EPIGENETICS

Content Statements:

D2.2.1	Gene expression as the mechanism by which information in genes has effects on phenotype
D2.2.2	Regulation of transcription by proteins that bind to specific base sequences in DNA
D2.2.3	Control of the degradation of mRNA as a means of regulating translation
D2.2.4	Epigenesis as the development of patterns of differentiation in cells of a multicellular organism
D2.2.5	Differences between the genome, transcriptome and proteome of individual cells
D2.2.6	Methylation of the promoter and histones in nucleosomes as examples of epigenetic tags
D2.2.7	Epigenetic inheritance through heritable changes to gene expression
D2.2.8	Examples of environmental effects on gene expression in cells and organisms
D2.2.9	Consequences of removal of most but not all epigenetic tags from the ovum and sperm
D2.2.10	Monozygotic twin studies
D2.2.11	External factors impacting the pattern of gene expression

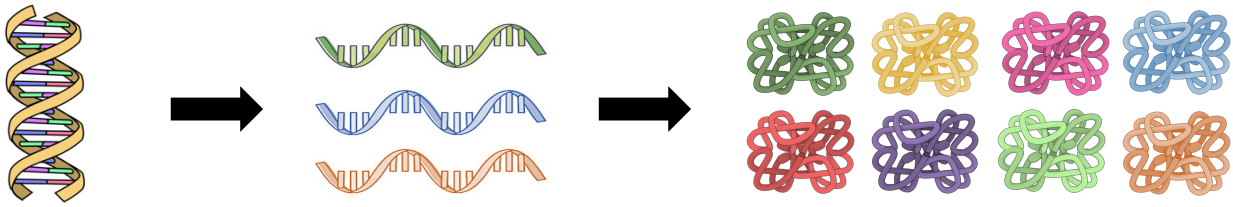
GENE EXPRESSION

Gene expression is the process by which information encoded within genes is used to create proteins. Gene expression involves two sequential steps – **transcription** (DNA → mRNA) and **translation** (mRNA → protein).

The products of gene expression collectively give rise to the genome, transcriptome and proteome of a cell.

- **Genome:** Totality of genetic information in a cell, tissue or organism (all genes and non-coding DNA)
- **Transcriptome:** All genetic instructions that have been actively transcribed to RNA (mRNA, tRNA, tRNA)
- **Proteome:** The complete set of proteins expressed within a cell, tissue or organism at a particular time

While a genome will be identical between all of cells in an organism, transcriptomes and proteomes vary.



GENOME	TRANSCRIPTOME	PROTEOME
~20,000 genes	~100,000 transcripts	>1,000,000 proteins

GENE REGULATION

Gene expression levels can be moderated at the level of the proteome, the transcriptome or the genome:

1. Expression levels can be controlled by regulating the amount of protein produced (translational control)
2. Expression can be controlled by regulating the amount or the activity of mRNA (transcriptional control)
3. Expression can be controlled by regulating the accessibility of the genes (histone or DNA modification)

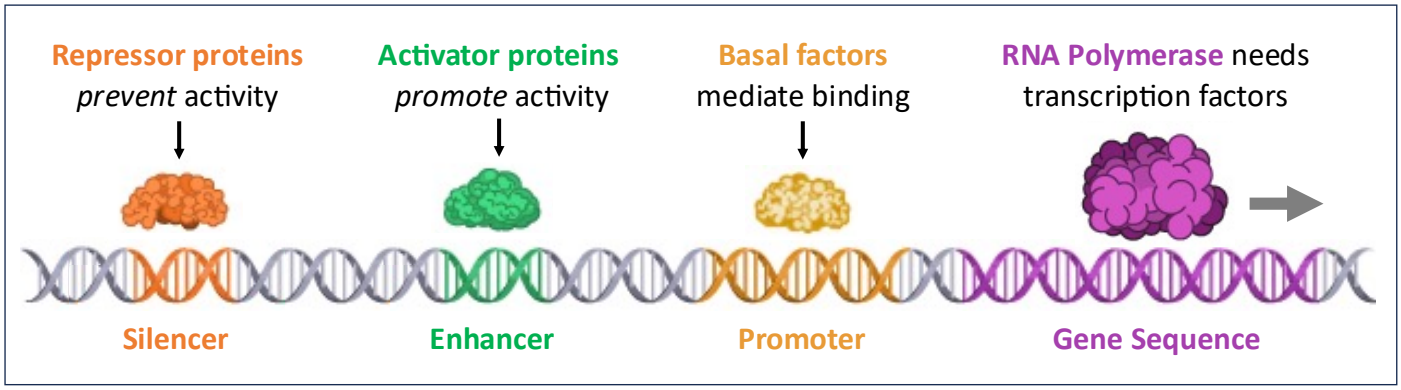
By regulating gene expression, cells can become specialised and differentiate despite having shared genetic instructions – this also allows the **phenotype** of organisms to change while the genotype remains the same.

TRANSLATIONAL CONTROL

Gene expression may be regulated by controlling the degradation of mRNA transcripts. Within human cells, mRNA may persist for time periods from minutes up to days, before being broken down. As mRNA provides the instructions required by the ribosomes for protein assembly, the longevity of the transcripts determines the expression levels of a gene – the longer a transcript lasts before degradation, the greater the amount of protein produced. The mRNA transcripts are broken down by **nucleases** to form single nucleotides that can be recycled by the cell. This degradation alters the transcriptome within a cell at a particular point of time.

TRANSCRIPTION FACTORS

Gene expression may also be regulated by controlling the rate at which mRNA transcripts are produced. The process of transcription is initiated by the binding of RNA polymerase to a promoter. Certain proteins – called transcription factors – mediate this binding. Transcription factors may either bind to the promoter directly, or to DNA sequences outside of the promoter. **Activator** proteins bind to *enhancer sites* and act to increase the rate of transcription. **Repressor** proteins bind to *silencer sites* and decrease transcription rates. The presence of certain transcription factors may be tissue specific, leading to a differentiation of cell types.

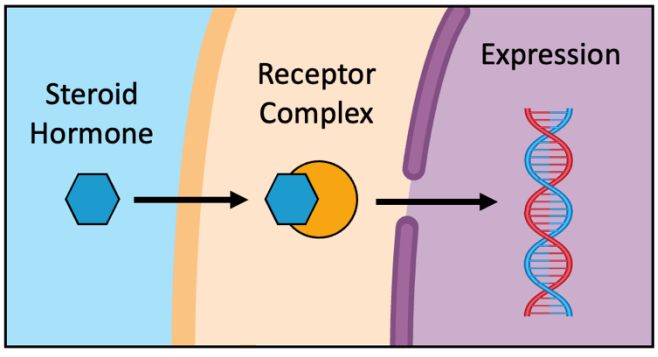


TRANSCRIPTION SIGNALS

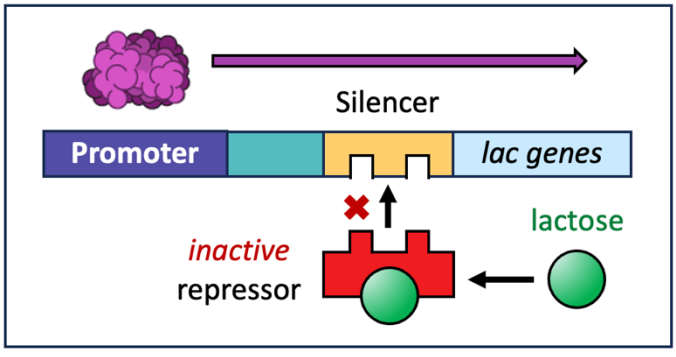
The activity of any given transcription factor can be controlled by either intracellular or extracellular signals.

An example of an *extracellular* signal is a **hormone** – a chemical messenger transported in blood. Steroid hormones bind to receptors within the cell to form a complex that will act as a transcription factor. Peptide hormones bind to receptors on plasma membranes and act to control transcription via second messengers.

An example of an *intracellular* signal is an **inducer molecule** – such as lactose. Lactose binds to a repressor protein that supresses the transcription of genes responsible for lactose metabolism. The binding of lactose prevents the repressor from functioning – hence lactose metabolism occurs when lactose is present in cells.



Hormonal Signalling (e.g. estrogen)



Inducer Molecule (e.g. lactose)

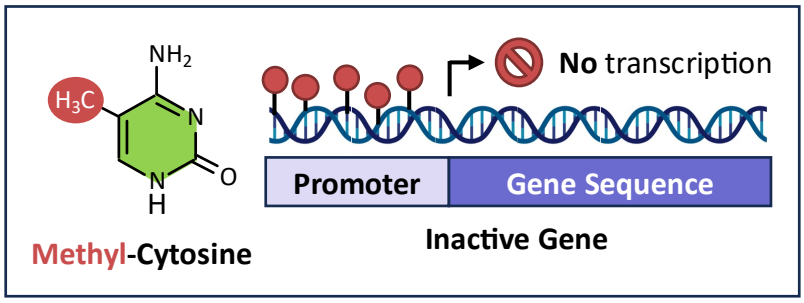
EPIGENESIS

Epigenesis describes the specific development of an organism from an undifferentiated zygote to a complex multicellular organism via differential gene expression. The activation of different genes in different cells of a multicellular organism will lead to different patterns of development in these cells. This causes these cells to have different characteristics even though their base sequences are identical (their phenotype is altered but not their genotype). Epigenetic changes are triggered by specific chemical modifications to DNA (called **epigenetic tags**). These epigenetic tags do not alter the DNA base sequence and are potentially reversible.

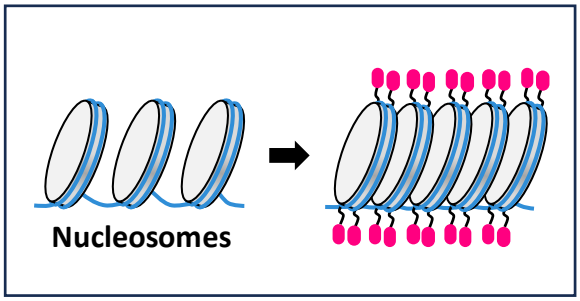
METHYLATION

One specific chemical modification that acts as an epigenetic tag is the addition of a methyl group ($-CH_3$) to either DNA (at the promoter) or histones (within the nucleosome). In eukaryotic cells, methylation of DNA predominantly occurs at cytosine bases that are immediately adjacent to a guanine base (a **CpG island**). A majority of eukaryotic genes have CpG islands within their promoter sequence and the direct methylation of the promoter impedes the downstream activity of RNA polymerase. Hence **DNA methylation** functions to reduce transcriptional activity in eukaryotic cells by inactivating transcription (the gene is switched 'off').

In eukaryotic organisms, the DNA is associated with **histone proteins** to form a condensed complex known as a nucleosome. The histone proteins have protruding tails that are positively charged, which allows the histone to associate with the negatively charged DNA. Adding a methyl group to the histone tail maintains the positive charge, making DNA more coiled and reducing transcription. Individual nucleosomes are linked together (like beads on a string) to form chromatin. When the histones are methylated, the DNA becomes supercoiled (tightly packed) and not accessible for transcription – existing as condensed **heterochromatin**. The removal of methyl tags will cause the DNA to become more loosely packed and therefore accessible to the transcription machinery (**euchromatin**). Different cell types will have different DNA segments packaged as heterochromatin and euchromatin depending on which genes are active within the cell. Some segments of DNA may be permanently supercoiled, while other segments may change over the life cycle of the cell.



DNA Methylation



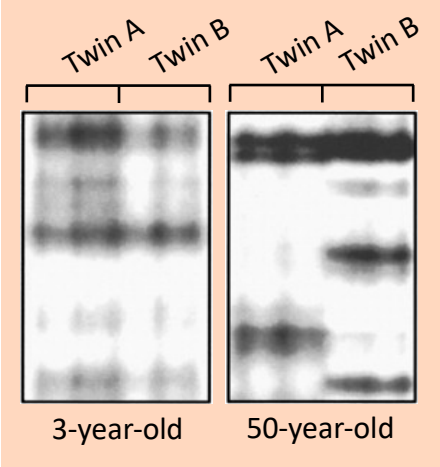
Histone Methylation

ENVIRONMENTAL TRIGGERS

Air pollutants are an example of environmental factors that can affect the methylation of DNA in a person. Several traffic-related pollutants have been associated with changes in DNA methylation – such as nitrogen oxides and polyaromatic hydrocarbons (PAH). In general, air pollution mediates a reduction in methylation across the genome, although methylation at some sites may be increased. Air pollution is specifically linked to the methylation of **immunoregulatory genes**, leading to increased inflammation and an altered immune profile. These methylation patterns result in the manifestation of cardiopulmonary complications – such as high blood pressure and asthma. A change in methylation is also a common cause for some types of cancer. Air pollutants may act by affecting the action of the methylating enzyme – **DNA methyltransferase** (DNMT).

MONOZYGOTIC TWIN STUDIES

The role of external environmental stimuli on gene expression can be demonstrated by comparing epigenetic profiles of monozygotic twins. Monozygotic twins result from the division of a fertilised egg into two distinct embryos that have an identical genome (i.e. they are clones). By comparing the methylation pattern of monozygotic twins, the role of the environment in phenotypic development can be assessed. DNA methylation patterns will differ between twins and will differ further over time as a result of exposure to unique environmental conditions. DNA methylation patterns can be used to identify genes that may be involved in the development of specific diseases (present in one twin).



EPIGENETIC INHERITANCE

Complex organisms develop from undifferentiated cells as a consequence of the programmed expression of genes via epigenetic tags. Different genes are switched 'on' or 'off' in specific cells to promote development of distinctive cell lines (tissues). As egg and sperm cells develop from differentiated germline cells, the tags that already exist in these cells must be erased to allow for epigenesis to occur upon fertilisation. Gametes must be reprogrammed via **epigenetic tag removal** to return them to a blank genetic slate. Reprogramming ensures that the early embryo can form every type of cell in the body by resetting an embryo's epigenome.

IMPRINTED GENES

A small proportion of genes do not undergo reprogramming during gamete production and will retain their epigenetic tags. These sequences are called **imprinted genes** and allow phenotypic changes to be passed on to offspring. The presence of imprinted genes can explain the size differences seen in ligers and tigons (lion-tiger hybrids). Lions live in social groups where a single lioness may mate with many males and birth cubs from multiple fathers. A male lion is incentivised to have the largest offspring (better for cub survival) and hence will produce sperm containing imprinted genes promoting growth. But a lioness is incentivised to have small offspring (lower birth risk) and produces eggs containing imprinted genes to restrict growth. The combination of these imprinted genes will cancel each other out and result in normal-sized cubs. Tigers are solitary animals and a tigress only produces a litter with cubs from a single male. As such, imprinted genes are not needed. When a lion and tiger successfully mate, the different imprinted genes present in the parents determines offspring size. In **ligers** (male lion × female tiger), the male lion's imprinted genes result in a large size. In **tigons** (male tiger × female lion), a lack of this imprinted gene prevents larger sizes.