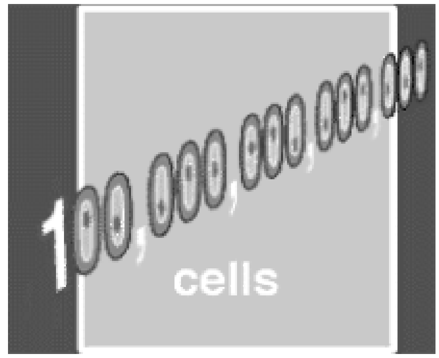
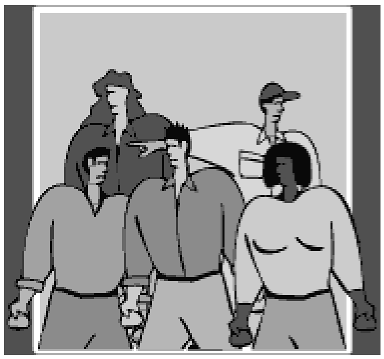
What Is A Genome?

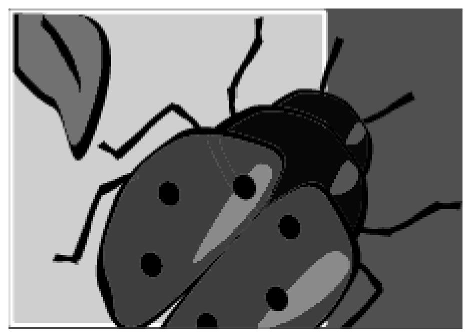
Your body is made up of about one hundred, million, million cells (100,000,000,000,000). Each of these cells has a complete set of instructions about how to make you. This set of instructions is your genome.





Your genome is quite similar to everyone else's genome, which is why we all turn out to be human beings. But every other living thing also has a genome.

For instance, beetles each have a beetle genome, which is a complete set of instructions for making a beetle. Cabbages have a cabbage genome and bacteria have a bacteria genome.

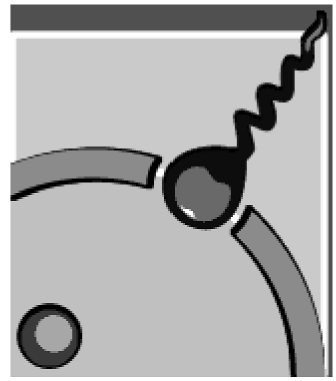


Dinosaurs would have had a dinosaur genome, which is why Michael Crichton (author of Jurassic Park) thought we might be able to grow a dinosaur, if we had a copy of its genome.

***<u>GENOME</u>*: One copy of all the DNA in a cell of an organism. Our genome is 3,000,000,000 base-pairs, packaged into 23 pairs of chromosomes: bacteria may have only 1,500,000 base-pairs in one chromosome.**

Where Do Genomes Come From?

We actually have two genomes each! We get one copy of our genome from each of our parents. A sperm cell has only one copy of a genome, and an egg cell also only has only one copy. At the moment of fertilisation, the sperm cell and the egg cell join together to make a cell containing two genomes.



The fertilized egg is like all the other cells in our body: it has two copies of the genome. These will act as a complete set of instructions to make a new human being!

What Is A Genome Made Of?

Your genome is made of a chemical called DNA. The letters stand for deoxyribonucleic acid, but that's not really important. What is important is the unique shape of DNA. DNA is shaped like a twisted ladder. Imagine a ladder made of rubber. If you held the bottom of the ladder firmly and twisted the top, the shape you would create is the same shape that your DNA has. Scientists call this shape a 'double helix'.

What Does DNA Look Like?

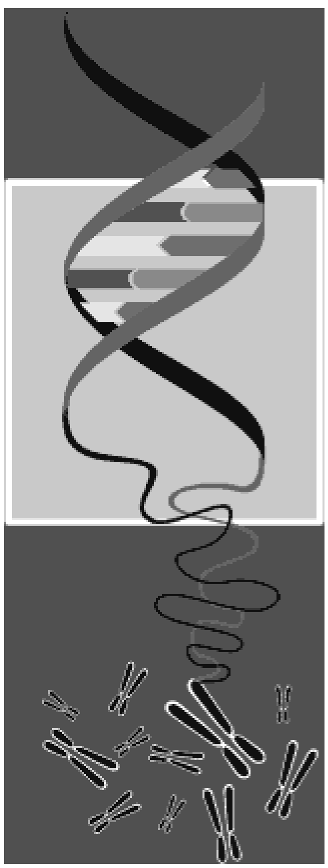
DNA has rungs, like a ladder. These rungs are called base pairs, and it is the base pairs that do the important job of being the instructions. These base pairs can break, allowing the sides of the helix to unravel. This special property is what allows DNA to copy itself, and to act as the instructions.

Of course, DNA is much smaller than a ladder. If the base pairs on your DNA were as far apart as the rungs on a real ladder, then the DNA from just one of your cells would stretch half way to the Moon!

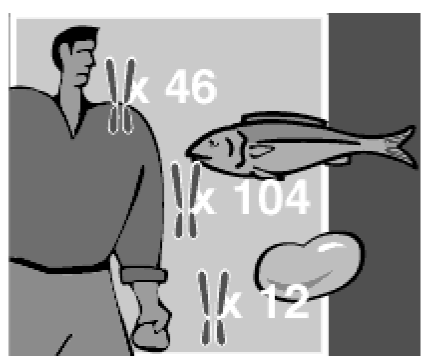
***DNA**: Deoxyribonucleic acid is the long molecule that holds our instructions. Two strands of DNA are twisted together into the double helix. The base-pairs between A and T and between C and G hold the two strands together. DNA is usually packaged into chromosomes.*

***BASE:** The DNA Code is 'written' using four chemical 'letters' (A, C, G, T) called bases. The bases can pair up, A with T and C with G.*

***A DNA CODE** is the set of instructions to build an organism. Much of the code is written in 'words' of three letters (such as ATG, CCG, TAA and so on) in DNA. This code must be 'translated' by the cell into the building blocks of proteins. Other parts of the code are 'switches' to turn genes on or off, up or down.*



But the DNA isn't spread out like a ladder. It's tightly packaged into bundles, surrounded by protein to protect it. If we look at DNA under a microscope, we can see these bundles of protein and DNA, arranged in strips called chromosomes. The number of chromosomes in a cell depends on what species of animal it is from. Humans have 46 chromosomes (23 pairs), but carp fish have 104 chromosomes (52 pairs) while broad beans have 12 chromosomes (6 pairs).



How Are The Instructions Written In The DNA Code?

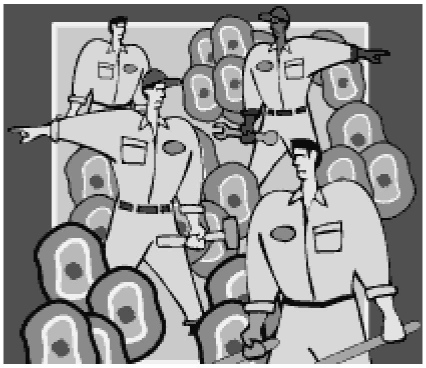
DNA is a code. The English alphabet is also a code. Let's take the word "koala". The word as it is typed on the screen is not actually a koala. It doesn't live in Australia or eat eucalyptus leaves. But the letters 'k', 'o', 'a', 'l' and 'a' in that particular order mean an animal that lives in Australia and eats eucalyptus leaves. If you didn't know any English, you wouldn't be able to guess what the word means from the letters that are in it. The letters 'k', 'o', 'a', and 'l' appear in lots of other words where they don't mean anything to do with koalas. The alphabet is a code.

What Does The DNA Language Look Like?

DNA is also a code. It is written in only four 'letters', called A, C, T and G. So DNA language looks like this...

The meaning of this code lies in the sequence of the letters A, T, C and G in the same way that the meaning of a word lies in the sequence of alphabet letters. Different languages use different alphabets to convey meaning.

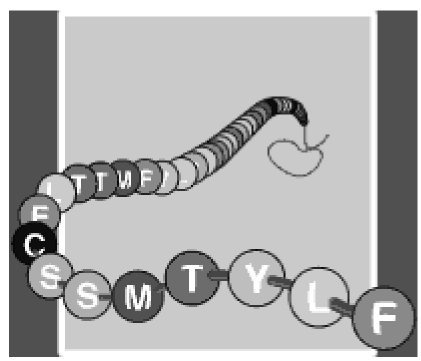
What Does DNA Code For?



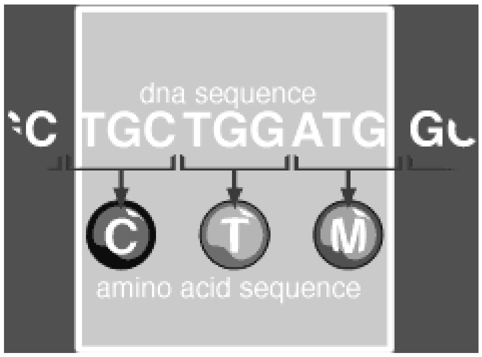
The DNA codes for protein. In our cells, proteins are the labour force. It is proteins that get everything done. Proteins make new cells and destroy old or diseased ones. Proteins break down our food to release energy. Proteins organise the transport of useful chemicals between cells. Often, these useful chemicals are themselves proteins. As well as doing things, proteins are the building blocks for most of your body. In the same way that a wall is made mostly of bricks, your body is made mostly of protein.

The ingredients of a protein are amino acids. To build a protein we need to build a long chain of amino acids. There are 20 different types of amino acids, so there are lots of different protein chains we can build.

Biologists give amino acids a code letter, as for DNA. This is much easier than writing out the whole name each time. For example, M is methionine, L is leucine, F is phenylalanine (because P is proline).



What Are Amino Acids?



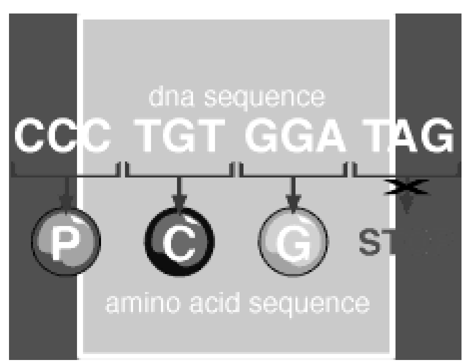
The DNA code uses groups of three 'letters' to make meaning. Most groups of three 'letters' codes for an amino acid (some code for 'punctuation - starts and stops). For instance, the DNA letters TGC code for an amino acid called cysteine, whereas the DNA letters TGG code for an amino acid called tryptophan. Each of these sequence of three DNA letters is called a DNA triplet, or codon. Since there are four different DNA letters (A, G, C and T), there are 4 x 4 x 4 = 64 different combinations that can be used to make a codon.

AMINO ACIDS:* *The building blocks of proteins. There are 20 naturally occurring amino acids. Different proteins have different numbers of amino acids in different orders. The order of amino acids is written in the DNA code.

CODON:* *A sequence of three DNA bases that specifies one building block of a protein. DNA code for proteins is written in 'words' of three letters: 61 of the 64 possible combinations code for an amino acid and three are 'stop' codons.

So There Are 64 Codons?

Some of these 64 codons code for the same amino acid. For instance, GAA and GAG both code for an amino acid called glutamic acid. Some of the 64 codons don't code for any of the amino acids. Instead they provide the grammar of the DNA sequence. For instance, the codon TAA mans 'full stop' or 'stop here'. Full stops are essential for when the cell is making protein from the DNA code , otherwise the cell wouldn't know where to stop



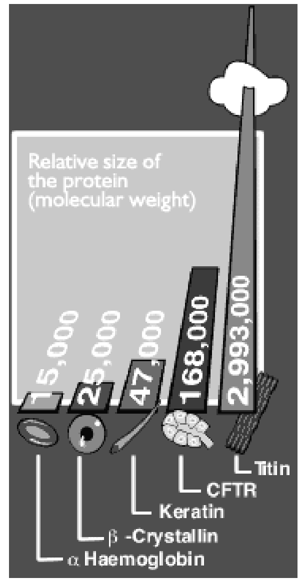
An Example

An example of a DNA code might be

CCCTGTGGAGCCACACCCTAG

This would make the protein:

Proline-Cysteine-Glycine-Alanine-Threonine-Prolin

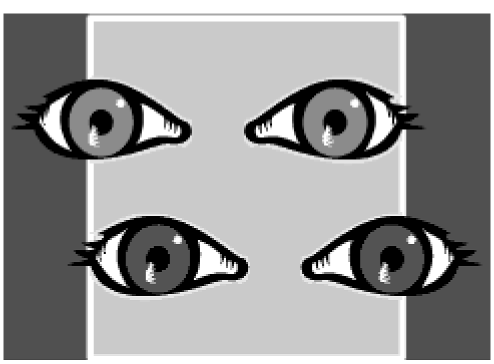


Most proteins are actually much longer than this. Many of the proteins that make your body makes contain hundreds of amino acids. So if we know what the DNA sequence is, we can work out which amino acids the protein must contain and in what order.

What Are Genes?

Each DNA sequence that can be used by your cells to make a protein is called a gene. At the beginning of a gene, there are some codons that mean 'start making your protein here' and at the end of a gene, there are codons that mean 'stop making your protein here', or 'full stop'.

What Other Things Can Genes Do?



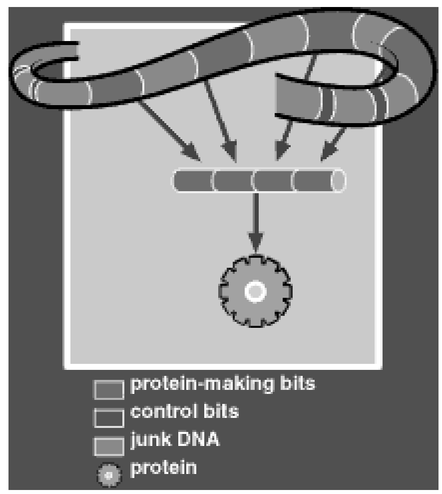
We talk about genes having different characteristics. For instance, we talk about 'a gene for eye colour'. What we mean is that this is the gene that codes for the protein that is the pigment in the iris of each of our eyes. Some genes come in different versions. Some people have a gene that codes for a protein that makes their eyes look blue while other people have a gene that makes a protein that makes their eyes look brown.

Does All DNA Code For Proteins?

GENE: A set of DNA letters that has all the information required to make a protein. It therefore includes instructions about the beginning and end of the protein (punctuation) and signals that tell the cell when to make the protein.

But not all of the DNA sequence in our genome is used to make protein (perhaps less than 10%). There is a lot of DNA that is never used to make protein: we know what some of this DNA does, but not all. The bits of DNA we don't understand are often called 'junk DNA'

Much of this DNA is repeated sequences - as if a printer had made a mistake and scattered lots of copies of one page of a book throughout the story. As much as 40% of our genome is made of repeated sequences.



The other 50% is DNA that lies outside the sequences used to make protein. Quite simply, we don't know what this DNA does. It may be important as a 'spacer' in the genome to make sure the active parts work properly. It may be important in evolution of complex animals such as humans.

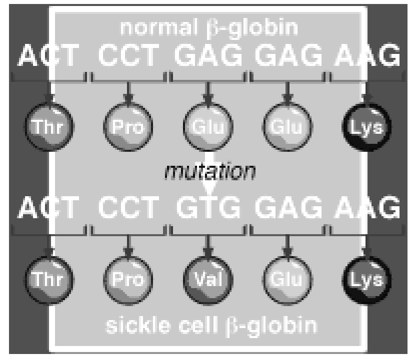
This DNA is made up of the same letters as coding DNA, but it doesn't have the same meaning. If we were to try to build a protein using 'junk' DNA sequence, we might find that the protein wouldn't fit together, or that it wasn't stable once we had built it. Not all sequences of amino acids will make useful proteins. A genome is all of the genes plus all of the other DNA.

Can Genomes Change?

Yes, and these changes are called mutations. Sometimes, one of the DNA letters is accidentally swapped for another letter. For example the codon GCT might be changed into GCA. This mutation might have a very serious effect, or it might have no effect at all. For example, if the letter T is swapped for an A in the codon GCT then the protein will still be the same, since both the old codon (GCT) and the new codon (GCA) code for the amino acid Alanine. Some mutations have a much more serious effect.

***Mutation:* A change in DNA sequence. Mutations are relatively common in our DNA - but most have no detectable effect.**

For example, swapping an A for a T in a gene for haemoglobin causes a serious disease called sickle cell anaemia. Haemoglobin does the job of carrying oxygen around the body in the blood. In people with sickle cell anaemia, the sequence of amino acids in haemoglobin is different, and so it doesn't work as well.



There are other types of mutations as well. Sometimes, a bit of the DNA sequence is missed out by mistake, or a new bit added in. Sometimes, parts of the sequence are swapped over, even between different chromosomes.

Each of our genes is a copy from either our mum or our dad. If there is a mutation in one of these genes, this will be passed on from parent to child along with the rest of the gene. This is why diseases often run in families. Mutations can happen in the junk DNA as well as in the genes. Because the junk DNA isn't used to make protein, mutations that happen in the junk DNA don't usually have any effect.

What is the Human Genome Project?

The Human Genome Project is a very big science experiment. Scientists from all over the world have joined in to help with what is probably the biggest biology experiment ever. What these scientists are trying to do is find out what each of the DNA letters in the human genome is. Once the Human Genome Project has been completed we will be able to write out the whole of the code in the human genome.

We can start with (say) G at the beginning, and write down all of the DNA letters in the order that they appear. There are 3,200,000,000 of these letters, and scientists don't expect to finish this experiment until 2003. Since the Human Genome Project started in the mid-1980s, this experiment will have taken over 15 years!

The results from the Human Genome Project will help to keep humans all over the world healthier. Scientists may be able to use the results to make new medicines to treat sick people with, or to identify people who have a higher risk of certain diseases so they can take preventative action.