Unit 6
Molecular Genetics

South Dakota State Standard

9-12.L.1.1 Students are able to relate cellular functions and processes to specialized structures within cells.
9-12. N.1.1 Students are able to evaluate a scientific discovery to determine and describe how societal, cultural, and personal beliefs influence scientific investigations and interpretations.

Prefix/Suffix	**Definition**
-ase	Enzyme
Deoxy-	Without oxygen
Poly-	Many
Pep-	Refers to proteins
Syn-	Together
-sis	Usually refers to a process
Ribo-	Refers to the sugar ribose

Top Vocabulary Terms:

1. Transcription
2. Translation
3. Polypeptide
4. Amino Acid
5. Enzyme
6. Protein Synthesis
7. Ribosome
8. Golgi Apparatus

I. DNA and RNA

Introduction

What tells the first cell of an organism what to do? How does that first cell know to become two cells, then four cells, and so on? Does this cell have instructions? What are those instructions and what do they really do? What happens when those instructions don't work properly? Are the "instructions" the genetic material? Though today it seems completely obvious that Deoxyribonucleic acid, or **DNA**, is the genetic material, this was not always known.

DNA and RNA

Practically everything a cell does, be it a liver cell, a skin cell, or a bone cell, it does because of proteins. It is your proteins that make a bone cell act like a bone cell, a liver cell act like a liver cell, or a skin cell act like a skin cell. In other words, it is the proteins that give an organism its traits. We know that it is your proteins that that make you tall or short, have light or dark skin, or have brown or blue eyes. But what tells those proteins how to act? It is the structure of the protein that determines what it does. And it is the order and type of **amino acids** that determine the structure of the protein. And that order and type of amino acids that make up the protein are determined by your DNA sequence.

The relatively large chromosomes that never leave the nucleus are made of DNA. And, as proteins are made on the ribosomes in the cytoplasm, how does the information encoded in the DNA get to the site of protein synthesis? That's where RNA comes into this three-player act.

DNA → RNA → Protein

That's known as the central dogma of molecular biology. It states that "DNA makes RNA makes protein." This process starts with DNA. And first DNA had to be identified as the genetic material.

The Hereditary Material

For almost 100 years, scientists have known plenty about proteins. They have known that proteins of all different shapes, sizes, and functions exist. For this reason, many scientists believed that proteins were the heredity material. It wasn't until 1928, when Frederick Griffith identified the process of transformation, that individuals started to question this concept. Transformation is a process in which a cell is altered by the uptake of DNA, usually done by bacteria. Griffith demonstrated that transformation occurs, but what was the material that caused the transforming process?

Discovery of DNA

For many years scientists performed tests to determine what the hereditary material in our cells was. In 1952, Alfred Hershey and Martha Chase were able to provide an explanation. They conclusively demonstrated that DNA is the genetic material. Later it was discovered that DNA is composed of **nucleotides**, each of which contains a nitrogen-containing base, a five-carbon sugar (deoxyribose), and a phosphate group. In these nucleotides, there is one of the four possible bases: adenine (A), guanine (G), cytosine (C), or thymine (T). In 1947, Erwin Chargaff performed some tests that proved that only certain bases bond together. Adenine always pairs with thymine, and guanine always pairs with cytosine. Because of this complementary nature of DNA, the bases on one strand determine the bases on the other strand. Adenine and guanine are known as **purines**. These bases consist of two ring structures. Purines make up one of the two groups of nitrogenous bases. Thymine and cytosine are **pyrimidines**, which have just one ring structure. By having a purine always combine with a pyrimidine in the DNA double helix, the distance between the two sugar-phosphate backbones is constant, maintaining the uniform shape of the DNA molecule.

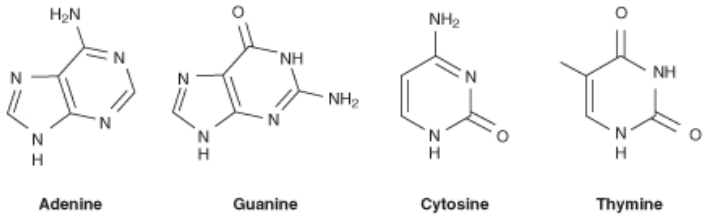


Figure 1*- Chemical structure of the four nitrogenous bases in DNA.*

What is the building block of DNA?

Figure 2- *DNA is shaped like a double helix.*

Shape of DNA

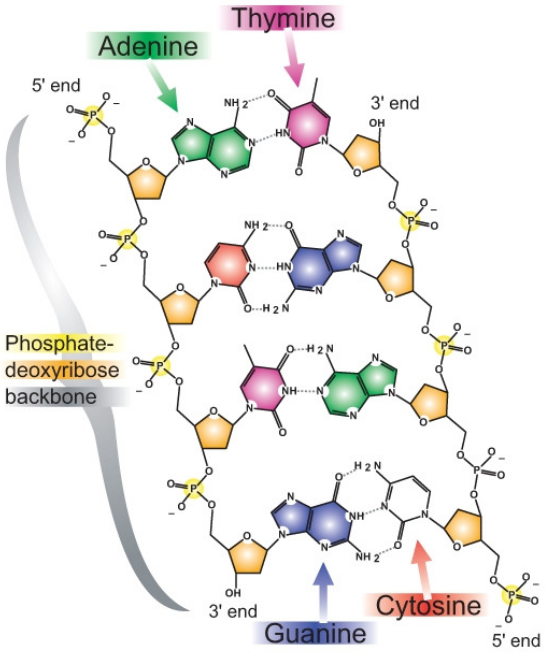
In the early 1950s, Rosalind Franklin started working on understanding the structure of DNA fibers. Franklin, together with Maurice Wilkins, used her expertise in x-ray diffraction photographic techniques to analyze the structure of DNA. In February 1953, Francis Crick and James D. Watson of the Cavendish Laboratory in Cambridge University had started to build a model of DNA. Watson and Crick indirectly obtained Franklin's DNA X-ray diffraction data demonstrating crucial information into the DNA structure. Francis Crick and James Watson then published their double helical model of DNA in *Nature* on April 25th, 1953.

DNA has the shape of a double helix, just like a spiral staircase. There are two sides, called the **sugar-phosphate backbone**, because they are made from alternating phosphate groups and deoxyribose sugars. The sugar-phosphate backbone is held together by covalent bonds. The "steps" of the double helix are made from the base pairs formed between the nitrogenous bases. The DNA double helix is held together by hydrogen bonds between the bases attached to the two strands. Adenine and thymine are bonded using a double hydrogen bond whereas guanine and cytosine are bonded using a triple hydrogen bond.

What type of bonds hold the backbone together?

a. Covalent bonds
b. Ionic bonds
c. Hydrogen bonds

Figure 3*-The base-pairing nature of DNA. Adenine always pairs with thymine, and they are held together with two hydrogen bonds. The guanine-cytosine base pair is held together with three hydrogen bonds.*



How many bonds are there between cytosine and guanine?

So it is this four letter code, made of just A, C, G, and T, that determines what the organism will become and what it will look like. How can these four bases carry so much information? This information results from the order of these four bases in the chromosomes. This sequence carries the unique genetic information for each species and each individual. Humans have about 3,000,000,000 bits of this information in each cell. A gorilla may also have close to that amount of information, but a slightly different sequence. For example, the sequence AGGTTTACCA will have different information than CAAGGGATTA. The closer the evolutionary relationship is between two species, the more similar their DNA sequences will be. For example, the DNA sequences between two species of reptiles will be more similar than between a reptile and an elm tree.

DNA sequences can be used for scientific, medical, and forensic purposes. DNA sequences can be used to establish evolutionary relationships between species, to determine a person's susceptibility to inherit or develop a certain disease, or to identify crime suspects or victims. Of course, DNA analysis can be used for other purposes as well. So why is DNA so useful for these purposes? It is useful because every cell in an organism has the same DNA sequence.

RNA

DNA → RNA → Protein

"DNA makes RNA makes protein." So what exactly is RNA? Ribonucleic acid, or **RNA**, is the other important nucleic acid in the three player act. When we say that "DNA makes RNA makes protein," what do we mean? We mean that the information in DNA is somehow transferred into RNA, and that the information in RNA is then used to make the protein.

To understand this, it helps to first understand the importance of RNA. Large chromosomes made of DNA are stored in the nucleus of a cell. These chromosomes store information on how to make a protein. However, a chromosome is too large to leave the nucleus. This is where RNA becomes useful. RNA is a small molecule that can enter and leave the nucleus. Therefore, RNA can be used as a carrier of information from the nucleus into the cytoplasm of a cell where protein synthesis occurs.

RNA Structure

RNA structure differs from DNA in three specific ways. Both are nucleic acids and made out of nucleotides; however, RNA is single stranded while DNA is double stranded. RNA contains the 5-carbon sugar ribose, whereas in DNA, the sugar is deoxyribose. Though both RNA and DNA contain the nitrogenous bases adenine, guanine and cytosine, RNA contains the nitrogenous base uracil instead of thymine. Uracil pairs with adenine in RNA, just as thymine pairs with adenine in DNA. A comparison of RNA and DNA is shown in **Table 1** and **Figure 4**.

What are the three main ways in which DNA is different than RNA?

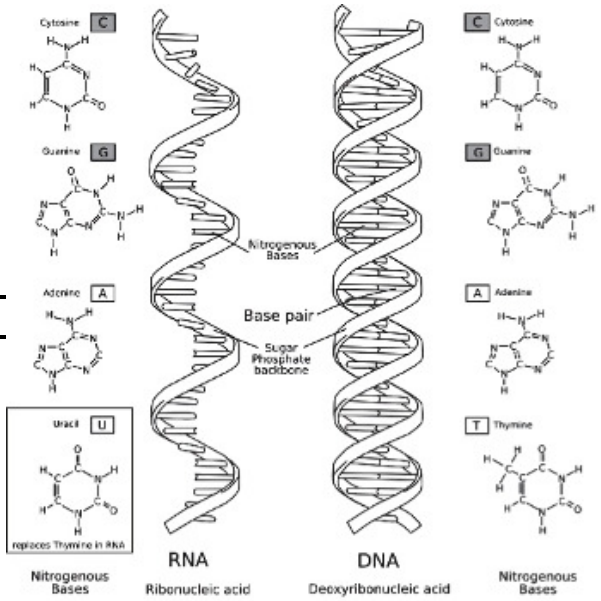


Figure 4- *A comparison of DNA and RNA structure.*

Table 1- A Comparison of RNA and DNA

	RNA	DNA
Make-up	single stranded	double stranded
Specific Base	contains uracil	contains thymine
Sugar	ribose	deoxyribose
Size	relatively small	big (chromosomes)
Location	moves to cytoplasm	stays in nucleus
Types	3 types: mRNA, tRNA, rRNA	generally 1 type

Two Types of RNA

Messenger RNA or **mRNA** is a single stranded molecule that is made in the nucleus of a cell and takes the genetic information out into the cytoplasm of the cell. Transfer RNA or **tRNA** has a "3 leaf clover" shape and is used to translate the mRNA message into a protein. They are specifically designed for this task because on one end of the molecule they carry an amino acid (a monomer of protein) and on the other end they carry an anti-codon which matches up to the codons of mRNA.

What is carried on each end of a tRNA molecule?

Which molecule can leave the nucleus?

a. tRNA *b. mRNA* *c. DNA*

II. DNA replication

DNA replication is the process in which a cell's entire DNA is copied, or replicated. This process occurs during the Synthesis (S) phase of the eukaryotic cell cycle. As each DNA strand has the same genetic information, both strands of the double helix can serve as templates for the reproduction of a new strand. This is called the **semi-conservative** model, which means that each new DNA double helix is made of one of the old strand and one of the new strand. The two resulting double helices are identical to the initial double helix.

Helicase and Polymerase

DNA replication begins as an enzyme, **DNA helicase**, breaks the hydrogen bonds holding the two strands together. The resulting structure has two branching strands of DNA backbone with exposed bases. These exposed bases allow the DNA to be "read" by another enzyme, **DNA polymerase**, which then builds the complementary DNA strand.

What is the function of DNA helicase?

Explain what semi-conservative means in your own words.

5' → 3'

Every molecule of DNA is made out of two strands of DNA that wind around each other. Each strand has a 5' and 3' end. A molecule of DNA is made of two strands of the DNA that are facing the opposite direction. Therefore the two new strands of DNA are "built" in opposite directions, through either a **leading strand** or a **lagging strand**. The leading strand is the DNA strand that DNA polymerase constructs in the 5' → 3' direction. This strand of DNA is made in a continuous manner. The lagging strand goes in the opposite direction, from 3' to 5'. DNA polymerase cannot build a strand in the 3' → 5' direction. Thus, this "lagging" strand is synthesized in short segments known as **Okazaki fragments**. Those short fragments are later fused together by the enzyme **DNA ligase**.

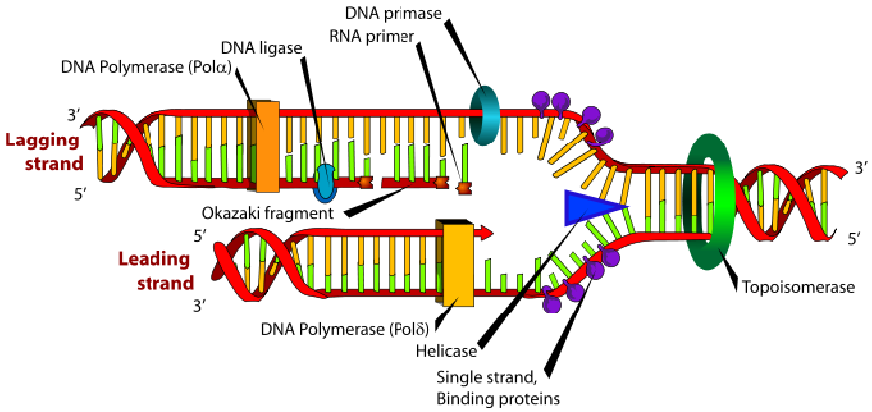


***Figure 5**-DNA replication. The two DNA strands are opened by helicase. DNA polymerase begins the replication of the leading strand. Meanwhile small fragments of DNA (Okasaki fragments) are formed on the lagging strand by DNA polymerase. Later, DNA ligase fuses the Okasaki fragments together.*

What is the difference between the lagging strand and the leading strand?

This process continues until all of the DNA in a chromosome has been copied. Each new strand that has formed is complementary to the strand used as the template. Each resulting DNA molecule is identical to the original DNA molecule. During prophase of mitosis or prophase I of meiosis, these molecules of DNA condense into a chromosome made of two identical chromatids. This process ensures that cells that result from cell division have identical sets of genetic material, and that the DNA is an exact copy of the parent cell's DNA.

What is the complementary DNA strand to the following DNA sequence? ATTACGGGA

III. Protein Synthesis- DNA → RNA → Protein

The central dogma of molecular biology describes the fundamental process that makes us all different. We all have different proteins. That is, though they may be the same types of proteins, such as we have the protein collagen found in bones, many of our proteins are slightly different and thus work slightly differently. If all our proteins acted the same way, we would all be exactly the same. But because we all have different DNA sequences, and DNA contains genes, and **genes** contain the information to encode an RNA molecule or a protein, we are all different.

So how does "DNA makes RNA makes protein" actually happen? The two processes necessary to make a protein from the information in DNA are transcription and translation. Transcription, which happens in the nucleus, uses the DNA sequence to make an RNA molecule. The RNA then leaves the nucleus and goes to the cytoplasm where translation occurs on a ribosome and produces a protein.

Transcription

Transcription is "DNA → RNA." In other words, **transcription** is the transfer of the genetic "instructions" from DNA to RNA. During transcription, a complementary copy of RNA is made. Whereas in DNA replication both strands of the DNA double helix are used as templates, in transcription only one strand is needed. RNA polymerase enzymatically "reads" a template strand of DNA, known as the coding strand, to synthesize the complementary RNA strand.

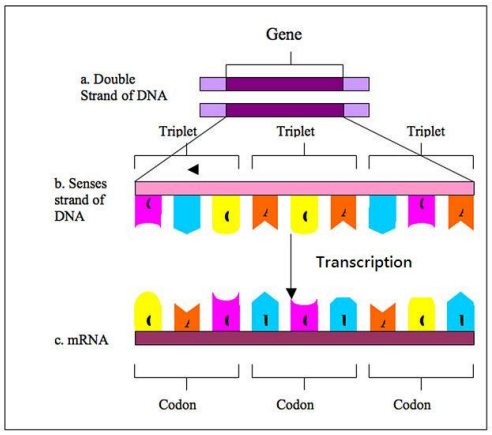


Figure 6- *Each gene (a) contains triplets of bases (b) that are transcribed into RNA (c). Every triplet, or codon, encodes for a unique amino acid.*

What is transcription?

Splicing

Humans have approximately 22,000 genes, yet make many more proteins. How? Genes can be combined in multiple ways to make many different proteins, just like letters can be arranged in many ways to make words.

Eukaryotic pre-mRNA contains introns and exons. An ***exon*** is the region of a gene that contains the code for producing a protein. Most genes contain many exons, with each exon containing the information for a specific portion of a complete protein. In many species, a gene's exons are separated by long regions of DNA that have no identified function. These long regions are **introns,** and must be removed prior to translation. ***Splicing*** is the process by which introns are removed (**Figure 7**).

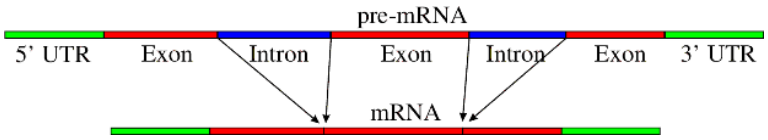


Figure 7-*Splicing, the process by which introns (blue) are removed from pre-mRNA. Exons (red) contain the information used to produce the polypeptide. There are untranslated regions (UTR) at both the beginning and end of the pre-mRNA.*

Why must mRNA be spliced before it leaves the nucleus?

The Genetic Code

So how exactly is the language of nucleotides used to code for the language of amino acids? How can a code of only As, Cs, Gs, and Us carry information for 20 different amino acids? The **genetic code** is the code in which the language of nucleotides is used to create the language of amino acids.

Cracking the Code

A code of at least three letters has to be the answer. A one letter code would only be able to code for four amino acids. A two letter code could only code for 16 amino acids. With a three letter code, there are 64 possibilities. As there are 20 amino acids, the answer must be a code of at least three letters. A code of three letters is called a codon. Each codon corresponds to a particular amino acid. The question is, how does a codon code for a particular amino acid? The answer lies with tRNA. This next step in protein synthesis is called **translation**.

What is a codon?

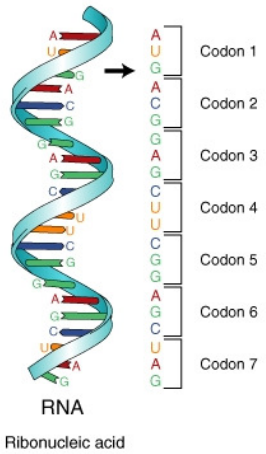


Figure 8*- The mRNA is divided into three-base segments called codons. A codon is the segment of nucleotides that codes for an amino acid, or for a start or stop signal. There are 64 codons.*

Translation

Translation is the second part of the central dogma of molecular biology: RNA → Protein. It is the process in which the genetic code in mRNA is read to make a protein. After mRNA leaves the nucleus, it moves to a **ribosome**, which consists of rRNA and proteins. The ribosome reads the sequence of codons in mRNA. Molecules of tRNA bring amino acids to the ribosome in the correct sequence. To understand the role of tRNA, you need to know more about its structure. Each tRNA molecule has an anticodon for the amino acid it carries. An anticodon is complementary to the codon for a amino acid. For example, the amino acid lysine has the codon AAG, so the anticodon is UUC. Therefore, lysine would be carried by a tRNA molecule with the anticodon UUC. Wherever the codon AAG appears in mRNA, a UUC anticodon of tRNA temporarily binds. While bound to mRNA, tRNA gives up its amino acid. Bonds form between the amino acids as they are brought one by one to the ribosome, forming a **polypeptide chain**. The chain of amino acids keeps growing until a stop codon is reached.

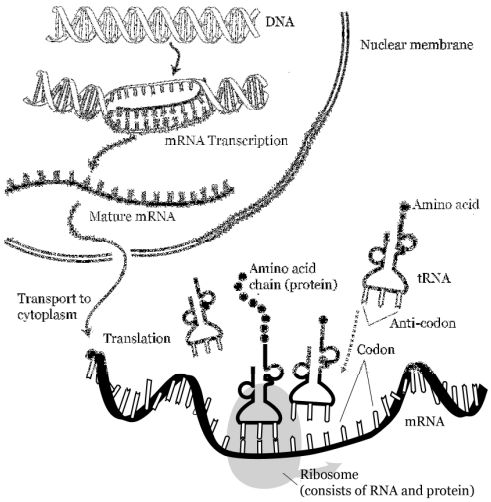


Figure 9- *Translation. Translation of the codons in mRNA to a chain of amino acids occurs at a ribosome. Find the three types of RNA in the diagram. What are their roles in translation?*

After a polypeptide chain is synthesized, it may undergo additional processes. For example, it may assume a folded shape due to interactions among its amino acids. It may also bind with other polypeptides or with different types of molecules, such as lipids or carbohydrates. Many proteins travel to the Golgi apparatus to be modified for the specific job they will do.

What is the anticodon to the codon AAU?

What amino acid does that anticodon code for? (Use the codon box below)

Amino acids that correspond to base triplets:

FIRST BASE	SECOND BASE OF A CODON				THIRD BASE
	U	C	A	G	
U	phenylalanine	serine	tyrosine	cysteine	U
	phenylalanine	serine	tyrosine	cysteine	C
	leucine	serine	STOP	STOP	A
	leucine	serine	STOP	tryptophan	G
C	leucine	proline	histidine	arginine	U
	leucine	proline	histidine	arginine	C
	leucine	proline	glutamine	arginine	A
	leucine	proline	glutamine	arginine	G
A	isoleucine	threonine	asparagine	serine	U
	isoleucine	threonine	asparagine	serine	C
	isoleucine	threonine	lysine	arginine	A
	methionine (or START)	threonine	lysine	arginine	G
G	valine	alanine	aspartate	glycine	U
	valine	alanine	aspartate	glycine	C
	valine	alanine	glutamate	glycine	A
	valine	alanine	glutamate	glycine	G

***Figure 10**-Codon Box: Use this box to determine which amino acid corresponds to each mRNA codon.*

Vocabulary

Amino acids: a monomer of proteins. Used to build polypeptide chains.
DNA: Deoxyribonucleic acid. The genetic material of a cell.
DNA helicase: an enzyme that breaks down hydrogen bonds between nitrogenous bases.
DNA ligase: an enzyme that fuses Okazaki fragments together.
DNA polymerase: an enzyme that builds complementary DNA strands by adding complementary bases.
DNA replication: a process in which a cell's entire DNA is copied.
Exon: a region of a gene that contains the code for producing a protein.
Genes: A portion of DNA that contains the information needed to make an RNA molecule.
Genetic Code: a combination of nucleotides that creates the language of DNA.
Intron: A region of a gene that does contain the code for producing proteins.
Lagging strand: The strand of DNA that is made in short segments from 3' to 5'.
Leading strand: The strand of DNA that is made in a continuous matter from 5' to 3'.
mRNA: a single stranded molecule that carries genetic information out of the nucleus into the cytoplasm.
Nucleotides: a monomer of nucleic acids. Consists of a sugar, a phosphate, and a nitrogenous base.
Okazaki fragment: short fragments of DNA that are made on the lagging strand during DNA replication.
Polypeptide chain: A chain of amino acids that can be turned into a protein.
Purines: Nitrogenous bases that have a two ring structure. Ex: Adenine and Guanine.
Pyrimidines: Nitrogenous bases that have a one ring structure. Ex: Thymine and cytosine.
Ribosom: An organelle in the cytoplasm where translation occurs.
RNA: Ribonucleic acid. Used as a messenger of genetic material.
Semi-conservative: during DNA replication, each strand serves as a template for a new strand of DNA.
Splicing: The process of removing introns from mRNA before it leaves the nucleus.
Sugar-phosphate backbone: Alternating molecules of sugar and phosphate that make up the backbone of DNA.
Transcription: The process of turning a DNA code into an mRNA code.
Translation: The process of turning an mRNA code into a polypeptide chain.
tRNA: an RNA molecule that carries an anticodon on one end and an amino acid on the other.

Works Cited:

Text and Images from: *www.ck12.org/flex*

Figure 10- Starr, Cecie, and Ralph Taggart. *Biology: The Unity and Diversity of Life*. 9th ed. Brooks/Cole, 2001. Print.