teachers'domain

Molecular Level of Genetics

Most of the molecules found in humans and other living organisms fall into one of four categories:

1. carbohydrates (sugars and starches)
2. lipids (fats, oils, and waxes)
3. proteins
4. nucleic acids

Proteins are large chain-like molecules that are twisted and folded back on themselves in complex patterns. They serve as structural material for the body, gas transporters, hormones , antibodies , neurotransmitters , and enzymes . In fact, when looking at someone, you mostly see proteins since skin and hair are primarily made of them. Proteins acting as enzymes are particularly important substances because they trigger and control the chemical reactions by which carbohydrates, lipids, and other substances are created. When you look at another human being, you mostly see proteins.

Our bodies produce about 90,000 thousand different kinds of proteins, all of which consist of simpler units called amino acids .

Proteins in all organisms are mostly composed of just 20 kinds of amino acids. Proteins differ in the number, sequence, and kinds of amino acids. Our bodies produce some of these amino acids, while others come directly from food that we consume.

AMINO ACIDS

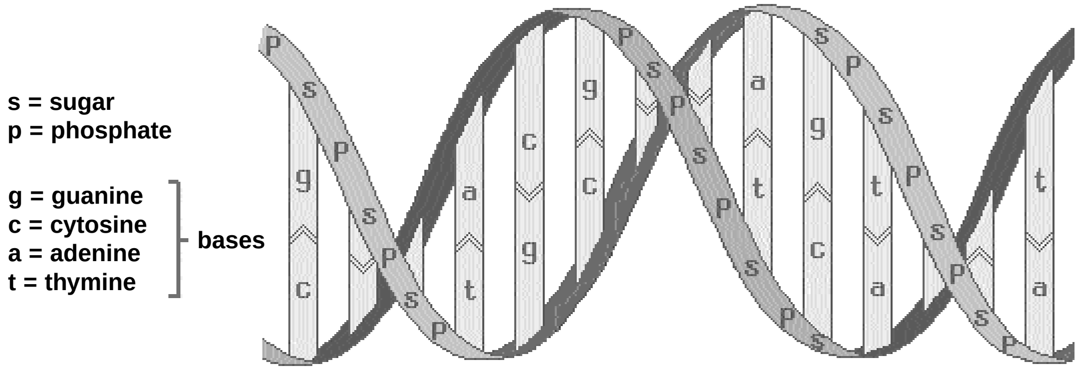
alanine	glutamic acid	leucine	serine
arginine	glutamine	lysine	threonine
asparagine	glycine	methionine	tryptophan
Aspartic acid	histidine	phenylalanine	tyrosine
cysteine	isoleucine	proline	valine

Proteins, and subsequently amino acids, are mostly made up of just four elements: carbon, oxygen, hydrogen, and nitrogen. In fact, 96.3% of your body is composed of these common elements.

The largest molecules in people and other organisms are nucleic acids. Like proteins, they consist of very long chains of simpler units. However, the components, shapes, and functions of nucleic acids differ significantly from those of proteins. There are two basic varieties of nucleic acids: DNA (deoxyribonucleic acid) and RNA (ribonucleic acid). Both play critical roles in the production of proteins.

A chromosome consists mainly of one or more very long DNA molecules. Each of these molecules contains the genetic codes, or genes, for the synthesis of many different proteins and for the regulation of other genes. In a sense, a DNA molecule is a linear sequence of permanently stored blueprints or recipes that are used regularly by our cells to make proteins out of amino acids.

DNA molecules in all living things have the shape of a double helix , which is like a twisted ladder. The sides of the ladder are composed of sugar and phosphate units, while the rungs consist of complementary pairs of four different chemical bases. Each combined sugar, phosphate, and base subunit is a nucleotide



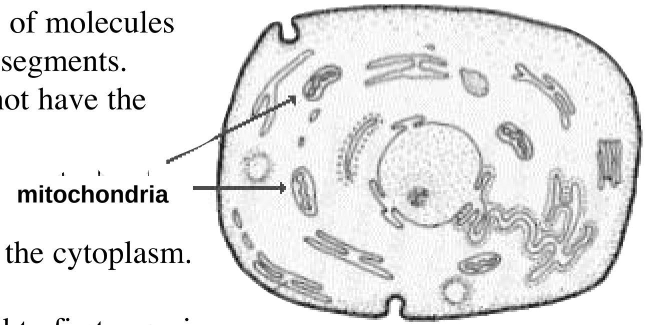
Section of a DNA molecule showing the double helix molecular shape

The sequence of bases from one nucleotide to the next in line is the code for the assembly of specific amino acids to make specific types of proteins. Therefore, a gene is essentially a specific sequence of these base pairs. The sequence need not be continuous but can be divided into different sections of a DNA molecule. Apparently, only 1.2-1.5% of the 2.9 billion base pairs in human DNA actually code for genes. These meaningful code sequences are called exons . The remaining 98+% of our DNA base pairs were in the past thought to consist merely of genetic "junk", referred to as introns . However, it is now becoming clear that much of this "junk" actually has important functions. Some of the introns act as subtle enhancers of genes. Others function as buffers against change by absorbing the mutagenic effect of viruses. Still others help determine the shape of chromosomes. It is likely that future research will discover that the "non-gene" intron code sections, that make up the bulk of DNA, perform still other important tasks.

NOTE: Textbooks written before 2001 most often indicated that there are 100,000 human genes and that 3+% of our DNA base pairs are parts of genes. These estimates were significantly reduced as a result of completion of the entire human genome mapping announced by spokesmen for the Human Genome Project in February 2001. It is now believed that there are only about 32,000 human genes. However, many of these genes apparently code for several different proteins. Now that the human genome "parts list" has been compiled, research will be focused on what these parts do--i.e., what proteins they code for and what those proteins do in our bodies.

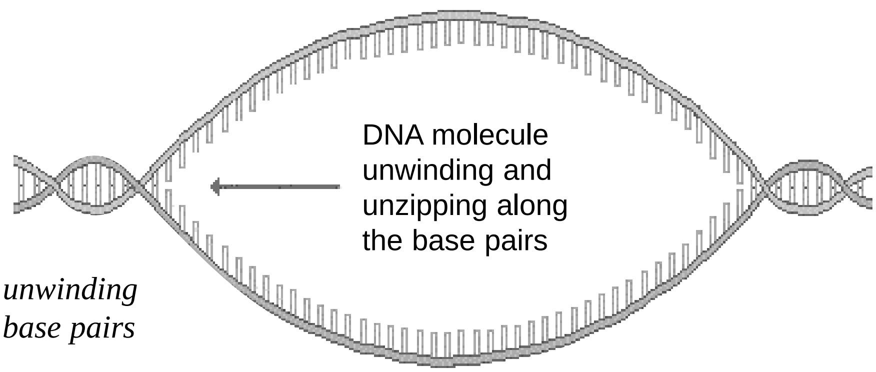
Not all of our DNA is in the cell nuclei. A small amount is in the mitochondria , which are located in the cytoplasm and mostly produce fuel for cell functions. Mitochondrial DNA (mtDNA) is normally inherited only from our mothers and is unrelated to the nuclear DNA (nDNA) in chromosomes. The 13 or more genes of mtDNA appear to have relatively few functions.

The second type of nucleic acid, RNA, consists of molecules that are single stranded copies of nuclear DNA segments. They are smaller than DNA molecules and do not have the double helix shape. In addition, the DNA base thymine is replaced by the RNA base uracil. The sugar component is also somewhat different. RNA is found in both the cell nucleus and the cytoplasm.



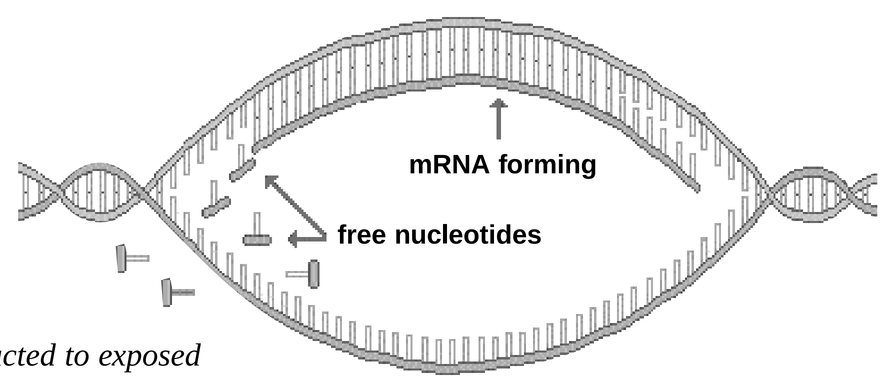
Generalized animal cell

In order to understand what RNA does, we need to first examine how the DNA code is transcribed, or copied, to RNA. The process begins by a section of a DNA moleculeunwinding and then unzipping in response to a specific enzyme. The separation occurs between the bases, as shown below.



DNA molecule partially unwinding and unzipping along the base pairs

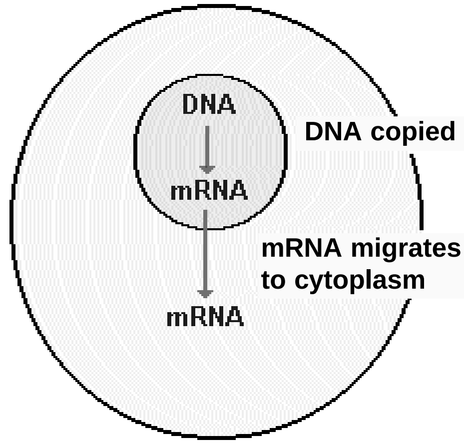
Free complementary nucleotides in the nucleus are attracted to the now unattached DNA bases on the exposed strands. The result is the formation of a messenger RNA (mRNA) molecule that is a copy, or transcription, of a specific section of the nuclear DNA moleculecorresponding to a gene. Many identical copies are made, one right after another.



Free nucleotides attracted to exposed bases of a partially unzipped DNA molecule

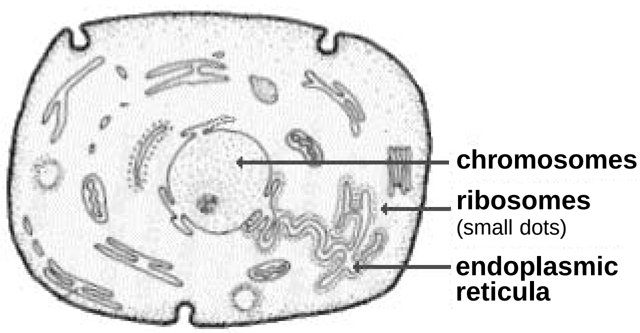
Free complementary nucleotides in the nucleus are attracted to the now unattached DNA bases on the exposed strands. The result is the formation of a messenger RNA (mRNA) molecule that is a copy, or transcription, of a specific section of the nuclear DNA molecule corresponding to a gene. Many identicopies are made, one right after another.

These new identical messenger RNA molecules then leave the nucleus and go out into the cytoplasm where the protein they are coded for is actually synthesized or assembled.



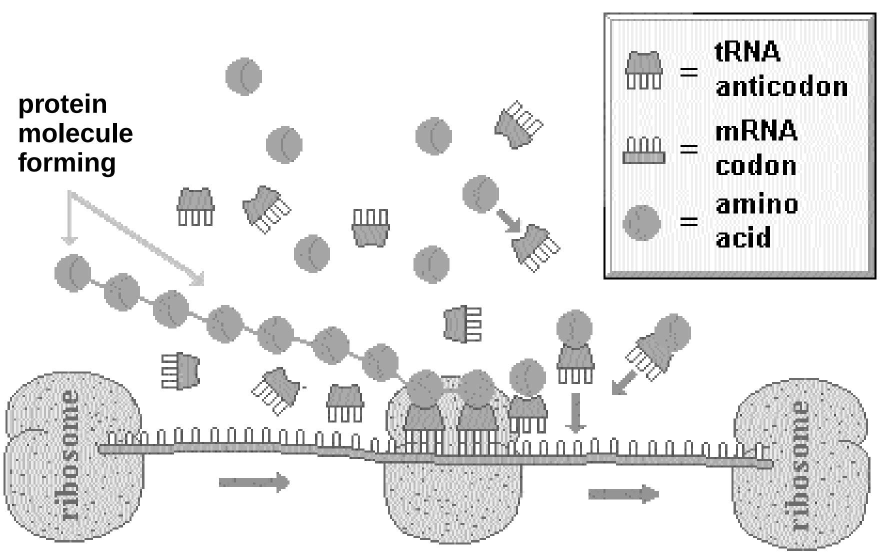
mRNA migrating out of the cell nucleus

Specifically, the messenger RNA molecules migrate from the chromosomes to the ribosomes, which are small graules in the cytoplasm. Some ribosome are on the surface of membrane networks called endoplasmic reticula , while others are free ribosomes. Assembly of proteins takes place at the site of the ribosomes



Generalized animal cell

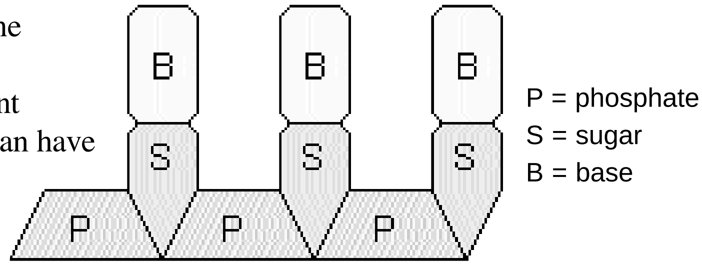
Protein synthesis begins as ribosomes move along the messenger RNA strand and attach transfer RNA (tRNA) anticodons (each with 3 bases) to triplets of complementary bases on the mRNA.



Protein synthesis at the ribosomes initiated by mRNA momentarily bonding with tRNA (Note: this schematic representation is a simplification of the actual process.)

Each transfer RNA attracts and brings a specific amino acid along with it. As a ribosome translates the messenger RNA code, a protein is assembled lineally, one amino acid at a time. Each kind of amino acid has a single codon that specifies it. A codon is a sequence of 3 nucleotide components chemically bound together (illustrated below). As mentioned above, every nucleotide consists of a sugar, a phosphate, and a base. Codons differ in terms of the sequence of their 3 bases. For example, the sequence CAG (cytosine-adenine-guanine) is a code for the amino acid glutamine.

This simple genetic code permits 64 different codons because each of the 3 nucle-otides can have 1 of the 4 bases (4 x 4 x 4 = 64). Since there are many fewer than 64 amino acids, the code system has built in redundancy--most amino acids can be attracted by transfer RNA having several different base triplets. In other words, some codons are functionally equivalent, as shown in the table below. For instance, asparagine is specified with the sequence AAU (adenine-adenine-uracil). However, AAC (adenine-adenine-cytosine) also works.



Sugar-phosphate-base chemical bond of a codon

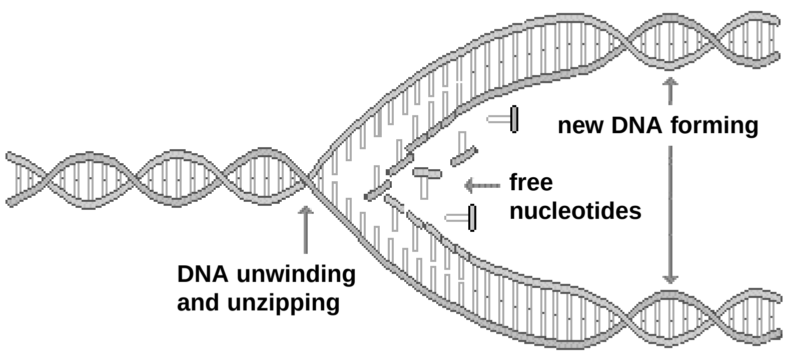
Amino Acids	DNA Codons	mRNA Codons
alanine	CGA, CGG, CGT, CGC	GCU, GCC, GCA, GCG
arginine	GCA, GCG, GCT, GCC, TCT, TCC	CGU, CGC, CGA, CGG, AGA, AGG
asparagine	TTA, TTG	AAU, AAC
aspartic acid	CTA, CTG	GAU, GAC
cysteine	ACA, ACG	UGU, UGC
glutamic acid	CTT,CTC	GAA, GAG
glutamine	GTT, GTC	CAA, CAG
glycine	CCA, CCG, CCT, CCC	GGU, GGC, GGA, GGG
histidine	GTA, GTG	CAU, CAC
isoleucine	TAA, TAG, TAT	AUU, AUC, AUA
leucine	AAT, AAC, GAA, GAG, GAT, GAC	UUA, UUG, CUU, CUC, CUA, CUG
lysine	TTT, TTC	AAA, AAG
methionine (start codon)	TAC	AUG
phenylalanine	AAA, AAG	UUU, UUC
proline	GGA, GGG, GGT, GGC	CCU, CCC, CCA, CCG
serine	AGA, AGG, AGT, AGC, TCA, TCG	UCU, UCC, UCA, UCG, AGU, AGC
threonine	TGA, TGG, TGT, TGC	ACU, ACC, ACA, ACG
tryptophan	ACC	UGG
tyrosine	ATA, ATG	UAU, UAC
valine	CAA, CAG, CAT, CAC	GUU, GUC, GUA, GUG
(stop codon)	ATT, ATC, ACT	UAA, UAG, UGA

(The DNA base thymine is replaced with uracil in the formation of mRNA.)

Not all codons specify amino acid components to be included in a protein. For instance, a start codon appears in DNA at thebeginning of the lineal code sequence for each gene and a stop codon is at the end. In other words, they indicate where a protein recipe begins and ends.

Most plant and animal cells have tens of thousands of ribosomes. Many ribosomes simultaneously translate identical strands of messenger RNA. As a result, the synthesis of proteins can be rapid and massive. These same processes can occur at the same time in millions of cells when a particular protein is needed.

In addition to keeping the blueprints for protein synthesis, DNA has one further function--it replicates, or duplicates, itself. At the beginning of this process, the parent DNA molecule unwinds and unzips along its bases beginning at one end. Then in response to an enzyme, free nucleotides pair up with corresponding bases on both of the DNA strands, as illustrated below. This results in the formation of two exact copies of the original molecule. Nuclear DNA replication occurs just before mitosis and meiosis.



DNA replication

Occasionally, an error is made in DNA replication. For example, an incorrect base pair may be included. This constitutes a mutation. If it occurs in the formation of sex cells, the mutation may be inherited and passed on in future generations. Such errors in replication are the ultimate sources of all new genes and are essential for the evolution of new species. They are also responsible for changes in somatic cells that result in the uncontrolled tumorgrowths of cancer.

It is important to realize that the genetic code system of humans is not unique but is shared by all living things. The same codons code for the same amino acids in people, dogs, fleas, and even bacteria. In addition, we share many genes with other creatures. For instance, about 90% of human genes are identical to those of a mouse. Even more surprising is the fact that more than 1/3 of our genes are shared with a primitive group of worm species known as nematodes. The universal nature of the genetic code is compelling evidence for the evolution of all organisms from the same early life forms.