Key Points

- Genes are composed of DNA arranged on chromosomes.
- Some genes encode structural or regulatory RNAs. Other genes encode proteins.
- Replication copies DNA; transcription uses DNA to make complementary RNAs; translation uses mRNAs to make proteins.
- In eukaryotic cells, replication and transcription take place within the nucleus while translation takes place in the cytoplasm.
- In prokaryotic cells, replication, transcription, and translation occur in the cytoplasm.

Key Terms

- **DNA**: a biopolymer of deoxyribonucleic acids (a type of nucleic acid) that has four different chemical groups, called bases: adenine, guanine, cytosine, and thymine
- **messenger RNA**: Messenger RNA (mRNA) is a molecule of RNA that encodes a chemical "blueprint" for a protein product.
- **protein**: any of numerous large, complex naturally-produced molecules composed of one or more long chains of amino acids, in which the amino acid groups are held together by peptide bonds

**The relationship between genes and proteins**

Genes and Proteins

- Since the rediscovery of Mendel's work in 1900, the definition of the gene has progressed from an abstract unit of heredity to a tangible molecular entity capable of replication, transcription, translation, and mutation.
- Genes are composed of DNA and are linearly arranged on chromosomes.
- Some genes encode structural and regulatory RNAs.
- There is increasing evidence from research that profiles the transcriptome of cells (the complete set all RNA transcripts present in a cell) that these may be the largest classes of RNAs produced by eukaryotic cells, far outnumbering the protein-encoding messenger RNAs (mRNAs), but the 20,000 protein-encoding genes typically found in animal cells, and the 30,o00 protein-encoding genes typically found in plant cells, nonetheless have huge impacts on cellular functioning.
- Protein-encoding genes specify the sequences of amino acids, which are the building blocks of proteins. In turn, proteins are responsible for orchestrating nearly every function of the cell. Both protein-encoding genes and the proteins that are their gene products are absolutely essential to life as we know it.

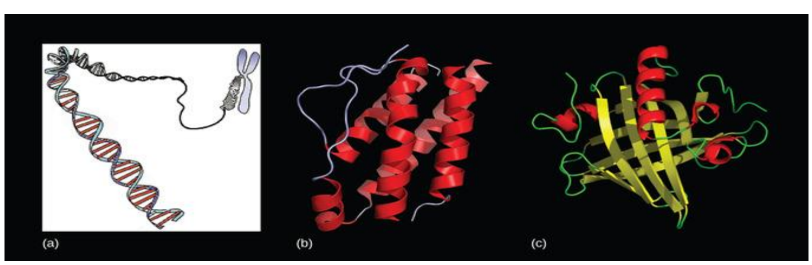



*Figure: Genes Encode Proteins: Genes, which are carried on (a) chromosomes, are linearly-organized instructions for making the RNA and protein molecules that are necessary for all of processes of life. The (b) interleukin-2 protein and (c) alpha-2u-globulin protein are just two examples of the array of different molecular structures that are encoded by genes.*

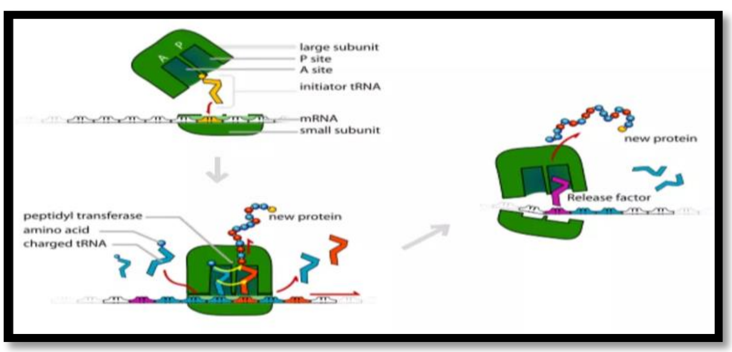
- Replication, Transcription, and Translation are the three main processes used by all cells to maintain their genetic information and to convert the genetic information encoded in DNA into gene products, which are either RNAs or proteins, depending on the gene.
- In eukaryotic cells, or those cells that have a nucleus, replication and transcription take place within the nucleus while translation takes place outside of the nucleus in cytoplasm. In prokaryotic cells, or those cells that do not have a nucleus, all three processes occur in the cytoplasm.
- **Replication** is the basis for biological inheritance. It copies a cell's DNA. The enzyme DNA polymerase copies a single parental double-stranded DNA molecule into two daughter double-stranded DNA molecules.
- **Transcription** makes RNA from DNA. The enzyme RNA polymerase creates an RNA molecule that is complementary to a gene-encoding stretch of DNA.
- **Translation** makes protein from mRNA. The ribosome generates a polypeptide chain of amino acids using mRNA as a template. The polypeptide chain folds up to become a protein.

**The protein synthesis: A translation**

**Protein synthesis**

- Is accomplished through a process called translation. After DNA is transcribed into a messenger RNA (mRNA) molecule during transcription, the mRNA must be translated to produce a protein. In translation, mRNA along with transfer RNA (tRNA) and ribosomes work together to produce proteins.

**Translation: Making Protein Synthesis Possible**




**Stages of Translation in Protein Synthesis**

1. **Initiation:** Ribosomal subunits bind to mRNA.
2. **Elongation:** The ribosome moves along the mRNA molecule linking amino acids and forming a polypeptide chain.
3. **Termination:** The ribosome reaches a stop codon, which terminates protein synthesis and releases the ribosome.

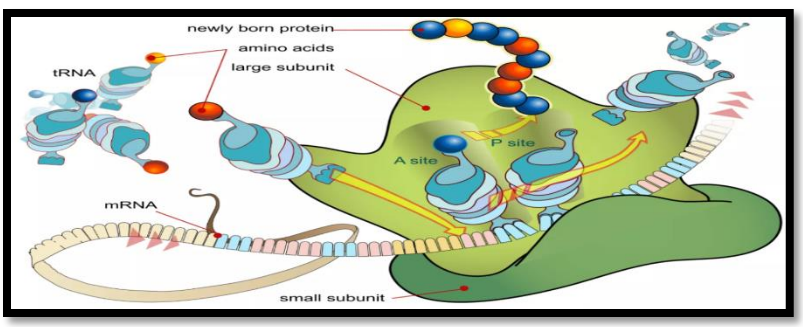
## Transfer RNA

- Transfer RNA plays a huge role in protein synthesis and translation. Its job is to translate the message within the nucleotide sequence of mRNA to a specific amino acid sequence.
- These sequences are joined together to form a protein.
- Transfer RNA is shaped like a clover leaf with three loops.
- It contains an amino acid attachment site on one end and a special section in the middle loop called the anticodon site.
- The anticodon recognizes a specific area on a mRNA called a codon.

## Messenger RNA Modifications

- Translation occurs in the cytoplasm. After leaving the nucleus, mRNA must undergo several modifications before being translated.
- Sections of the mRNA that do not code for amino acids, called introns, are removed.
- A poly-A tail, consisting of several adenine bases, is added to one end of the mRNA, while a guanosine triphosphate cap is added to the other end.
- These modifications remove unneeded sections and protect the ends of the mRNA molecule. Once all modifications are complete, mRNA is ready for translation.

## Translation




***In translation, mRNA along with tRNA and ribosomes work together to produce a protein.***
***Mariana Ruiz Villarreal/Wikimedia Commons***
***Once messenger RNA has been modified and is ready for translation, it binds to a specific site on a ribosome. Ribosome consists of two parts, a large subunit and a small subunit. They contain a binding site for mRNA and two binding sites for transfer RNA (tRNA) located in the large ribosomal subunit.***

## Initiation

- During translation, a small ribosomal subunit attaches to a mRNA molecule.
- At the same time an initiator tRNA molecule recognizes and binds to a specific codon sequence on the same mRNA molecule.
- A large ribosomal subunit then joins the newly formed complex.
- The initiator tRNA resides in one binding site of the ribosome called the **P** site, leaving the second binding site, the **A** site, open.
- When a new tRNA molecule recognizes the next codon sequence on the mRNA, it attaches to the open **A** site.
- A peptide bond forms connecting the amino acid of the tRNA in the **P** site to the amino acid of the tRNA in the **A** binding site.

**Elongation**

- As the ribosome moves along the mRNA molecule, the tRNA in the **P** site is released and the tRNA in the **A** site is translocated to the **P** site.
- The **A** binding site becomes vacant again until another tRNA that recognizes the new mRNA codon takes the open position.
- This pattern continues as molecules of tRNA are released from the complex, new tRNA molecules attach, and the amino acid chain grows.

**Termination**

- The ribosome will translate the mRNA molecule until it reaches a termination codon on the mRNA.
- When this happens, the growing protein called a polypeptide chain is released from the tRNA molecule and the ribosome splits back into large and small subunits.
- The newly formed polypeptide chain undergoes several modifications before becoming a fully functioning protein.
- Proteins have a variety of functions. Some will be used in the cell membrane, while others will remain in the cytoplasm or be transported out of the cell.
- Many copies of a protein can be made from one mRNA molecule.
- This is because several ribosomes can translate the same mRNA molecule at the same time.
- These clusters of ribosomes that translate a single mRNA sequence are called polyribosomes or polysomes.

**Transcription and RNA processing in Eukaryotic cells**

**Key points:**

- When an RNA transcript is first made in a eukaryotic cell, it is considered a **pre-mRNA** and must be processed into a **messenger RNA** (**mRNA)**.
- A **5' cap** is added to the beginning of the RNA transcript, and a 3' **poly-A tail** is added to the end.
- In **splicing**, some sections of the RNA transcript (**introns**) are removed, and the remaining sections (**exons**) are stuck back together.
- Some genes can be **alternatively spliced**, leading to the production of different mature mRNA molecules from the same initial transcript.

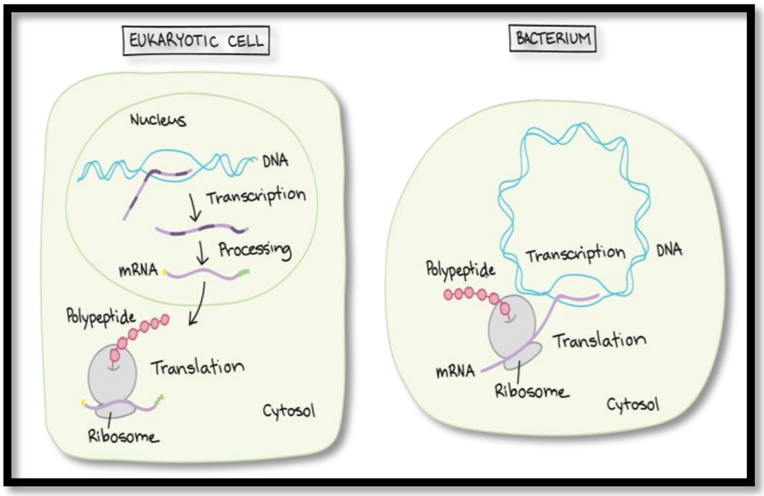
In humans and other eukaryotes, a freshly made RNA transcript (hot off the RNA polymerase "presses") is not quite ready to go. Instead, it's called a **pre-mRNA** and has to go through some processing steps to become a mature **messenger RNA** (**mRNA**) that can be translated into a protein. These include:

- Addition of cap and tail molecules to the two ends of the transcript. These play a protective role, like a book's front and back covers.

- Removal of "junk" sequences called **introns**. Introns are sort of like blank or messed-up pages made during a book's printing, which have to be removed in order for the book to be readable.
- 

**Overview of pre-mRNA processing in eukaryotes**

As a quick review, gene expression (the "reading out" of a gene to make a protein, or chunk of a protein) happens a little bit differently in bacteria and eukaryotes such as humans.




***Left panel: eukaryotic cell. In the nucleus, a pre-mRNA is produced through transcription of a region of DNA from a linear chromosome. This transcript must undergo processing (splicing and addition of 5' cap and poly-A tail) while it is still in the nucleus in order to become a mature mRNA. The mature mRNA is exported from the nucleus to the cytosol, where it is translated at a ribosome to make a polypeptide.***
***Right panel: bacterium. The DNA takes the form of a circular chromosome and is located in the cytosol. While the DNA is being transcribed to make an RNA, the RNA (which is already considered a mRNA at this point) can associate with a ribosome and start being translated to make a polypeptide.***

In bacteria, RNA transcripts are ready to act as messenger RNAs and get translated into proteins right away. In eukaryotes, things are a little more complex, though in an pretty interesting way. The molecule that's directly made by transcription in one of your (eukaryotic) cells is called a **pre-mRNA**, reflecting that it needs to go through a few more steps to become an actual messenger RNA (mRNA). These are:

- Addition of a **5' cap** to the beginning of the RNA
- Addition of a **poly-A tail** (tail of A nucleotides) to the end of the RNA
- Chopping out of **introns**, or "junk" sequences, and pasting together of the remaining, good sequences (**exons**)

Once it's completed these steps, the RNA is a mature mRNA. It can travel out of the nucleus and be used to make a protein.

**5' cap and poly-A tail**

- Both ends of a pre-mRNA are modified by the addition of chemical groups.
- The group at the beginning (5' end) is called a cap, while the group at the end (3' end) is called a tail.

- Both the cap and the tail protect the transcript and help it get exported from the nucleus and translated on the ribosomes (protein-making "machines") found in the cytosol^11start superscript, 1, end superscript.
- The **5' cap** is added to the first nucleotide in the transcript during transcription.
- The cap is a modified guanine (G) nucleotide, and it protects the transcript from being broken down.
- It also helps the ribosome attach to the mRNA and start reading it to make a protein.

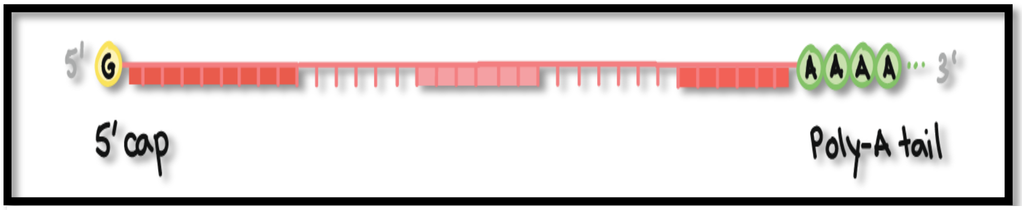



***Image of a pre-mRNA with a 5' cap and 3' poly-A tail. The 5' cap is on the 5' end of the pre-mRNA and is a modified G nucleotide. The poly-A tail is on the 3' end of the pre-mRNA and consists of a long string of A nucleotides (only a few of which are shown).***

### How is the poly-A tail added?

- The 3' end of the RNA forms in kind of a bizarre way.
- When a sequence called a **polyadenylation signal** shows up in an RNA molecule during transcription, an enzyme chops the RNA in two at that site.
- Another enzyme adds about 100100100 - 200200200 adenine (A) nucleotides to the cut end, forming a **poly-A tail**. The tail makes the transcript more stable and helps it get exported from the nucleus to the cytosol.

### RNA splicing

- Is a process that removes the intervening, non-coding sequences of genes (introns) from pre-mRNA and joins the protein-coding sequences (exons) together in order to enable translation of mRNA into a protein.

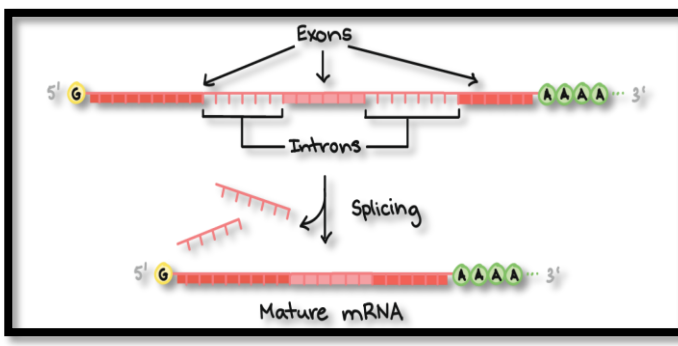



***Diagram of a pre-mRNA showing exons and introns. Along the length of the mRNA, there is an alternating pattern of exons and introns: Exon 1 - Intron 1 - Exon 2 - Intron 2 - Exon 3. Each consists of a stretch of RNA nucleotides. During splicing, the introns are revmoved from the pre-mRNA, and the exons are stuck together to form a mature mRNA that does not contain the intron sequences.***
***A key point here is that it's only the exons of a gene that encode a protein. Not only do the introns not carry information to build a protein, they actually have to be removed in order for the mRNA to encode a protein with the right sequence. If the spliceosome fails to remove an intron, an mRNA with extra "junk" in it will be made, and a wrong protein will get produced during translation.***
***[More about the spliceosome and how introns are recognized]***

### Alternative splicing

- Is the process of selecting different combinations of splice sites within a messenger RNA precursor (pre-mRNA) to produce variably spliced mRNAs. These multiple mRNAs can encode proteins that vary in their sequence and activity, and yet arise from a single gene.

- In alternative splicing, one pre-mRNA may be spliced in either of two (or sometimes many more than two!) different ways. For example, in the diagram below, the same pre-mRNA can be spliced in three different ways, depending on which exons are kept. This results in three different mature mRNAs, each of which translates into a protein with a different structure.

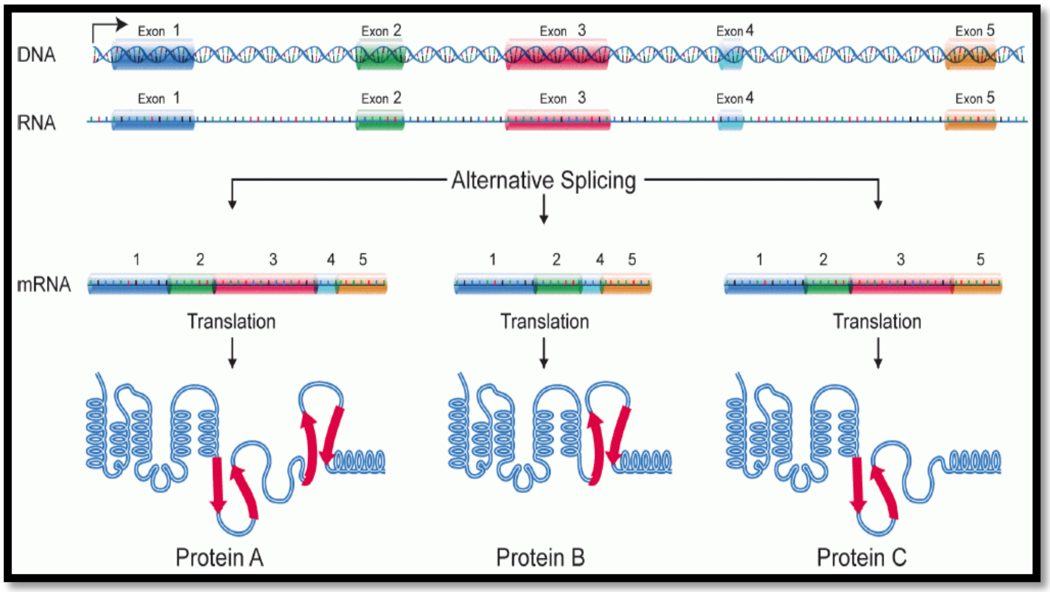



***Diagram of alternative splicing.***

*A sequence of DNA encodes a pre-mRNA transcript that contains five regions that may potentially be used as exons: Exon 1, Exon 2, Exon 3, Exon 4, and Exon 5. The exons are arranged in linear order along the pre-mRNA and have introns in between them.*

*In splicing event #1, all five exons are retained in the mature mRNA. It consists of Exon 1 - Exon 2 - Exon 3 - Exon 4 - Exon 5. When it is translated, it specifies Protein A, a protein with five domains: Coil 1 (specified by Exon 1), Coil 2 (specified by Exon 2), Loop 3 (specified by Exon 3), Loop 4 (specified by Exon 4), and Coil 5 (specified by Exon 5).*

*In splicing event #2, Exon 3 is not included in the mature mRNA. It consists of Exon 1 - Exon 2 - Exon 4 - Exon 5. When it is translated, it specifies Protein, B a protein with four domains: Coil 1 (specified by Exon 1), Coil 2 (specified by Exon 2), Loop 4 (specified by Exon 4), and Coil 5 (specified by Exon 5). It does not contain Loop 3 because Exon 3 is not present in the mRNA.*

*In splicing event #3, Exon 4 is not included in the mature mRNA. It consists of Exon 1 - Exon 2 - Exon 3 - Exon 5. When it is translated, it specifies Protein C, a protein with four domains: Coil 1 (specified by Exon 1), Coil 2 (specified by Exon 2), Loop 3 (specified by Exon 3), and Coil 5 (specified by Exon 5). It does not contain Loop 4 because Exon 4 is not present in the mRNA.*