Genetics & molecular biology

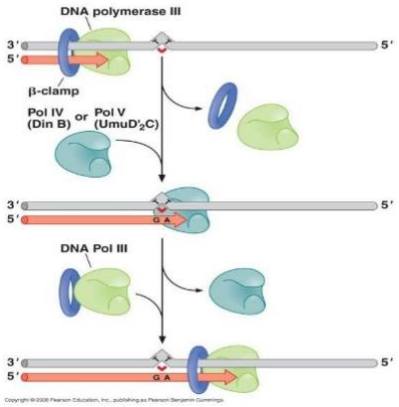
◉ Sheet ○ Slide

Number:	- 9
Done by:	- Nadin Rayyan & Yusser Raja
Corrected by:	- Abdullah Alaraj
Doctor:	- Mamoun Ahram

Translesion DNA synthesis

- ***Translesion DNA synthesis*** is a mechanism used by both prokaryotic and eukaryotic cells to bypass DNA mutations when in a hurry to carry out DNA replication.

  Ex.) The presence of a pyrimidine dimer which creates lesions in the DNA and with no time for cells to fix it.
- How it occurs: DNA polymerase responsible for DNA synthesis gets dissociated from the DNA and a group of specialized DNA polymerases jump over the lesions and fill in the gaps by adding a few bases.
- Although they have the ability to recognize what base is present and display some selectivity in base insertion, they still make mistakes. So, they are said to have **low fidelity** and are **error prone.**
- These specialized DNA polymerases then dissociate, and the original DNA polymerase returns.

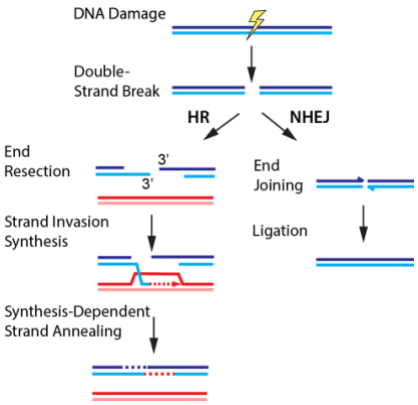


Recombination Repair

- Recombination repair is another DNA repair system that takes place when the DNA strand is broken up into two pieces. It can happen in 1 of 2 ways:

  1.) **Non-homologous end joining (NHEJ):** system that connects the two strands back together but **creates mutations** despite its efficiency.

2.) Homologous Repair System: system where the missing gaps (on damaged strand) are filled in by the opposite strand (normal strand).



- This system involves a protein called **Rad51**.

Breast Cancer

- In breast cancer specifically, there is a protein known as **BRCA1** that is mutated and causes hereditary breast cancer.
- BRCA1 and BRCA2 (a protein that is related to it) are originally involved in DNA repair, specifically in activating **homologous recombination repair** by recruiting Rad51 to the ssDNA.
- BRCA1 is also involved in **transcription** and **transcription coupled DNA repair**. (the details of this are not important as it is still under research)
- Mutations in BRCA1 account for 2% of all breast cancers and 5% of ovarian cancers.
- Screening for BRCA1 creates ethical issues as it may cause a woman to live in fear of developing breast cancer or may cause social issues/concerns for her and her family.

A Controversial Issue: Gene Repair

- 2 scientists discovered a system in bacteria known as **CRISPR Cas9** which is similar to the human immune system. Bacteria use this system to protect themselves from viral infections.
- The 2 scientists used this system to fix mutations by putting it into a cell to remove the mutated part of the DNA and replace/repair it with another part.
- Furthermore, it was used to genetically modify human embryos creating ethical problems.
- Has 50% accuracy.

**This is the end of the DNA repair mechanisms; the rest of this sheet will discuss transcription in eukaryotes.*

Transcription in Eukaryotes

RNA Polymerases:

- Unlike in bacteria which contain only 1 RNA polymerase, eukaryotic cells have **3: RNA polymerase I, II, III.**
    1. **RNA Polymerase I:** transcribes **rRNA** genes.
    2. **RNA Polymerase II:** transcribes protein encoding genes **(MRNA)** and microRNA. *the important one*
    3. **RNA Polymerase III:** transcribes **tRNA** genes and one type of **rRNA** gene.
- Also unlike bacteria, RNA polymerase II needs the help of additional proteins in Eukaryotic cells, while bacteria can initiate transcription on their own.
- These additional proteins are called **General Transcription Factors**. "General" because they don't work on a specific type of gene, they work on all genes or assemble on all promoters used by RNA polymerase II.
- They are designated as TFII (Transcription Factor for Polymerase II) and listed as TFIIA, TFIIB, and so on. (TFI and TFIII for RNA polymerases 1 & 3 also exist).

General Transcription Factors: Functions

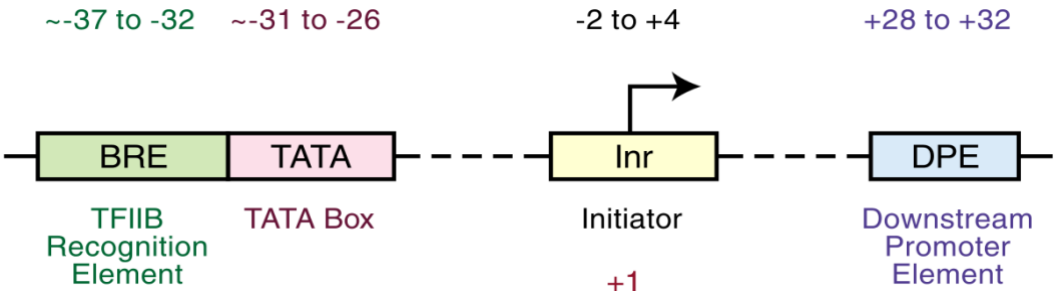
- Help **position the RNA polymerase** correctly at the **promoter.**
- Aid in **pulling apart the two strands of DNA** to allow transcription to begin. (acting as helicases)
- **Push the RNA polymerase forward** to begin transcription.

Core Components of Promotors:

- In Eukaryotic cells there are different types of promotors and not just one.
- A gene may have 1 or 2 of these different promotors.
- One promotor known as the **TATA box**, lies on **position -25** and looks like the -10 region consensus sequence found in bacteria. It is the binding site of **TFIID**. Present in only 20% of genes.

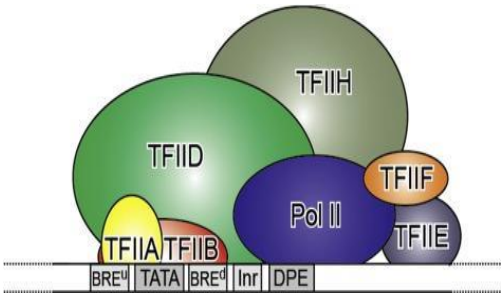
- Another one is referred to as the upstream element **(BRE)** and is the binding site of TFIIB.
- The **initiator element (Inr)** which surrounds (contains) the **+1 site.**
- Multiple **downstream promotor elements** (DPE)

** Not all of them exist at once but genes can have a combination of these promotor elements.*

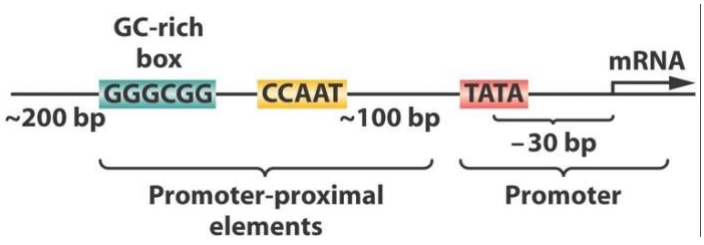


Formation of the preinitiation complex:

The RNA poly II binds to the promoter with the help of transcription factors, the most imp are TFIID and TFIIH

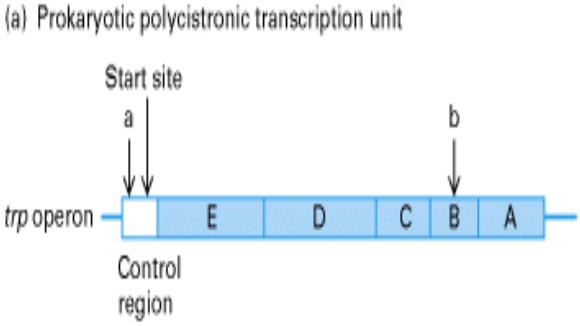


- The preinitiation complex stimulates **basal transcription**. Basal transcription is the **minimum** level of transcription occurring slowly.
- This needs activation and can be activated by other proteins which we will discuss later, that bind to **promotor proximal elements.**
- Promotor proximal elements are located **upstream** from the
- classical or core promotor.
- Promotor proximal elements have sequences that are **gene specific or tissue specific** and are bound to regulatory proteins according to a cell's needs (energy) or conditions (temperature).

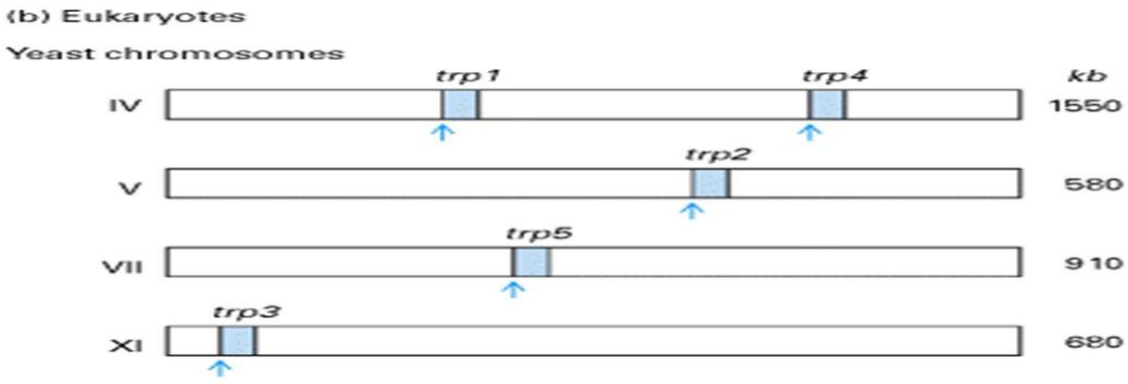


Operons vs. Promoter Proximal Elements:

- Unlike bacteria, transcription in eukaryotes does not involve **operons**, where genes that encode enzymes that are involved in related functions are located next to each other.



- Instead, the genes encoding enzymes that participate in the same mechanism (for example, metabolism of lactose or synthesis of tryptophan) are scattered in **different chromosomes**, but they have **the same promoter proximal elements**.
- Therefore, when a certain transcription factor binds to these elements, the genes are all transcribed at the **same time** - in harmony - despite their different localization, because they have the same regulatory subunit (the same PPE).

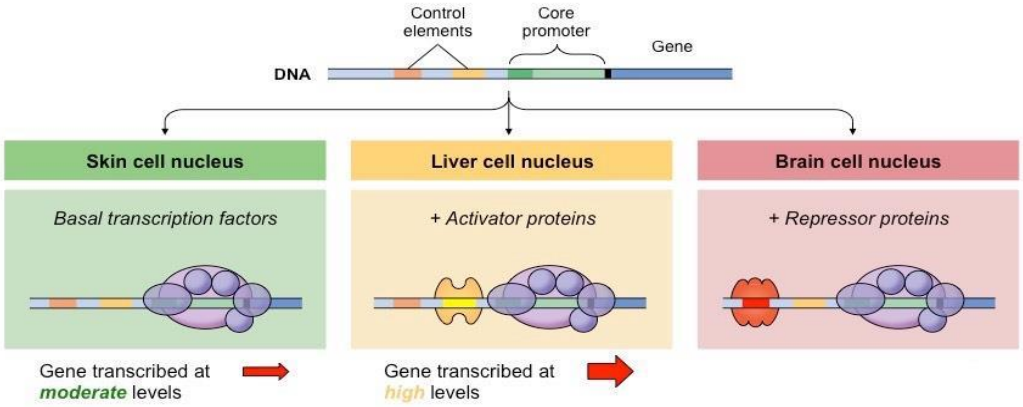


Tissue-Specific Transcription Factors:

- Tissues are distinguished by gene expression patterns; each cell has the same 21,000 genes, but not all of these genes are **expressed**.
- For example, insulin can be produced by pancreatic cells, but not by brain cells, although brain cells do have the gene for production of insulin. The gene is **present**, but it is **not active**.

- This is because the proteins, or **transcription factors**, that bind the promoter proximal element and **regulate transcription** of the

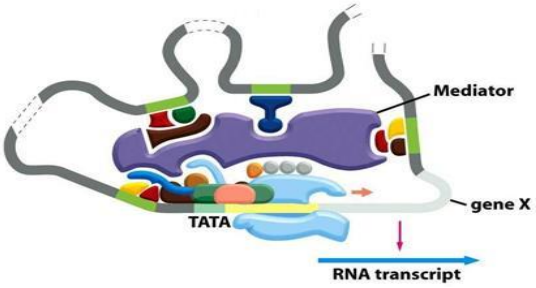
genes responsible for production of insulin are found in the pancreas, but not in the brain; the genes will not be transcribed.



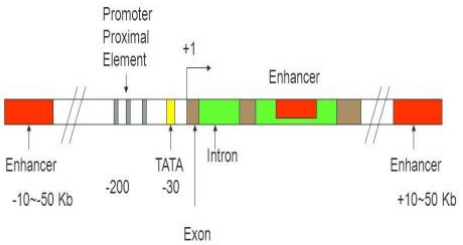
- **Differential expression of transcription factors (tissue-specific transcription factors) determine gene expression.**

Enhancers:

- Many genes are regulated by cis-acting **regulatory sequences** called **enhancers**, which are binding sites for **gene-specific** transcription factors that regulate RNA polymerase II such as a protein called the *Mediator*.
- They may be located at a **farther distance** (sometimes hundreds of kilobases) from the transcription starting site.
- The ability of enhancers to function even when separated by long distances from transcription initiation sites is possible because of **DNA looping**.



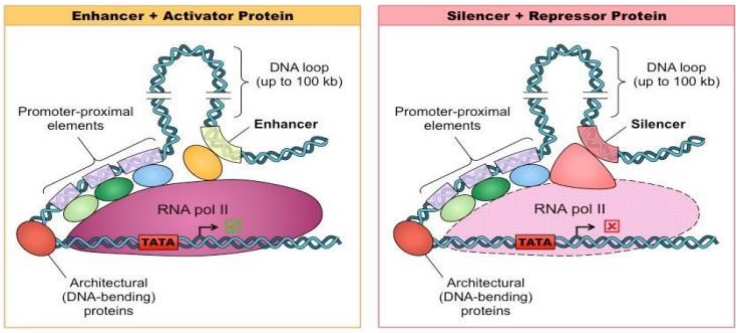
- Enhancers, like promoters, function by binding **transcription factors** that then regulate RNA polymerase.
- DNA looping allows transcription factors bound at a distant enhancer to **interact with Mediators or general transcription factors at the promoter**, stimulating transcription.
- They can regulate transcription **regardless of orientation or**



location due to DNA looping – as long as there's no mutations.

Enhancers are **not required to initiate transcription; without an enhancer, the gene is still transcribed at a low basal level. The addition of an enhancer stimulates transcription.*

- Transcription in eukaryotes can also be regulated by **silencers**, which work the same way enhancers do (through DNA looping), except they **repress** transcription instead of stimulating it.



Mechanism of Transcription

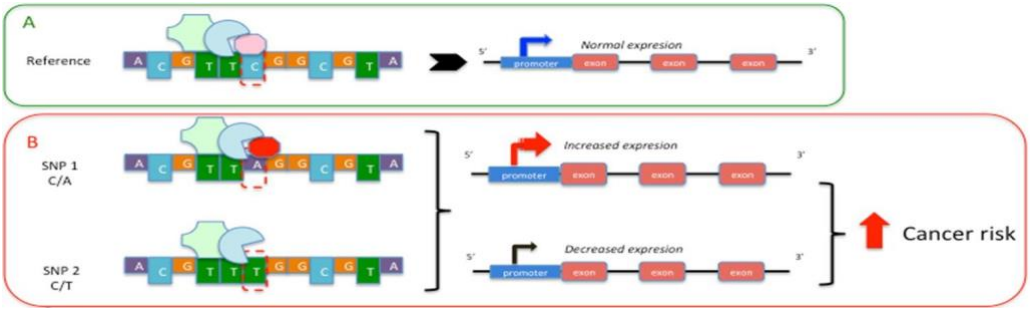
1. Initiation:

- The first step of initiation is the binding of a general transcription factor called **TFIID** to the promoter.
- The binding is followed by **recruitment of other proteins** to the promoter and formation of the **preinitiation complex**.
- Some of the aforementioned recruited proteins are RNA polymerase II and **TFIIH**.
- TFIIH can function as a **helicase** and **unwind the DNA** around the initiation site, **exposing the DNA** template to RNA polymerase and creating an **open promoter complex**.

SNPs in Promoters:

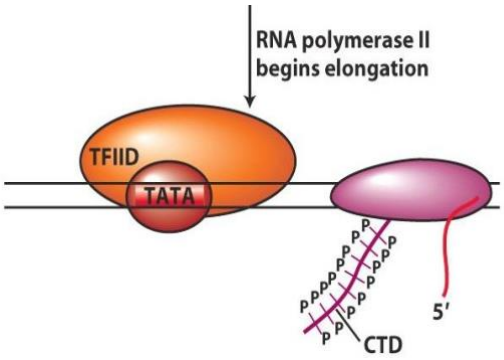
- **Single nucleotide polymorphisms** (SNP) are one of 2 types: **a**. latent or **b**. causative (causes a certain phenotype).
- Causative SNPs can either be in the **coding region** (in which case it will alter the amino acids) or the **noncoding region** (can be in the regulatory sequence, for example).
- SNPs in the promoter region can **affect the binding** of transcription factors required for expression of genes.

- These variations may lead to an **increase** or a **decrease** of the expression of the affected gene, which eventually can influence the risk of developing a disease so it affects the phenotype.



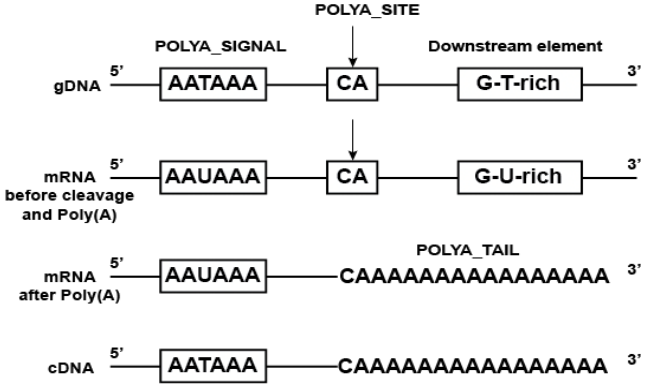
2. Elongation:

- Movement of RNA polymerase and beginning of elongation is catalyzed, again, by **TFIIH**.
- TFIIH has **protein kinase subunits**, which means it can also function as kinase and **phosphorylate the tail of RNA polymerase**.
- This allows the RNA polymerase to then dissociate and move forward.



3. Termination:

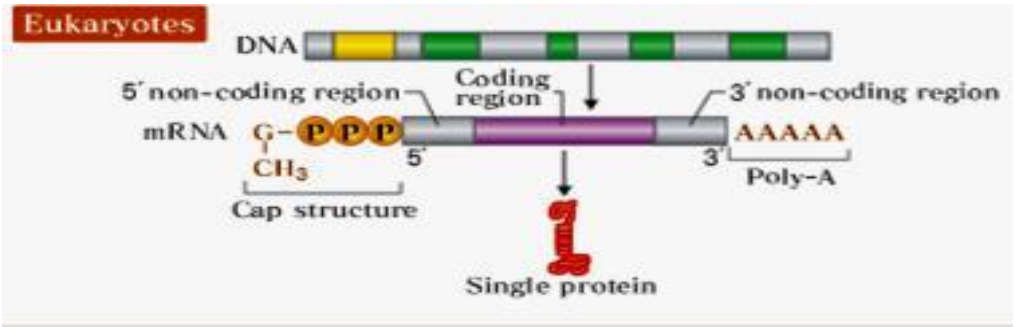
- Termination is determined by a **consensus sequence** (**not** a stop codon) for termination, which is AAUAAA, followed by CA, and 10-30 nucleotides downstream by a GU-rich sequence.
- These sequences are recognized by a complex of proteins that **cleave the RNA chain at the CA sequence**.



- An enzyme called poly-A polymerase then adds a **poly-A tail of about 200 nucleotides** to the 3'-end of the transcript. This process is called **polyadenylation**.
- Termination is coupled to the process that cleaves and polyadenylates at 3' end of the transcript**.**

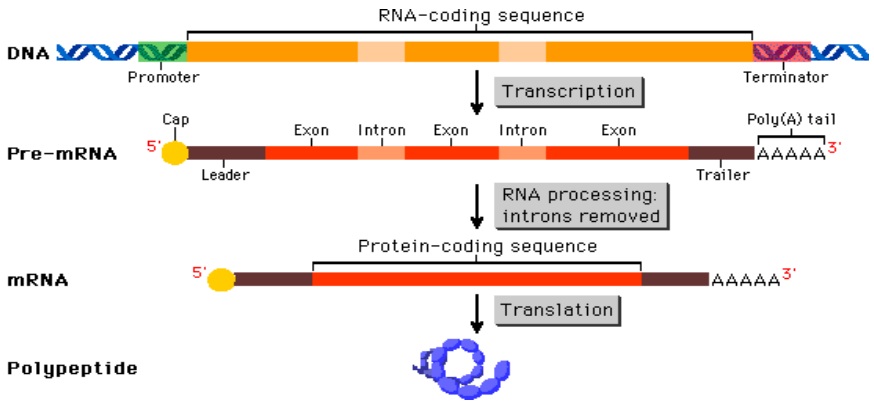
Eukaryotic Genes:

- Eukaryotic transcription units or genes produce mRNAs that encode only one polypeptide at a time, thus termed **monocistronic**.

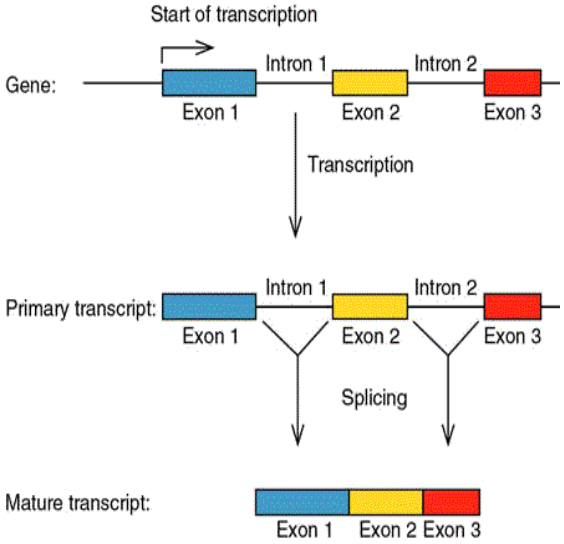


Introns vs. Exons:

- The genomes of eukaryotic cells contain specific DNA sequences that do not code for proteins known as **introns**.
- The protein-coding regions are called **exons**.
- When the RNA is synthesized, it contain **both introns and exons** and is known as **pre-mRNA**.



RNA Splicing:



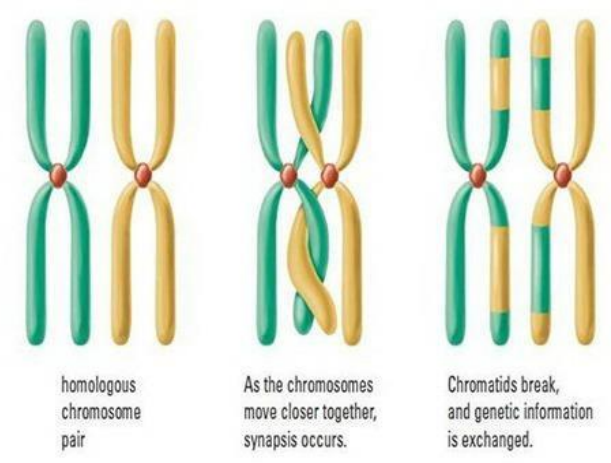
- In order to produce a mature RNA, the **intron sequences must be removed from the pre-mRNA**, allowing the remaining exons to connect to each other.
- The introns sequences are removed from the newly synthesized RNA through the process of **RNA splicing**.
- Now the RNA molecule is known as **mRNA (mature transcript)**.

Significance of Introns:

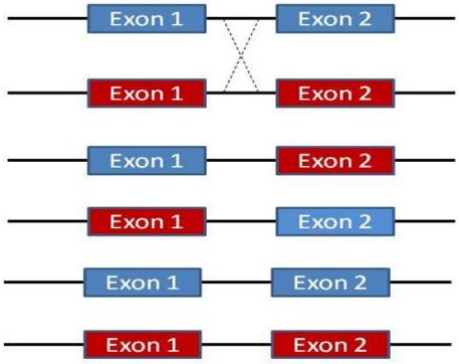
- They can **encode functional RNAs** such as nucleolar RNA (that functions in ribosomal processing) as well as **microRNAs**.
- They contain **regulatory sequence**s of gene expression, such as enhancers.
- **Variation** among individuals can be generated via **genetic recombination**.
- The exon-intron arrangement may facilitate the emergence of new proteins via **alternative splicing**.

Introns and Genetic Recombination:

- During production of haploid gametes, homologous chromosomes pair together in meiosis and **crossing over** between exons occurs.
- This will result in the production of **different germ cells**.



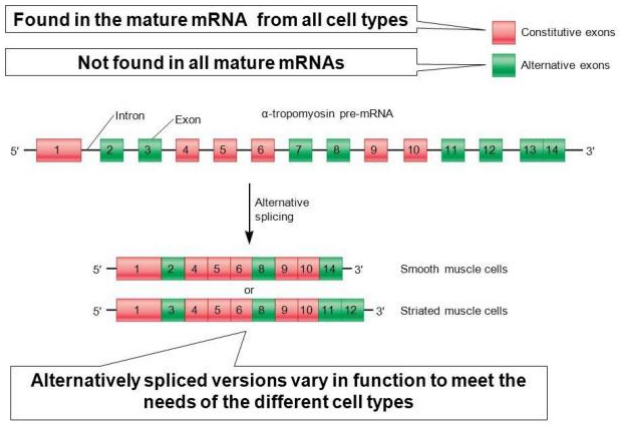
- For example, the resulting germ cell may have one exon from each parent, and it also may have two exons from the same parent, and so on.



Alternative Splicing:

- The transcripts are spliced in different ways to produce different mRNAs and **different proteins** (known as protein **isoforms**, which are highly related gene products that perform essentially the same biological function).
- It is **tissue specific**.
- An example of this is muscle cells: tropomyosin is an actin binding protein found in muscle cells. Its regulation and activity level is different among different types of muscle cells, so there are different isoforms of this protein in each type of muscle cell. This happens through alternative splicing.

**Exons that are 3' to another exon are never placed 5' to it after splicing.*



The End