Genetics

Genetics (from Ancient Greek γενετικός *genetikos*, "genitive" and that from γένεσις *genesis*, "origin"[1][2][3]), a discipline of biology, is the science of heredity and variation in living organisms.[4][5] The fact that living things inherit traits from their parents has been used since prehistoric times to improve crop plants and animals through selective breeding. However, the modern science of genetics, which seeks to understand the process of inheritance, only began with the work of Gregor Mendel in the mid-nineteenth century.[6] Although he did not know the physical basis for heredity, Mendel observed that organisms inherit traits in a discrete manner—these basic units of inheritance are now called genes.

The DNA
DNA, the molecular basis for inheritance consist of two strands. Each strand of DNA is a chain of nucleotides, matching each other in the center to form what look like rungs on a twisted ladder.

Genes correspond to regions within DNA, a molecule composed of a chain of four different types of nucleotides—the sequence of these nucleotides is the genetic information that organisms inherit. DNA naturally occurs in a double stranded form, with nucleotides on each strand complementary to each other. Each strand can act as a template for creating a new partner strand—this is the basis of the physical method for making copies of genes that can be inherited.

The sequence of nucleotides in a gene is translated by cells to produce a chain of amino acids, creating proteins—the order of amino acids in a protein corresponds to the order of nucleotides in the gene. This is known as the genetic code. The amino acids in a protein determine how it folds into a three-dimensional shape; this structure is, in turn, responsible for the protein's function. Proteins carry out almost all the functions needed for cells to live. A change to the DNA in a gene, can change a protein's amino acids, thus changing its shape and function: This can have a dramatic effect in the cell and on the organism as a whole.

Although genetics plays a large role in the appearance and behavior of organisms, it is the combination of genetics with what an organism experiences that determines the ultimate outcome. For example, while genes play a role in determining a person's height, the nutrition and health that person experiences in childhood also have a large effect on the growth of the child.

History of genetics
Sexlinked inheritance
Morgan's observation of sex-linked inheritance of a mutation causing white eyes in *Drosophila* led him to the hypothesis that genes are located upon chromosomes.

Although the science of genetics began with the applied and theoretical work of Gregor Mendel in the mid-1800s, other theories of inheritance preceded Mendel. A popular

theory during Mendel's time was the concept of blending inheritance: the idea that individuals inherit a smooth blend of traits from their parents. Mendel's work disproved this, showing that traits are composed of combinations of distinct genes rather than a continuous blend. Another theory that had some support at that time was the inheritance of acquired characteristics: the belief that individuals inherit traits strengthened by their parents. This theory (commonly associated with Jean-Baptiste Lamarck) is now known to be wrong—the experiences of individuals do not affect the genes they pass to their children.[7] Other theories included the pangenesis of Charles Darwin (which had both acquired and inherited aspects) and Francis Galton's reformulation of pangenesis as both particulate and inherited.[8]

Mendelian and classical genetics

The modern science of genetics traces its roots to Gregor Johann Mendel, a German-Czech Augustinian monk and scientist who studied the nature of inheritance in plants. In his paper "Versuche über Pflanzenhybriden" ("Experiments on Plant Hybridization"), presented in 1865 to the *Naturforschender Verein* (Society for Research in Nature) in Brünn, Mendel traced the inheritance patterns of certain traits in pea plants and described them mathematically.[9] Although this pattern of inheritance could only be observed for a few traits, Mendel's work suggested that heredity was particulate, not acquired, and that the inheritance patterns of many traits could be explained through simple rules and ratios.

The importance of Mendel's work did not gain wide understanding until the 1890s, after his death, when other scientists working on similar problems re-discovered his research. William Bateson, a proponent of Mendel's work, coined the word *genetics* in 1905.[10][11] (The adjective *genetic*, derived from the Greek word *genesis* - γένεσις, "origin" and that from the word *genno* - γεννώ, "to give birth", predates the noun and was first used in a biological sense in 1860.)[12] Bateson popularized the usage of the word *genetics* to describe the study of inheritance in his inaugural address to the Third International Conference on Plant Hybridization in London, England, in 1906.[13]

After the rediscovery of Mendel's work, scientists tried to determine which molecules in the cell were responsible for inheritance. In 1910, Thomas Hunt Morgan argued that genes are on chromosomes, based on observations of a sex-linked white eye mutation in fruit flies.[14] In 1913, his student Alfred Sturtevant used the phenomenon of genetic linkage to show that genes are arranged linearly on the chromosome.[15]

Molecular genetics

JamesDWatson.
James D. Watson and Francis Crick determined the structure of DNA in 1953.

Although genes were known to exist on chromosomes, chromosomes are composed of both protein and DNA—scientists did not know which of these was responsible for inheritance. But in 1928, Frederick Griffith discovered the phenomenon of transformation.

He discovered that dead bacteria could transfer genetic material to "transform" other still-living bacteria. Sixteen years later, in 1944, Oswald Theodore Avery, Colin McLeod and Maclyn McCarty identified the molecule responsible for transformation as the DNA.[16] The Hershey-Chase experiment in 1952 also showed that DNA (rather than protein) was the genetic material of the viruses that infect bacteria, providing further evidence that DNA was the molecule responsible for inheritance.[17]

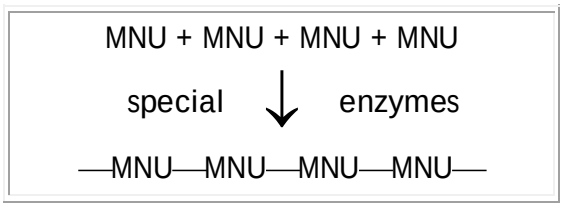
James D. Watson and Francis Crick determined the structure of DNA in 1953, using the X-ray crystallography work of Rosalind Franklin who showed that DNA had a helical structure (i.e., shaped like a corkscrew).[18][19] Their double-helix model had two strands of DNA with the nucleotides pointing inward, each matching a complementary nucleotide on the other strand to form what looks like rungs on a twisted ladder.[20] This structure showed that genetic information exists in the sequence of nucleotides on each strand of DNA. The structure also suggested a simple method for duplication: if the strands are separated, new partner strands can be reconstructed for each based on the sequence of the old strand.

Although the structure of DNA showed how inheritance worked, it was still not known how DNA influenced the behavior of cells. In the following years, scientists tried to understand how DNA controls the process of protein production. It was discovered that the cell uses DNA as a template to create matching messenger RNA (a molecule with nucleotides, very similar to DNA). The nucleotide sequence of a messenger RNA is used to create an amino acid sequence in protein; this translation between nucleotide and amino acid sequences is what is known as the genetic code.

With this molecular understanding of inheritance, an explosion of research became possible. One important development was chain-termination DNA sequencing in 1977 by Frederick Sanger: this technology allows scientists to read the nucleotide sequence of a DNA molecule.[21] In 1983, Kary Banks Mullis developed the polymerase chain reaction, providing a quick way to isolate and amplify a specific section of a DNA from a mixture.[22] Through the pooled efforts of the Human Genome Project and the parallel private effort by Celera Genomics, these and other techniques culminated in the sequencing of the human genome in 2003.[23]

Nucleotide Structure

Nucleic acids (both RNA and DNA) are **polymers** made up of monomers called **mononucleotide units** (MNU in the diagram). These mononucleotide units are joined together by **intermolecular dehydration** reactions that form phosphate ester bonds. Those reactions are, of course, catalyzed by specialized enzymes.

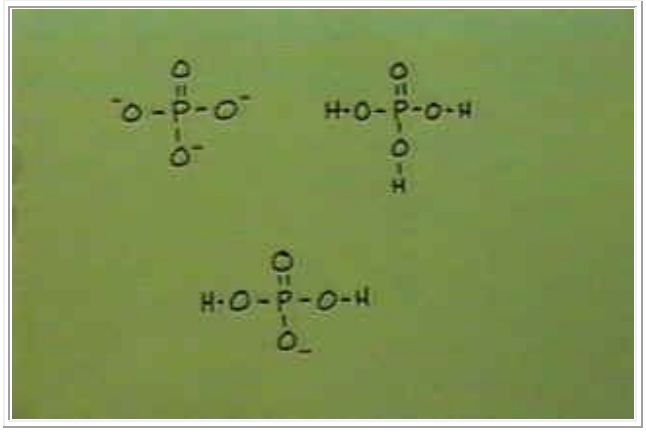


The nucleotide units themselves are made up of smaller types of components. Each nucleotide contains a **phosphate** unit, a **sugar** unit, and a **heterocyclic base** unit. (Also look at the diagram at the bottom of Example 4 in your workbook.)

MNU = Phosphate—Sugar—Base

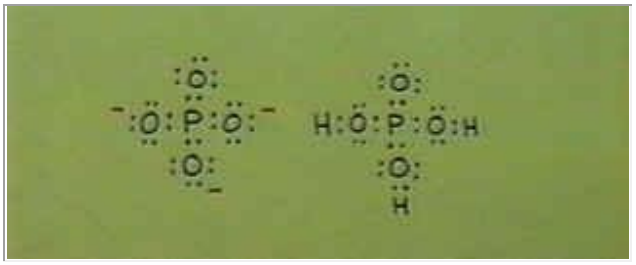
Phosphate

The phosphate unit can be represented either as a **phosphate ion** or as **phosphoric acid** molecule. As you look at various representations of this in different places, you will see both approaches used. Because of the acid strength of phosphoric acid and the base strength of phosphate ion, something part way in between, such as the **dihydrogen phosphate ion**, is probably closer to the truth as far as the form in which the phosphate actually exists in solution.



Note that these representations show the top oxygen atom bonded to the phosphorous by what appears to be a double bond. Its not really a double bond, but it serves to satisfy our presumed requirement that oxygen has two bonds.

If you look at the electron dot representation of either the ion or the molecule, you will see that there are two electrons (one pair) not four electrons (two pairs) shared between the oxygen and the phosphorous. So, it is really not a double bond, but yet the oxygen does have all the eight electrons that it is supposed to have in its valence shell.



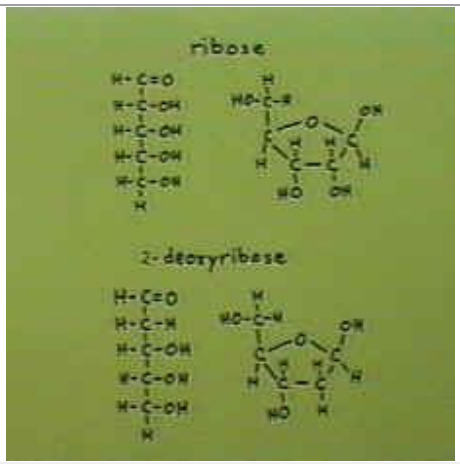
Note also, that when you look at the phosphate ion in this way, there's really no difference between the three oxygen atoms that are presumed to "have a

charge" and the one oxygen atom that "doesn't have a charge."

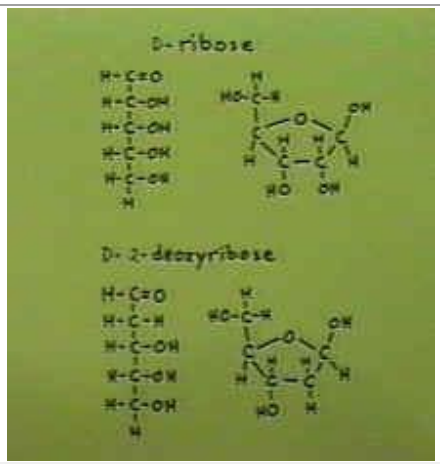
Another important thing to keep in mind as we look at the structure of nucleic acids is that **two** of the oxygens (or **OH**s) of the phosphate group (or phosphoric acid molecule) will be used to bond this unit to sugar molecules.

Sugar

The **sugar** that is part of a nucleotide is a 5-carbon atom sugar in its ring form. It will either be **ribose in RNA** or **deoxyribose in DNA**. The "deoxy" simply means that the ribose molecule has lost an oxygen. That missing oxygen happens to be from the second carbon, so the more correct name for deoxyribose is **2-deoxyribose**. (These are also shown in Examples 4 and 5 in your workbook.)

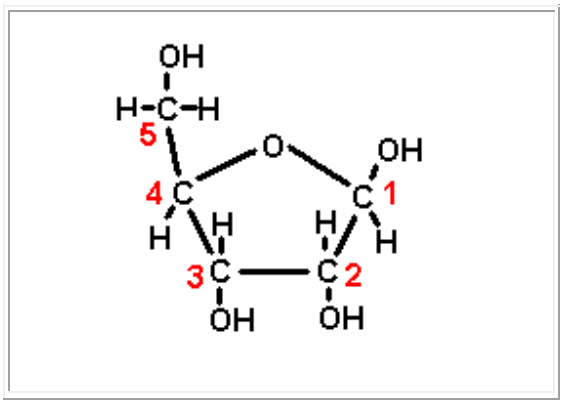


Because the last asymmetric carbon atom has an OH to the right, these molecules are sometimes given the more complete names of **D-ribose** and also **D-2-deoxyribose**. More commonly they are simply referred to as ribose and deoxyribose.



It is the ring form of the ribose and deoxyribose that are used in the nucleotide units. Each of the **OH**s in these molecules serves a particular function. Ribose is shown here, but these comments apply to deoxyribose as well.

- The double-bonded oxygen on the first carbon in the linear form becomes the beta OH that is used to bond to a base unit.
- The OH in the second position serves to distinguish between the ribose in RNA and the deoxyribose in DNA.
- The OH on the third carbon will bond to the phosphate group of other nucleotides.
- The OH group on the fourth carbon is involved in the closure of the ring.
- The OH group on the fifth carbon is what bonds to the phosphate unit of this particular nucleotide.

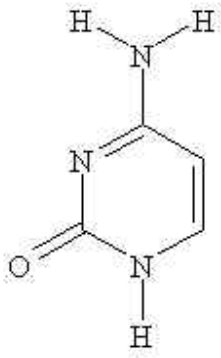
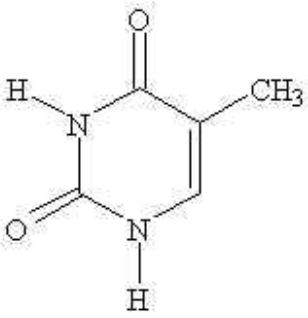
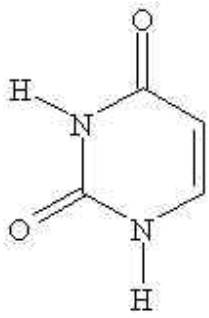


Heterocyclic Bases

Several different bases are found in nucleotides. They are heterocyclic bases or sometimes referred to as nitrogenous bases because they contain nitrogen within the rings. The fact that they are bases is actually irrelevant for the function that they serve and we really won't be paying attention to their base properties. (Structures for these compounds are also shown at the top of Example 4 in your workbook.)

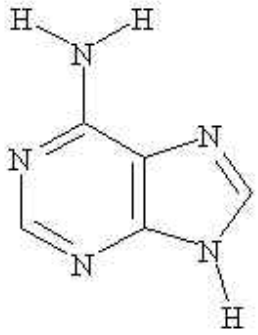
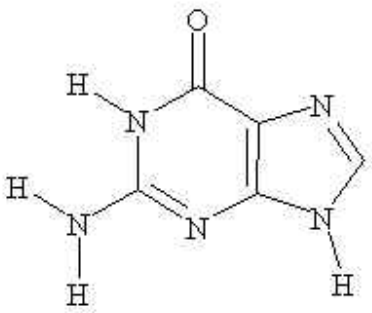
Pyrimidines

Some have **one ring** and are similar in structure to the compound **pyrimidine** and, because of that, they are called the pyrimidines or the pyrimidine bases. There are three of them and they are called **cytosine** which is found in both DNA and RNA, **thymine**, which is found only in DNA, and **uracil**, which is found only in RNA. The abbreviations **C**, **T** and **U** will be used extensively to refer to these compounds.

		
cytosine (C)	**thymine (T)**	**uracil (U)**

Purines

Some of these heterocyclic bases have **two rings** like the compound **purine** and, therefore, they are called the purines or the purine bases. They are **adenine** and **guanine**, represented by the letters **A** and **G**, and they are both found in DNA and RNA.

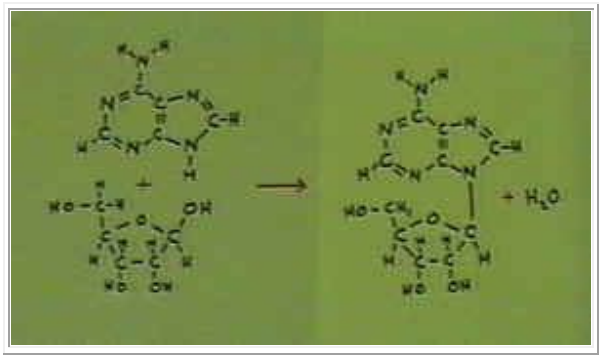
	
adenine (A)	**guanine (G)**

For future reference, it is important to note that within these compounds **one NH** will be involved in the bond **to a sugar molecule** (bottom right in these diagrams). The **other NH's** (and also Ns and double-bonded oxygens) **will hydrogen bond to other bases**. Similarly, in the pyrimidines, one NH (bottom in these diagrams) will be used to bond to the sugar molecule, either ribose or the deoxyribose. The other NHs and also double-bonded oxygens and nitrogen will be used to form hydrogen bonds to other bases.

Nucleosides

Now let's put the parts together. One of these heterocyclic bases bonded to a sugar molecule makes a **nucleoside**.

For example, when the heterocyclic base **adenine** bonds with the sugar molecule **ribose** by an intermolecular dehydration the nucleoside **adenosine** is formed. A water molecule is formed and a bond is formed between a nitrogen atom in the base and a carbon atom in the ribose. (This is also shown in Exercise 6 in your workbook.)



Practice

To get some practice working with the combination of a heterocyclic base with a sugar molecule to **make a nucleoside**, write equations using structural formulas in which you will combine uracil with sugar and then also thymine with sugar to make their corresponding nucleosides. (These are also shown in Exercise 7 in your workbook.)

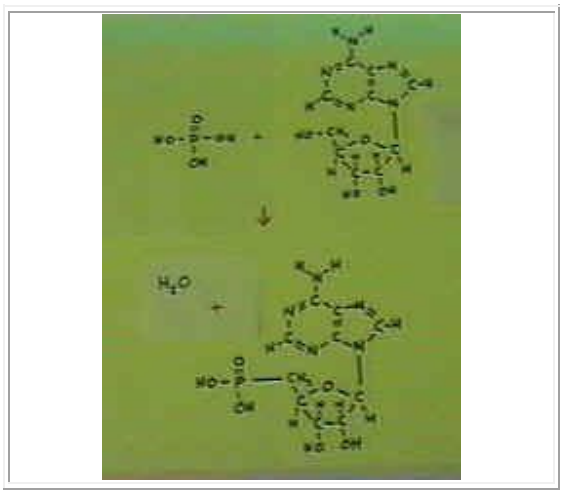
uracil + ribose → uridine
thymine + deoxyribose → deoxythymidine

Take some time now to draw the structural formulas that represent those reactions. Check with the instructor when you come to the lab to make sure that you have drawn these correctly.

Nucleotides

The next step is to combine the nucleoside with a phosphate to form a nucleotide.

In this case adenosine combines with the phosphate (or phosphoric acid) to form the nucleotide adenosine monophosphate. This is a **dehydration reaction** in water is released and a phosphate ester bond is formed. (This reaction equation is also shown in Example 8 in your workbook. The structure in that example shows the oxygen atom that is missing from the phosphoester bond in this diagram. See if you can find where the missing oxygen atom should be.)



This nucleotide is often called **adenosine monophosphate** because it is made

from adenosine with one phosphate group attached. It is quite often abbreviated as **AMP**. It can also be called **adenylic acid**.

Practice

To get practice working with the **formation of nucleotides**, write equations using structural formulas for the following reactions, which involve the formation of two additional nucleotides. (These are also shown in Exercise 9 in your workbook.) Try to avoid being distracted by the names, they are not our focus here. Instead, concentrate on the structures and the manner in which parts join together to make the nucleotides.

cytidine	+	phosphate or phosphoric acid	→	cytidine monophosphate (CMP) or cytidilic acid
guanine + ribose	+	phosphate or phosphoric acid	→	guanosine monophosphate (GMP) or guanilic acid

Check your answers with the instructor when you come into the lab.

Mendelian inheritance APH202

Punnett square Mendel's flowers
A Punnett square depicting a cross between two pea plants heterozygous for purple (B) and white (b) blossoms

At its most fundamental level, inheritance in organisms occurs by means of discrete traits, called genes.[24] This property was first observed by Gregor Mendel, who studied the segregation of heritable traits in pea plants.[9][25] In his experiments studying the trait for flower color, Mendel observed that the flowers of each pea plant were either purple or white - and never an intermediate between the two colors. These different, discrete versions of the same gene are called alleles.

In the case of pea plants, each organism has two alleles of each gene, and the plants inherit one allele from each parent.[26] Many organisms, including humans, have this pattern of inheritance. Organisms with two copies of the same allele are called homozygous, while organisms with two different alleles are heterozygous.

The set of alleles for a given organism is called its genotype, while the observable trait the organism has is called its phenotype. When organisms are heterozygous, often one allele is called dominant as its qualities dominate the phenotype of the organism, while

the other allele is called recessive as its qualities recede and are not observed. Some alleles do not have complete dominance and instead have incomplete dominance by expressing an intermediate phenotype, or codominance by expressing both alleles at once.[27]

When a pair of organisms reproduce sexually, their offspring randomly inherit one of the two alleles from each parent. **These observations of discrete inheritance and the segregation of alleles are collectively known as Mendel's first law or the Law of Segregation.**

Notation and diagrams

Pedigree-chart-example.
Genetic pedigree charts help track the inheritance patterns of traits.

Geneticists use diagrams and symbols to describe inheritance. A gene is represented by a letter (or letters)—the capitalized letter represents the dominant allele and the recessive is represented by lowercase.[28] Often a "+" symbol is used to mark the usual, non-mutant allele for a gene.

In fertilization and breeding experiments (and especially when discussing Mendel's laws) the parents are referred to as the "P" generation and the offspring as the "F1" (first filial) generation. When the F1 offspring mate with each other, the offspring are called the "F2" (second filial) generation. One of the common diagrams used to predict the result of cross-breeding is the Punnett square.

When studying human genetic diseases, geneticists often use pedigree charts to represent the inheritance of traits.[29] These charts map the inheritance of a trait in a family tree.

Interactions of multiple genes

Galton-height-regress.

Human height is a complex genetic trait. Francis Galton's data from 1889 shows the relationship between offspring height as a function of the mean of the parents' height. While correlated, remaining variation in offspring heights indicates environment is also an important factor in this trait.

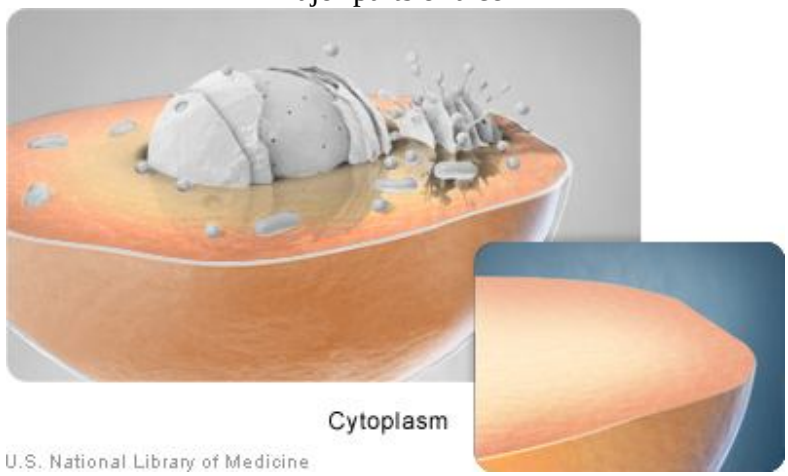
Organisms have thousands of genes, and in sexually reproducing organisms assortment of these genes are generally independent of each other. This means that the inheritance of an allele for yellow or green pea color is unrelated to the inheritance of alleles for white or purple flowers. **This phenomenon, known as "Mendel's second law" or the "Law of independent assortment", means that the alleles of different genes get shuffled between parents to form offspring with many different combinations.** However, some genes do not assort independently, demonstrating genetic linkage, as discussed later

Often different genes can interact in a way that influences the same trait. In the white Leghorn and white broiler breeds, for example, there exists a gene with alleles that determine the color of the feathers. Another gene, seem to control whether the feathers are white or colored. When a bird has two copies of this white allele, its feathers are white - regardless of whether the one of the parents carry colored alleles. This interaction between genes is called epistasis, with the white gene epistatic to color .

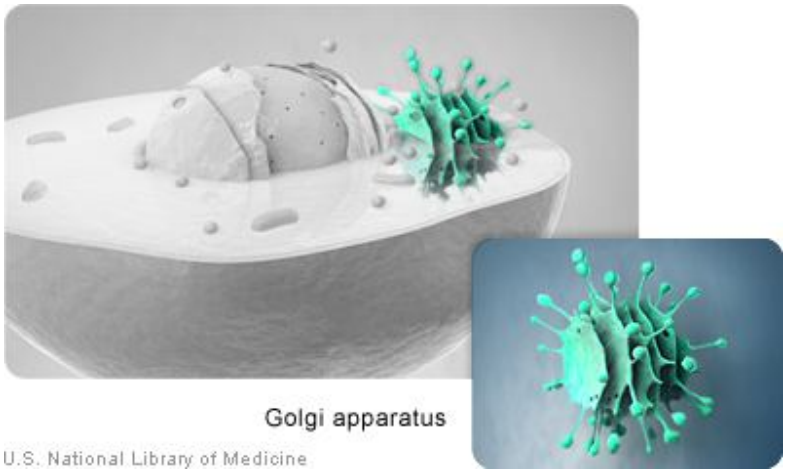
Many traits are not discrete features (eg. purple or white flowers) but are instead continuous features (eg. human height and skin color). These complex traits are the product of many genes.[31] The influence of these genes is mediated, to varying degrees, by the environment an organism has experienced. The degree to which an organism's genes contribute to a complex trait is called heritability.[32] Measurement of the heritability of a trait is relative - in a more variable environment, the environment has a bigger influence on the total variation of the trait. For example, human height is a complex trait with a heritability of 89% in the United States. In Nigeria, however, where people experience a more variable access to good nutrition and health care, height has a heritability of only 62%.[33]

THE CELL ABG504/ABG701/ANS705

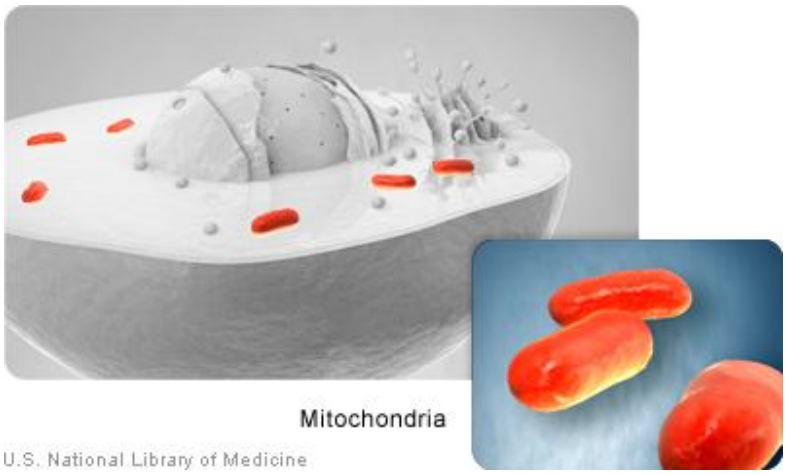
Major parts of a cell



The cytoplasm surrounds the cell’s nucleus and organelles.



The Golgi apparatus is involved in packaging molecules for export from the cell.

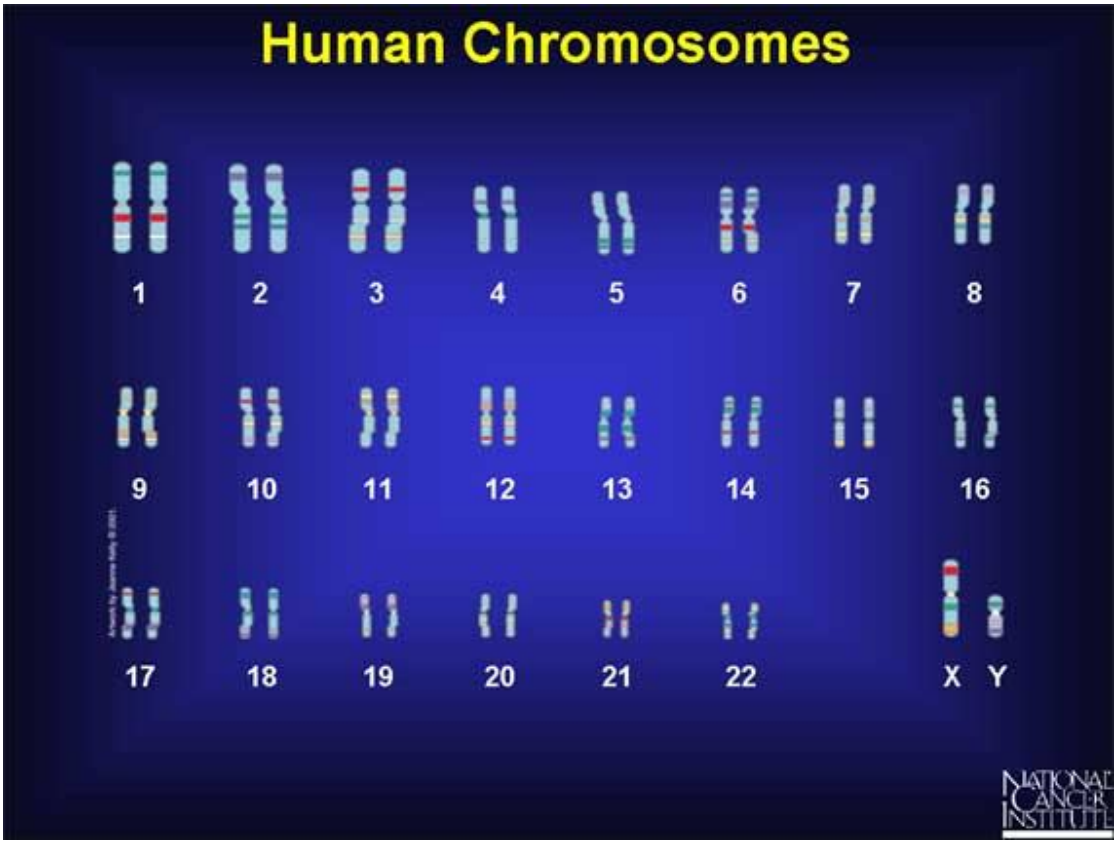


Mitochondria provide the cell’s energy.

What are Chromosomes

Human cells contain two sets of chromosomes, one inherited from the mother and one from the father.

Each set has 23 single chromosomes--22 autosomes and a sex-determining chromosome, either X or Y. The set shown here is from a male, since it contains an X and a Y chromosome; if the chromosome set were from a female, it would contain an X and an X.



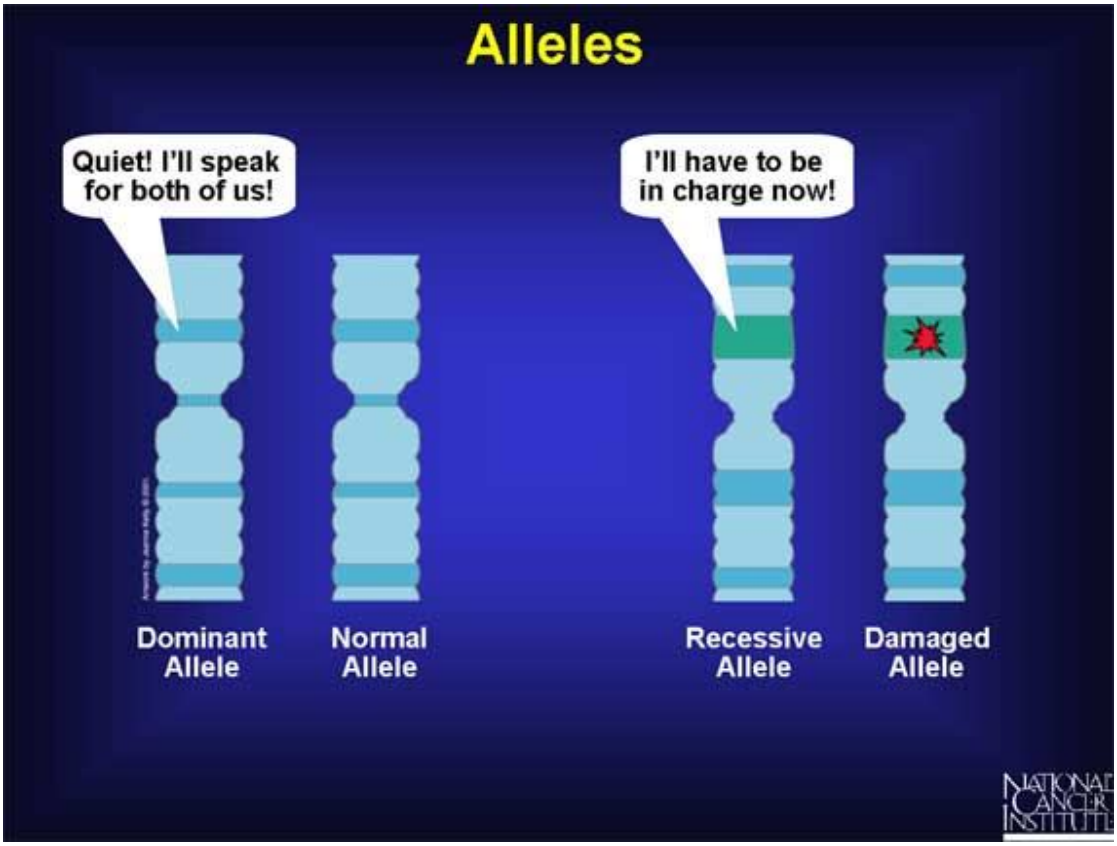
BIOINFORMATICS

Over the past few decades, major advances in the field of molecular biology, coupled with advances in genomic technologies, have led to an explosive growth in the biological information generated by the scientific community. This deluge of genomic information has, in turn, led to an absolute requirement for computerized databases to store, organize, and index the data and for specialized tools to view and analyze the data.

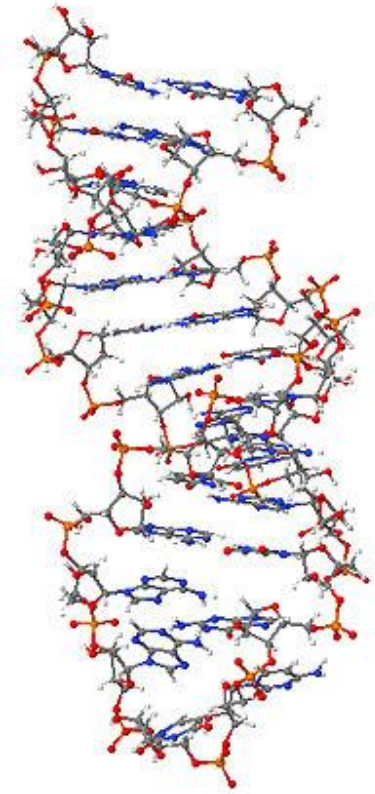
The completion of a "working draft" of the human genome--an important milestone in the Human Genome Project--was announced in June 2000 at a press conference at the White House and was published in the February 15, 2001 issue of the journal *Nature*.

Genes come in pairs, with one copy inherited from each parent.

Many genes come in a number of variant forms, known as alleles. A dominant allele prevails over a normal allele. A recessive allele prevails if its counterpart allele on the other chromosome becomes inactivated or lost.


Alleles
Quiet! I'll speak for both of us!
I'll have to be in charge now!
Dominant Allele
Normal Allele
Recessive Allele
Damaged Allele
NATIONAL CANCER INSTITUTE

RNA ----Proteomics/ Transcriptomics ABG 706/708



A hairpin loop from a pre-mRNA. Notice the single strand with its nitrogen-rich (blue) bases extending from its oxygen-rich (red) backbone.

Ribonucleic acid or **RNA** is a polymer or chain of nucleotide units, each comprising a nitrogenous base (adenine, cytosine, guanine, or uracil), a five-carbon sugar (ribose), and a phosphate group. The sugar and phosphate groups form the polymer's backbone, while the nitrogenous bases extending from the backbone provide RNA's distinctive properties.

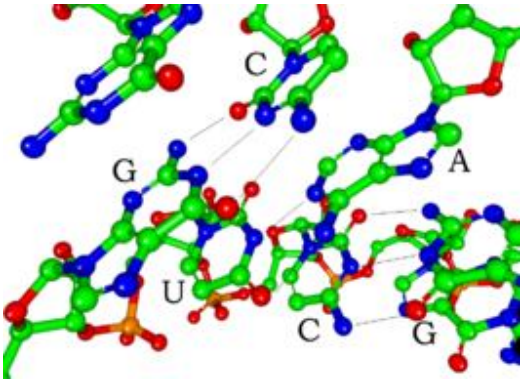
In living cells, RNA in different configurations fulfills several important roles in the process of translating genetic information from deoxyribonucleic acid (DNA) into proteins. One type of RNA (messenger(m) RNA) acts as a messenger between DNA and the protein synthesis complexes known as ribosomes; a second type (ribosomal(r) RNA) forms vital portions of the structure of ribosomes; a third type (transfer(t) RNA) is an essential guide to deliver the appropriate protein building blocks, amino acids, to the ribosome; and other types of RNA, microRNAs (miRNAs) play a role in regulating gene expression, while small nuclear(sn) RNA helps with assuring that mRNA contains no nucleotide units that would lead to formation of a faulty protein. RNA also serves as a genetic blueprint for certain viruses, and some RNA molecules (called ribozymes) are also involved in the catalysis of biochemical reactions.

RNA is very similar to DNA, but differs in a few important structural details. RNA is usually single stranded, while DNA naturally seeks its stable form as a double stranded molecule. RNA nucleotides contain ribose while DNA nucleotides contain the closely related sugar deoxyribose. Furthermore, RNA uses the nucleotide uracil in its composition, instead of the thymine that is present in DNA. RNA is transcribed from DNA by enzymes called RNA polymerases and is generally further processed by other enzymes, some of them guided by non-coding RNAs.

Single-stranded RNA is similar to the protein polymer in its natural propensity to fold back and double up with itself in complex ways assuming a variety of biologically useful configurations.

The connectedness of living organisms can be seen in the ubiquitousness of RNA in living cells and in viruses throughout nature, and in the universal role of RNA in protein synthesis.

Chemical and stereochemical structure



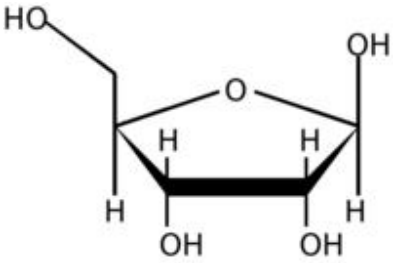
Base-pairing in a siRNA (small interfering RNA) segment, a double-stranded type of RNA. Hydrogen atoms are not shown.

RNA is a **nucleic acid**, a complex, high-molecular-weight macromolecule composed of nucleotide chains whose sequence of bases conveys genetic information.

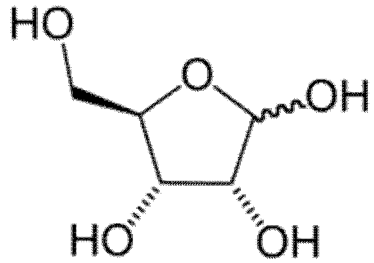
A **nucleotide** is a chemical compound comprising three components: a nitrogen-containing base, a pentose (five-carbon) sugar, and one or more phosphate groups. The nitrogen-containing base of a nucleotide (also called the **nucleobase**) is typically a derivative of either purine or pyrimidine. The most common nucleotide bases are the purines adenine and guanine and the pyrimidines cytosine and thymine (or uracil in RNA).

Nucleic acids are polymers of repeating units (called monomers). Specifically, they often comprise long chains of nucleotide monomers connected by covalent chemical bonds.

RNA molecules may comprise as few as 75 nucleotides or more than 5,000 nucleotides, while a DNA molecule may comprise more than 1,000,000 nucleotide units.



Ribose in acyclic form



A conventional skeletal formula

In RNA, the sugar component, **ribose** is a water-soluable, pentose sugar (monosaccharide with five carbon atoms). Ribose has the chemical formula $C_5H_{10}O_5$.

Ribose is an aldopentose, which means a pentose sugar with an aldehyde functional group in position one. An aldehyde group comprises a carbon atom bonded to a hydrogen atom and double-bonded to an oxygen atom (chemical formula O=CH-). Ribose forms a five-member ring with four carbon atoms and one oxygen. Hydroxyl (-OH) groups are attached to three of the carbons. The fourth carbon in the ring (one of the carbon atoms adjacent to the oxygen) has attached to it the fifth carbon atom and a hydroxyl group.

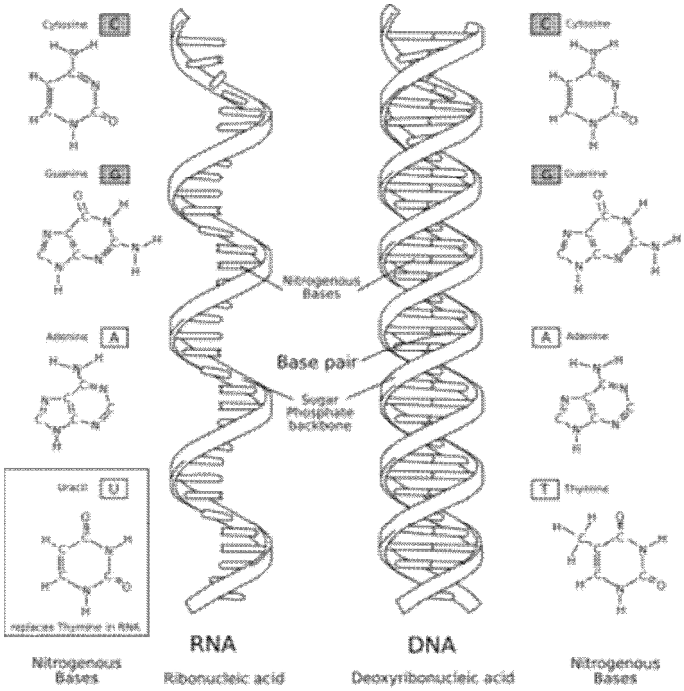
The RNA polymer features a ribose and phosphate backbone with one of four different nucleotide bases—adenine, guanine, cytosine, and uracil—attached to each ribose-phosphate unit.

There are also numerous modified bases and sugars found in RNA that serve many different roles. Pseudouridine (Ψ), in which the linkage between uracil and ribose is changed from a C–N bond to a C–C bond, and ribothymidine (T), are found in various places (most notably in the TΨC loop of tRNA). Another notable modified base is hypoxanthine (a deaminated guanine base whose nucleoside is called inosine). Inosine plays a key role in the Wobble Hypothesis of the genetic code. There are nearly 100 other naturally occurring modified nucleosides, of which pseudouridine and nucleosides with 2'-O-methylribose are by far the most common. The specific roles of many of these modifications in RNA are not fully understood. However, it is notable that in ribosomal

RNA, many of the post-translational modifications occur in highly functional regions, such as the peptidyl transferase center and the subunit interface, implying that they are important for normal function.

The most important structural feature of RNA that distinguishes it from DNA is the presence of a hydroxyl group at the 2'-position of the ribose sugar. The presence of this functional group enforces the C3'-endo sugar conformation (as opposed to the C2'-endo conformation of the deoxyribose sugar in DNA) that causes the helix to adopt the A-form geometry rather than the B-form most commonly observed in DNA. This results in a very deep and narrow major groove and a shallow and wide minor groove. A second consequence of the presence of the 2'-hydroxyl group is that in conformationally flexible regions of an RNA molecule (that is, not involved in formation of a double helix), it can chemically attack the adjacent phosphodiester bond to cleave the backbone.

Comparison with DNA



Left: An RNA strand, with its nitrogenous bases. **Right:** Double-stranded DNA.

The most common nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). The main role of DNA is the long-term storage of genetic information. DNA is often compared to a blueprint, since it contains instructions for constructing other components of the cell, such as proteins and RNA molecules. The DNA segments that carry genetic information are called *genes*, but other DNA sequences have structural purposes or are involved in regulating the expression of genetic information. RNA, also, may serve more than one purpose, but it is most commonly identified as the intermediate

between the DNA blueprint and the actual workings of the cell, serving as the template for the synthesis of proteins from the genetic information stored in DNA.

RNA and DNA differ in three main ways.

First, unlike DNA which is double-stranded, RNA is intrinsically a single-stranded molecule in most of its biological roles and has a much shorter chain of nucleotides. (While RNA is usually single-stranded, the RNA molecule also quite commonly forms double-helical regions where a given strand has folded back on itself. Double-stranded RNA is found also in certain viruses.)

Secondly, while DNA contains *deoxyribose*, RNA contains *ribose*. There is no hydroxyl group attached to the pentose ring in the 2' position in DNA, whereas RNA has two hydroxyl groups. These hydroxyl groups make RNA less stable than DNA because it is more prone to hydrolysis. ("Deoxy" simply indicates that the sugar lacks an oxygen atom present in ribose, the parent compound.)

Thirdly, the complementary nucleotide to adenine is not thymine, as it is in DNA, but rather uracil, which is an unmethylated form of thymine.

Most biologically active RNAs, including tRNA, rRNA, snRNAs, and other non-coding RNAs (such as the signal recognition particle(SRP) RNAs), contain extensively base paired regions that have folded together to form double stranded helices. Structural analysis of these RNAs reveals that they are highly structured with tremendous variety with collections of short helices packed together into structures much more akin to proteins than to DNA, which is usually limited to long double-stranded helices. Through such a variety of structures, RNAs can achieve chemical catalysis, like enzymes. For instance, determination of the structure of the ribosome—an enzyme that catalyzes peptide bond formation—revealed that its active site is composed entirely of RNA.

Synthesis

Synthesis of RNA is usually catalyzed by an enzyme, RNA polymerase, using DNA as a template. Initiation of synthesis begins with the binding of the enzyme to a promoter sequence in the DNA (usually found "upstream" of a gene). The DNA double helix is unwound by the helicase activity of the enzyme. The enzyme then progresses along the template strand in the 3' -> 5' direction, synthesizing a complementary RNA molecule with elongation occurring in the 5' -> 3' direction. The DNA sequence also dictates where termination of RNA synthesis will occur (Nudler and Gottesman 2002).

There are also a number of RNA-dependent RNA polymerases as well that use RNA as their template for synthesis of a new strand of RNA. For instance, a number of RNA viruses (such as poliovirus) use this type of enzyme to replicate their genetic material (Hansen et al. 1997). Also, it is known that RNA-dependent RNA polymerases are required for the RNA interference pathway in many organisms (Ahlquist 2002).

Biological roles

RNA's great variety of possible structures and chemical properties permits it to perform a much greater diversity of roles in the cell than DNA. Three principal types of RNA are involved in protein synthesis:

- Messenger RNA (mRNA) serves as the template for the synthesis of a protein. It carries information from DNA to the ribosome.
- Transfer RNA (tRNA) is a small chain of nucleotides that transfers a specific amino acid to a growing polypeptide chain at the ribosomal site of synthesis. It pairs the amino acid to the appropriate three-nucleotide codon on the mRNA molecule.
- Ribosomal RNA (rRNA) molecules are extremely abundant and make up at least 80 percent of the RNA molecules found in a typical eukaryotic cell. In the cytoplasm, usually three or four rRNA molecules combine with many proteins to perform a structural and essential catalytic role, as components of the ribosome.

RNA also may serve as a catalyst for reactions and as a genetic blueprint, rather than DNA, in various viruses. Some RNA, including tRNA and rRNA, is non-coding in that it is not translated into proteins.

Messenger RNA (mRNA)

Messenger RNA

Messenger RNA is RNA that carries information from DNA to the ribosome sites of protein synthesis in the cell. In eukaryotic cells, once mRNA has been transcribed from DNA, it is "processed" before being exported from the nucleus into the cytoplasm, where it is bound to ribosomes and translated into its corresponding protein form with the help of tRNA. In prokaryotic cells, which do not have nucleus and cytoplasm compartments, mRNA can bind to ribosomes while it is being transcribed from DNA. After a certain amount of time the message degrades into its component nucleotides, usually with the assistance of ribonucleases.

Non-coding RNA

RNA genes (also known as non-coding RNA or small RNA) are genes that encode RNA that is not translated into a protein. The most prominent examples of RNA genes are those coding for transfer RNA (tRNA) and ribosomal RNA (rRNA), both of which are involved in the process of translation. Two other groups of non-coding RNA are microRNAs (miRNA) which regulate the expression of genes through a process called RNA interference (RNAi), and small nuclear RNAs (snRNA), a diverse class that includes for example the RNAs that form spliceosomes that excise introns from pre-mRNA (Berg et al. 2002).

Transfer RNA (tRNA)

Transfer RNA

Transfer RNA is a small RNA chain of about 74-95 nucleotides that transfers a specific amino acid to a growing polypeptide chain at the ribosomal site of protein synthesis, during translation. It has sites for amino-acid attachment and an anticodon region for codon recognition that binds to a specific sequence on the messenger RNA chain through hydrogen bonding. It is a type of non-coding RNA.

Ribosomal RNA (rRNA)

Ribosomal RNA

Ribosomal RNA is the catalytic component of the ribosomes, the protein synthesis factories in the cell. Eukaryotic ribosomes contain four different rRNA molecules: 18S, 5.8S, 28S, and 5S rRNA. Three of the rRNA molecules are synthesized in the nucleolus, and one is synthesized elsewhere. rRNA molecules are extremely abundant and make up at least 80 percent of the RNA molecules found in a typical eukaryotic cell.

Catalytic RNA

Ribozyme

Although RNA contains only four bases, in comparison to the twenty-odd amino acids commonly found in proteins, certain RNAs (called ribozymes) are still able to catalyze chemical reactions. These include cutting and ligating other RNA molecules, and also the catalysis of peptide bond formation in the ribosome.

Genetic blueprint in some viruses

Some viruses contain either single-stranded or double-stranded RNA as their source of genetic information. Retroviruses, for example, store their genetic information as RNA, though they replicate in their hosts via a DNA intermediate. Once in the host's cell, the RNA strands undergo reverse transcription to DNA in the cytosol and are integrated into the host's genome. Human immunodeficiency virus (or HIV) is a retrovirus thought to cause acquired immune deficiency syndrome (AIDS), a condition in which the human immune system begins to fail, leading to life-threatening opportunistic infections.

Double-stranded RNA (dsRNA) is RNA with two complementary strands, similar to the DNA found in all cells. dsRNA forms the genetic material of some viruses called dsRNA viruses. In eukaryotes, long RNA such as viral RNA can trigger RNA interference, where short dsRNA molecules called siRNAs (small interfering RNAs) can cause enzymes to break down specific mRNAs or silence the expression of genes. siRNA can also increase the transcription of a gene, a process called RNA activation (Doran 2007). siRNA is often confused with miRNA; siRNAs are double-stranded, whereas miRNAs are single-stranded.

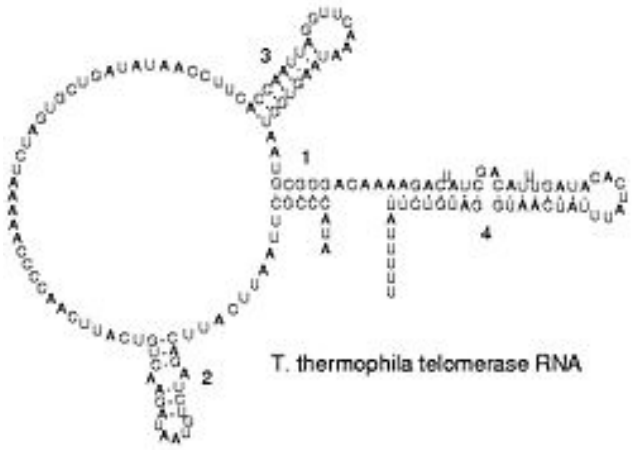
RNA world hypothesis

The RNA world hypothesis proposes that the earliest forms of life relied on RNA both to carry genetic information (like DNA does now) and to catalyze biochemical reactions like an enzyme. According to this hypothesis, descendants of these early lifeforms gradually integrated DNA and proteins into their metabolism.

In the 1980s, scientists discovered that certain RNA molecules (called *ribozymes*) may function as enzymes, whereas previously only proteins were believed to have catalytic ability. Many natural ribozymes catalyze either their own cleavage or the cleavage of other RNAs, but they have also been found to catalyze the aminotransferase activity of the ribosome.

The discovery of ribozymes provides a possible explanation for how early RNA molecules might have first catalyzed their own replication and developed a range of enzymatic activities. Known as the RNA world hypothesis, this explanation posits that RNA evolved before either DNA or proteins from free-floating nucleotides in the early "primordial soup." In their function as enzymes, RNA molecules might have begun to catalyze the synthesis of proteins, which are more versatile than RNA, from amino acid molecules. Next, DNA might have been formed by reverse transcription of RNA, with DNA eventually replacing RNA as the storage form of genetic material. Although there are remaining difficulties with the RNA world hypothesis, it remains as a possible key to understanding the origin and development of the multi-functional nature of nucleic acids, the interconnectedness of life, and its common origins.

RNA secondary structures



Secondary structure of an RNA from a telomerase.

The functional form of single stranded RNA molecules, just like proteins, frequently requires a specific tertiary structure. The scaffold for this structure is provided by

secondary structural elements, which arise through the formation of hydrogen bonds within the interfolded molecule. This leads to several recognizable "domains" of secondary structure like hairpin loops, bulges, and internal loops. The secondary structure of RNA molecules can be predicted computationally by calculating the minimum free energies (MFE) structure for all different combinations of hydrogen bondings and domains (Mathews et al. 2004). There has been a significant amount of research directed at the RNA structure prediction problem.

History

Nucleic acids were discovered in 1868 by Johann Friedrich Miescher (1844-1895), who called the material 'nuclein' since it was found in the nucleus. It was later discovered that prokaryotic cells, which do not have a nucleus, also contain nucleic acids.

The role of RNA in protein synthesis had been suspected since 1939, based on experiments carried out by Torbjörn Caspersson, Jean Brachet, and Jack Schultz. Hubert Chantrenne elucidated the messenger role played by RNA in the synthesis of proteins in ribosomes. Finally, Severo Ochoa discovered RNA, winning Ochoa the 1959 Nobel Prize for Medicine. The sequence of the 77 nucleotides of a yeast RNA was found by Robert W. Holley in 1964, winning Holley the 1968 Nobel Prize for Medicine. In 1976, Walter Fiers and his team at the University of Ghent determined the complete nucleotide sequence of bacteriophage MS2-RNA (Fiers et al. 1976).

List of RNA types

Type	Function	Distribution
mRNA	Codes for protein	All cells
rRNA	Translation	All cells
tRNA	Translation	All cells
snRNA	RNA modification	All cells
snoRNA	RNA modification	All cells
miRNA	Gene regulation	Eukaryotes
piRNA	Gene regulation	Animal germline cells
siRNA	Gene regulation	Eukaryotes
Antisense mRNA	Preventing translation	Bacteria
tmRNA	Terminating translation	Bacteria
SRP RNA	mRNA tagging for export	All cells
Ribozyme	Catalysis	All cells
Transposon	Self-propagating	All cells
Viroid	Self-propagating	Infected plants

In addition, the genome of many types of viruses consists of RNA, namely:

- Double-stranded RNA viruses
- Positive-sense RNA viruses
- Negative-sense RNA viruses
- Retroviruses
- Satellite viruses

How Genes direct the production of proteins

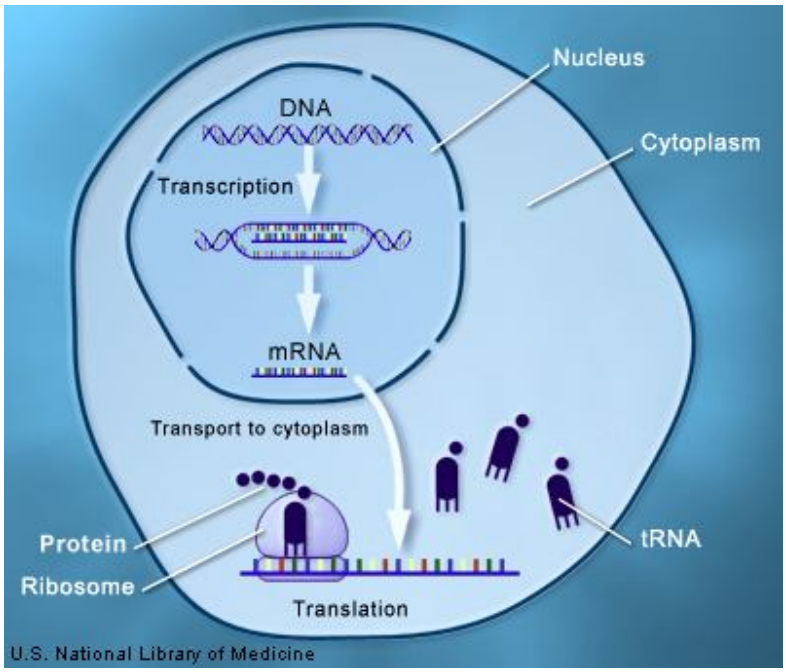
Most genes contain the information needed to make functional molecules called proteins. (A few genes produce other molecules that help the cell assemble proteins.) **The journey from gene to protein is complex and tightly controlled within each cell. It consists of two major steps: transcription and translation. Together, transcription and translation are known as gene expression.**

During the process of transcription, the information stored in a gene's DNA is transferred to a similar molecule called RNA (ribonucleic acid) in the cell nucleus. Both RNA and DNA are made up of a chain of nucleotide bases, but they have slightly different chemical properties. **The type of RNA that contains the information for making a protein is called messenger RNA (mRNA) because it carries the information, or message, from the DNA out of the nucleus into the cytoplasm.**

Translation, the second step in getting from a gene to a protein, takes place in the cytoplasm. **The mRNA interacts with a specialized complex called a ribosome, which "reads" the sequence of mRNA bases. Each sequence of three bases, called a codon, usually codes for one particular amino acid. (Amino acids are the building blocks of proteins.) A type of RNA called transfer RNA (tRNA) assembles the protein, one amino acid at a time. Protein assembly continues until the ribosome**

encounters a “stop” codon (a sequence of three bases that does not code for an amino acid).

The flow of information from DNA to RNA to proteins is one of the fundamental principles of molecular biology. It is so important that it is sometimes called the “central dogma.”



Through the processes of transcription and translation, information from genes is used to make proteins.

INTRODUCTORY PROTEOMICS: MUSCLE ISOFORM DETERMINATION

Introduction to RNA extraction by Dr Wick assisted by Monica and Funmi Adebambo

RNA Extraction

- Switch on the Lamina Flow

- Wipe the work top with 70% Ethanol followed by distilled H2O and RNase Away
- Arrange on work top all equip and reagents for the day's work i.e.
    1. 200, 1000ul pipette tips
    2. 2-20; 20-200; 100 – 1000 micropipettes
    3. RNase Away; 75% Ethanol, 100% ethanol; Chloroform; Isopropanol; Trizol® Reagent; HPLC grade water; Sterile Scalpel Blades; Omni International Tissue Master; sterile cryo tubes; corning tubes; autoclaved microtubes, get the 50ul of Formamide in DEPC water 50:50
    4. Gloves; KIM wipes
    5. Thermoline flask with Liquid Nitrogen
    6. refrigerated micro centrifuge in cold room (-4^{o}C)
    7. IJJB089y work book- (yellow manual, page 89 labelled IJJB)
- Wear your gloves (single or double)
- Label 2 cryo tubes; into one add 1ml of HPLC grade water; into the second add 1ml of Trizol
- Label 2 microtubes leave empty
- Label 2 15ml corning tubes, into one add 5ml HPLC grade water; into the second add 3mls Trizol
- Get 2 petri dishes, into one place Kim wipe pour on it 2ml RNase Away; Open a sterile scalpel blade carefully drop it onto the wet Kim wipe make sure the holding edge is on the edge of the Petri dish (may open the second on it as well and cover immediately with the edge of the wet Kim wipe and the Petri dish lid. Into the second Petri dish pour 1000ul of Trizol and cover.
- Connect the Tissue Master to the Hood mains, sterilize it by flicking the tip inside the 5ml HPLC water several times and then inside the 3ml Trizol solution inside the large Corning tubes.
- Pour 100ul of Trizol into a cryotube
- Pour enough quantity of liquid N2 from Thermo flask into Polystyrene box ready for use.
- Take one corning tube of Broiler sample from Liquid N2 Tank labeled 049W. This is a 20d.p.h. broiler muscle. (label = Broiler 20d.p.h)
- Drop bottle in liquid Nitrogen and as your sampling progresses always ensure that the corning tube is dropped into Liquid Nitrogen not necessarily covered when sample is being taken from the tube
- Take about 50ng of muscle from the tube with a scalpel into the Trizol in the Petri dish. Finely chop it and mix with the Trizol.

- Suck all the mixture into the 100ul Trizol in cryo tube.
- Homogenize sample with the tissue master very well.
- Keep at room temp for 5-10 mins

Should extra muscle drop into the Trizol don't return it into the sample tube, rather put it into a cryo tube and drop into liquid Nitrogen in the polystyrene because it contains Trizol

- Return the remaining muscle sample into the LN2 tank
- **Wipe your hand with the RNase Away. Do this as often as you touch anything before touching your sample because RNA is all around you, from your forehead and bacteria around. Take Absolute precaution**
- **Label a corning tube from the -20^{o}C fridge IJJB089y, pour into it 50ul of 50:50 DEPC Water with formamide as the tube into which you will reconstitute your RNA after extraction.**
- **If the mixture is more than 1.5 mls, reduce volume by transferring only 1ml into Cryo tube allow to rest for 5 mins.**
- **Spin down at 15,000rpm for 10mins in a refridgerated Microcentrifuge inside the cold room at -4^{o}C**
- **While waiting, clean and rinse your homogenizer in the HPLC Water, wipe, cover and disconnect and put away.**
- **Check your sample, remove the supernatant (pink layer only) into another cryo tube. Reduce the volume if it is too full to be able to accommodate 200ul of chloroform without spilling over.**
- **Add 200ul of chloroform, mix well by hand**
- **Leave at room temp for 5-10 mins, the supernatant begins to separate**
- **Spin down at 15,000 rpm for 15mins at -4^{o}C in the cold room**
- **Remove clear supernatant into a new tube, if you mistakenly suck up the trizol with the liquid, re spin down again for 2 mins**
- **Get as much of the supernatant if it means using 10ul pipette.**
- **Add 0.5ml (500ul) isopropanol**
- **Leave at room temp for 5-10mins**
- **Spin down at 15,000 rpm for 18 mins**
- **Discard isopropanol layer supernatant**
- **Wash pellet with 1ml 75% ethanol**
- **Briefly vortex**

- **Spin down at 15,000 for 5 mins at -4^{o}C**
- **Discard supernatant by decanting (decanting means pouring gently across a Kimwipe) followed by one flick**
- **Wash pellet again with 1 ml of 75% ethanol**
- **Briefly vortex**
- **Spin down at 15,000 rpm for 5 mins at -4^{o}C**
- **Decant supernatant and place tube upside down on Kim Wipe rest tube by the side of the pack**
- **Leave for 10-15 mins, never exceed 15 mins because if it overdries it will never come off the tube.**
- **Re suspend in 50ul of 50:50 DEPC water and Formamide already in the labeled corning tube.**
- **Should you make a mistake, re constitute the solution with 1ml 100% ethanol**
- **Spin down as with DNA, decant supernatant and re suspend airdried RNA in 50ul 50:50 DEPC/Formamide.**
- **Label properly and store away in sample box inside -20^{o}C fridge**

Running the Agarose Gel

Run Agarose gel of sample on 1JJB089y (049W broiler 20d.p.h)

- **Bring out the Agarose gel elecrophoresis system**
- **Pour into the tank enough 1 X TAE buffer to cover the running plate**
- **Bring out the already prepared gel from the cold room (note prepared 7% agarose gel could be kept at -4^{o}C in plastic container ready for use)**
- **Cut the number of lanes required in the middle but with half cut comb edges note that an extra lane must be available for the RNA ladder.**
- **Place on the gel plate covered by the buffer.**
- **Rinse the wells with the buffer**
- **Drop 3ul of loading dye for the number of samples and the ladder on a cut strip of parafilm**
- **Bring out the RNA ladder lot # 414498 from the -20^{o}C fridge**
- **Pipette out 3ul of ladder mix thoroughly with the drop of loading buffer, suck up, rinse the pipette tip with the buffer in the tank, load into middle or noted well on your diagram in your notepad**

- **Take also 3ul of the sample, mix thoroughly with the 3ul loading buffer, suck up, clean the pipette tip in the buffer and load into the well**
- **Cover and Connect the tank (black to black, red to red) and to the power pack.**
- **Switch on the pack, set to 140volts ;400mAmp.**
- **Check your time record the start: 12.15**
- **Allow gel to run till almost end of the gel following the track of the runs**
- **Stop time =12.44 approx. 30mins run.**
- **Stop and view under UV light on the UV box**
- **Take sample in polytene bag to Animal Science (Dr Hughes lab) for photography Tele 8000 view exposure at turn 2.**
- **Put your RNA sample away in the -80^{o}C freezer**

REPORT: no band i.e. No RNA might be due to pptn/suspension problem

Extract another RNA Sample 2 = 4JMR Chicken SCWL Adult 1yr+ P.major

- **Wipe the LAMINA Flow top with Ethanol, followed by Sterile distilled water and RNA Away.**
- **Switch on the Laminahood**
- **Clean all pipettes and tissue master with RNA Away.**
- **Connect the tissue master**
- **Put the Petri dishes on the work top, into one put a soaked Kim wipe with RNA Away, open two scalpel blades onto the soaked wipe and cover the blades with the wipe.**
- **Into the second put 1ml of Trizol cover**
- **Take out your sample from the -80oC freezer or the LN2 tank (adult chicken)**
- **Cut about 50ng chop and mince well with the Trizol in the Petri dish**
- **Suck out all mixture into a corning or cryo tube and cover**
- **Clean the tip of Tissue Master first in the HPLC water in the corning tube and second in the Trizol in the Corning tube. Make sure you thoroughly clean the tip.**
-