Eukaryotic Gene Expression

**Gene Expression:** The process in which the information within a gene is used, first to synthesise RNA, through transcription, and then to a protein, through translation, eventually to affect the phenotype of an organism.

**Central Dogma of Molecular Biology:** The principle of directional informational flow from DNA to RNA to protein.

**Transcription:** The process, in which a complementary RNA copy is made under the direction of the template strand of a specific region of the DNA molecule, catalysed by the enzyme RNA polymerase.

**Translation:** The process, in which a polypeptide chain is synthesised by ribosomes using genetic information encoded in an mRNA template

DNA $\rightarrow$ pre-mRNA $\rightarrow$ mature mRNA $\rightarrow$ tRNA & rRNA $\rightarrow$ polypeptide

**Gene:** A section of the DNA that contains the information in the form of a specific sequence of nucleotides to direct the synthesis of one polypeptide chain or RNA. It is a unit of inheritance located in the locus on the chromosome which specifies a particular character of an organism.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>RNA</th>
<th>DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substituent</td>
<td>Made of polynucleotides, basic units: phosphate -P to pentose sugar, nitrogenous base</td>
<td></td>
</tr>
<tr>
<td>Bond</td>
<td>Have a sugar-phosphate backbone linked by phosphodiester bonds</td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>Polymerised through condensation synthesis</td>
<td>Synthesised by complementary base pairing of nucleotides using a template</td>
</tr>
<tr>
<td>Size &amp; mass</td>
<td>Smaller molecular mass (20k to 2000k Da)</td>
<td>Larger molecular mass (100k to 150000k Da)</td>
</tr>
<tr>
<td>No. of subunits</td>
<td>1 polynucleotide chain</td>
<td>2 polynucleotide chains</td>
</tr>
<tr>
<td>3D structure</td>
<td>Almost always single-stranded, helical molecule, which can be folded into a complex tertiary structure eg. tRNA</td>
<td>Always a double-stranded helical molecule which forms a double helix. Coiled around histone proteins</td>
</tr>
<tr>
<td>Monomers</td>
<td>Ribonucleotides</td>
<td>Deoxyribonucleotides</td>
</tr>
<tr>
<td>Pentose sugar</td>
<td>OH group attached on 2’Carbon</td>
<td>H attached on 2’ Carbon</td>
</tr>
<tr>
<td>Chemical stability</td>
<td>Less stable – more reactive partly due to ribose having an additional reactive 2’ OH group</td>
<td>More stable – more resistant to spontaneous enzymatic breakdown due to deoxyribose lacking 2’OH group</td>
</tr>
<tr>
<td>Purines : pyrimidines</td>
<td>$A:U \neq G:C \neq 1:1$ (Ratio cannot be predicted as RNA is single-stranded, without a complementary strand)</td>
<td>$A:T = G:C = 1:1$ (Chargaff’s rule)</td>
</tr>
<tr>
<td>Basic forms</td>
<td>Several basic forms: messenger RNA, transfer RNA, ribosomal RNA, small nuclear RNA, small interfering RNA</td>
<td>Only one basic form</td>
</tr>
</tbody>
</table>
are removed and the remaining group is added to the free 3' OH group of the growing RNA chain via phosphodiester bond

| Re-annealing of DNA and proofreading | • RNA polymerase reanneals the unwound DNA behind it, dissociating the growing RNA chain from the template  
• It carries out proofreading functions and is responsible for the removal of incorrectly inserted ribonucleotides |
|------------------------------------|
| Termination                        | • RNA polymerase transcribes a terminator sequence in the DNA  
• Triggers the release of the RNA chain and dissociation of the RNA polymerase from the DNA  
• Transcribed terminator codes for a polyadenylation sequence (AAUAAA) |

**Exons**: protein-coding sequence in the gene

**Introns**: long sequences of nucleotides inserted between exons that do not code for any portion of the polypeptide, i.e. are non-coding sequences

**Post-transcriptional Modification**

<table>
<thead>
<tr>
<th>Modification</th>
<th>Process</th>
<th>Function</th>
</tr>
</thead>
</table>
| Addition of 5’ Methylguanosine cap | The 5’ end of the new pre-mRNA molecule is modified by addition of a cap that consists of a methylated guanine (G) nucleotide/methylguanosine triphosphate | • Protects mRNA from degradation by nucleases and phosphatases that degrade the RNA from the 5’ end during its transport from the nucleus to the cytoplasm  
• 5’ cap signals the 5’ end of the mRNA which serves as the assembly point to recruit the small subunit of the ribosome for translation to begin  
• Helps distinguish mRNAs from the other types of RNA molecules |
| Addition of 3’ poly(A) tail | Immediately after the pre-mRNA is cleaved by an endonuclease at a site 10-35 nucleotides after the AAUAAA poly(A) sequence, the 3’ end of the pre-mRNA is modified by addition of a series of ~200 adenine (A) nucleotides, referred to as the poly(A) tail. Catalysed by poly(A)-polymerase | • 3’ poly(A) tail protects the mRNA from degradation by nucleases  
• Make mRNA a more stable template for degradation  
• Required to facilitate the export of mRNA out of the nucleus via nuclear pores |
| RNA splicing | RNA splicing occurs after the release of pre-mRNA from RNA polymerase, during which, introns are removed while remaining exons are spliced/ligated together to form mature mRNA. Requires hydrolysis of ATP | • Provides for variation as different combination of exons could result in different types of polypeptides synthesised |

**Genetic code:**

- consists of information in the form of **3 nucleotide bases** called **codons** of mRNA
- also the **triplet bases** in the **non-template/non-transcribed strand** of DNA