is regulated, and state two DNA sequences and two types of protein involved in that regulation.

3. Describe the state of the *lac* operon when lactose is absent, naming the components involved.

4. Lactose is added to a culture of *E. coli*. Explain the sequence of events that switches the *lac* operon on.

5. Define epigenetics and give the two main mechanisms of epigenetic change.

6. Explain the effect of (a) methylation of a gene's promoter and (b) acetylation of histones on gene expression.

7. Explain how a skin cell and a nerve cell from the same person can differ so much although their DNA is identical.

8. Outline two ways in which changes to gene regulation can contribute to cancer.

Answers

1. Gene expression is a gene being used to make a functional product (a protein or functional RNA). Not all genes are expressed in all cells because different cell types switch on different sets of genes (differential expression), so each makes only the proteins it needs; expressing every gene everywhere would be

wasteful and would prevent specialisation.

2. Mainly at **transcription**. DNA sequences: promoters and enhancers/silencers (operators in prokaryotes). Proteins: transcription factors acting as activators and repressors.

3. Lactose absent → operon **OFF**. The regulatory gene makes a **repressor** protein that binds the **operator**, blocking RNA polymerase at the **promoter** from transcribing the **structural genes**, so no lactose-digesting enzyme is made.

4. Lactose (as allolactose) binds the repressor and changes its shape; the repressor can no longer bind the operator; RNA polymerase is now free to move from the promoter through the structural genes and transcribe them; the enzymes are made and lactose is broken down. Lactose is acting as an inducer.

5. Epigenetics is the study of heritable changes in gene expression that do not change the DNA base sequence. Two mechanisms: **DNA methylation** and **histone modification** (e.g. acetylation).

6. (a) Promoter methylation generally **silences** the gene (switches it off) by making it harder for the transcription machinery to bind. (b) Histone acetylation loosens chromatin so DNA is more accessible, tending to **activate** (switch on) expression.

7. Both cells contain the same complete genome, but they show differential gene expression: different transcription factors and different epigenetic marks (methylation and histone modification) switch on the genes needed for skin function in one and nerve function in the other, keeping the rest silent. Each makes a different set of proteins, so has different structure and function.

8. (i) **Mutations** in the DNA sequence of genes that regulate the cell cycle (e.g. leaving a growth-promoting gene permanently on, or disabling a gene that restrains division). (ii) **Epigenetic changes** - such as abnormal methylation silencing a protective gene, or loss of methylation switching on a growth-promoting gene - that alter expression without changing the sequence. Either can cause uncontrolled division.