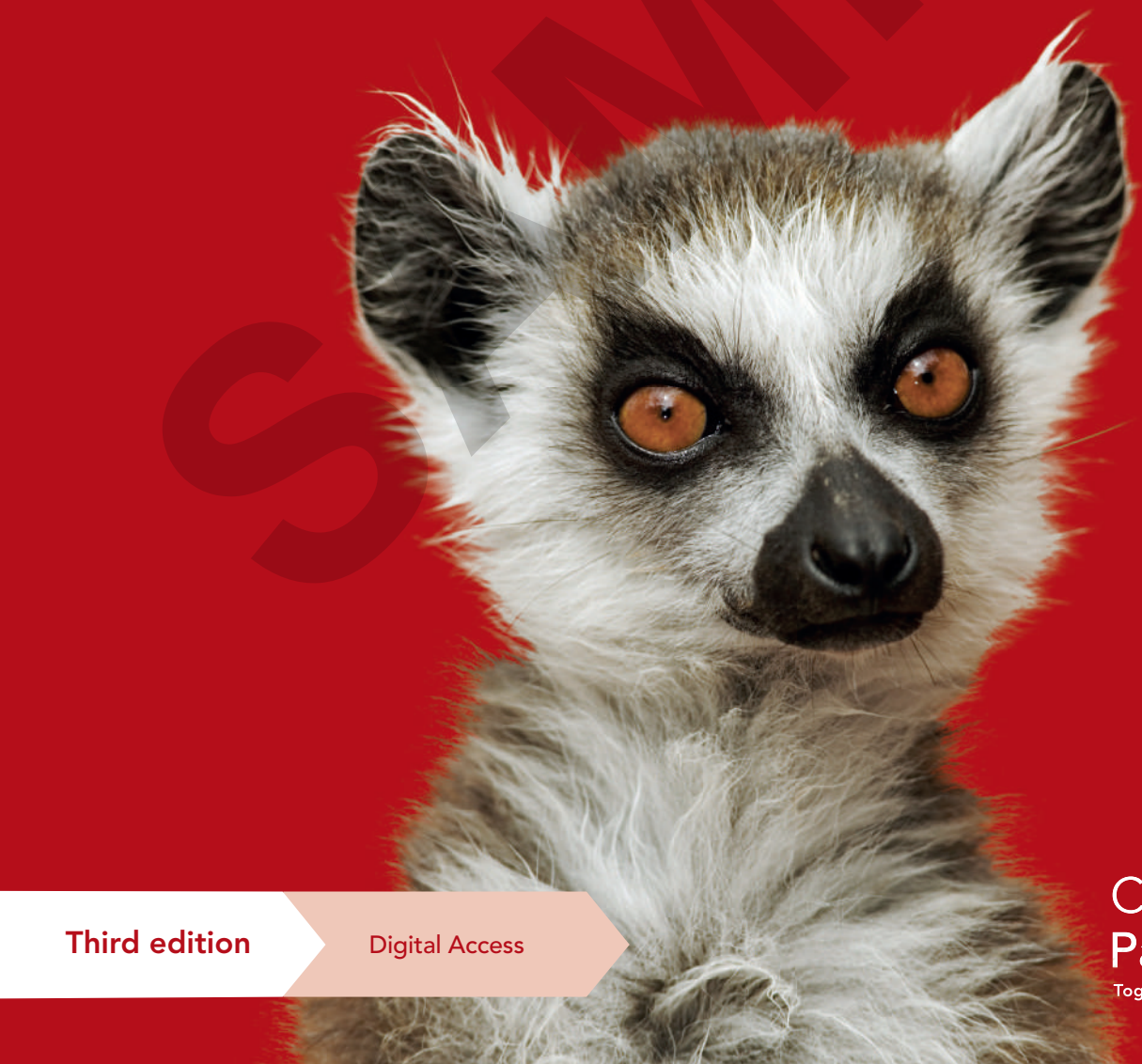
CAMBRIDGE
UNIVERSITY PRESS

Executive Preview

# Biology

## for the IB Diploma

MULTI-COMPONENT SAMPLE





Third edition

Digital Access

Cambridge Panel
Together with IB teachers

Brighter Thinking
Better Learning

Dear Teacher,

Welcome to the new edition of our *Biology for the IB Diploma* series, providing full support for the new course for examination from 2025. This new series has been designed to flexibly meet all of your teaching needs, including extra support for the new assessment. This preview will help you understand how the coursebook, the workbook and the teacher's resource work together to best meet the needs of your classroom, timetable and students.

This Executive Preview contains sample content from the series, including:

- A guide explaining how to use the series
- A guide explaining how to use each resource

In developing this new edition, we carried out extensive global research with IB Biology teachers – through lesson observations, interviews and work on the Cambridge Panel, our online teacher research community. Teachers just like you have helped our experienced authors shape these new resources, ensuring that they meet the real teaching needs of the IB Biology classroom.

The coursebook has been specifically written to support English as a second language learners with key subject words, glossary definitions in context and accessible language throughout. We have also provided new features that help with active learning, assessment for learning and student reflection. Numerous exam-style questions with answers in the digital coursebook, which accompanies the print coursebook, ensure your students feel confident approaching the assessment and have all the tools they need to succeed in their examination.

Core to the series is the brand-new digital teacher's resource. It will help you support your learners and confidently teach to the new IB Biology guide, whether you are new to teaching the subject or more experienced. For each topic there are lesson ideas and activities, common misconceptions to look out for, worksheets, PowerPoint presentations, answers to the coursebook, extra wrap-up activities and more. Also included is a practical guide to help your students develop their academic writing.

Please take five minutes to find out how our resources will support you and your learners. To view the full series, you can visit our website or speak to your local sales representative. You can find their contact details here:

**cambridge.org/gb/education/find-your-sales-consultant**

Best wishes,

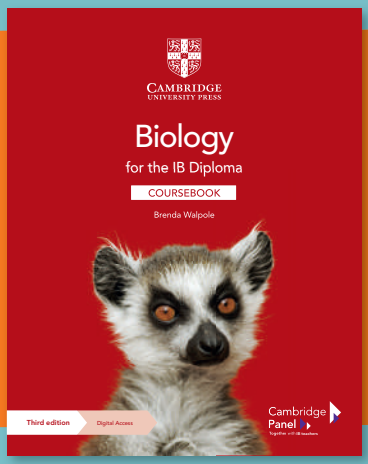
**Micaela Inderst**

Senior Commissioning Editor for the IB Diploma

Cambridge University Press

# How to use this series

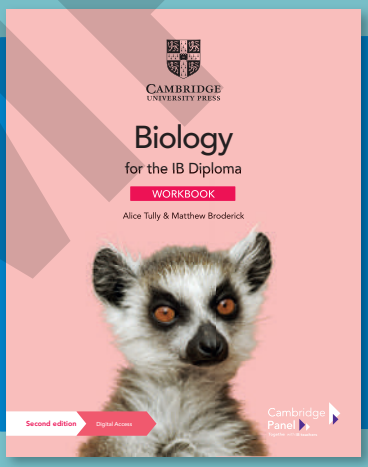
This suite of resources supports students and teachers of the IB Biology Diploma course. All of the books in the series work together to help students develop the necessary knowledge and scientific skills required for this subject.

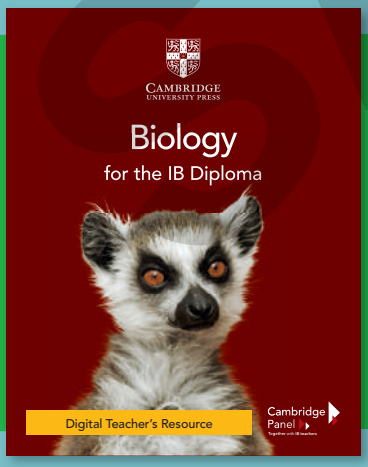



The coursebook with digital access provides full coverage of the latest IB Biology Diploma course.

It clearly explains facts, concepts and practical techniques, and uses real world examples of scientific principles. A wealth of formative questions within each chapter help students develop their understanding, and own their learning. A dedicated chapter in the digital coursebook helps teachers and students unpack the new assessment, while exam-style questions provide essential practice and self-assessment. Answers are provided on Cambridge GO, supporting self-study and home-schooling.

The workbook with digital access builds upon the coursebook with digital access with further exercises and exam-style questions, carefully constructed to help students develop the skills that they need as they progress through their IB Biology Diploma course. The exercises also help students develop understanding of the meaning of various command words used in questions, and provide practice in responding appropriately to these.




The Teacher's resource supports and enhances the coursebook with digital access and the workbook with digital access. This resource includes teaching plans, overviews of required background knowledge, learning objectives and success criteria, common misconceptions, and a wealth of ideas to support lesson planning and delivery, assessment and differentiation. It also includes editable worksheets for vocabulary support and exam practice (with answers) and exemplar PowerPoint presentations, to help plan and deliver the best teaching.

CAMBRIDGE
UNIVERSITY PRESS

# Biology

## for the IB Diploma

COURSEBOOK

Brenda Walpole



Third edition

Digital Access

Cambridge Panel
Together with IB teachers

# Table of contents

## Chapter 10 Defence against disease



## Chapter 11 Evolution, speciation and Ecosystems



## Chapter 12 Ecological relationships

# How to use this book

Throughout this book, you will find lots of different features that will help your learning. These are explained below.

## UNIT INTRODUCTION

A unit is made up of a number of chapters. The key concepts for all the chapters covered in a unit are summarised in the unit opening chapter as the introduction.

## LEARNING OBJECTIVES

Each chapter in the book begins with a list of learning objectives. These set the scene for each chapter, help with navigation through the coursebook and indicate the important concepts in each topic.
A bulleted list at the beginning of each section clearly shows the learning objectives for the section.

## GUIDING QUESTIONS

This feature contains questions and activities on subject knowledge you will need before starting this chapter.

The content in this book is divided into Standard and Higher Level material. A vertical line runs down the margin of all Higher Level material, allowing you to easily identify Higher Level from Standard material.

## Links

These are a mix of questions and explanation that refer to other chapters or sections of the book.

Key terms are highlighted in **orange bold** font at their first appearance in the book so you can immediately recognise them. At the end of the book, there is a glossary that defines all the key terms.

## KEY POINTS

This feature contains important key learning points (facts) to reinforce your understanding and engagement.

## EXAM TIPS

These short hints contain useful information that will help you tackle the tasks in the exam.

## SCIENCE IN CONTEXT

This feature presents real-world examples and applications of the content in a chapter, encouraging you to look further into topics. You will note that some of these features end with questions intended to stimulate further thinking prompting you to consider some of the benefits and problems of these applications.

## NATURE OF SCIENCE

Nature of Science is an overarching theme of the IB Biology Diploma course. The theme examines the processes and concepts that are central to scientific endeavour, and how science serves and connects with the wider community. Throughout the book, there are 'Nature of Science' features that discuss particular concepts or discoveries from the point of view of one or more aspects of Nature of Science.

### THEORY OF KNOWLEDGE

This section stimulates thought about critical thinking and how we can say we know what we claim to know. You will note that some of these feature end with questions intended to get you thinking and discussing these important Theory of Knowledge issues.

### INTERNATIONAL MINDEDNESS

Throughout this Biology for the IB Diploma course, the international mindedness feature highlights international concerns. Science is a truly international endeavour, being practised across all continents, frequently in international or even global partnerships. Many problems that science aims to solve are international and will require globally implemented solutions.

### TEST YOUR UNDERSTANDING

These questions appear within each chapter and help you develop your understanding. The questions can be used as the basis for class discussions or homework assignments. If you can answer these questions, it means you have understood the important points of a section.

### WORKED EXAMPLE

Many worked examples appear throughout the text to help you understand how to tackle different types of questions.

### REFLECTION

These questions appear at the end of each chapter. The purpose is for you as a learner to reflect on the development of your skills proficiency and your progress against the objectives. The reflection questions are intended to encourage your critical thinking and inquiry-based learning.

### EXAM-STYLE QUESTIONS

Exam-style questions at the end of each chapter provide essential practice and self-assessment. These are signposted in the print coursebook and can be found in the digital version of the coursebook.

### SELF-ASSESSMENT CHECKLIST

These appear at the end of each chapter/section as a series of statements. You might find it helpful to rate how confident you are for each of these statements when you are revising. You should revisit any topics that you rated 'Needs more work' or 'Almost there'.

| I can... | Subsection | Needs more work | Almost there | Confident to move on |
|---|---|---|---|---|
| | | | | |

## Free online material

Additional material to support the Biology for the IB Diploma course is available online.

This includes Assessment guidance – a dedicated chapter in the digital coursebook helps teachers and students unpack the new assessment and model exam specimen papers. Additionally, answers to the Test your understanding and Exam-style questions are also available.

Visit Cambridge GO and register to access these resources.

# Chapter 3

# DNA and protein synthesis



## INTRODUCTION

Nucleic acids are very large macromolecules composed of a backbone of sugar and phosphate molecules each with a nitrogenous base attached. In Chapter 2 the basic structure of these molecules was considered. Here we will look at the vital role of nucleic acids in producing the proteins we need for life and how the genetic information contained in DNA is passed from one generation to the next.

# 3.1 DNA replication

## LEARNING OBJECTIVES

In this section you will:

- understand that DNA replication is a semi-conservative process that produces two identical new molecules
- learn that the enzyme helicase unwinds the double helix and separates the strands
- understand that the polymerase chain reaction is a laboratory process that amplifies small quantities of DNA
- recognise that gel electrophoresis separates DNA fragments by their charge and size and is useful in paternity and forensic investigations

- recognise that DNA strands are antiparallel and are orientated in opposite directions
- learn that DNA replication is regulated by a series of enzymes: primase, polymerase and ligase
- understand that DNA polymerase can only work in a 5′ to 3′ direction
- discover that DNA polymerases proofread new DNA strands.

## GUIDING QUESTIONS

- How is inherited material copied?
- How do laboratory techniques enable us to analyse DNA?

## KEY POINTS

**DNA replication** is copying DNA so that two identical new molecules are produced.

**replication fork** is the point where the DNA double helix is being separated to expose the two strands as templates for replication.

**semi-conservative replication** happens when both strands of a DNA double helix are used as templates for replication so that new DNA molecules contain one original and one new strand.

## 3.1.1 DNA replication

An essential feature of DNA is that it must be able to replicate itself accurately, so that when a cell divides, the genetic code it carries can be passed on to the daughter cells. DNA replication copies DNA precisely so that new molecules are produced with exactly the same sequence of bases as the original strands. DNA replication takes place in the nucleus during interphase of the cell cycle when DNA is not tightly coiled (Section 6.5).

As Figure 3.1.1 shows, DNA replication does not occur in a haphazard manner. An enzyme called DNA helicase unzips one region of the DNA molecule and nucleotides are added in a step-by-step process that links them to one another and to their complementary bases in an area known as the replication fork.

1. The first step in the process is the ‘unzipping’ of the two strands. **DNA helicase** moves along the double helix, unwinding the two strands, which separate from one another as the relatively weak hydrogen bonds between the bases are broken.
2. The unpaired nucleotides are exposed and each single strand now acts as a template for the formation of a new complementary strand. Free nucleotides move into place: C pairs with G and A pairs with T.
3. The free nucleotide bases form complementary pairs with the bases on the single DNA strands. **DNA polymerase** is the enzyme involved in linking the new nucleotides into place. Finally, the two new DNA molecules are rewound, each one forming a new double helix.

The two new DNA strands that are produced are absolutely identical to the original strands. Complementary base pairing between the template strand and the new strand ensures that an accurate copy of the original DNA is made every time replication occurs. DNA replication is said to be semi-conservative replication because no DNA molecule is ever completely new. Every double helix contains one ‘original’ and one ‘new’ strand.

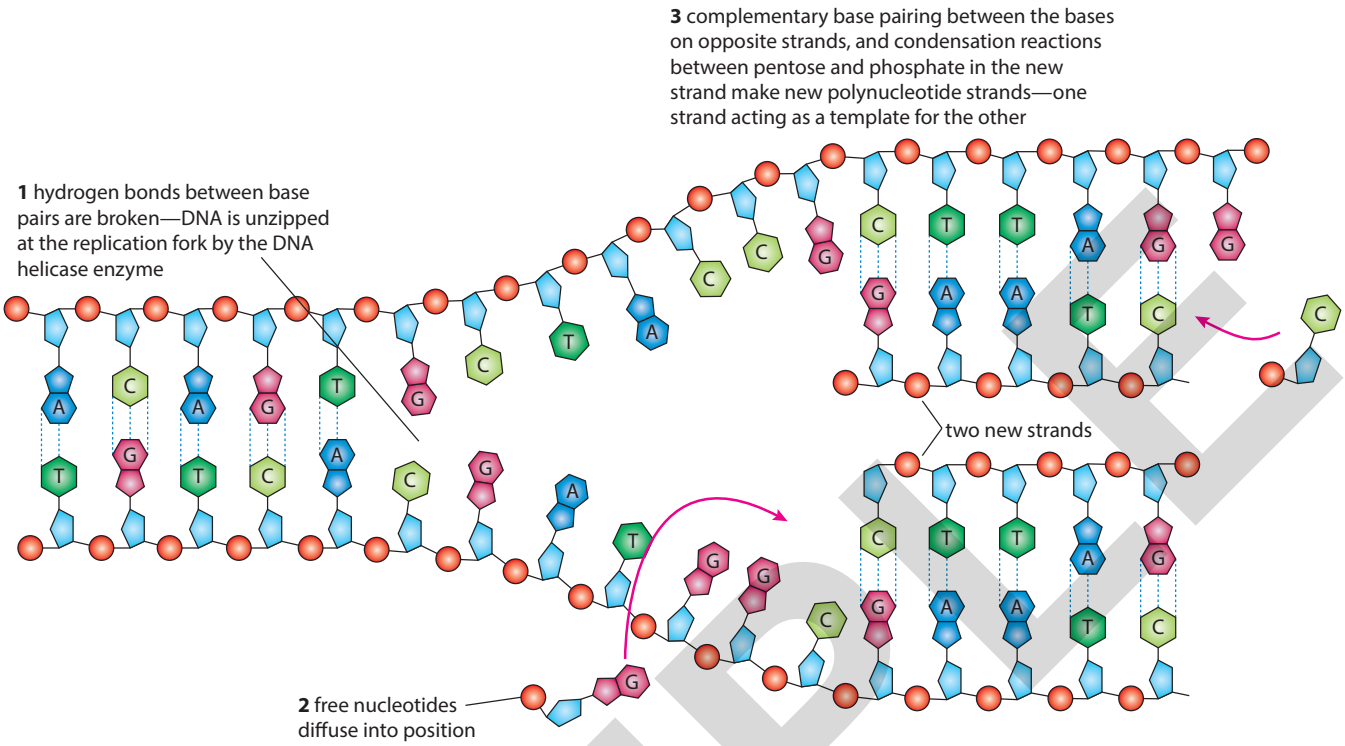



**Figure 3.1.1:** DNA replication.

## NATURE OF SCIENCE

### Obtaining evidence: Meselson and Stahl's experiment and semi-conservative replication of DNA

The research of Meselson and Stahl demonstrates the importance of making and testing a hypothesis in science. They investigated the two hypotheses about DNA replication that were current in the 1950s. The first hypothesis proposed that when DNA is replicated the original helix is conserved unchanged and the newly produced helix contains all new material. This conservative hypothesis was in contrast to the semi-conservative hypothesis, which proposed that one of the original DNA strands from a helix would always be found as one-half of the new double helix produced after replication. Meselson and Stahl designed their experiments using *Escherichia coli*. The bacteria were grown on a medium containing nitrogen $^{15}N$, which is a heavy **isotope** of the normal $^{14}N$. These isotopes were essential to Meselson and Stahl's experiments. After many generations the bacteria have incorporated $^{15}N$ into their cells, so that their DNA became 'labelled' with the heavy isotope and could be identified easily. The bacteria were then transferred to a new medium containing the lighter isotope $^{14}N$, and allowed to grow for a period of time that corresponded to the length of a generation. Figure 3.1.2 shows how the labelled DNA would be distributed among the daughter molecules after one and two replications, according to the semi-conservative theory and the conservative theory. Meselson and Stahl's careful measurements of the amounts of $^{15}N$ in the daughter molecules after one replication showed that all the helices contained one strand of labelled DNA and one strand of normal DNA. Their results therefore supported the theory of semi-conservative replication.

CONTINUED

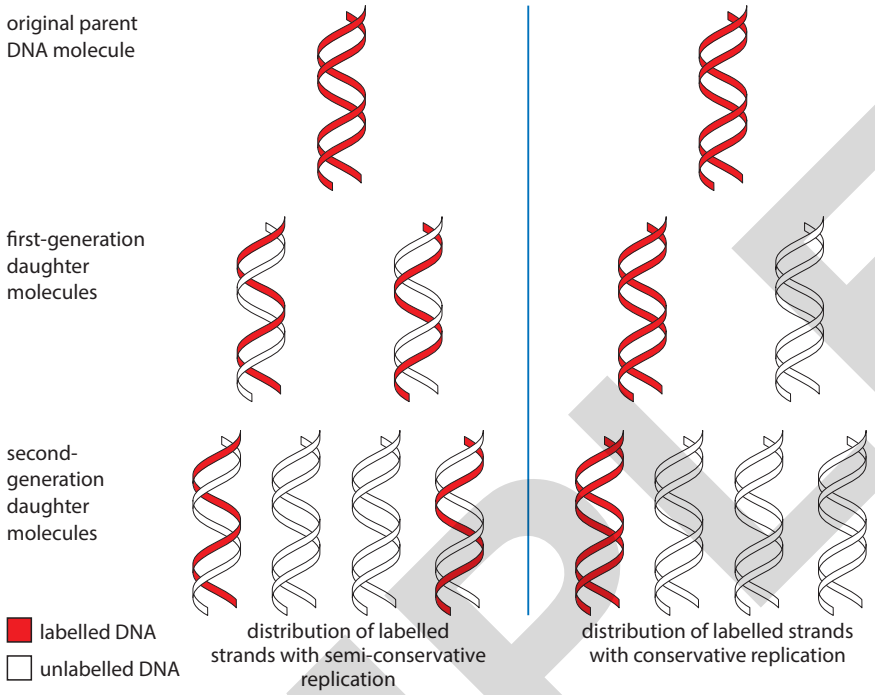



**Figure 3.1.2:** The distribution of labelled DNA in daughter molecules after replication, according to the semi-conservative theory and the conservative theory of replication.

## 3.1.2 DNA sequencing

DNA sequencing is a technique that analyses sequences of DNA bases to work out the sequence of individual genes, groups of genes or even entire chromosomes and genomes. Since the advances in technology made during the Human Genome Project at the turn of the century many new and automated methods of sequencing have been developed. Geneticists and forensic scientists use these techniques to analyse and compare sequences in DNA, some of which are common to different species and others that are repeated many times. The results are used in criminal investigations, for establishing family relationships and in medical diagnoses.

KEY POINTS

**chromosome** in eukaryotes, a structure consisting of a long thread of DNA and protein that carries the genetic information of the cell; in bacteria, the DNA molecule that contains the genetic information of the cell.

**genome** refers to the whole of the genetic information of an organism.

### DNA profiling

At a crime scene, forensic scientists check for fingerprints because a person's fingerprint is unique and can be used to identify them. Forensic scientists also collect samples of hair, skin, blood and other body fluids left at a crime scene because they all contain a person's DNA and that too is a unique record of their presence.

Matching the DNA from a sample to a known individual is called DNA profiling. In forensic science,

KEY POINT

**DNA profiling** is the process of producing a specific DNA pattern, called a profile, from a sample of DNA.

DNA profiles from crime scenes can be used to establish the possibility of guilt or to prove a suspect innocent (Figure 3.1.3). DNA profiling can also be used to determine paternity. For example, a woman might believe that a particular man is the father of her child. By comparing DNA samples from all three individuals – the woman, the man and the child – paternity can be established.

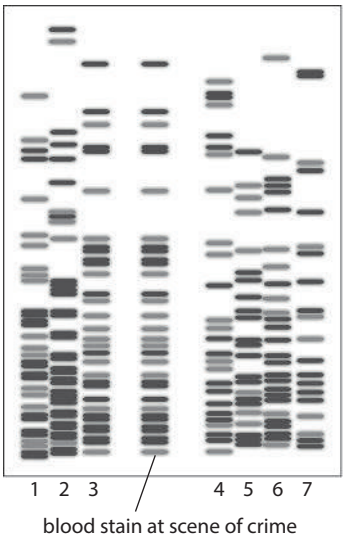



**Figure 3.1.3:** DNA profile of a blood stain found at the scene of a crime compared with profiles from seven suspects. Which suspect was at the scene of the crime? What is the evidence to support your answer?

### The polymerase chain reaction

The polymerase chain reaction (PCR) is an automated method used to amplify (copy) segments of DNA. To study DNA in forensic or genetic analysis large amounts of a sample of DNA are needed. In most cases only small amounts are available, so PCR is a vital tool.

**KEY POINT**

**polymerase chain reaction (PCR)** a process in which small quantities of DNA are artificially amplified for research and diagnosis.

Sometimes, at a crime scene or when a body is found after a very long time, only a minute amount can be collected. The PCR can make millions of copies of tiny amounts of DNA so there is a sufficient amount to produce a profile or study a gene of interest. Technicians must take great care when handling the original sample so that it is not contaminated with their own or other DNA. Only the DNA region that is of interest will be amplified. A geneticist might be studying gene function or a forensic scientist could want to match crime scene DNA with that of a suspect. In medicine, the DNA of bacteria or viruses can be used in diagnosis. DNA amplified by PCR may be used for sequencing or to produce DNA profiles using gel electrophoresis (see the following stages).

The stages in the process are:

1 Denaturation – Heat the DNA sample to 95 °C so that the double strands of DNA separate into two single strands.

2 Annealing and extension – Cool to 68 °C and add the enzyme *Taq* polymerase, DNA primers and DNA nucleotides that are needed to build duplicate copies of the original DNA using the two strands of DNA as templates.

3 Repeat the cycle of separation and synthesis of new DNA 30–40 times so that eventually more than a billion exact copies of the original DNA segment are produced (Figure 3.1.4).

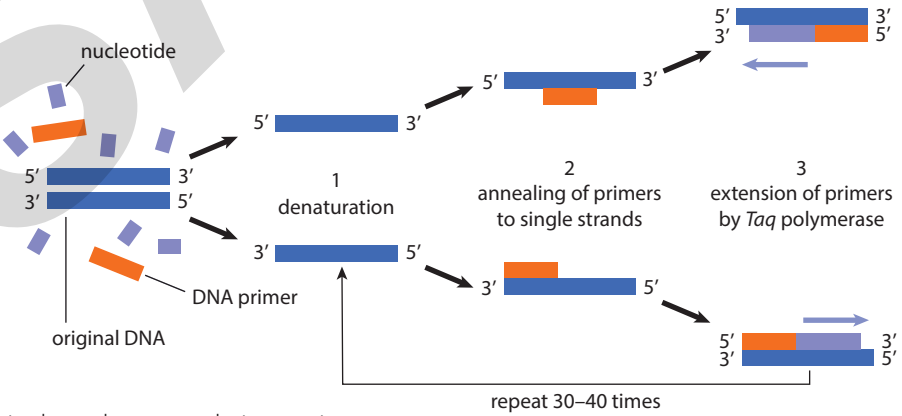



**Figure 3.1.4:** Stages in the polymerase chain reaction.

PCR is automatically controlled in a machine called a thermocycler that alters the temperature of the reaction every few minutes, firstly to cause DNA separation and then for synthesis of new strands.

### *Taq* polymerase

The PCR needs the enzyme DNA polymerase to build new strands of DNA from the existing template strands, just as a cell does when it copies its DNA. The DNA polymerase used in PCR is *Taq* polymerase. It has been obtained from the heat-tolerant bacterium *Thermus aquaticus*, which is found in thermal vents and in hot springs. *Taq* polymerase is unaffected by high temperature and is most active around 70 °C, a temperature at which a human DNA polymerase would not work.

**KEY POINT**

**Taq polymerase** a heat-stable DNA polymerase named after the microorganism *Thermus aquaticus* used to amplify DNA in the polymerase chain reaction.

### PCR primers

Primers are short sequences of nucleotides that provide a starting point for DNA synthesis by *Taq* polymerase. The technician using PCR will add primers to select the region of DNA that is be copied, or amplified. PCR primers are pieces of single-stranded DNA, usually around 20 nucleotides in length.

Two primers are used in each PCR reaction. The sequences bind by complementary base pairing to opposite strands of DNA at the ends of the region to be copied. The *Taq* polymerase will add nucleotides to the primers so that the region between them is copied (Figure 3.1.5).

### PCR in diagnosis

PCR can be used in medical diagnosis to detect genetic sequences of bacteria and viruses and thus identify active infections. Using specific primers, PCR can be used to amplify known sequences that only exist in certain viruses or bacteria. If that sequence is not found and there is no infection, then no amplification will take place and no DNA will be produced. Pathogens that are difficult, or take a long time, to culture can be identified quickly. The detection of the presence of bacteria in clinical specimens such as spinal fluid, blood and urine can enable doctors to make speedy diagnoses and give the correct treatments.

Genetic disorders are caused by mutations (changes in DNA) that can be just a few base sequences, or be changes in large sequences of DNA or, sometimes, whole chromosomes. PCR enables geneticists to study just a small segment of DNA at a specific region of chromosome. The sequences of bases in a gene can be amplified and disorders detected, diagnosed and monitored. In some cases, gene therapy (Section 3.3) is available to rectify these disorders, and PCR can be used to monitor the functioning of the relevant genes and gene segments.

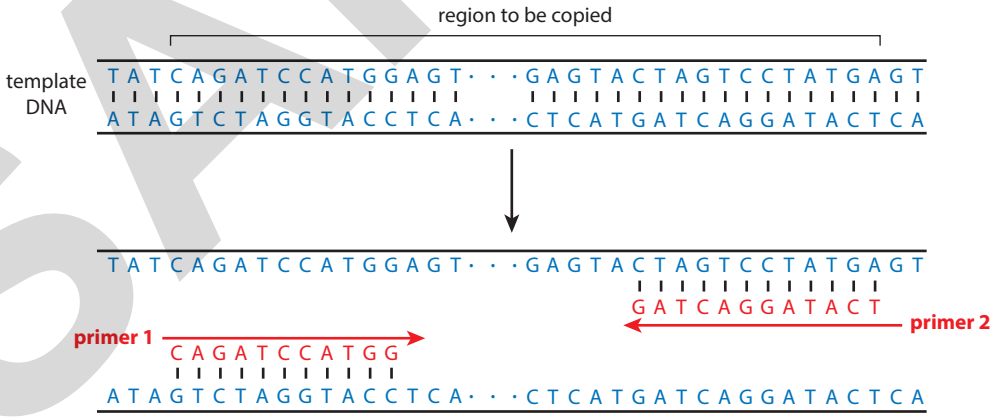



**Figure 3.1.5:** Positions of primers on single strands of DNA that are to be copied.

## SCIENCE IN CONTEXT

### Hunting for coronaviruses

There are two main types of coronavirus test: those that can detect the presence of the virus that is active in the body and those that detect a previous response to the virus by the immune system.

The PCR test is used in the first type of test, and looks for evidence that the COVID-19 virus, SARS-CoV-2, is present in a person's body by detecting the presence of its RNA in a swab sample from their nose or throat. The PCR test detects the genetic material from the virus by amplifying tiny amounts that may be present. PCR can only tell us if the virus is currently present in a person's body.

PCR tests involve several stages so errors are possible between sampling and analysis. False negatives (that is, a result that is negative when in fact the patient has the virus) do occur but estimates are that 80–85% of the results are correct.

## Gel electrophoresis

Gel electrophoresis is a method used to separate fragments of DNA on the basis of size and the electric charge they carry. It can identify natural variations found in every individual's DNA.

Any DNA sample usually contains long molecules that are too large to be used for profiling so DNA profiling often examines repetitive sequences of so-called 'satellite' DNA that vary in their degree of repetitiveness from person to person. These are called variable number tandem repeats (VNTRs) and short tandem repeats (STRs). These regions have repeated sequences of DNA that are very similar in close relatives but so variable in unrelated people that non-relatives are extremely unlikely to have the same repeated sequences.

The DNA fragments are placed in a well in a plate of gel (a jelly-like material) and one well is reserved for a reference fragment known as a DNA ladder, this is a DNA molecule of a known length, used as a reference to estimate the size of the unknown DNA molecules in the sample.

## KEY POINTS

**DNA ladder** DNA molecules of different lengths used in gel electrophoresis, used as a reference to estimate the size of unknown DNA molecules.

**gel electrophoresis** a technique which separates DNA fragments according to their size and charge.

An electric field is applied and because each DNA fragment has a small negative charge, it will move in the electric field, through the gel. The distance a fragment can move depends on its size; smaller fragments move most easily through the gel matrix and travel further, while larger fragments are left behind close to their starting point. After the fragments have been separated in the gel, they are stained and produce a unique pattern of bands called a DNA profile (Figures 3.1.3 and 3.1.6).

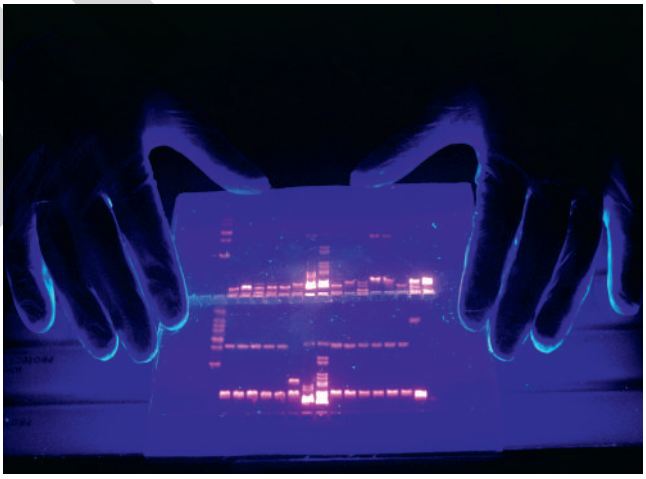


**Figure 3.1.6:** Scientist examining an agarose electrophoresis gel used to prepare a DNA profile. The sample of DNA is marked with a radioactive substance, so the DNA banding pattern appears pink under ultraviolet light. The pattern is preserved by applying radiographic film to the gel.

## SCIENCE IN CONTEXT

### Short tandem repeats and tracing our ancestors

Any one STR will be shared by between 5 and 20% of people who are not related. But in forensic science many STRs are examined at the same time; the more STR regions that are examined, the more accurate the test becomes. The pattern of repeats can identify an individual with a high degree of accuracy. In the world of genealogy (tracing family history), DNA profiling and STRs are also used as vital tools. Today, if you want to prove that you are descended from a certain line then you may be able to use genetics to prove it. Genealogists research the ancestry of families, looking for groups of people who share the same STRs and who can be identified as being related to each other over hundreds or even thousands of years.

## THEORY OF KNOWLEDGE

### DNA profile databases

In the USA, the Federal Bureau of Investigation (FBI) has a national database of DNA profiles from convicted criminals, suspects, missing persons and crime scenes. The data that are held may be used in current investigations and to solve unsolved crimes. There are many commercial laboratories that carry out DNA profiling analysis on behalf of the law enforcement agencies. Many of them check 13 key STR sequences in DNA samples, which vary considerably between individuals. The FBI has recommended that these should be used because they provide odds of one in one thousand million that two people will have the same results.

CODIS is the acronym for the Combined DNA Index System, a computer software program that operates the national database of DNA profiles. Every US state has a statutory right to establish a DNA database that holds DNA profiles from offenders convicted of particular crimes. CODIS software enables laboratories to compare DNA profiles electronically, linking serial crimes to each other and identifying suspects from profiles of convicted offenders. CODIS has contributed to thousands of cases that have been solved by matching crime scene evidence to known convicted offenders.

**To consider:**

1. DNA profiles do not show individual base sequences but only identify repeated sequences. How much confidence should be placed on DNA evidence?
2. How secure is DNA profiling?
3. What are the implications for society if the authorities were to hold a DNA profile for every person?
4. What safeguards should be in place to protect the rights of individuals whose DNA profiles have been placed on a database but who have not been convicted of a crime?
5. Is it right to convict a person on DNA evidence alone?

## TEST YOUR UNDERSTANDING

1. Outline what is meant by the term semi-conservative replication.
2. State the role of the enzyme helicase.
3. Give two examples of the use of DNA profiles.

# 3.1.3 The detailed process of DNA replication

DNA replication ensures that exact copies of existing molecules are produced before a cell divides. The process is said to be semi-conservative and each strand of an existing DNA molecule acts as a template for the production of a new strand (see Nature of Science, Obtaining evidence: Meselson and Stahl's experiment and semi-conservative replication of DNA, earlier in this section).

In eukaryotes, replication is controlled through interactions between proteins, including cyclins and CDKs (Section 6.5) and takes up to 24 hours to complete. Each of the original DNA strands acts as a template to build up a new strand (Figure 3.1.7). The DNA double helix is unwound to expose the two strands for replication by the enzyme DNA helicase, at a region known as a replication fork. The action of helicase creates single-stranded regions, which are less stable than the double-stranded molecule. To stabilise these single strands, single-stranded binding proteins (SSBs) are needed. SSBs protect the single-stranded DNA and allow other enzymes involved in replication to function effectively upon it.

Replication must occur in the 5′→3′ direction (and also in transcription and translation, described in Section 3.2), because the enzymes involved only work in a 5′→3′ direction (adding new nucleotides to the 3′ end of the newly forming DNA molecule). As the two strands are antiparallel, replication has to proceed in opposite directions on the two strands. However, the replication fork where the double helix unwinds moves along in one direction only. This means that on one of the strands replication can proceed in a continuous way, following the replication fork along, but on the other strand the process has to happen in short sections, each moving away from the replication fork (Figure 3.1.8). The strand undergoing continuous synthesis is called the leading strand. The other strand, in which the new DNA is built up in short sections, is known as the lagging strand.

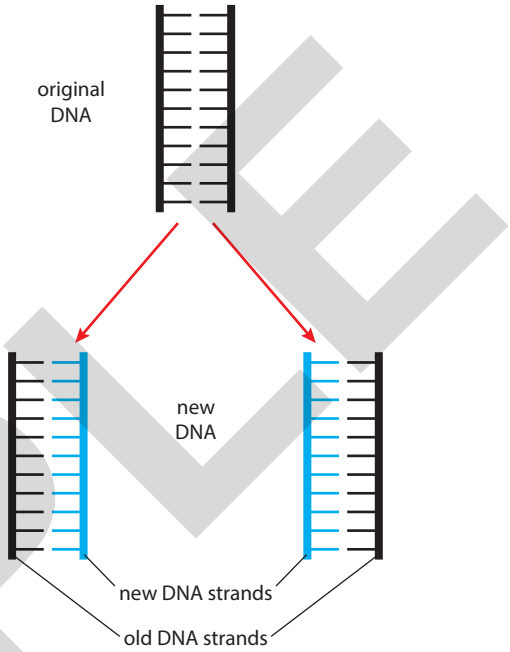



**Figure 3.1.7:** DNA replication is semi-conservative. As it is copied one original strand becomes paired with one new strand. One of the two strands in each new DNA molecule is conserved, hence 'semi-conservative'.

## Copying the leading strand

Replication to produce the leading strand begins at a point on the molecule known as the 'origin of replication' site. First RNA primase adds a short length of RNA, attached by complementary base pairing, to the template DNA strand. This acts as a primer, allowing the enzyme

### KEY POINTS

**lagging strand** is the new strand that is synthesised in short fragments in the opposite direction to the movement of the replication fork.

**leading strand** is the new strand that is synthesised continuously and follows the replication fork.

**single-stranded binding protein** is the protein which binds to single-stranded regions of DNA to protect them from digestion and remove secondary structure.

### KEY POINTS

**DNA polymerase III** extends the new DNA strand in a 5′ → 3′ direction from the RNA primer.

**primer** is short strand of nucleic acid that forms a starting point for DNA synthesis.

**RNA primase** is an enzyme that catalyses the synthesis of RNA primers as the starting point for DNA synthesis.

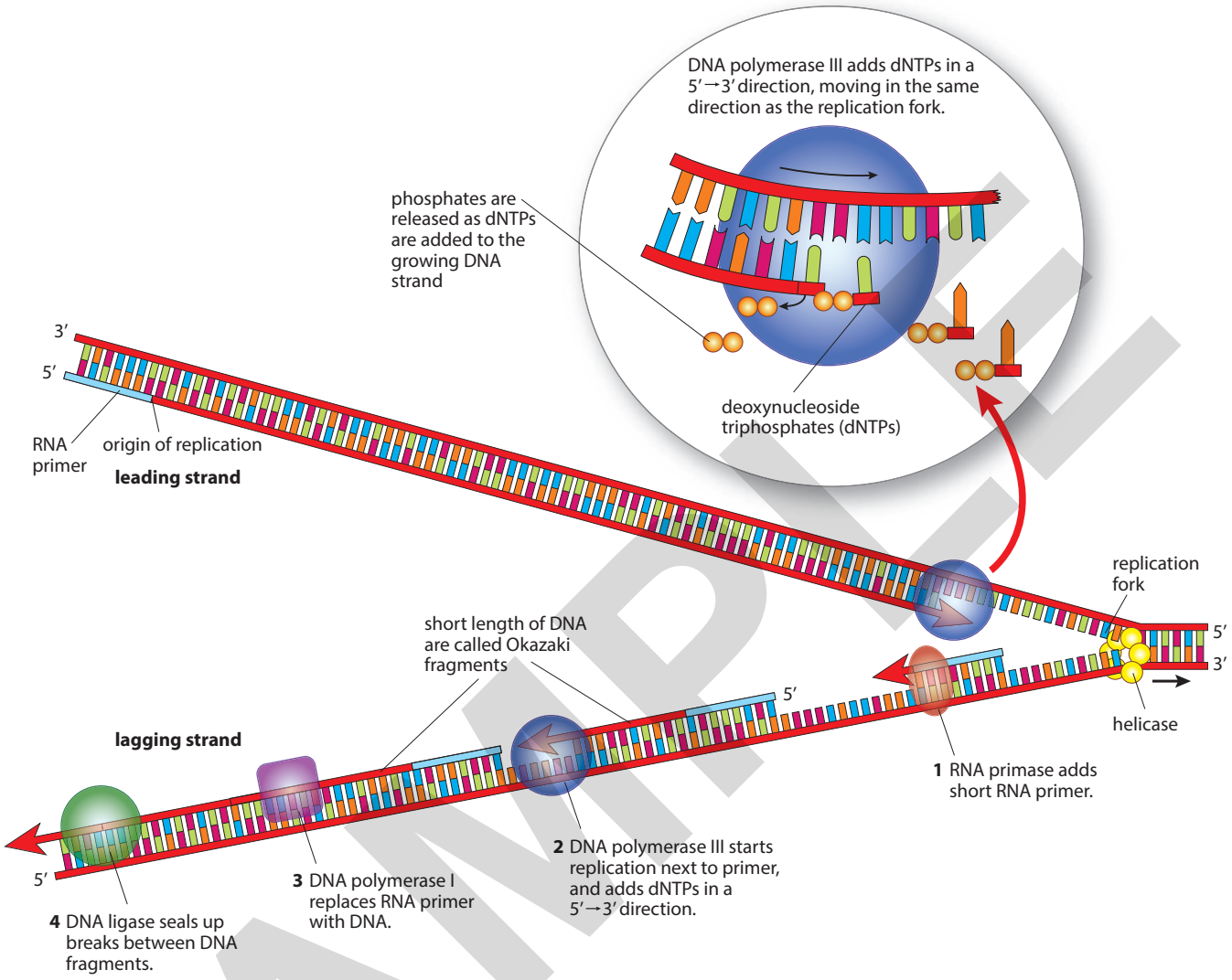



**Figure 3.1.8:** DNA replication showing the leading and lagging strands, and the direction of DNA synthesis.

## KEY POINTS

**deoxynucleoside triphosphate (dNTP)** is a building block for DNA: deoxyribose, three phosphate groups and one of the four bases.

**DNA polymerase I** removes the RNA nucleotides of the primers on the lagging strand and replaces them with DNA nucleotides.

DNA polymerase III to bind. DNA polymerase III adds free 'building units' called deoxynucleoside triphosphates (dNTPs) to the 3′ end of the primer and then to the forming strand of DNA. In this way the new molecule grows in a 5′→3′ direction, following the progress of helicase as it moves the replication fork along the DNA double helix. The RNA primer is later removed by DNA polymerase I. In this way, a continuous new DNA strand is built up on the leading strand.

## EXTENSION

The dNTPs have two extra phosphate groups attached, and are said to be 'activated'. They pair up with their complementary bases on the exposed DNA strand and DNA polymerase III then link together the sugar and the innermost phosphate groups of adjacent nucleotides. The two extra phosphate groups are broken off and released.

## Copying the lagging strand

Synthesis of the lagging strand is a little more complicated, as it has to occur in discontinuous sections, which are then joined together.

1 As for the leading strand, **DNA primase** first synthesises a short RNA primer, complementary to the exposed DNA. This happens close to the replication fork.

2 DNA polymerase III starts replication by attaching at the 3′ end of the RNA primer and adding dNTPs in a 5′→3′ direction. As it does so, it moves away from the replication fork on this strand.

3 DNA polymerase I now removes the RNA primer and replaces it with DNA using dNTPs. Short lengths of new DNA called Okazaki fragments are formed from each primer. The new fragment grows away from the replication fork until it reaches the next fragment.

4 Finally, **DNA ligase** seals up each break between the Okazaki fragments by making sugar–phosphate bonds so that a continuous strand of new DNA is created.

## Proofreading new DNA

DNA polymerases 'proofread' their work as they build up new DNA strands. If the polymerase enzyme detects that an incorrect nucleotide has been added and does not pair up correctly the enzyme will remove and replace it.

### KEY POINTS

**DNA primase** a type of RNA polymerase catalyses the production of a short length of RNA called a primer which is base-paired to the parent DNA strand. The primer is removed when replication is complete and replaced by DNA.

**DNA ligase** joins adjacent Okazaki fragments by forming a covalent bond between adjacent nucleotides.

**Okazaki fragments** short fragments of a DNA strand formed on the lagging strand.

### EXAM TIP

There are several important enzymes to remember in the process of replication so it is helpful to keep a list of them and their jobs.

### TEST YOUR UNDERSTANDING

4 Outline what is meant by antiparallel.

5 State the direction in which DNA replication occurs.

6 Outline the role of DNA primase.

7 Summarise the differences between forming the leading and lagging strands.

### REFLECTION

Could you explain DNA profiling to someone who had never heard of it? Reflect on its importance to your understanding of ourselves.

## Links

- How is the molecular structure of DNA linked to its function? (Chapter 1)
- Why must the genetic code carried by DNA be copied exactly? (Chapter 3.4)
- How is replication involved in cell division? (Chapter 6)

# 3.2 Protein synthesis

## LEARNING OBJECTIVES

In this section you will:

- define transcription as the synthesis of mRNA by RNA polymerase
- learn that complementary base pairing between DNA and mRNA ensures that the polypeptides produced function properly
- define translation as the production of polypeptides from mRNA using tRNA
- recognise how complementary base pairing between codons and anticodons ensures accurate translation
- learn that ribosomes are the sites of translation; free ribosomes synthesise proteins for use within the cell, whereas bound ribosomes synthesis proteins for secretion or use in lysosomes

- understand the directional nature of transcription and translation
- recognise that transcription begins at a promoter region
- understand that in prokaryotes, translation occurs immediately after transcription but eukaryotes modify mRNA by removing introns to form mature mRNA composed of exons
- recognise that exons can be spliced in different ways to produce different proteins from a single gene
- understand how nucleosomes regulate transcription in eukaryotes
- understand that a large portion of the eukaryotic genome consists of non-coding sequences
- learn that non-coding DNA persists for many generations and has important functions
- recognise that polysomes allow many polypeptides to be made at the same time
- understand that translation does not always result in functional protein and that polypeptides are modified before they can function
- learn that amino acids are recycled in the cell by proteasomes.

## GUIDING QUESTIONS

- How does complementary base pairing contribute to the resilience of the genetic code?
- How are enzymes involved in protein production?

## 3.2.1 Transcription

The main role of DNA is to direct the activities of the cell. It does this by controlling the proteins that the cell produces. Enzymes, hormones and many other important biochemical molecules are the proteins that control what the cell becomes, what it synthesises and how it functions. Protein synthesis can be divided into two sets of reactions: the first is transcription and the second is translation. In eukaryotes, transcription occurs in the nucleus and translation in the cytoplasm.

The sections of DNA that code for particular proteins are known as genes. Genes contain specific sequences of bases in sets of three, called triplets. Some triplets control where transcription begins and ends.

## KEY POINTS

**gene** is a particular section of a DNA strand that codes for a specific polypeptide; a heritable factor that controls a specific characteristic.

**transcription** means copying a sequence of DNA bases to mRNA.

**translation** decoding mRNA at a ribosome to produce a polypeptide.

**triplet** a sequence of three bases that code for an amino acid.

## Copying the DNA message to RNA

The first stage in producing a protein is the production of messenger RNA from a segment of DNA so that the genes that code for the required polypeptides can be moved to the cytoplasm. After this the message is translated and the necessary amino acids are used to build polypeptide chains.

The first stage in the synthesis of a protein is the production of an intermediate molecule that carries the coded message of DNA from the nucleus into the cytoplasm where the protein can be produced. This intermediate molecule is called messenger RNA (mRNA). RNA (or ribonucleic acid) has similarities and differences with DNA and these are shown in Table 2.5.1.

**KEY POINT**

**messenger RNA (mRNA)** is a single-stranded transcript of one strand of DNA, which carries a sequence of codons for the production of protein.

The building blocks for RNA are the RNA nucleotides that are found in the nucleus. Complementary base pairing of RNA to DNA occurs in exactly the same way as in the replication process but this time uracil (U) pairs with adenine since the base thymine (T) is not found in RNA. Transcription results in the copying of one section of the DNA molecule, not its entire length. Figure 3.2.1 describes the process.

1 DNA is unzipped by the enzyme RNA polymerase and the two strands uncoil and separate.

2 Free nucleotides move into place along one of the two strands.

3 The same enzyme, RNA polymerase, assembles the free nucleotides in the correct places using complementary base pairing. As the RNA nucleotides are linked together, a single strand of mRNA is formed. This molecule is much shorter than the DNA molecule because it is a copy of just one section, a gene. The mRNA separates from the DNA and the DNA double helix is zipped up again by RNA polymerase.

Once an mRNA molecule has been transcribed, it moves via the pores in the nuclear envelope to the cytoplasm where the process of translation can take place. In prokaryotes, translation occurs immediately after transcription because there is no nuclear envelope.

## Initiating transcription

Before transcription begins, RNA polymerase must attach to a promoter region of DNA. This process is different in prokaryotes and eukaryotes. In prokaryotes

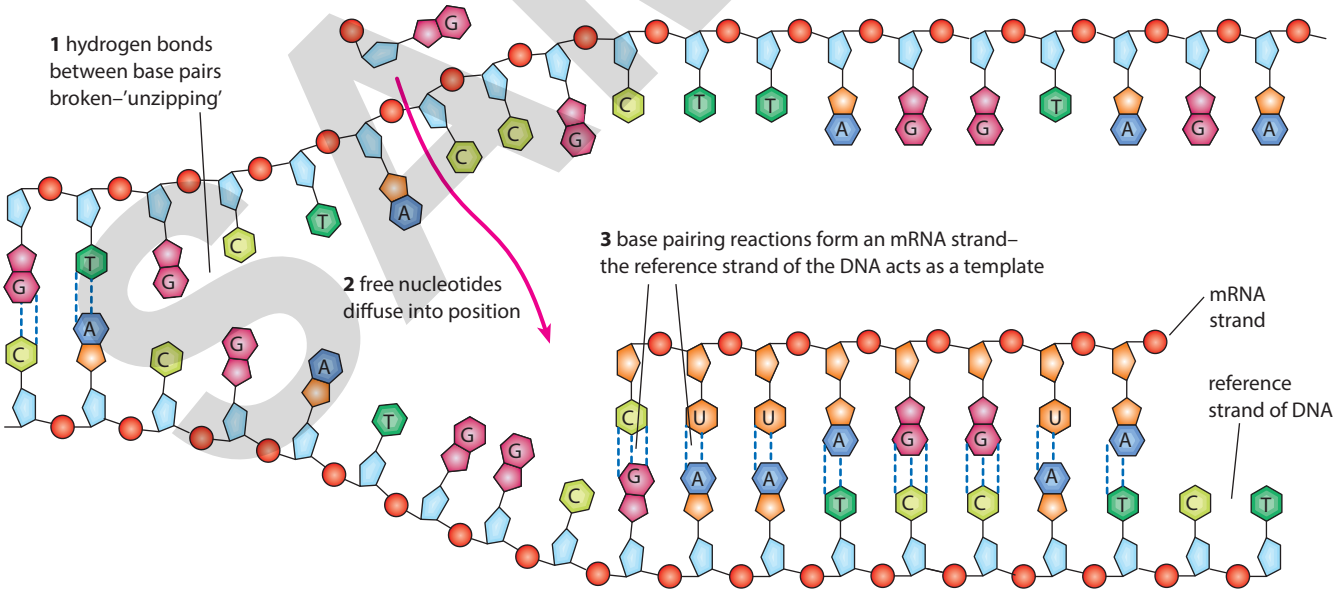



**Figure 3.2.1:** Transcription.

the RNA polymerase binds directly to a promoter region, close to the region that will be transcribed. Transcription is controlled by another section of DNA called the regulator gene, which can produce a repressor molecule that binds to the operator region and prevents transcription (Figure 3.2.2).

Eukaryotes have several transcription factors that are needed to bind the RNA polymerase to the promoter region. Some of these bind to an enhancer region, away from the promoter. The transcription factors (labelled the transcription-initiation complex in Figure 3.2.3) bring the enhancer region close to the promoter and RNA polymerase can then bind and begin transcription (Figure 3.2.3).

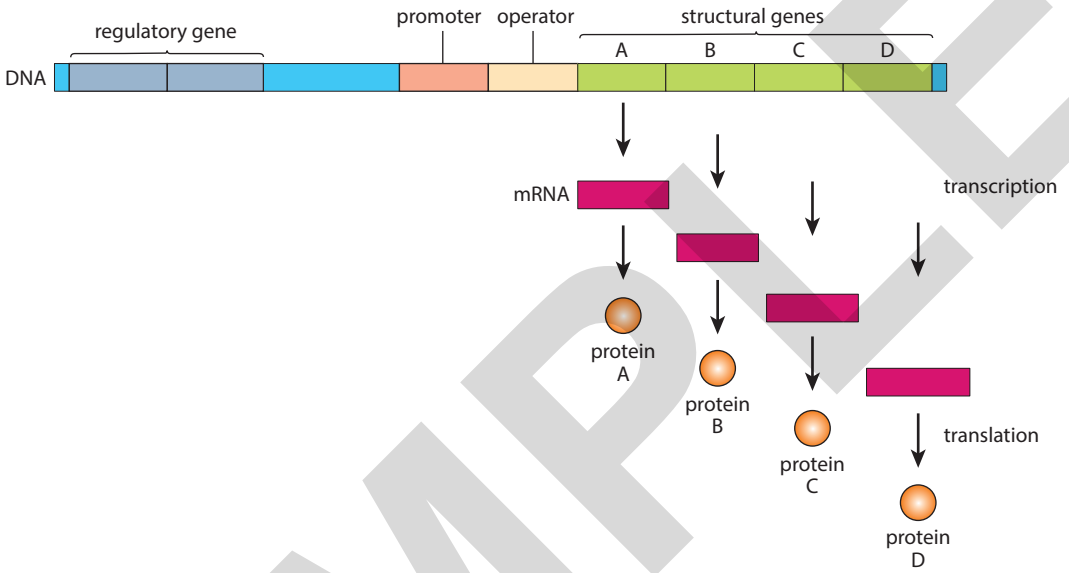



**Figure 3.2.2:** Controlling transcription in a prokaryote.

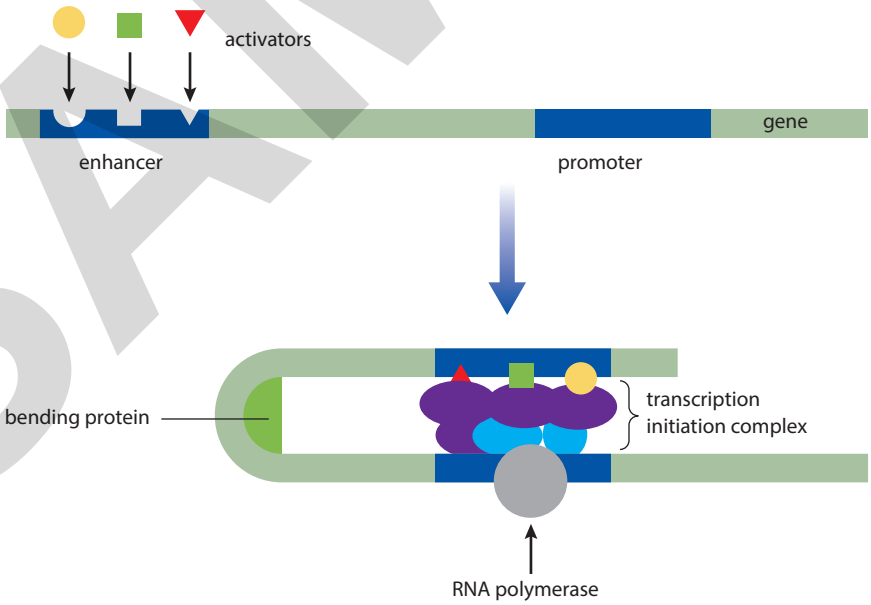



**Figure 3.2.3:** In eukaryotes transcription is controlled by proteins that bind to specific sequences of DNA. Bending proteins are discussed in the Higher Level Section 3.2.3.

Both prokaryotes and eukaryotes have regions of DNA that do not code for protein. In prokaryotes these include promoter regions that serve regulatory functions.

### Regulation of transcription by nucleosomes

DNA in eukaryotes is incorporated into nucleosomes so that the genetic material can be stored in a compact form (Section 2.5). Nucleosomes are important because they can either inhibit or allow transcription by controlling whether the necessary molecules can bind to DNA.

**KEY POINT**

**nucleosome** a part of a eukaryotic chromosome important in regulating transcription, made up of DNA wrapped around histone molecules and held in place by another histone protein.

In order to transcribe genes, activators and enzymes involved in transcription must be able to gain access to DNA. In all eukaryotic species, the regions of DNA that contain the promoters and regulators, which are the binding sites for RNA polymerase and the starting point for transcription, have fewer nucleosomes than other areas, allowing greater access for binding proteins. Conversely, the regions that are transcribed have a higher density of nucleosomes. This suggests that nucleosomes have an important role in determining which genes are transcribed. This in turn can influence other factors such as cell variation and development.

DNA does not need to be completely released from a nucleosome to be transcribed and, although nucleosomes are very stable protein–DNA complexes, they are not static. They can undergo different structural rearrangements including so-called 'nucleosome sliding' and DNA site exposure: if a nucleosome is 'unwrapped' there is a significant period of time during which DNA is accessible. They can also be modified by methylation (Section 3.4.2). The new transcript of mRNA before modification is called pre-mRNA and it becomes known as mature mRNA after removal of introns'

## 3.2.2 Translation

Translation is the process by which the information carried by mRNA is decoded and used to build the sequence of amino acids that eventually forms a protein molecule. During translation, amino acids are joined together in the order dictated by the sequence of codons on the mRNA to form a polypeptide. This polypeptide eventually becomes the protein coded for by the original gene.

Complementary base pairing ensures that the sequence of bases along the mRNA molecule corresponds to the sequence on the original DNA molecule. Each sequence of three mRNA bases is called a **codon** and codes for one specific amino acid, so the order of these codons determines how amino acids will be assembled into polypeptide chains in the cytoplasm. The completed polypeptide chains will be folded to make functioning proteins. Translation is carried out in the cytoplasm by structures called ribosomes and molecules of another type of RNA known as transfer RNA or tRNA.

The mRNA codons that code for each amino acid are shown in Table 3.2.1. From the table you should be able to deduce which amino acid corresponds to any codon.

The genetic code is said to be a **degenerate code** because there are many codons that specify the same amino acid. It is also said to be **universal** because all living things use the same triplet code to specify the same amino acids.

Mutations are changes in sequence of bases that affect protein structure, you can read more about mutations in section 3.3.

### Transfer RNA

The process of translation requires a type of nucleic acid known as transfer RNA (tRNA). tRNA is made of a single strand of nucleotides that is folded and held in place by base pairing and hydrogen bonds (Figure 3.2.4). There are many different tRNA molecules but they all have a characteristic 'clover leaf' appearance with some small differences between them.

At one position on the molecule is a triplet of bases called the anticodon, which pairs by complementary

| First base | | Second base | | | | | | | | Third base |
|---|---|---|---|---|---|---|---|---|---|---|
| | | U | | C | | A | | G | | |
| U | | UUU | phenylalanine | UCU | serine | UAU | tyrosine | UGU | cysteine | U |
| | | UUC | | UCC | | UAC | | UGC | | C |
| | | UUA | leucine | UCA | | UAA | 'stop' | UGA | 'stop' | A |
| | | UUG | | UCG | | UAG | | UGG | tryptophan | G |
| C | | CUU | leucine | CCU | proline | CAU | histidine | CGU | arginine | U |
| | | CUC | | CCC | | CAC | | CGC | | C |
| | | CUA | | CCA | | CAA | glutamine | CGA | | A |
| | | CUG | | CCG | | CAG | | CGG | | G |
| A | | AUU | isoleucine | ACU | threonine | AAU | asparagine | AGU | serine | U |
| | | AUC | | ACC | | AAC | | AGC | | C |
| | | AUA | | ACA | | AAA | lysine | AGA | arginine | A |
| | | AUG | methionine or 'start' | ACG | | AAG | | AGG | | G |
| G | | GUU | valine | GCU | alanine | GAU | aspartic acid | GGU | glycine | U |
| | | GUC | | GCC | | GAC | | GGC | | C |
| | | GUA | | GCA | | GAA | glutamic acid | GGA | | A |
| | | GUG | | GCG | | GAG | | GGG | | G |

**Table 3.2.1:** Amino acids and their associated mRNA codons.

base pairing with a codon on the mRNA strand. At the 3′ end of the tRNA molecule is a base sequence CCA, which is the attachment site for an amino acid.

An amino acid is attached to the specific tRNA molecule that has its corresponding anticodon, by an activating enzyme. As there are 20 different amino acids, there are also 20 different activating enzymes in the cytoplasm.

**KEY POINTS**

**activating enzyme** an enzyme that catalyses the attachment of an amino acid to the appropriate tRNA.

**anticodon** a triplet of bases in tRNA that pair with a complementary triplet (codon) in mRNA.

**transfer RNA (tRNA)** short lengths of RNA that carry specific amino acids to ribosomes during protein synthesis.

## Ribosomes

**Ribosomes** are the site of protein synthesis. Some ribosomes occur free in the cytoplasm and synthesise proteins that will be used within the cell. Others are bound to the endoplasmic reticulum, forming rough endoplasmic reticulum (RER) (Section 5.2), and synthesise proteins that will be secreted from the cell or used within lysosomes.

Ribosomes are composed of two subunits, one large and one small. The subunits are built of protein and ribosomal RNA (rRNA). On the surface of the ribosome are three tRNA-binding sites (site 1, site 2 and the exit site), and one mRNA-binding site (Figure 3.2.5). Two tRNA molecules carrying amino acids can bind to a ribosome at one time. Polypeptide chains are built up in the groove between the two subunits.

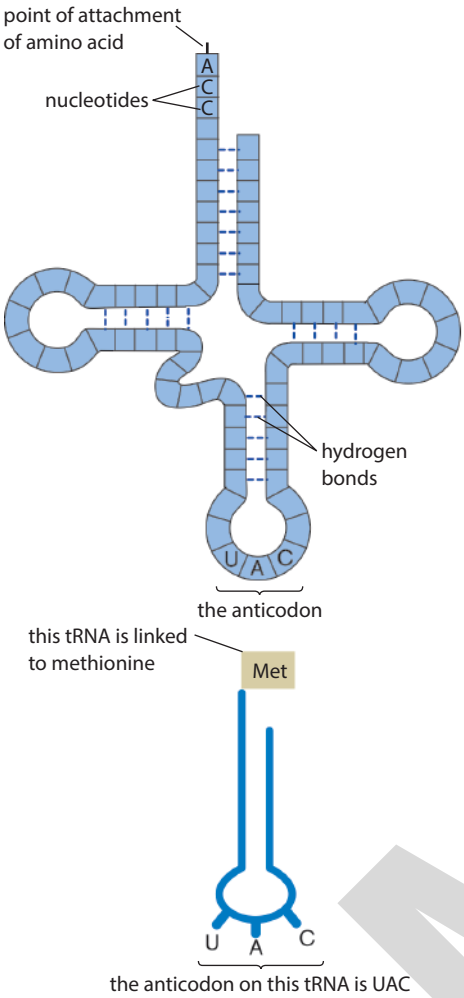



**Figure 3.2.4:** Transfer RNA (tRNA) has a 'clover leaf' shape.

## Building a polypeptide

Ribosomes have binding sites for both the mRNA molecule and tRNA molecules. The ribosome binds to the mRNA and then draws in specific tRNA molecules with anticodons that match the mRNA codons. Only two tRNA molecules bind to the ribosome at once. Each one carries with it the amino acid specified by its anticodon. The anticodon of the tRNA binds to the complementary codon of the mRNA molecule with hydrogen bonds.

When two tRNA molecules are in place on the ribosome, a peptide bond forms between the two amino acids they carry to form a dipeptide. A peptide bond links the amino group of one amino acid to the carboxyl group of the next.

Once a dipeptide has been formed, the first tRNA molecule detaches from both the amino acid and the ribosome. The ribosome moves along the mRNA one triplet to the next codon.

These processes, shown in Figure 3.2.5, are repeated over and over again until the complete polypeptide is formed. The final codon that is reached is a 'stop' codon, which does not code for an amino acid but tells the ribosome to detach from the mRNA. As it does so, the polypeptide floats freely in the cytoplasm or into the RER.

### EXTENSION

A single chromosome contains DNA that codes for many proteins. Most genes are about 1000 nucleotides long, a few are longer and a very small number are less than 100 nucleotides. The size of a gene corresponds to the size of the polypeptide for which it codes.

### TEST YOUR UNDERSTANDING

8 Define transcription.

9 Where does transcription take place in a prokaryotic cell and a eukaryotic cell?

10 Outline the structure of a ribosome.

11 Where in the cell are proteins for secretion or use in lysosomes produced?

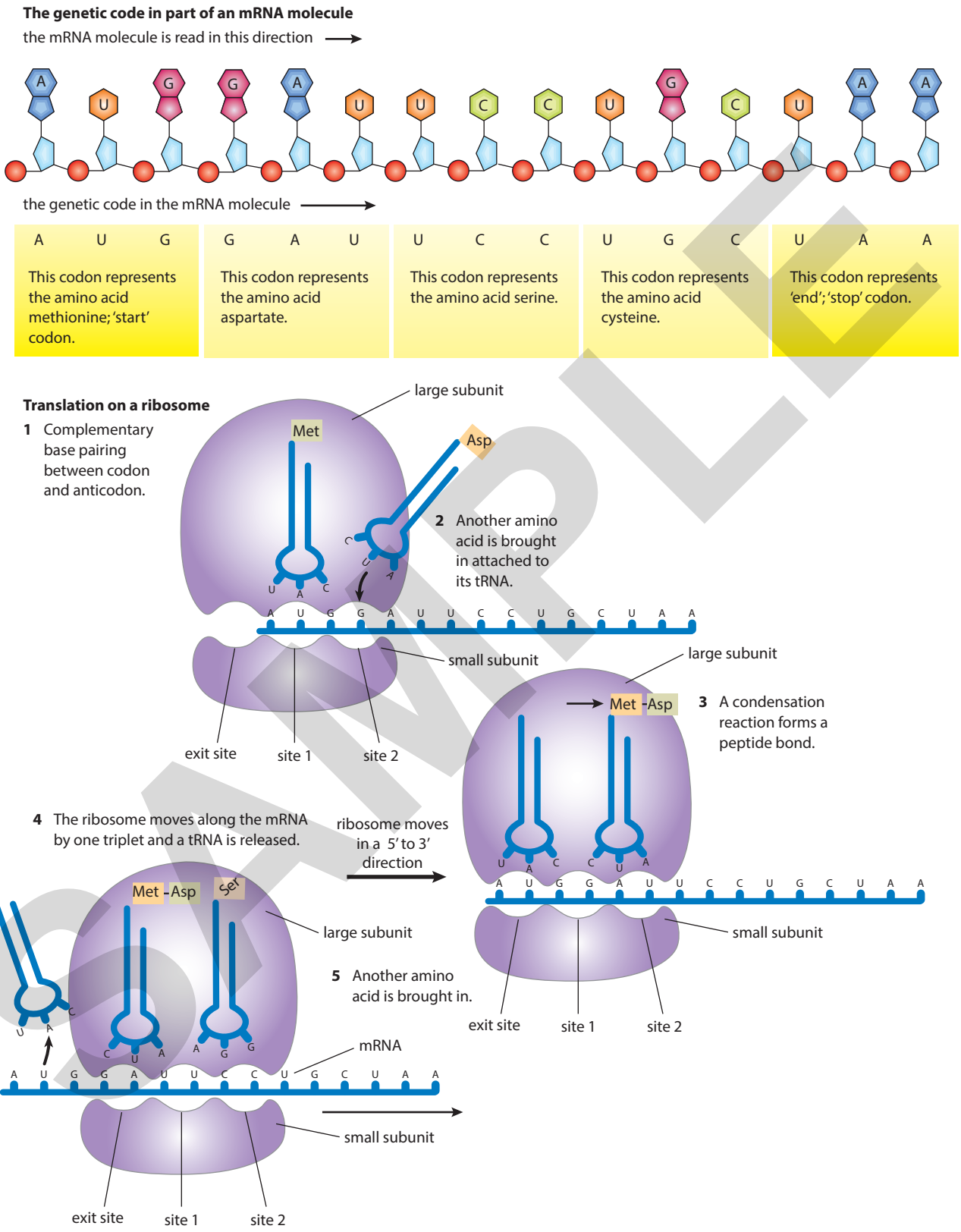



**Figure 3.2.5:** The stages of translation.

# 3.2.3 Non-coding regions of DNA

## Repeated sequences

DNA molecules are very long but every strand has regions that do not code for proteins. For many years these regions were poorly understood but many of them have now been found to have important functions in the regulation of gene expression and other cell activities. Eukaryotic genomes contain mostly non-coding DNA; in fact, nearly 99% of the human genome is non-coding. About 7% of this DNA is thought to have regulatory functions but the exact proportion is not fully known. You may hear this type of DNA being referred to as 'junk DNA', but it is not a term that is now accepted by the scientific community because at least some of this DNA has a function.

When the highly repetitive sequences of non-coding DNA, called variable number tandem repeats (VNTRs), were first discovered they appeared to have no function but as genomes were mapped and compared it was found that several long repeated sequences in humans, mouse and rat DNA were common to all three species. These repeated non-coding sequences regulate and control the activity of genes and possibly embryo development. Studies of the genomes of many species have shown that non-coding DNA is conserved over hundreds of millions of years, suggesting that these regions have been conserved through evolution. It is likely that they give advantages in preserving certain vital genetic characteristics.

**KEY POINT**

**tandem repeat** A tandem repeat is a repeated sequence of DNA base pairs where multiple repeats lie side by side on a chromosome. Tandem repeats are generally associated with non-coding DNA. The number of times the DNA sequence is repeated is variable. Variable tandem repeats are used in DNA profiling because they are very similar in close relatives but very different in unrelated individuals.

Some regions of non-coding DNA act as 'switches' that determine when and where genes are expressed by controlling where and when transcription can begin. Others may be essential for chromosome structure and play a role in cell division. Introns are also transcribed but not translated into proteins.

## Genes for transfer RNA

In humans, the genes that code for tRNA molecules are found on all chromosomes except 22 and Y. These genes do not code for protein but code either for cytoplasmic tRNA or for mitochondrial tRNA. The number of genes that code for tRNA is related to evolutionary history, so that organisms in the Archaea and Eubacteria domains have fewer than those in the domain Eukarya. This seems to be due to the duplication of the genes over time.

## Promoter regions

Promoter regions are DNA sequences that define where transcription of a gene by RNA polymerase begins. They are usually found at the 5′ end of the area where transcription begins.

RNA polymerase requires the presence of a class of proteins known as 'general transcription factors' before transcription can begin. Interactions between the transcription factors, RNA polymerase and the promoter region allow the polymerase to move along the gene so that transcription can occur. Many different transcription factors have been found and each one is able to recognise and bind to a specific nucleotide sequence in DNA. A specific combination of transcription factors is necessary to activate a particular gene.

Other DNA sequences, known as enhancer sequences, are also important and provide a place for regulatory proteins, called activators, to bind.

The role of binding proteins in gene expression is shown in Figure 3.2.3 and discussed in Section 3.4. The proteins bind to the enhancer, which may be some distance from the gene. 'Bending proteins' may then assist in bending the DNA so that the enhancer region is brought close to the promoter. Activators, transcription factors and other proteins attach, so that an 'initiation complex' is formed and transcription can begin. Some activator proteins affect the transcription of multiple genes.

Transcription factors are regulated by signals produced from other molecules such as hormones that are able to activate transcription factors and thus control transcription. Many other molecules in the environment of a cell or an organism can also have an impact on gene expression and protein production.

## Telomeres

Telomeres are regions of repeated nucleotide sequences of non-coding DNA at each end of every eukaryotic chromosome. They protect chromosomes from damage and from fusing with adjacent chromosomes. They have been likened to the protection that a plastic tip on the end of a shoelace gives to protect a lace from fraying.

When chromosomes are replicated during cell division, the enzymes cannot copy the sequences at the end of the chromosomes. Telomeres act to protect important genes from being lost, by capping the end sequences. Cells contain enzymes called telomerases, which can replenish the repeated sequences after cell division in stem cells, but telomeres tend to shorten over time as cells replicate. Telomere shortening can block cell division and, by limiting the number of cell divisions, they protect cells from losing genetic information. The average cell will divide between 50 and 70 times before chromosomes are shortened too much and the cell dies.

**EXTENSION**

You may like to read about the work of Leonard Hayflick (b. 1928), a United States scientist who worked on cloning cells for vaccine production. He discovered that the number of times a cell can divide is not infinite, it is limited to a number called the Hayflick Limit.

## Polysomes

Translation occurs at many places along an mRNA molecule at the same time. The electron micrograph in Figure 3.2.6 shows transcription and translation

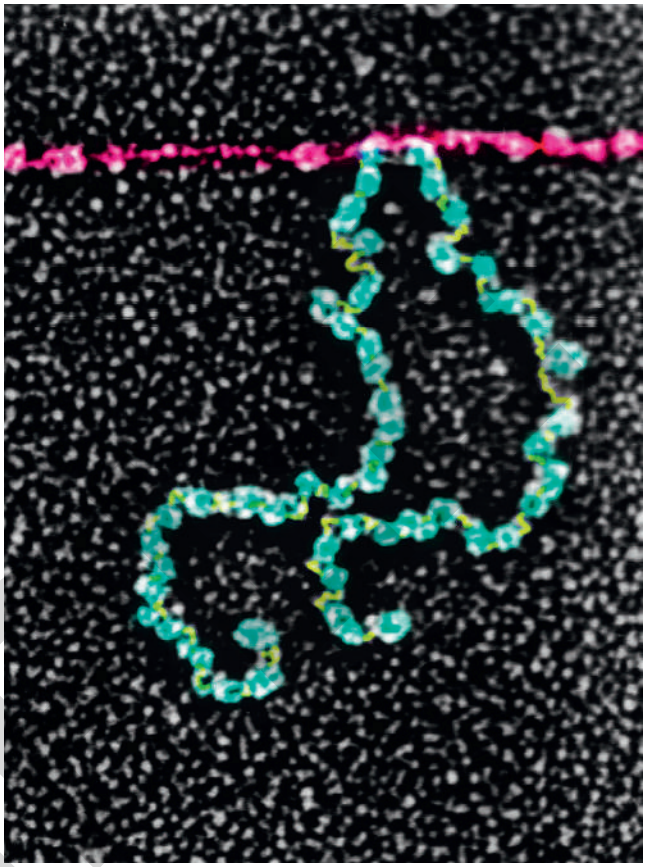


**Figure 3.2.6:** Electron micrograph of polysomes in a bacterium (× 150 000).

occurring simultaneously in a bacterium. A polysome is a group of ribosomes along one mRNA strand (Figure 3.2.7). Part of the bacterial chromosome can be seen as the fine pink line running horizontally along the top of the micrograph and two growing polypeptide chains are shown forming below it. DNA is being transcribed by RNA polymerase and the newly formed mRNA is being immediately translated by the ribosomes. In eukaryotes, the two processes occur in the nucleus and cytoplasm, respectively, and so are separated not only in time but also in location.

**KEY POINT**

**polysome** an arrangement of many ribosomes along a molecule of mRNA so that multiple copies of the same polypeptide are produced at the same time.

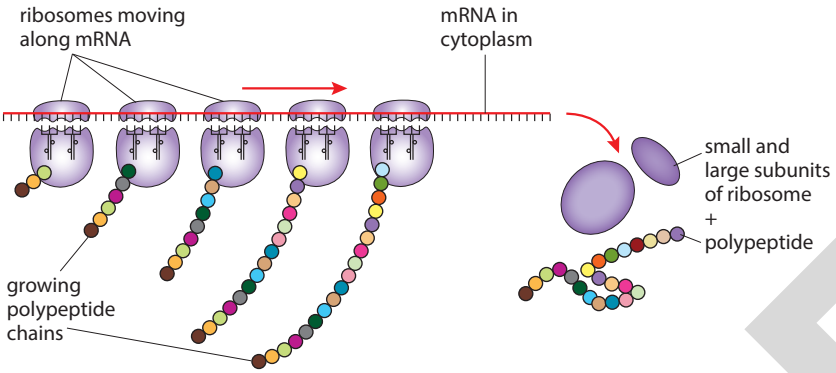



**Figure 3.2.7:** Diagram of a polysome in a eukaryotic cell. Polysomes appear like beads moving along a string of mRNA.

## Post-transcriptional modification: introns and exons

### 1. Introns and exons

In eukaryotes many genes contain sequences of DNA that are transcribed but not translated. These sequences, known as introns, appear in mRNA but are removed before it is translated. Introns (or intragenic regions) are sequences of nucleotides *within* genes. After transcription of a gene, the introns are removed in a process known as post-transcriptional modification. Introns are removed in the nucleus before the mRNA moves to the cytoplasm for translation. Once introns have gone, the remaining sequences of bases, known as exons, are spliced together to form mature mRNA that will then be translated (Figure 3.2.8). Mature mRNA leaves the nucleus via the nuclear pores and moves to the cytoplasm.

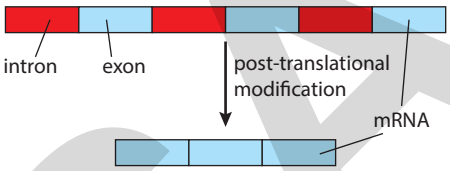



**Figure 3.2.8:** Introns and exons in mRNA.

Introns occur in many genes in all eukaryotic organisms and the number of introns per gene varies considerably between species. For example, introns are common in humans and mice, where genes almost always contain introns but are rare in other eukaryotes such as the yeast (*Saccharomyces cerevisiae*).

The human genome has been found to have about eight introns per gene but other organisms such as fungi may have fewer than 20 in their entire genome.

It also seems that there are more introns in species with smaller populations, and evolutionary and biological factors are thought to influence this.

Exons may be spliced together in different ways so that a number of different, but similar, protein sequences can be produced from a single gene. This alternative splicing means that a single gene can code for more than one protein. Different exons may be included or excluded from the final mRNA produced and as a result, the proteins produced from alternatively spliced mRNAs will contain different amino acid sequences and can have different biological functions. Mature mRNAs containing various combinations of exons from one original precursor mRNA increases the diversity of the proteins that can be produced. One example of this is in the production of immunoglobulins by B cells. A cell can splice together different exons and produce different immunoglobulins (antibodies) in response to different antigens that are present in the body (Chapter 10).

Splicing is controlled by molecules that respond to signals from both inside and outside the cell. Alternative splicing allows the human to genome to synthesise many more proteins from its 20 000 genes than it could if there was no splicing.

#### KEY POINTS

**alternative splicing** including different exons in processed mRNA so that a cell can produce different proteins from the same gene.

**exons** sequences of bases in mRNA that are spliced together and translated after introns have been removed.

**intron** sequences of bases in mRNA that are removed after transcription.

### 2. 5' Caps and 3' polytails

Two other modifications made to mRNA before translation are:

1. Adding a 5' cap – a 5' cap ( a modified G nucleotide) is attached to the 5' end of the mRNA
2. Adding a 3' poly-A tail – poly-A tail is attached to the 3' end of the mRNA and consists of a long string of A nucleotides

Both these modifications help new mRNA strands leave the nucleus. They also help to protect the mRNA from damage. In the cytoplasm, the addition of the 5' cap and poly-A tail, help the ribosomes attach to the 5' end of the mRNA.

**EXAM TIP**

Remember: introns intervene in genes, but only exons are expressed.

### Producing functional proteins – Post-translational modification

Translation produces polypeptides which are the starting point for forming the working proteins that we need. But proteins are large, complex molecules, usually made up of hundreds of amino acid subunits linked in polypeptides. But it is the folding and linking of polypeptide chains into secondary, tertiary and, in some cases, quaternary structures that leads to the formation of functional proteins (Section 2.4). Folding takes place in the cytoplasm. Once amino acids have been linked together, they will begin to fold, but partially folded proteins can become entangled with other molecules in the cytoplasm. Special chaperone proteins protect folding polypeptides from interacting with other macromolecules (Section 2.4). Chaperones prevent associations with incorrect folding partners by surrounding a polypeptide as it folds and until the folding is finished and the protein completed. Chaperones also assist in refolding proteins that have become unfolded. Chaperone protein refolding is a more energy-efficient process than synthesising a new polypeptide.

Another modification made to polypeptides and proteins after translation is the addition of a prosthetic group. Prosthetic groups are not polypeptides but they bind to different proteins or parts of them to enable them to function. An example is the respiratory pigment hemoglobin, which contains four polypeptide chains, each one containing a prosthetic heme group (Section 2.4).

## 3.2.4 Prosthetic groups

Many proteins contain **prosthetic groups** and those that do are called **conjugated proteins**. Prosthetic groups are non-protein groups that are able to bind to different proteins or parts of them. We can see two examples of prosthetic groups in the respiratory pigments myoglobin (Figure 3.4.4) and hemoglobin which both contain a prosthetic heme group. Hemoglobin consists of four polypeptide chains, each one containing a heme group. The heme group (Figure 3.4.1) consists of a central Fe (iron) atom and a porphyrin ring. The prosthetic heme group is vital to the structure of hemoglobin because the shape of the whole protein is changed as oxygen binds to it. The iron group not only allows oxygen to bind but also holds the compact structure with four subunits in place and allows for progressively easier oxygenation as more oxygen molecules bind to the protein.

## 3.2.5 Chaperone proteins

Chaperone proteins (also known as molecular chaperones) are proteins that help other proteins fold, either during their formation or afterwards, and help to protect them from denaturation. If a protein is partly denatured then chaperone proteins help it to refold. The amount of a chaperone protein present in a cell may increase when a cell is stressed by factors such as heat, salinity or the presence of heavy metals. There are chaperone proteins in all cells, tissues and organs.

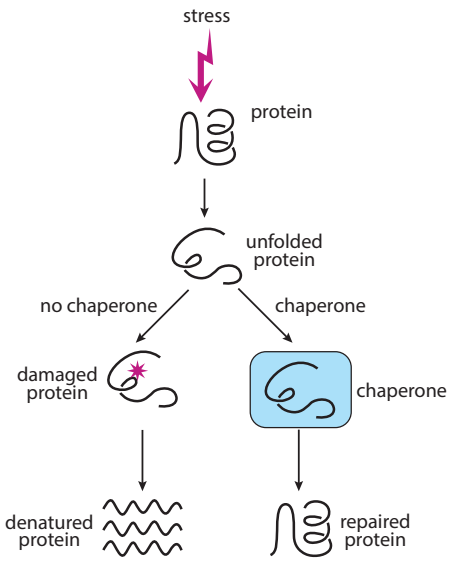



**Figure 3.2.9:** A chaperone protein can assist if a protein becomes unfolded and intervene before the protein is damaged beyond repair.

## Recycling amino acids

Proteins synthesis requires a constant supply of amino acids. Some are provided from our diet but many come from recycled amino acids that have formed part of proteins in the body. Proteasomes are protein complexes which degrade unneeded or damaged proteins by proteolysis, a chemical reaction that breaks peptide bonds. Proteasomes are present in the cytoplasm and in the nuclei of all eukaryotic cells. The amino acids that are released can be used to build new polypeptides and proteins.

### EXTENSION

The proteins present in human cells are far greater than the number that are coded for by our genes and this is due in part to modifications made to the proteins after they are formed. You can find out more about the proteins that are modified using the SWISS-PROT protein sequence database. Recent studies have identified more than 8000 proteins that are phosphorylated, more than 3000 that have acetyl groups added to them and around 5000 that have carbohydrates added to them.

# 3.2.6 Protein transport molecules

Some proteins are able to operate as transport molecules and signalling systems inside cells. The bonds that hold these proteins together can, in some cases, be broken and remade to enable them to do so. For example, when molecular signals are received from outside a cell by receptors on its membrane surface they can be processed and transferred to the nucleus by the modification of transport proteins.

The enzyme protein kinase is important in this process because it can phosphorylate a transport protein. Phosphorylation changes the function of the protein by changing its activity, its location or the way it interacts with other proteins. A protein can be phosphorylated at the cell surface by protein kinase and dephosphorylated later in the cytoplasm. These changes in the protein's structure activate and then deactivate it while it acts as a signalling mechanism because phosphorylated proteins can be moved along many different pathways in the cell.

### KEY POINT

**protein kinase** an enzyme that regulates the biological activity of proteins by phosphorylating specific amino acids using ATP.

### NATURE OF SCIENCE

**Looking for trends and discrepancies: do all organisms use only 20 common amino acids in their proteins?**

Humans can make 10 of the 20 amino acids we need to build proteins but we do not have the enzymes needed for the biosynthesis of the others. Plants, on the other hand, must be able to make all the amino acids they require.

Researchers have also investigated the trends in amino acid compositions of proteins found in species of the important kingdoms of Archaea, Bacteria and Eukaryotes. International databases ProteomicsDB and SWISS-PROT (which contain information about the structure and composition of proteins) can compare amino acid frequencies for 195 known proteomes and all recorded sequences of proteins. They discovered that the amino acid compositions of proteins do differ substantially for different kingdoms.

In addition to the variations in amino acids in proteins, some microorganisms and plants are able to make so called 'non-standard' amino acids by modifying standard amino acids. Some species are also able to synthesise many uncommon amino acids. For example, some microbes synthesise lanthionine, which is a modified version of the amino acid alanine. Many other proteins are modified after they have been produced. This 'post-translational modification' involves the addition of extra side groups to the amino acids in a protein.

Considering all the evidence, it seems that, although we can observe many similar proteins in different species, we cannot always say that the same amino acids are used in their construction. The range of amino acids in proteins can vary considerably from species to species.

## CONTINUED

**To consider:**

1 What contribution have international databases made to our understanding of protein structure?

2 How can comparing proteomes and amino acids in different organisms help our understanding of evolutionary relationships?

## TEST YOUR UNDERSTANDING

**12.** Explain the difference between introns and exons.

**13** Name three types of non-coding DNA sequence.

**14** Where are telomeres located and what is their role?

**15** Why are polysomes important in cells?

## REFLECTION

Reflect on the areas of this topic that you found particularly interesting. What was it about them that caught your attention?

## Links

- How does the variety of proteins produced contribute to the functioning of a cell? (Chapter 6)
- How does the degenerate genetic code protect a cell against mutations? (Chapter 3.3)

# 3.3 Mutations

## LEARNING OBJECTIVES

In this section you will:

- learn that mutations ae structural changes to genes at the molecular level learn how new alleles form by mutation; changes may be neutral, harmful or beneficial
- understand that mutations in germ cells can be passed to offspring
- define a mutagen as a substance that can cause genetic change
- recognise that mutations can add, delete or substitute bases in genes
- recall examples of insertion and deletion mutations learn that the genetic code is degenerate and so it is resistant to some changes caused by mutations

- understand that DNA polymerases can make proofreading errors, and the errors remain permanently
- recognise that mutations do not always cause changes to a protein's function
- understand the technique of 'gene knockout' and its use in investigating gene function
- learn how CRISPR sequences are used in gene editing
- recognise the importance of highly conserved sequences in genes

## GUIDING QUESTIONS

- Why are variation and mutation essential for evolution? (Chapter 11)
- Why must the cell cycle be regulated and controlled? (Chapter 6)

# 3.3.1 Chromosomes, genes and mutations

A DNA molecule comprises a pair of strands, each strand consisting of a linear sequence of nucleotides, with weak hydrogen bonds between the bases holding the two strands together. This linear sequence of bases contains the genetic code in the form of triplets of bases. A gene is a particular section of a DNA strand that, when transcribed and translated, forms a specific polypeptide (Figure 3.3.1). Some of the polypeptides will form structural proteins, while others become enzymes or pigments such as hemoglobin (Section 2.4) and it is the translation of the genes which gives each individual organism its own specific characteristics. (Transcription and translation are described in Section 3.2.) Each gene is found at a specific position on a chromosome and so, for example, it is possible to say that the gene for human insulin is always found on chromosome number 11.

Organisms that reproduce sexually almost always have pairs of chromosomes, with one of each pair coming from each parent. The members of the pair carry equivalent genes, so that – for example – in humans, both versions of chromosome number 11 carry the insulin gene. But there may be slight differences in the version of the gene on each chromosome. These slightly different forms of the gene are known as alleles. Alleles differ from one another by one or only a few bases and it is these differences in alleles that give rise to the variation we observe in living organisms.

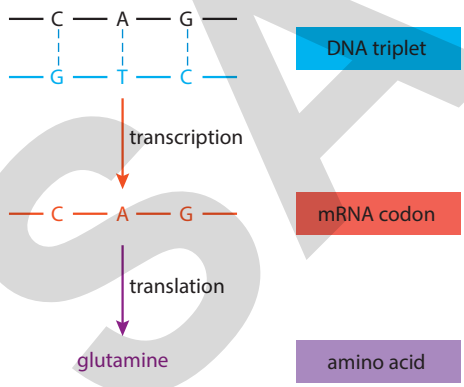



**Figure 3.3.1:** The base sequence in DNA is decoded via transcription and translation.

## What are mutations?

The process of DNA replication is complex and mistakes sometimes occur – a nucleotide may be left out, an extra one may be added or the wrong one inserted. These mistakes are known as gene **mutations**. Mutations may occur spontaneously, as a result of errors in copying DNA, or they can be caused by factors in the environment known as **mutagens**, described in Section 3.3.2. A mutation that occurs in **germ cells** (**gametes**) that will go on to form a new offspring) will be passed to every cell in those offspring, including their germ cells. As a result, the offspring may have a genetic condition that is not present in either of their parents. A mutation that occurs in **somatic cells** (body cells) will not be inherited. A mutation involving the change of a single nucleotide is called a **base substitution mutation**. When the DNA containing an incorrect nucleotide is transcribed and translated, errors may occur in the polypeptide that is produced. Errors may be beneficial, neutral or harmful.

**Beneficial mutations** change DNA and allow the synthesis of new proteins that may work slightly differently. One example of a new, recently discovered, beneficial mutation has been found in a gene that codes for a receptor protein on the cell surface in the plasma membrane. Only a very few people carry this mutation but the change to their receptor protein gives them total immunity to infection by human immunodeficiency virus (HIV) because the virus cannot bind to their cells.

In 2009, United States researchers located another beneficial mutation in a gene (*SLC30A8*), which affects insulin. They found that subjects who were both overweight and elderly who carried the altered gene had considerable protection from developing type II diabetes. Evolutionary biologists also believe that our ability to discriminate three colours – red, green and blue – is due to a beneficial mutation that occurred in our primate ancestors' DNA millions of years ago.

However, beneficial mutations can be closely associated with harmful ones. Sickle-cell anemia, also called sickle-cell disease (SCD), is caused by a mutation that causes red blood cells to develop a crescent, or sickle shape; this abnormal shape can lead to a number of health problems but also has some advantages, which are described in the section on sickle-cell anemia.

## How do mutations occur?

A mutation can involve the addition, deletion or substitution of a base in DNA or the inversion of a section of DNA so that it is turned backwards in the sequence.

Table 3.2.1 shows the amino acids that are specified by different mRNA codons. Most amino acids are coded for by more than one codon and so many substitution mutations have no effect on the final polypeptide that is produced. These are said to be neutral (or silent) mutations. For example, a mutation in the DNA triplet CCA into CCG would change the codon in the mRNA from GGU to GGC but it would still result in the amino acid glycine being placed in a polypeptide. Other examples of neutral mutations are those that affect non-coding regions of the chromosome, or which result in changes to features such as blood type or eye colour in humans that do not adversely affect a person.

Some substitution mutations, however, do have serious effects. For example, one important human condition that results from a single base substitution is sickle-cell anemia.

**EXAM TIP**

You may be asked to explain inversions, deletions, substitutions or additions to a DNA sequence. When you do this, first remember to check whether you need to convert the DNA sequences to mRNA sequences before identifying the amino acids. Most tables used to decode the genetic code are shown as tables of mRNA codons.

## Sickle-cell anemia: the result of a base substitution mutation

Sickle-cell anemia is a blood disorder in which red blood cells become sickle shaped and cannot carry oxygen properly (Figure 3.3.2). It occurs most frequently in people with African ancestry, about 1% suffer from the condition and between 10 and 40% are carriers of it. Sickle-cell anemia is due to a single base substitution mutation in one of the genes that make hemoglobin, the oxygen-carrying pigment in red blood cells.

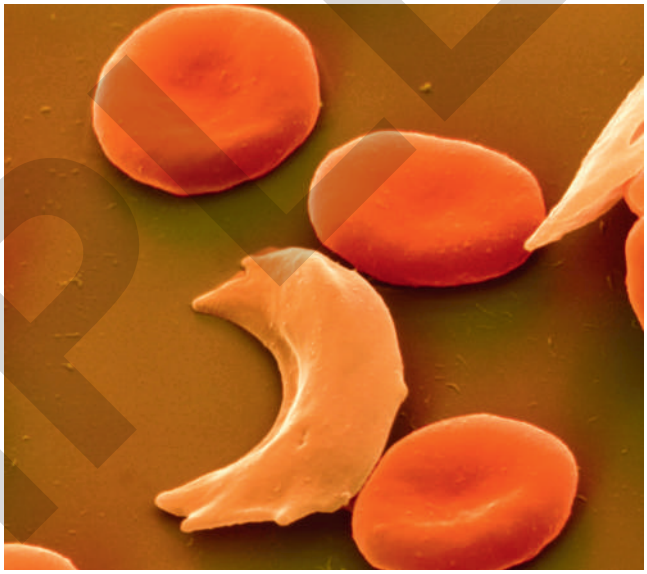


**Figure 3.3.2:** Coloured scanning electron micrograph showing a sickle-cell and normal red blood cells (× 7400).

**WORKED EXAMPLE 3.3.1**

A mutation can involve the addition, deletion or substitution of a base in DNA.
Consider the effect of these changes on this short length of DNA:

CTG GGG GGT GTG AAC

The sequence of amino acids produced by this sequence should be

Leu Gly Gly Val Asn

If the base highlighted in red above is *deleted* the consequence would be:

CTG GGG GGT TGA AC

This would result in the amino acids

Leu Gly Gly STOP

because TGA is a stop codon. The polypeptide produced will be shorter than it should have been.

CONTINUED

**Question**

What type of mutation has occurred in these examples and what are the consequences?

**a** CTG GGG GGT AGT GAA C

**Answer**

In this case a base has been *added* to the sequence after the third codon. This will result in serine, coded for by AGT, being added as the fourth amino acid instead of valine. All the subsequent amino acids in the sequence will also be incorrect. This is known as a frameshift mutation because the 'reading' of the DNA in sets of three bases is changed completely. New amino acids will be inserted and produce a different translation of the code from the inserted bases onwards. Deletion of a base can also cause a frameshift mutation.

**b** CTG GGG GGT GTG CAA

**Answer**

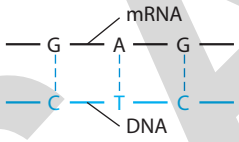
Here the final triplet has been *inverted* (turned around, so it is backwards) and is CAA instead of AAC. In this case glutamine will be inserted into the amino acid chain instead of asparagine.

**c** CTG GGG GGG GTG AAC

**Answer**

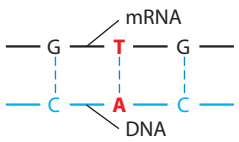
In this example a *substitution* of a base has occurred. G has replaced T. In this case there will be no effect on the amino acid sequence that is assembled because GGT and GGG both code for the amino acid glycine. This is an example of a neutral (or silent) mutation.

Hemoglobin is made up of four subunits, as shown in Figure 3.3.3 – two a-chains and two b-chains. The b-chains are affected by the sickle-cell mutation. To form a normal b-chain, the particular triplet base pairing in the DNA is:




The C–T–C on the coding strand of the DNA (in blue here) is transcribed into the mRNA triplet G–A–G, which in turn is translated to give glutamic acid in the polypeptide chain of the b-chain.

If the sickle-cell mutation occurs, the adenine base A is substituted for thymine base T on the DNA coding strand, so the triplet base pairing becomes:




C–A–C on the coding strand of the DNA is now transcribed into the mRNA triplet G–U–G, which in turn is translated to give the amino acid valine. Valine replaces glutamic acid in the b-chain.

Valine has different properties from glutamic acid and so this single change in the amino acid sequence has very serious effects. The resulting hemoglobin molecule is a different shape and it is less soluble and, when in low oxygen concentrations, it deforms the red blood cells to give them a sickle shape. Sickle cells carry less oxygen, which results in anemia. They are also rapidly removed from the circulation, leading to a lack of red blood cells and other symptoms such as jaundice, kidney problems and enlargement of the spleen.

People who have one sickle-cell allele and one normal allele are said to have sickle-cell trait and have some resistance to malaria. This benefit explains why the mutation persists in areas where malaria is endemic. But, in parts of the world where malaria is not a problem, the mutation no longer provides a survival advantage. Instead, it poses the threat of sickle-cell disease, which occurs in the children of carriers who inherit the sickle-cell gene from both their parents.

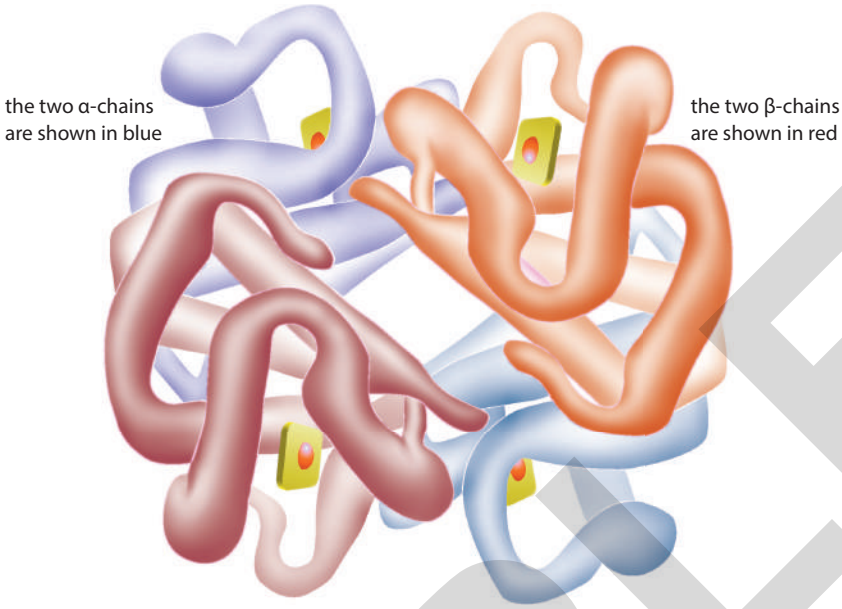



**Figure 3.3.3:** The structure of a hemoglobin molecule showing the 3D arrangement of the subunits that make it up.

## Degeneracy in the genetic code

As Table 3.2.1 shows, there are 64 combinations of three-letter nucleotide sequences that can be made from the four nucleotides. Of these, 61 represent amino acids and three are stop signals. Although each codon is specific for only one amino acid or one stop signal, the genetic code is described as degenerate because a single amino acid may be coded for by more than one codon.

Degeneracy in the genetic code makes the code more resistant to changes. It means that single base substitutions can occur without disrupting protein synthesis or functioning of the organism.

### Consequences of insertions and deletions

Inserted and deleted bases can have serious consequences for protein production in cells. Sometimes changes in bases can cause serious illness but in other cases they lead to benefit. Two examples are the mutations in the *HTT* gene and in the CCR5 gene.

### Huntington's disease: result of insertions - trinucleotide repeats

The *HTT* gene provides instructions for making a protein called huntingtin which plays an important role in nerve cells in the brain. One region of the *HTT* gene contains a particular DNA segment known as a CAG trinucleotide repeat. This segment is made up of a series of three DNA bases (cytosine, adenine, and guanine) that appear multiple times in a row. Normally, the CAG segment is repeated 10 to 35 times within the gene. But if the number of repeats increases to more than 40 the result will be Huntington's disease. If more than 60 repeats are present a more severe form of the disease develops. The greater the number of repeats, the sooner the disease appears.

Huntington's disease stops parts of the brain working properly causing memory loss, difficulty with movement, mood swings and depression. It develops over a period of time and is usually fatal after about 20 years.

### HIV resistance: result of a deletion

The human CCR5 gene located on chromosome 3 codes for a protein called CCR5. The protein is found on surface of lymphocytes and other cells of the immune system. The proteins form part of the receptors that are involved in cell signalling and the coordination of immune responses. The CCR5 receptors provide a point of entry for the HIV-1 virus to infect the cells. Some people have inherited a mutation called Delta 32 and have part of the CCR5 gene deleted. The gene changes alter the structure of the CCR5 receptors and make it difficult for the HIV virus to enter cells. Individuals who have this deletion mutation live normal lives, but if they inherit copies of the mutation from their both parents, they are naturally immune to HIV.

## NATURE OF SCIENCE

### Tests for genetic diseases

There are many commercially available tests for potential health and disease risks. For example, the most effective and accurate method of testing for Huntington's disease is called a direct genetic test. DNA from a blood sample is taken and analysed for the number of CAG repeats it contains. The presence of 36 or more repeats is an indication that the person has, or will develop Huntington's disease. In a few cases, the test result is not clear and a definite answer is not possible. Any person who decides to take this test must be prepared to face not only the emotional effect it will have on themselves and their family, but also the affect on other aspects of their life. Life insurance and some job opportunities may become unavailable to them if the test is positive.

If any genetic test is taken it is important that the outcomes are interpreted correctly by an expert who can explain the consequences of the results. A negative result can eliminate the need for check-ups while a positive result can help direct a person to monitoring and treatment options. Some tests can help people decide about whether or not to have children and both genetic and non-genetic tests can provide information about a person's health in the future.

**To consider:**

1 Why do some people think that genetic test results can cause family discord, psychological distress and stigmatisation?

2 Why might some people decide not to have children as a result of a genetic test?

## Gene knockout to investigate gene function

Gene knockout is a method that is used to damage or 'knock out' specific genes so that they no longer function and are not expressed. Gene knockout is used with model organisms such as mice and yeast which have specific genes knocked out so they can be used to study how those genes function and investigate what happens when the genes are lost. Researchers can draw inferences from the difference between the knockout organism and normal individuals with a similar genetic background. Knockout organisms are also used in the development of new drugs which target specific biological processes or genetic deficiencies. A library of the genomes of model organisms is available to researchers. The loss of gene activity often causes the phenotype of the model organisms such as mice to change so that living organisms can be used to study gene function. For example, 'Metheuselah' is a knockout model mouse which lives for far longer than an average animal and 'Frantic' is a model mouse which is used for studying anxiety disorders. The loss of the knocked out genes provides valuable information about what the gene normally does. Mice are useful model organisms because humans share many genes with mice.

## CRISPR

CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats. Repeated DNA sequences, called CRISPR, were first noticed in bacteria. They have 'spacer' DNA sequences in between the repeats that exactly match sequences found in viruses. These sequences which contain short repetitions of base sequences are involved in the defence mechanisms of prokaryotic organisms such as bacteria and archaea to viruses that infect them. The sequences have come from DNA fragments of viruses that have previously infected the prokaryotes and the organisms use them to detect and destroy DNA from similar viruses that might infect them.

CRISPR is now used as a genetic engineering tool that uses repeated sequences of DNA to edit genes. Using CRISPR it is possible to find a specific section of DNA inside a cell so that a gene can either be modified or even turned on or off without altering the DNA sequences. The key to CRISPR is the many variations of 'Cas' proteins found in bacteria. These are the proteins produced by prokaryotes which help defend them against viruses. A protein called Cas9 is the most widely used by scientists. This protein can easily be programmed to find and bind to almost any desired target sequence, simply by giving it a piece of RNA to guide it.

When the CRISPR Cas9 protein is added to a cell along with a piece of guide RNA, the Cas9 protein links to the guide RNA and then moves along the strands of DNA until it finds and binds to a 20-DNA-letter long sequence that matches part of the guide RNA sequence.

One successful use of CRISPR has been in the treatment of human papillomavirus (HPV). This very common virus has more than 100 different strains; some of them affect the skin causing warts and are barely noticed, but others contribute to 99% of cervical cancers. Researchers have been able to turn off two genes in the virus and knock out the production of two viral oncoproteins. A constant supply of these viral proteins is needed to transform normal cells into cancer cells. Without the proteins, cells infected with the virus go into senescence which means that they stop dividing. Targeting the genes for the proteins can potentially treat HPV-related cancers. CRISPR and Cas9 has also been used in the treatment of hepatitis B, in this case the ends of certain repeated sequences in the Hepatitis B viral genome are targeted. It has also been used experimentally to repair the mutations that cause cataracts in mice.

## NATURE OF SCIENCE

Since the human genome was first sequenced, genetic research has expanded rapidly. The cost of sequencing the entire genome of one person has dropped from about US$1 billion to $1,000, and the speed of sequencing has become many times faster. Scientists around the world have a new and powerful way of understanding how genetic variation may affect not only organisms that humans make use of, but also human health and disease. CRISPR technology can be used to edit genes and has the potential to change genomes.

The CRISPR method was first used in 2012 and replaced costly methods of gene editing that had been used in some plants and animals previously. CRISPR has made gene editing cheap and easy. The technique is widely used in research and already has the potential to alter plants and animals on our farms. The technology also has the potential to treat and prevent many diseases.

## CONTINUED

It could even change the genomes of future generations, although many people think this is unethical. CRISPR is already being used to fingerprint cells and observing what happens inside them, and for directing evolution.

The most common use of CRISPR involves using Cas9 protein to cut the DNA at a target area. When the cut is repaired, mutations can be introduced that disable a gene. But CRISPR can also be used to make precise changes such as replacing faulty genes. At present this is much more difficult. The knowledge gained from studying human gene knockouts gives scientists a tool to identify potential new targets for medical treatments and better understand safety concerns of treatments that are being developed. But the technique could have the potential to make permanent changes to a person's genome. Scientists around the world are subject to different rules in the use of genome technology. For this reason, there is an ongoing effort to make a system of regulation for all scientists working in this rapidly growing field of research.

### Questions

1 Why is it important to have clear rules about what should and should not be attempted using technology such as CRISPR?

2 What are the potential benefits and dangers in gene editing and replacement?

## Why are some gene sequences conserved in a species?

## KEY POINTS

**Conserved sequence** a base sequence in a DNA molecule (or an amino acid sequence in a protein) that has remained relatively unchanged throughout evolution.

**Conserved sequences** are sequences of DNA (or protein) that are identical or very similar across a species or group of species. A highly conserved sequence is

one that has remained relatively unchanged for long periods of evolutionary history. Examples of highly conserved sequences include those for the RNA found in ribosomes that are present in all domains of life, and the homeobox sequence which is a DNA sequence of around 180 base pairs that regulates the early stages of embryonic development. This sequence is found in many eukaryotes. Studies of sequence conservation are now form part of investigations in genomics, proteomics and evolutionary biology.

An explanation for the presence of these conserved sequences was put forward in 1965, by Emile Zuckerandl and Linus Pauling who proposed the hypothesis of a molecular clock. Amino acid sequences can be conserved to maintain the structure or function of a protein. Conserved proteins undergo fewer amino acid replacements or are more likely to substitute amino acids with similar biochemical properties. The molecular clock theory went on to suggest that steady rates of amino acid replacement could be used to estimate the time when two organisms diverged in evolution (cross ref ). Many phylogenetic relationships worked out from studies of the fossil record seemed to support this theory but some other genes were found to evolve at different rates. This led to the development of theories of molecular evolution. In 1966 Margaret Dayhoff compared ferrodoxin (small proteins involved in a range of metabolic reactions) sequences in many organisms and proposed that natural selection will act to conserve protein sequences that are essential to life. This hypothesis explains why conserved sequences of DNA are found for many important proteins in many species.

## 3.3.2 Harmful mutations and mutagens

New cells are needed to replace cells that have died or to allow an organism to grow. The nucleus and cytoplasm of a cell divide in processes known as mitosis and cytokinesis, which are phases in a series of events known as the cell cycle (Section 6.5).

In normal circumstances the cell cycle is strictly controlled with cell division (mitosis) occurring to form new cells to replace damaged or dying cells.

In most cases, mitosis continues until a tissue has grown sufficiently or repairs have been made to damaged areas.

Most normal cells also undergo a programmed form of death known as **apoptosis** as tissues develop. Apoptosis can be caused when a cell experiences stress or if it receives signals that indicate it should die.

But sometimes mitosis does not proceed normally. Cell division may continue unchecked and produce an excess of cells, which clump together. This growth is called a **tumour**. Tumours can be either **benign**, which means they are restricted to that tissue or organ, or **malignant** (cancerous), in which some of the abnormal cells migrate to other tissues or organs and continue to grow further tumours there. If they are allowed to grow without treatment, tumours can cause obstructions in organs or tissues and interfere with their functions.

Mutagens are physical, chemical or biological agents that can cause mutations and modify DNA. Mutagens include ionising radiation – such as X-rays, gamma rays and ultraviolet light – and also chemical compounds, such as those found in tobacco smoke and aflatoxins produced by certain fungi. The DNA changes caused by mutagens are not all harmful. However, because some of them cause cancer, some mutagens are said to be **carcinogens** (cancer causing). The development of a primary tumour can also be caused by mistakes in copying DNA, or a genetic predisposition as a result of inheritance. (You can learn more about the control of cell division and the development of cancer in Section 6.5)

### Smoking and cancer

Smoking is a major cause of several types of cancer. There is strong evidence to show that it increases the risk of cancer of the bladder, cervix, kidney, larynx and stomach, and smokers are seven times more likely to die of these cancers than non-smokers. In the UK, approximately 70% of lung cancers in both males and females are related to smoking.

## SCIENCE IN CONTEXT

### Smoking and lung cancer

All tobacco products contain various amounts of carcinogenic substances. Tobacco smoke contains more than 70 chemicals, including many which are known to initiate or promote cancer. Recently the role of nicotine, the addictive drug in tobacco, has come under scrutiny as more people are turning to e-cigarettes and other non-tobacco sources of nicotine as substitutes for smoking. There is no clear evidence that nicotine is a direct carcinogen, but it seems to act as a promoter and may inhibit anti-tumour immune responses. In experiments, nicotine has been shown to induce breaks in DNA and enhance the growth of existing cancers.

The link between smoking and lung cancer was recognised in the 1940s and 1950s, with evidence from epidemiology, animal experiments, examination of cells and chemical analysis. Cigarette manufacturers disputed the evidence, as part of a campaign to maintain cigarette sales. Their propaganda was successful in the short term and, as late as 1960, only one-third of all US doctors believed that the case against cigarettes had been established.

Today it is understood that the risk of contracting lung cancer increases with the number of cigarettes that a person smokes and the number of years that they continue to smoke. If a person gives up smoking, their risk of developing cancer decreases (Figure 3.3.4).

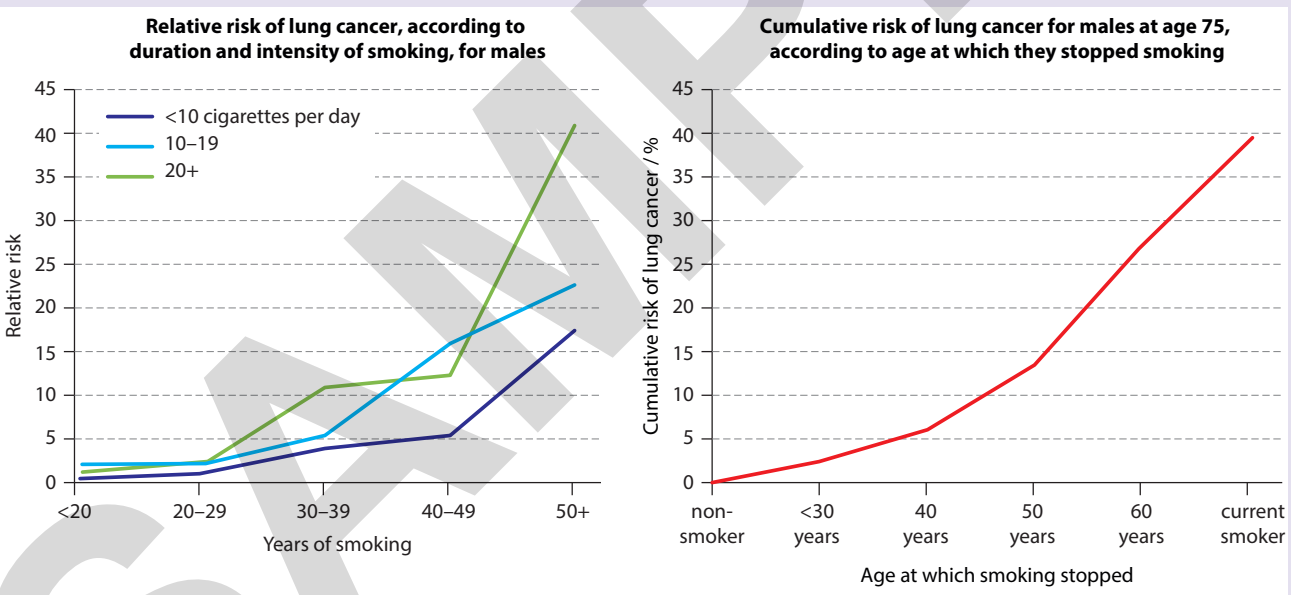



**Figure 3.3.4:** Graphs to show the relationship between smoking and lung cancer, and the cumulative risk of lung cancer among men in the UK at age 75 according to the age at which they stopped smoking (data from Cancer Research UK).

## CONTINUED

Lung cancer develops slowly and it takes years before the effects of the carcinogens become obvious. The number of males who suffered lung cancer in the UK was at its highest levels in the early 1970s. This was as a result of a peak in smoking 20–30 years earlier. Cancer in females increased through the 1970s and 1980s because more females smoked in the 1950s and 1960s. Statisticians predicted that cancer in females would increase to reach the same levels as those in males over the next decade. New government education campaigns have persuaded people give up smoking and new laws have limited smoking in public places, and so the number of deaths have started to decrease. Figure 3.3.5 shows the incidence of lung cancer in the UK 1978–2017 and the incidence of smoking since 1948. As had been predicted, the rate for females increased by 15% but for males rates decreased by 11%.

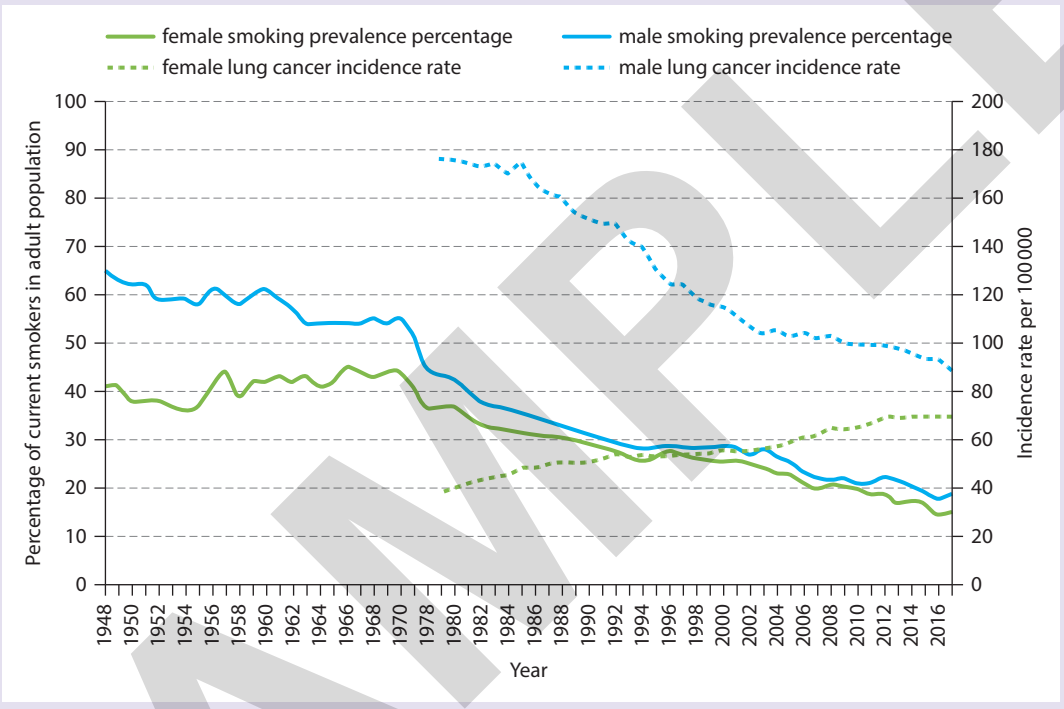



**Figure 3.3.5:** The incidence of smoking in UK males and females 1948–2017 and the incidence of lung cancer 1978–2017.

**To consider:**

- Lung cancer affects the economic performance of a country in health care costs and loss of production when people are unwell. Why do you think that governments and officials were reluctant to accept the link between smoking and health in the 1950s?
- How important were laws and restrictions in persuading people to modify their smoking behaviour?

# 3.4 Gene expression and epigenetics

## LEARNING OBJECTIVES

In this section you will:

- define gene expression as the mechanism in which genetic information affects the phenotype
- define epigenetics as the study of changes to gene activation in differentiated cells
- learn that gene expression is regulated by proteins that bind to base sequences in DNA and that degradation of mRNA can regulate translation
- understand that epigenetic changes modify the activation of certain genes but do not change their base sequences so that the phenotype will change but genotype does not
- learn that DNA methylation inhibits transcription
- understand that epigenetic changes are faster than changes caused by natural selection
- recognise that environment has an impact on gene expression and can trigger heritable changes in epigenetic factors
- learn that most epigenetic tags in gametes are removed from the embryo genome, but some remain and are inherited leading to the appearance of phenotypic differences in organisms such as lion-tiger hybrids
- learn that external factors such as hormones and nutrients can affect the pattern of gene expression
- understand how environmental effects on DNA methylation can be studied using monozygotic twins.

## GUIDING QUESTIONS

- How is DNA modified to influence gene expression?
- How is differentiation of cells brought about by epigenetics?

## 3.4.1 Epigenetics and gene expression

As our genes are transcribed and translated and proteins are built from the information they carry, an organism's appearance, enzymes and metabolism are all controlled and decided. **Gene expression** is the mechanism by which information carried in DNA has its effects on the phenotype of an organism. The stages of this process include transcription, translation and the function of the protein that is produced. Gene expression is regulated and controlled so that the genes that are expressed match the needs of the organism. Transcription can be regulated by proteins that binds to specific base sequences; these may be promoters, enhancers or transcription factors which either allow or prevent transcription of a gene. Translation can be regulated by the length of time that mRNA is present in a cell and this too is controlled in the cytoplasm. In addition, epigenetic changes can influence patterns of development and differentiation of cells without changing the genotype of the cell.

## KEY POINTS

**gene** a particular section of a DNA strand that codes for a specific polypeptide; a heritable factor that controls a specific characteristic.

**Gene expression** the mechanism by which genetic information affects the phenotype of an organism

**Genome** the entire set of DNA instructions found in a cell

**Proteome** the complete set of proteins expressed by an organism

**Transcriptome** all the mRNA molecules expressed from the genes of an organism

## Regulation of transcription by proteins that bind to DNA

### Gene expression and binding proteins

Before transcription can begin and mRNA production start, RNA polymerase requires the presence of a class of proteins known as general transcription factors. Transcription factors (TFs) are regulatory proteins whose function is to activate transcription of DNA by binding to specific DNA sequences. TFs specifically bind to target sequences which are highly conserved. These sequence specific transcription factors are probably the most important mechanism of gene regulation in cells. In eukaryotes gene expression requires the co-ordinated interaction of several of these proteins. Interactions between the transcription factors, RNA polymerase and the promoter region of the DNA molecule allow the RNA polymerase to attach and move along a gene so that transcription can occur. There will be greater transcription of certain genes when specific transcription factors are present so that the proteins that a cell needs can be assembled. Many different transcription factors have been found and each one is able to recognise and bind to a specific nucleotide sequence in DNA.

KEY POINT

**transcription factors (TFs)** proteins that bind to particular sites on DNA and activate transcription. Together with RNA polymerase and other proteins (activators), TFs form the transcription apparatus and have a key role in regulating genes.

**Promoter** a region of DNA to which proteins (RNApolymerase and TFs) bind to initiate transcription of a gene

The role of activators is shown in Figure 3.2.3. These proteins bind to a region of the DNA called the **enhancer**, which may be some distance from the gene. 'Bending proteins' may then assist in bending the DNA, so that the enhancer region is brought close to the promoter. Activators, transcription factors and other proteins attach, so that a 'transcription-initiation complex' is formed and transcription can begin.

Transcription factors are regulated by signals produced from other molecules. For example, hormones are able to activate transcription factors and thus control transcription of certain genes. Many other molecules in the environment of a cell or an organism can also have an impact on gene expression and protein production.

### Regulation of translation by mRNA degradation

Once mRNA has reached the cytoplasm, translation can be regulated by mRNA persistence (length of time it is present). mRNA is degraded by nuclease enzymes and the degradation of mRNA and the efficiency with which it is translated are another essential stage in determining gene expression. Individual mRNA molecules can exist in an active state, a silent state or a state that is targeted for decay. In general, RNA is degraded at the end of its useful life, which is very short for introns and spacer fragments, but longer for other sections of mRNA. The time varies between a few minutes to a few days. RNA molecules with defects in processing or assembly are rapidly identified and degraded by the nuclease enzymes. mRNA lifespan can be shortened if translation is incorrect affected or made more stable if translation elongation or termination ae inhibited. mRNA lifespan can also be altered in response to developmental, environmental and metabolic signals.

### RNA silencing

RNA silencing is one method of gene silencing that includes several pathways to control and regulate gene expression. Small non-coding strands of RNA (such as microRNAs and RNAi) may either block sections of transcribed mRNA by pairing with it, or degrade the mRNA in the cytoplasm. In both cases the mRNA is not translated and both methods therefore can prevent the translation of some genes. RNA silencing is also used to silence genes in research into the production of medicines to combat cancer and other diseases. This is because RNA silencing is used in the cells of most organisms to fight RNA viruses which are destroyed in the cytoplasm after transcription.

KEY POINT

**gene silencing** interruption or suppression of the expression of a gene either at transcription or translation

Epigenetics is a relatively new area of investigation in biology. It is the study of how the expression of DNA can be changed without changing the structure of DNA itself. The phenotype (characteristics) of an organism may be changed but its genotype (sequences of DNA) remain the same. Epigenetic changes can affect how cells read their genetic code and a few of the changes can be passed on to the next generation.

## Environment and gene expression

The expression of genes can be influenced by the environment - not only the organism's external

environment, but also its internal environment, which is affected by chemicals such as hormones and various products of metabolism. Temperature, light and chemicals are just some of the environmental factors that can cause some genes to be turned on or off and influence how an organism functions or develops.

## 3.4.2 Epigenetic changes

The activity of genes can be influenced by certain DNA modifications that do not change an organism's DNA sequence but which do affect which genes are active and which are not. These changes to gene activity will influence the phenotype of an individual but not its genotype. Chemical compounds attached to single genes can produce modifications known as epigenetic changes. When chemical compounds are attached to the genome, it is referred to as an epigenome. The additions or 'tags' are not part of the DNA sequence, but remain attached to DNA as cells divide and, in some cases, can be passed on to offspring. One group of chemical tags are methyl groups which attach to the DNA base cytosine when it is followed by a guanine, and it is their location that influences the expression of the associated gene. Tagging patterns vary from one cell to the next.

### KEY POINTS

**epigenetics** the study of changes to gene activation in differentiated cells.

**epigenome** all the chemical compounds that have been added to a genome to regulate the expression of all the genes within the genome.

All cells in an organism contain the same DNA, but the many different cell types function differently, so that muscle cells, for example, produce different proteins from cells of the intestine. Epigenetic changes help to determine whether genes are turned on or off and determine which proteins are transcribed in each cell. Cells are different because some of their genes are turned off, while others are turned on. Epigenetic silencing turns genes off so that only necessary proteins are produced and enable different cells to behave differently. Environmental influences, such as diet and exposure to pollutants, can also affect the epigenome. (See Science in Context, Nutrition and epigenetics). Here we discuss the three most important types of epigenetic change: DNA methylation, histone modification and RNA silencing.

### SCIENCE IN CONTEXT

#### Nutrition and epigenetics

We know that an organism's development is influenced by genes being switched on or off at specific times and there has been much debate about how environmental factors can lead to such epigenetic modifications. The environment's effects can influence human health, and some of these effects can be inherited.

In the early 21st century, Swedish scientists investigated whether nutrition affected the death rate associated with cardiovascular disease and diabetes in a number of Swedish families. They were interested to learn whether the effects were passed from parents to their children and grandchildren.

Researchers examined records of harvests and food prices in Sweden from the 1890s onwards and the medical records of three generations of the families. Their studies revealed that if a father did not have sufficient food in his pre-pubescent years, his sons were less likely to suffer from cardiovascular disease. For diabetes the picture was different: the children's death rate due to diabetes was unaffected if their father had had a plentiful supply of food at the same critical period. But if the children's grandfather had been well nourished at the same critical period of his life, the incidence of diabetes in his grandchildren was increased.

This suggests that diet can cause epigenetic changes to human genes that affect likelihood of disease. Furthermore, the changes have been passed on through males in a family in a similar way to the inheritance of coat colour in mice described in section 3.4.3.

**To consider:**

1 Why do you think that researchers used records from the 1890s in their investigations?

2 What other factors might be important in the development cardiovascular disease and diabetes?

## DNA methylation

DNA methylation is the most common type of epigenetic modification. It involves attaching methyl groups, consisting of one carbon atom and three hydrogen atoms, to segments of DNA (Figure 3.4.1). When methyl groups are added to a particular gene, that gene is turned off or silenced, and no protein is produced from it.

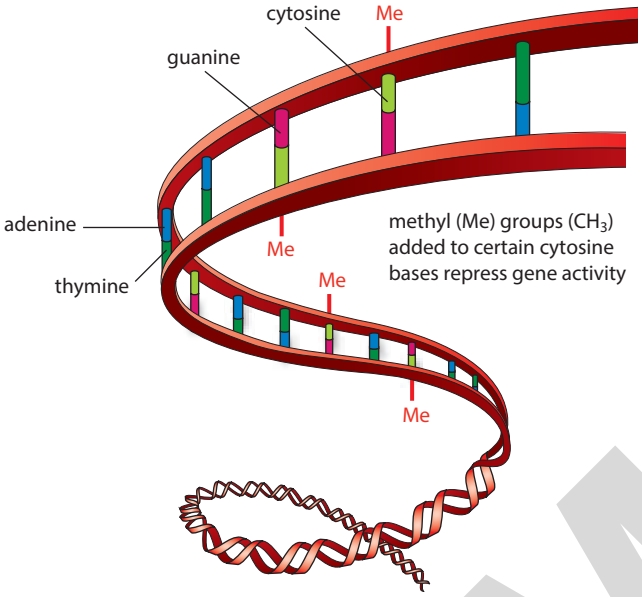



**Figure 3.4.1:** Methylation of DNA is one epigenetic factor affecting gene expression.

**KEY POINT**

**DNA methylation** a process that adds a methyl group to the DNA base cytosine. It is an example of an epigenetic marker.

DNA methylation always happens in a region known as a CpG site where a cytosine nucleotide is located next to a guanine nucleotide that is linked by a phosphate. The enzyme DNA methyl transferase adds methylation markers to the base cytosine. Inserting methyl groups here changes the appearance and structure of DNA, and modifies the interactions between transcription factors that determine whether the gene will be expressed. Promoter regions of genes often lie within areas known as 'CpG islands' and if CpG is methylated, the gene will not be expressed.

Special binding proteins can recognise and 'read' these epigenetic markers. If binding protein is missing or a mutation occurs in it, genes which should not be expressed will transcribed.

**EXTENSION**

Rett syndrome, a neurological disorder which affects brain development in girls, has been linked to the absence of binding proteins which recognise methylated markers in DNA.

## Histone modifications

DNA in eukaryotes is packaged around histones and incorporated into nucleosomes so that the genetic material can be stored in a compact form (Figure 3.4.2) known as chromatin (Section 3.1). In order to transcribe genes, enzymes involved in transcription must be able to gain access to DNA. In all eukaryotes, the regions of DNA that contain promoters and regulators have fewer nucleosomes than other areas, allowing greater access for binding proteins, while regions that are transcribed have a higher density of nucleosomes. Nucleosomes have an important role in determining which genes are transcribed and can influence cell variation and development.

**KEY POINT**

**chromatin** an association of histone proteins and DNA which help to package DNA in a compact form in the cell nucleus.

DNA does not need to be completely released from a nucleosome to be transcribed. Nucleosomes are very stable protein–DNA complexes but they are not static. They can undergo structural rearrangements, including 'nucleosome sliding' and DNA site exposure. Nucleosomes are important because they can either inhibit or allow transcription by controlling whether binding proteins can access DNA.

Histones can be modified so that they influence the arrangement of chromatin about the chromosome and thus also DNA transcription. If chromatin is compact (condensed) it will prevent DNA transcription but if it is loose (active) DNA can be transcribed. Histones can either be methylated or acetylated by the addition of a methyl or acetyl group to the amino acid lysine in the histone. Acetylation produces active, less condensed chromatin so that proteins involved in transcription can bind to DNA and a gene can be transcribed. Histone methylation can indicate either active or inactive regions of chromatin (Figure 3.4.3).

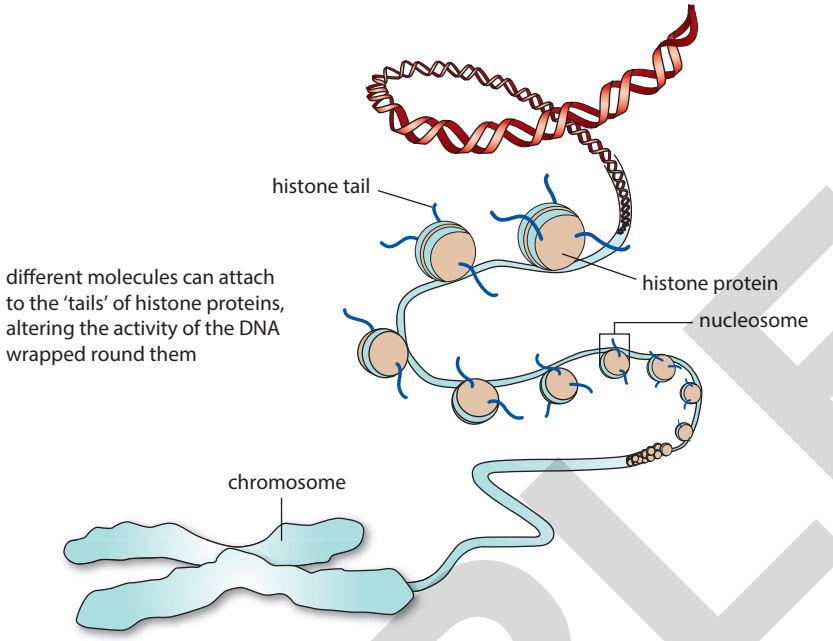



**Figure 3.4.2:** Histone modification is another epigenetic factor affecting gene expression.

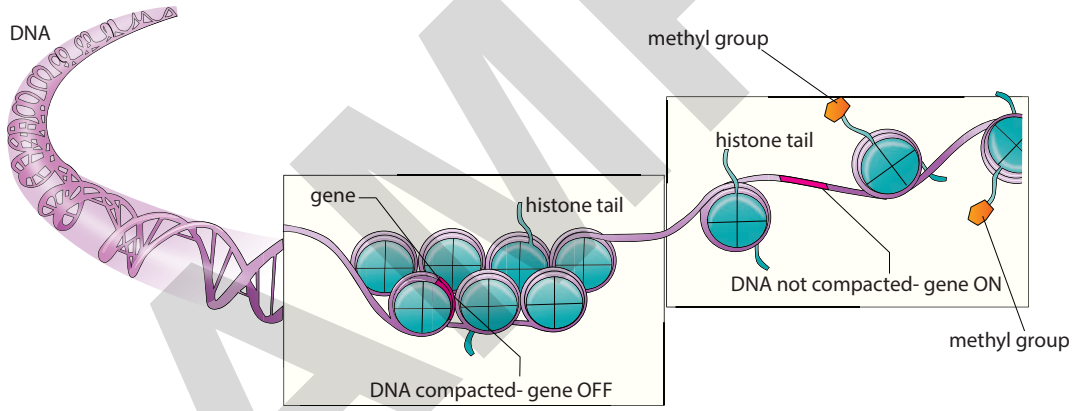



**Figure 3.4.3:** Effect of methylation on chromatin.

Histone methylation and deacetylation are also important in female mammals who have two X chromosomes. One of the two chromosomes is inactivated so that females do not produce twice as many X-chromosome gene products as males.

## 3.4.3 Epigenetic markers and offspring

The genes in egg and sperm cells from the same species contain different epigenetic markers, which cause them to be differentially expressed in the zygote (the cell produced by the fusion two gametes) and developing embryo. These genes are known as imprinted genes. The expression of these genes depends upon which parent contributed them. Most epigenetic markers (tags) found in egg and sperm cells are removed from the epigenome of the embryo and so are not inherited. But in mammals a few do remain and are passed on. One example of this has been investigated in agouti mice. The *Avy* (agouti variable yellow) gene in mice influences the animal's coat colour. Research into the *Avy* gene established that coat colour was related to the degree of methylation of the gene. A high degree of methylation inactivates the gene so the mouse has a dark coat. Without methylation, the gene is

active and the coat is yellow. An active gene is also linked to an increased likelihood of obesity and diabetes. Later research showed that if pregnant mice were fed with increased the levels of methylated molecules, such as folic acid and zinc, DNA methylation at the agouti locus was found in their offspring. Baby mice were born with darker coats and leaner bodies (Figure 3.4.4). Furthermore, the mother's diet affected not only characteristics of her own offspring, but also of her daughters' offspring in the next generation.

Scientists had previously believed that methylation markers were always removed from DNA as sperm and egg cells were produced but these experiments suggest that markers must remain, for at least some genes.

Another example of imprinted genes affecting phenotype can be seen in the cat family. Lions and tigers do not normally meet in nature, but in captivity they may mate and sometimes produce hybrid offspring. The offspring look different, depending on which animal is the mother. A male lion and a female tiger produce a liger – the largest of the big cats. A male tiger and a female lion produce a tigon, a cat that is about the same size as its parents. The difference in size and appearance between ligers and tigons is due in part to the parents' differently imprinted genes.

## Genomic imprinting

Even though both parents contribute equally to the genetic content of their offspring, genomic imprinting sometimes leads to expression of certain genes from only one parent. A marker or imprint can affect particular genes on the maternal and paternal chromosomes in such a way that only one copy of those genes is expressed in the offspring.

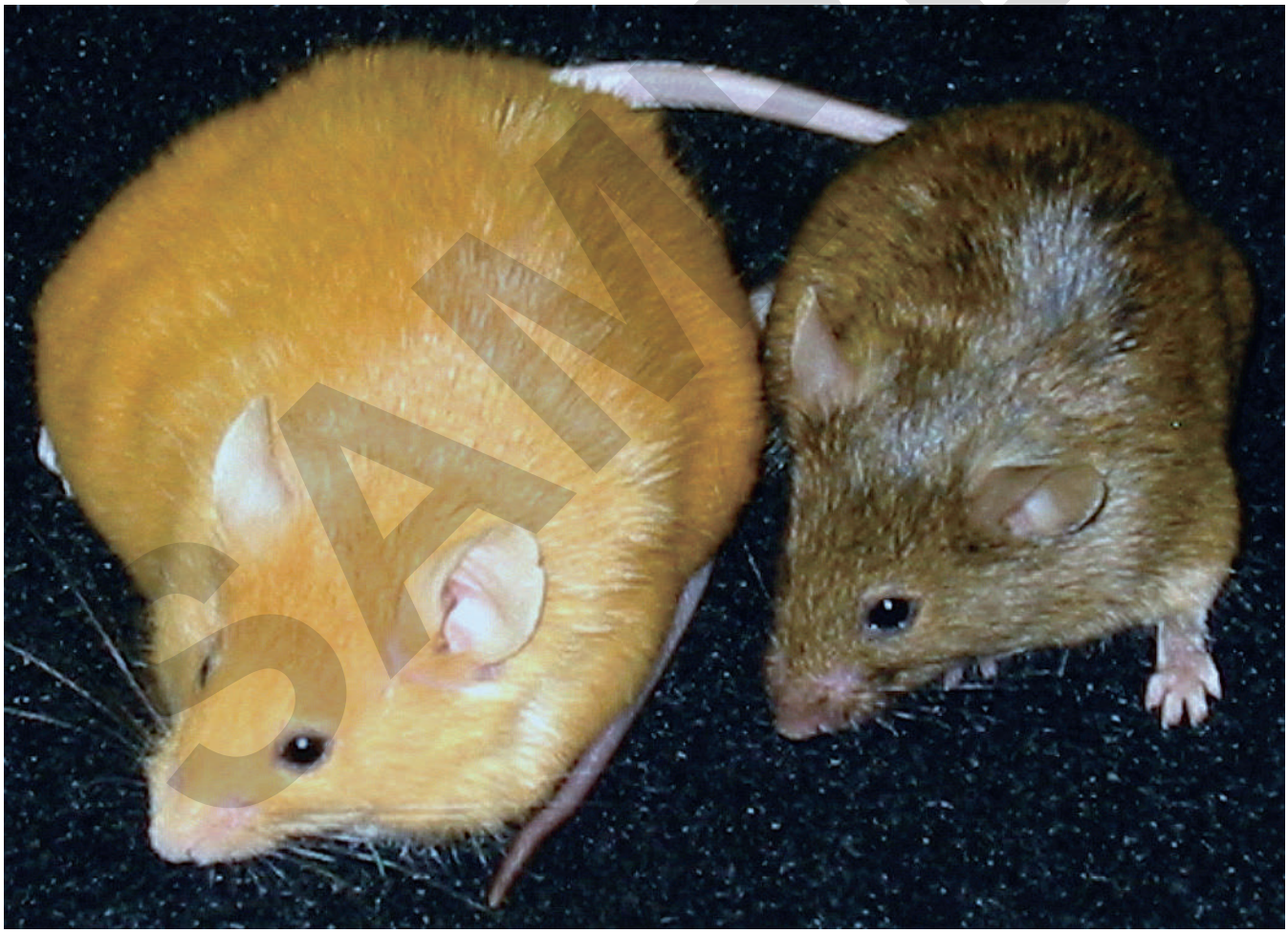


**Figure 3.4.4:** These mice are genetically identical and the same age. The mother of the left-hand mouse received a normal 'mouse diet' during pregnancy, while the mother of the mouse on the right was fed supplements including folic acid.

KEY POINT

**genomic imprinting** inheritance that is not controlled in a Mendelian way. Genes are silenced through DNA methylation but the pattern of gene expression is different depending on whether the gene comes from the father or mother.

Prader–Willi syndrome and Angelman syndrome are human genetic disorders caused by the same mutation on chromosome 15. Prader–Willi syndrome is a disorder that causes behavioural and cognitive problems, deficiencies in sexual development and obesity. It occurs when a mutation is inherited from the child's father and the gene from the mother is imprinted or silenced. Angelman syndrome occurs when the mutated gene from the mother is active. Although the same mutation is involved, Angelman syndrome causes developmental problems, sleep disorders and hyperactivity, but people with the condition have a normal life expectancy and laugh readily.

## 3.4.4 Rate of epigenetic change

Epigenetics shows us that gene expression can change in a more complex way than simple changes to the DNA sequence. If epigenetic changes occur in sperm or egg cells, the changes are inherited by offspring, as in the case of the *Avy* gene in mice.

Epigenetic changes occur more rapidly than changes due to natural selection (Section 11.2). The rate of changes, such as DNA methylation, is much higher than rates of mutations transmitted genetically and they are also easily reversed. This provides a way for variation within a species to increase quickly, especially if the environment is rapidly changing. This will be the case for both epigenetic effects within a single generation, as well as those which persist into the next generation. Epigenetic changes may also create new heritable variation that will enable organisms to adapt over a longer period of time. If a population is very small and has little genetic variation then epigenetic variation can help organisms adapt to new or changing environments. Many epigenetic effects are caused by the environment (refer to Section 3.4.5) and influence phenotypes in many ways. If the environment is changing slowly, the environment of a parent may serve to predict what their offspring may encounter.

Recent data also suggest that epigenetic patterns may change during the course of life, so that key genes in vital processes may be affected with age.

## 3.4.5 Pollution, methyl tags and twin studies

Imprinted genes are very sensitive to environmental signals because they have only a single active copy so any epigenetic changes will have a greater impact on gene expression.

Environmental signals can also affect the imprinting process itself. Imprinting happens during egg and sperm formation, when epigenetic tags are added to silence specific genes. Diet, hormones and toxins can all affect this process and the expression of genes in the next generation.

Chemicals in the environment are being linked to many processes that affect DNA methylation and histone modification. Investigations have identified a range of environmental chemicals that affect epigenetic markers and these include the metals cadmium, arsenic and nickel, and air pollutants including particulates and benzene.

Air pollution from car exhausts is an important cause of breathing problems and other respiratory disorders. High levels of the tiny particles (less than 2.5 μm in diameter) found in exhaust fumes not only irritate the lungs but can also enter the bloodstream and cause inflammation in the body. Inhaling these fine particles has been linked to DNA methylation in the T-helper cells of the immune system. These cells have a key role in our response to inflammation.

Monozygotic (identical) twins have the same genomes as they develop from a single fertilised ovum. But in many cases monozygotic twins are not identical in their phenotypes and in the diseases and conditions that affect them. When their genomes are investigated, results show that there are differences in methylation patterns in the twins' DNA. In studies, twins have shown substantial differences in the occurrence of schizophrenia, autism and diabetes, and these differences have been show to be related to methylation patterns of the twins' DNA. Other common diseases that may also be linked to epigenetic effects include heart disease and cancers, which are often influenced by environmental factors such as pollution.

Disease-discordant identical twins make good subjects for studying differences in disease linked to pollution and methylation patterns because their genes are matched and many non-genetic effects, such as their early environment, maternal influences and age, are also the same.

## THEORY OF KNOWLEDGE

### Nature or nurture?

Studies of identical and fraternal twins are used to separate the influences of genes and the environment on particular characteristics. If a characteristic is more common in identical twins than fraternal twins, it is likely that genetic factors are at least partly responsible. This is because identical twins have the same genes, whereas fraternal twins are likely to share only 50% of their genes. Twin studies allow scientists to study the influence of 'nature versus nurture', a phrase that was first used by British scientist Francis Galton. Galton came to realise how important studying twins could be and in 1875 he wrote 'The History of Twins' and tried to quantify the relative effects of nature versus nature on human intelligence. He believed strongly that intelligence is largely inherited and so, to improve humanity, the ablest and healthiest people should be encouraged to have more children. But his ideas were used by eugenicists who took them further and proposed that the human species could be improved by preventing the least able or those with 'undesirable' characteristics from having children at all.

In the early 21st century, Eric Turkheimer, a United States professor of psychology, looked again at the inheritance of IQ and studies involving twins. He noticed that most of the studies that reported that IQ is an inherited characteristic involved twins from affluent homes. When he looked at twins from low-income families, he found that the IQ of identical twins varied just as much as the IQ of fraternal twins. From his research he deduced that income can affect a child's natural intelligence. More recently, a much larger study showed that the relationship between income, genetics and IQ is not straightforward and there are many more variables and influencing factors.

Today twin studies are being used to investigate a range of factors including eating disorders, obesity and sexual orientation.

**To consider:**

1. To what extent should science be used to provide evidence for complex human characteristics such as intelligence?
2. How difficult is it to assess a person's intelligence?
3. How are such assessments influenced by the background and preconceptions of the assessor?

# 3.4.6 External factors affecting the pattern of gene expression

As well as factors such as TFs inside the nucleus, mRNA degradation and epigenetic factors, many external factors also influence the pattern of gene expression. These include hormones and chemicals which are present in the external environment of the cell and influence what happens inside it.

**Hormones** are factors produced from outside a cell that can affect the cell's gene expression. Steroid hormones such as estrogen influence the transcription of many genes because they interact with receptors inside cells. Estrogen binds to estrogen receptors in the plasma membrane and from there activates signalling pathways. Steroid hormones pass through the plasma membrane of a target cell and bind to intracellular receptors in the cytoplasm or in the nucleus. The cell signalling pathways induced by the steroid hormones regulate specific genes in the cell's DNA. The hormones and receptor complex act as transcription regulators by binding to the promoter region of the gene, this stimulates RNA polymerase binding and gene transcription and thus increasing or decreasing the synthesis of mRNA molecules of specific genes. This, in turn, determines the amount of corresponding protein that is synthesized. (You can read more about the details hormone activity in section 7.3.1)

In bacteria **biochemical factors** affect gene expression. One example is lactose which affects the expression of genes needed for lactose metabolism in *E coli* bacteria. The *lac* operon of *E. coli* contains genes involved in lactose metabolism. The bacteria are able to break down lactose, but if glucose is present, they prefer to use it as an energy source. Glucose metabolism involves fewer steps and needs less energy to metabolise. But if lactose is the only sugar available, the *E. coli* will use it instead. The lac operon is only expressed if lactose is present, and glucose is absent.

Two regulators turn the operon on and off in response to lactose and glucose levels. The *lac* repressor acts as a lactose sensor. It usually blocks transcription of the operon but stops acting as a repressor when lactose

is present. The other regulator is catabolite activator protein (CAP), which acts as a glucose sensor. CAP will bind to specific DNA sites in or near promoter regions and enhance the ability of RNA polymerase to bind and initiate transcription. It activates transcription of the operon, but only when glucose levels are low. (Fig 3.4.5)

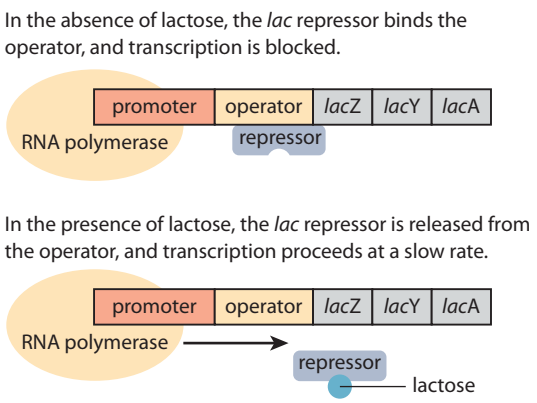



**Figure 3.4.5**

## TEST YOUR UNDERSTANDING

**19** Define the term epigenetics.

**20** Outline how DNA methylation affects activation of genes.

**21** Suggest what effect pollution may have on DNA methylation.

**22** Give an example of an epigenetic tag that is not removed from an embryo's epigenome.

**23** Why are monozygotic twins useful in epigenetic studies?

## REFLECTION

Epigenetics is a new area of scientific research. How much did you know about it before working on this section?

## SELF-ASSESSMENT CHECKLIST

Think about the topics covered in this section. Which parts are you most confident with? Which topics require some extra practice?

| I can... | Subsection | Needs more work | Almost there | Confident to move on |
|---|---|---|---|---|
| explain that DNA replication is semi-conservative and produces two identical molecules | 3.1.1 | | | |
| state the two roles of the enzyme helicase | 3.1.1 | | | |
| outline the function of the PCR and list two examples of its use | 3.1.2 | | | |
| outline the technique of gel electrophoresis and its use | 3.1.2 | | | |
| explain the orientation of DNA strands and how DNA polymerases work in a 5′→3′ direction | 3.1.3 | | | |
| distinguish the leading and lagging strands | 3.1.3 | | | |
| describe the process of DNA replication in eukaryotes and the functions of primase, polymerases and ligase | 3.1.3 | | | |
| define and outline the process of transcription | 3.2.1 | | | |
| describe how nucleosomes regulate transcription | 3.2.1 | | | |
| explain the importance of introns and exons | 3.2.1 | | | |
| define translation and the importance of complementary base pairing | 3.2.2 | | | |

## SELF-ASSESSMENT CHECKLIST

| I can... | Subsection | Needs more work | Almost there | Confident to move on |
|---|---|---|---|---|
| explain the role of ribosomes in the formation of polypeptides | 3.2.2 | | | |
| recall that free ribosomes synthesise proteins for use within the cell | 3.2.2 | | | |
| summarise the types and importance of non-coding DNA | 3.2.3 | | | |
| outline the roles of promoter regions and telomeres | 3.2.3 | | | |
| describe a polysome and identify them in electron micrographs | 3.2.3 | | | |
| outline how functional proteins are produced after translation | 3.2.3 | | | |
| state that new alleles are formed by mutation and changes may be harmful, beneficial or neutral | 3.3.1 | | | |
| explain that mutations may add, delete or substitute a base, or invert a section of dna, and that substitution causes sickle-cell disease, addition of repeated sequences leads to Huntington's disease. | 3.3.1 | | | |
| describe how the genetic code is degenerate and gives resilience to changes | 3.3.1 | | | |
| state that tumours are groups of cells that grow out of control and may be benign or malignant | 3.3.2 | | | |
| define mutagen and give some examples | 3.3.2 | | | |
| outline the importance of apoptosis | 3.3.2 | | | |
| recall that DNA polymerase can proofread errors in new DNA strands | 3.3.4 | | | |
| define epigenetics | 3.3.1 | | | |
| explain how gene expression is regulated by binding proteins | 3.4.1 | | | |
| state that epigenetic changes affect phenotype but not genotype | 3.4.2 | | | |
| describe how epigenetic changes can be due to DNA methylation and modification of histones | 3.4.2 | | | |
| give an example of a heritable epigenetic change | 3.4.3 | | | |

## CONTINUED

| I can... | Subsection | Needs more work | Almost there | Confident to move on |
|---|---|---|---|---|
| state that most epigenetic changes are not inherited but some can affect the epigenome of the offspring | 3.4.4 | | | |
| recall that epigenetics can cause variation more quickly than natural selection in a changing environment | 3.4.4 | | | |
| outline the importance of pollution in epigenetics | 3.4.5 | | | |
| describe the importance of monozygotic twins in the study of epigenetics | 3.4.5 | | | |
| outline the genetic condition PKU caused by a mutation | 3.6.7 | | | |

## REFLECTION

Reflect upon the content of this chapter and identify those areas of strength and weakness in your understanding. How can you improve in those topics you have found difficult?

## EXAM-STYLE QUESTIONS

You can find questions in the style of IB exams in the digital coursebook.

CAMBRIDGE
UNIVERSITY PRESS

# Biology

## for the IB Diploma

WORKBOOK

Alice Tully & Matthew Broderick

# Table of contents

## Chapter 3 DNA and protein synthesis

- **3.1** DNA replication
  - **3.1.1** DNA replication
  - **3.1.2** DNA sequencing
  - **3.1.3** The detailed process of DNA replication
- **3.2** Protein synthesis
  - **3.2.1** Transcription
  - **3.2.2** Translation
  - **3.2.3** non-coding regions of DNA
  - **3.2.4** Prosthetic groups
  - **3.2.5** Chaperone proteins
  - **3.2.6** Protein transport molecules
- **3.3** Mutations
  - **3.3.1** Chromosomes, genes, and mutations
  - **3.3.2** Harmful mutations and mutagens
- **3.4** Epigenetics
  - **3.4.1** Epigenetics and gene expression
  - **3.4.2** Epigenetic changes
  - **3.4.3** Epigenetic markers and offspring
  - **3.4.5** Pollution, methyl tags and twin studies
  - **3.4.4** Rate of epigenetic change

## Chapter 4 Genetics

- **4.1** Inheritance
  - **4.1.1** The genome
  - **4.1.2** Chromosome structure
  - **4.1.3** Genes and alleles
  - **4.1.4** Karyotyping
  - **4.1.5** Determination of sex
- **4.2** Genetic inheritance
  - **4.2.1** Principles of inheritance
  - **4.2.2** Determining genotypes and phenotypes
  - **4.2.3** Codominance and multiple alleles
  - **4.2.4** Incomplete dominance
  - **4.2.5** Sex chromosomes and autosomes
  - **4.2.6** Pedigree charts
  - **4.2.7** Genetic diseases
  - **4.2.8** Polygenes
  - **4.2.9** Variation in phenotypes without change to genotype
  - **4.2.10** Dihybrid crosses and linked genes
  - **4.2.11** The chi-squared test and dihybrid crosses

# Unit 2 Cellular organisation

## Chapter 5 Cell structure

- **5.1** Origins of life
  - **5.1.1** Forming organic molecules in the early Earth
  - **5.1.2** Cell theory
  - **5.1.3** The Miller–Urey experiments
  - **5.1.4** The deep-sea vent hypothesis and a source of energy for primitive life
  - **5.1.5** Micelles
  - **5.1.6** Comets
  - **5.1.7** Last universal common ancestor
- **5.2** Cell structure
  - **5.2.1** Cells and their structure
  - **5.2.2** The endosymbiosis theory
  - **5.2.3** Developments in microscopy
- **5.3** Viruses
  - **5.3.1** The structure of viruses
  - **5.3.2** Diversity and origins of viruses
  - **5.3.3** Rapid evolution in viruses

## Chapter 6 Cell function

- **6.1** Membranes and organelles
  - **6.1.1** Membrane structure
  - **6.1.2** Organelles
  - **6.1.3** Organelles and interactions between them
- **6.2** Movement across membranes
  - **6.2.1** Diffusion, facilitated diffusion and osmosis
  - **6.2.2** Active transport
  - **6.2.3** Membranes and transmission of nerve impulses
- **6.3** Water potential
  - **6.3.1** Water potential in plants and animals
  - **6.3.2** Advanced water potential
- **6.4** Limitations to cell size
  - **6.4.1** Surface area to volume ratio
  - **6.4.2** Cell growth and division

## Chapter 11 Evolution, speciation and Ecosystems

**11.1** Classification
- **11.1.1** The binomial system of classification
- **11.1.2** Using a dichotomous key
- **11.1.3** Cladistics
- **11.1.4** Finding evidence for clades and constructing cladograms
- **11.1.5** The shapes of cladograms

**11.2** Selection
- **11.2.1** A mechanism for evolution
- **11.2.2** Natural selection and the evidence for evolution
- **11.2.3** Artificial selection
- **11.2.4** Gene pools
- **11.2.5** Types of selection
- **11.2.6** The Hardy–Weinberg principle
- **11.2.7** Changing allele frequencies due to artificial selection

**11.3** Evolution
- **11.3.1** What is evolution?
- **11.3.2** Evidence for evolution
- **11.3.3** How new species arise
- **11.3.4** Effects of isolation on the gene pool

**11.4** Ecological niches and adaptations
- **11.4.1** Niches and community structure
- **11.4.2** Niches and the effects of competition
- **11.4.3** Convergent and divergent evolution and changes in structure

**11.5** Biodiversity
- **11.5.1** Competition in identical niches
- **11.5.2** Adaptations to different niches

## Chapter 12 Ecological relationships

**12.1** Modes of nutrition
- **12.1.1.** Feeding groups
- **12.1.2.** Complexities in feeding relationships
- **12.1.3** Adaptations for feeding

**12.2** Transfer of energy and matter
- **12.2.1** Energy flow
- **12.2.2** Nutrient recycling
- **12.2.3** Quantifying energy flow

**12.3** Ecological relationships and populations
- **12.3.1** Interactions and relationships between organisms and populations
- **12.3.2** Estimating population sizes
- **12.3.3** Growth of new populations
- **12.3.3** Features of relationships between predators, prey and plants
- **12.3.4** Cooperative interactions
- **12.4.5** Keystone species

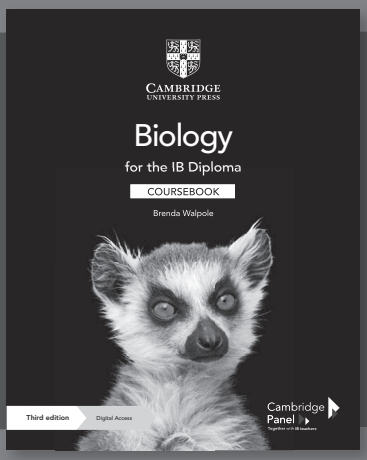
**12.4** Stability and change in ecosystems
- **12.4.1** Stability, change and succession
- **12.4.2** The impact of agriculture
- **12.4.3** Impact on biogeochemical cycles
- **12.4.4** The processes of succession
- **12.4.5** Human impacts on ecosystems
- **12.4.6** Pioneer species and succession

**12.5** Mass extinction and Biodiversity
- **12.5.1** Conservation of Biodiversity
- **12.5.2** Human activities and the 6th mass extinction
- **12.5.3** Causes of the Biodiversity crisis
- **12.5.4** Approaches to conservation
- **12.5.5** Eutrophication and Biomagnification

**12.6** Climate change
- **12.6.1** Causes and consequences of climate change
- **12.6.2** Timing of biological events and global warming

# How to use this series

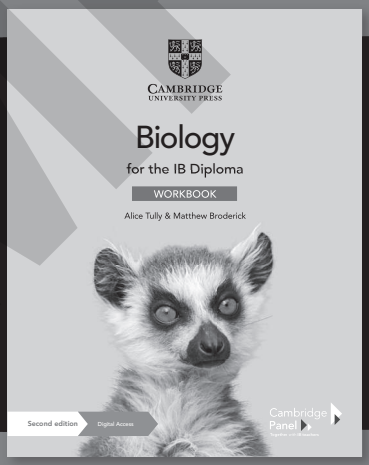
This suite of resources supports students and teachers of the IB Biology Diploma course. All of the books in the series work together to help students develop the necessary knowledge and scientific skills required for this subject.

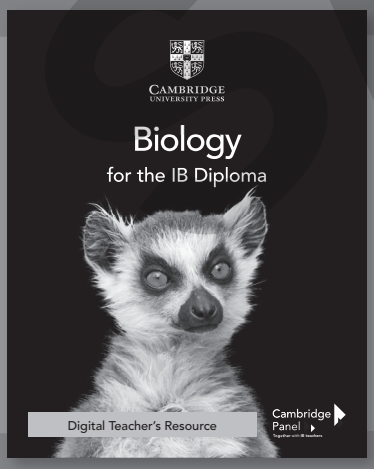



The coursebook with digital access provides full coverage of the latest IB Biology Diploma course.

It clearly explains facts, concepts and practical techniques, and uses real world examples of scientific principles. A wealth of formative questions within each chapter help students develop their understanding, and own their learning. A dedicated chapter in the digital coursebook helps teachers and students unpack the new assessment, while exam-style questions provide essential practice and self-assessment. Answers are provided on Cambridge GO, supporting self-study and home-schooling. With clear language and style, the coursebook with digital access is designed for international students.

The workbook with digital access builds upon the coursebook with digital access with further exercises and exam-style questions, carefully constructed to help students develop the skills that they need as they progress through their IB Biology Diploma course. The exercises also help students develop understanding of the meaning of various command words used in questions, and provide practice in responding appropriately to these.




The Teacher's resource supports and enhances the coursebook with digital access and the workbook with digital access. This resource includes teaching plans, overviews of required background knowledge, learning objectives and success criteria, common misconceptions, and a wealth of ideas to support lesson planning and delivery, assessment and differentiation. It also includes editable worksheets for vocabulary support and exam practice (with answers) and exemplar PowerPoint presentations, to help plan and deliver the best teaching.

# How to use this book

A chapter outline appears at the start of every chapter to introduce the learning aims and help you navigate the content.

**CHAPTER OUTLINE**

In this chapter you will:

- understand how different elements are required in different organisms
- explain how elements have different functions in different organisms
- describe how the properties of carbon result in a variety of stable compounds

**KEY TERMS**

Definitions of key vocabulary are given at the start of the chapter and are emboldened where they first occur in the chapter. You will also find definitions of these words in the end of book glossary.

## Exercises

Exercises help you to practice skills that are important for studying Standard Level and Higher Level Biology. A vertical line runs down the margin of all Higher Level exercises, allowing you to easily identify Higher Level from Standard Level exercises.

**EXAM-STYLE QUESTIONS**

Questions at the end of the each chapter are more demanding example questions, some of these may require use of knowledge from previous chapters. Answers to these questions can be found in digital form on Cambridge GO. A vertical line runs down the margin of all Higher Level exam-style questions, allowing you to easily identify Higher Level from Standard Level questions.

**TIP**

Tip boxes will help you complete the exercises, and given you support in areas that you might find difficult.

## Answers

Answers to the exercises and exam-style questions in the workbook are available online and are free to access. Visit Cambridge GO and register to access these resources.

Chapter 3

# DNA and protein synthesis

## CHAPTER OUTLINE

In this chapter you will:

- describe the process of DNA replication
- explain the stages of Polymerase Chain Reaction
- describe how gel electrophoresis is used to create a DNA profile, and applications for its use

**describe how DNA replication takes place on the leading and lagging strands**

- describe how RNA polymerase is used to make mRNA during transcription
- describe how translation occurs to make a polypeptide from mRNA
- understand how post translational modification produces mature mRNA in eukaryotic cells
- compare the processes of transcription and translation

**outline the roles of polysomes and telomeres**

- describe the role of mutagens in causing mutations
- compare the different types of gene mutation
- explain the significance of the genetic code showing degeneracy

**outline how proofreading is able to detect errors during DNA replication**

**describe ways in which gene expression can be controlled**

**describe examples of how the environment can cause epigenetic change**

**outline how DNA methylation is studied in twin case studies**

## KEY TERMS

**Make sure you understand the following key terms before you do the exercises.**

**allele:** an alternative form of a gene found at a specific locus (site) on a chromosome

**apoptosis:** programmed cell death which takes place in multicellular organisms. It removes cells that need to be removed during development or those which could damage the organism if they are allowed to survive

**carcinogen:** Anything that is associated with causing cancer

**codon:** a sequence of three mRNA bases that code for an amino acid

**diploid:** a diploid nucleus contains two copies of each chromosome in homologous pairs

**DNA methylation:** a process that adds a methyl group to the DNA base cytosine. It is an example of an epigenetic marker

**DNA replication:** copying DNA so that two identical new molecules are produced

**epigenetics:** the study of changes to gene activation in differentiated cells

**exons:** sequences of bases in mRNA that are spliced together and translated after introns have been removed

**gel electrophoresis:** a technique which separates DNA fragments according to their size and charge

**gene:** a particular section of a DNA strand that codes for a specific polypeptide; a heritable factor that controls a specific characteristic

**genome:** the whole of the genetic information of an organism

**haploid:** a cell containing one of each of a pair of homologous chromosomes

**homologous:** a pair of chromosomes with the same genes but not necessarily the same alleles of those genes

**lagging strand:** the DNA strand that is replicated in short fragments away from the replication fork

**leading strand:** the DNA strand that is replicated continuously towards the replication fork

**locus:** the specific position of a gene on a homologous chromosome; a gene locus is fixed for a species, for example, the insulin gene is always found at the same position on chromosome 11 in humans

**mutagen:** an agent that causes mutation

**mutation:** a change in the amount or sequence of bases in DNA

CONTINUED

**phenotype:** the characteristics of an organism which may be physical appearance or biochemical feature

**polymerase chain reaction (PCR):** a process in which small quantities of DNA are artificially amplified for research and diagnosis

**semi-conservative replication:** both strands of a DNA double helix are used as templates for replication so that new DNA molecules contain one original and one new strand

**somatic cell:** any cell forming part of the body of an organism

**transcription:** copying a sequence of DNA bases to mRNA

**translation:** decoding mRNA at a ribosome to produce a polypeptide

**triplet:** a sequence of three bases that code for an amino acid

**tumour:** a disorganised mass of cells

# Exercise 3.1 DNA replication

**DNA replication** produces identical DNA molecules used for growth, repair and reproduction. A variety of enzymes are involved to ensure efficient and accurate replication.

1 a Distinguish between a DNA and an RNA molecule. You should be able to come up with 4 distinct differences.

b Compare and contrast the bonding between the base pairs of T and A, and G and C.

c DNA replication is **semi-conservative**. Explain what is meant by semi-conservative when referring to DNA replication.

2 **PCR** is a technique that allows the rapid production of multiple copies of DNA.

a State the full name of PCR.

b State the name of the enzyme used in PCR.

c If there is one copy of a **gene**, calculate how many copies will be produced after eight cycles of PCR.

d Outline the three main stages that occur during PCR.

e PCR is used in medical diagnosis to amplify known sequences that only exist in certain viruses or bacteria. Outline the advantages of this use of PCR.

f State two factors that affect how far DNA fragments will travel during **gel electrophoresis**.

3 DNA profiling takes advantage of highly repetitive sequences found withing non-coding regions of DNA. The length of these regions vary from person to person, creating a unique 'fingerprint' for each individual. This is done by completing the following stages:

- DNA is extracted from sample.
- Enzymes cut the DNA into fragments.
- Fragments are separated by length.
- Fragments are treated with a radioactive probe that shows up on an X-ray film and looks like a bar code.

**a** Suggest suitable tissue samples that DNA can be extracted from in humans.

**b** State how and why the DNA fragments are separated during gel electrophoresis

**c** Look at Figure 3.1. Using your knowledge of DNA profiling, suggest which sample is a probable match for the sample from the crime scene.

**d** Explain why you chose the suspect for the previous question.

**e** Outline another possible use of DNA profiling.

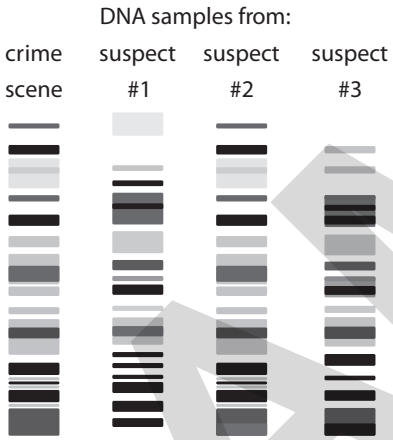



**Figure 3.1:** A forensic DNA profile comparison.

4 **a** Compare and contrast how DNA replication occurs on the **leading strand** and the **lagging strand** by completing the table below:

| | Replication of Leading Strand | Replication of Lagging strand |
|---|---|---|
| DNA is used as a template | YES | NO |
| DNA is created 5′ to 3′ | | |
| Replication is continuous | | |
| Replication occurs away from the replication fork | | |
| Okazaki fragments are involved | | |
| Only 1 RNA primer is required | | |

**b** DNA replication is carried out by a complex system of enzymes. Outline the role of each of the important enzymes as follows:

- **i** Helicase
- **ii** DNA primase
- **iii** DNA polymerase I
- **iv** DNA polymerase III

**c** Explain what is meant by 'antiparallel strands of DNA'.

# Exercise 3.2 Protein synthesis

The central dogma of biology is DNA → mRNA → protein. This exercise will travel the journey from DNA to protein synthesis.

**1** **Transcription** is the first step in protein synthesis. Answer the following questions about transcription.

**a** State the name of the enzyme that is responsible for transcribing a sequence of DNA bases into mRNA.

**b** State the name of the DNA strand that is used as a template during transcription.

**c** Name the RNA base that pairs with adenine during transcription.

**d** There are four different mRNA bases (A, U, C, G) that can be used to make the triplet code of an amino acid. Determine how many different triplet combinations can be made with these 4 bases.

**TIP**

Look out for questions that include DNA and RNA to make sure you are identifying the correct sequence.

**2** Deduce the complementary DNA base sequences for these mRNA sequences.

**a** GUC CCU AGA UUG

**b** CGA CCU CAC AAC

**c** CCC GCU GGA GUG

**3** **a** At the site of which organelle does **translation** occur?

**b** Define the term codon.

**c** Outline the role of a 'stop **codon**'

**4** Transcription occurs in three main stages.

**a** Name the stage of transcription that is being described in parts **i–iii**.

- **i** RNA polymerase binds to the DNA and the double helix unwinds.
- **ii** The mRNA becomes longer as RNA nucleotides are added to the 3′ OH group.
- **iii** When mRNA strand is complete, the DNA, RNA polymerase and mRNA disassociate from each other.

**b** Define the following key terms related to transcription and gene expression.

**i** Antisense strand

**ii** Sense strand

**iii** Exon

**iv** **mRNA** Splicing

**c** Identify the key terms from the following descriptions:

**i** Short sequence of non-coding DNA that acts as a binding site for RNA polymerase.

**ii** DNA wrapped around a core of eight histone proteins

**iii** Reversible reaction when $CH_3$ is added to a cytosine base in DNA.

**5** How many different amino acids are involved in protein synthesis?

**6** Compare and contrast transcription and translation by completing the table below.

| | Transcription | Translation |
|---|---|---|
| In which direction does it occur? | | |
| What is used as a template? | | |
| What is the main product | | |
| Where in the cell does it occur? | | |

**7** Four main stages of translation are outlined in **A–D**.

**A** The ribosome moves one codon along the mRNA strand in the 5′ to 3′ direction.

**B** A peptide bond forms between two adjacent amino acids

**C** Reaching a stop codon signals the polypeptide chain and the mRNA to release from the ribosome, which separates into its subunits.

**D** The components of translation are assembled: the mRNA, the small subunit of a ribosome, an activated tRNA molecule and a large ribosomal subunit.

**a** List the stages in the correct order, adding the name of each stage.

**b** Ribosomes have three tRNA binding sites. What are their names?

**c** The ribosome is separated into subunits during the termination stage of translation. Name the two subunits that the ribosome separates into.

**d** Deduce the name of the structure being described: A number of ribosomes can move along the same mRNA strand at the same time. This structure increases the efficiency of protein synthesis by working simultaneously on building the same polypeptide.

8 Telomeres prevent important genes from being lost.

a Define the term *telomere*.

b What happens to the length of a chromosome's telomeres as it is repeatedly replicated.

9 a Preproinsulin is converted to proinsulin, and then to insulin. The conversion of preproinsulin is carried out by the action of two different proteases. Explain why the modification of some polypeptides is important.

b Name of the intramolecular bond that links the polypeptide chains together in insulin.

# Exercise 3.3 Mutations

**Gene mutations** result in a change in the sequence or number of bases in DNA. They are often beneficial as they can create advantageous characteristics.

1 Identify the type of gene mutation described in parts **a–d**.

a The removal of one nucleotide and replacement with a different one

b A nucleotide is added into the DNA sequence.

c A mutation that changes the sequence GATC to CTAG.

d The removal of a nucleotide from a DNA sequence.

2 The genetic code is degenerate because a single amino acid may be coded for by more than one mRNA codon. What is the advantage of the genetic code being degenerate.

3 Sickle-cell anemia is a blood disorder that affects the ability of the blood to carry oxygen properly. People with sickle-cell anemia have hemoglobin that is abnormally shaped due to a genetic mutation.

Explain how a single base substitution mutation leads to sickle-cell anemia by completing the table below:

| | Normal version of gene | Mutated version of gene |
|---|---|---|
| DNA (haemoglobin gene) | CTC | |
| mRNA | GAG | |
| Amino acid | Glutamic Acid | |
| Shape of red blood cells | Biconcave disc | |

**TIP**

There are many important terms in biology. Make your own glossary and list the key terms in alphabetical order so that it is easier to keep track of them.

4 Distinguish between a benign and a malignant **tumour**.

5 Distinguish between the terms **mutagen** and carcinogen.

6 Use the information in Figure 3.2 to answer the following questions:

a State the percentage prevalence of smokers aged 20–24 in 1974.

State the percentage prevalence of smokers aged 20–24 in 2012.

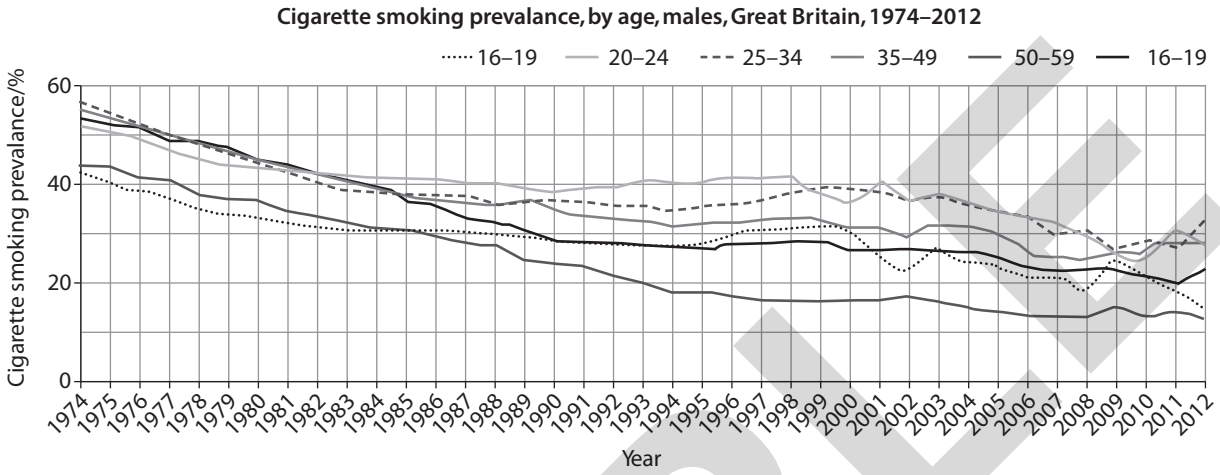



**Figure 3.2:** Trends in smoking in Great Britain.

Use the previous answers to calculate the percentage decrease of smoking prevalence in males aged 20–24 between 1974 and 2012.

Outline the general trend that you observe in Figure 3.2 for the prevalence of smoking in males aged 20–24.

What would be the difference in your answer to part d if the command term was 'describe' instead of 'outline'?

7 Figure 3.3 shows a similar graph for the prevalence of both men and women smokers over time.

a Compare the trend in smoking prevalence for males and females between 1974 and 2012.

b Explain why the prevalence of smoking has decreased in males and females.

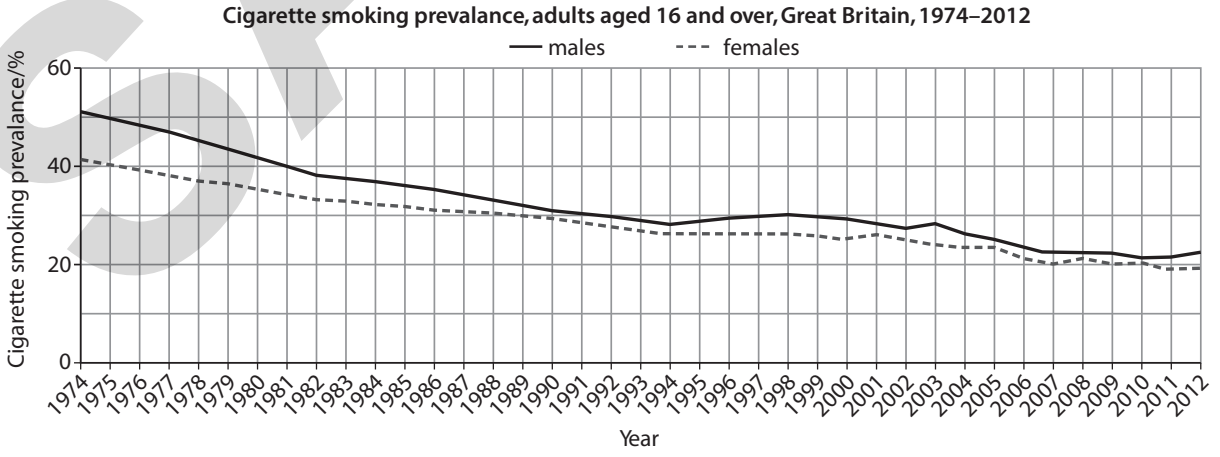



**Figure 3.3:** Trends in smoking prevalence for men and women over 16 in Great Britain.

**8** Read the following article about smoking before answering the questions that follow.

**To vape, or not to vape?**

E-cigarettes are battery-operated devices that provide inhaled doses of a vaporised solution and liquid nicotine. The assumption is that smoking electronic cigarettes (e-cigarettes) could be much safer than traditional tobacco-filled cigarettes. There is however no proof to this.

Manufacturers claim that the devices offer the following advantages:

- a safer alternative to tobacco cigarettes
- flavoured to taste better for customers
- help smokers to stop smoking and reduce the damaging effects of tar, nicotine and other carcinogenic ingredients.

Fewer people, especially children, are now smoking tobacco cigarettes. However, various authorities, such as the Food and Drug Administration (FDA) and the Centers for Disease Control and Prevention (CDC) in the USA have raised the following concerns.

- the products might not be safe to use.
- there are traces of nicotine and other toxic chemicals in the e-cigarettes.
- use among children has doubled.
- the ingredients have not been sufficiently studied or regulated.
- the voltage of the device and the temperature of the vapour can affect the level of emissions.
- new, probable carcinogens have been identified in vapour: propylene oxide and glycidol.

The e-cigarette industry is worth over US$3 billion in the USA alone. Their use has caused the recent decreasing numbers of people who smoke, so e-cigarettes could be a good thing. Most e-cigarette companies are owned by the tobacco companies, which produced the tobacco and cigarettes that e-cigarettes are intended to replace. Are the messages unbiased, and can e-cigarette users feel safe?

**a** Identify the evidence from this article that supports e-cigarettes as being safer than tobacco cigarettes.

**b** State the names of the authorities named in the article that offer a different viewpoint to the manufacturers.

**c** Using the article, discuss whether you think that students at your school should be allowed to smoke e-cigarettes.

9 The harvesting, production and sale of tobacco is banned in Bhutan and the consumption of tobacco is largely prohibited in public places. The government of Bhutan has taken these steps as part of its drive for spiritual content and happiness (Gross National Happiness). Tobacco is recognised as being harmful to both spiritual and social health. Bhutan is the only country in the world with such wide-ranging control over tobacco. However, the penalties for infringement of these laws has been criticised as being too harsh.

Using the information given, and your own knowledge of the effects of smoking, evaluate whether Bhutan is right to enforce anti-tobacco laws.

10 DNA proofreading and repair systems detect replication errors in DNA.

a State the name of the enzyme that is responsible for the addition of new nucleotides.

Outline what happens during mismatch repair.

# Exercise 3.4 Epigenetics

**Epigenetics** is the study of changes to the activation of genes in differentiated cells. These changes can lead to an alteration in the **phenotype** but not the **genotype**. Studying how identical twins develop different characteristics is an important theme in epigenetics.

1 a i **DNA methylation** inactivates a gene by reducing its level of expression. Explain what this means.

ii List two of the main causes of DNA methylation in humans.

iii State the name and formula of the component that is attached to DNA during DNA methylation.

b Human development is influenced by both nature and nurture.

i Explain what is meant by *nature* and *nurture*.

ii Discuss whether nature or nurture affects the phenotypic features of an organism.

2 Identical twins share the same DNA, but often experience different environments, particularly as the twins get older. Scientists can study twins and try to find out which genetic, or environmental, factors are responsible for the onset of different diseases and conditions. A comparison of traits between identical twins and fraternal twins (twins who came from two separate maternal eggs) can help to further identify the influence of genetic factors. Use Figure 3.5 to answer the questions that follow.

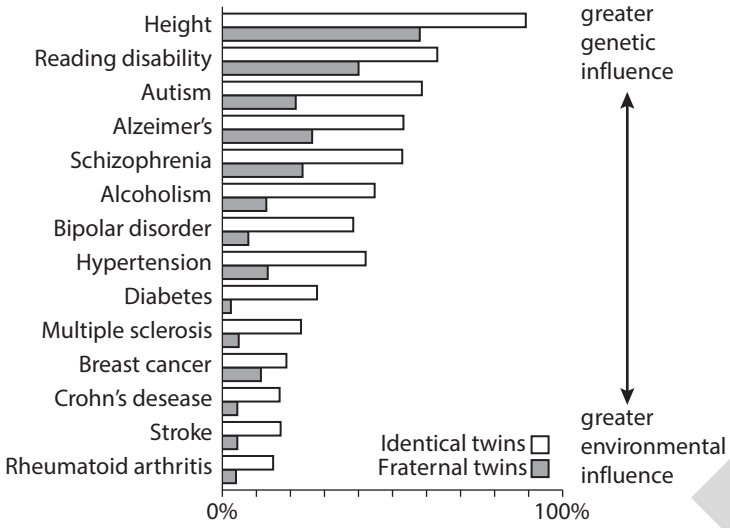



**Figure 3.5:** Comparison of identical and fraternal twins.

**a** **i** Estimate the percentage of identical twins in which both twins suffer from a reading disability.

**ii** Estimate the percentage of identical twins in which both twins suffer from a bipolar disorder.

**iii** Estimate the percentage of identical twins in which both twins suffer from diabetes.

**iv** Estimate the percentage of identical twins in which both twins suffer from breast cancer.

**b** Suggest what can be determined when a high percentage of identical twins share a particular condition.

**c** Height is determined by genetic factors. Evaluate this statement.

**d** Explain why identical twins may not express the same genes at the same time.

**3** Recent research into epigenetics suggests a link between poor diet in women whilst they are pregnant and the incidence of attention deficit hyperactivity disorder (ADHD) in their children. The studies have shown that a diet of high fat and high sugar may alter epigenetic markers in the unborn child's DNA.

Annotate Figure 3.6 to show where a methyl group is added to cytosine during methylation.

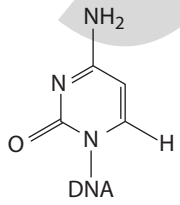



**Figure 3.6:** Cytosine

## EXAM-STYLE QUESTIONS

1 Which of the following roles is undertaken by DNA Polymerase III during the process of DNA replication?

I Joining Okazaki fragments together
II Proofreading the new DNA strands
III Joining complimentary DNA nucleotides to the template strands

A **I** only
B **II** and **III** only
C **III** and **I** only
D **I**, **II** and **III**

2 Why is Taq Polymerase used in the Polymerase Chain Reaction?

A It originates from the human body, so functions optimally with our DNA
B It can add DNA nucleotides in both the 3' - 5' and 5'- 3' directions
C It can withstand very high temperatures without denaturing
D It works faster than DNA Polymerase, so makes the process of PCR more efficient

3 Which of the following is not an application for DNA Profiling?

A Paternity testing
B To sequence the DNA of extinct and endangered species
C To identify the perpetrator in forensic scene of crime investigations
D To identify viral and bacterial pathogens for use in medical diagnosis

4 Areas of non-coding DNA varies in its level of repetitiveness from person to person. Which process takes advange of VNTR and STR sequences being specific to an individual?

A Mutation
B Transcription
C Polymerase chain reaction
D DNA profiling

5 Which of the folllowing describes DNA replication?

A Conservative
B Semi conservative
C Dispersive
D Helical

6 Which pair of scientists determined the method of DNA replication?

A Franklin and Wilkins
B Hershey and Chase
C Meselson and Stahl
D Watson and Crick

7 Where are the first DNA nucleotides added by DNA Polymerase during replication?

A To the 3' end of the RNA primer
B To the 5' end of the template strand
C To the 3' end of an okazaki fragment
D To the 5' end of the non-coding strand

CONTINUED

8 Which of the following best describes transcription?
- **A** DNA directed polypeptide synthesis
- **B** mRNA directed protein synthesis
- **C** DNA directed mRNA synthesis
- **D** Protein directed DNA synthesis

9 Which of the following is not a feature of the genetic code?
- **A** Read in triplets
- **B** Degenerate
- **C** Universal
- **D** Spliced

10 Why is the mRNA used for translation shorter than the gene it was created from?
- **A** The poly A tail is removed during post transcriptional modification
- **B** The 5' cap is removed during post transcriptional modification
- **C** Introns are removed during post transcriptional modification
- **D** Telomeres are removed during post transcriptional modification

11 Which enzyme controls the process of transcription?
- **A** RNA polymerase
- **B** Helicase
- **C** DNA primase
- **D** Taq polymerase

12 Why do prokaryotic cells not undergo mRNA splicing?
- **A** Their DNA does not have introns
- **B** Their DNA does not contain exons
- **C** Their DNA is not wrapped around histones
- **D** Their DNA is not contained in a nucleus

13 What occurs as a result of alternative splicing?
- **A** Increase in the size of the genome
- **B** Increase in the size of the proteome
- **C** Increase in the size of the cell
- **D** Increase in the surface area to volume ratio

14 Which of the following correctly describes protein synthesis?
- **A** All proteins contain 20 different amino acids
- **B** Each polypeptide is created by transcription
- **C** Ribosomes create protein by joining tRNAs together
- **D** Peptide bonds join amino acids during the process of translation

15 Which of the following correctly describes an aspect of tRNA?
- **A** A clover leaf shape is created due to peptide bonding between complementary bases
- **B** An amino acid is attached to the 5' end at the AAC sequence
- **C** The codon is carried at the base of the structure, which is a sequence of 3 RNA nucleotides
- **D** tRNA is a single stranded RNA molecule consisting of A, U, C and G nucleotides

CONTINUED

16 Which of the following correctly describes a feature of ribosome structure?

A There are 3 tRNA binding sites
B There are 3 mRNA binding sites
C There are 3 subunits
D There are 3 rRNA binding sites

17 Which of the following is not true of sickle-cell anemia?

A It affects the gene for haemoglobin
B The gene is carried on the X chromosome
C It is a dominant trait
D It is caused by a base substitution mutation

18 Which of the statements A–D, correctly describes what can be found in a nucleosome?

I Histones
II DNA
III RNA
IV Nucleoid

A II only
B I and IV
C II and III
D I and II

19 Identify the process described. Programmed cell death that takes place in multicellular organisms.

A Methylation
B Malignant
C Apoptosis
D Metastasis

20 Identify the structure shown in Figure 3.6.

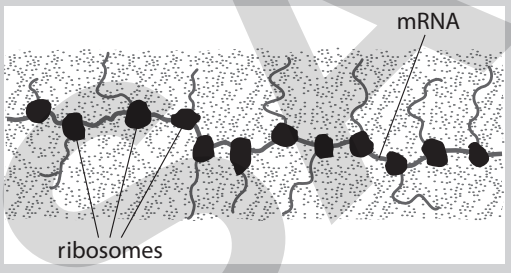



Figure 3.6

A Polysome
B Globular protein
C Amino acid
D Eukaryotic cell

CONTINUED

**21** What does the term 'splicing' refer to?
- **A** Breaking
- **B** Joining
- **C** Cutting
- **D** Creating

**22** Alternative splicing creates:
- **A** DNA with only a proportion of its original coding sequence
- **B** Variants of a protein from the same mRNA, thereby increasing the proteome of an organism
- **C** Different polypeptides, each with a different function in the body
- **D** Polypeptides that include a number of introns and exons

**23** What is the role of nucleases in protein synthesis?
- **A** They catalyse peptide bond formation during translation
- **B** They bind to nucleosomes and prevent both transcription and translation
- **C** They bind to promoters and increase the rate of translation
- **D** They regulate translation by breaking down mRNA strands

**24** Which is the best description of a proteasome?
- **A** A protein complex that recycles unnecessary or non-functioning proteins in the cell
- **B** An enzyme that catalyses the formation of peptide bonds during translation
- **C** A structure involved in the removal of introns during post transcriptional modification
- **D** The structure formed when multiple ribosomes attach to the same piece of mRNA

**25** Which of the following is an application of alternative splicing?
- **A** The creation of slightly different antibodies in response to different antigens
- **B** The production of starch in plants and glycogen in animals
- **C** The creation of millions of copies of target DNA during PCR
- **D** Forensic investigations and paternity testing

**26** Which of the following is most likely to affect gene expression?
- **A** Substitution of one base for another in the middle of the gene
- **B** Insertion of a base at the start of the gene
- **C** Deletion of a base towards the end of the gene
- **D** Methylation of a cytosine base in the centre of the gene

**27** Which of the following is most likely to decrease the expression of a gene?
- **A** Binding of a general transcription factor to the promoter region
- **B** Increase in the number of polysomes
- **C** Binding of an activator protein to the enhancer region of a gene
- **D** Increase in the number of nucleases within a cell

CONTINUED

**28** Which of the following describes a purpose of non-coding regions of DNA?

**I** To form telomeres at the end of chromosomes
**II** To regulate transcription
**III** To code for the production of rRNA and tRNA

**A** I only
**B** I and II only
**C** II and III only
**D** D.I, II and II

**29** How many tRNA molecules can bind simultaneously to the large subunit of a ribosome?

**A** A.0
**B** B.1
**C** C.2
**D** D.3

**30** Which of the following proteins are most likely to be made by bound ribosomes at the RER rather than by free ribosomes in the cytoplasm?

**A** A hormone
**B** A membrane protein
**C** An enzyme
**D** A histone protein

**31** Which of the following correctly describes an aspect of mutation?

**A** All mutations are harmful
**B** All mutagens are carcinogens
**C** Mutations to DNA in germ cells are likely to be passed to offspring
**D** Mutations remove parts of the nucleotide sequence for a particular gene

**1** Describe the importance of DNA replication in cellular division. **[3]**
**2** How can gene and chromosome mutations occur during mitotic cell division? **[4]**
**3** How do genetic diseases caused by dominant alleles persist in humans? **[4]**
**4** **a** Outline the method used to conduct a DNA profile. **[5]**
**b** Outline three uses of DNA profiling. **[6]**
**5** Outline the roles of the enzymes involved in DNA replication **[4]**
**6** Compare the replication of the leading and lagging strand **[2]**

CONTINUED

**7** The image shows a diagram of a tRNA molecule. Draw onto the diagram to show the position of:

**a** The amino acid

**b** The anticodon

**c** The 3' and 5' ends

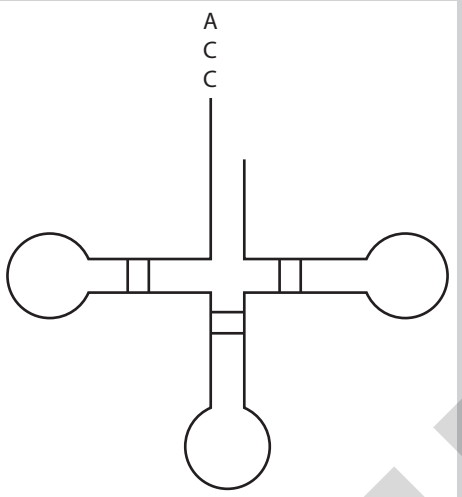



**Figure 3.7:** tRNA molecule **[5]**

**8** The image shows a diagram of one of the stages of translation. Draw onto the diagram to indicate:

**a** A peptide bond

**b** a tRNA molecule

**c** mRNA

**d** an amino acid

**e** the direction in which the ribosome is moving

**f** the missing anticodon

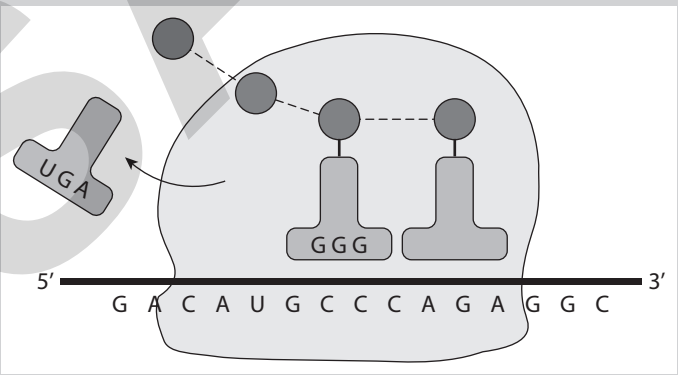



**Figure 3.8:** Stages of translation. **[6]**

CONTINUED

**9** The image shows a diagram of a section of DNA about to be transcribed. Draw onto the diagram to indicate:

**a** the name of the enzyme attached to DNA

**b** the template strand **[2]**

**10** Deduce the sequence of nucleotides on the mRNA strand that will be produced

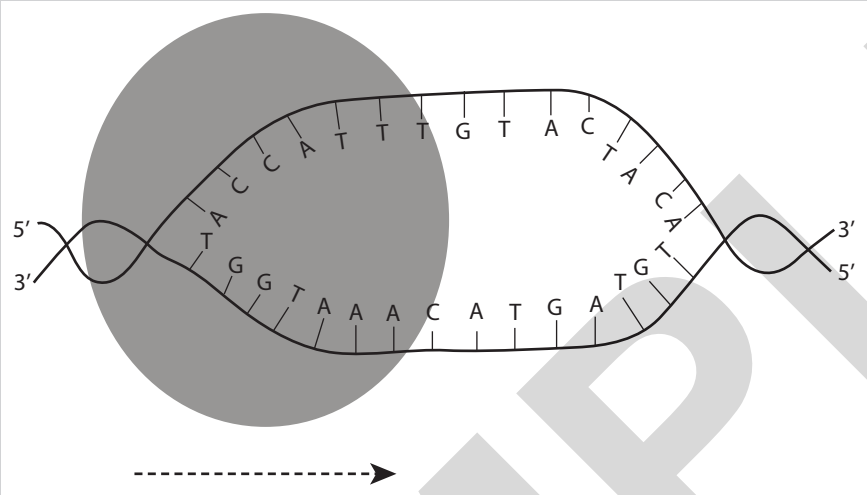



**Figure 3.9:** Nucleotides on the mRNA. **[2]**

**11** The image shows the stages of protein production

**a** Identify processes 1 and 2

**b** Identify structures a, b and c

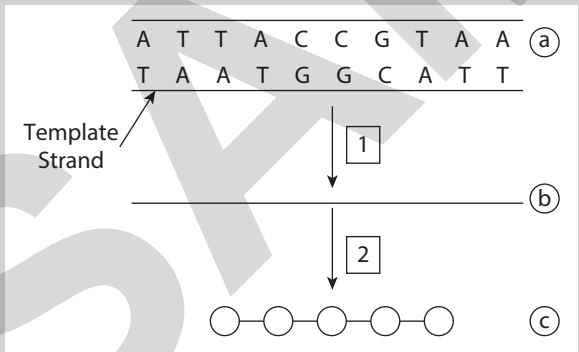



**Figure 3.10:** stages of protein production **[2]**

CONTINUED

**12** After the death of a wealthy person, there were no named inheritors due to lack of a will, and no known relatives. It was decided that only the closest living relative would be considered. 3 people came forward as potential relatives with a claim to the fortune. Samples of DNA were taken from each claimant and compared to the wealthy's DNA using gel electrophoresis. The results are shown below:

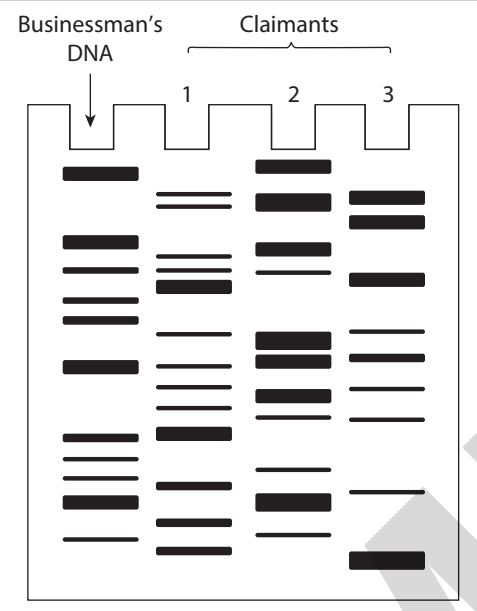



**Figure 3.11:** DNA using gel electrophoresis.

Deduce which claimant should inherit the fortune **[2]**

Answer: Claimant 3 has no bands in common, so should not inherit

Claimant 1 has a couple of bands in common, but is not likely to be a close relative

Claimant 2 has over half of their bands in common, so is most likely to be a closer relative. They should inherit.

**13** The following table shows the number of bases involved in the production stages of protein Y

| | |
|---|---|
| DNA gene sequence | 15,000 bases |
| mRNA transcript | 14,500 bases |
| mature mRNA | 8,500 bases |

**a** Calculate the percentage change between the number of bases in the DNA sequence and the mRNA transcript **[1]**

**b** What is the reason that the mRNA transcript and DNA sequence are not the same length? **[1]**

**c** Calculate the percentage difference between the mRNA transcript and the mature mRNA **[1]**

**d** Why is the mature mRNA shorter than the initial mRNA transcript? **[1]**

CAMBRIDGE
UNIVERSITY PRESS

# Biology

## for the IB Diploma

Digital Teacher's Resource

Cambridge Panel
Together with IB teachers

# Table of contents

## Chapter 6 Cell function

- 6.1 Membranes and organelles
  - **6.1.1** Membrane structure
  - **6.1.2** Organelles
  - **6.1.3** Organelles and interactions between them
- 6.2 Movement across membranes
  - **6.2.1** Diffusion, facilitated diffusion and osmosis
  - **6.2.2** Active transport
  - **6.2.3** Membranes and transmission of nerve impulses
- 6.3 Water potential
  - **6.3.1** Water potential in plants and animals
  - **6.3.2** Advanced water potential
- 6.4 Limitations to cell size
  - **6.4.1** Surface area to volume ratio
  - **6.4.2** Cell growth and division
- 6.5 Cell division
  - **6.5.1** Binary fission in single-celled organisms
  - **6.5.2** The cell cycle
  - **6.5.3** Meiosis
  - **6.5.4** Non-disjunction
  - **6.5.5** Chromosome behaviour and Mendel's laws

## Chapter 7 Cell control and communication

- 7.1 Principles of cell signalling
  - **7.1.1** Principles of cell signalling and cell interaction
  - **7.1.2** Cell signalling in unicellular organisms
  - **7.1.3** Cell signalling in multicellular organisms
- 7.2 Neural transmission
  - **7.2.1** The structure of nervous systems
  - **7.2.2** Transmission of nerve impulses
  - **7.2.3** Synapses and synaptic transmission
  - **7.2.4** Myelination of nerve fibres
  - **7.2.5** Excitatory and inhibitory neurotransmitters
  - **7.2.6** Effects of chemicals on synaptic transmission
  - **7.2.7** Perception of pain and consciousness
- 7.3 Chemical signalling in animals and plants
  - **7.3.1** Hormones in animals
  - **7.3.2** Insulin and glucagon, and control of blood glucose
  - **7.3.3** Using hormones in medical treatments
  - **7.3.4** Mode of action of steroid and amino acid-derived
  - **7.3.5** Effects of phytohormones (plant growth regulators)

# Unit 3: Organisation of organisms

## Chapter 8 Physiology



## Chapter 9 Co-ordination, muscles and motility



## Chapter 10 Defence against disease

## Chapter 11 Evolution, speciation and Ecosystems

- 11.1 Classification
  - **11.1.1** The binomial system of classification
  - **11.1.2** Using a dichotomous key
  - **11.1.3** Cladistics
  - **11.1.4** Finding evidence for clades and constructing cladograms
  - **11.1.5** The shapes of cladograms
- 11.2 Selection
  - **11.2.1** A mechanism for evolution
  - **11.2.2** Natural selection and the evidence for evolution
  - **11.2.3** Artificial selection
  - **11.2.4** Gene pools
  - **11.2.5** Types of selection
  - **11.2.6** The Hardy–Weinberg principle
  - **11.2.7** Changing allele frequencies due to artificial selection
- 11.3 Evolution
  - **11.3.1** What is evolution?
  - **11.3.2** Evidence for evolution
  - **11.3.3** How new species arise
  - **11.3.4** Effects of isolation on the gene pool
- 11.4 Ecological niches and adaptations
  - **11.4.1** Niches and community structure
  - **11.4.2** Niches and the effects of competition
  - **11.4.3** Convergent and divergent evolution and changes in structure
- 11.5 Biodiversity
  - **11.5.1** Competition in identical niches
  - **11.5.2** Adaptations to different niches

## Chapter 12 Ecological relationships

- 12.1 Modes of nutrition
  - **12.1.1.** Feeding groups
  - **12.1.2.** Complexities in feeding relationships
  - **12.1.3** Adaptations for feeding
- 12.2 Transfer of energy and matter
  - **12.2.1** Energy flow
  - **12.2.2** Nutrient recycling
  - **12.2.3** Quantifying energy flow
- 12.3 Ecological relationships and populations
  - **12.3.1** Interactions and relationships between organisms and populations
  - **12.3.2** Estimating population sizes
  - **12.3.3** Growth of new populations
  - **12.3.3** Features of relationships between predators, prey and plants
  - **12.3.4** Cooperative interactions
  - **12.4.5** Keystone species

# How to use this Teacher's resource

This digital Teacher's resource contains both general guidance and teaching notes that help you deliver the content for the IB Biology course. You will find answers to the coursebook and the workbook questions on the supporting resources area of Cambridge GO – they are freely available to teachers only.

There are **teaching notes** for each sub-chapter of the coursebook. You can see an overview of where all topics are covered in the teaching plan, in the 'Resources' column. Each set of teaching notes contains various features to help you deliver the topics covered in a unit/chapter.

At the start of each chapter there is a teaching plan for the chapter. This summarises the topics covered in the chapter, including the number of learning hours recommended for each topic, an outline of the learning content, and the resources from this series that can be used to deliver the topic.

## Teaching plan

| Sub-chapter | Approximate number of learning hours | Learning content | Resources |
|---|---|---|---|
| | | | |

This icon in the resources section indicates material that is available from Cambridge GO.

Each chapter also includes information on any **background knowledge** that learners should have before studying content covered in the chapter.

**BACKGROUND KNOWLEDGE**

- Explain what an electromagnetic spectrum represents, the different types of radiations and their uses, the quantitative relationship between wavelength, frequency and energy of the radiations.

## Syllabus overview

- At the start of each unit there is a syllabus overview, which gives a brief outline of the content knowledge, practical skills and opportunities to cover assessment objectives covered in that section of the syllabus. It also provides links to related topic areas in other parts of the syllabus.
- The **learning plan** will enable you to identify the related learning intentions and success criteria from the coursebook chapter.

**LEARNING PLAN**

| Learning objectives | Success criteria |
|---|---|
| | |

There is also a feature highlighting any **common misconceptions** associated with particular learning topics. Potential misunderstandings are identified, along with methods of eliciting evidence of these misconceptions from your class and suggestions on how to overcome them.

| Misconception | How to identify | How to overcome |
|---|---|---|
| | | |

For each topic, there is a selection of **lesson starter ideas**, **main teaching ideas** and **plenary ideas**. You can pick out individual ideas that meet the needs of your class. The activities include suggestions for how they can be differentiated or used for assessment.

**Differentiation ideas** are provided for each topic, with 'stretch and challenge' activities and ideas to extend learning opportunities and 'support' activities, ideas and modifications for learners who need extra practice or help.

The **cross-curricular links** feature provides suggestions for linking to other areas of study within the Standard Level and Higher Level IB curriculum. Cross-topic links allow students to make connections between the different syllabus sections of the IB Biology course. They encourage students to approach Biology as a holistic topic and help them develop the skills required for approaching exam questions, which often draw on several areas of the course. Syllabus sections key – CD = Commonality and Diversity, FF = Form and Function, II = Interaction and Interdependence, EI = Equilibrium and Resilience.

| CROSS-CURRICULAR LINKS |
|---|
| |

This Teacher's resource includes a range of digital materials that you can download from Cambridge GO. (For more information about how to access and use your digital resource, please see inside front cover.)

You will find the **glossary** of terms for the coursebook and workbook and also answers to **activities**, **worksheets** and **end of chapter tests** within and at the end of this resource.

To help with lesson planning, a blank lesson plan template is available to download from Cambridge GO (as part of this digital Teacher's resource).

More information about these approaches to learning and teaching is available to download from Cambridge GO (as part of this digital Teacher's resource).

# 3 DNA and protein synthesis

## Teaching plan

| Sub-chapter | Approximate number of learning hours | Learning content | Resources |
|---|---|---|---|
| 3.1 DNA replication | 2–4 | Students learn about the process of semi-conservative replication, linking this to polymerase chain reactions and gel electrophoresis. | **Coursebook**<br>Section 31<br>Test your understanding 1–3<br>Test your understanding 4–7<br>Exam-style questions 1, 3 and 7<br>**Workbook**<br>Exercise 3.1<br>Exam-style question 6<br>**Teacher's resource**<br>PowerPoint 3 slides 1 and 2 |
| 3.2 Protein synthesis | 3–6 | Students explore the concept of protein synthesis through transcription, including its regulation and translation. | **Coursebook**<br>Section 3.2<br>Test your understanding 8–12<br>Test your understanding 13–15<br>Exam-style questions 2, 4 and 11<br>**Workbook**<br>Exercise 3.2<br>Exam-style questions 3 and 5<br>**Teacher's resource**<br>PowerPoint 3, slides 1 and 2 |

| Sub-chapter | Approximate number of learning hours | Learning content | Resources |
|---|---|---|---|
| 3.3 Mutations | 2–4 | Students learn that there are different types of mutation and that they can have varying effects, including modifying the cell cycle and leading to tumours. | **Coursebook**<br>Section 3.3<br>Test your understanding 16–20<br>Test your understanding 21–23<br>Exam-style question 8<br>**Workbook**<br>Exercise 3.3<br>Exam-style questions 2, 4 and 8<br>**Teacher's resource**<br>PowerPoint 3, slides 4–7 |
| 3.4 Epigenetics | 1–2 | Students to consider protective mechanisms such as the degeneracy of the genetic code and apoptosis as well as artificial treatment in the form of gene therapy. | Coursebook<br>Section 3.4<br>Test your understanding 24–28<br>Exam-style questions 9 and 12<br>**Workbook**<br>Exercise 3.4<br>**Teacher's resource**<br>PowerPoint 3, slides 4–7 |

## BACKGROUND KNOWLEDGE

- Students require a strong understanding of both DNA and RNA structure to understand the processes of DNA replication and protein synthesis.
- To fully understand how genes code for proteins, ensure students have a detailed understanding of primary, secondary, tertiary and quaternary levels of protein structure.
- Students should have a basic understanding of enzyme function, including how they are affected by temperature, as they will encounter several examples in this section of the course.
- Students will need an understanding of cell structure and organelle function, especially ribosomes, to fully comprehend the mechanism of protein synthesis.
- Students should study the cell cycle before learning about the impact of mutations, so that they gain a deeper understanding of how mutations can lead to tumours by impacting the cell cycle.
- A basic understanding of natural selection and evolution will help students better understand the impact of mutations.

# Syllabus overview

- This section covers all parts of the Continuity and Change/Molecules section of the syllabus and the second part of the Continuity and Change/cells section of the syllabus. They all have Standard Level and Higher Level components. Students learn about the processes of DNA replication, protein synthesis and factors affecting gene expression.

# 3.1 DNA replication and 3.2 Protein synthesis

## LEARNING PLAN

| Learning objectives | Success criteria |
|---|---|
| • Understand that DNA replication is a semi-conservative process that produces two identical new molecules.<br>• Learn that the enzyme helicase unwinds the double helix and separates the strands by breaking hydrogen bonds between base pairs.<br>• Understand that the polymerase chain reaction is a laboratory process that amplifies small quantities of DNA.<br>• Recognise that gel electrophoresis separates DNA fragments by their charge and size.<br>• Define transcription as the synthesis of mRNA by RNA polymerase.<br>• Learn that complementary base pairing between DNA and mRNA ensures that the polypeptides produced function properly.<br>• Define translation as the production of polypeptides from mRNA.<br>• Understand that in prokaryotes, translation occurs immediately after transcription but eukaryotes modify mRNA by removing introns to form mature mRNA composed of exons.<br>• Recognise that exons can be spliced in different ways to produce different proteins from a single gene.<br>• Both prokaryotes and eukaryotes have non-coding sequences. In prokaryotes, these serve regulatory functions such as promotors. | • Students will be able to explain that DNA replication is semi-conservative and produces two identical DNA molecules.<br>• Students will be able to state the two roles of the enzyme helicase.<br>• Students will be able to outline the function of the PCR and list two examples of its use.<br>• Students will be able to outline the technique of gel electrophoresis and its use.<br>• Students will be able to define and outline the process of transcription.<br>• Students will be able to explain the importance of complementary base pairing.<br>• Students will be able to define translation.<br>• Students will be able to state that translation occurs immediately after transcription in prokaryotes as they lack a nuclear membrane.<br>• Students will be able to explain the importance of introns and exons.<br>• Students will be able to describe how nucleosomes regulate transcription.<br>• Students will be able to explain the importance of complementary base pairing.<br>• Students will be able to use molecular visualisation software to analyse the structure of a tRNA molecule.<br>• Students will be able to explain the role of ribosomes and their structure in the formation of polypeptides. |

CONTINUED

| Learning objectives | Success criteria |
| --- | --- |
| • Understand how nucleosomes regulate transcription in eukaryotes.<br>• Recognise how complementary base pairing between codons and anticodons ensures accurate translation.<br>• Learn that ribosomes are the sites of translation; free ribosomes synthesise proteins for use within the cell, whereas bound ribosomes synthesise proteins for secretion or use in lysosomes. | • Students will be able to recall that free ribosomes synthesise proteins for use within the cell.<br>• Students will be able to use molecular visualisation software to analyse the structure of eukaryotic ribosomes. |
| Recognise that DNA strands are antiparallel and are orientated in opposite directions.<br>The leading strand is replicated continuously whilst the lagging strand is replicated discontinuously.<br>Learn that DNA replication is regulated by a series of enzymes: DNA primase, DNA polymerase I, DNA polymerase III and DNA ligase.<br>Learn that DNA polymerase can only work in a 5′→ 3′ direction.<br>Discover that DNA polymerases proofread new DNA strands.<br>A large portion of the eukaryotic genome consists of non-coding sequences.<br>Non-coding DNA persists for many generations and has important functions.<br>Polysomes allow many polypeptides to be made at the same time.<br>Translation does not always result in functional protein. | Students will be able to explain the orientation of DNA strands and how DNA polymerases work in a 5′→ 3′ direction.<br>Students will be able to distinguish between the leading and lagging strands.<br>Students will be able to describe the process of DNA replication in eukaryotes and the functions of primase, polymerases and ligase.<br>Students will be able to state the direction in which DNA polymerase works.<br>Students will be able to outline the role of DNA polymerases in proofreading new DNA strands.<br>Students will be able to summarise the types and importance of non-coding DNA.<br>Students will be able to outline the roles of promotor regions and telomeres.<br>Students will be able to describe polysomes and identify them in electron micrographs.<br>Students will be able to outline how functional proteins are produced after translation. |

## Common misconceptions

| Misconception | How to identify | How to overcome |
|---|---|---|
| Students might describe a molecule of DNA as a strand of DNA. | This often occurs when students describe semi-conservative replication, so ask student to do this. | Ask students to describe the shape of DNA or ask them to define the term double helix. Review the key differences between RNA and DNA. Both exercises highlight DNA as a double-stranded molecule. |
| Some students use the words unwinding and separating interchangeably when describing the role of DNA helicase. | Ask students to outline the role of DNA helicase. Students who only use one of unwind or separate are likely to think these mean the same thing. | The unwinding and separating of DNA strands can be simply modelled with two pieces of intertwined string. |
| Students might rote learn that replication, transcription and translation 'happen' in a 5'→ 3' direction without understanding what this means. | Ask students to describe the direction of movement of DNA polymerase (III) along both the template strand and along the strand being synthesised. | Students might find it easier to understand that nucleotides can only be added to 3' end of another nucleotide if they understand that the joining of nucleotides is a condensation reaction, which requires a free hydroxyl group. This could be shown on a molecular diagram of a nucleotide. |
| Students may confuse DNA replication with cell division. | Ask students what happens to a cell during DNA replication. | Review the cell cycle, focussing on the S-phase where DNA replication takes place. |
| Students often struggle to differentiate between DNA, genes and chromosomes. | Ask students to put these in order of size and distinguish between them. | Showing students diagrams that depict the relative sizes of these structures could help. Match-up activities to determine their functions could be a useful exercise. Analogies to explain their functions may also aid understanding (see Differentiation section below). |
| Students might think that translation makes amino acids. | Ask students to describe how amino acids are made and/or their role in translation. | Review the structure of proteins. |

## Starter ideas

### 1 Complementary base pairing memory aids (10 minutes)

**Resources:** A diagram of the structure of DNA that depicts complementary base pairing, which includes the names of the four bases: adenine, thymine, guanine and cytosine (see PowerPoint 2, slide 23).

**Description and purpose:** The purpose of this starter is to help students remember complementary base pairing. Tell students to work in pairs or threes to invent as many memorable ways as possible of remembering which bases pair with which, e.g. Apple Tart, Chocolate Gateau. Students could share their favourite ideas with the class and the most popular ones could be displayed in the classroom to refer to in subsequent activities.

### 2 DNA and RNA think-pair-share (10 minutes)

**Resources:** Writing equipment.

**Description and purpose:** The purpose of this activity is to remind students about the key similarities and differences between DNA and RNA, so you can address any misconceptions they might have had. Ask students to work individually to write down brief statements regarding what they can remember about the similarities and differences between DNA and RNA. This is a good opportunity for students to practise exam technique by using the word 'whereas' to highlight differences in a sentence. Then, ask students to exchange their ideas with a partner, adding any missing points to their list. Finally, ask students to volunteer one of their ideas at a time to share with the whole class, until all points have been covered.

### 3 Nucleic acids word ladder (10 minutes)

**Resources:** Writing equipment.

**Description and purpose:** The purpose of this activity is to discover students' prior knowledge of the topic. Ask students to write the words 'nucleic acid' down. Explain that they are going to work in pairs or threes to create a list of key terms called a word ladder. Each key term in the list needs to start with the last letter of the previous word and be linked to the topic of nucleic acids. After a set time, ask students to stop writing and count their words. Students with the longest word ladders should be asked to read their word ladder out slowly and clearly to the rest of the class. If a student in the class believes a key term does not belong in the word ladder, they should raise their hand to question it. If the students reading out their word ladder can justify all key terms questioned in their word ladder, they win the word ladder competition. If not, play passes to the team with the next longest word ladder, until a winner is found. An example of a word ladder is included below:

1 Nucleic acid
2 DNA
3 Adenine
4 Epigenetics
5 Sugar-phosphate backbone
6 Elongation
7 Nitrogen
8 Nucelotide

### 4 Protein structure review (10 minutes)

**Resources:** Plasticine.

**Description and purpose:** The purpose of this activity is for students to review how proteins are formed from amino acids so that this can be linked to the role of nucleic acids in protein synthesis. Provide pairs or threes with plasticine of a variety of colours and a word bank, including the words: monomer, polymer, polypeptide, protein, amino acid, peptide bond, hydrogen bond, folding, primary/secondary/tertiary structure. Ask students to create a model of protein structure, using the plasticine, which relates to the word in the word bank. After 5 minutes, students show their model to another group. Ask one or two groups to demonstrate their model to the class and ask the class to suggest possible amendments or corrections if needed.

## Main teaching ideas

### 1 Modelling DNA replication with clothes pegs (30 minutes)

**Resources:** Per student group - four colours of clothes pegs with springs (at least eight of each colour), two lengths of 30 cm string, two clamps, two stands.

**Description:** The purpose of this activity is for students to visualise the process of semi-conservative DNA replication.

Tell students to work in pairs or threes. Ask them to set up the two clamps and stands just under 30 cm apart. Tell them they will create a model of DNA by suspending the two lengths of string next to each other, horizontally, from one clamp to the other and adding clothes pegs. When the string is set up, students should thread eight of their clothes pegs onto one of the strings by passing the string through the hole of the spring. This represents one strand of the DNA double helix. Students should create the second strand in a similar fashion, this time ensuring that complementary bases (represented by the different peg colours) match. The complementary pegs can be pegged together, representing hydrogen bonding.

Using the coursebook, students should read about the process of semi-conservative replication and use this to help them model the process with their pegs, taking particular care to understand the meaning of semiconservative and the importance of complementary base pairing in the process.

**Safety:** Care should be taken to ensure students set up heavy clamp stands carefully, so they do not fall over.

› **Assessment ideas:** Groups could volunteer to talk through their model to the rest of the class. The reminder of the class should be encouraged to volunteer any missing details or improvements.

› **Differentiation:** Less confident students might find it easier to watch a teacher demonstration of this model, with targeted questioning from the teacher to assess student understanding and opportunities for questions from students. You could challenge more confident Higher Level students to include a representation of the leading and lagging strands and the variety of enzymes involved, using additional classroom materials available (e.g. erasers, paperclips, pencil sharpeners).

› **Reflection:** Students could consider how modelling helps develop understanding and what limitations models might have.

## 2 Protein synthesis code cracking (30 minutes)

**Resources:** Beads of four different colours, string to thread the beads on to and an RNA codon table, which includes the one-letter abbreviations of each amino acid (projected on the board would work well).

**Description:** The purpose of this activity is to help students practice using a codon table as well as consolidate their understanding of transcription and translation. Provide students with an RNA genetic codon table, which includes the one-letter abbreviations representing each amino acid. Tell the students that the process of forming a protein with these amino acids is a bit like cracking a secret code, which is contained in DNA. They are going to model this. First, students should decide on the secret message they would like to encode to send to a classmate, by forming a short phrase using the amino acid one-letter abbreviations. They will need to be creative, as six letters of the alphabet are missing! Tell students that their secret phrase represents a protein made of amino acids (letters).

To encode their phrase, they need to produce a piece of DNA that contains the codons for their message. To do so they will thread different coloured beads onto a piece of string. Each different bead colour represents a different base (adenine, thymine, cytosine and guanine). Emphasise that the RNA codon table shows RNA codons, so these need to be converted into complementary DNA codons before the beads are added to the string.

When students have prepared their beads on a string (representing a gene), tell them they are now going to model transcription and translation. To do so, they should swap their 'genes' with a partner. First, students need to transcribe the DNA of the gene to determine the mRNA that would be formed (they could note this down). Then, they should use the codon table to translate the mRNA into the correct sequence of amino acids to reveal the secret message protein.

› **Assessment ideas:** Peer assessment would work well here. Students can confirm whether the phrase deduced is correct. If not, students could work together to identify the error in transcription or translation.

› **Differentiation:** Less confident students might benefit from being given pre-prepared 'genes' to transcribe and translate, rather than making their own. More confident students could experiment with errors (mutations) in the genes formed to determine their effects on the protein produced. This could lead to a discussion regarding the redundancy of the genetic code. More confident students could also explore the universality of the genetic code and the idea of leaving 'watermarks' in synthetic or genetically modified organisms DNA by preparing a short presentation on Craig Venter's use of the genetic code to produce a synthetic bacterium.

› **Reflection:** Students could explore their initial feelings after completing this activity. Are they mostly positive or negative? Why? What can they learn from this?

## 3 Nessa Carey's marshmallow nucleosomes model (20 minutes)

**Resources:** Per student group - 18 marshmallows, eight toothpicks, soft, rounded sweets e.g. Jelly Tots/ gumdrops, one long strawberry lace/pencil sweet.

**Description:** The purpose of this activity is to help students visualise the structure of nucleosomes and their role in regulating transcription in eukaryotes. Using the information in Chapter 2.5 of the coursebook, students should work in pairs or threes to create a model of two nucleosomes, using the marshmallows to represent histone proteins, toothpicks for the histone tails and the strawberry lace as a molecule of DNA

wrapped around the core histone structure. Students could then be provided with questions, such as those below, to encourage them to explore the functions of nucleosomes by using their model:

1 State what each part of your model represents.
2 As well as the core histone proteins making up the nucleosome, you should have a ninth histone protein on the outside of your nucleosome model. It is called the H1 histone protein. What is its function?
3 The histone proteins have tails that extend out and can be modified to alter gene activity. The tails of the core proteins can be modified so they link up. Model this using the gum drop sweets to represent the chemical modifications. What effect does this have on the DNA double helix?
4 Using your previous answer, state which part of the mitotic cell cycle the linking of histone protein tails might occur in.
5 Enzymes such as RNA polymerase require access to the DNA double helix for transcription to occur. Using your previous answers, state two ways in which nucleosomes could help to regulate which genes are transcribed in eukaryotes.

**Practical guidance**: Providing students with protective bench paper helps to keep desks clean.

**Safety:** Care must be taken by students when handling sharp toothpicks. No food should be consumed in a laboratory environment and sweets handled for model making should not be eaten.

**Answers:**
1 Marshmallows represent histone proteins, toothpicks represent the histone tails and the strawberry lace represents a molecule of DNA wrapped around the core histone structure.
2 The protein binds the DNA tightly to the nucleosome core.
3 The DNA would supercoil as the nucleosomes get pulled together.
4 Prophase of mitosis.
5 Nucleosomes can prevent enzymes accessing the DNA by supercoiling. To allow enzymes access to the DNA, the histone protein tails will disassociate and the H1 histone protein will be removed from the nucleosome. This will loosen the DNA around the core histone proteins.

› **Assessment ideas:** Hands-up or -down questioning could be used to answer the questions and students could mark their own work.

› **Differentiation:** More confident students could use online research to further develop their models, for example by modelling acetylation and methylation. Less confident students could build their models by following a step-by-step teacher demonstration, before attempting the questions.

### 4 Gel electrophoresis (60 minutes)

**Resources:** Various, depending on kit bought (see description below).

**Description:** The purpose of this activity is for students to carry out a gel electrophoresis practical, to visualise and understand the process. Gel electrophoresis kits for schools can be bought online although can be expensive. Alternatively, 'homemade' kits can be created using household materials. There are some online websites and videos outlining possible materials and instructions for use or those described in the journal 'Biochemistry and Molecular Biology Education', Vol. 40, No. 3, pp. 198-203, 2012, could be used. Whilst waiting for the electrophoresis gel to run, students could complete question 2 parts f and g from Exercise 4.1 of the workbook, to consolidate their understanding of gel electrophoresis.

**Practical guidance:** Gel electrophoresis moulds could be prepared in advance to ensure the practical runs efficiently.

**Safety:** Particular care should be taken as this practical involves the use of liquids near electricity.

› **Assessment ideas:** Self-assessment could be used here to mark the workbook questions.

› **Reflection:** Working individually, students could make a list of concepts they already knew that helped them understand the practical; new concepts they now understand; and concepts they are still struggling. They could create an action plan of next steps to further their understanding.

## Plenary ideas

### 1 Replication, transcription or translation? (10 minutes)

**Resources:** Statements about replication, transcription and translation (see PowerPoint 4, slide 3), writing equipment.

**Description:** The purpose of this activity is to determine whether students can distinguish between the processes of replication, transcription and translation and to address any misconceptions that are revealed. Ask students to work in pairs or threes to write down, for each statement, whether they think it applies to replication, transcription, translation or a combination of these.

**Assessment ideas:** Use hands-down questioning to go through the answers with students. For each student answer given, ask for a show of hands as to whether the class agrees or disagrees. This gives a clear idea of how many students in the class have got the answer right. If few students have obtained the correct answer, be sure to find out what the misconceptions are and address them.

### 2 PCR vs DNA replication (15 minutes)

**Resources:** Writing equipment.

**Description and purpose:** The purpose of this activity is to assess students' understanding of the processes of natural DNA replication and PCR. Students should work individually, in pairs or threes (depending on confidence) to construct an answer to the question: Compare and contrast the processes of natural DNA replication and PCR. Students could start constructing an answer from memory then use their class notes if they get stuck.

**Assessment ideas:** Use hands-up or -down questioning to ask students for similarities and differences between these processes. Students could self-mark their work as answers are volunteered. Incorrect responses can be explored through class discussion and any omitted points should also be addressed.



# 3.3. Mutations and 4.4 Epigenetics

## LEARNING PLAN

| Learning objectives | Success criteria |
|---|---|
| • Learn how new alleles form by mutation; changes may be neutral, harmful or beneficial. | Students will be able to state that new alleles are formed by mutation and that changes may be harmful, beneficial or neutral. |
| • Understand that mutations in germ cells can be passed to offspring. | Students will be able to state that mutations in germ cells can be passed to offspring. |
| • Discover how the cell cycle is usually strictly controlled. | Students will be able to state that the cell cycle is usually strictly controlled. |
| • Recognise that tumours may be benign or malignant. | Students will be able to summarise that tumours are a group of cells that grow out of control and may be benign or malignant. |
| • Define a mutagen as a substance that can cause genetic change. | Students will be able to define a mutagen and give some examples. |
| • Recognise that mutations can add, delete, invert or substitute bases in genes. | Students will be able to explain that mutations may be caused by addition, deletion or substitution of a base, or the inversion of a section of DNA, and that sickle cell disease (SCD) is caused by a substitution mutation. |
| • Learn that the genetic code is degenerate and so it is resistant to some changes caused by mutations. | Students will be able to describe how the genetic code is degenerate and explain how this gives it resilience to changes. |
| • Understand that DNA polymerases can make proofreading errors, and the errors remain permanently. | |

CONTINUED

| Learning objectives | Success criteria |
| --- | --- |
| Learn that some mutations can be corrected or repaired. | Students will be able to recall that DNA polymerase can proofread errors in new DNA strands and that mutations can be corrected or repaired. |
| Recognise that mutations do not always cause changes to a protein's function. | Students will be able to explain why mutations do not always cause changes to a protein's function, even if the amino acid is incorrect. |
| Define epigenetics as the study of changes to gene activation in differentiated cells. | Students will be able to define epigenetics. |
| Learn that gene expression is regulated by proteins that bind to base sequences in DNA. | Students will be able to explain how gene expression is regulated by binding proteins. |
| Understand that epigenetic changes modify the activation of certain genes but do not change their base sequences so that the phenotype will change but genotype does not. | Students will be able to state that epigenetic changes affect phenotype but not genotype. |
| Learn that DNA methylation inhibits transcription. | Students will be able to describe how epigenetic changes can be due to DNA methylation and modification of histones so that transcription is affected. |
| Understand that epigenetic changes are faster than changes caused by natural selection. | Students will be able to state that epigenetic changes are faster than changes caused by natural selection. |
| Recognise that environment has an impact on gene expression and can trigger heritable changes in epigenetic factors. | Students will be able to give an example of a heritable epigenetic change. |
| Learn that most epigenetic tags in gametes are removed from the embryo genome, but some remain and are inherited. | Students will be able to summarise how most epigenetic changes are not inherited but that some can affect the epigenome of the fetus. |
| Understand how environmental effects on DNA methylation can be studied using monozygotic twins. | Students will be able to describe the importance of monozygotic twins in the study of epigenetics. |

## Common misconceptions

| Misconception | How to identify | How to overcome |
|---|---|---|
| All mutations are harmful. | All of the misconceptions listed in this section could be identified by asking students to fill in a short rating scale questionnaire, where students indicate on a line whether they think the statements are always true (at one end of the line), always false (at another end of the line), sometimes true and sometimes false (in the middle of the line), or somewhere in between. | Explore a range of beneficial alleles in humans (lactose tolerance, HIV resistance, LRP5 amplification) with students and ask them how these alleles arose. |
| Mutations always result in a drastic changes to structures/ functions of the body. (Think Spiderman or X-men.) | | Students should be encouraged to link the idea of protein synthesis to genes, and then asked how a mutation in a gene would affect the body. |
| Mutations are unnatural. | | It can be helpful to situate the idea of mutations in the context of evolution for students to understand that they are a natural phenomenon. Without mutations, the natural process of evolution would not occur. |
| Mutations are only caused by radiation or chemicals. | | Ask students to consider how many times DNA replicates in a person's lifetime (lots and lots!). Ask students to imagine that they had to type out a sentence billions of times over and over again. How many mistakes might they make? This analogy might help them understand that mutations happen by chance. |

## Starter ideas

### 1 Mutated sentences (10 minutes)

**Resources:** PowerPoint 3, slide 4.

**Description and purpose:** The purpose of this starter is to encourage students to think about what are mutations and what their effects might be. A simple sentence, made of three letter words, is used to represent codons in the genetic code. Changes to this sentence are used to model single nucleotide mutations.

Ask students to discuss, in pairs or threes, what has happened to the original sentence in the PowerPoint for each mutation. They should also consider the effects that each mutation has had and whether they can link this to the idea of protein synthesis. Ask student volunteers to share their thoughts with the class.

### 2 Nature (genes) vs. Nurture (environment) Venn diagram (10 minutes)

**Resources:** PowerPoint 3, slide 8.

**Description and purpose:** The purpose of this starter is for students to explore ideas regarding factors that influence an organism's characteristics and to what extent these are determined by genes, the environment or both. Ask students to draw a Venn diagram with one circle for genes and another for environment. The overlapping area is for characteristics influenced by both. Ask students to sort the words on the presentation into the Venn diagram and compare their answers with their classmates then share ideas in a whole class discussion. This might highlight misconceptions that can then be addressed. It can also stimulate a discussion about epigenetics, or the fact that biological sex in some organisms is determined by the environment.

### 3. Retrieval practice grid (15 minutes)

**Resources:** PowerPoint 3, slide 7.

**Description and purpose:** The purpose of this activity is to help students assimilate key concepts into their long-term memory and to review important concepts for this topic. An example of a retrieval grid is included in the PowerPoint. You will need to customise this to your students, depending on when students learnt concepts. Tell students they will have 10 minutes to answer as many questions as they can. They get more points for correctly answering questions about topics covered further in the past. After 10 minutes, go through the answers with students and ask them to add up their scores. Ask for a show of hands for students who achieved more than 20 points. Address any misconceptions encountered as students answer questions and review any forgotten concepts.

### 4. Word association game (10 minutes)

**Description and purpose:** The purpose of this activity is to elucidate students' prior knowledge and uncover potential misconceptions. Ask a student to state a biology key term they associate with the word 'gene'. Then, ask the next student to associate a biological key word with that just stated by their peer. Continue until all students have stated an associated biological key term. If terms stated by students seem dubious, ask them to justify their answer. Address any misconceptions as they are encountered.

› **Language focus:** Students become familiar with key terms and definitions associated with chromosomes, genes and inheritance.

## Main teaching ideas

### 1 Mutations broken telephone game (20 minutes)

**Resources:** Depending on version of game: mini whiteboards and pens for drawings, PowerPoint 4, slides 5 and 6, enough cards for each student in the class to have one card. They should be labelled: no mutation, 90% of the set of cards; inversion, one or two cards; deletion, one or two cards; addition, one or two cards; substitution, one or two cards.

**Description and purpose:** The purpose of this activity is to help students visualise the difference between mutations which add, delete, substitute or invert bases in genes. It can also be used to simulate the effect of mutagens and DNA repair mechanisms (see Differentiation section).

Students should form groups of approximately 10 and stand in single file, with enough space between them so that they can only easily hear the students stood next to them (alternatively students can whisper to each other). Explain that the teacher is a section of DNA in a cell, called a gene, containing the code to instruct the cell how to make the protein. Explain that when cells divide to form more cells, DNA molecules (and therefore genes) are copied and passed on to the new cells. Tell the students they are going to model this process of replicating the code for proteins, with each student representing a new cell containing new copies of DNA (formed by semi-conservative replication). The teacher passes the code to the first student in the line of 10 by whispering it to the student (or handing it to them on a piece of paper), simulating the first round of cell division and DNA replication. The code could be a fairly long sentence to be repeated out loud, a set of actions (e.g. pat head twice, hop on one leg, clap hands three times) or a drawing (e.g. a row of 10 shapes: triangle, triangle, square, circle, etc.). The code should be whispered from student to student until the last student is reached, without any repetition. Start with the last student and work back to the first student, each student should show the 'protein' they have been told to make by the code passed to them (e.g. by carrying out the actions suggested). The class should watch carefully and identify any differences between the 'proteins' being made. A discussion can then be encouraged, with students explaining why any observed differences occurred (linking this to the idea of mutations) and what types of differences they are (was a step added, deleted or substituted?). If no differences occur, this could be an ideal opportunity to reinforce the role of complementary base pairing in DNA replication before repeating the game (maybe with some of the modifications from the Differentiation section).

› **Assessment ideas:** Students could be shown images of different types of mutation, such as those in the PowerPoint, to identify.

› **Differentiation:** Groups of more confident students could be given cards labelled no mutation: 90% of the set of cards; inversion: one or two cards; deletion: one or two cards; addition: one or two cards; or substitution: one or two cards. Students should modify the code (or not) according to the card they

have been given. When students 'show' their 'proteins' at the end of the activity, the class should try to guess which student has which type of card. More confident classes could also repeat the simulation with background noise (e.g. a song playing in the background) to demonstrate the effect of mutagens on the likelihood of mutations occurring. Higher Level students could also simulate DNA repair mechanisms by allowing repetitions between whispers.

› **Reflection:** Students could reflect on the role they played in facilitating their learning as well as their peers' during this activity. Did they explain something to someone else? Did they encourage their peers to ask questions (even if they knew the answer themselves)? Did they listen attentively to their peers, as well as the teacher, so that they could build on others' discussion points? How did their actions promote a productive learning atmosphere and how can these be applied (or improved) in future activities?

## 2 The effects of mutations ( 30 minutes)

**Resources:** PowerPoint 3, slides 7 and 8, RNA codon table, pot/cup of coloured counters.

**Description:** The purpose of this activity is to help students understand that mutations can be harmful, beneficial or neutral (linking the latter to the degeneracy of the genetic code).

Part 1: Show students the DNA sequences from the PowerPoint. Use the original base sequence to remind students how to determine the mRNA codons formed during transcription and the amino acids coded for, using the RNA codon table. The transcription and translation of this first sequence could be done as a class, with hands-down questioning to ensure students understand the concept. Then, ask students to work in pairs to transcribe and translate the remaining two, mutated sequences. Students should compare the amino acid sequences produced and work together to explain why one mutation caused a change in the amino acid sequence whilst the other did not. Use this as an opportunity to introduce the term 'degeneracy'.

Part 2: Tell students that they are going to model the effects of mutations on organisms. For each round of the model, students should each select a coloured counter from the pot at random. These are variations of a protein for a similar trait that have been created by mutations in DNA. For each round, describe a particular environment and the type of protein being produced (e.g. hot, desert environment; protein which ...). Then, list the qualities of proteins produced by going through each colour (e.g. red, no change in amino acid sequence; blue, change in amino acid leads to a protein for thick fur; green, change in amino acid leads to a protein for thin fur, etc.). Students need to decide whether the mutation has been neutral, beneficial or harmful to them in the environment described. Students who have a harmful mutation could sit down, those with neutral mutations could stay standing and those with advantageous mutations could put their hands on their heads. After a few rounds of the simulation, invite students to explain why base substitution mutations can have a significant impact on a living organism, asking them to refer to protein structure and function.

› **Assessment ideas:** Hands-up and -down questioning throughout these activities should help assess student understanding.

› **Differentiation:** Less confident students might benefit from reviewing protein synthesis before attempting this task.

› **Reflection:** Ask students what they found most challenging about this activity, why they found that particular aspect challenging and what they can learn from this.

## 3 Gene therapy research (60 minutes)

**Resources:** internet access.

**Description:** The purpose of this activity is for students to develop their understanding of gene therapy. Tell students to imagine they have a recessive genetic disease, such as cystic fibrosis and that they have been offered the chance to participate in clinical trials for gene therapy for this. They need to carry out research to decide whether this is something they would like to go ahead with. Invite students to use the coursebook and the internet to find out more about gene therapy, using the prompts below, and to make notes on their findings. After 30-40 minutes of research, engage students in a class discussion to divulge their findings and decide whether they would participate in the proposed clinical trials, justifying their decisions.

**Research prompts:**

- Find a definition of gene therapy.
- Describe the types of genetic condition that could be treated.
- Outline where the genetic material being introduced is delivered (there are different techniques).

- Outline different methods of delivering the genetic material, including references to retroviruses and adenoviruses. Are there any advantages or disadvantages of different techniques?
- In what situations is gene therapy temporary or permanent?
- To what extent are gene therapy techniques ready for medical use?

› **Assessment ideas:** Students could produce a summary sheet of their findings for the teacher to collect in and mark.

› **Differentiation:** Less confident students might like to work in small groups to divide the research questions between them and then share their findings. More confident students could research CRISPR gene editing techniques.

› **Reflection:** Students could reflect on what they found challenging in the research process. Why was it challenging? What could they change about their approach in the future?

## 4 Using databases (30 minutes)

**Resources:** Internet access.

**Description:** The purpose of this activity is to allow students to determine different base sequences in a gene in two species. This is a syllabus-specific skill.

Ask students to access the United States National Center for Biological Information (NCBI) GenBank website (see Links to digital resources section) and explain that this is a publicly available database of sequences from several sources. Tell students that they are going to find a gene sequence for a human protein of their choice (e.g. lactase) and find out how much this differs from the gene sequence for the same protein in another species. Ask students to use the search bar to enter the name of their chosen protein, including the species it is found in - *Homo sapiens* (the drop-down menu should say 'nucleotide'). Students should select the result that most closely matches what they are searching for. It can help to check the descriptions to ensure the result gives a DNA (rather than mRNA) sequence. Students might also need to be reminded that some proteins with quaternary structure consist of more than one polypeptide chain, so may be coded for by more than one gene and genes for proteins may have different alleles. Once they have selected their protein, ask students to find the FASTA information for the protein on the webpage loaded. This is a text-based format that represents nucleotide (or amino acid) sequences. Students should be able to copy this text. Now that they have their gene sequence, tell students they are going to use a website called BLAST to compare this sequence with those of other species. Ask students to access the NCBI BLAST website and select 'nucleotide blast'. Here, they should be able to paste their copied sequence into a textbox. When they click 'BLAST', the database searches for similar sequences. Students can choose species' sequences to compare by checking the relevant tick boxes and clicking 'view alignments'. Encourage students to spot the percentage similarity between their selected species and analyse the sequences to identify different types of mutation (e.g. substitution, deletion or addition).

› **Assessment ideas:** Students could take a screenshot of the alignments of two species and annotate it to show the key differences between them, to be marked by their teacher.

› **Reflection:** Students could make a list of strengths and weaknesses with regards to the IT skills. What did they already know that helped them with this task and what did they learn that they didn't know before?

## 5 Epigenetics paper model (20 minutes)

**Resources:** Cut-out sheets from the Teach Genetics Utah website (see Links to digital resources section), paperclips, scissors, audio-visual facilities.

**Description and purpose:** This activity allows students to visualise the effects of acetylation and methylation on histones and, consequently, gene expression.

Provide students with the cut-out sheets, scissors and paper clips. Show them the tutorial video (see Links to digital resources section), pausing after each step so that students can follow along.

**Safety:** Ensure scissors are classroom safety scissors.

› **Assessment ideas:** Students could complete Exercise 4.4 from the workbook and self-mark their answers in class.

› **Reflection:** Ask students to consider what they have learnt today in terms of their day to day lives. Will the lesson change anything for them? Why/why not?

## Plenary ideas

### 1 Memory mission (10 minutes)

**Description:** The purpose of this activity is to encourage students to memorise key concepts of the lesson. The first student in the class states one fact they have learnt. The next student repeats the first student's fact, then gives a fact of their own. The third student repeats the first two facts then adds a fact of their own. This continues until the last student is reached (they may need a little help from their peer's especially if it is a large class!). Listen carefully whilst students state their facts, in case any misconceptions arise that need to be addressed.

### 2 Just a minute (10 minutes)

**Resources:** Timer.

**Description and purpose:** The purpose of this activity is to encourage students to speak confidently about recently covered concepts. Set the timer to 1 minute. Name a key term (e.g. genomes) and ask a student volunteer to start speaking about this. They cannot say the key term, repeat themselves, say anything that is biologically incorrect or hesitate in their speech, otherwise the play passes to another student. Other students should listen carefully and raise their hands if they think the student volunteer has made a mistake. Whoever is speaking when the timer reaches zero is the winner!

› **Language focus:** Students become familiar with key terms and definitions associated with mutations and epigenetics and speaking fluently about them.

### 3. Pictionary (10 minutes)

**Resources:** Pens/pencils and sheet of paper or mini whiteboards with pens and sponges.

**Description:** The purpose of this activity is to allow students to review key terms for the topic.

Divide students into groups of three or four. Ensure each group has something to draw with and something to draw on. Tell students you have a list of 10 words that students need to work through. The first team to reach word number 10 wins. Ask one student from each group to meet you at the front of the classroom. Tell the class that you will tell these students the word in secret. Then, they will return to their group and try to communicate their word to their team by drawing only (no speaking, miming, writing letters or any other form of communication allowed). Once their teammates have guessed the word, they should send a different student from their group to tell you the word in exchange for the next one in the list.

› **Language focus:** Students become familiar with key terms and definitions associated with genes, transcription and translation.

## Assessment ideas

- To determine the extent to which they have a detailed understanding of the processes of DNA replication, transcription and translation, students could be asked to fill in a Venn diagram, made of three circles to represent each process. This could be given as a very open-ended task.
- To assess students' understanding of epigenetics and mutations, they could be asked to explain reasons why identical twins' proteomes might differ more as they grow older.
- Students could assess their understanding using the Test your understanding questions from the coursebook.
- Students can complete the self-assessment grid from the coursebook to indicate their confidence to themselves and their teacher. Categories in the 'needs work' and 'maybe' sections should be reviewed by students and/or re-taught by the teacher.
- The teacher could assess student understanding using the Exam-style questions from the coursebook and workbook.
- There are test questions available at the end of this chapter.

# Differentiation

## Stretch and challenge

- Students could be asked to create an analogy for protein synthesis to explain the process to younger students.
- Students could read the coursebook information about Meselson and Stahl then use the information to draw the bands for centrifuge tubes for the first three generations of bacteria.
- Students could use research and apply mathematical skills to determine why codons need to consist of three rather than two bases.
- Student could read about the work of Leonard Hayflick, an American scientist who discovered that the number of times a cell can divided is limited. Hayflick worked on cell cloning for the production of vaccines.
- 'Code of a Killer' is an excellent, three-part British police drama TV series that explores Alec Jeffrey's discovery and application of DNA fingerprinting. Students might enjoy watching this. (UK Certificate age rating: 15.)
- Students could read Nessa Carey's 'Epigenetic Revolution' popular science book.
- Students could find out about chromosomal translocations, including balanced chromosomal translocations, and produce a visual explanation of this phenomenon as a poster.
- Students could research beneficial mutations in humans (e.g. apolipoprotein A-$I_{Milano}$, amplifying LRP5, tetrachromatic vision, resistance to HIV).
- Students could read about microproteins and find out why they are a relatively new discovery.

## Support

- Students might benefit from being given an analogy (see below) to make the abstract process of protein synthesis more tangible. They could answer comprehension questions using this analogy. For example (Q1) Use the analogy to help you put the following in order of size: nucleus, gene, DNA. (Q2) Translation is like a chef making a cake. Use this analogy to help you write a short definition of transcription. (Q3) Use the analogy to help you explain how proteins are made:
  - The nucleus is like a library.
  - DNA is like a master copy of a cookbook.
  - A gene is like a printed recipe for a specific cake.
  - mRNA is like a photocopied recipe.
  - Ribosomes are like chefs.
  - Proteins are like cakes
- Much of this topic relies on students having a strong understanding of DNA structure, transcription and translation. Students might find it helpful to have visual reminders of these processes up on walls to refer to.

  This abstract topic can be difficult for students to visualise and therefore understand. Analogies, models and diagrams can all help students develop their understanding.

  Students will find it useful to practise Punnett squares and pedigree chart questions, so provide them with as many exercises as possible.

## Links to digital resources

- RNA codon table

  There are many websites that have examples of RNA codon table. For an example, search for keywords: RNA codon table.

- Animations of DNA replication and protein synthesis

  Animated videos of these processes are extremely helpful in developing student understanding. There is a huge variety of online resources for this. For example, in the 'Resources' Biology Animations section of the DNA Learning Centre website or the 'Videos' section of the Your Genome website. Search for keywords: DNA replication video; Transcription video; Translation video

- Nucleosome marshmallow model

  There are a few videos online that demonstrate this process.

  Search for keywords: Nessa Carey marshmallow nucleosome epigenetics

- Electrophoresis kits

  These can be ordered online through several websites. There are also websites and videos that explain how to make gel electrophoresis kits.

  Useful search terms: online order electrophoresis kit; homemade DIY electrophoresis kit

- Mutations video introduction

  There are many online videos that provide helpful visual introductions to the concept of mutations.

  For example, a good video is available from the Brightstorm website by searching their site for 'mutation' and click on the 'AP Biology' video.

  Search for keywords: mutations video

- Gene databases

  There are some databases that contain publicly available gene sequences. NCBI's GenBank is a useful example.

  Search for keywords: NCBI GenBank

### CROSS-CURRICULAR LINKS

- How is the molecular structure of DNA linked to its function? (Chapter 1)
- Why must the genetic code carried by DNA be copied exactly? (Chapter 3.4)
- How is replication involved in cell division? (Chapter 6)
- How does the variety of proteins produced contribute to the functioning of a cell? (Chapter 6)
- How does the degenerate genetic code protect a cell against mutations? (Chapter 3.3)
- How can mutations be beneficial for the survival of a species?
- How do mutations occur during cell division?
- Why is important that not all mutations are corrected?
- How does DNA fit into the tiny volume of a nucleus? (FF 1.5)
- Why is compartmentalisation in cells important? (FF 2.1)
- How does chromosome behaviour during cell division lead to variation? (Chapter 6)
- What other factors affect the inheritance of genes? (Section 4.2) (CC 3.2)
- What other types of mutation are there? (CC 1.2 and CC 1.3)

## Homework ideas

1 Preparing for an extended response

Ask students to use the coursebook and their class notes to revise the process of either DNA replication, transcription or translation in preparation for a timed, written extended response question at the start of the next lesson. Students can self-mark this in class with a mark scheme projected on the board. Repeat until each process has been revised. Breaking up the revision of this topic can really help students consolidate their understanding and help them distinguish between the processes.

**2** Smoking factsheet

Students could be asked to produce an A4 factsheet to discuss the health implications of smoking e-cigarettes. Students should include:

- References to nicotine as a mutagen.
- An overview of tumour formation and the cell cycle.
- A balanced argument.

**3** Revising for a vocabulary quiz

Students could revise the spellings and definitions of the words in this chapter in preparation for a quiz the following lesson.

**4** Coursebook questions

Read Chapter 3 sections and answer the Test your understanding questions.

Name _____ Date _____

# End of Chapter 3 test

*This test and its sample answers have been written by the author. IB may award marks differently.*

**1** By which process does DNA replicate?

**A** Conservative replication
**B** Dispersive replication
**C** Semi-conservative replication
**D** Hemi-conservative replication **[1]**

**2** Which of the following enzymes is used in PCR?

**A** DNA ligase
**B** *Taq* polymerase
**C** DNA polymerase III
**D** DNA helicase **[1]**

**3** How are DNA profiles formed?

**A** By separating DNA fragments by length
**B** By sequencing DNA
**C** By mutating DNA
**D** Using DNA databases **[1]**

**4** Which sequence correctly shows the mRNA molecule produced from the DNA sequence below?

**TAC TAG AGT AAC CAT ATT**

**A** ATG ATC TCA TTG GTA TAA
**B** AUG AUC UCA UUG GUA UAA
**C** UAG UAC ACU AAG GAU AUU
**D** TAC TAG AGT AAC CAT ATT **[1]**

**5** What are functions of nucleosomes?

**I** Regulating transcription **II** Packaging DNA **III** Translating DNA

**A** **I** and **II**
**B** **II** and **III**
**C** **I** and **III**
**D** **II** and **III** **[1]**

**6** Where are anticodons located?

**A** On mRNA

**B** On DNA

**C** On tRNA

**D** On amino acids **[1]**

**7** Identify the type of mutation shown below:

Original sequence: TAC GTC GTC ATG CAT ATT

Mutated sequence: TAC GTC ATG CAT ATT

**A** Deletion

**B** Inversion

**C** Addition

**D** Substitution **[1]**

**8** Which of the following can be used as a vector in gene therapy?

**A** Syringe

**B** Virus

**C** Radiation

**D** Ribosome **[1]**

**9** What is the purpose of gene 'knockout'

To determine the phenotype produced a specific gene

**I** To remove a harmful mutation

**II** To damage specific genes so they no longer function

**V** To determine the base sequence of a gene

**A** **I**, and **II**

**B** **I**, **II** and **II**

**C** **I** and **III**

**D** **I**, **II**, **III** and **V** **[1]**

**10** What is the role of DNA primase in DNA replication?

**A** To add an RNA primer.

**B** To add a DNA primer.

**C** To remove an RNA primer

**D** To remove a DNA primer. **[1]**

**11** Outline alternative splicing.

**[2]**

**12** Outline two roles of DNA, aside from coding for proteins.

**[2]**

**13** Explain why translation does not always result in a functional protein.

**[4]**

**14** Explain how epigenetic factors can affect gene expression.

**[5]**

**15** Explain why a base substitution mutation is usually less harmful than a deletion mutation.

**[2]**

**16** Explain why mRNA used for translation is a different length from that created in the nucleus during transcription

**[2]**



**END OF TEST**

# Chapter 3

## DNA and protein synthesis

# Replication, transcription or translation?

For each statement write down whether it applies to replication, transcription or translation or a combination of these.

- A semi-conservative process.
- Involves a polymerase enzyme.
- Involves DNA gyrase.
- Involves complementary base pairing.
- Results in the formation of Okazaki fragments.
- Requires a ribosome.
- Involves RNA.
- Involves the formation of a nucleotide strand in a 5′ to 3′ direction.
- Can be prevented by methylation of DNA.
- Consists of initiation, elongation, translocation and termination.

# Mutated sentences

## What has happened to the original sentence for each mutation?

**Original sentence:** The big fat cat ran off his mat.

**First mutation:** The big fat rat ran off his mat.

**Second mutation:** Thb igf atc atr ano ffh ism at.

**Third mutation:** The bbi gfa tca tra nof fhi sma t.

**Fourth mutation:** The big fat tac ran off his mat.

# Mutations broken telephone game

## Types of mutation: Are these deletions, additions, substitutions or inversions?

**Original DNA sequence:** TAC GGT GCA TTG ACT ACT ACT ATC

**Mutation 1:** TAC GGT GCA TTG ACT ACT ACT ACT ACT ATC

**Mutation 2:** TAC GGT GCA ACT ACT ACT ATC

**Mutation 3:** TAC GGT GCA GTG ACT ACT ACT ATC

**Mutation 4:** TAC GTT ACG TGG ACT ACT ACT ATC

# Mutations broken telephone game

## Types of mutation: Answers

**Original DNA sequence:** TAC GGT GCA TTG ACT ACT ACT ATC

**Mutation 1:** TAC GGT GCA TTG ACT ACT ACT ACT ACT ATC – Addition

**Mutation 2:** TAC GGT GCA --- ACT ACT ACT ATC – Deletion

**Mutation 3:** TAC GGT GCA GTG ACT ACT ACT ATC – Substitution

**Mutation 4:** TAC GTT ACG TGG ACT ACT ACT ATC – Inversion

# The effects of mutations

## Transcribe and translate the following base sequences.

**Base substitution:** A mutation in which there is an insertion of the incorrect nucleotide in the base sequence of a gene.

**e.g. Original DNA sequence:** TAC GTC GAG CAT TTC TAG ATT

**Base substitution 1:** TAC GTC GAG CAT TTC TAA ATT

**Base substitution 2:** TAC GTC GAG CAT TTC TAC ATT

# The effects of mutations

## Answers

**Original sequence:**
mRNA: AUG CAG CUC GUA UUG AUC UAA
Polypeptide: Met – Glutamine – Leucine – Valine – Leucine – Isoleucine – Stop

**Substitution 1:**
mRNA: AUG CAG CUC GUA UUG AUU UAA
Polypeptide: Met – Glutamine – Leucine – Valine – Leucine – Isoleucine – Stop

**Substitution 2:**
mRNA: AUG CAG CUC GUA UUG AUG UAA
Polypeptide: Met – Glutamine – Leucine – Valine – Leucine – Methionine – Stop

# Retrieval practice grid

## How many of these questions can you answer in 10 minutes?

| | | | |
|---|---|---|---|
| Describe the arrangement of chromosomes in metaphase 1 of meiosis. | Describe the structure of a nucleosome. | What is a gene? | Which cellular structure synthesizes protein? |
| What is formed in the process of transcription? | Name the bond that connects two amino acids. | What is formed during the process of translation? | What is a zygote? |
| In what stage of the cell cycle does DNA replicate? | Define the term epigenetic. | What is an allele? | Explain the degeneracy of the genetic code. |

| | | | |
|---|---|---|---|
| Last lesson (1 point) | Last week (2 points) | Two weeks ago (3 points) | Further back (4 points) |

# Nature vs Nurture

Draw a Venn diagram with one circle for genes overlapping a circle for the environment. Sort the words below into your Venn diagram.

| | |
|---|---|
| Eye colour | Skin colour |
| Biological sex | Weight |
| Blood group | Accent |
| Hair colour | Height |

Name _____ Date _____

# 2.1 Definitions – true or false?

There are many of important key terms in Biology. Make sure you know what they mean so that you can use them correctly.

For each of the definitions below, state whether it is true or false. If it is false, write the correct definition out. Can you add any more important key terms and definitions?

| Key term | Proposed definition | True or false? Remember to correct the false ones |
|---|---|---|
| Chromatid | The structure into which DNA is packaged. | |
| Chromosome | A molecule of replicated DNA which makes up part of a chromosome. | |
| Codon | A sequence of three mRNA bases that code for an amino acid. | |
| Epigenome | All the chemical compounds that have been added to a genome to regulate the expression of all the genes within the genome. | |
| Gene | A section of DNA that makes a specific amino acid. | |
| Genome | All of the genes in a cell or organism. | |
| Homozygous | Describes a pair of chromosomes of the same length with genes at the same loci. | |
| Lagging strand | The new strand that is synthesised in short fragments in the opposite direction to the movement of the replication fork. | |
| Leading strand | The new strand that is synthesised continuously and follows the replication fork. | |
| Oncogene | A gene that leads to cancer. | |
| PCR | A process in which small quantities of DNA are artificially amplified for research and diagnosis. | |

| Key term | Proposed definition | True or false? Remember to correct the false ones |
|---|---|---|
| Gene knockout | A method that is used to damage specific genes so that they no longer function | |
| Transcriptome | All the mRNA molecules expressed from the genes of an organism | |
| Proteome | the complete set of proteins expressed by an organism | |
| Gene expression | the mechanism by which genetic information affects the phenotype of an organism | |