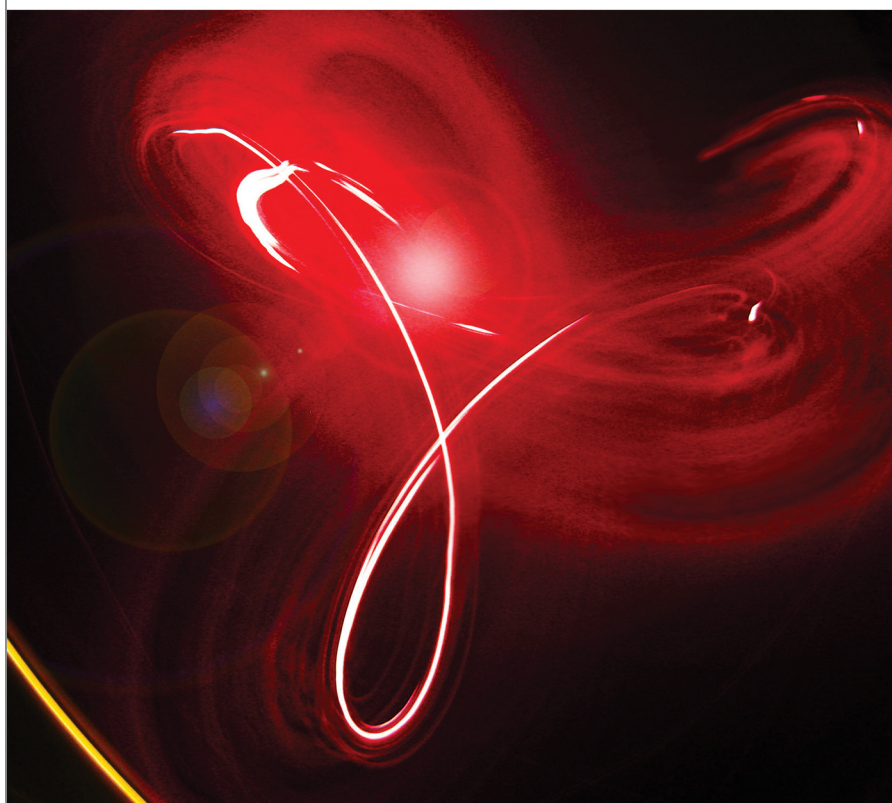


What is the genome made of?



What is the genome made of?



OpenLearn

Free learning from
The Open University

This item contains selected online content. It is for use alongside, not as a replacement for the module website, which is the primary study format and contains activities and resources that cannot be replicated in the printed versions.

About this free course

This free course provides a sample of Level 1 study in Science:

<http://www.open.ac.uk/courses/find/science>.

This version of the content may include video, images and interactive content that may not be optimised for your device.

You can experience this free course as it was originally designed on OpenLearn, the home of free learning from The Open University -

www.open.edu/openlearn/science-maths-technology/science/biology/what-the-genome-made/content-section-0.

There you'll also be able to track your progress via your activity record, which you can use to demonstrate your learning.

The Open University, Walton Hall, Milton Keynes, MK7 6AA

Copyright © 2016 The Open University

Intellectual property

Unless otherwise stated, this resource is released under the terms of the Creative Commons Licence v4.0 http://creativecommons.org/licenses/by-nc-sa/4.0/deed.en_GB. Within that The Open University interprets this licence in the following way:

www.open.edu/openlearn/about-openlearn/frequently-asked-questions-on-openlearn. Copyright and rights falling outside the terms of the Creative Commons Licence are retained or controlled by The Open University. Please read the full text before using any of the content.

We believe the primary barrier to accessing high-quality educational experiences is cost, which is why we aim to publish as much free content as possible under an open licence. If it proves difficult to release content under our preferred Creative Commons licence (e.g. because we can't afford or gain the clearances or find suitable alternatives), we will still release the materials for free under a personal end-user licence.

This is because the learning experience will always be the same high quality offering and that should always be seen as positive – even if at times the licensing is different to Creative Commons.

When using the content you must attribute us (The Open University) (the OU) and any identified author in accordance with the terms of the Creative Commons Licence.

The Acknowledgements section is used to list, amongst other things, third party (Proprietary), licensed content which is not subject to Creative Commons licensing. Proprietary content must be used (retained) intact and in context to the content at all times.

The Acknowledgements section is also used to bring to your attention any other Special Restrictions which may apply to the content. For example there may be times when the Creative Commons Non-Commercial Sharealike licence does not apply to any of the content even if owned by us (The Open University). In these instances, unless stated otherwise, the content may be used for personal and non-commercial use.

We have also identified as Proprietary other material included in the content which is not subject to Creative Commons Licence. These are OU logos, trading names and may extend to certain photographic and video images and sound recordings and any other material as may be brought to your attention.

Unauthorised use of any of the content may constitute a breach of the terms and conditions and/or intellectual property laws.

We reserve the right to alter, amend or bring to an end any terms and conditions provided here without notice.

All rights falling outside the terms of the Creative Commons licence are retained or controlled by The Open University.

Head of Intellectual Property, The Open University

Edited and designed by The Open University.

Printed and bound in the United Kingdom at the University Press, Cambridge.

Contents

Introduction	5
Learning Outcomes	6
1 Overview	7
1.1 DNA and the genome	7
1.3 Chromosome structure and DNA replication	16
Conclusion	18
Take the next step	19
Acknowledgements	19

Introduction

Genomes are composed of DNA, and a knowledge of the structure of DNA is essential to understand how it can function as hereditary material. DNA is remarkable, breathtakingly simple in its structure yet capable of directing all the living processes in a cell, the production of new cells and the development of a fertilized egg to an individual adult.

DNA has three key properties: it is relatively stable; its structure suggests an obvious way in which the molecule can be duplicated, or replicated; and it carries a store of vital information that is used in the cell to produce proteins. The first two properties of DNA are analysed in this course.

This OpenLearn course provides a sample of Level 1 study in [Science](#).

Learning Outcomes

After studying this course, you should be able to:

- understand the basic composition and structure of DNA
- understand what is meant by complementary DNA base pairing
- understand how base pairing allows a mechanism for DNA replication
- understand the number of DNA molecules within a chromosome.

1 Overview

1.1 DNA and the genome

1.1.1 The chemical structure of DNA

This course explores the chemical nature of the genome. Genomes are composed of DNA, and a knowledge of the structure of DNA is essential to understand how it can function as hereditary material. DNA is remarkable, breathtakingly simple in its structure yet capable of directing all the living processes in a cell, the production of new cells and the development of a fertilized egg to an individual adult.

DNA illustrates beautifully the precise relationship between molecular structure and biological function. DNA has three key properties: it is relatively stable; its structure suggests an obvious way in which the molecule can be duplicated, or replicated; and it carries a store of vital information that is used in the cell to produce proteins. In this course, we will examine the chemical nature of DNA, which accounts for both its stability and the way it can be replicated – the first two of these three key properties.

Click below to view the video.

Video content is not available in this format.

It was in 1953 that James Watson and Francis Crick ([Figure 1](#)) working in the UK, deduced and published the three-dimensional structure of DNA. It was the year that might be described as the dawn of molecular biology, for their publication was to have far-reaching consequences in terms of our understanding of the nature of life, i.e. how cells function at the molecular level. So monumental was this work that they were awarded, together with Maurice Wilkins, the Nobel Prize for Medicine with Physiology in 1962. In this section we shall examine in some detail the structure of the DNA molecule.



Figure 1 James Watson and Francis Crick in 1953 with the double helix model of DNA, which they built to help them to determine its structure; the model was assembled from metal clamps used to hold test-tubes and other bits of laboratory apparatus.

Watson and Crick showed that **DNA** or **deoxyribonucleic acid** has a *double helix* structure: two helical strands like spiral staircases, coiled round one another. It is this structure that accounts for the stability of DNA, one of its key features. Its simplest representation is shown in [Figure 2](#). We will first consider the composition of each separate strand by taking it apart to reveal the building blocks of this molecule. Then we will show how the two strands interact to form the characteristic double helix structure of DNA.

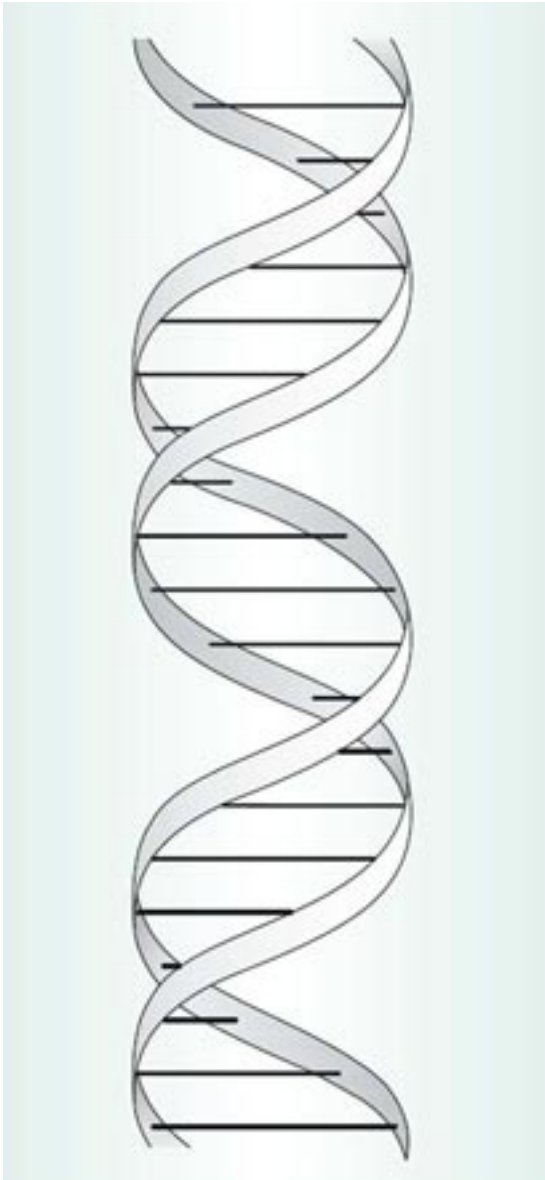


Figure 2 A simplified model showing the double helix structure of DNA.

Much as a protein is a string of amino acids, so each strand of the double helix is a string of building blocks, called **nucleotides**. Each nucleotide consists of three components: phosphate, a sugar molecule and a **base** ([Figure 3](#)). (The type of sugar is deoxyribose, hence the molecule is known as deoxyribonucleic acid.) The phosphate and the sugar are the same in each nucleotide but the base can differ. There are four different bases in DNA: *adenine*, *guanine*, *cytosine* and *thymine*. When illustrating a length of DNA we can simplify the name of each base to a single capital letter, so that **A**=adenine, **G**=guanine, **C**=cytosine and **T**=thymine, as shown in [Figure 4](#).

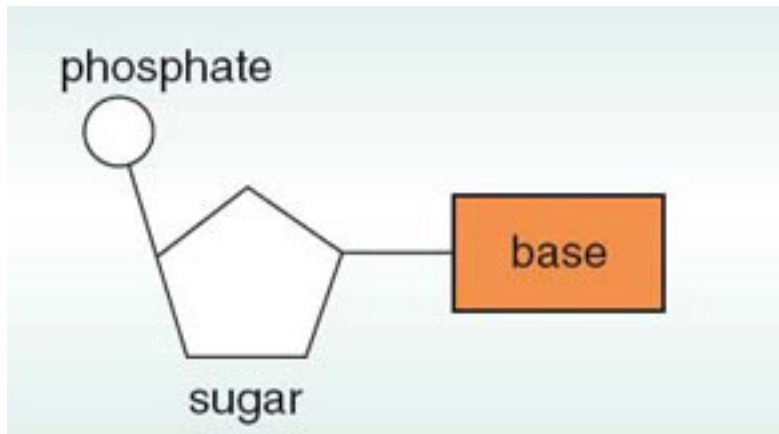


Figure 3 A nucleotide composed of phosphate, a sugar molecule and a base.

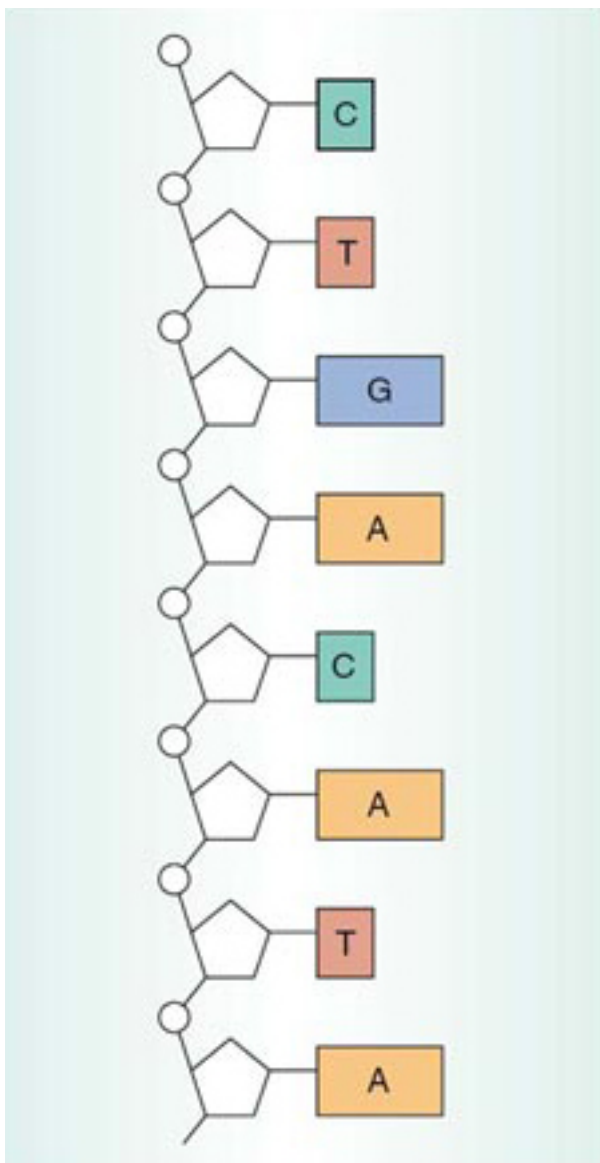


Figure 4 A short segment of a single, straightened strand of DNA, in which the bases, adenine, guanine, cytosine and thymine, are denoted simply as A, G, C, and T, respectively. The sugar—phosphate backbone is shown in white. (This figure represents one strand of the double helix.)

In each strand, the phosphate of one nucleotide is joined to the sugar of another nucleotide, and so on down the strand, as shown on the left (in white) of the single strand shown in [Figure 4](#). The strand of alternating phosphates and sugars is known as the *sugar—phosphate backbone*, and the bases protrude away from this towards the other strand of the helix.

SAQ 1

How many nucleotides are shown in the DNA segment in [Figure 4](#)?

Answer

The simplest way of answering this question is to count the number of rectangles that represent the bases; there are eight.

There is a simplified way of describing the structure of DNA in text, and that is to write out the sequence of bases in a single strand, representing each base by its initial letter A, G, C or T. We will use this shorthand extensively from now on.

SAQ 2

What is the base sequence in the portion of the DNA molecule shown in [Figure 4](#), starting from the top?

Answer

CTGACATA.

These sequences are usually written or printed in lines like the letters or characters on this page. This is a convention, like reading left to right and top to bottom of this page. There may be many hundreds of thousands of bases within a single strand of a DNA molecule, in a long linear sequence. Such a sequence might appear as follows (reading from left to right):

...AAACGCGCGTATATAAATCGCTAGCTTCAACGACTGCTGACGTAGTTCCC....

You may be surprised to learn that DNA molecules are by far the largest known molecules on Earth. One DNA molecule runs the full length of the chromosome. If the DNA molecules of all the chromosomes in the nucleus of a single human cell were uncoiled, stretched out straight and laid end to end, they would measure about two metres. If all the DNA in all your cells were stretched out end-to-end, it would reach to the Moon and back about 10 000 times!

If you look at the comparatively short sequence of 50 bases printed above, and think about how many ways a simple coding language of just four letters could be rearranged in such a sequence, you will gain some appreciation of the huge variety of sequences that is possible. Take, for example, a short chain of just eight nucleotides. Since there are four different bases, there are four options for each position, and therefore $4 \times 4 \times 4 \times 4 \times 4 \times 4 \times 4 \times 4 = 65\,536$ possible different sequences for a DNA molecule of just eight nucleotides! A DNA molecule consisting of thousands of nucleotides therefore represents a vast store of potential information.

So far we have considered a single strand of DNA, but [Figure 2](#) showed that DNA has a *double* helix structure. Here each of the two 'ribbons' spiralled around each other represents the sugar—phosphate backbone of [Figure 4](#), whilst the horizontal bars represent the bases of the two strands.

The key to understanding the structure of DNA and how it functions in the cell lies in the interaction between the bases in each strand at the core of the molecule. Along the length of a strand within the double helix, each base makes a specific pairing with a corresponding base in the other strand. These interactions are known as **base-pairing**, for which there are very precise rules.

T pairs only with A, and C pairs only with G. These pairs are called *complementary base pairs*.

These pairs of complementary bases sit 'flat' within the double helix, rather like the steps of a spiral staircase, as shown in [Figure 2](#). The relative proportions of each base in DNA is related to the base-pairing rules just outlined.

SAQ 3

If you were to extract some DNA from cells, and isolate and purify the four bases, how many A bases would you expect to find relative to T bases? Similarly, how many C bases would you find relative to G?

Answer

A consequence of the base-pairing rules is that the amount of A in a DNA molecule is always equal to the amount of T; the same applies to the amount of C relative to that of G.

Click below to view the video.

Video content is not available in this format.

The alignment of base pairs within a DNA molecule is shown in [Figure 5](#) (on next page). Here the helix is shown unwound, with the two sugar—phosphate backbones now parallel but still on the outside; the complementary base pairs form the core of the molecule. Since the sequence of bases on one strand is complementary to the sequence of bases on the other strand, the two strands of the double helix are described as *complementary*.

The structure of DNA can be summarized as follows. The complementary base pairs sit at the core of the molecule, and are arranged flat like the steps of a spiral staircase. The two sugar—phosphate backbones lie to the outside of the helix, each spiralled around the other.

Watson and Crick, in their famous 1953 paper published in the journal *Nature*, wrote:

We wish to suggest a structure for deoxyribose nucleic acid (DNA). This structure has novel features which are of considerable biological interest.... It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

On the following page we consider how new DNA molecules are produced (the scientific term for this is *synthesized*), and show how true Watson and Crick's prediction was. But first watch the following video then try question 1.

Click below to view the video.

Video content is not available in this format.

Question 1

In a fragment of double-stranded DNA, there is a total of 100 bases, of which 30 are C. Calculate the total number of each of the following items in the DNA fragment: (a) complementary base pairs; (b) G bases; (c) T bases; (d) A bases.

Answer

- (a) 100 bases would form 50 complementary base pairs.
- (b) Since C always pairs with G, the number of G bases is the same as the number of C bases, i.e. 30.
- (c) and (d) 60 of the 100 bases are either C or G, so the remaining 40 bases are either T or A. As A always pairs with T, then half of this number, i.e. 20 are T and 20 are A.

1 1.2 How DNA is replicated

Cell division involving the nuclear division of mitosis produces two progeny cells, which contain identical genetic material, which is also identical with that of the original parent cell. This is how a fertilized egg grows into an adult many-celled organism. For one cell to become two new ones with identical genetic material, the DNA in each chromosome must undergo a process in which an identical copy is made.

As noted above, Watson and Crick postulated that DNA base-pairing provides a mechanism by which the DNA might be copied (see Figure 5). This DNA copying mechanism, usually referred to as **DNA replication**, is the process we consider here.

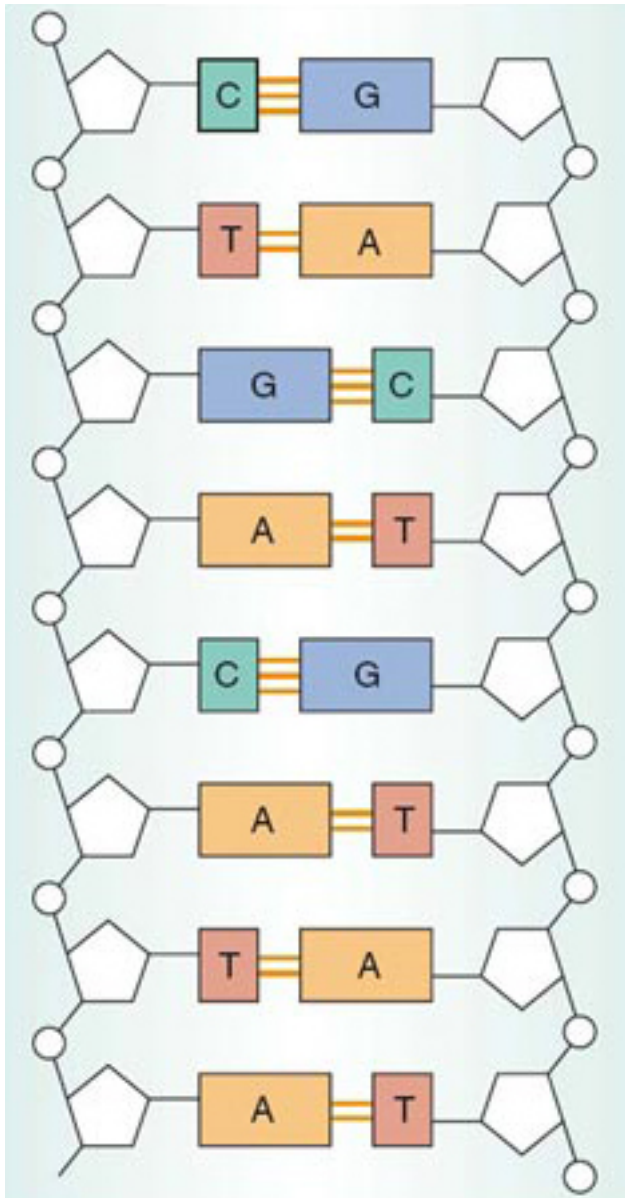


Figure 5 A portion of a DNA molecule with the helix unwound, showing the complementary base pairs between the two strands held together by base-pairing interactions (shown as orange lines).

The separation of the two strands of DNA is an early event in the process of DNA replication. Once the strands have been separated, new DNA strands are synthesized; the enzyme that brings about this process is called *DNA polymerase*, which adds nucleotides to each separated strand according to the base-pairing rules.

[Figure 6](#) shows the principal stages of DNA replication. The two strands of the double helix shown in [Figure 6a](#) unwind, starting at one end, to expose the bases on each strand. The two complementary single strands are shown separated in [Figure 6b](#). Each of these strands now acts as a *template*, a mould, for DNA replication. The base-pairing rules are the basis of this process; that is, the nucleotides are added in a manner that places complementary bases opposite each other – C always opposite G and vice versa, A always opposite T and vice versa. At the same time, the two new sugar—phosphate backbones are formed. The result is the production of two identical double-stranded DNA molecules ([Figure 6c](#)). Initially each double-stranded DNA molecule is unwound, as

shown in [Figure 6c](#); later, the paired strands wind around each other to form the characteristic double helix structure, as shown in [Figure 6d](#).

This process has been termed *semi-conservative replication*, meaning 'half-conserved replication'. This is because in each new DNA double helix, one of the two original strands is conserved, i.e. is unchanged, from the original parent molecule; these are labelled as the parent strands in [Figure 6b](#). The second strand in each daughter DNA double helix has been newly synthesized in its entirety; these are labelled as the new strands in [Figure 6c](#). Or, to put it crudely, each daughter double helix is only 'half new'; each has one parent strand and one new strand, as shown in [Figure 6d](#).

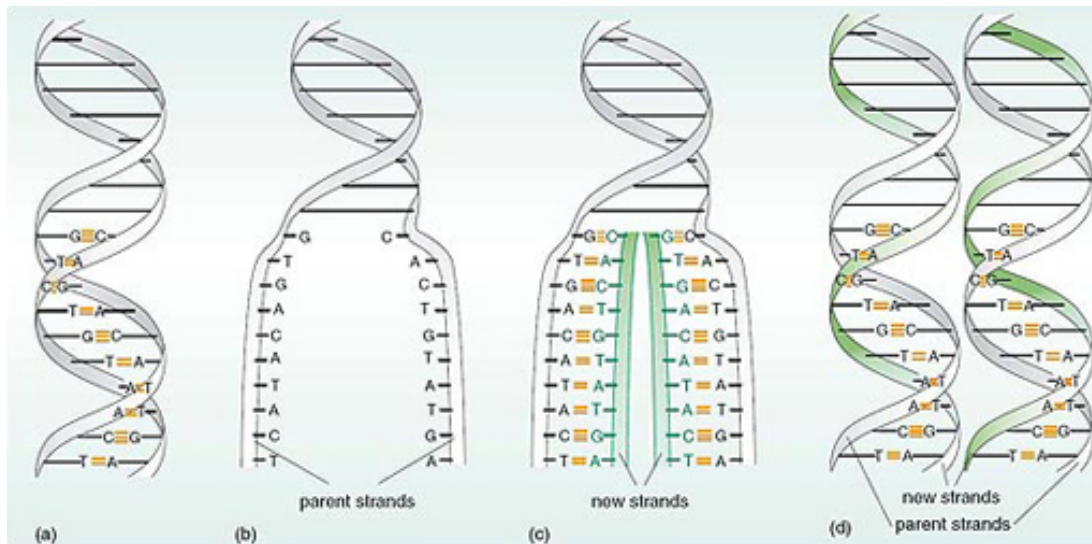


Figure 6 The process of DNA replication. (a) A portion of a DNA double helix showing 10 labelled complementary base pairs. (b) Part of the double helix has unwound and come apart at one end, revealing two single strands. (c) Part of each strand has been replicated, but the two double-stranded DNA molecules have not yet wound into two daughter double helices. The new DNA strands are shown in green. The process continues until all of the parent DNA molecule has been replicated. (d) The two double-stranded DNA molecules wind into two daughter double helices. Each double helix is composed of an original, parent strand and a newly synthesized strand.

[Figure 6](#) shows just a small portion of DNA being replicated. The process continues until the whole of the DNA molecule has been replicated, and the two daughter DNA molecules form the characteristic double-helix structures shown in [Figure 10.6d](#), as opposed to the unwound products of replication shown in [Figure 6c](#). Before the cell can divide to produce identical progeny cells, all of the DNA molecules in the cell have to replicate to produce two identical copies.

This, in outline, is how DNA is copied during most cycles of cell division. If you compare [Figures 6a](#) and [6d](#) you will see that both DNA molecules in (d) are a faithful copy of the sequences of bases of the parent molecule in (a).

Click below to view the video.

Video content is not available in this format.

An important consequence of DNA structure is that the genetic information it contains is copied into more DNA.

1.3 Chromosome structure and DNA replication

DNA replication is closely linked to chromosome replication, which in turn is linked to cell division, which includes the nuclear division of mitosis. This raises an intriguing question: how many DNA molecules are present in a chromosome?

Chromosomes are composed of DNA intimately associated with proteins. When the chromosomes become visible at the beginning of mitosis, the DNA has already been replicated, and the chromosomes are double structures; that is, each chromosome consists of two chromatids. During mitosis, the two chromatids of each pair separate to opposite ends of the cell so that at the end of mitosis each chromosome is a single unit. Evidence suggests that such a single (unreplicated) chromosome contains one continuous DNA double-stranded molecule running along its length, as shown in [Figure 7a](#). That makes a very long molecule! This suggests that, because genes are linked together along the length of a chromosome, each gene must be a short section of a DNA double helix molecule.

Before the next round of cell division begins, each chromosome forms paired chromatids when the DNA is replicated. Each chromatid contains one DNA double helix along its length, as shown in [Figure 7b](#).

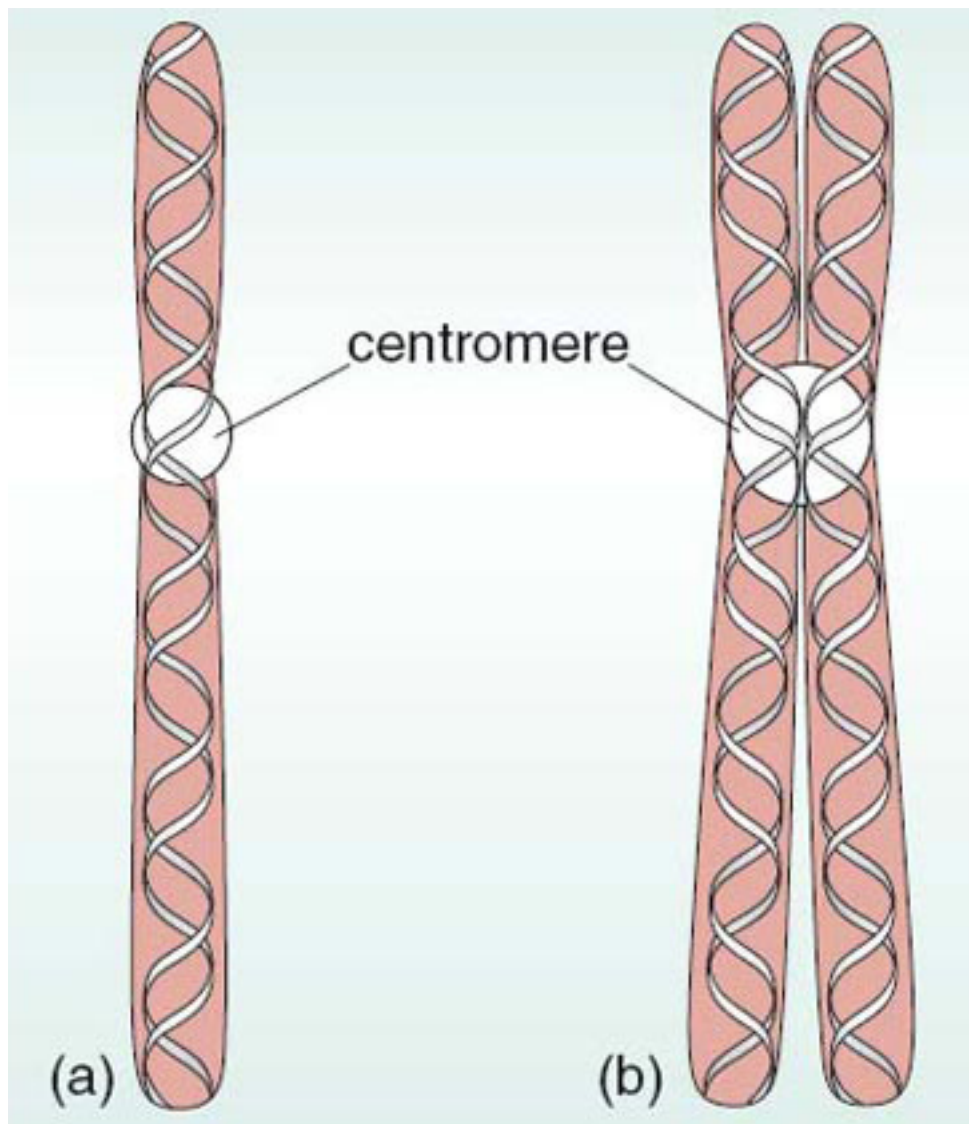


Figure 7 The number of DNA molecules in a chromosome: (a) a chromosome prior to replication contains a single DNA molecule; (b) a chromosome that has been replicated and consists of two chromatids, each comprising a single DNA double helix molecule. In reality, in each chromatid each DNA double helix is associated with protein, which condenses into a coil and then into a supercoil.

Conclusion

This free course provided an introduction to studying Science. It took you through a series of exercises designed to develop your approach to study and learning at a distance and helped to improve your confidence as an independent learner.

Take the next step



If you enjoyed this course, why not explore the subject further with our paid-for short course, *Science: human genetics and health issues*?

[Find out more about *Science: human genetics and health issues*.](#)

Acknowledgements

The content acknowledged below is Proprietary (see terms and conditions) and is used under licence.

Grateful acknowledgement is made to the following sources for permission to reproduce material in this course:

Course image: [Robert Couse-Baker](#) in Flickr made available under [Creative Commons Attribution-ShareAlike 2.0 Licence](#).

All other materials included in this course are derived from content originated at the Open University.

Don't miss out:

If reading this text has inspired you to learn more, you may be interested in joining the millions of people who discover our free learning resources and qualifications by visiting The Open University - www.open.edu/openlearn/free-courses