TRANSCRIPTION

Content Statements:

D1.2.12 Directionality of transcription and translation

D1.2.13 Initiation of transcription at the promoter

D1.2.14 Non-coding sequences in DNA do not code for polypeptides

D1.2.15 Post-transcriptional modification in eukaryotic cells

D1.2.16 Alternative splicing of exons to produce variants of a protein from a single gene

GENETICS

DNA functions as the genetic blueprint for cells. DNA codes for protein – which function to manifest specific characteristics. **Genes** are sequences of DNA that code for particular proteins. Each gene will have a unique base sequence that is transcribed into RNA and, in turn, translated into a polypeptide in response to signals within a cell. Not all DNA sequences consist of protein-encoding genes – the majority of DNA is non-coding.

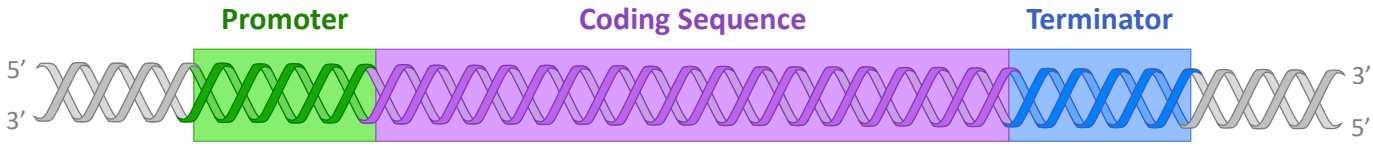
NON-CODING DNA

Non-coding DNA has been found to serve a variety of different purposes within cells:

- **Satellite DNA** sequences (e.g. short tandem repeats) are used for DNA profiling
- **Telomeres** (chromosome ends) function to prevent chromosomal deterioration
- **Introns** are non-coding sequences *within* protein-encoding genes in eukaryotes
- **Non-coding genes** produce RNA that does not code for proteins (tRNA and rRNA)
- **Gene regulatory sequences** moderate transcription (e.g. enhancers or silencers)

GENES

A gene is composed of three key sections. The **promoter** is the site to which the enzyme RNA polymerase will bind – it is responsible for initiating transcription. The **coding sequence** is the region of DNA that is transcribed into RNA, while the **terminator** sequence functions to stop transcription by RNA polymerase. Basically, genes consist of start regions (promoter), copying regions (coding) and stop regions (terminator).

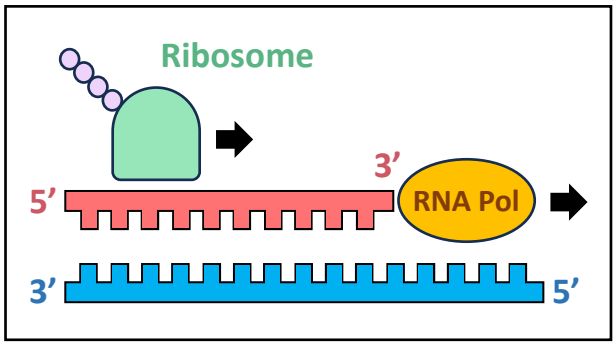


TRANSCRIPTION

Transcription is the process by which a DNA sequence (gene) is copied into complementary RNA sequences by **RNA polymerase**. This enzyme binds to the promoter and then separates the double-stranded DNA of the coding sequence (by breaking the hydrogen bonds between base pairs). **Free RNA nucleotides** then align opposite their complementary base partner and RNA polymerase joins them together with covalent bonds (between the sugar-phosphate backbone). When the enzyme reaches the terminator sequence the synthesised RNA transcript is released and the double helix reforms. Transcription occurs in the nucleus.

DIRECTIONALITY

DNA is a double-stranded molecule and genes can occur on either strand. The carbon atoms in the pentose sugar of a nucleotide are numbered. The phosphate group is always attached to the 5'-carbon and will be connected to another nucleotide via the 3'-carbon. In transcription, new strands are made in a **5' → 3' direction**. Ribosomes also translate by reading the mRNA in a 5' → 3' direction.



TRANSCRIPTION FACTORS

The expression of genes is coordinated by **transcription factors**, which are produced by regulatory genes. Transcription factors either mediate or impede the binding of RNA polymerase to the promoter sequence:

- **Activator proteins** bind to *enhancer sites* and mediate promoter binding (gene 'on' – ↑ transcription)
- **Repressor proteins** bind to *silencer sites* and impede promoter binding (gene 'off' – ↓ transcription)

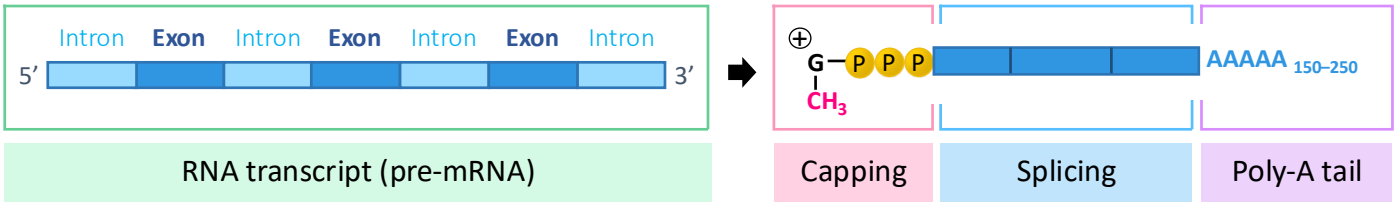
The presence of certain transcription factors may be tissue-specific or regulated by certain chemical signals. The study of changes in organisms as a result of variations in levels of gene expression is called **epigenetics**.

RNA PROCESSING

In eukaryotes, there are three post-transcriptional events that must occur in order to form messenger RNA.

- **Capping:** A methyl group is added to the 5'-end of the transcript to prevent degradation by nucleases
- **Polyadenylation:** A poly-A tail is added to the 3'-end in order to improve the stability of the transcript
- **Splicing:** Non-coding sequences are removed and the expressing sequences are then fused together

Non-coding sequences within a gene are called **introns**, whereas the coding sequences are known as **exons**.



ALTERNATIVE SPLICING

Splicing typically results in the removal of introns by a complex called the **spliceosome** – however, exons can also be selectively removed via a process known as *alternative* splicing. The removal of exon segments will result in the formation of different polypeptides from a single gene sequence. This is one way a proteome can be larger than the genome, as multiple protein variants can be produced from a genetic sequence.

An example of an application for alternative splicing is the production of different versions of the same protein. For instance, an enzyme can be cytosolic or membrane-bound based on the presence or absence of an anchoring motif. Alternative splicing in fruit flies results in the production of different versions of a particular protein in females and males (contributing to morphological differences between the sexes).