

No products in the basket.

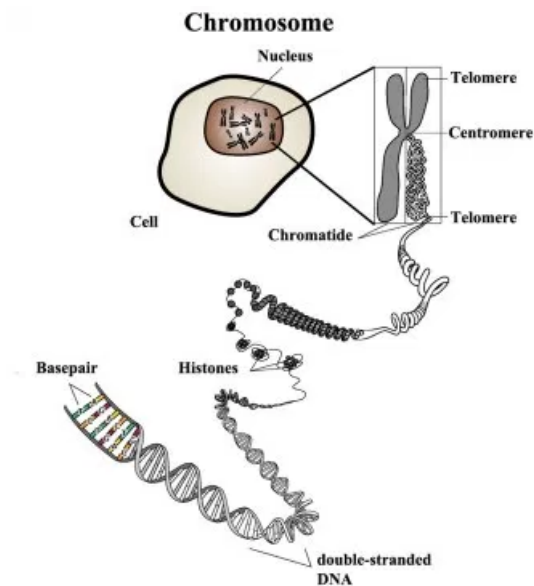
[Home](#)[About](#)[Blog](#)[Background](#)[Donate to Science!](#)

# DNA, RNA and protein – the Central Dogma

You are here: [Home](#) / [Theory](#) / DNA,

## Chromosomes

For many research questions, I use the genetic information that is available on my research species. But to understand why this information is useful I need to explain a bit about genetics first. Every organism consists of cells, all multicellular organisms have a cell and a cell nucleus. This nucleus contains the DNA, the hereditary material. But this DNA does not float around randomly through the nucleus, it is neatly packed into something we call chromosomes (see the figure below). We people have 46 chromosomes, fruit flies have 8 and the flour beetle I work on has 20. Almost all animals are diploid, which means that you have 2 copies of each chromosome. So of the 46 chromosomes you have, 23 originate from your mother and 23 originate from your father. These chromosomes contain all the hereditary information in the form of double stranded DNA.



## DNA

**DNA** is short for **D**eoxyribo**n**ucleic **a**cid. The DNA in every nucleus of an organism is exactly the same in all cells. The only exceptions are the sperm cells and the eggs, they only contain half of the DNA that a normal cell contains (sperm and eggs in humans contain only 23 of the 46 chromosomes).

DNA is made up from 4 different bases (nucleotides), adenine (A), thymine (T), guanine (G) and cytosine (C). This is true for plants, animals, bacteria, in fact it is true all life forms on earth that contain DNA. The bases on one strand of DNA form base pairs with a second strand of DNA to form the double helix. But the base pairs that can be formed are limited; adenine (A) can only form a base pair with thymine (T) and guanine (G) can only form a base pair with cytosine (C). So when we know the sequence of bases on 1 strand of DNA, we also know the sequence of bases on the other strand of DNA. The order of bases is referred to as **the sequence**. An example of a short sequence of a single strand of DNA is: ATTGCTCAT

### High-coverage custom NGS panel - 2.5-hr amplicon library prep

Remove multiplex PCR by-products. Improve NGS data specificity & uniformity. [paragongenomics.com/customdesign](http://paragongenomics.com/customdesign)



### Human Cells

Quality Primary Cells Isolated from Many Tissues, Quality Tested  
[cellapplications.com/primary-cells](http://cellapplications.com/primary-cells)



Because we know the sequence of this strand we also know which bases are on the other strand:

Strand 1:ATTGCTCAT

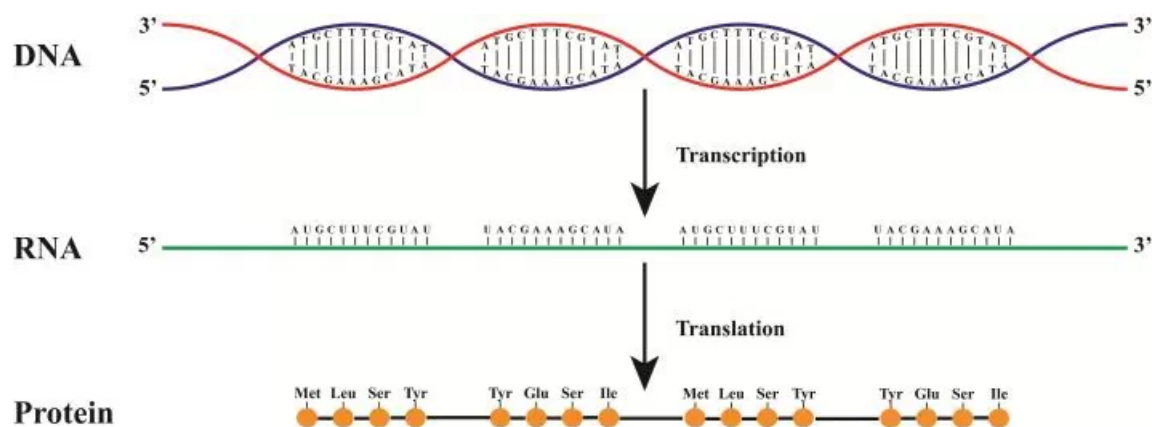
Strand 2:TAACGAGTA

I will often talk about sequences; this is because the sequence of DNA codes for the type of protein that is being made and these proteins are the important in all aspects of life. The way DNA encodes a protein is something I will get back to in a bit. First it is important to know that the information of the sequence of DNA gives us the opportunity to “read” the DNA. A lot of extra information is needed to properly read DNA, but I will not go into detail here. Modern technology has provided us with the complete sequence of a couple of different organisms already. So we know the sequence of all the DNA in all the chromosomes of this organism! This complete sequence is called a genome and this genome is freely accessible through this website: <http://www.ncbi.nlm.nih.gov/>

I want to note here that the DNA is not identical for all individuals of a species, so the human genome that is available online is not identical to your genome. But, we can learn a lot from the genome that is available online. This is because the most important parts of the genome vary considerably less than the less important parts. Take for example eye color, it is not important for survival whether you have blue or brown eyes, so this is a less important character. The red blood cells that are able to transport oxygen on the other hand are very important; people with red blood cells that are unable to transport oxygen will not survive. So variation in a character as important as the ability to transport oxygen is sort of “not tolerated”. I will later write more about selection.

## From DNA to protein

But how does DNA code for protein? (also referred to as the Central Dogma) To make protein from DNA we first need to take a different step. That is to make RNA from DNA. RNA is important for a lot of different functions but I will only talk about messenger RNA here, which is used to synthesize protein from. RNA (Ribonucleic Acid) is synthesized in the nucleus and is very similar to DNA. The synthesis of RNA also involves the use of bases, but in RNA synthesis no thymine (T) is used but uracil (U) is used instead. The sequence of RNA corresponds to the sequence of DNA from which the RNA is synthesized (see the figure below).



The synthesis of RNA from DNA is called transcription (the DNA is transcribed into RNA). In this figure the RNA is being synthesized from the red strand of DNA (which serves as template), this strand of DNA starts with the base T. The RNA strand starts with the only base that can form a base pair with this T, the A. This continues until the complete sequence of RNA is synthesized. Because the red strand serves as template, the sequence of RNA will be identical to the blue strand of DNA, only with the base U instead of the base T.

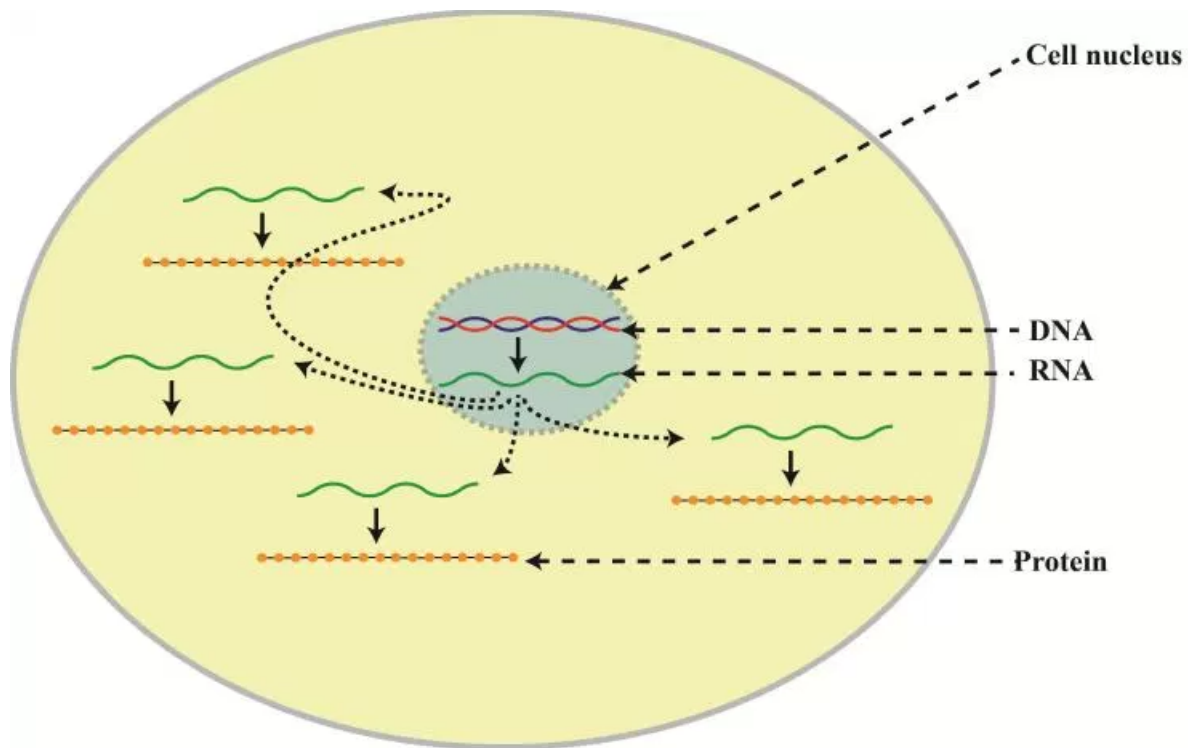
So now we have an RNA strand. From this strand the protein will be synthesized, this is called translation (RNA is translated into protein). A protein is made from amino acids, these form a strand. I show the protein strand as a linear line, but in reality complex interactions between amino acids lead to 3 dimensional forms that are essential for the functioning of the protein. The translation of RNA to protein is different than the synthesis of RNA from DNA (transcription). When the DNA was transcribed into RNA, one base of DNA corresponded to one base of RNA, this 1 to 1 relation is not used in the translation to protein. During this translation, 1 amino acid is added to the protein strand for every 3 bases in the RNA. So a RNA sequence of 48 bases codes for a protein strand of 16 amino acids. A certain combination of 3 bases always gives the same amino acids, so we can put the translation into a table (see below). We take the first 3 bases from the figure above as example, which are AUG. The first base is A, we look it up on the left side of the table, which shows us that we have to look in the 3<sup>rd</sup> row of the table. The second base is U, we look it up on the top of the table which shows us that we have to look in the 1<sup>st</sup> column and 3<sup>rd</sup> row. There we see our third base and our combination. We can see that the combination of AUG codes for the amino acid Methionine (Met). In this way we can translate the complete RNA sequence into the protein sequence.

		Second Position				
		U	C	A	G	
U	UUU	Phe	UCU	UAU	UGU	U
	UUC					Ser
	UUA	Leu	UCA	UAA	UGA	
	UUG					UCG
C	CUU	Leu	CCU	CAU	CGU	
	CUC					Pro
	CUA	CCA	CAA	CGA	A	
	CUG				CCG	CAG
A	AUU	Ile	ACU	AAU		
	AUC				Thr	AAC
	AUA	ACA	AAA	AGA		
	AUG				ACG	AAG
G	GUU	Val	GCU	GAU		
	GUC				Ala	GCC
	GUA	GCA	GAA	GGA		
	GUG				GCG	GAG

### In the cell

But how does this work in an actual cell? And why make RNA first and then protein? Why not make protein from the DNA directly? Well the DNA is located in the nucleus of the cell, here RNA is transcribed but protein is not translated. After transcription the RNA is relocated to the cytoplasm of the cell, here it is translated into protein. So the separation of nucleus and cytoplasm prevents protein from being made directly from DNA. But there are other reasons why RNA is made. I will name a few, but not all (there are so many).

First, the DNA is well protected in the nucleus against everything that floats around in the cytoplasm, which prevents the DNA from getting damaged. The transcription of DNA to RNA prevents that the DNA has to be translated itself in the cytoplasm and thereby prevents DNA damage. Another reason is that we only have 1 copy of DNA in each cell, but sometimes we need a lot of the same protein. Therefore it would be convenient if we could make more than one copy of the same protein at the same time. When the DNA is transcribed into RNA 10x, there are 10 RNA templates to make protein from. So protein can be made 10x as fast. So making RNA prevents DNA damage and provides flexibility in the amount and speed of protein synthesis (see the figure below).



These proteins are essential in all living organisms, proteins are involved in DNA synthesis, RNA synthesis, the immune response, cell structure and for a lot more! So proteins are important for almost everything in living organisms. There are several steps to get from DNA to protein and I have not talked about a lot of processes involved, but these are only modulations on the general process I just explained. I hope that I could make it clear how DNA is translated to protein and why DNA sequences can be powerful tools in research. Later, I will write more specific about my research and then it will become clear how this story is related to my research.

Here is a video by Nature of the same process:

