SCHOLAR Study Guide

Higher Biology
Unit 1: DNA and the Genome

Authored by:

Bryony Clutton (North Berwick High School)

Reviewed by:

Fiona Stewart (Perth Grammar School)

Previously authored by:

Jaquie Burt

Eileen Humphrey

Lorraine Knight

Nadine Randle

Fergus Forsyth

Patrick Hartie

Heriot-Watt University

Edinburgh EH14 4AS, United Kingdom.

First published 2018 by Heriot-Watt University.

This edition published in 2018 by Heriot-Watt University SCHOLAR.

Distributed by the SCHOLAR Forum.

SCHOLAR Study Guide Higher Biology: Unit 1

Higher Biology Course Code: C807 76

ISBN 978-1-911057-36-9

Print Production and Fulfilment in UK by Print Trail www.printtrail.com

Acknowledgements

Thanks are due to the members of Heriot-Watt University's SCHOLAR team who planned and created these materials, and to the many colleagues who reviewed the content.

We would like to acknowledge the assistance of the education authorities, colleges, teachers and students who contributed to the SCHOLAR programme and who evaluated these materials.

Grateful acknowledgement is made for permission to use the following material in the SCHOLAR programme:

The Scottish Qualifications Authority for permission to use Past Papers assessments.

The Scottish Government for financial support.

The content of this Study Guide is aligned to the Scottish Qualifications Authority (SQA) curriculum.

Contents

Topic 1

Structure and organisation of DNA

Contents

Prerequisites

You should already know that:

- DNA takes the form of a double-stranded helix;
- the two strands of DNA are held together by complementary base pairs;
- DNA contains the four bases A, T, G and C which make up the genetic code.

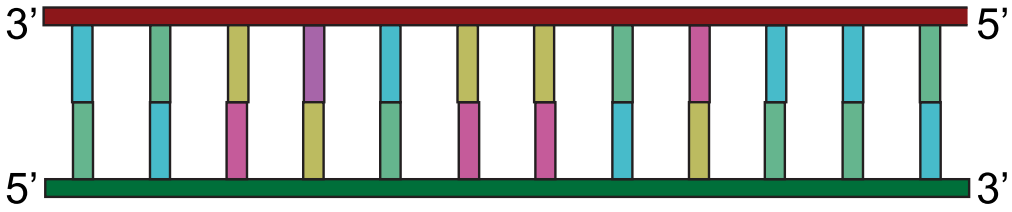
Learning objective

By the end of this topic, you should be able to:

- appreciate that DNA is found in all living organisms and is the chemical which carries hereditary information;
- understand that the sequence of bases within DNA is the genetic code;
- appreciate that the sum total of all the DNA bases is the genome;
- describe the structure of a nucleotide;
- describe the structure of DNA;
- describe how nucleotides combine to form a backbone and that base pairing holds the two strands together, forming a double helix;
- describe the base pairing rule;
- name the bonds which hold bases together;
- explain that the two strands of DNA lie anti-parallel to each other and are read in different directions;
- understand that DNA can be organised in a number of ways, either circular or linear;
- state that circular forms of DNA can be found in prokaryotes, mitochondria and chloroplasts;
- state that bacteria and yeast also contain smaller rings of DNA, called plasmids;
- state that the linear form of DNA found in eukaryotes is tightly packaged with associated proteins called histones.

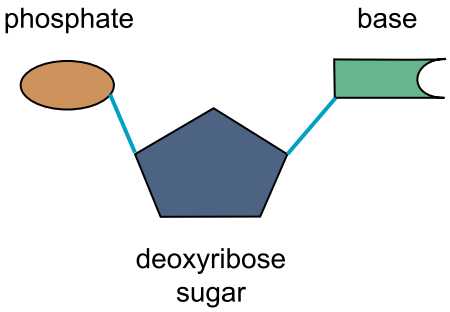
1.1 The structure of DNA

In 1953, the structure of the DNA molecule was explained for the first time by two scientists, James Watson and Francis Crick, at the Cavendish Laboratory in Cambridge.



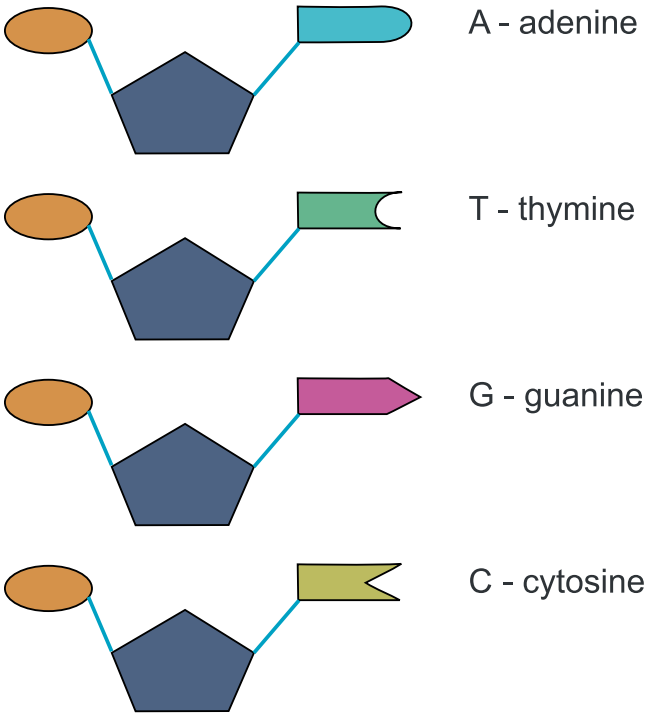
Double helix structure of DNA

They found that a molecule of DNA consists of two strands of repeating units called nucleotides. Nucleotides are composed of phosphate, deoxyribose sugar and a base.



A DNA nucleotide

There are four bases in DNA: (A) adenine, (T) thymine, (G) guanine and (C) cytosine. Each nucleotide has a different base.



Nucleotide bases

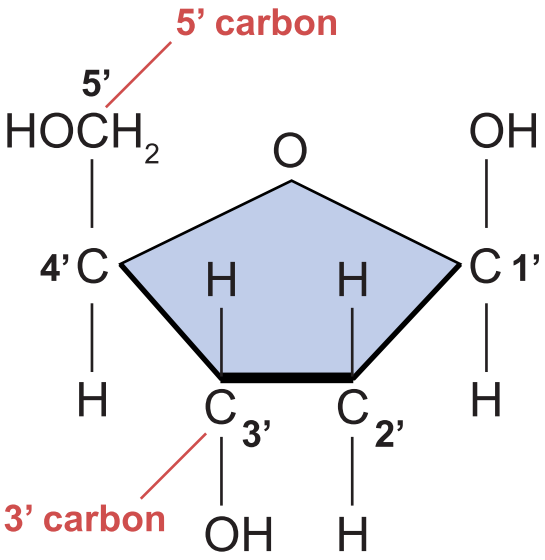
These bases follow the base pairing rules: A always pairs with T, and G always pairs with C. The bases are held together by hydrogen bonds.

DNA nucleotide		DNA nucleotide
A	pairs with	T
T	pairs with	A
G	pairs with	C
C	pairs with	G

Base pairing rules

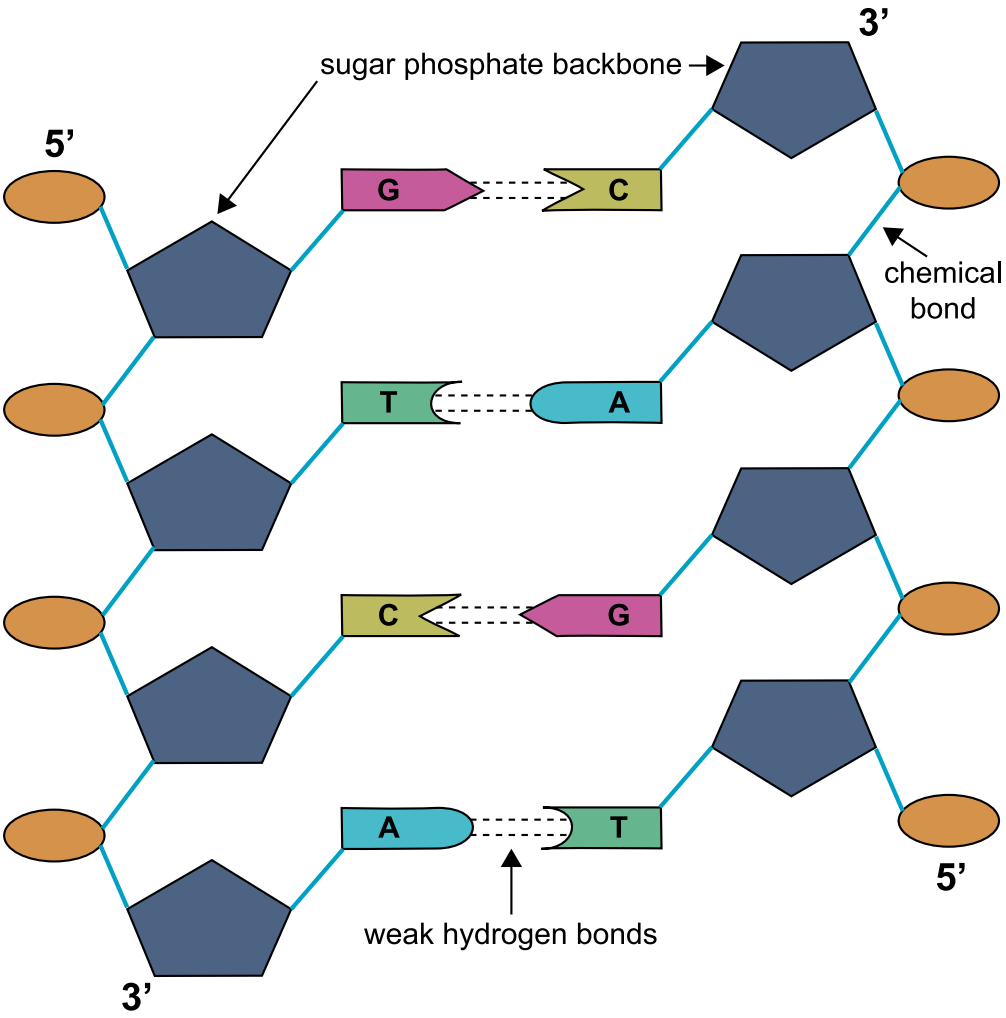
The phosphate of one nucleotide is attached by a strong chemical bond to the deoxyribose sugar of the next. This forms the sugar-phosphate backbone of the DNA.

DNA takes the form of a **double helix**. DNA is made up of two **antiparallel** strands. This means the strands run in opposite directions to each other. The carbon atoms on the deoxyribose sugar are numbered as shown below.



Deoxyribose sugar with carbon atoms numbered

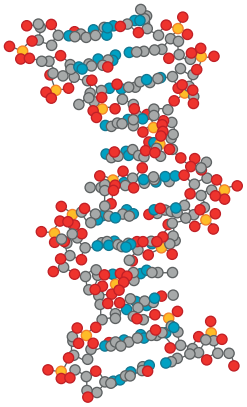
Using this system, each end of the DNA can be labelled to show the antiparallel strands. The structure of DNA is shown as follows.



DNA structure

The structure of DNA: Interactive 3D model Go online

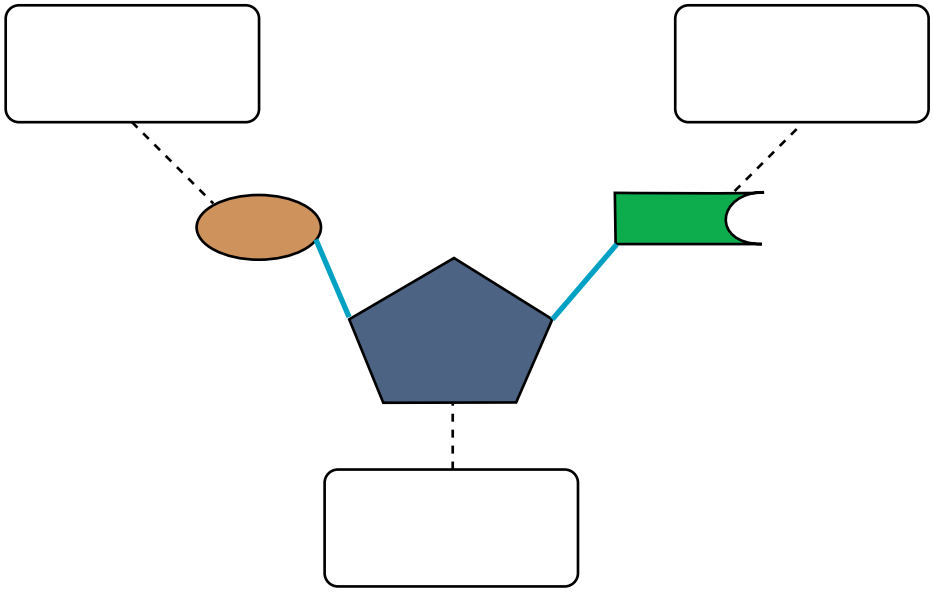
An activity that shows the double helix structure of DNA as an interactive 3D model is available in the online materials at this point. The following illustration gives an idea of what to expect.



Double helix structure of DNA - molecular model

The structure of DNA: Questions Go online

Q1: Complete the diagram using the words from the list.



Word list: base, deoxyribose sugar, phosphate.

. . . . . .

Q2: Complete the blanks using the base pairing rules.

- The nucleotide guanine pairs with _____.
- The nucleotide thymine pairs with _____.
- The nucleotide cytosine pairs with _____.
- The nucleotide adenine pairs with _____.

Q3: What shape is the DNA molecule?

. . . . . .

Q4: What type of bonding holds two DNA strands together?

. . . . . .

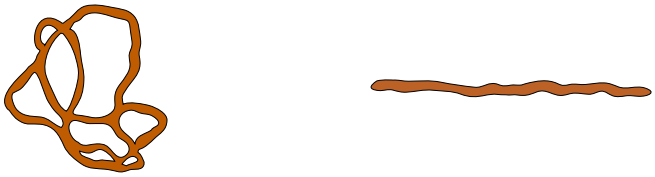
Q5: Which three components make up a DNA nucleotide?

1.2 The organisation of DNA in prokaryotes and eukaryotes

Prokaryotes are organisms which lack a true membrane-bound nucleus. Bacteria are an example of a prokaryote. Their DNA is found in the cytoplasm of the cell.

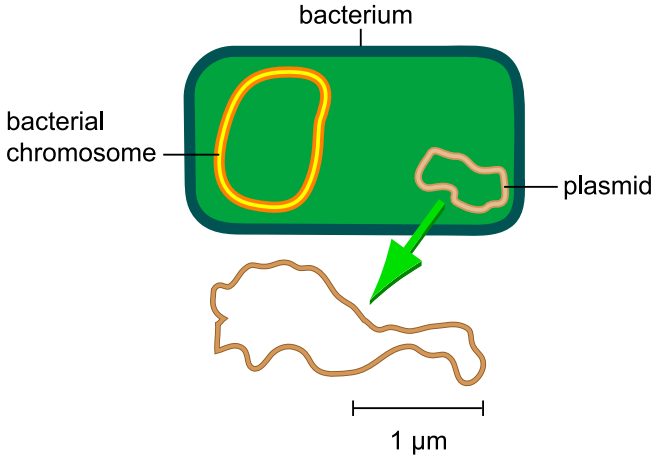
Eukaryotes are organisms which have a membrane bound nucleus that stores their genetic material. Animals, plants and fungi are examples of eukaryotes.

DNA is a double-stranded molecule that can either be circular or linear.



Circular and linear DNA

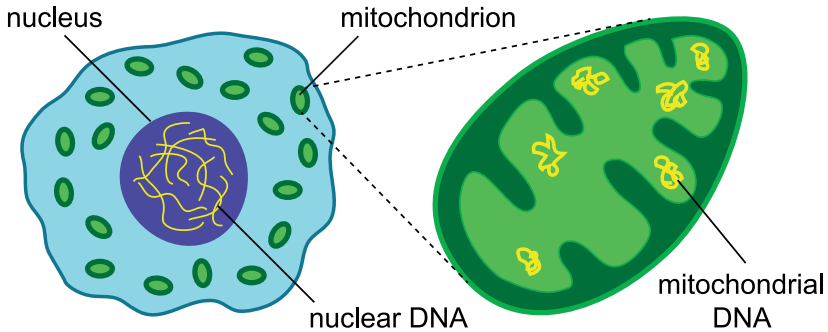
Prokaryotes have a large circular chromosome. They may also have smaller rings of DNA called **plasmids**.



A bacterial (prokaryotic) cell

Circular plasmids may also be found in yeast (a fungus), which is classified as a eukaryote.

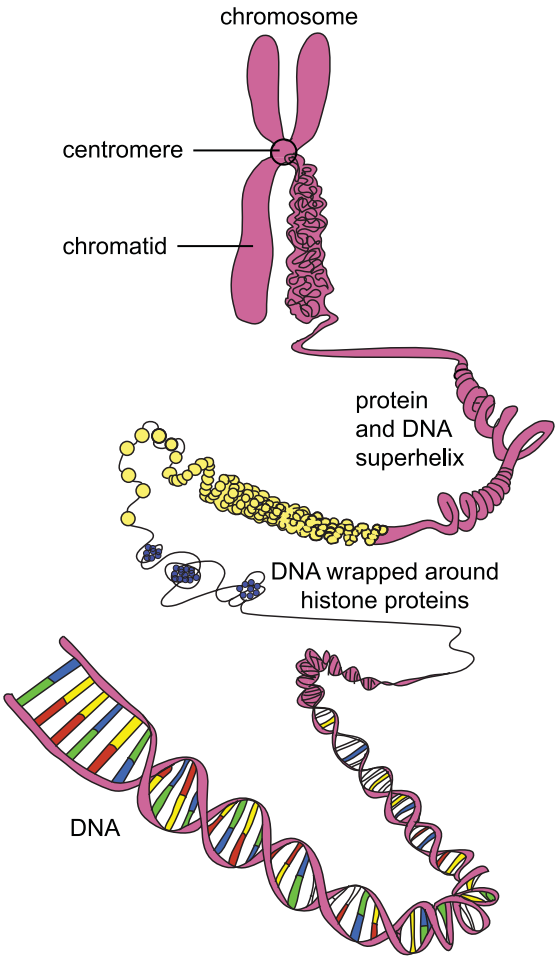
In eukaryotes, DNA is found tightly coiled into linear chromosomes. DNA is also found within mitochondria (mtDNA) where it forms circular chromosomes. It is sometimes described as the smallest chromosome and is inherited from the mother in humans.



A eukaryotic cell

DNA can also be found within the chloroplasts of plant cells. It is usually larger than mitochondrial DNA and takes the form of circular chromosomes containing the genes involved in the photosynthetic process. Where circular DNA is found in eukaryotes, it is thought that it has been incorporated from early bacteria or prokaryotes.

Typically, the DNA content of a single human cell, if completely unravelled, would measure around two metres in length. This DNA must be packaged so it can fit inside the nucleus. DNA found in the linear chromosomes of the nucleus of eukaryotes is tightly coiled and packaged with associated proteins called histones. This is shown in the diagram below.



DNA packaging

The organisation of DNA in prokaryotes and eukaryotes: Questions Go online

Q6: In which type of organism is circular DNA mainly found?

. . . . . .

Q7: In which structures of eukaryotic cells might circular DNA be found?

. . . . . .

Q8: What is the term for sections of extra-chromosomal DNA sometimes found in organisms?

1.3 Learning points

Summary

- DNA encodes hereditary information in a chemical language.
- All cells store their genetic information in the base sequence of DNA.
- The structure of a DNA nucleotide is composed of deoxyribose sugar, a phosphate and a base.
- Nucleotides bond to form a sugar-phosphate backbone.
- Base pairs (adenine, thymine, guanine and cytosine) hold the two strands together by hydrogen bonds, forming a double helix.
- Adenine always pairs with thymine, and guanine always pairs with cytosine.
- DNA is a double-stranded, antiparallel (one strand goes from 3′ to 5′, the other from 5′ to 3′) structure with a deoxyribose and phosphate backbone held together by internal base pairs.
- The DNA molecule can be circular or linear.
- Circular chromosomal DNA and plasmids are found in prokaryotes.
- Circular plasmids are found in yeast.
- Circular chromosomes are in the mitochondria and chloroplasts of eukaryotes.
- The DNA found in the linear chromosomes of the nucleus of eukaryotes is tightly coiled and packaged with associated proteins called histones.
- The sum total of all genetic material is the genome.
- Chromosomes in eukaryotes are contained within a membrane-bound nucleus.

1.4 Extension materials

The material in this section is not examinable. It includes information which will widen your appreciation of this section of work.

Extension materials: The Discovery of DNA

The significance of DNA and its role in hereditary can be traced from the work of Griffiths, who in 1928 demonstrated the "transforming principle" in bacteria. He, and later others (Avery, McCartney & McLeod, 1944), would show this "transforming principle" to be DNA.

Later, in the 1950s, Hershey & Chase, working with bacteriophage and radioactive forms of phosphorus and sulfur, would confirm DNA as the genetic material and eliminated protein as the carrier of genetic information.

Edwin Chargaff, using paper chromatography and ultraviolet spectroscopy techniques, demonstrated two findings, now known as Chargaff's rules: firstly, that adenine and thymine always occur together, and similarly that cytosine and guanine pair up - this is called base pairing; secondly, that DNA sequences vary between species.

In the early 1950s, work by Maurice Wilkins and Rosalind Franklin uncovered some characteristic features of the DNA molecule. Using a method called X-ray crystallography, it was shown that the molecule had a helical structure. Using this, and other evidence, Francis Crick and James Watson were able to construct the model of DNA that we recognise today.

The main features of the model are that it not only shows the configuration of the molecule, but it allows for the explanation of two processes: the first is the mechanism for DNA replication (semi-conservative) and the second is how it codes for proteins.

The story of the discovery of the structure of DNA, and an excellent book about how science works, is: *The Double Helix* by James D. Watson, published by Penguin Books.

1.5 End of topic test

End of Topic 1 test Go online

Q9: What is the term for the sum total of DNA in an organism?

. . . . . .

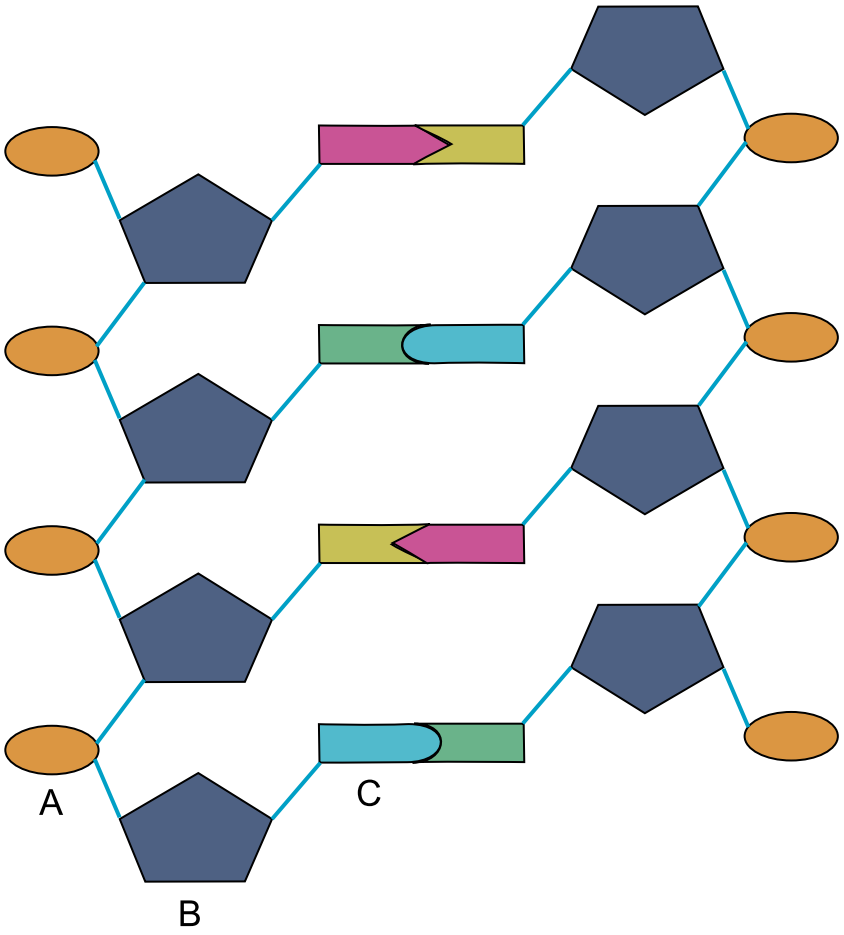
Q10: What is the complementary base sequence for the following DNA sequence: ATGGACTTTAGGT?

. . . . . .

Q11: A DNA molecule contains 1500 bases, of which 450 are adenine. What percentage of the bases in the DNA molecule are guanine?

. . . . . .

Q12: The following diagram shows part of a DNA molecule.



Which of the following correctly identifies **base**, **phosphate** and **sugar**, in that order?

1. A, B, C
2. A, C, B
3. B, A, C
4. C, A, B

. . . . . .

Q13: What is the name of the sugar that is found in DNA?

. . . . . .

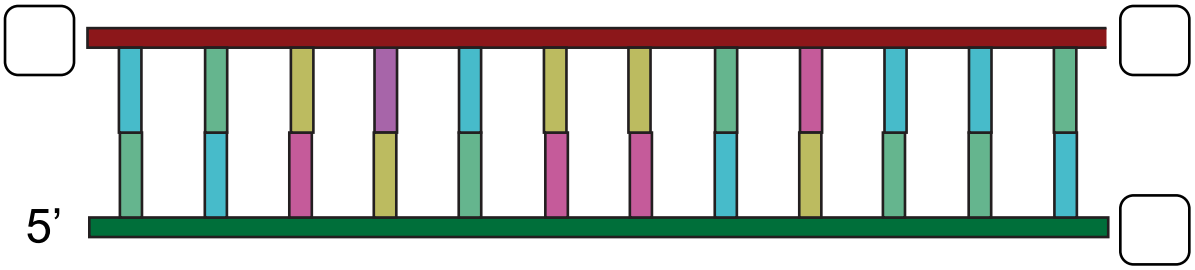
Q14: Base pairs in a DNA molecule are linked by weak bonds. What is the name for these?

. . . . . .

Q15: What name is given to the backbone of DNA?

. . . . . .

Q16: The diagram below illustrates the structure of a DNA molecule.



Fill in the missing labels to complete the diagram.

. . . . . .

Q17: Name two sources of circular DNA in eukaryotes.

. . . . . .

Q18: Name the substance which DNA is packaged with in the nucleus of eukaryotes.

Topic 2

Replication of DNA

Contents

Prerequisites

You should already know that:

- chromosomes (and therefore DNA) are replicated during mitosis;
- the two strands of DNA are held together by complementary base pairs;
- DNA contains the four bases A, T, G and C which make up the genetic code.

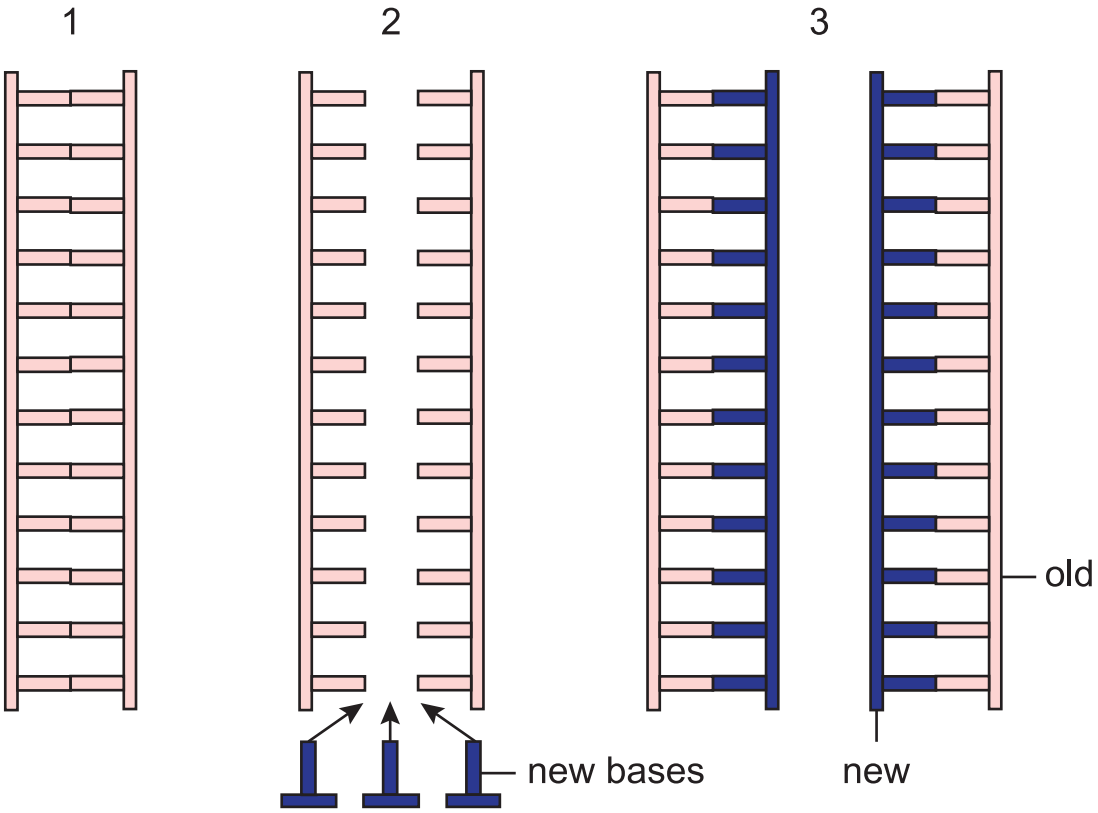
Learning objective

By the end of this topic, you should be able to:

- explain the function of DNA polymerase, when it acts, and what conditions are necessary for it to function;
- describe how DNA replicates in terms of DNA unwinding and DNA polymerase adding complementary nucleotides;
- explain the importance and significance of adding new nucleotides to the 3′ end of an existing DNA chain;
- state that replication occurs at various points on a DNA molecule;
- outline the process where DNA polymerase can act continuously on the leading strand, but discontinuously on the lagging strand;
- describe the action and significance of the enzyme ligase;
- state the purpose of the polymerase chain reaction (PCR);
- describe the action and purpose of primers in PCR;
- describe the sequence of PCR in terms of heating DNA, adding primers, and cooling DNA;
- explain why heat tolerant DNA polymerase is required;
- explain the significance and outcome of 'cycling' sequences;
- describe the practical applications of PCR.

2.1 DNA replication

DNA replication takes place prior to cell division. The replication of DNA is semi-conservative. Each strand acts as a template for the synthesis of a new DNA molecule by the addition of complementary base pairs, thereby generating a new DNA strand that is the complementary sequence to the parental DNA. Each daughter DNA molecule ends up with one of the original strands and one newly synthesised strand.



DNA replication

DNA replication is an enzyme controlled process which relies on the activities of **DNA polymerase** and **DNA ligase**.

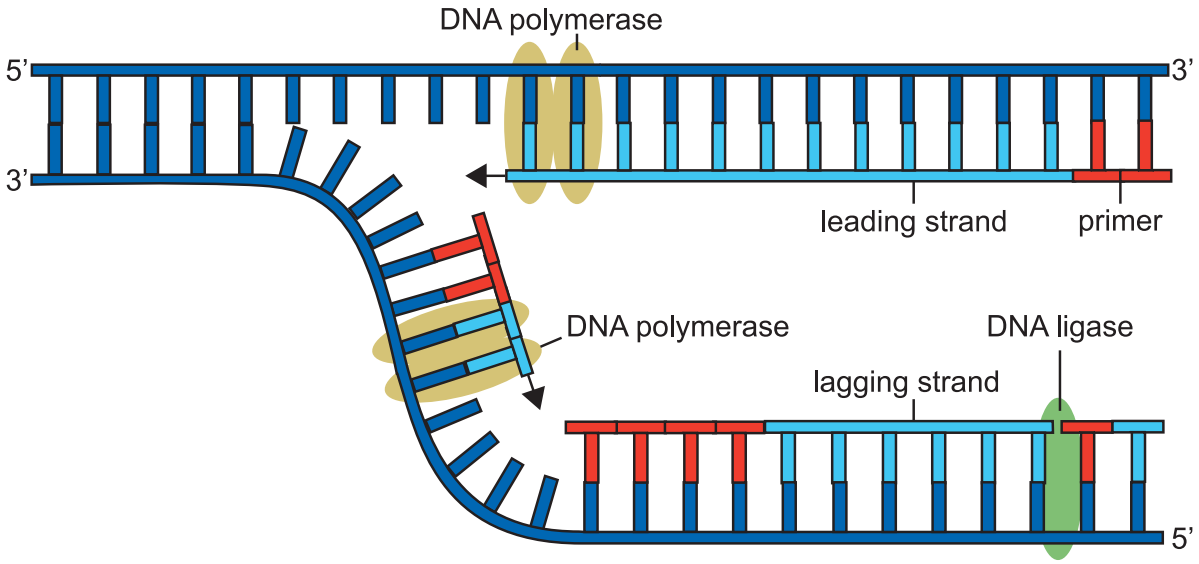
- **DNA polymerase**: adds DNA nucleotides, using complementary base pairing, to the deoxyribose (3′) end of the new DNA strand which is forming.
- **DNA ligase**: joins fragments of DNA together.

Before replication begins, there must be a pool of free nucleotides present; however, DNA polymerase cannot start adding nucleotides on its own. Short sections of RNA nucleotides called **primers** are added to the DNA and the enzyme extends from them. A primer is a short strand of nucleotides which binds to the 3′ end of the template DNA strand allowing polymerase to add DNA nucleotides.

Due to the action of the enzyme DNA polymerase, the two strands of DNA are copied differently. The **leading strand** is made continuously while the **lagging strand** is made in fragments, which are then joined together.

DNA replication: Steps Go online

The following provides a summary of the steps involved in DNA replication.



- DNA is unwound and hydrogen bonds between bases are broken to form two template strands.
- two replication forks form and open the double-strand in opposite directions, exposing the bases.
- on the leading strand, a primer binds to the DNA and DNA polymerase adds nucleotides to the 3′ end. DNA polymerase catalyses the formation of a chemical bond between nucleotides and continues to add nucleotides to the 3′ end of the growing strand.
- on the lagging strand, a primer binds to the DNA once it is exposed and DNA polymerase adds nucleotides to the 3′ end. As more DNA is exposed, a new primer is added. DNA polymerase extends the new strand from this primer until it meets the previous fragment. The old primer is replaced by DNA and the enzyme DNA ligase joins the fragments together. As the DNA unzips further, another fragment will be made and connected to the previous one.

2.2 The polymerase chain reaction (PCR)

PCR can be used to amplify a desired DNA sequence of any origin (virus, bacteria, plant or animal) millions of times in a matter of hours. It is especially useful because:

- it is highly specific;
- it is easily automated;
- it is capable of amplifying minute amounts of sample.

To amplify a target DNA sequence, several components are required:

- buffer;
- nucleotides;
- **primers**;
- Taq polymerase;
- template DNA.

The section of DNA which is to be amplified must be added to the reaction mixture. It acts as a template to copy from. The buffer keeps the reaction mixture at the correct pH to ensure the reaction will proceed.

Polymerase enzyme is found in all animals. It has an optimum temperature of 37°C. PCR requires polymerase to operate at high temperatures. *Thermus aquaticus* is a bacterium that lives in hot springs and hydrothermal vents. Taq polymerase, an enzyme which adds nucleotides to DNA, was first isolated from this bacteria. It is special type of polymerase which is stable at high temperatures, having an optimum temperature of 70°C.

Polymerase can only add nucleotides to an existing strand of DNA, therefore, PCR requires primers. In PCR, primers are short strands of nucleotides which are complementary to specific target sequences at the two ends of the region of DNA to be amplified.

The PCR process involves repeated cycles of the following steps:

- the DNA molecule which is to be amplified is denatured by heating to between 92 and 98°C, breaking the hydrogen bonds between base pairs to separate the strands;
- the solution is cooled to between 50 and 65°C to allow the primers to bind to target sequences;
- the solution is heated to between 70 and 80°C for heat-tolerant DNA polymerase to replicate the region of DNA.

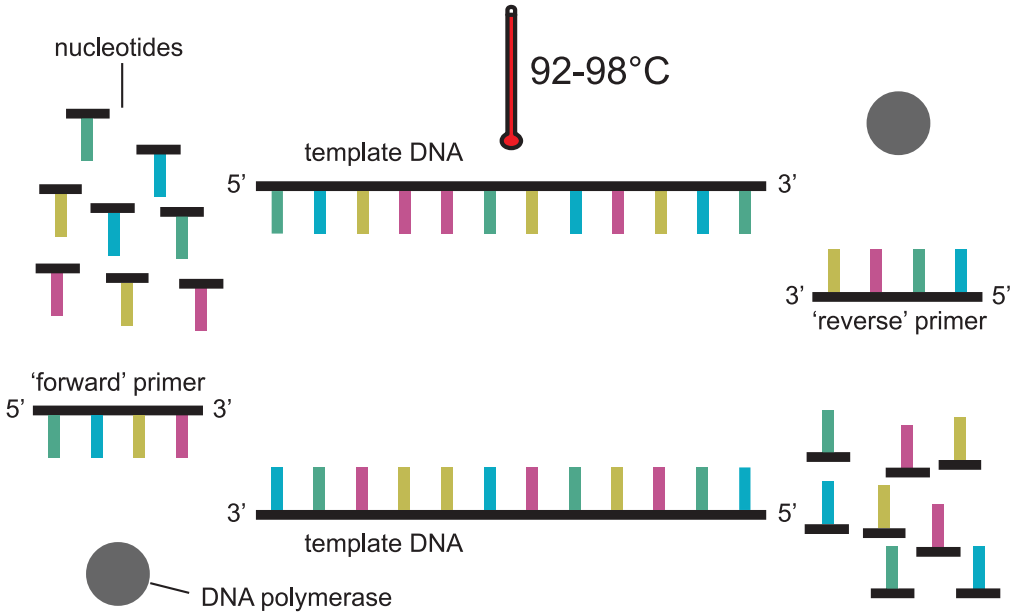
This cycle is usually repeated at least 30 times.

Polymerase chain reaction allows DNA to be amplified *in vitro*; this means it is performed outwith a living organism. The opposite, *in vivo*, means carried out within an organism.

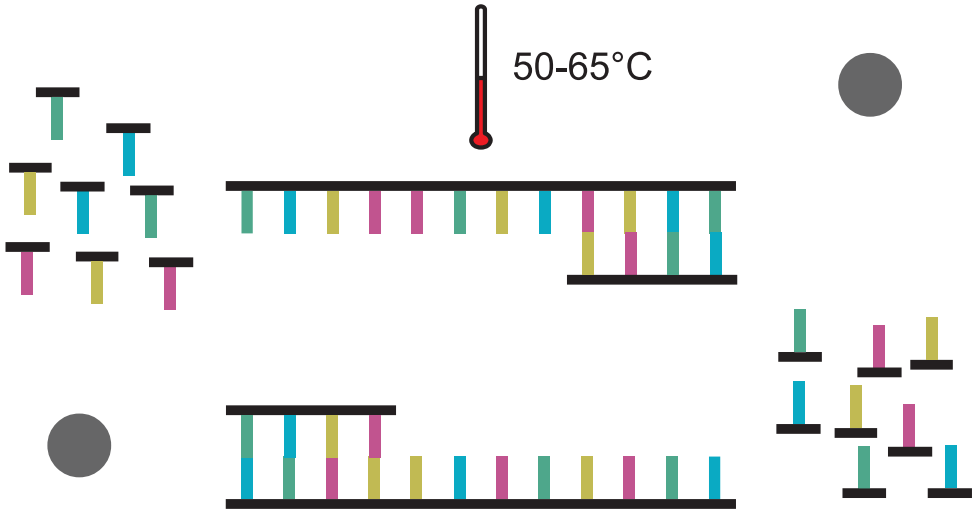
The polymerase chain reaction (PCR): Stages Go online

The following provides a summary of the stages of PCR.

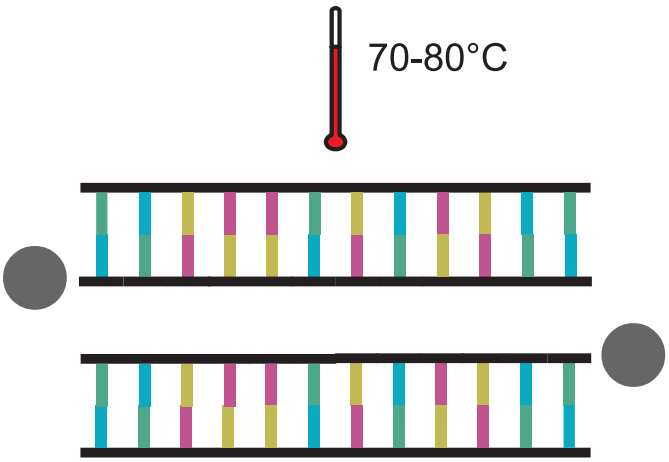
Stage 1, Heating (92-98°C): DNA is heated to separate the strands.



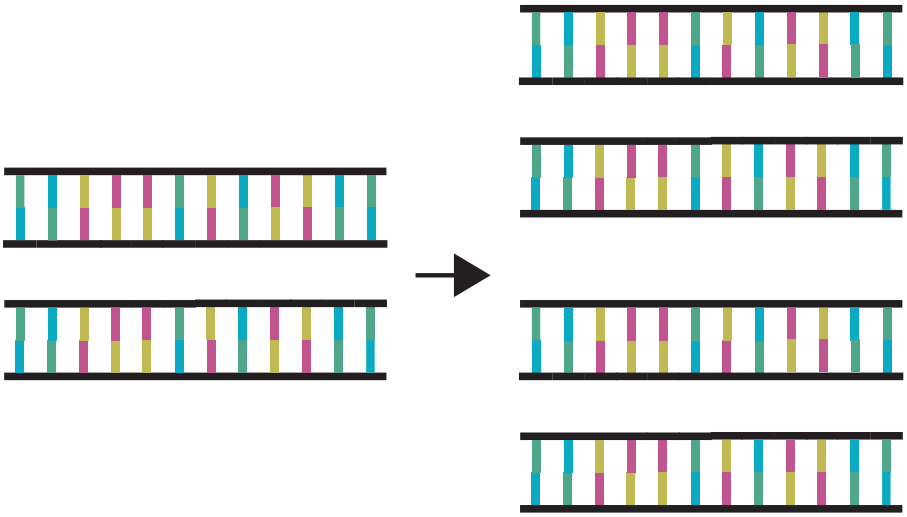
Stage 2, Annealing (50-65°C): DNA is then cooled to allow primers to bind to target sequences.



Stage 3, Extension (70-80°C): DNA is then heated for heat-tolerant DNA polymerase to replicate the region of DNA.



Result: The amount of DNA has now doubled.



Each repetition of PCR doubles the amount of DNA.

PCR has revolutionised many areas of research science. PCR is involved in the process of DNA sequencing and has allowed amplification of DNA from ancient sources, such as Neanderthal bones, enabling in-depth DNA analysis. PCR also has forensic applications, allowing minute quantities of DNA from a crime scene to be amplified, sequenced, and compared to DNA sequences from suspects. PCR has medical applications, for example the diagnosis of human immunodeficiency virus (HIV). Finally, PCR can be used to settle paternity disputes by amplifying and comparing a child's DNA to their potential father.

2.3 Learning points

Summary

- Prior to cell division, DNA polymerase replicates a DNA strand precisely using DNA nucleotides.
- DNA polymerase needs a primer to start replication.
- A primer is a short strand of nucleotides which binds to the 3′ end of the template DNA strand allowing polymerase to add DNA nucleotides.
- DNA unwinds to form two template strands.
- DNA polymerase adds DNA nucleotides, using complementary base pairing, to the deoxyribose (3′) end of the new DNA strand which is forming.
- This process occurs at several locations on a DNA molecule.
- DNA polymerase can only add DNA nucleotides in one direction resulting in the leading strand being replicated continuously and the lagging strand replicated in fragments.
- Fragments of DNA are joined together by ligase.
- The polymerase chain reaction (PCR) is a technique for the amplification of DNA *in vitro*.
- In PCR, primers are short strands of nucleotides which are complementary to specific target sequences at the two ends of the region of DNA to be amplified.
- PCR is a three step process:
  - heating (92-98°C) separates the DNA strands;
  - annealing (50-65°C) is the binding of the primers to target sequences;
  - extension (70-80°C) of the primers to complete the complementary strand is carried out by heat stable DNA polymerase.
- Repeated cycles of heating and cooling amplify this region of DNA.
- PCR can amplify DNA to help solve crimes, settle paternity suits and diagnose genetic disorders.

2.4 Extended response question

The activity which follows presents an extended response question similar to the style that you will encounter in the examination.

You should have a good understanding of DNA structure and replication before attempting the question.

You should give your completed answer to your teacher or tutor for marking, or try to mark it yourself using the suggested marking scheme.

Extended response question: DNA structure and replication

Give an account of DNA structure and replication. *(8 marks)*

2.5 Extension materials

The material in this section is not examinable. It includes information which will widen your appreciation of this section of work.

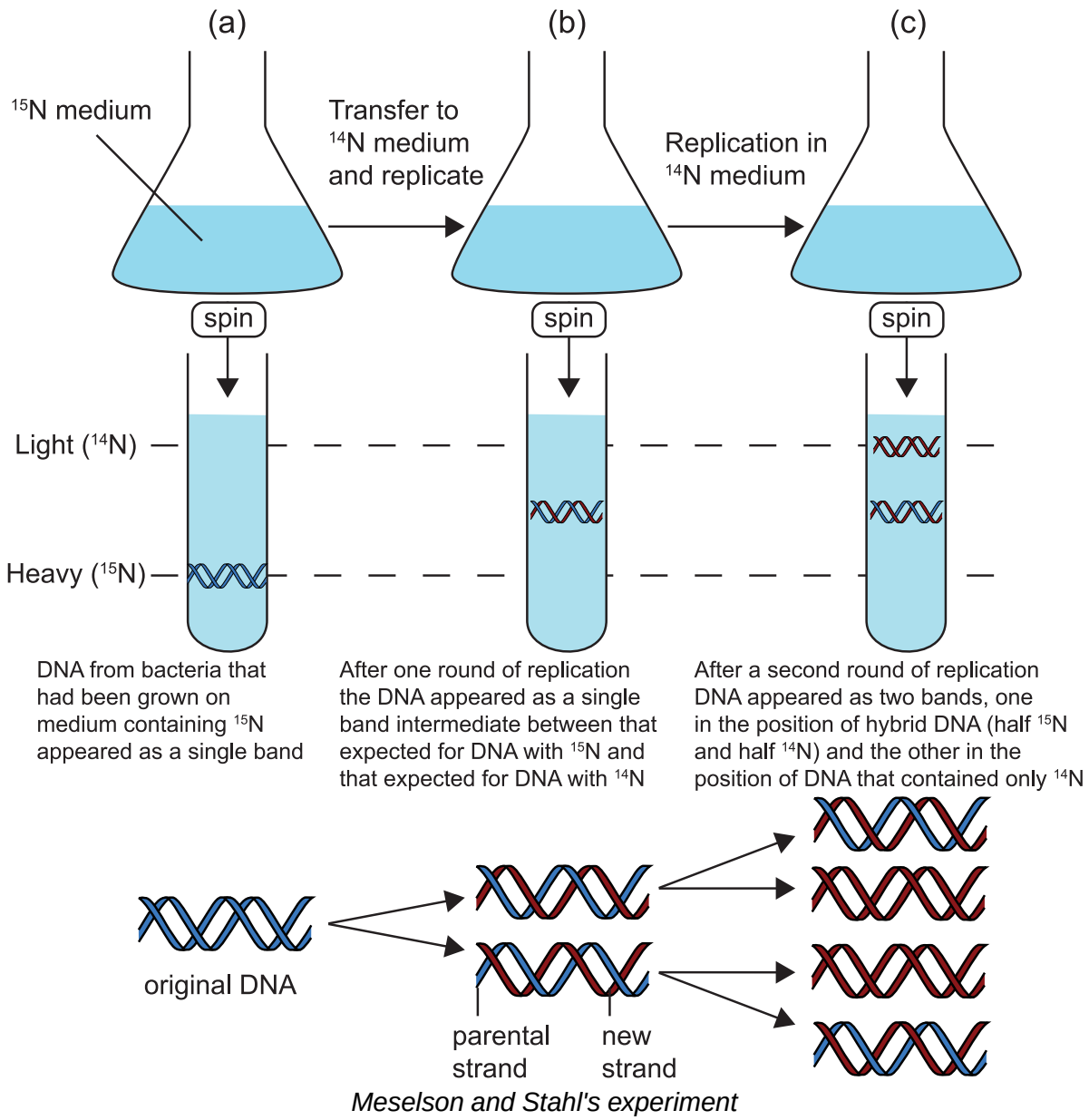
Extension materials: Meselson and Stahl's experiment

The mechanism of DNA replication is said to be semi-conservative. That is, after replication, each of the two resulting DNA molecules is composed of one original (or conserved) strand and one new strand. This hypothesis, put forward by Watson and Crick, was proved experimentally by Meselson and Stahl in the late 1950s. Their experiment involved growing a culture of the bacterium *Escherichia coli* in a growth medium containing heavy nitrogen (^{15}N). As the bacteria grew, they incorporated the heavy nitrogen into their nitrogenous bases. The bacteria were then inoculated into growth media containing light nitrogen (^{14}N) and three classes of DNA were subsequently extracted after the change to the light nitrogen medium.

The three classes of DNA were:

- parental DNA;
- first generation DNA;
- second generation DNA.

The results of the experiment showed that parental DNA grown in heavy medium was 'heavier' than when grown in light medium. First generation growth showed that the DNA was all of medium density. Lastly, the second generation showed DNA of both medium and light intensities. Second generation growth supported the semi-conservative model of DNA replication since there were two bands of growth (one with both conserved and new DNA, and a band of light DNA).



Meselson and Stahl's experiment

2.6 End of topic test

End of Topic 2 test Go online

Q1: Name the enzyme that is required to add nucleotides to a growing strand of DNA.

. . . . . .

Q2: What name is given to the short strand of nucleotides added to DNA at the start of the replication process?

. . . . . .

Q3: Name the enzyme that is required to join fragments of DNA together.

. . . . . .

Q4: The following four stages occur during DNA replication:

a) Base pairing occurs between free nucleotides and each of the DNA strands.
b) The hydrogen bonds between DNA strands break.
c) The DNA molecules coil up to form double helices.
d) Nucleotides are bonded together by DNA polymerase.

Which of the following gives these stages in the correct order?

1. a, b, d, c
2. b, a, d, c
3. c, b, a, d
4. d, a, b, c

. . . . . .

Q5: The DNA polymerase used in the polymerase chain reaction possesses a particular characteristic that makes it ideally suited to the purpose. What is it?

a) The enzyme is relatively stable at high temperatures.
b) The enzyme synthesises DNA at a very rapid rate.
c) The enzyme is very accurate at copying DNA from a template.
d) The enzyme can seal together fragments of DNA.

. . . . . .

Q6: The steps involved in the polymerase chain reaction are given below.

a) Temperature of the reaction adjusted to 70-80°C.
b) Temperature of the reaction adjusted to 50-65°C.
c) The DNA strands separate.
d) Temperature of the reaction adjusted to 92-98°C.
e) Synthesis of DNA by the enzyme DNA polymerase.
f) Annealing of the primers to the single-stranded DNA.

Which of the following describes the correct order in which the steps would occur?

1. d, c, b, e, a, f
2. d, c, a, f, b, e
3. d, c, b, f, a, e
4. d, c, a, e, b, f

. . . . . .

Q7: _____ occurs at 70-80°C.

Word list: extension, annealing, heating, exponential.

. . . . . .

Q8: PCR leads to an _____ amplification of desired DNA sequences.

Word list: extension, annealing, heating, exponential.

. . . . . .

Q9: Starting with a single molecule of DNA, the polymerase chain reaction was allowed to go through three complete cycles. How many molecules of DNA would be produced?

a) 4
b) 8
c) 16
d) 32

Topic 3

Gene expression

Contents

Prerequisites

You should already know that:

- DNA carries the genetic information for making proteins;
- the base sequence of DNA determines the amino acid sequence in proteins;
- messenger RNA (mRNA) is a molecule which carries a copy of the code from the DNA in the nucleus to a ribosome in the cytoplasm;
- at the ribosome, proteins are assembled from amino acids.

Learning objective

By the end of this topic, you should be able to:

- explain that the genetic code is universal to all forms of life;
- explain that the phenotype is a combination of the genotype and environmental factors;
- describe the structure of RNA;
- describe the function of mRNA, tRNA and rRNA;
- describe the process of transcription including the role of RNA polymerase and complementary base pairing;
- describe the process of RNA splicing;
- describe the process of translation, including the role of tRNA and ribosomes;
- state that codons are found on mRNA and anticodons are found on tRNA;
- name the bonds which hold amino acids together in a protein;
- explain how different proteins can be expressed from one gene;
- state that polypeptide chains fold to give the final structure of the protein;
- describe the role of hydrogen bonds and the interactions between amino acids in the 3D shape of a protein.

3.1 Introduction

The many thousands of proteins that our cells use are synthesised inside the cells during a complex process involving DNA and RNA (the nucleic acids), as well as **ribosomes**. You will remember from previous studies that a protein is composed of amino acids joined together in a specific sequence.

The information to determine the sequence of amino acids in a protein is contained in the DNA in the nucleus of our cells. In this topic, we will study how the instructions for making a protein are transferred into the cytoplasm using RNA, and how a protein is actually constructed on the ribosomes.

Gene expression involves two major stages.

- The first process is **transcription**, during which the DNA is used to produce an RNA molecule that is called a primary transcript. This RNA has the same sequence as the gene. Human genes can be divided into **exons** and **introns**, but it is only the exons that carry the information needed for protein synthesis.
- The next stage of gene expression is known as **translation**, which allows amino acids to come together in a certain order at the **ribosome**, where they form a polypeptide chain.

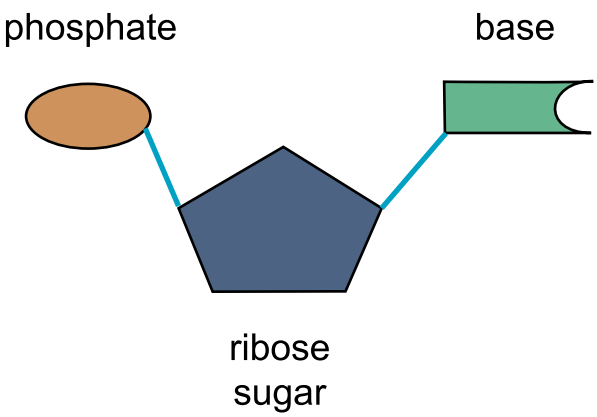
DNA	*makes*	**mRNA**	*makes*	**protein**
Process:	transcription		translation	
Occurs in:	the nucleus		cytoplasm	

Protein synthesis summary

3.2 The structure and functions of RNA

Ribonucleic acid (RNA) provides a bridge between DNA and protein synthesis.

RNA consists of nucleotides that are composed of phosphate, ribose sugar and a base.



An RNA nucleotide

Important points to remember about RNA structure are that:

- RNA nucleotides contain the sugar ribose;
- RNA has the base (U) uracil rather than (T) thymine (as found in DNA);
- RNA molecules are usually single-stranded.

There are three main types of RNA involved in protein synthesis:

1. **mRNA** (messenger RNA), which carries a copy of the DNA code from the nucleus to the ribosome;
2. **tRNA** (transfer RNA), which are molecules found in the cytoplasm that become attached to specific amino acids, bringing them to the ribosomes where they are joined together;
3. **rRNA** (ribosomal RNA), which forms a complex with protein molecules to make the ribosome.

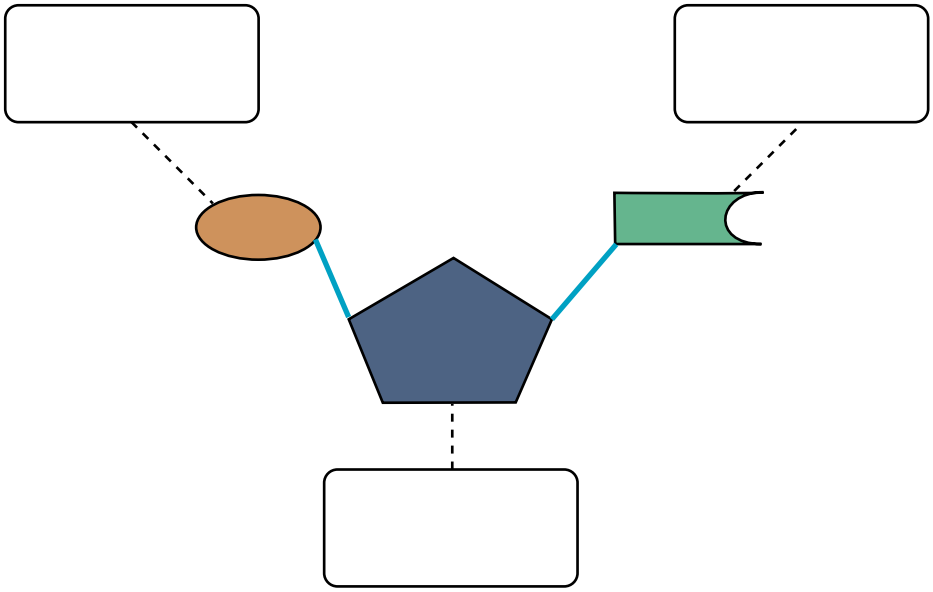
Messenger RNA (mRNA)	For the synthesis of a protein, the particular sequence of bases on the DNA is first transcribed into the complementary sequence of mRNA. This messenger RNA can then carry the information for a protein through the nuclear envelope to the sites of protein synthesis (the ribosomes). Each triplet of bases on the mRNA molecule is called a codon and codes for a specific amino acid.
Transfer RNA (tRNA)	This type of RNA is responsible for the transport and transfer of individual amino acids during protein synthesis. Amino acids are transported by specific tRNA molecules, which recognise the genetic code presented by the mRNA. The three bases exposed at the bottom form the anticodon. This is the complementary base sequence to the base sequence on mRNA coding for a particular amino acid.
Ribosomal RNA (rRNA)	This type of RNA is bound to structural proteins to form a ribosome. The ribosome is used in the synthesis of proteins.

The three forms of RNA

The structure and functions of RNA: Questions

Go online

Q1: Complete the diagram using the labels from the list.



Labels: base, phosphate, ribose sugar.

. . . . . .

Q2: Complete the table using the properties of RNA and DNA from the list.

	RNA	DNA
Structure		
Preferred form		
Number of types		
Present in		
Bases		

Properties: >1; 1; adenine, cytosine, guanine and thymine; adenine, cytosine, guanine and uracil; double-stranded; double helix; not a double helix; single-stranded; the cytoplasm and the nucleus; the nucleus.

Q3: What are the components of an RNA nucleotide?

. . . . . .

Q4: List the four types of bases found in an RNA molecule.

. . . . . .

Q5: Name the three types of RNA found in the cell.

. . . . . .

Q6: What are main functional differences between mRNA and tRNA?

. . . . . .

Q7: Nucleotides are the building blocks of:

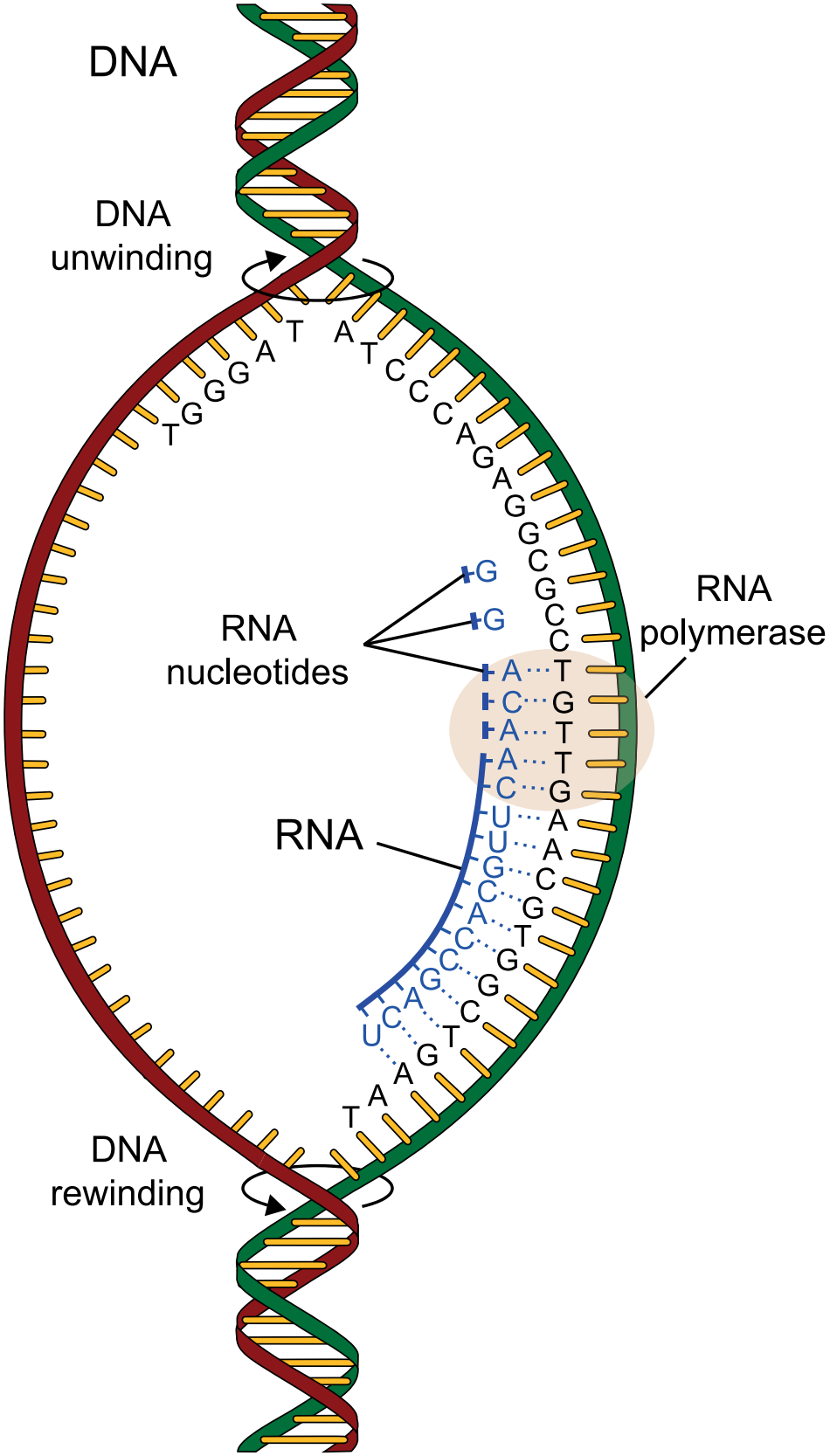
a) DNA only
b) RNA only
c) both DNA and RNA
d) neither DNA nor RNA

3.3 Transcription

Transcription is the first step in protein synthesis. Information from DNA is copied into an RNA molecule, a process which takes place in the nucleus. The RNA polymerase enzyme moves along the DNA, unwinding the double helix and breaking the hydrogen bonds that hold the base pairs together. Free RNA nucleotides bond with the complementary base pairs on the DNA. The base pairing rules are summarised in the table below. The RNA nucleotides are held in place by hydrogen bonds while strong bonds form between the phosphate of one nucleotide and the ribose sugar of the adjacent nucleotide. When transcription is complete, the RNA polymerase enzyme and the mRNA strand that has been constructed are released. The mRNA that has been produced at this stage is known as the primary transcript.

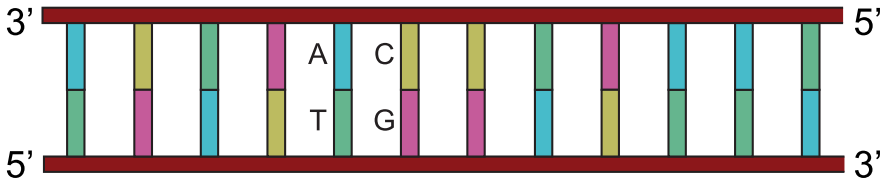
DNA nucleotide		**RNA nucleotide**
A	pairs with	U
T	pairs with	A
G	pairs with	C
C	pairs with	G

RNA base pairing rules

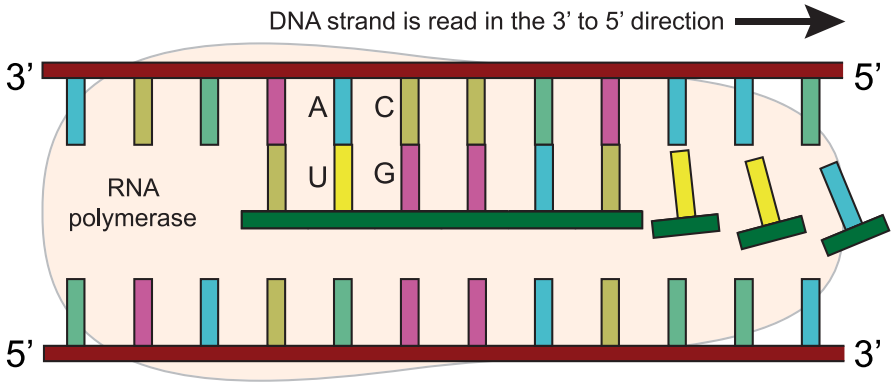


An overview of the transcription process

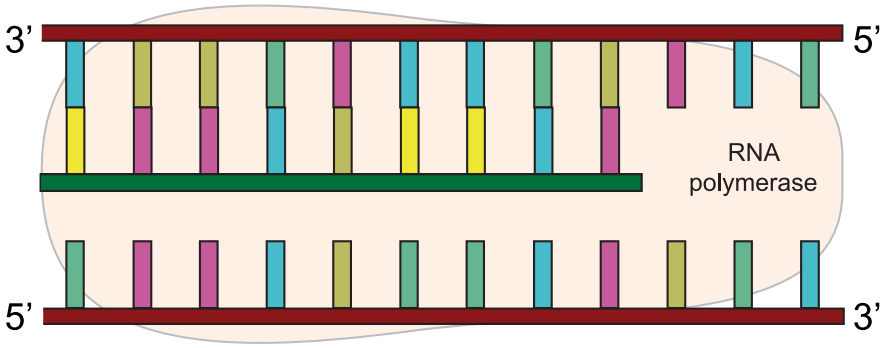
A section of DNA containing the genetic instructions for the production of a protein.



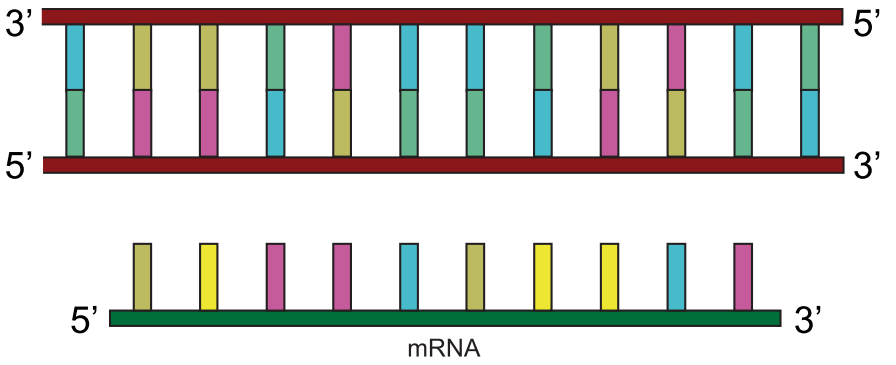
RNA polymerase unwinds the DNA and breaks the hydrogen bonds between the bases. This causes the strands to separate and expose their bases. Free RNA nucleotides find and align with complementary DNA nucleotides by hydrogen bonding. Strong chemical bonds form between the sugar of one RNA nucleotide and the phosphate of the next.



This process continues along the length of the gene.



The weak hydrogen bonds between the DNA and RNA bases break, allowing the mRNA to separate from the DNA and then move away from the DNA. The weak hydrogen bonds between the DNA strands re-unite and the molecule winds up into a double helix again.



The steps involved in RNA transcription

After a eukaryotic cell transcribes a protein coding gene, the RNA transcript, called the primary transcript, is processed. One type of processing is called **RNA splicing**. This takes place in the nucleus, after which the mature mRNA is released into the cytoplasm where ribosomes translate the mRNA transcript into protein.

RNA splicing forms a mature mRNA transcript. The **introns** of the primary transcript are non-coding regions and are removed. The **exons** are coding regions and are joined together to form the mature transcript. The order of the exons is unchanged during splicing.

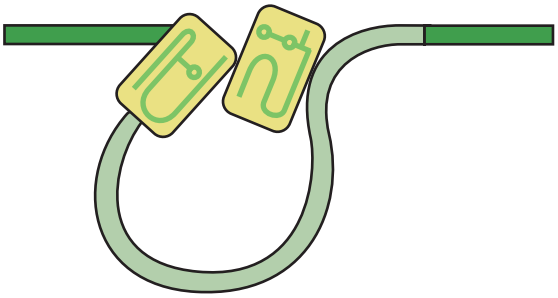
RNA splicing: Steps Go online

The following provides a summary of the steps involved in RNA splicing.

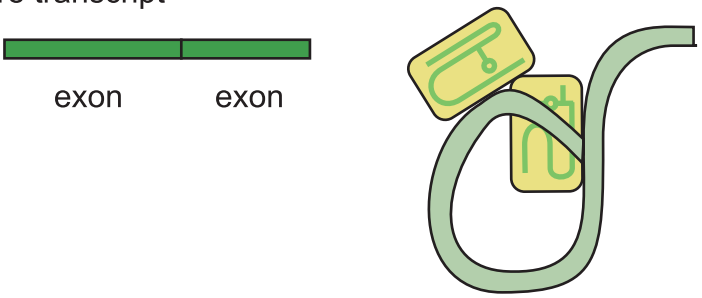
primary transcript

molecules which bring about splicing attach to primary transcript

splicing takes place



mature transcript



Transcription: Questions Go online

Q8: Name the enzyme which is responsible for producing mRNA.

. . . . . .

Q9: Give the location of transcription.

. . . . . .

Q10: Name the bonds between DNA bases which must be broken during transcription.

. . . . . .

Q11: What name is given to the mRNA molecule which has just been released from the DNA?

. . . . . .

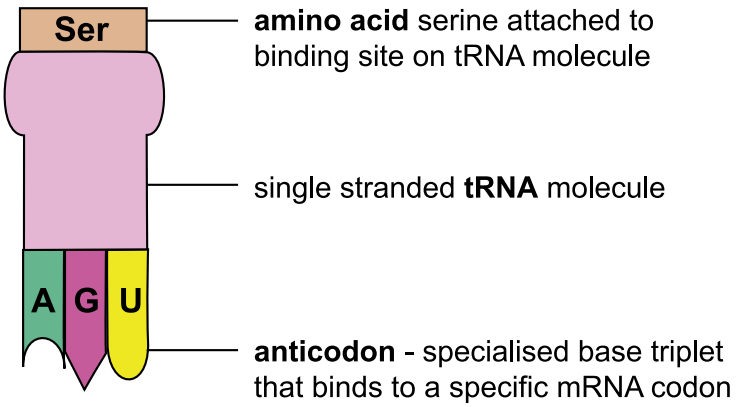
Q12: Name the process which removes introns from the primary transcript.

3.4 Translation

Once the DNA in a gene has been transcribed into mRNA, **translation** can take place. First, the mRNA molecules pass through the nuclear pores. Translation of mRNA into protein takes place on **ribosomes** in the cytoplasm and requires a second type of RNA, **transfer RNA (tRNA)**.

tRNA Go online

The following provides structural information about a tRNA molecule, in this case with the amino acid serine attached to it.



Amino acids are attached to tRNA at the amino acid attachment site at the top of the molecule. Unlike mRNA, there are some regions of base pairing in a tRNA molecule. The three bases exposed at the base of the tRNA molecule is called the **anticodon**. (Each group of three bases on the mRNA which codes for an amino acid is called a **codon**).

As an mRNA molecule passes through a ribosome, each **codon** is translated into an amino acid. The genetic code table indicates which amino acid corresponds to each mRNA codon. The tRNA molecule carrying the complementary anticodon binds briefly to the mRNA codon. The amino acid attached to the tRNA is then added to the polypeptide chain being synthesised. Amino acids are joined together by strong peptide bonds. After the amino acid has been added to a polypeptide chain during translation, the tRNA is free to pick up another amino acid in the cytoplasm.

		Second letter					
		U	C	A	G		
First letter	U	phe	ser	tyr	cys	U	Third letter
	U	phe	ser	tyr	cys	C	
	U	leu	ser	stop	stop	A	
	U	leu	ser	stop	trp	G	
	C	leu	pro	his	arg	U	
	C	leu	pro	his	arg	C	
	C	leu	pro	gln	arg	A	
	C	leu	pro	gln	arg	G	
	A	ile	thr	asn	ser	U	
	A	ile	thr	asn	ser	C	
	A	ile	thr	lys	arg	A	
	A	met	thr	lys	arg	G	
	G	val	ala	asp	gly	U	
	G	val	ala	asp	gly	C	
	G	val	ala	glu	gly	A	
	G	val	ala	glu	gly	G	

Genetic code table

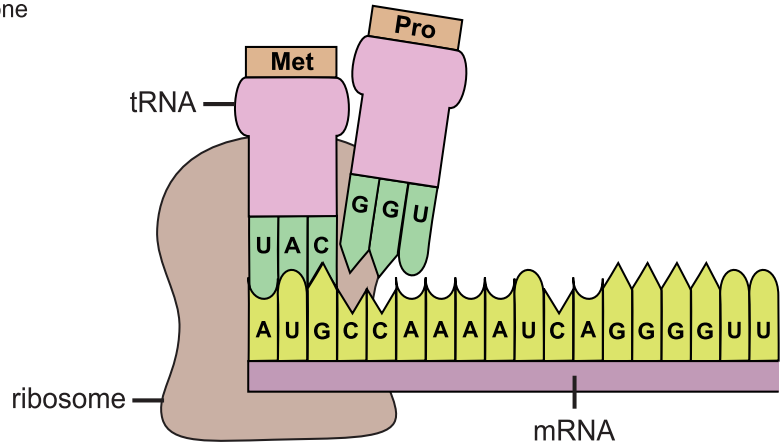
When the polypeptide chain is completed, it is released from the ribosome. Further processing, such as folding and binding to other polypeptide chains, results in the formation of a mature protein. The mRNA molecule is usually reused to produce more identical polypeptide chains.

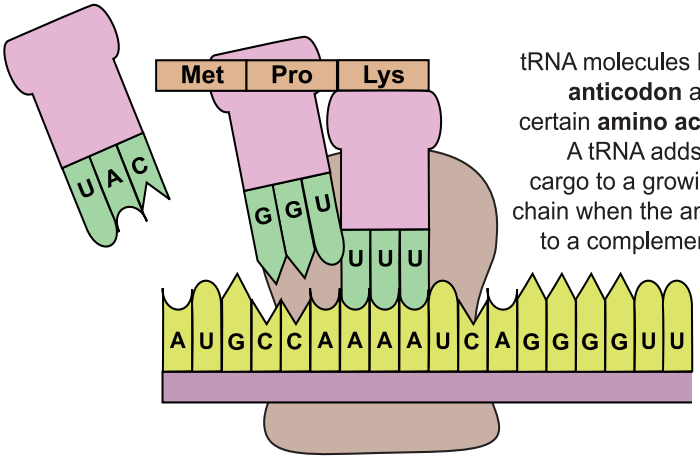
There are three codons that do not code for amino acids: UGA, UAA and UAG. These codons are known as stop codons and signal where translation ends. The genetic code also includes start codons where translation begins. In eukaryotes this is almost always AUG, which also codes for the amino acid methionine.

Translation: Steps

Go online

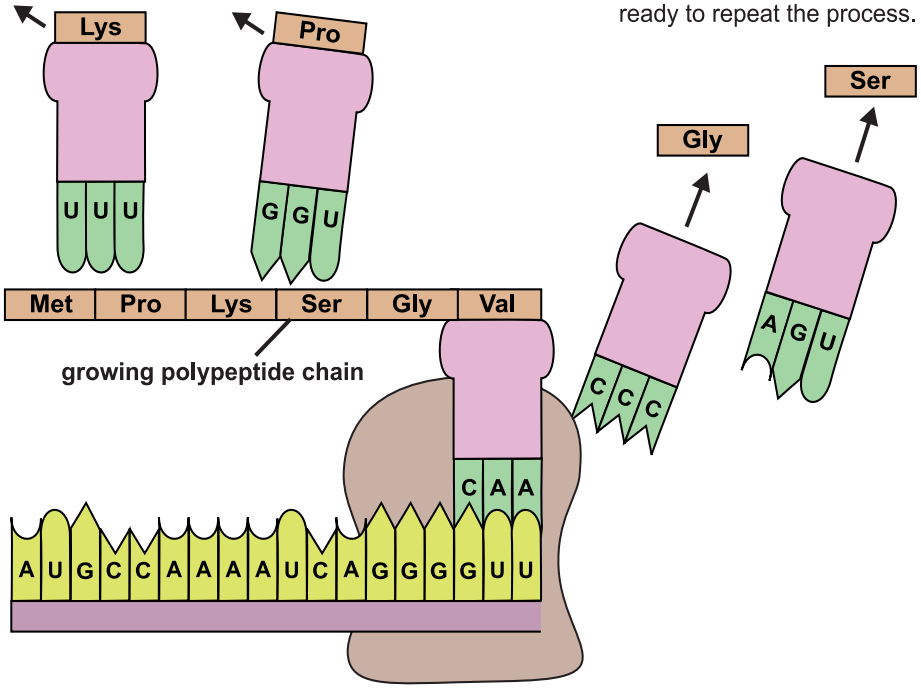
As a molecule of mRNA slides through a ribosome, **codons** are translated into amino acids, one by one

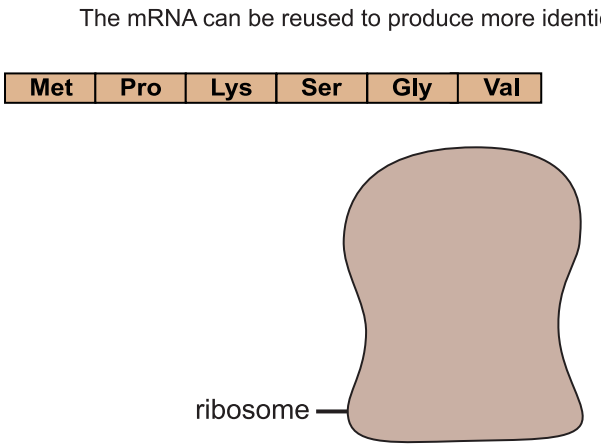




tRNA molecules have a specific **anticodon** at on end and a certain **amino acid** at the other. A tRNA adds its amino acid cargo to a growing polypeptide chain when the anticodon bonds to a complementary codon on the mRNA

Strong **peptide bonds** form between amino acids in the growing peptide chain. Each tRNA molecule becomes attached to another molecule of its amino acid ready to repeat the process.





The steps involved in translation

Translation: Questions

Go online

The following steps describe the role of messenger RNA in the cell, providing a summary of protein synthesis transcription to translation.

Q13: The following steps describe the role of messenger RNA in the cell, providing a summary of protein synthesis from transcription to translation. List the steps in the correct order.

- Hydrogen bonds are formed between the first codon of the mRNA and the complementary anticodon on a tRNA
- Hydrogen bonds form between the bases on the RNA and the DNA nucleotides
- The second tRNA binds to the mRNA
- The first tRNA leaves the ribosome, and another tRNA enters and base-pairs with the mRNA
- The double-stranded DNA unwinds, hydrogen bonds in the DNA break and the DNA strands separate
- A second peptide bond is then formed. The process continues, with the ribosome moving along the mRNA
- The mRNA leaves the nucleus and enters the cytoplasm
- A peptide bond forms between the amino acids carried by the tRNA molecules
- A ribosome attaches to the mRNA. Two transfer RNA (tRNA) molecules are also contained within the ribosome
- As each mRNA codon is exposed, incoming tRNA pairs with it and polypeptide synthesis continues until completed
- When synthesis of the mRNA is completed, the mRNA separates from the DNA
- An RNA nucleotide binds to a complementary nucleotide on one of the DNA strands
- The RNA nucleotides are linked together to form messenger RNA (mRNA)

. . . . . .

Q14: What name is given to the three bases exposed at the bottom of a tRNA molecule?

. . . . . .

Q15: Where does translation take place?

. . . . . .

Q16: Which cellular organelle is required for translation?

. . . . . .

Q17: Name the bond which holds amino acids together in a protein.

. . . . . .

Q18: Three mRNA codons do not code for amino acids. What is their role in translation?

Q19: Complete the sequence of bases encoded in the mRNA and then determine the sequence of amino acids in the protein. Use the genetic code table below to help.

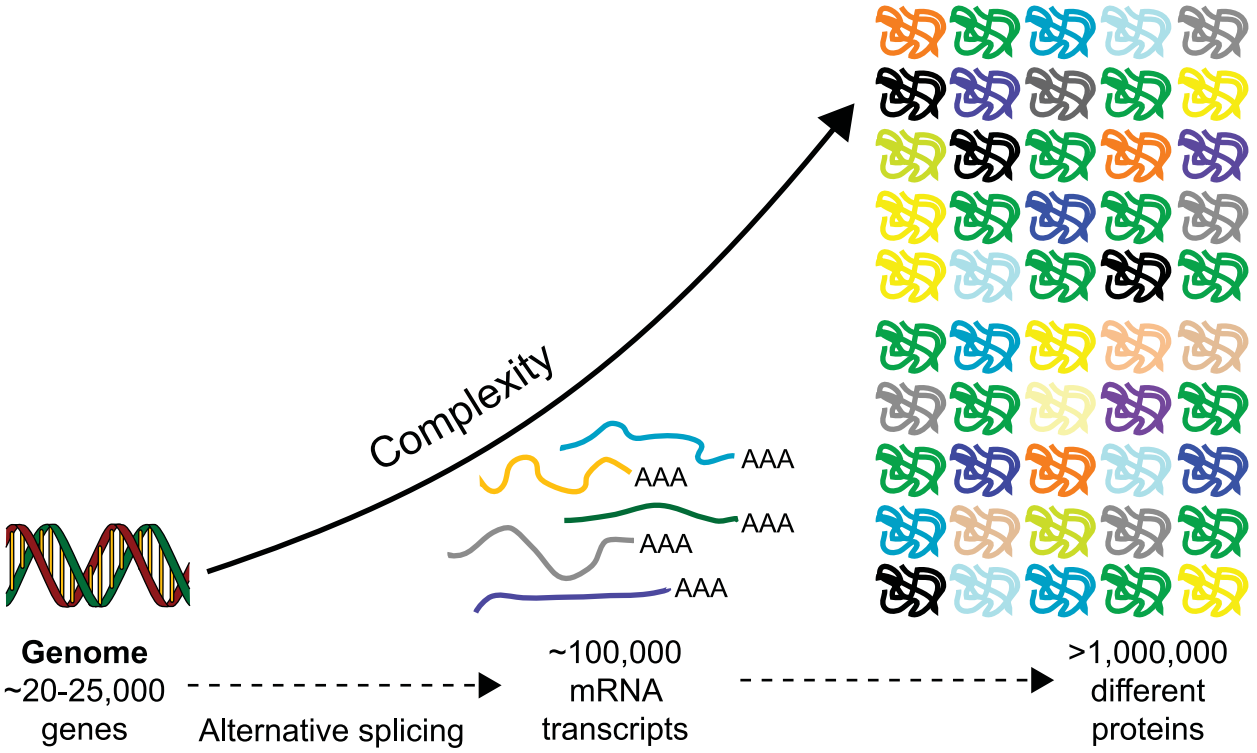
Second letter

First letter		U	C	A	G		Third letter
	U	phe	ser	tyr	cys	U	
	U	phe	ser	tyr	cys	C	
	U	leu	ser	stop	stop	A	
	U	leu	ser	stop	trp	G	
	C	leu	pro	his	arg	U	
	C	leu	pro	his	arg	C	
	C	leu	pro	gln	arg	A	
	C	leu	pro	gln	arg	G	
	A	ile	thr	asn	ser	U	
	A	ile	thr	asn	ser	C	
	A	ile	thr	lys	arg	A	
	A	met	thr	lys	arg	G	
	G	val	ala	asp	gly	U	
	G	val	ala	asp	gly	C	
	G	val	ala	glu	gly	A	
	G	val	ala	glu	gly	G	

DNA	C	A	C	A	G	T	G	T	T	T	G	T	C	C	G
mRNA															
protein															

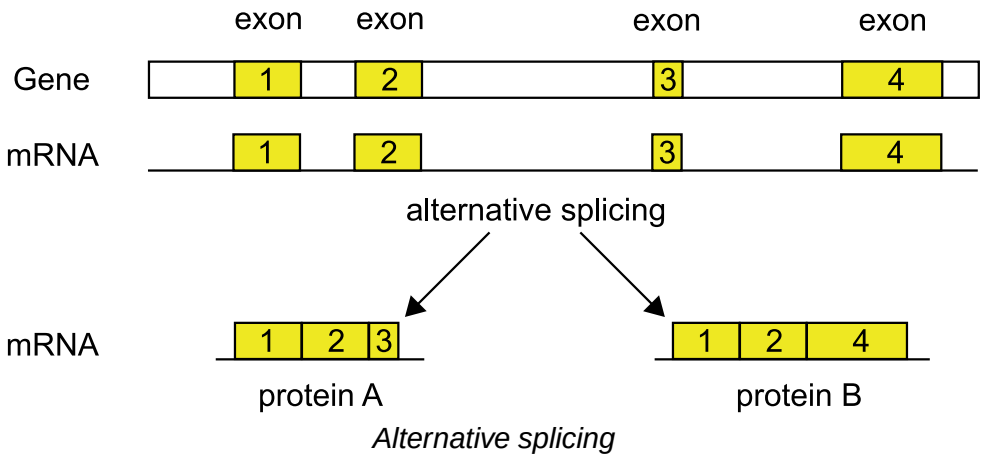
3.5 One gene, many proteins

Until quite recently, there was a theory that stated "one gene, one protein". This, however, has been superseded. With the completion of the Human Genome Project it is now accepted that the human genome contains between 20,000 and 25,000 genes, and yet it is also accepted that there are in excess of one million proteins in humans. Clearly there must be some mechanism that allows the genes to be expressed in a variety of ways.



One gene, many proteins

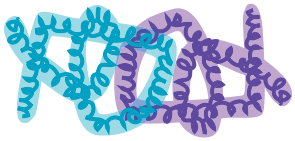
The mRNA (sometimes called pre-mRNA) can be edited in different ways by assembling a different sequence of exons for translation. As a result, many different mature transcripts of mRNA can be derived from one section of DNA. This process is known as alternative splicing. As a result of alternative splicing, different mature mRNA transcripts are produced from the same primary transcript depending on which exons are retained.



Alternative splicing

3.6 Protein structure and function

Proteins do not simply exist and function as strings of amino acids. After translation, the protein is folded to produce its final 3D shape. The folded polypeptide chains of a protein are held in place by hydrogen bonds and other interactions between individual amino acids.

Description	Diagram
Chain of amino acids linked by strong peptide bonds	
Polypeptide structure determined by weak hydrogen bonds	
Strong bonds form between special groups of amino acids	
More than one polypeptide makes up the final structure	

Protein structure

Different types of proteins have a variety of different functions within a living organism, examples of which are given in the following table.

Type of protein	Example
Structural	collagen, elastin
Contractile	actin and myosin in muscle cells
Hormones	insulin
Receptors	insulin receptor in liver cells (forming part of the structure of the plasma membrane)
Transport proteins	transporter of glucose into the cell
Defence proteins	immunoglobulins
Enzymes	lipase, pepsin, maltase

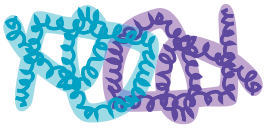
Protein functions

Protein structure and function: Questions Go online

Q20: Complete the table by selecting from the listed images.

Description	Diagram
Chain of amino acids linked by strong peptide bonds	
Polypeptide structure determined by weak hydrogen bonds	
Strong bonds form between special groups of amino acids	
More than one polypeptide makes up the final structure	

a) 

b) 

c) 

d) 

3.7 Learning points

Summary

- Gene expression involves the transcription and translation of DNA sequences.
- Only a fraction of the genes in a cell are expressed.
- RNA is single stranded and is composed of nucleotides containing ribose sugar, phosphate and one of four bases: cytosine, guanine, adenine and uracil.
- Transcription and translation involves three types of RNA (mRNA, tRNA and rRNA).
- Messenger RNA (mRNA) carries a copy of the DNA code from the nucleus to the ribosome.
- mRNA is transcribed from DNA in the nucleus and translated into proteins by ribosomes in the cytoplasm. Each triplet of bases on the mRNA molecule is called a codon and codes for a specific amino acid.

Summary continued

- Transfer RNA (tRNA) folds due to complementary base pairing. Each tRNA molecule carries its specific amino acid to the ribosome.
- A tRNA molecule has an anticodon (an exposed triplet of bases) at one end and an attachment site for a specific amino acid at the other end.
- Ribosomal RNA (rRNA) and proteins form the ribosome.
- RNA polymerase moves along DNA unwinding the double helix and breaking the hydrogen bonds between the bases. RNA polymerase synthesises a primary transcript of mRNA from RNA nucleotides by complementary base pairing.
- Uracil in RNA is complementary to adenine.
- RNA splicing forms a mature mRNA transcript.
- The introns of the primary transcript are non-coding regions and are removed.
- The exons are coding regions and are joined together to form the mature transcript.
- The order of the exons is unchanged during splicing.
- tRNA is involved in the translation of mRNA into a polypeptide at a ribosome.
- Translation begins at a start codon and ends at a stop codon.
- Anticodons bond to codons by complementary base pairing, translating the genetic code into a sequence of amino acids. Peptide bonds join the amino acids together. Each tRNA then leaves the ribosome as the polypeptide is formed.
- Different proteins can be expressed from one gene, as a result of alternative RNA splicing. Different mature mRNA transcripts are produced from the same primary transcript depending on which exons are retained.
- Amino acids are linked by peptide bonds to form polypeptides.
- Polypeptide chains fold to form the three-dimensional shape of a protein, held together by hydrogen bonds and other interactions between individual amino acids.
- Proteins have a large variety of shapes which determines their functions.
- Phenotype is determined by the proteins produced as the result of gene expression.
- Environmental factors also influence phenotype.

3.8 Extended response question

The activity which follows presents an extended response question similar to the style that you will encounter in the examination.

You should have a good understanding of protein synthesis before attempting the question.

You should give your completed answer to your teacher or tutor for marking, or try to mark it yourself using the suggested marking scheme.

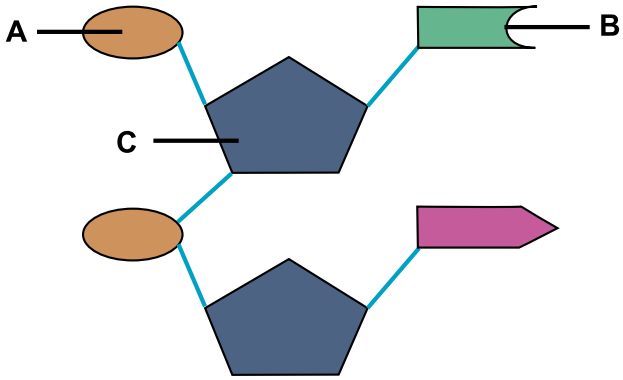
Extended response question: Protein synthesis

Describe the role of RNA in protein synthesis. *(7 marks)*

3.9 End of topic test

End of Topic 3 test Go online

Q21: The diagram shows two nucleotides that form part of a messenger RNA molecule. Which of the following correctly identifies phosphate, sugar and base, in order?



1. A, B and C
2. B, A and C
3. A, C and B
4. B, C and A

. . . . . .

Q22: Which of the following correctly describes the structure of an RNA molecule?

a) It is single-stranded; it has deoxyribose in its backbone; it contains the base thymine.
b) It is double-stranded; it has ribose in its backbone; it contains the base uracil.
c) It is single-stranded; it has ribose in its backbone; it contains the base uracil.
d) It is double-stranded; it has deoxyribose in its backbone; it contains the base thymine.

. . . . . .

Q23:

a) What is the messenger RNA (mRNA) sequence encoded by the following DNA sequence:
GCATTCATTGCA?

b) What name is given to the above process?

c) Name the enzyme required to carry out the above process.

. . . . . .

Q24: Complete the following sentence by selecting the correct options from the word list.

The mRNA produced after transcription is called the _____; the _____ are removed, leaving only the _____ in the final _____.

Word list: exons, introns, mature transcript, primary transcript.

. . . . . .

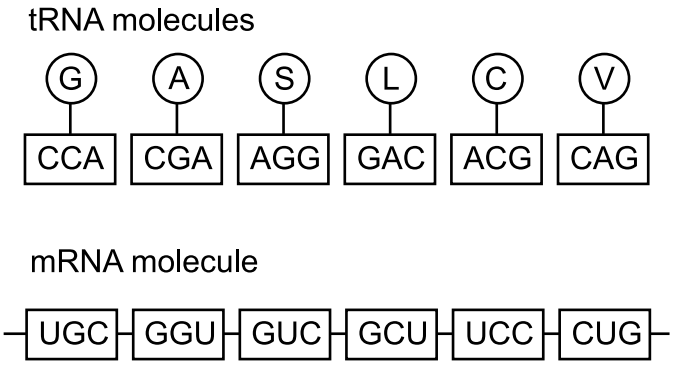
Q25: What name is given to the process which removes introns from mRNA?

. . . . . .

Q26: What name is given to the process by which ribosomes use messenger RNA to produce a polypeptide chain?

. . . . . .

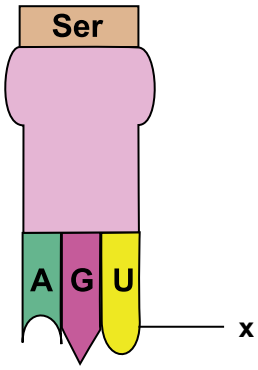
Q27: The diagram below shows six transfer RNA (tRNA) molecules, each of which carries a different amino acid (indicated by the letters G, A, S, L, C and V). Part of a messenger RNA (mRNA) molecule is also shown.



What is the sequence of amino acids encoded by the mRNA molecule?

. . . . . .

Q28: What name is given to the site indicated by letter X in the diagram below?

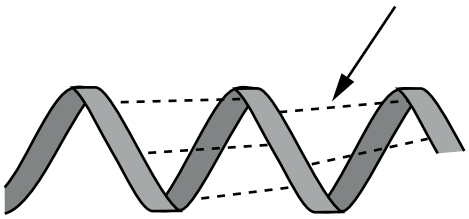


. . . . . .

Q29: Name the process which allows different mature mRNA transcripts to be produced from the same primary transcript.

. . . . . .

Q30: The following diagram shows part of a peptide chain from a protein.



What type of bond is indicated by the arrow?

Topic 4

Differentiation in multicellular organisms

Contents

Prerequisites

You should already know that:

- stem cells have the potential to become different types of cell;
- stem cells are involved in growth and repair.

Learning objective

By the end of this topic, you should be able to:

- understand and explain the term differentiation;
- describe the role of plant meristems and how they generate new tissue during plant growth;
- describe the role of animal stem cells;
- describe some therapeutic uses of stem cells;
- describe the importance of stem cell research and take into account the ethical issues.

4.1 Introduction

All living things are characterised by levels of organisation that are hierarchical. The cell is the lowest level of organisation that can exist independently. Multicellular organisms have cells organised into groups of cells called tissues, the next level of organisation. Tissues are formed from specialised cells that carry out a particular function. The columnar cells in the lining of the intestine, for example, are specialised for absorption of nutrients. Tissues can themselves become grouped together to form an organ. Most organs, such as the heart, lungs and liver, are also specialised for a certain function. The final level of organisation is the organ system where a group of organs work together at a particular function. Examples are the nervous system, the endocrine system and the vascular system.

Although the vast majority of cells contain identical genomes, cells within the same organism differ from one another because they express different genes and make different proteins.

Differentiation is the process by which cells or tissues undergo a change towards a more specialised function.

4.2 Meristems

Growth is restricted to specific regions (the meristems) of a plant, but it can occur throughout the plant's lifetime. In animals, growth can occur throughout the animal's body, but it stops when the animal reaches maturity. Animals do not have meristems; they are exclusive to plants.

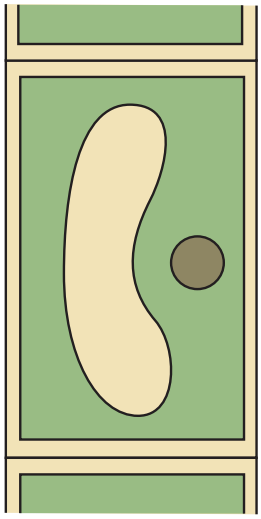
Meristems are regions of unspecialised cells in plants that can divide (self-renew) and/or differentiate. These cells divide rapidly by **mitosis** to differentiate and form new plant tissues.

Apical meristems are located at the tips of the roots and shoots of a plant. The name is derived from the position at the tip, which is also known as the apex. They contain a cluster of actively dividing cells that increase the length of the plant. Therefore, in order for a plant to increase in length, it has to produce new cells at the apical meristems.

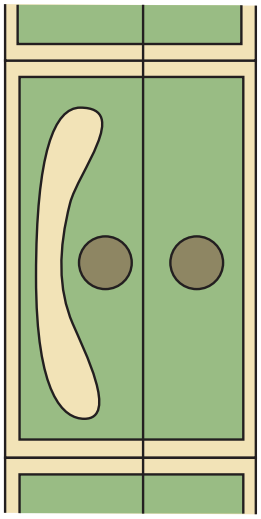
Differentiation: Phloem and xylem vessels Go online

The following illustrates how cells differentiate to form phloem and xylem vessels.

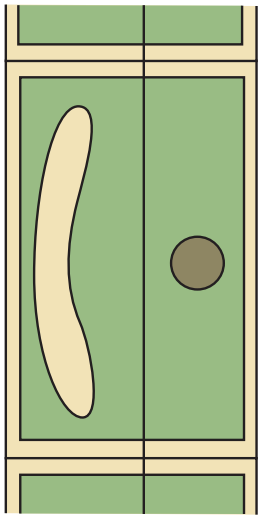
1. Phloem sieve tubes transport carbohydrates throughout a plant.



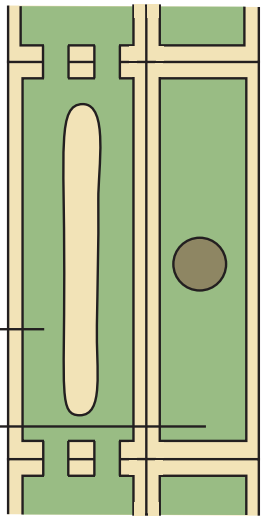
2. The cell beings to divide longitudinally and a second nucleus is formed.



3. After the cell division is complete, the nucleus in the sieve tube disintegrates and a companion cell is formed.

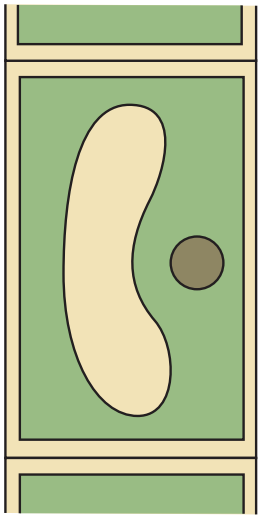


4. A cell wall forms between the sieve tube and companion cell. The end walls of the sieve tube become perforated and the cell contents die.

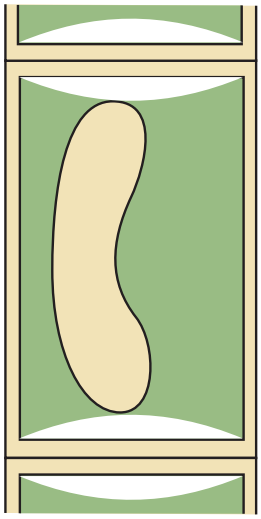


Differentiation within phloem vessels

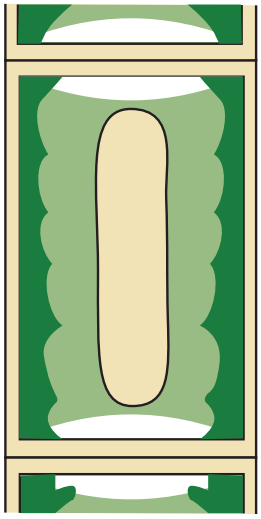
1. Xylem vessels transport water throughout a plant and provide it with support.



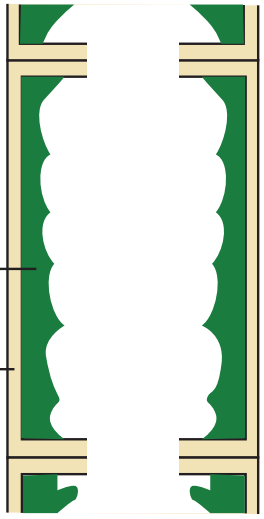
2. The nucleus disintegrates.



3. A strengthening material called **lignin** is deposited in a spiral on the inside of the cell wall. This gives support to the plant.



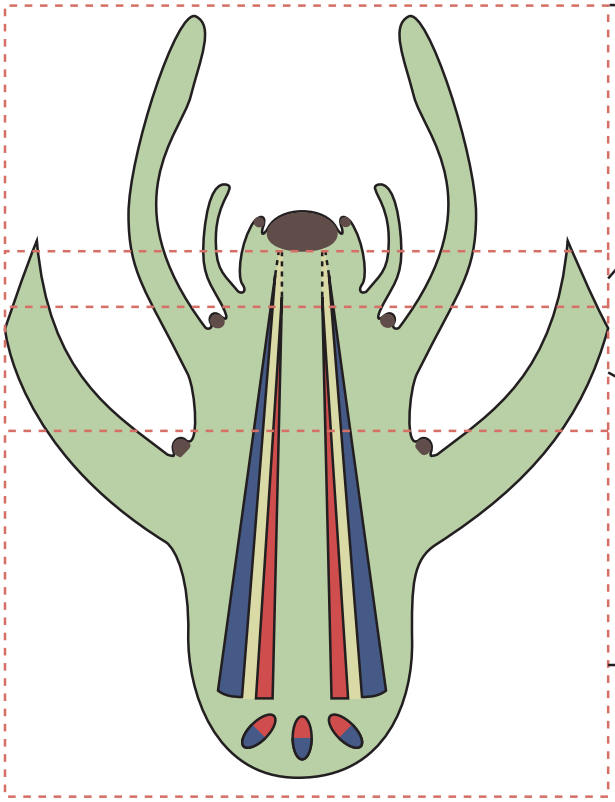
4. The cell contents die and the end walls disintegrate. This forms a hollow tube for water to be transported through.



Differentiation within xylem vessels

Shoot growth: Zones Go online

The following provides a summary of the zones involved in shoot growth.



Zone of cell division and mitosis
New cells are formed in the shoot apical meristem when the undifferentiated cells multiply by cell division and mitosis

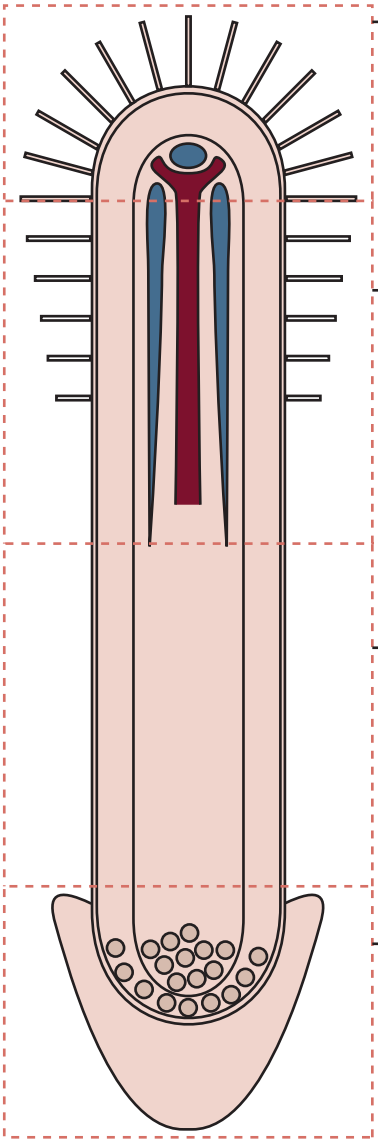
Zone of elongation and vacuolation
The undifferentiated cells produced in the meristem elongate and vacuoles form in them. This is the only zone in which growth occurs.

Zone of differentiation
No growth actually occurs in this zone, but undifferentiated cells differentiate. This forms specialised cells with particular functions.

Permanent tissues
The cells have fully differentiated to form specialised cells and do not undergo any more changes.

Root growth: Zones Go online

The following provides a summary of the zones involved in root growth.



Permanent tissues
The cells have fully differentiated to form specialised cells and do not undergo any more changes

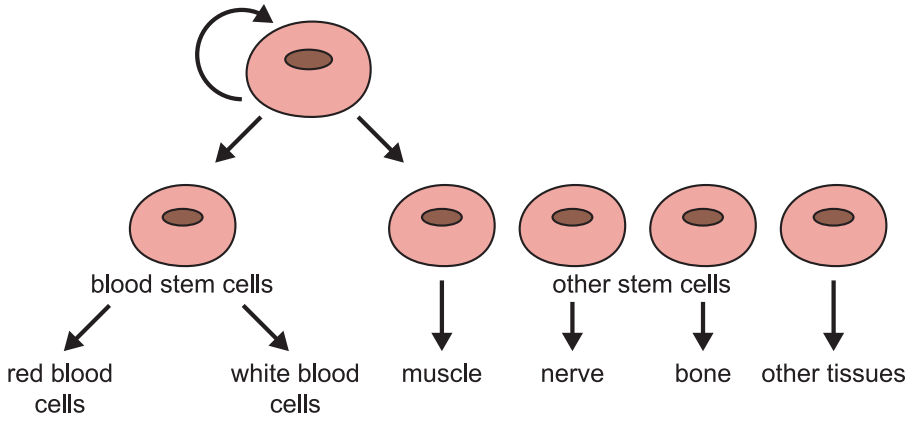
Zone of differentiation
No growth occurs in this zone but, the undifferentiated cells differentiate. This forms specialised cells with particular functions.

Zone of elongation and vacuolation
The undifferentiated cells produced in the meristem elongate and vacuoles form in them. This is the only zone in which the cells actually grow.

Zone of cell division and mitosis
New cells are formed in the root apical meristem when the undifferentiated cells multiply by cell division and mitosis. A lubricated root cap protects the apical meristem in the root tip. It also helps the root move through the soil.

4.3 Stem cells

Stem cells are unspecialised cells. When a stem cell divides, each new cell has the potential to remain a stem cell. This process is called **self-renewal**. In addition to self-renewal, stem cells can **differentiate** to become another type of cell with a more specialised function, such as a muscle cell, a red blood cell or a brain cell.



Stem cells

Stem cells are different from other body cells as they have the following characteristics:

1. Undifferentiated (unspecialised cell type), allowing them to divide and maintain a supply of stem cells for the body.
2. Found in all multicellular organisms.
3. Self-renewing and can differentiate; in some organs like the gut, stem cells regularly divide to repair and replace worn out or damaged tissues.

The two types of stem cells found in humans are:

1. Embryonic stem cells.
2. Tissue (adult) stem cells.

Stem cells: Questions Go online

Q1: What do you understand by self-renewal in stem cells?

. . . . . .

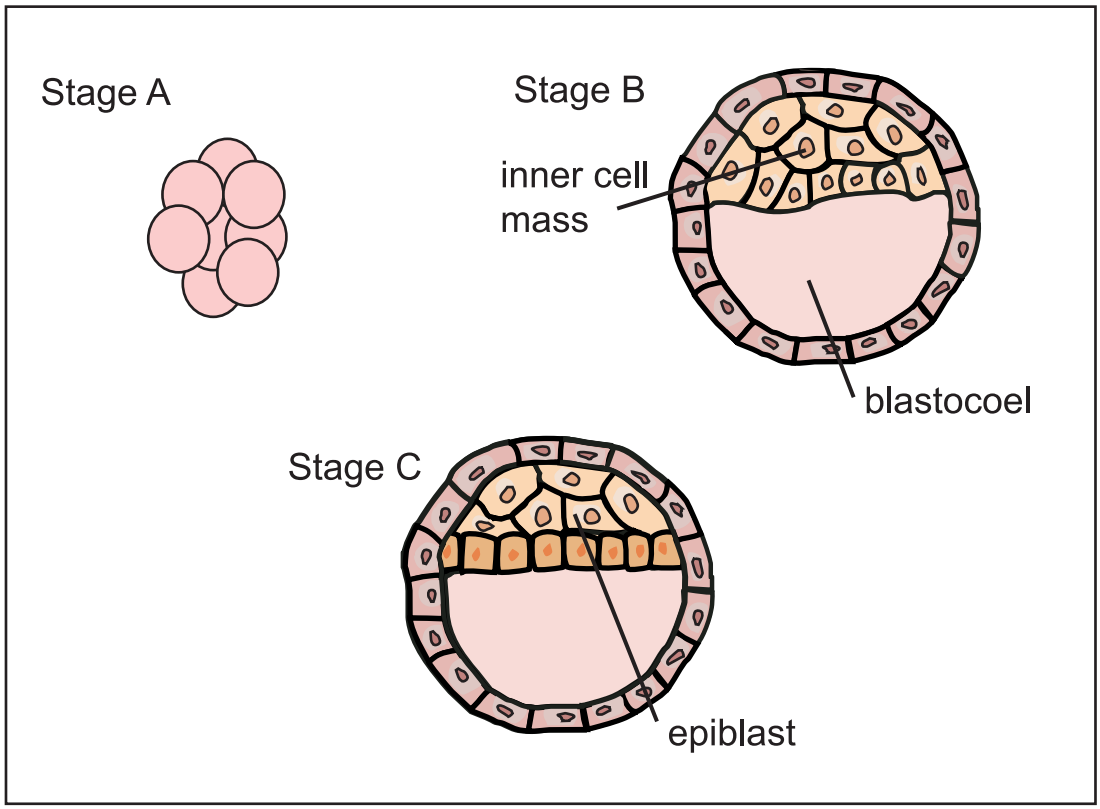
Q2: What is the unique property of stem cells which makes them different from a specialised cell?

. . . . . .

Q3: Can you give examples of differentiated cells and their functions?

4.4 Embryonic stem cells

Most embryonic stem cells are derived from embryos that develop from eggs that have been fertilised, but before preimplantation.



Stages in the early development of the human embryo

After the process of fertilisation, the zygote undergoes rapid cell division (stage A) and produces a multicellular ball called a blastula. The blastula contains a fluid-filled cavity called the blastocoel (stage B). In humans, the blastula becomes implanted in the uterus, and the cells of the inner cell mass begin to differentiate (stage C).

Embryonic stem cells are derived from embryos at the **blastocyst** stage. All the genes in embryonic stem cells can be switched on so these cells can differentiate into any type of cell. These cells are described as **pluripotent** as they can differentiate into all the cell types that make up the organism.

Human embryonic stem cells can be formed by transferring cells from a preimplantation-stage embryo into a plastic laboratory culture dish that contains a nutrient broth, known as culture medium. In more recent research into embryonic stem cells, scientists have reliably directed the differentiation of embryonic stem cells into specific cell types by using different culture conditions. They are able to use the resulting, differentiated cells to treat certain diseases.

Embryonic stem cells: Question Go online

Q4: Put the following stages of the process of using human embryonic stem cells (hESCs) to form specialised cells into the correct order.

- Formation of mass of undifferentiated stem cells
- Stem cell cultured in the laboratory
- Embryo stem cell removed
- Undifferentiated stem cells cultured in different culture conditions
- Formation of specialised cells: nerve cell, muscle cell, gut cells
- Early human embryo Blastocyst

4.5 Tissue (adult) stem cells

Early work on tissue stem cells started in the 1950s. Researchers discovered that bone marrow contains at least two kinds of stem cells. These were haematopoietic stem cells, which form all of the types of blood cells in the body, and bone marrow stem cells, which can generate bone, cartilage and fat cells that support the formation of blood and fibrous connective tissue.

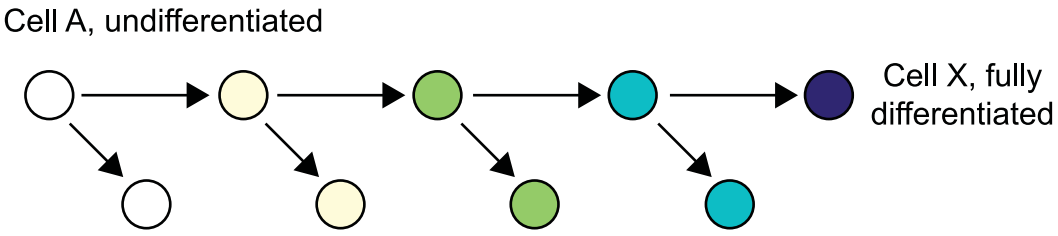
Tissue stem cells are involved in the growth, repair and renewal of the cells found in that tissue. They are multipotent. Tissue stem cells are multipotent as they can differentiate into all of the types of cell found in a particular tissue type. For example, blood stem cells located in bone marrow can give rise to all types of blood cell.

Tissue stem cells have been identified in many organs and tissues, including brain, bone marrow, peripheral blood, blood vessels, skeletal muscle and skin. It is also worth knowing that these stem cells are also found in foetuses and babies. Although found in many types of tissues, only a very small number of stem cells actually occur in these tissues.

4.6 Differentiation

Cellular differentiation is the process by which a cell develops more specialised functions by expressing the genes characteristic for that type of cell. For example, a white blood cell only expresses genes which relate to its function, such as those which produce antibodies.

Cellular differentiation is the result of gene expression, which is under the influence of many factors. Not all cells complete the process of differentiation; some cells pause at the stage where they can still undergo duplication. This allows them to generate replacement cells which may die or be damaged - it allows for growth and repair. These cells are known as tissue (or adult) stem cells.



Differentiation

The original cell has the power to differentiate into several different varieties of cell. Cells differentiate gradually over several rounds of division until the final differentiated cells can no longer reproduce themselves.

4.7 Therapeutic use of stem cells

Some types of stem cells have been used in medicine for a number of years to repair damaged or diseased organs. Some examples are listed below:

- Skin: a rich source of tissue stem cells. Patients with serious burns can be treated using a technique which grows new skin in the lab from skin stem cells.
- Blood: a type of stem cell found in the bone marrow is capable of giving rise to all of the different types of blood cells. Bone marrow transplants have been carried out for many years as a treatment for diseases such as leukaemia and other blood disorders.
- Cornea: corneal stem cells are removed and cultured in a laboratory. They are then transplanted onto the diseased cornea to repair it.

4.8 Research involving stem cells

Almost all animals contain a very small population of cells that retain the ability to reproduce daughter cells which will, in turn, replace differentiated tissues that have become worn, diseased or damaged. However, the power of the stem cell to regenerate could be potentially dangerous and, if poorly regulated, give rise to many forms of cancer. An understanding of how stem cells regulate their own growth and development is therefore extremely important.

To begin studying stem cells, it is necessary to develop a stem cell line. A stem cell line is a group of constantly dividing cells from a single parent group of stem cells. Stem cell lines are grown in culture dishes, allowing them to divide and grow as undifferentiated cells for many years.

Stem cell research allows scientists to discover more information about key cell processes such as growth, differentiation and gene regulation. Stem cells can also be used to study how diseases develop. Stem cells may be a viable alternative to animals for testing new drugs in the future. The potential benefits of using stem cells in medicine seem endless. In the future, scientists hope that stem cells will be used to cure conditions such as Alzheimer's disease, Parkinson's disease, diabetes, traumatic spinal cord injury, vision and hearing loss, Duchenne's muscular dystrophy, stroke and heart disease.

4.9 Ethical issues regarding stem cells

Stem cell research and therapy are regulated in this country by the Human Fertilisation and Tissue Authority, and the Human Tissue Authority. Researchers and clinicians have to work within strict guidelines outlined in the Human Fertilisation and Embryology Act 1990, which was revised in 2000 and updated in 2008. In some other countries, all work in this area is outlawed.

Stem cell research and the sourcing of stem cells have produced a great deal of argument and discussion. Many, from principally religious and moral stand-points, have argued against this work. This is because they believe that life begins at the point of fertilisation and that the zygote or **blastocyst** should be thought of, and treated, as if it was the same as a living being. Currently, embryos up to 14 days may be used. Other groups, such as patients awaiting treatment, are supportive of stem cell research.

Other areas of research are attempting to grow stem cells from adult and differentiated tissue using a technique known as **induced pluripotentency**. In this process, adult differentiated cells are taken and re-engineered back to embryonic-like cells. This could be a way of expanding research in stem cells by avoiding the ethical issues associated with the use of embryonic stem cells.

A study of ethics

The ethical matrix (designed by Professor Ben Mepham, Centre for Applied Bioethics at the University of Nottingham) is a tool to help people analyse an ethical issue and make an informed choice. It is based on three key ethical principles (for further information on these see Mepham: *Bioethics: an introduction for the Biosciences* Oxford University Press (2008), now in 2nd edition):

1. Wellbeing: the safety, welfare and health of an individual or group.
2. Autonomy: an individual's right to be free to choose and make their own decisions.
3. Justice: to what extent a situation is just or fair for an individual or group.

You are asked to consider the following questions with reference to the information in the ethical matrix to help explore your own opinions and feelings using, as far as possible, the evidence here and any other source you feel appropriate.

a) What do you think might be the priority of each of the interest groups?

b) In what way do you think that the three principals apply to each interest group?

c) To what extent do you think others might agree or disagree with you?

d) Might your decision be influenced by the thoughts and beliefs of others?

e) Can you suggest any way round some of the ethical issues that are raised by others?

Interest Groups	Wellbeing (safety, welfare and health)	Autonomy (freedom and choice)	Justice (fairness)
Patients - people who are hoping that stem cell therapies will treat an illness, disease or injury.			
Scientists - people working in stem cell research, developing stem cell therapies to treat patients.			
Embryo - the source of embryonic stem cells for research.			
Society - issues for wider society, such as social priorities, research and medical priorities, and how money should be allocated.			

4.10 Learning points

Summary

- Cellular differentiation is the process by which a cell expresses certain genes to produce proteins characteristic for that type of cell. This allows a cell to carry out specialised functions.
- Meristems are regions of unspecialised cells in plants that can divide (self-renew) and/or differentiate.
- Stem cells are unspecialised cells in animals that can divide (self-renew) and/or differentiate.
- In the very early embryo, a group of cells, the inner cell mass, can differentiate into almost all body tissues and so are pluripotent.
- Tissue stem cells are involved in the growth, repair and renewal of the cells found in that tissue. They are multipotent.
- Tissue stem cells are multipotent as they can differentiate into all of the types of cell found in a particular tissue type. For example, blood stem cells located in bone marrow can give rise to all types of blood cell.
- Stem cell research provides information about how cell processes, such as cell growth, gene expression and gene regulation, occur.
- Stem cells are used in therapies to repair and replace damaged organs and tissues; more therapies are being developed to treat diseases such as diabetes.
- Research uses involve stem cells being used as model cells to study how diseases develop or being used for drug testing.
- There are major ethical issues surrounding stem cell use and research; for example, the use of embryonic stem cells involves the destruction of embryos which some people believe is akin to murder.

4.11 Extended response question

The activity which follows presents an extended response question similar to the style that you will encounter in the examination.

You should have a good understanding of stem cells before attempting the question.

You should give your completed answer to your teacher or tutor for marking, or try to mark it yourself using the suggested marking scheme.

Extended response question: Stem cells

Describe the differences between and similarities of embryonic stem cells and tissue stem cells. *(6 marks)*

4.12 Extension materials

The material in this section is not examinable. It includes information which will widen your appreciation of this section of work.

Extension materials: The Jacob-Monod hypothesis

Although bacterial gene control is not strictly needed for the exam, the simpler mechanisms in prokaryotes lead to better understanding of control in complex cells.

There are many different ways in which gene expression is controlled. The mechanisms of gene expression are very complicated and are not fully understood in higher organisms. However, a lot is known about gene expression in bacteria thanks to work carried out by two French scientists. In the 1950s, Francois Jacob and Jacques Monod, two scientists at the Pasteur Institute in Paris, performed a series of experiments and determined how the production of the enzyme β-galactosidase was controlled in the bacterium *Escherichia coli*. They later won the Nobel prize for their work.

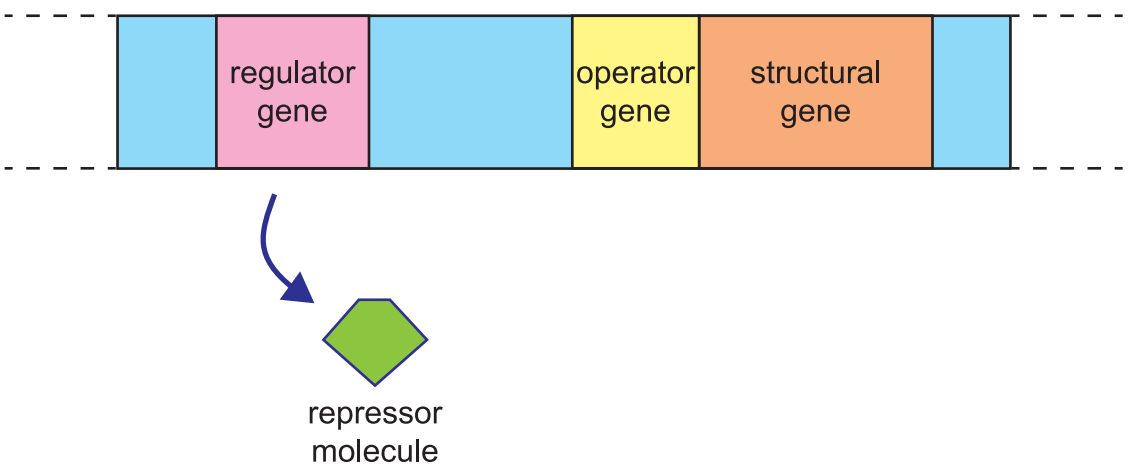
E. coli uses glucose during respiration to release energy. It will always use glucose if it is present in the environment. If no glucose is available, then it will utilise the glucose found in other energy sources, such as lactose. However, it is only able to use the glucose in lactose once it has been separated from galactose. The enzyme β-galactosidase breaks down lactose into glucose and galactose. Jacob and Monod found that *E. coli* only produces β-galactosidase when lactose is present in the nutrient medium in which the bacteria are growing. If lactose is absent, then no β-galactosidase is produced.

The gene that controls β-galactosidase production is 'switched on' (expressed) in the presence of lactose and 'switched off' (not expressed) when lactose is absent. Enzyme induction occurs, meaning that the gene is only switched on when the enzyme it codes for is required. In other words, *E. coli* is able to regulate the expression of the genes needed for lactose metabolism.

The Jacob-Monod hypothesis suggests that the production of β-galactosidase is controlled by an operon. An operon is a region of DNA that contains an operator gene that controls the expression of a structural gene which make the proteins or enzyme. A regulator gene that is found further along the DNA strand codes for a repressor molecule that interacts with the operator gene, preventing the expression of the structural gene.

When lactose is absent, the regulator gene produces a repressor molecule. This repressor molecule binds to the operator gene and this means that the structural gene is 'switched off'.

When lactose is present, it binds to the repressor molecule which means the operator gene is free. The operator gene then 'switches on' the structural gene which produces the enzyme β-galactosidase. This enzyme breaks down lactose into glucose and galactose. When the lactose is used up, the repressor molecule binds to the operator gene and the structural gene is once again 'switched off'.



The lac operon

4.13 End of topic test

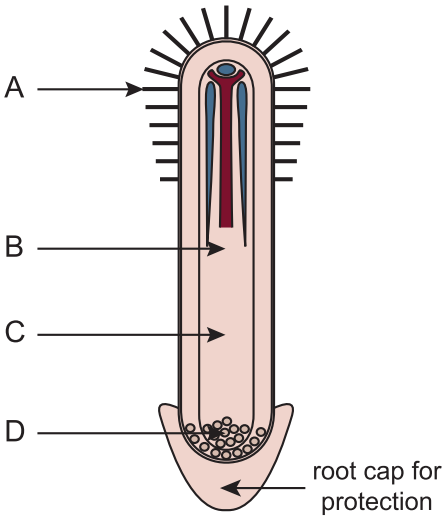
End of Topic 4 test Go online

Q5: Which of the following statements is **not** true of a plant meristem?

a) Growth can only occur at a meristem in a plant.
b) It contains unspecialised cells that differentiate.
c) It contains specialised cells that differentiate.
d) The roots and shoots have meristems.

. . . . . .

Q6: The following diagram shows a cross section through a root.



Which letter indicates the meristem?

. . . . . .

Q7: What is a stem cell?

. . . . . .

Q8: State two properties of stem cells.

. . . . . .

Q9: Name two sources of tissue (adult) stem cells.

. . . . . .

Q10: Why do our bodies need stem cells?

. . . . . .

Q11: Describe a current medical use of stem cells mentioned in this topic.

. . . . . .

Q12: What is the main ethical consideration regarding the use of embryonic stem cells?

Topic 5

Structure of the genome

Contents

Prerequisites

You should already know that:

- DNA carries the genetic information for making proteins.

Learning objective

By the end of this topic, you should be able to:

- understand the meaning of the term genome;
- describe that a section of DNA which produces a polypeptide is regarded as a gene;
- recognise that some DNA sequences code for proteins, while other sections do not;
- describe how some of the non-coding DNA has a role in controlling and regulating transcription;
- recognise that not all of the functions of non-coding DNA are as yet known or understood;
- describe the functions of RNA, including tRNA and rRNA.

5.1 Introduction to the Genome

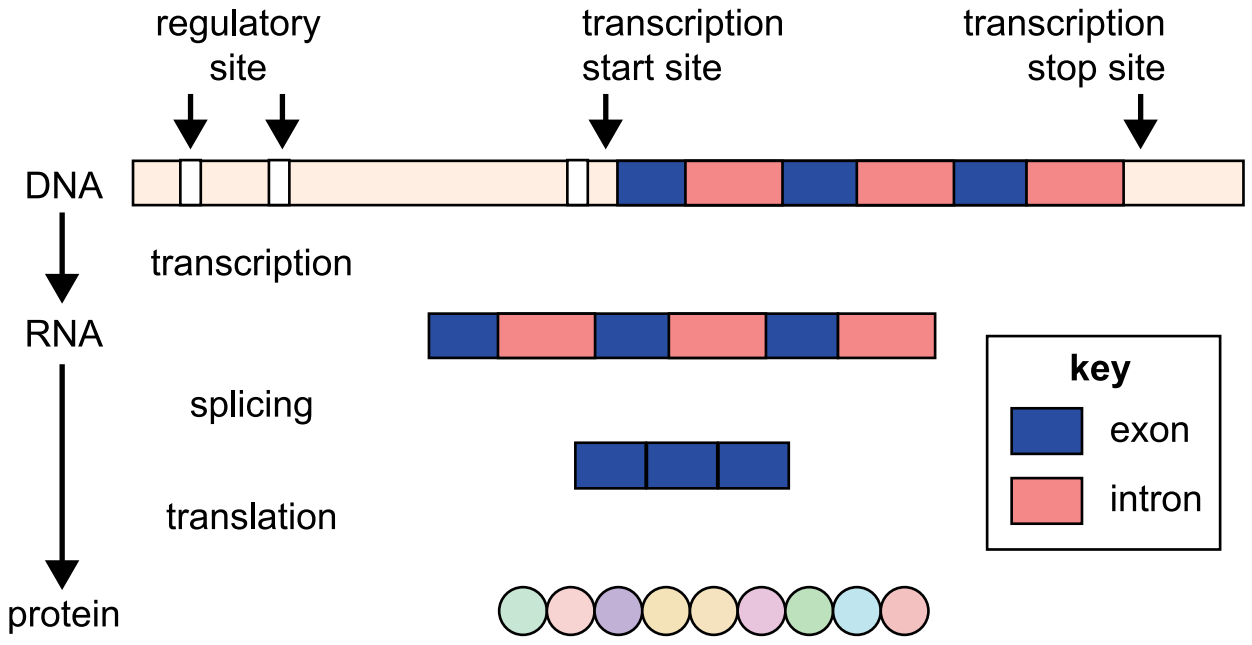
The **genome** is the sum total of all the hereditary material within an organism. It is usually taken to be the complete complement of DNA, although in some viruses this could be DNA or RNA.

The majority of the genetic information is carried within the nuclear DNA (linear chromosomes), but other sources exist. In bacteria, the DNA is found as a circular strand, sometimes called a chromosome, but lacking the associated packaging proteins. Bacteria often contain plasmids, which are small circular sections of DNA.

In eukaryotes, organelles such as **chloroplasts** and **mitochondria** also contain circular sections of DNA. Mitochondrial DNA is of great importance in hereditary studies as it is only passed down the maternal line. More often, the term genome is used to refer to nuclear DNA only. The mitochondrial DNA may be referred to as the mitochondrial genome and the DNA of the chloroplast as the **plastome**.

The human genome is recognised as consisting of 3×10^9 nucleotides. These are found as approximately 20 - 25,000 genes, arranged on 22 autosomal chromosomes, and a pair of sex chromosomes, either two X chromosomes or an X and a Y chromosome. The study of the properties of genomes is referred to as **genomics**, compared to the study of single genes or groups of genes, which is **genetics**.

5.2 The genome



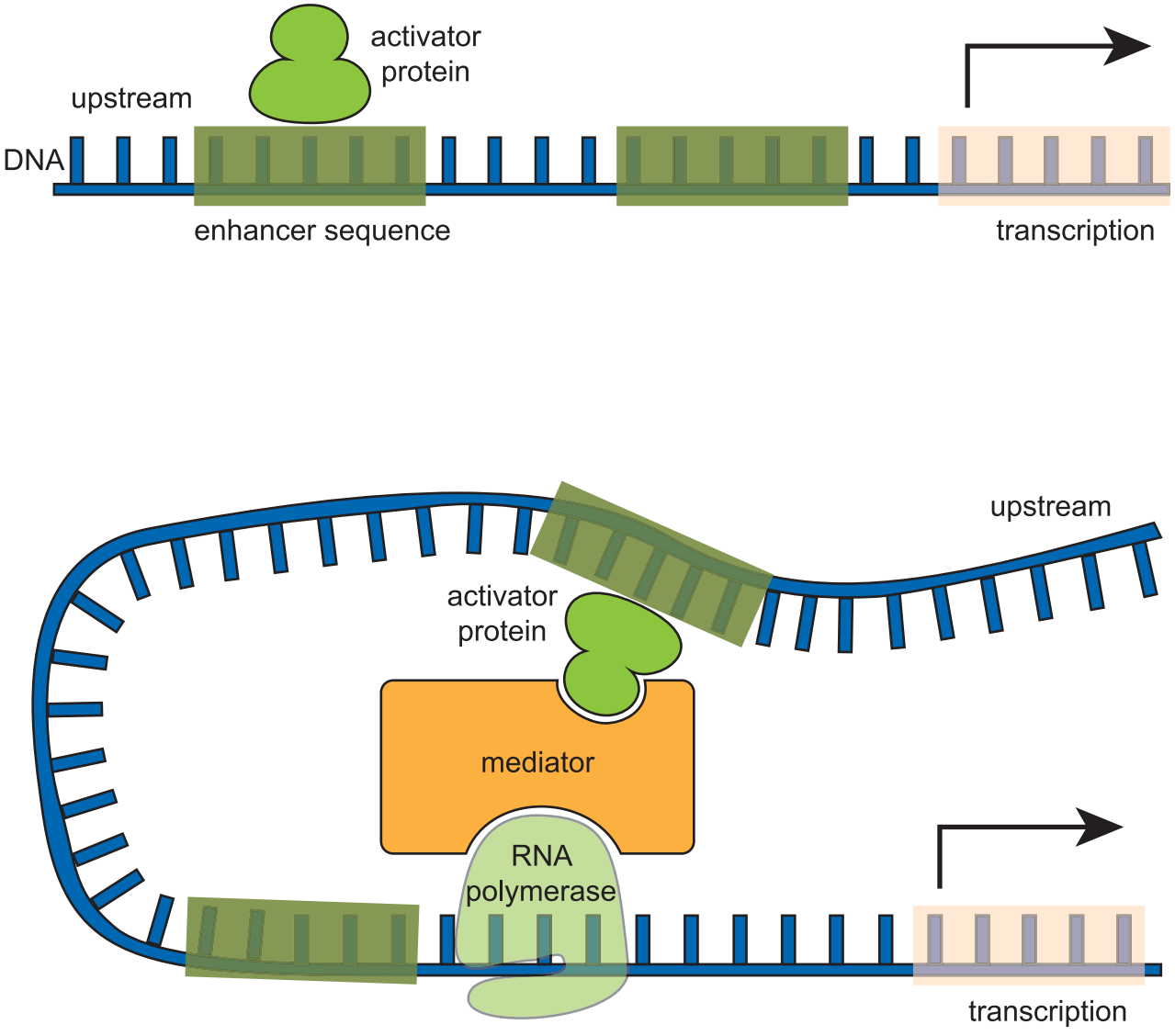
Processing from DNA to protein

An organism's genome is its genetic information encoded in its DNA. The genome contains many genes which carry instructions for making all of the proteins found in an organism. These regions of DNA are known as coding regions. The genome contains both coding and non-coding regions.

In fact, most of the genome is made up of non-coding sequences. These non-coding regions can have several functions:

- regulating transcription,
- transcription of RNA,
- no known function.

Some non-coding sections of DNA are used to regulate transcription. This means they can bind proteins which promote or prevent transcription of a gene. The diagram below illustrates how a sequence of DNA can regulate transcription of a gene.



Regulation of transcription

Some sections of DNA are transcribed into RNA but are not translated, for example tRNA, which carries specific amino acids to the ribosome during translation, and rRNA, which together with protein forms the ribosome. Another type of RNA which is not translated is RNA fragments. These are small sections of RNA which are involved in splicing, and other processes such as post-transcriptional regulation of genes.

The function of large sections of the genome are still unknown. It was once referred to as junk DNA, but it is now widely acknowledged that it serves a purpose.

The genome: Question Go online

Q1: Complete the table of genome terms by matching the descriptions from the list with the processes.

Process	Description
Transcription	
Splicing	
Translating	

Descriptions:

- DNA copied to RNA
- Exons pass to ribosome where polypeptides are assembled
- Introns removed from pre-mRNA

5.3 Learning points

Summary

- The genome of an organism is its hereditary information encoded in DNA.
- Coding regions are sections of DNA which contain a gene.
- Much of the genome is non-coding in that it does not contain genes.
- Some regions of the genome contain regulatory sequences which control the transcription of genes.
- Other non-coding regions contain sequences for producing non-translated forms of RNA such as tRNA, rRNA and RNA fragments.
- While some sections of non-coding DNA assist in the control and regulation of gene expression, there are sections whose function remains unknown.

5.4 Extension materials

The material in this section is not examinable. It includes information which will widen your appreciation of this section of work.

Extension materials: The Human Genome Project

The Human Genome Project began in October 1990 and was completed in 2003. The project involved the discovery of all the estimated 20,000 - 25,000 human genes, making them available for further studies. The project also led to the discovery of the complete sequence of the 3 billion DNA sub-units (the bases in the human genome). In April 2003, the completion of the human DNA sequence coincided with the 50th anniversary of Crick and Watson's description of the DNA structure in 1953.

Only about 3 percent of the human genome is actually used as the set of instructions. These regions are called coding regions. At present, little is known functionally for most of the remaining 97 percent of the genome; these regions are called non-coding regions.

Single nucleotide polymorphisms, or SNPs (pronounced "snips"), are DNA sequence variations where a single nucleotide (A,T,C,or G) in the genome sequence is altered. For example, a SNP might change the DNA sequence ACGGCTCA to ATGGCTCA.

Remember that there are 20 different amino acids. DNA is made up of four nucleotides, so if each were to specify (or code for) a single amino acid, only four amino acids could be coded for. A two letter code would give 16 (4^2) possible arrangements - still not enough to code for all 20 amino acids. The shortest unit that can code for all amino acids is a triplet code or codon. A triplet code produces 64 (4^3) possibilities - more than enough!

A variation is considered a SNP when it occurs in at least 1% of the population. SNPs make up about 90% of all human genetic variation and can occur every 100 to 300 bases along the 3 billion base human genome. About 66% of SNPs involve the replacement of cytosine (C) with thymine (T). SNPs can occur in coding and non-coding regions of the genome. They can act as biological markers, helping scientists in locating genes that are associated with certain diseases.

SNPs have no effect on cell function; scientists believe SNP maps will help them identify the multiple genes associated with complex ailments such as cancer, diabetes, vascular disease, and some forms of mental illness.

Craig Venter's genome was published in 2007. His genome contains 4.1 million variations; 3.2 million were SNPs. The following year, James Watson's genome was published, costing about £8 million and taking only 4 months. In 2010, the first personal genome machine came onto the market. This machine can sequence an individual's genome in about 12 days at a cost of £6,000!

Although SNPs do not cause diseases, they can help determine the likelihood that someone will develop a particular illness. One of the genes associated with Alzheimer's disease, *apolipoprotein E* (or ApoE), is a good example of how SNPs affect disease development. Scientists believe that SNPs may help them to discover and catalogue the unique sets of changes involved in different types of cancers. They are confident that SNPs can play an important role in the different methods used in the treatment of cancer.

Scientists are trying to identify all of the different SNPs in the human genome. They are sequencing the genomes of a large number of people and then comparing the base sequences to discover SNPs. The sequence data is being stored in computers that can generate a single map of the human genome, containing all possible SNPs.

5.5 End of topic test

End of Topic 5 test Go online

Q2: Complete the sentence using the words from list.

The _____ of an organism is its _____ information encoded in _____.

Word list: DNA; genome; hereditary

. . . . . .

Q3: DNA sequences that code for proteins are:

a) cistrons
b) exons
c) genes
d) introns

. . . . . .

Q4: All DNA codes for proteins. True or false?

. . . . . .

Q5: Name two types of RNA.

. . . . . .

Q6: The function of all DNA was worked out with the completion of the Human Genome Project.

True or false?

. . . . . .

Q7: By working out the nucleotide sequence of a genome, it is possible to describe all the organism's genes.

True or false?

Topic 6

Mutations

Contents

Prerequisites

You should already know that:

- DNA carries the genetic information for making proteins;
- a mutation is a random change to genetic material;
- mutations may be neutral, or confer an advantage or a disadvantage;
- mutations are spontaneous and are the only source of new alleles;
- environmental factors, such as radiation and chemicals, can increase the rate of mutation.

Learning objective

By the end of this topic, you should be able to:

- understand and explain the term mutation;
- describe and explain single gene mutations and their consequences;
- describe the impact of mutations on evolution;
- describe and explain chromosome structure mutations and their consequences.

6.1 Single gene mutations

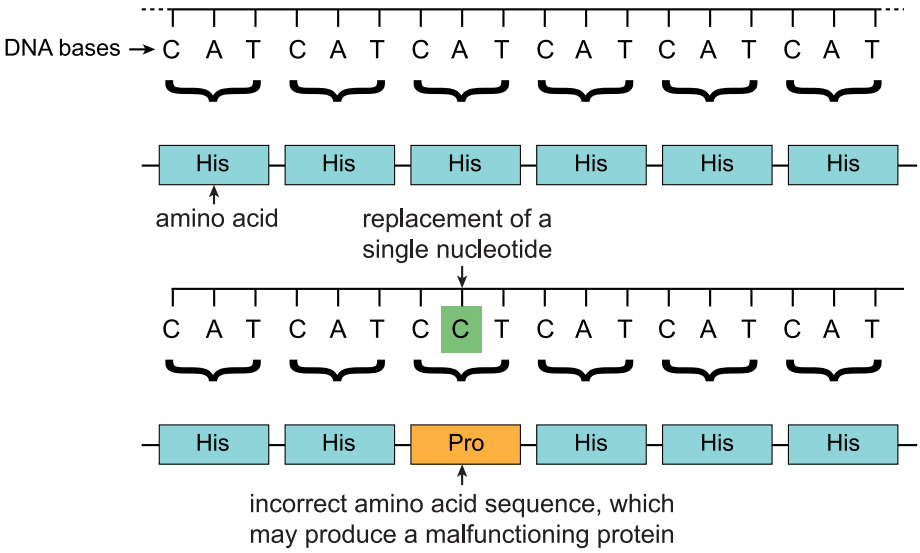
This section considers the types and effects of single gene mutations.

6.1.1 Types of mutation

Mutations are changes in the DNA that can result in no protein or an altered protein being synthesised. Single gene mutations involve changes to the number or the sequence of nucleotides within a single gene. There are three types of gene mutations: **substitution**, **insertion** and **deletion**.

Substitution mutation

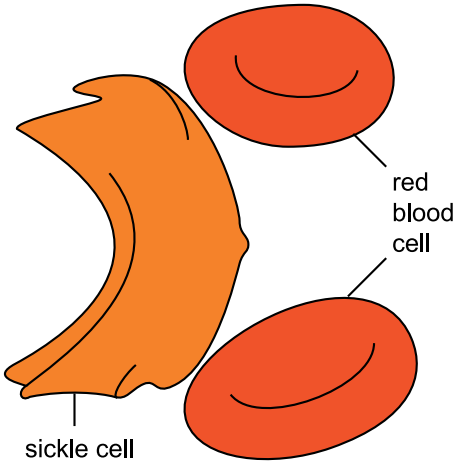
A substitution mutation means that one nucleotide is substituted for another and an incorrect amino acid may be inserted into a protein. Usually these changes are minor, but they can cause major problems in some cases, i.e. sickle cell anaemia.



A substitution mutation

A substitution mutation within a protein coding gene may not always lead to a change in the amino acid sequence of the encoded protein. In this case, the mutation is described as silent. For example, the amino acid leucine might be encoded in a protein by the base triplet GAT, for which the corresponding mRNA codon would be CUA. If the T in the DNA base triplet mutates to C (to give GAC), the mRNA codon becomes CUG, and this still encodes leucine. Substitution mutations affect only one amino acid (if any) in the encoded protein.

An example of a disease caused by a substitution mutation is sickle-cell anaemia. The affected gene encodes beta-globin, a protein that forms part of haemoglobin. A GAG codon is changed to GUG. The result is that the amino acid valine is coded for instead of glutamic acid. In individuals affected by this disease, many of the red blood cells form a characteristic sickle shape and can get trapped in blood vessels. This can cause extensive tissue and organ damage. The life span of the red blood cells is considerably reduced and, because they cannot be replaced quickly enough, the individual develops anaemia.



A sickle cell

(a)	**DNA**	TGA	GGA	**CTC**	CTC	TTC
	mRNA	ACU	CCU	**GAG**	GAG	AAG
	protein	thr	pro	**glu**	glu	lys

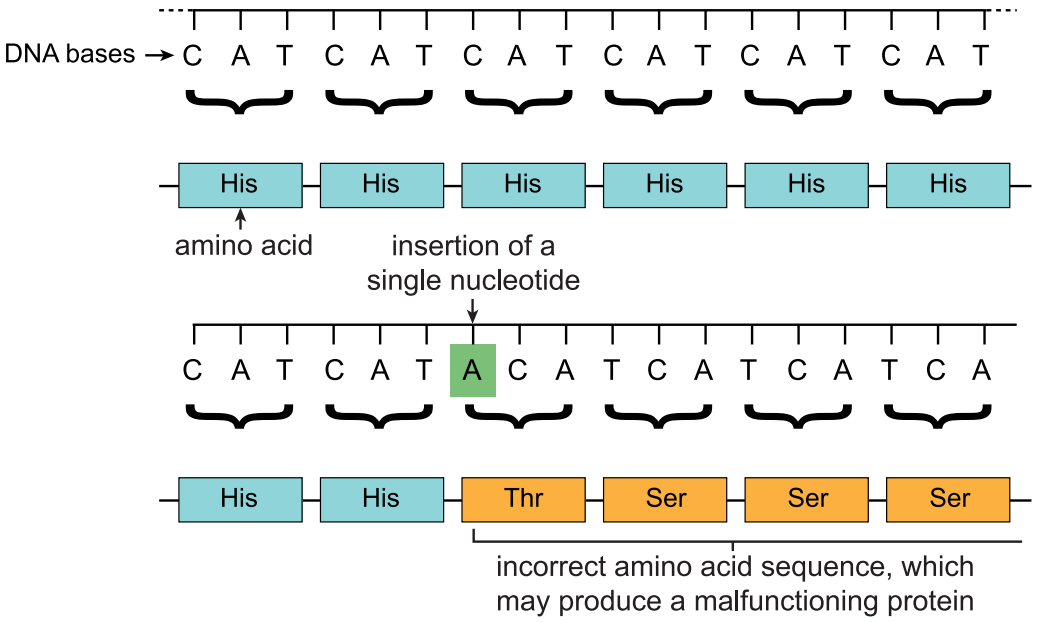
T mutates to A

(b)	**DNA**	TGA	GGA	**CAC**	CTC	TTC
	mRNA	ACU	CCU	**GUG**	GAG	AAG
	protein	thr	pro	**val**	glu	lys

The sickle cell genetic mutation

Insertion mutation

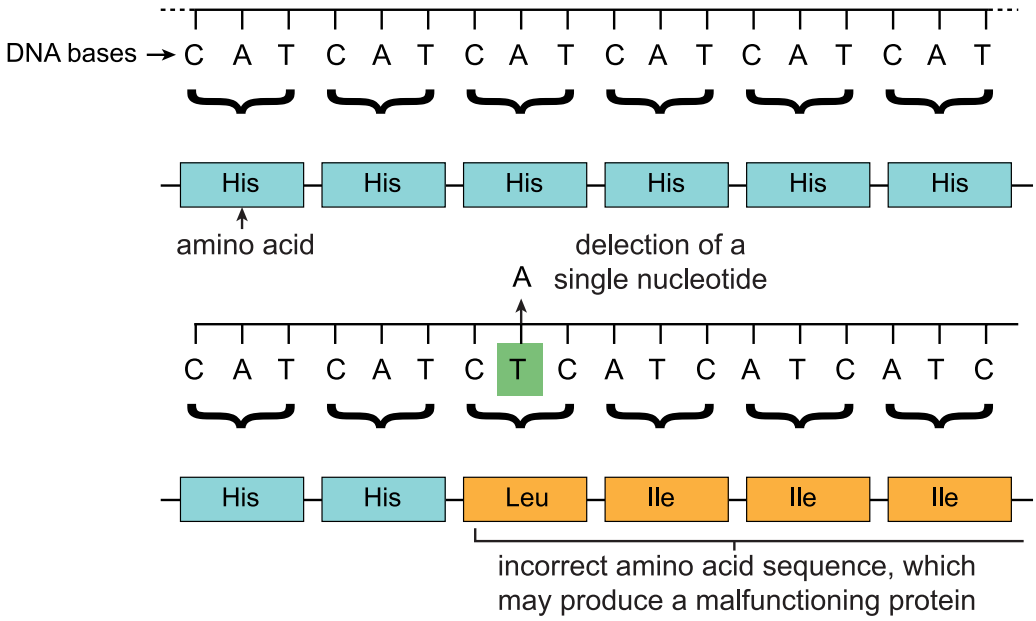
Insertion mutations are caused by the addition of one or more nucleotides into a section of DNA. If one or two nucleotides are inserted into a protein coding gene, this can have drastic effects on the protein which is produced because all of the subsequent triplets are read incorrectly. The protein which is made is therefore likely to have many different amino acids and may not work at all.



An insertion mutation

Deletion mutation

A deletion mutation refers to the removal of one or more nucleotides from the DNA. As with an insertion mutation, a deletion mutation alters the pattern of base triplets in the DNA. This means that deletions of one or two nucleotides are likely to cause drastic changes to a protein if they occur in a section of DNA containing a gene.



A deletion mutation

Gene mutations: Examples Go online

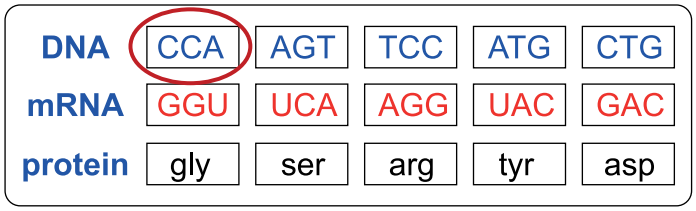
The following provides examples of the three types of gene mutations.

DNA	CCA	AGT	TCC	ATG	CTG
mRNA	GGU	UCA	AGG	UAC	GAC
protein	gly	ser	arg	tyr	asp

deletion mutation

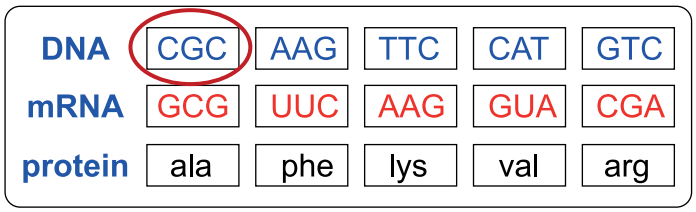
The base at position 7 in the DNA (T) has been deleted.

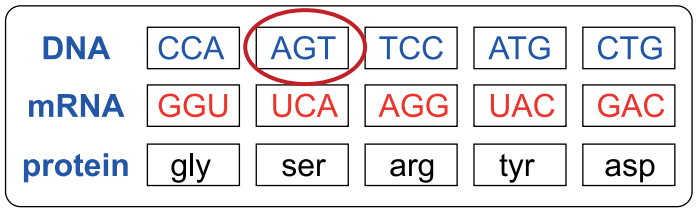
DNA	CCA	AGT	CCA	TGC	TGA
mRNA	GGU	UCA	GGU	ACG	ACA
protein	gly	ser	gly	thr	thr



insertion mutation

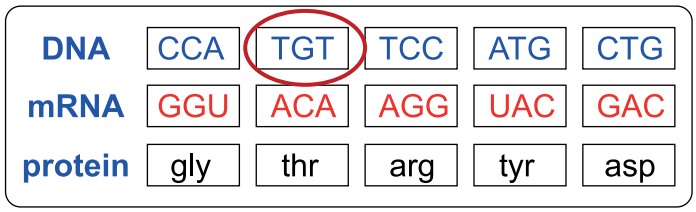
A new base (G) has been inserted at position 2 in the DNA.





substitution mutation

The base at position 4 in the DNA has changed from A to T.



Gene mutations: Questions Go online

Complete the sequences of proteins after mutation in each of the examples by referring to the genetic code table.

		Second letter					
		U	C	A	G		
First letter	U	phe	ser	tyr	cys	U	Third letter
	U	phe	ser	tyr	cys	C	
	U	leu	ser	stop	stop	A	
	U	leu	ser	stop	trp	G	
	C	leu	pro	his	arg	U	
	C	leu	pro	his	arg	C	
	C	leu	pro	gln	arg	A	
	C	leu	pro	gln	arg	G	
	A	ile	thr	asn	ser	U	
	A	ile	thr	asn	ser	C	
	A	ile	thr	lys	arg	A	
	A	met	thr	lys	arg	G	
	G	val	ala	asp	gly	U	
	G	val	ala	asp	gly	C	
	G	val	ala	glu	gly	A	
	G	val	ala	glu	gly	G	

Genetic code table

Q1:

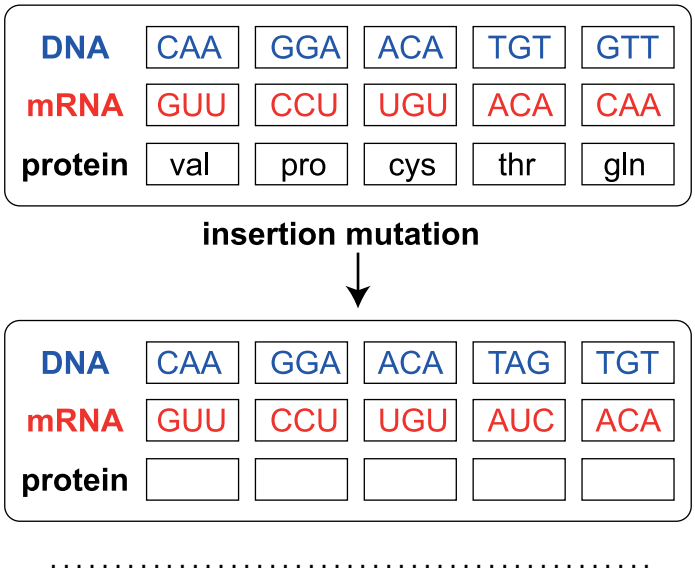
DNA	CAA	GGA	ACA	TGT	GTT
mRNA	GUU	CCU	UGU	ACA	CAA
protein	val	pro	cys	thr	gln

substitution mutation

DNA	CAA	GGA	ACC	TGT	GTT
mRNA	GUU	CCU	UGG	ACA	CAA
protein					

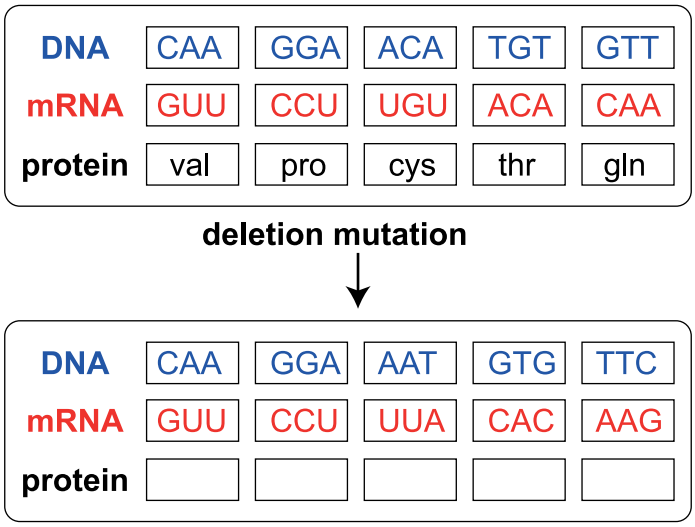
.....

Q2:



.....

Q3:

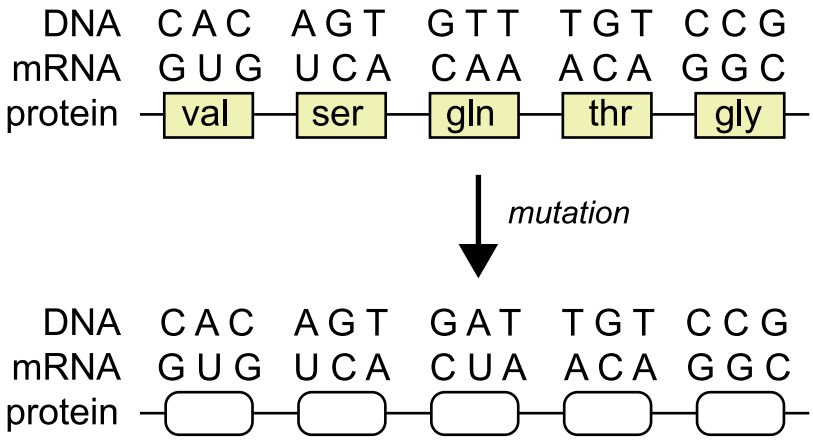


Complete the sequences of proteins after mutation in each of the examples by referring to the genetic code table and identify what kind of mutation has taken place.

		Second letter					
		U	C	A	G		
First letter	U	phe	ser	tyr	cys	U	Third letter
	U	phe	ser	tyr	cys	C	
	U	leu	ser	stop	stop	A	
	U	leu	ser	stop	trp	G	
	C	leu	pro	his	arg	U	
	C	leu	pro	his	arg	C	
	C	leu	pro	gln	arg	A	
	C	leu	pro	gln	arg	G	
	A	ile	thr	asn	ser	U	
	A	ile	thr	asn	ser	C	
	A	ile	thr	lys	arg	A	
	A	met	thr	lys	arg	G	
	G	val	ala	asp	gly	U	
	G	val	ala	asp	gly	C	
	G	val	ala	glu	gly	A	
	G	val	ala	glu	gly	G	

Genetic code table

Q4:

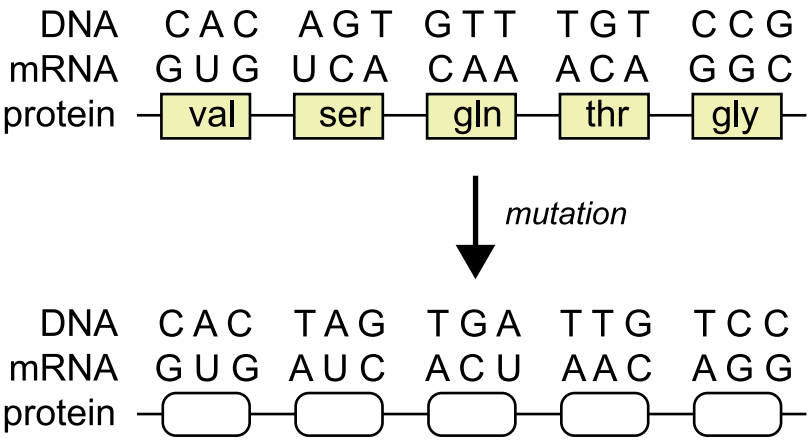


Which kind of mutation is shown? ____

(deletion, insertion or substitution?)

. . . . . .

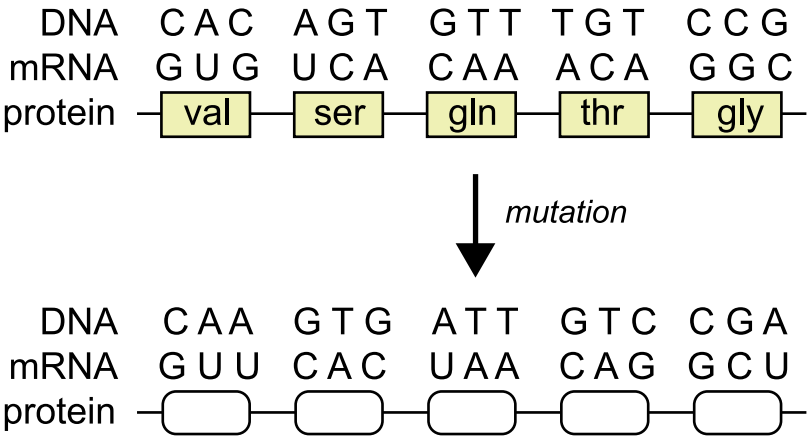
Q5:



Which kind of mutation is shown? _____
(deletion, insertion or substitution?)

.....

Q6:



Which kind of mutation is shown? _____
(deletion, insertion or substitution?)

6.1.2 Effects of mutations

The effect of a mutation will depend on its type and location. A protein requires the correct sequence of amino acids to function properly. If the base sequence of a gene is disrupted, the amino acid sequence may be disrupted as well.

A substitution mutation occurs when one base is swapped for another. The effects of this type of mutation will vary depending on where they occur. Some effects of substitution mutations include:

- missense,
- nonsense,
- splice site mutations.

A missense mutation results in a single incorrect amino acid being inserted into a protein. The effect this altered amino acid has on the function of the protein will vary depending on its location and chemical properties. A nonsense mutation results in the code for an amino acid being changed to a stop codon. This can result in an abnormally short protein which may not function properly.

Splice-site mutations result in some introns being retained and/or some exons not being included in the mature transcript and may result in a non-functional protein. Splice site mutations such as these alter post-transcriptional processing.

Insertions and deletions usually have greater effects than substitutions, especially if 1 or 2 bases are inserted or deleted. Nucleotide insertions or deletions result in frame-shift mutations.

Remember, mRNA is read in groups of three (codons). If 1 or 2 DNA nucleotides are inserted or deleted, all the bases downstream are moved up or down from their place; this means the reading frame is altered. Frame-shift mutations cause all of the codons and all of the amino acids after the mutation to be changed. This has a major effect on the structure of the protein produced.

If the mutation occurs early in the sequence, then the overall effect is far greater than if it occurred later. In addition to coding for different amino acids, the 'stop' sequence will become misplaced which could result in the polypeptide being either too long or too short, but in any event greatly altered.

(a)	**DNA**	GAT	GAT	GAT	GAT	GAT
	mRNA	CUA	CUA	CUA	CUA	CUA
	protein	leu	leu	leu	leu	leu
			C ↓			
(b)	**DNA**	GAT	GAT	GAT	GAT	GAT
(c)	**DNA**	GAT	CGA	TGA	TGA	TGA
	mRNA	CUA	GCU	ACU	ACU	ACU
	protein	leu	ala	thr	thr	thr

A frame shift mutation in DNA caused by insertion of one nucleotide

Effects of gene mutations on amino acid sequences: Questions

Go online

A gene is a region of DNA which consists of a specific sequence of nucleotide bases arranged in triplets. Every amino acid is coded for by one or more of these triplets. Therefore, the sequence of bases determines the sequence in which amino acids are joined together to form a polypeptide or protein. Thus, a gene codes for a particular protein or polypeptide.

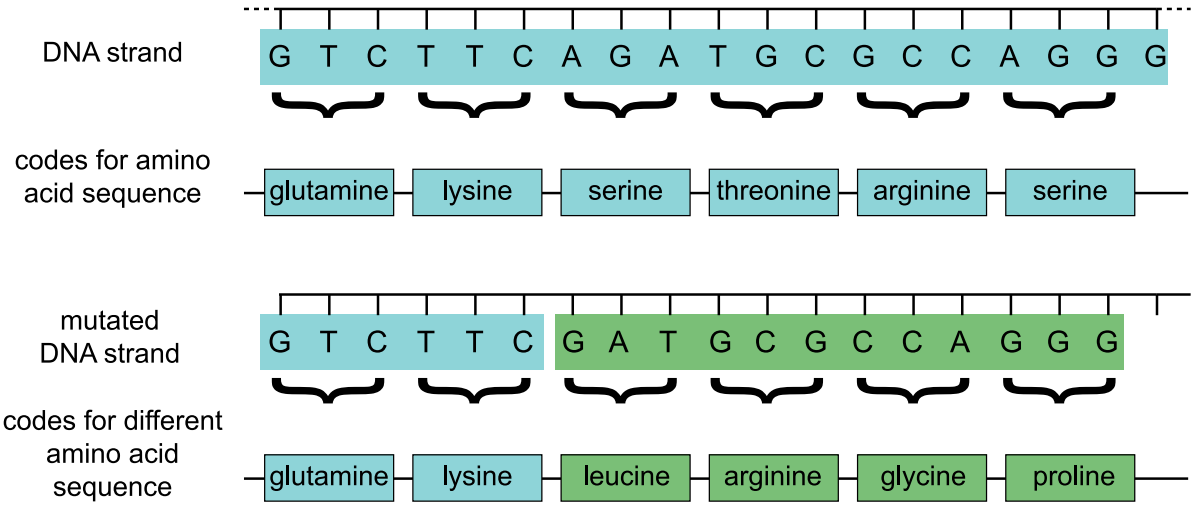
A gene mutation is a change in the sequence or type of nucleotide bases in a strand of DNA. This can lead to a change in the sequence of amino acids, and thus to a change in the protein which is synthesised in the ribosomes.

Sometimes a mutation causes only a minor change, perhaps affecting only one amino acid. This is known as a *point mutation* and the protein produced may be only slightly altered and still able to carry out its function.

Sometimes, however, a mutation can cause a major change affecting the coding for many amino acids. Such a mutation is known as a *frameshift mutation* and leads to a completely different protein being produced, which cannot carry out the required function.

Each of the four types of gene mutation are described in the following diagrams, with particular respect to changes in the amino acid sequences, which show the codons of part of a DNA strand and the amino acids which are coded for by them. Remember that the same amino acid can be coded for by more than one DNA triplet.

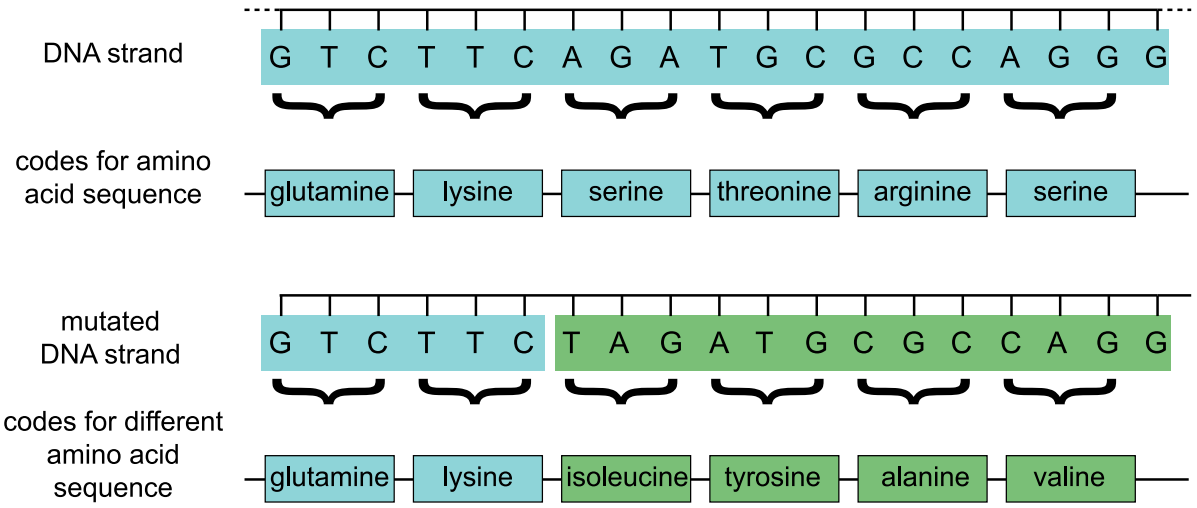
Q7: This is a deletion mutation. The first 'A' nucleotide in the original DNA strand has been removed (or deleted).



How many amino acids in this sequence have been changed? Is a deletion mutation a point or a frameshift mutation?

.....

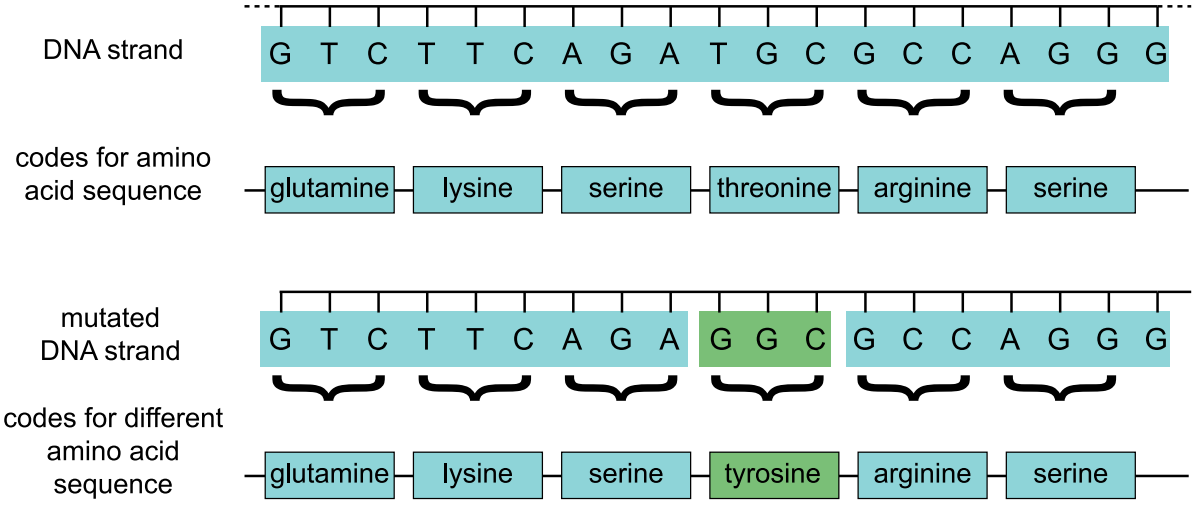
Q8: This is an insertion mutation. The 'T' nucleotide has been inserted between the 'C' and 'A' nucleotides in the original DNA strand.



How many amino acids in this sequence have been changed? Is an insertion mutation a point or a frameshift mutation?

. . . . . .

Q9: This is a substitution mutation. The 'T' nucleotide in the middle of the original DNA strand has been replaced (or substituted) by a 'G' nucleotide.



How many amino acids in this sequence have been changed? Is a substitution mutation a point or a frameshift mutation?

Q10: A deletion of three nucleotides from the middle of a gene may not lead to such a major change in the structure of the encoded protein as the deletion of one nucleotide. Why?

. . . . . .

Q11: Sometimes, a mutation (even the change of one base) will lead to a protein being produced that is shorter than the normal protein. Why is this?

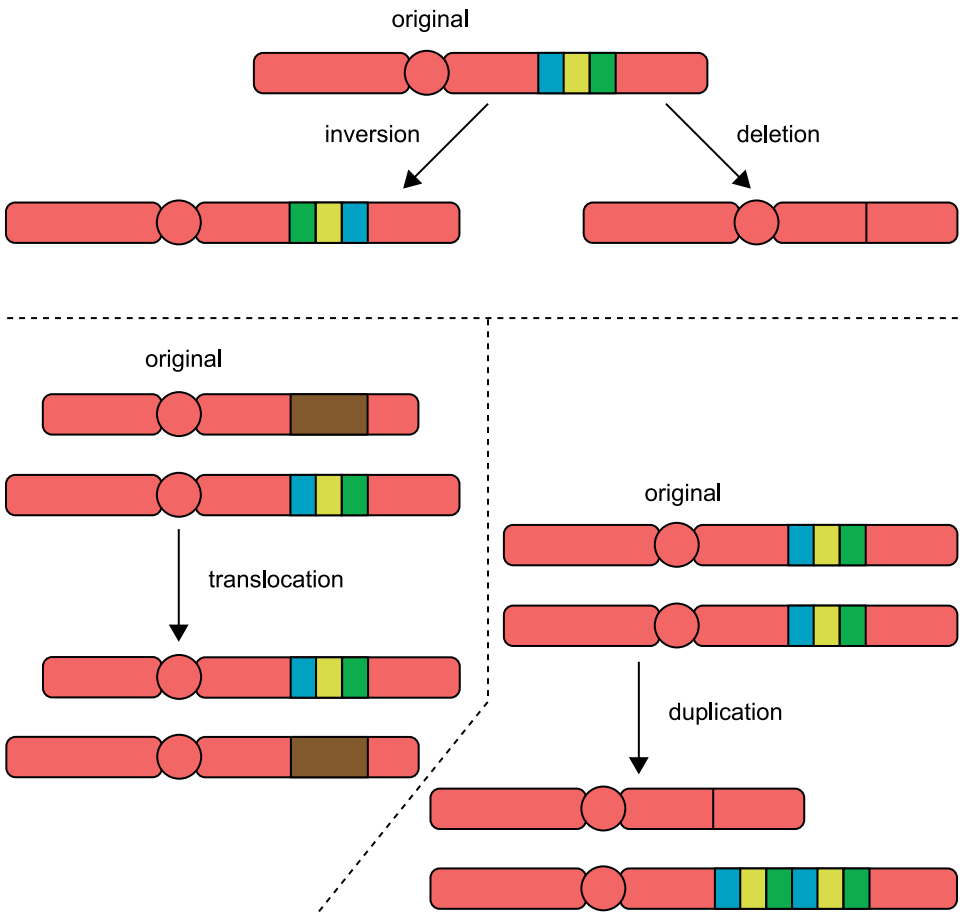
6.2 Chromosome structure mutations

Many types of mutations can arise through changes in chromosome structure. A change in chromosome structure starts when a chromosome breaks. The cell then attempts to repair the break, but in doing so may not restore the chromosome to its original structure. Changes in chromosome structure can be very large, with the result that many genes may be affected. Each type of chromosome mutation is briefly described below:

- **Translocation**: a section of a chromosome is added to another chromosome, not its homologous partner.
- **Deletion**: a section of a chromosome is removed.
- **Duplication**: a section of a chromosome is added from its homologous partner.
- **Inversion**: where a section of chromosome is reversed.

Chromosome mutation: A visual representation Go online

The following provides a visual representation of the changes that occur in chromosome structure during the three types of chromosome mutations.



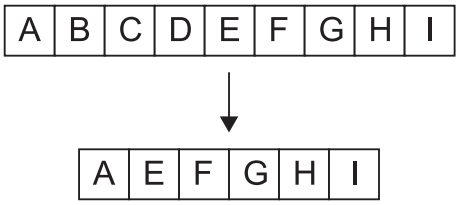
Differences between gene and chromosome mutation: Questions

Go online

A mutation is a change in the genetic material of an organism. Such changes can take place in the number or structure of chromosomes (known as chromosome mutations), or in the sequence of nucleotide bases in genes (known as gene mutations).

In each of the following examples, consider first of all whether it is a chromosome or a gene mutation, and then which particular type of mutation it is. Answer the questions by typing your answers in the boxes provided.

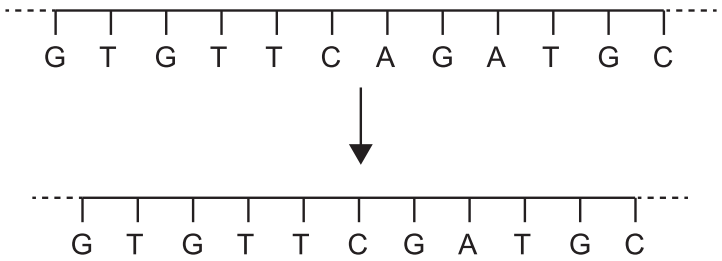
Q12:



Is this a chromosome or a gene mutation and what is the name of this type of mutation?

.....

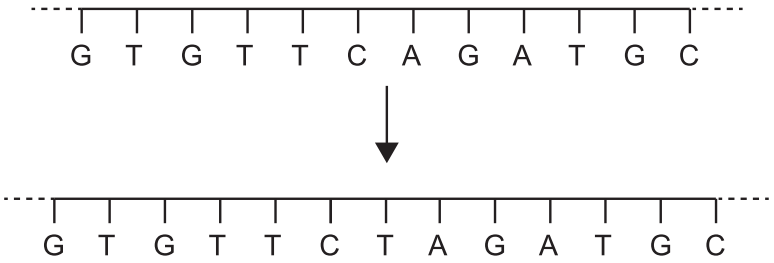
Q13:



Is this a chromosome or a gene mutation and what is the name of this type of mutation?

.....

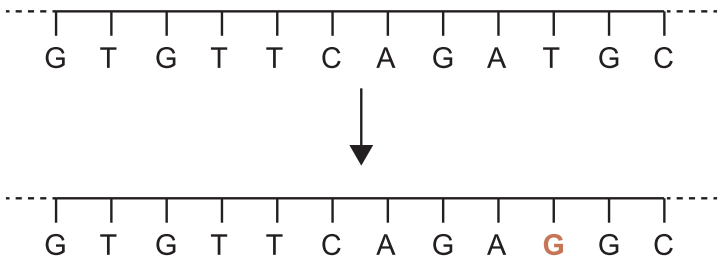
Q14:



Is this a chromosome or a gene mutation and what is the name of this type of mutation?

.....

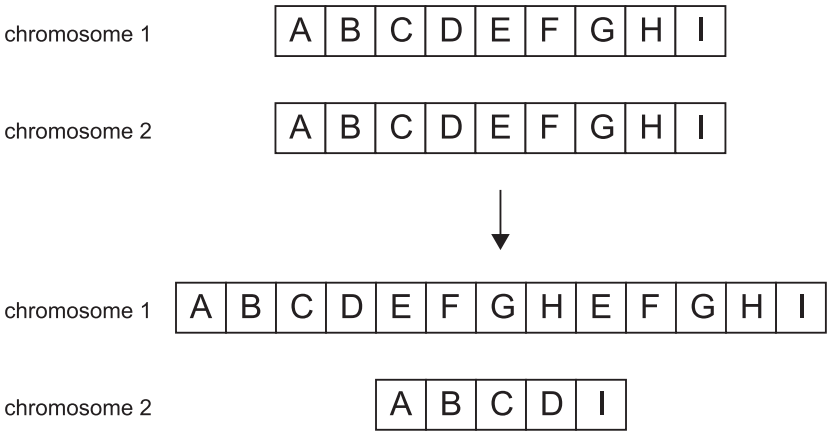
Q15:



Is this a chromosome or a gene mutation and what is the name of this type of mutation?

.....

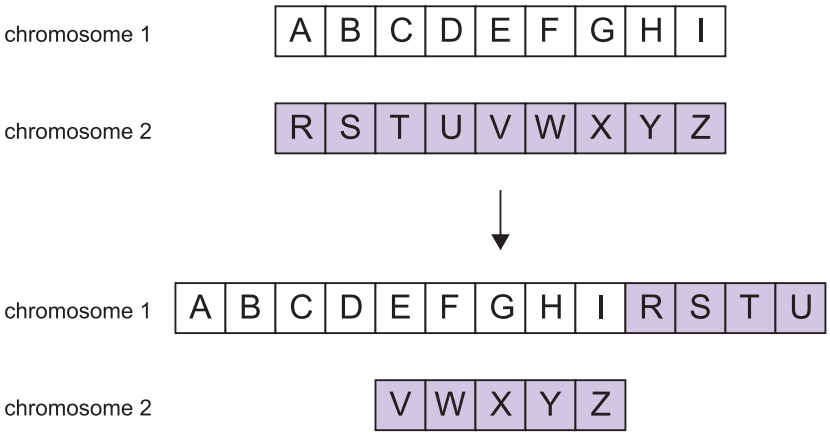
Q16:



Is this a chromosome or a gene mutation and what is the name of this type of mutation (that occurs in chromosome 1)?

.....

Q17:



Is this a chromosome or a gene mutation and what is the name of this type of mutation (that occurs in chromosome 1)?

6.3 The importance of mutations and gene duplication

Mutations provide new variation. Mutations bring about new variation by the production of new alleles. Without mutations there would be no new variation.

Occasionally a mutation can result in the duplication of an entire gene. The second copy of the gene can become altered and provide new DNA sequences. This gene duplication is thought to be an important driving force in evolution.

6.4 Learning points

Summary

- A mutation is any change in the DNA sequence of the genome.
- Mutations can occur at random, but may also be induced by a number of agents.
- Mutations can occur at various levels. It could be change in a single nucleotide, allele, gene or chromosome.
- Point mutations are changes at the single nucleotide level, in that one nucleotide is substituted for another or a nucleotide is inserted or deleted from the DNA sequence.
- As a result of the change in the DNA sequence there may be a change in the amino acid sequence, and hence the structure of the final protein.
- Missense mutations result in one amino acid being changed for another. This may result in a non-functional protein or have little effect on the protein.
- Nonsense mutations result in a premature stop codon being produced which results in a shorter protein.
- Splice-site mutations result in some introns being retained and/or some exons not being included in the mature transcript.
- Frame-shift mutations cause all of the codons and all of the amino acids after the mutation to be changed. This has a major effect on the structure of the protein produced.
- As a result of mutations producing altered proteins, organisms may change or evolve.
- Mutations may occur at the chromosomal level, either within the chromosome or between chromosomes.
- The sequence of genes on a chromosome may be changed by duplication, deletion, translocation or inversion.
- Duplication is where a section of a chromosome is added from its homologous partner.
- Deletion is where a section of a chromosome is removed.
- Inversion is where a section of chromosome is reversed.
- Translocation is where a section of a chromosome is added to a chromosome, not its homologous partner.

Summary continued

- The substantial changes in chromosome mutations often make them lethal.
- Duplication allows potential beneficial mutations to occur in a duplicated gene whilst the original gene can still be expressed to produce its protein.

6.5 Extended response question

The activity which follows presents an extended response question similar to the style that you will encounter in the examination.

You should have a good understanding of gene mutations before attempting the question.

You should give your completed answer to your teacher or tutor for marking, or try to mark it yourself using the suggested marking scheme.

Extended response question: Gene mutations

Describe gene mutations and outline some of their consequences. *(8 marks)*

6.6 Extension materials

The material in this section is not examinable. It includes information which will widen your appreciation of this section of work.

Extension materials: Lactose tolerance

In most parts of the world, the ability of humans to digest lactose declines rapidly after infancy. Lactose is a sugar that is found in both human breast milk and cow's milk. However, many adults in Northern Europe, and to some extent around the Mediterranean, retain the ability to digest lactose, and are thought to have developed this through a mutation followed by evolutionary adaption. It has been proposed that this ability may have arisen in response to living at 'high' latitude which, as a result of diminished sunlight, may lead to reduced levels of vitamin D. The vitamin is associated with calcium uptake and its absence could lead to bone malformation - rickets for example.

By drinking fresh milk, which is high in calcium, or ingesting other dairy products, the problem of calcium deficiency could be overcome.

Through research and analysis, it has been suggested that the most probable explanation to the lactose tolerance is an evolutionary adaptation to millennia of milk drinking from domestic livestock. It has also been shown that the process of milking predated the evolution of lactose digestion. This would suggest that, as the lactose was present in the environment, there was an adaptation to exploit it.

6.7 End of topic test

End of Topic 6 test Go online

Q18: Which of the following correctly describes the characteristics of mutations?

a) They are non-random, frequent occurrences.
b) They are random, frequent occurrences.
c) They are non-random, infrequent occurrences.
d) They are random, infrequent occurrences.

. . . . . .

Q19: What name is given to a mutation where one part of a chromosome becomes attached to another?

a) Deletion
b) Duplication
c) Inversion
d) Translocation

. . . . . .

Q20: Which of the following chromosome mutations is most likely to be lethal?

a) Deletion
b) Duplication
c) Inversion
d) Translocation

. . . . . .

Q21: The figure below shows part of the normal nucleotide sequence of a gene. N refers to an unknown nucleotide. The figure below also shows the effects of different types of mutations, indicated by A to C, on the nucleotide sequence.

Normal sequence	N	N	**C**	**A**	**C**	**G**	**T**	**A**	**A**	**C**	**G**	**T**	N	N
A	N	N	**C**	**A**	**C**	**G**	**T**	**A**	**A**	**C**	**C**	**G**	**T**	N
B	N	N	**C**	**A**	**G**	**T**	**A**	**A**	**C**	**G**	**T**	N	N	N
C	N	N	**C**	**A**	**C**	**G**	**A**	**A**	**A**	**C**	**G**	**T**	N	N

In the order A to C, what are the different types of mutations shown in the figure?

a) Substitution, insertion, deletion.
b) Insertion, substitution, deletion.
c) Insertion, deletion, substitution.
d) Substitution, deletion, insertion.

. . . . . .

Q22: Which of the following gene mutations may result in a frameshift mutation?

a) Insertion or substitution.
b) Substitution or deletion.
c) Insertion or deletion.

. . . . . .

Q23:

1. What word is used to describe a mutation which results in the addition of a premature stop codon?
2. What word is used to describe a mutation which results in the addition of an incorrect amino acid into a protein?

Topic 7

Evolution

Contents

Prerequisites

You should already know that:

- a mutation is a random change to genetic material;
- mutations may be neutral, or confer an advantage or a disadvantage;
- mutations are spontaneous and are the only source of new alleles;
- an adaptation is an inherited characteristic that makes an organism well suited to survival in its environment/niche;
- variation within a population makes it possible for a population to evolve over time in response to changing environmental conditions;
- natural selection/survival of the fittest occurs when more offspring are produced than the environment can sustain, only the best adapted individuals survive to reproduce, passing on the genes that confer the selective advantage;
- speciation occurs after a population becomes isolated whereas natural selection follows a different path due to different conditions/selection pressures.

Learning objective

By the end of this topic, you should be able to:

- understand that changes in organisms take place over a long period of time and are caused by mutations;
- describe how genetic information is passed vertically during sexual and asexual reproduction, from parent to offspring;
- describe how genetic material can be passed horizontally in prokaryotes, resulting in very rapid evolution;
- explain that selection is a non-random process that leads to the increased presence of a particular gene or genes in a population;
- describe the process of natural selection;
- understand that deleterious (harmful) sequences are selected against, and removed from the population;
- understand that speciation is the process of generating a new species;
- explain how speciation is influenced by allopatric and sympatric factors.

7.1 Evolution

Evolution describes the changes that occur to a **species** over time, leading to offspring that are better adapted to survive in their environment than the previous generation. For evolution to occur, there must be changes to the **gene pool** and hence changes to the frequency of genes. The gene pool refers to all of the different genes of a particular species.

The **allele frequency** refers to the frequency of any **allele** in the population. The allele frequency is sometimes called the gene frequency, but this can be misleading as the term is used to describe the frequency of alleles, not genes. Remember that a gene may have several different alleles. While the frequency of a gene may not change, the frequency of each allele can. For example, the gene for eye colour may remain at a constant level within a population, but the frequency of the allele for blue eyes may increase or decrease.

Changes in the allele frequency can occur by several different mechanisms, as described in the next table.

Mechanism	Definition
Mutation	Creates multiple alleles for many genes in the gene pool.
Gene migration	The movement of alleles between populations by individuals arriving from a different population and breeding. These individuals have a different gene pool and therefore introduce new alleles into the population.
Genetic drift	Tends to occur in small populations and describes the change in allele frequency due to a chance event. Small populations that are isolated from each other can vary greatly due to changes in allele frequencies.
Non-random mating	Does not change the frequency of the alleles, but increases the number of homozygous individuals. Inbreeding is the most common form.
Natural selection	The frequency of an allele increases in a population if it provides a selective advantage.
Chance	Changes to the allele frequency due to random loss. For example, an individual possessing a certain allele may die or fail to reproduce so that allele is lost from the population.

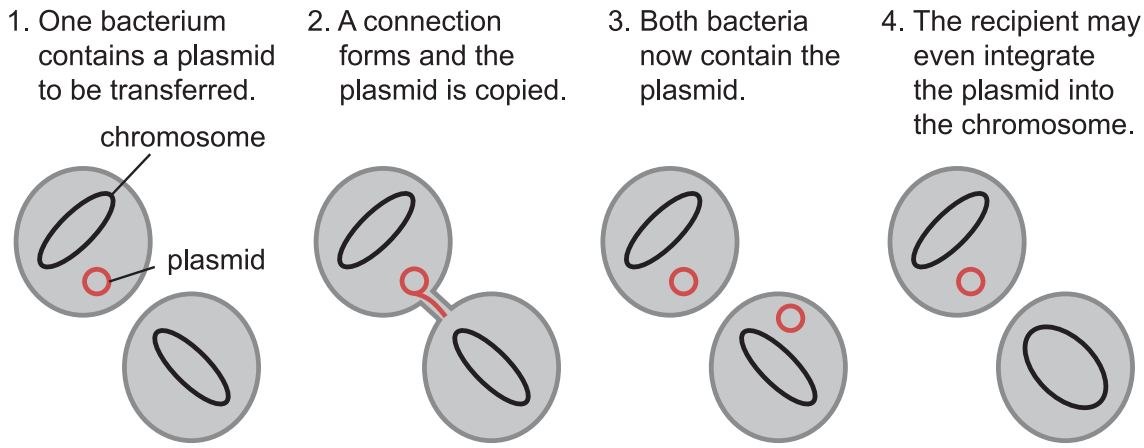
Factors affecting allele frequency

7.2 Gene transfer

Eukaryotes can reproduce by sexual or asexual reproduction. During sexual reproduction, the genetic material of two parents is combined to produce a new organism. During asexual reproduction, a new organism is produced from a single parent. Both sexual and asexual reproduction are examples of vertical gene transfer, a process by which genes are transferred from parent(s) to offspring.

In prokaryotes, reproduction is most frequently carried out asexually by a form of mitosis called binary fission. Again, this is an example of vertical gene transfer. However, there are occasions where prokaryotes can pass genetic material between themselves. The genetic material may be part of the single circular chromosome or a plasmid. This type of inheritance is called horizontal gene transfer because genes are passed between members of the same generation, not between parents and offspring. Viruses can also carry out horizontal gene transfer.

As with any other organism, prokaryotes are subjected to environmental pressures that sometimes cause mutations. Mutations can appear quite rapidly because prokaryotes exist in massive numbers. Once in the population, mutations can be passed between members with ease by horizontal gene transfer. Horizontal gene transfer allows prokaryotes to experience rapid evolutionary change.



7.3 Selection

This section considers natural selection and the effects of stabilising, disruptive and directional selection.

7.3.1 Natural selection

The environment surrounding all living organisms is never static and, as a result, they are constantly under pressure to respond to the changes in order to survive. Through sexual reproduction, the genetic material being passed between generations is subjected to constant change and rearrangement. As a result, after many generations the genome will be altered.

Natural selection is the mechanism by which evolution occurs. It is a process that selects the phenotypes that are best suited to the survival of a **species** in its particular environment. This means that the organisms which are most suited to their environment survive at the expense of

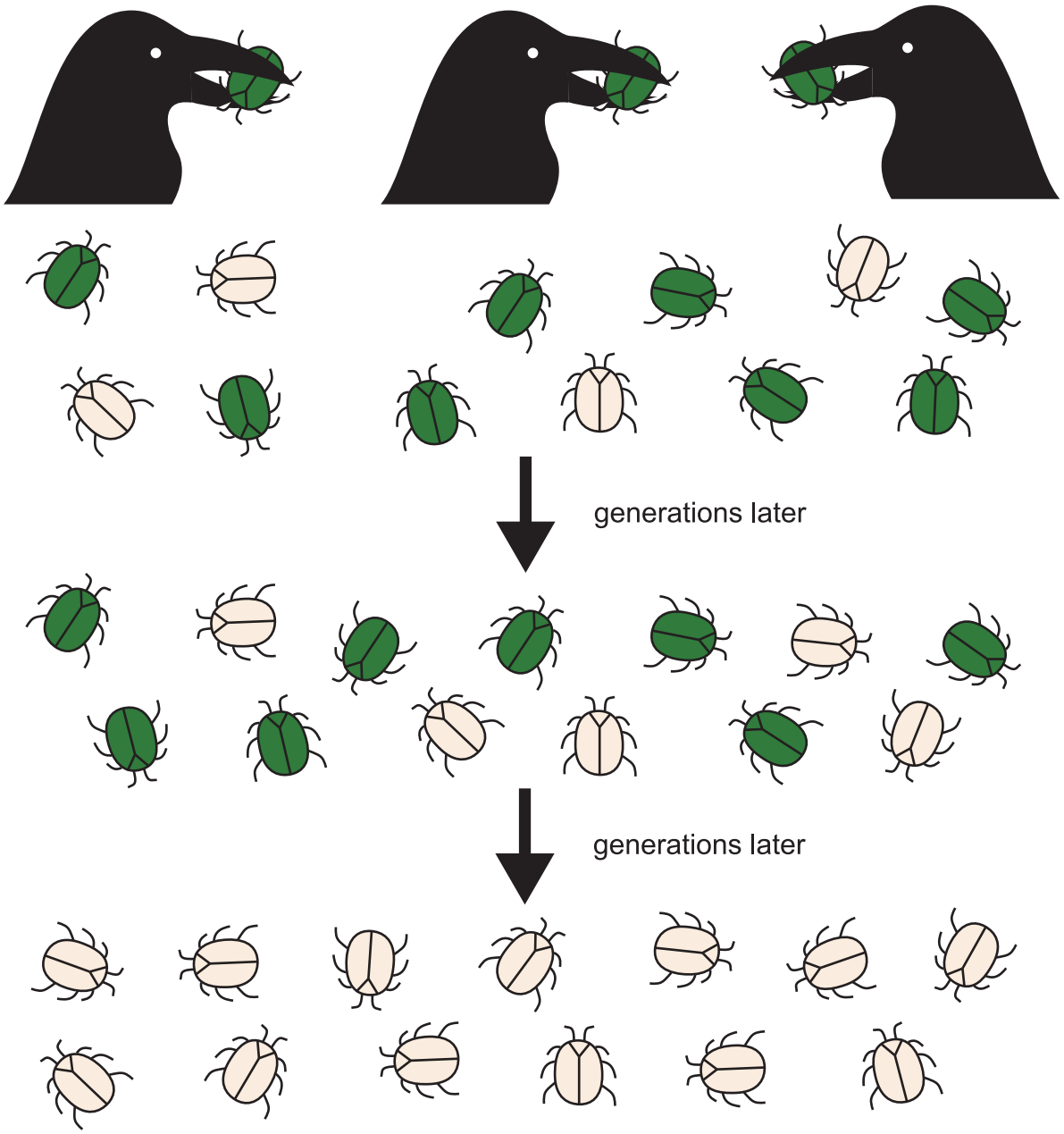
those which are less well adapted. As the environment is changing continuously, natural selection is an ongoing process.

Working individually, Charles Darwin and Alfred Wallace suggested the same theory of evolution which they published in a joint paper in 1858. They proposed that natural selection was the mechanism by which evolution occurred. In his book *On The Origin of the Species*, Charles Darwin provided extensive evidence to support the theory of natural selection.

It is often assumed that evolution occurs over thousands, if not millions, of years, but in some cases it can be readily studied in organisms that have evolved over much shorter periods of time.

The theory of natural selection states that:

- In each generation, more offspring are produced than it is possible for the environment to support. Therefore, each individual in the offspring has to compete and struggle to survive so that it can reproduce and pass on its genes.
- Every individual in a population displays slightly different phenotypes. It is the individuals that possess characteristics that are most useful and better adapted to their environment that are most likely to survive. Less beneficial phenotypes are gradually removed from the general population as individuals displaying these characteristics have a reduced survival rate and a reduced chance of reproducing.
- This process continues over many generations, increasing the numbers of individuals displaying the advantageous characteristics for that environment so that they dominate the population. In this way the phenotypes that are beneficial to the organisms in their particular environment are selected and preserved within the **species**.
- The ability of an individual to reach adulthood and reproduce is described as its fitness. The "fitter" an organism is, the more likely it is to survive and produce offspring that go on to reproduce.

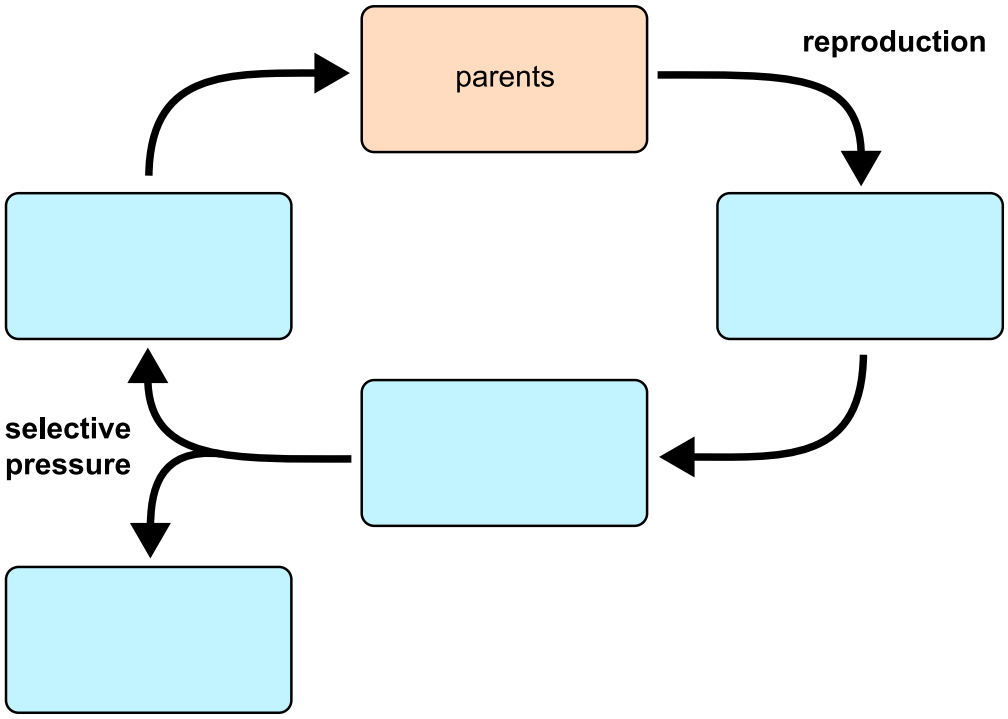


Natural selection in a nutshell: dark bodied beetles are selected against over a period of time whereas light bodied beetles flourish

Natural selection: Questions

Go online

Q1: Place the stages into the correct places to complete the diagram of the key stages that occur during natural selection:



individuals **best** adapted to the habitat **survive**	individuals compete with each other for survival	individuals **least** adapted to the habitat **die**	many offspring showing various phenotypes

Q2: Which of the following statements is **not** true of evolution by natural selection?

a) Natural selection is an ongoing process.
b) The frequency of beneficial genes increases within a population.
c) All individuals in a population display the same phenotypes.
d) High levels of competition exist between individuals.

. . . . . .

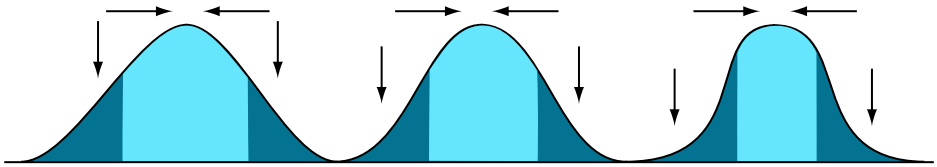
Q3: What does fitness of an individual describe in terms of evolution?

a) The number of times it mates during its lifetime.
b) The reproductive success it experiences during its lifetime.
c) The length of time it lives for.
d) How well it adapts to its environment.

7.3.2 Effects of Selection

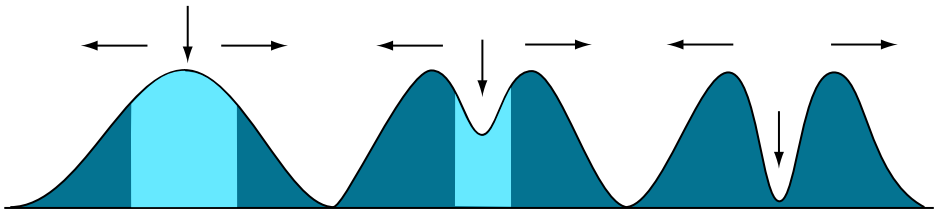
Changes in phenotype frequency can occur as a result of stabilising, directional and disruptive selection.

1. *Stabilising Selection*
   In this case, the average phenotype is selected and the extremes survive much less well, possibly even disappearing. As an example, birds in a particular environment may have a range of colourings from light to dark. If the climate were to change to dull and overcast, then the white and black individuals would stand out and become prey to predators. The result would be an increase in grey birds because their grey (average) colour was selected for.



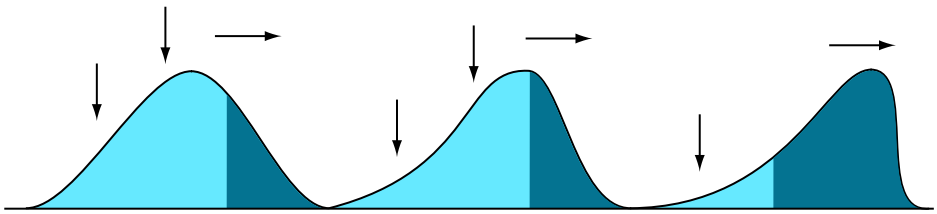
Stabilising selection

2. *Disruptive Selection*
   In this case, it is the extreme values of phenotypes that are chosen and those with average fitness are selected against. As an example, assume the same bird population as before, but now the climate changes and becomes colder with snow persisting in part of the habitat. White birds will be well hidden from predators in the snow and the black birds will blend into the dark background below the snowline. However, the grey individuals will stand out in both conditions and will thus be susceptible to predation. Now it is the extreme values or phenotypes that are selected for.



Disruptive selection

3. *Directional Selection*
   In this final case, one extreme value or phenotype is selected over both the average and the other extreme value. Based on the same bird population again, let us assume that the snow has gone and left a dark earth-scape. Now, both white and grey varieties will stand out and become victims of predation. The dark phenotype is selected for and the numbers of these birds rise as a result.



Directional selection

7.4 Speciation

This section considers the concept of species, the process leading to the formation of new species and its mechanisms, the barriers that can affect how species develop and the concept of hybrids.

7.4.1 Species

A **species** is described as a population of organisms that have the same characteristics and are capable of interbreeding to produce fertile offspring. For example, lions, tigers and jaguars all possess similar traits, such as body shape, facial and paw structure, and their ability to roar, but they are unable to mate with each other to produce fertile offspring, making them three distinct species.

Each member of the species has the same number of chromosomes and the same **gene pool**. As the gene pool is comprised of the sum of all of the different genes of a particular species, it follows that if there are changes to the gene pool, then evolution will occur.

Due to organisms moving and breeding with different populations of the same species, genes are continually moving between populations. When this stops, and populations become isolated, different species may emerge.

Species: Questions Go online

Q4: What is a population?

. . . . . .

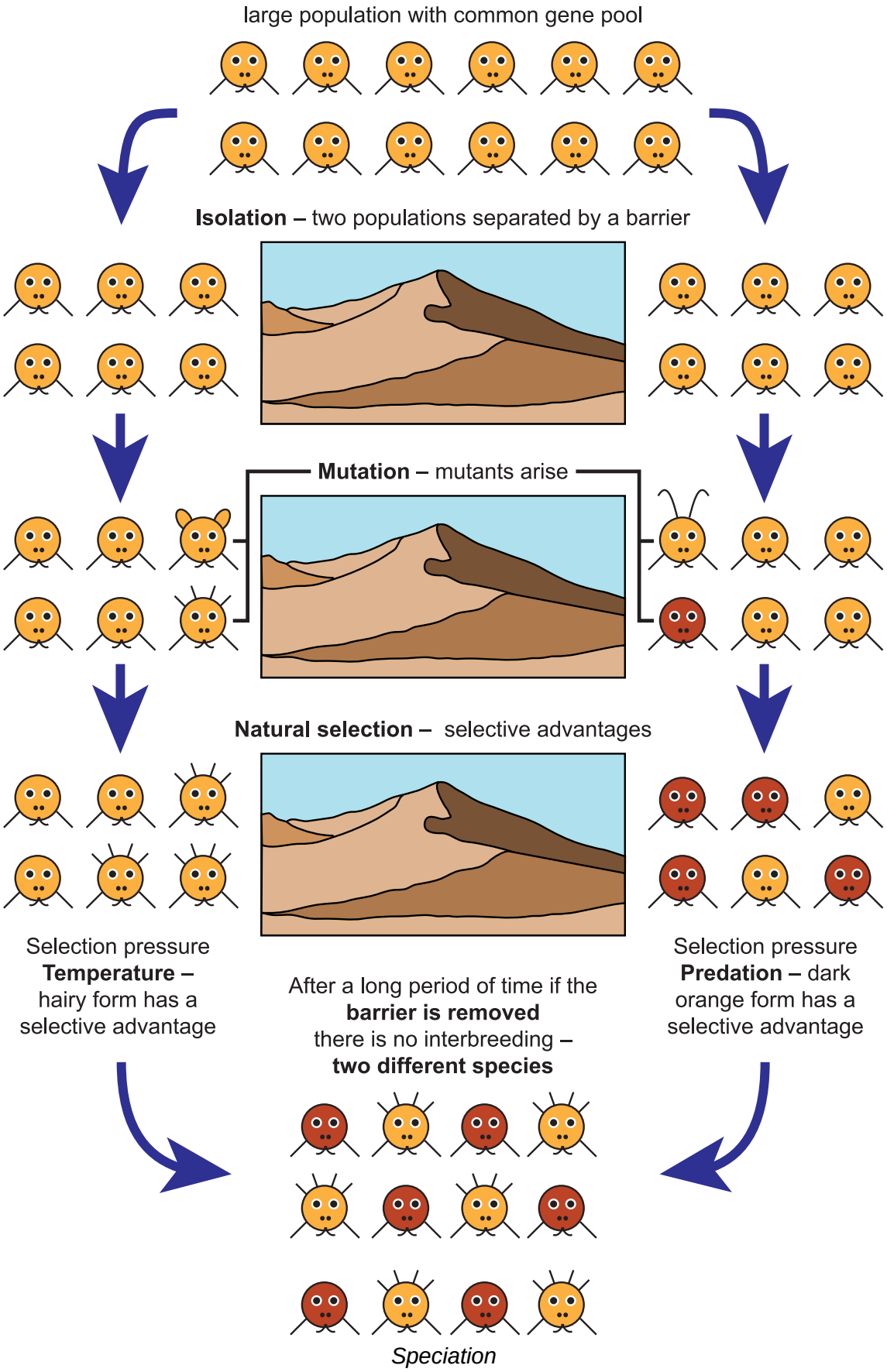
Q5: What is the difference between the gene pool and allele frequency?

7.4.2 Speciation

Speciation is the generation of new biological species by evolution as a result of isolation, mutation and selection. Populations of an existing **species** can become isolated from each other, with the result that their gene pools diverge. The isolated populations experience different selection pressures and adapt to their particular niche, developing different characteristics. Individuals from the different populations will eventually no longer be able to breed with each other. A new species will then have been formed.

When interbreeding populations become separated from each other the flow of **alleles** between them is prevented. This means that the **gene pool** of each sub-population is no longer influenced by the gene pools of other sub-populations of the same species.

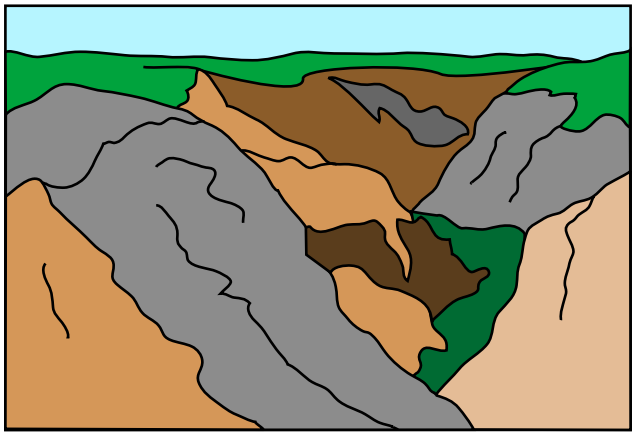
Within a population, any mutations that arise which are beneficial to the population are favoured by **natural selection**. In separated populations, new alleles may be introduced that cause the sub-populations to evolve in a slightly different way, eventually leading to the creation of new species.


large population with common gene pool
Isolation – two populations separated by a barrier
Mutation – mutants arise
Natural selection – selective advantages
Selection pressure
Temperature –
hairy form has a
selective advantage
After a long period of time if the
barrier is removed
there is no interbreeding –
two different species
Selection pressure
Predation – dark
orange form has a
selective advantage
Speciation

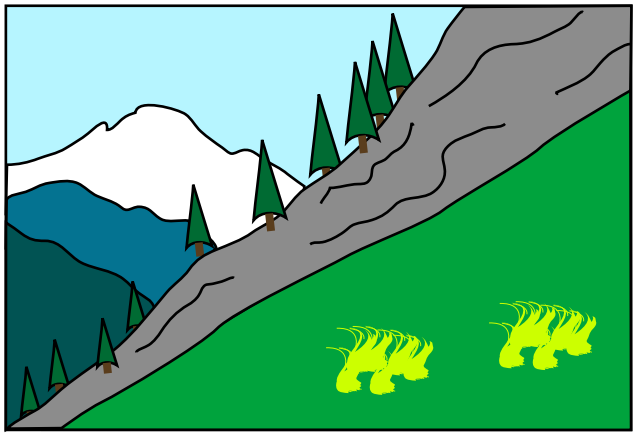
7.4.3 Barriers to species

There are several types of barriers which can bring about speciation:

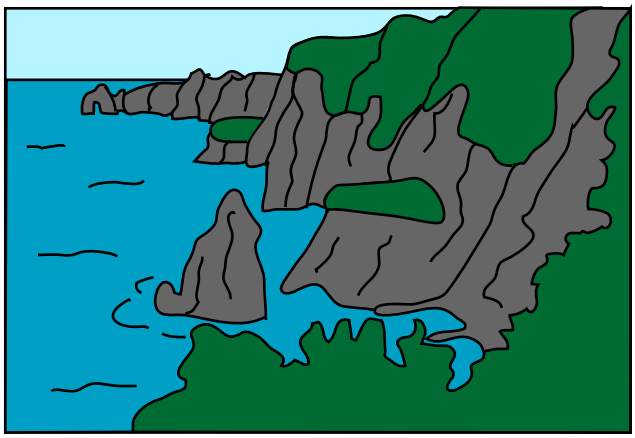
- *geographical barriers*: these include mountains, deserts, oceans and rivers that physically separate organisms and prevent populations from interbreeding. Geographical isolation may also occur if a habitat is lost, such as the destruction of a forest to form an arid landscape or a river drying up.
- *ecological barriers*: factors such as temperature, pH, salinity, humidity and altitude also act to separate populations. For example, many species have evolved to inhabit regions of different pH or salinity.
- *behavioural barriers*: if individuals in a population become fertile at different times of the year, their sexual organs change, or their courtship behaviour is different or unattractive, then the individuals cannot mate.



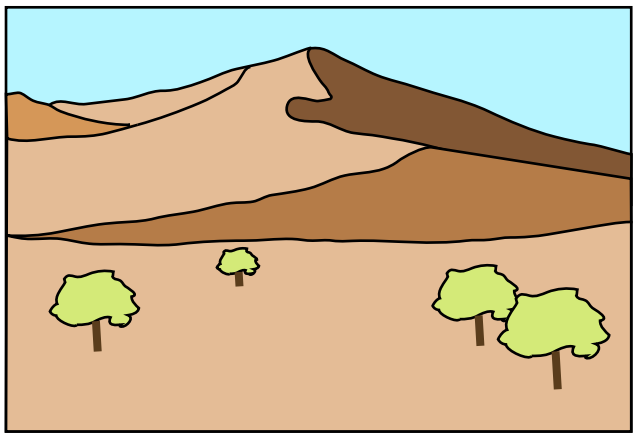
1. This canyon has been isolated by mountains.



2. Different terrains on neighbouring mountains.



3. Water creates a difficult barrier to cross.



4. Deserts are hostile environments.

Examples of geographical barriers to species

7.4.4 Speciation mechanisms

Allopatric speciation occurs due to populations becoming physically separated; it is brought about by geographical barriers. After separation, the now isolated populations may be subjected to different selective pressures or develop and maintain mutations which benefit that group. If, for any reason, the barriers are removed and the populations can freely intermingle but still cannot reproduce and produce fertile offspring, then speciation will have occurred and a new species will have been formed.

Sympatric speciation is a form of speciation where two species arise within the same habitat. For this to happen, other isolating mechanisms must be at work. Sympatric speciation occurs as a result of behavioural or ecological barriers.

Sympatric speciation is much more common in plants compared to animals. If parent plants produce offspring that are **polyploids**, this means that while these plants remain in the same habitat, they are now incompatible for breeding.

In a rare example of animal sympatric speciation, two groups of *Orcinus orca* (killer whale) live in the same habitat in the northeast Pacific ocean. Of these two groups, one is 'resident' and the other is 'transient'. Studies show that they stay apart from each other and do not interbreed; they have different diets, vocal behaviour and social structures. Although the two groups of whales currently belong to the same species, if this situation continues, speciation may occur in the future.

7.5 Learning points

Summary

- Evolution is the changes in organisms over generations as a result of genomic variations.
- Inheritance usually involves the passing of genetic material from parents to offspring. This is vertical inheritance.
- The passing of genetic material is brought about by sexual or asexual reproduction.
- Prokaryotes usually reproduce by binary fission, a form of vertical inheritance.
- Prokaryotes and viruses can also exchange genetic material between members of the same generation. This is horizontal inheritance.
- Horizontal inheritance frequently involves the exchange of whole or parts of a plasmid.
- Plasmid exchange will bring about rapid evolutionary change.
- Natural selection is the non-random increase in frequency of DNA sequences that increase survival and the non-random reduction in the frequency of deleterious sequences.
- Changes in phenotype frequency can be as a result of stabilising, directional and disruptive selection:
  - in stabilising selection, an average phenotype is selected for and extremes of the phenotype range are selected against;
  - in directional selection, one extreme of the phenotype range is selected for;
  - in disruptive selection, two or more phenotypes are selected for.
- A species is a group of organisms capable of interbreeding and producing fertile offspring, and which does not normally breed with other groups.
- Speciation is the generation of new biological species by evolution as a result of isolation, mutation and selection.
- Allopatric speciation occurs when populations become physically separated. This may be due to geographical barriers such as oceans or mountain ranges.
- Sympatric speciation is brought about by behavioural or ecological barriers. The process will usually occur within the same habitat.
- After a barrier separates a population, different mutations may arise within each subpopulation; natural selection will act differently on each group depending on the selection pressures present and, eventually, the two subpopulations become separate species.

7.6 End of topic test

End of Topic 7 test Go online

Q6: Evolution can be described as a change in a _____ over time, and is driven by _____.

. . . . . .

Q7: Inheritance can be described as the passing of _____ between generations.

. . . . . .

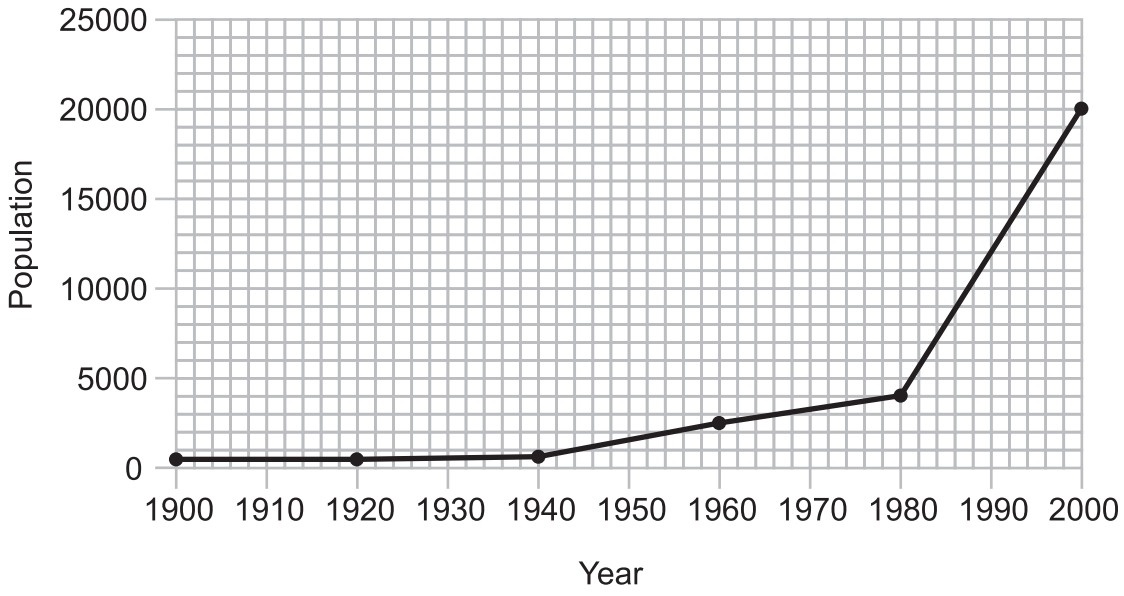
Q8: The exchange of plasmids between bacteria is an example of _____ inheritance.

. . . . . .

Q9: *Population change*

The Galapagos tortoises are very much at risk of extinction. The following table and graph show the rise in the human population in the Galapagos Islands over the last 100 years.

Year	Population size
1900	500
1920	500
1940	600
1960	2500
1980	4000
2000	20000

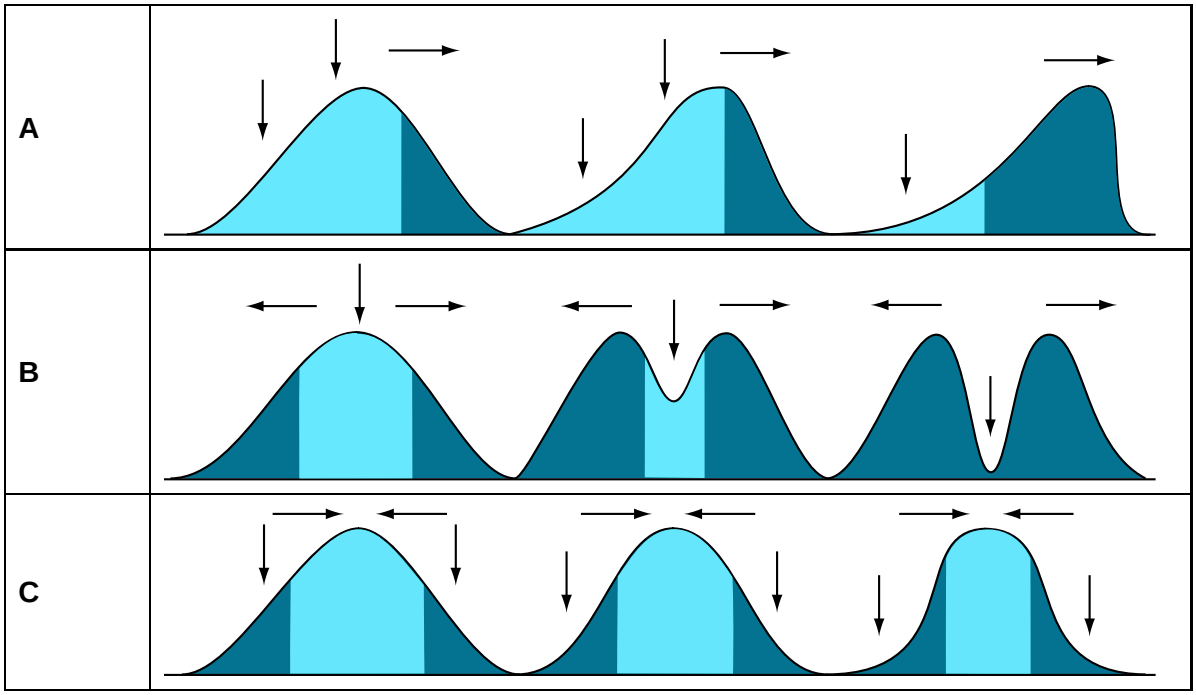


i. State the population size in 1990.

ii. By what percentage did the human population increase between 1980 and 2000?

. . . . . .

Q10: Identify which graph represents stabilising selection, directional selection and disruptive selection.



. . . . . .

Q11: Match each of the following characteristics with one of the three selection types (stabilising selection, directional selection and disruptive selection).

i. Average phenotype selected.
ii. Extreme phenotype selected.
iii. One extreme selected instead of the other one or the average.

. . . . . .

Q12: Speciation is characterised by changes in gene frequency.

True or False?

. . . . . .

Q13: In what way does sympatric speciation differ from allopatric speciation?

Topic 8

Genomics

Contents

Prerequisites

You should already know that:

- the four bases A, T, G and C make up the genetic code;
- a mutation is a random change to genetic material;
- mutations may be neutral, or confer an advantage or a disadvantage.

Learning objective

By the end of this topic, you should be able to:

- understand that the complete nucleotide sequence for either a gene or a complete genome can be sequenced, requiring the use of bioinformatics;
- understand that, by using sequence data, it is possible to trace the evolution of organisms, how they are related to each other, and when they may have diverged from each other;
- understand that, by using sequence data in conjunction with the fossil record, the sequence of evolution can be established;
- understand that many genomes have been sequenced, including the human genome and other important species; comparisons of genomes show that large areas are conserved;
- understand that an individual's genome can be analysed to assess the likelihood of disease and, if necessary, design tailored treatment.

8.1 Genomic sequencing

DNA sequencing is the process of determining the order of nucleotides in a section of DNA. It is now possible to determine the sequence of nucleotides in relatively small sections, i.e. a gene, or very large sections, i.e. a complete genome.

The process by which sequencing can be achieved involves several techniques, and it is only relatively recently that two fields, in particular, have become sufficiently advanced to be incorporated together. PCR and related techniques now produce nucleic acids in sufficient quantity and purity, and computer science and statistical analysis can now cope with the quantity of evidence to rationalise the data into a comprehensible form. The use of computers and statistical analysis is known as **bioinformatics**.

Sequencing techniques

The development of sequencing techniques began sometime after the development and refinement of the Crick and Watson model of DNA.

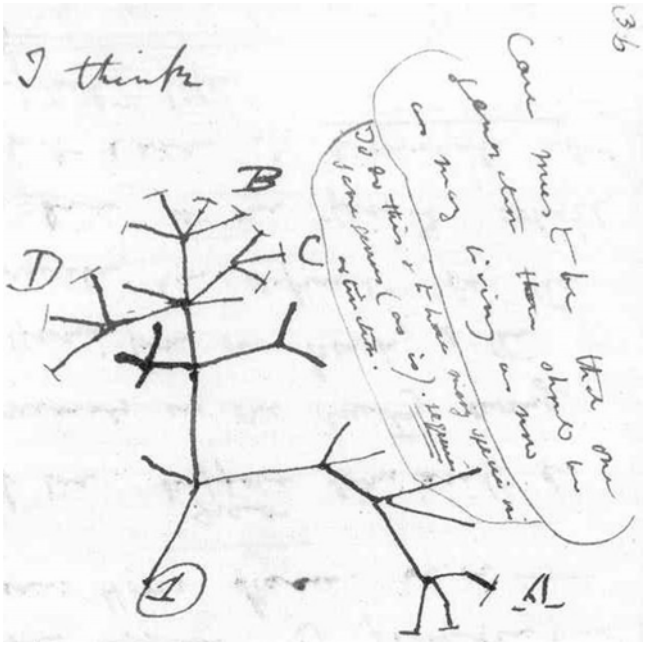
By the early 1970s, recombinant DNA technology was becoming established. From this, defined fragments of DNA could be generated, as opposed to samples from bacteriophage or viruses. By 1977, the first complete DNA **genome** had been established. At the time, two teams were establishing techniques of sequencing.

Early attempts at sequencing were developed by Sanger and Coulson in 1975, called the plus-minus method. This was overtaken by Maxam and Gilbert in 1977, by a method based on chemical modification of DNA followed by cleavage at specific bases and, while accurate, it proved difficult to scale up and involved extensive use of toxic chemicals.

At about the same time, Sanger and his team developed the so-called "Chain Termination" method; this, and its subsequent developments, has become the technique favoured by many. Originally, radioactive materials were used which posed a hazard, but they were replaced first with UV detection and then fluorescent dyes.

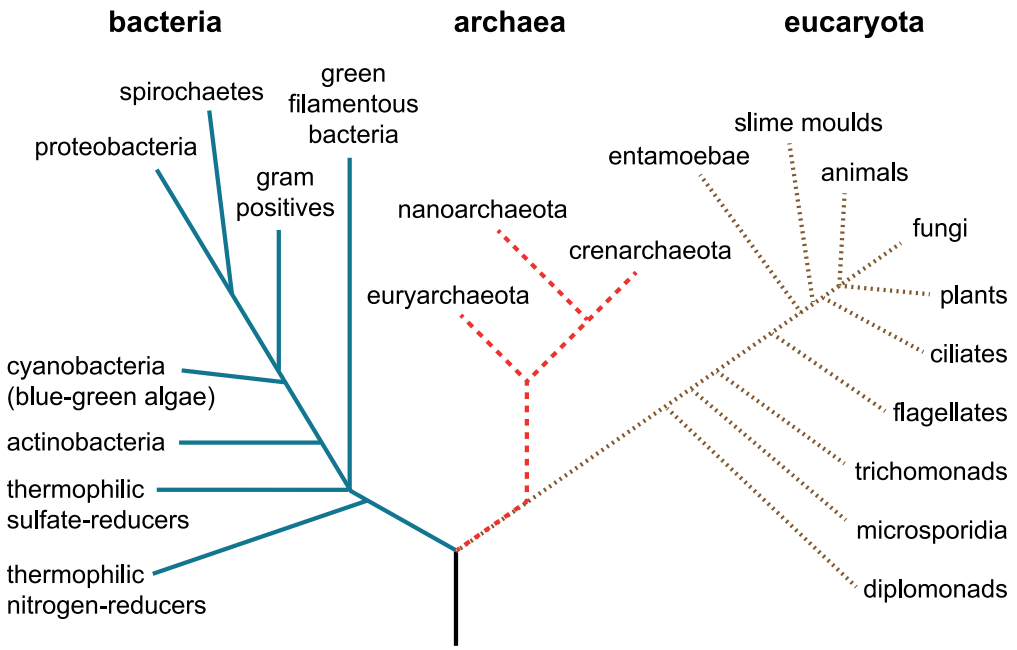
8.2 Phylogenetics

Phylogenetics has been described as the field of biology that deals with identifying and understanding the relationship between the different kinds of life on Earth, others use the term 'evolutionary relatedness'. It has become an essential tool in tracing the evolutionary tree of life as first proposed by Darwin.



A diagram and note written by Charles Darwin

Originally, the relationship between organisms was traced by comparison of physical characteristics, embryology and examination of fossil records. Some of this is still in use but, more recently, advances in DNA sequencing have become much more prevalent. These days, sequence data and fossil evidence are both used to determine the main sequence of events in the evolution of life. Based on this relatively modern approach, life is now classified based on the three domains using a system first proposed by Carl Woese in 1977.



Three domains in the tree of life

It has been over 150 years since Darwin first put forward his idea that all life is related, and it is now widely accepted that all life originated in the sea some 3000 million years ago. The first cells developed and replicated themselves in two ways. Some formed into chains, which gave rise to algae and plant life, while others developed into hollow balls, which became sponges and the origins of animal life.

These groups grew and differentiated until about 450 million years ago with the development of land animals. As time progressed, greater divergence arose. Several mass extinction events eliminated many types of organism, but also gave those surviving the opportunity to flourish.

Over long periods of time (millions of years), mutations built up at a steady rate in any section of DNA. If the rate could be shown to be reliable, it could be used as a mechanism to estimate the time between mutations, i.e. a clock. This, in turn, can be used to estimate both where and when organisms diverged.

Molecular clocks are used to show when species diverged during evolution. They assume a constant mutation rate and show differences in DNA sequences or amino acid sequences. Therefore, differences in sequence data between species indicate the time of divergence from a common ancestor.

However, rates of molecular evolution can vary between organisms and so molecular clocks have to be calibrated. To do this, it is necessary to know the absolute age of some evolutionary divergence which can usually be determined from the fossil record.

Phylogenetics: Question Go online

Q1: Complete the paragraph using the words from the list.

All life forms are now described as belonging to one of three _____. This is largely based on a comparison of _____. The three main groups are _____, _____ and _____. Phylogenetic clocks need to be calibrated by using _____.

Word list: archaea, bacteria, DNA, domains, eukaryota, fossil records.

8.3 Comparative genomics

The **genome** is the sum total of all the hereditary material within an organism. **Genomics** is the science of interpreting genes: the study of an organism's genome using information systems, databases and computerised research tools.

Many genomes have been sequenced, particularly those of disease-causing organisms, pest species and species that are important model organisms for research. The following table shows the size of genomes in some organisms.

Organism	Estimated size (base pairs)	Chromosome number	Estimated gene number
Human (*Homo sapiens*)	3 billion	46	ca. 25,000
Mouse (*Mus musculus*)	2.9 billion	40	ca. 25,000
Fruit fly (*Drosophila melanogaster*)	165 million	8	13,000
Plant (*Arabidopsis thalania*)	157 million	10	25,000
Roundworm (*Caenorhabditis elegans*)	97 million	12	19,000
Yeast (*Saccharomyces cerevisiae*)	12 million	32	6,000
Bacteria (*Escherichia coli*)	4.6 million	1	3,200

Comparison of some genomes

Comparative genomics is the process whereby the genomes of different species are compared. When comparing genomes from different species, scientists noted that many parts of the genome are highly conserved. This means that some sections of DNA are identical (or almost identical) between different species. These conserved regions of DNA are useful in determining evolutionary relationships.

One of the earlier organisms to be sequenced was the puffer fish. The attractions of the puffer fish, or fugu, are that it has one of the most compact genomes of all vertebrates. Roughly speaking, it contains a similar number of genes to humans, but they are contained in only 400 megabases compared to the 3.1 gigabases of the human genome.

By comparing the fugu genome to the human genome, it is possible to establish common functional elements of genes and regulatory sequences. By contrast, the non-functioning genes show where evolutionary divergence has occurred from a common ancestor, approximately 450 million years ago.

8.4 Personal genomics

One of the aims of **genomics** is to explain why some individuals are susceptible to disease while others seem unaffected. By understanding the interaction between genes and the environment, it may be possible to prevent the onset of some of these complex diseases in individuals.

Personal genomics is the sequencing and analysis of an individual's **genome**. Once an individual genotype (or part of it) is known it is compared to references in the published literature. From this, any mutations or sequences likely to give rise to disease can be identified. This is now referred to as predictive medicine, which in turn can lead to the use of an appropriate drug treatment if required, a process know as pharmacogenetics.

Key to personal genomics has been cost, which has been falling rapidly. When the first genome was sequenced, it cost in the region of three billion dollars; this genome was a composite of several individuals. Currently (May 2018), individuals can have their own genome (or sections of it) sequenced for £700 and completed within days or hours. However, after sequencing there must follow an analysis which may incur further costs and time.

As a result of advances in this field, a question of ethics has also arisen. Insurance companies, banks and others may decline services or increase premiums as a result of finding less desirable traits, e.g. Alzheimer's or other degenerative diseases. This has been termed genetic discrimination. As yet, regulations in this and associated fields are not clearly laid out.

Personal genomics could bring about greater understanding of the varying effects of drugs between different individuals. One example is the group of genes responsible for drug (and other metabolites) metabolism - cytochrome P450 (CYP) genes. Depending on which alleles have been inherited, an individual may be described as an extensive metaboliser (that is to say, normal), and can successfully metabolise certain compounds. Others, however, may be found to be intermediate or poor. As a consequence, patients may find that the drug of choice may be ineffective or cause severe adverse reactions. Clearly, it would be of significant advantage to have this information prior to treatment. It could save the patient from potential danger, and save the often considerable cost of medication.

8.5 Learning points

Summary

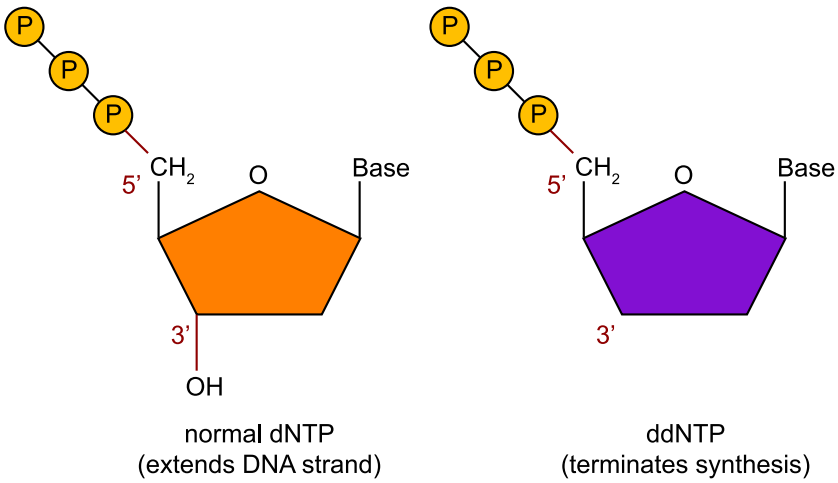
- It is possible to determine the sequence of bases in any section of DNA. It could be a relatively short sequence, for example a gene, up to and including a whole genome.
- To elucidate a DNA sequence, several different scientific disciplines are required.
  - First, DNA needs to be sequenced, which could be achieved by a variety of methods.
  - Second, the data collected must be analysed; this is achieved through mathematical, statistical and computing sciences (bioinformatics).
- From the DNA sequencing of different species' genomes, phylogenetic trees are formed. These show the relationship between organisms and their common ancestors.
- These comparisons are made by comparing DNA sequences and looking for mutations within comparable regions of a genome.
- Knowing the rate of mutation, it is possible to pinpoint the time at which divergence occurred.
- Mutation rates can be calibrated by cross reference to the fossil record.
- It is on this basis that classification of organisms is now based on the Three Domain system.
- Based on sequencing and fossil records, the path of evolution can be supported from the emergence of cells through prokaryotes, eukaryotes to higher plants and animals.
- The genomes of many organisms have been completed, and many more are being undertaken.
- Other than the human genome, organisms that cause disease in man, domestic animals and food crops have been sequenced. In addition, some organisms have been sequenced as they have proved to be most useful for comparison and modelling during research.
- It has become apparent when comparing genomes that sections are highly conserved, or similar.
- By studying an individual's genome, it is possible to determine errors or gather evidence to support the likelihood of ailments. Some examples include the predisposition to cancer, mental illness or drug dependency.
- It may be possible, armed with this knowledge, to treat or alleviate symptoms. Personalised medical care of this nature is called pharmocogenetics.

8.6 Extension materials

The material in this section is not examinable. It includes information which will widen your appreciation of this section of work.

Extension materials: Chain termination DNA sequencing method

The key principal of the chain termination DNA sequencing method is the use of dideoxynucleotide triphosphates (ddNTPs) as chain terminators.



Dideoxynucleotide triphosphates or ddNTPs

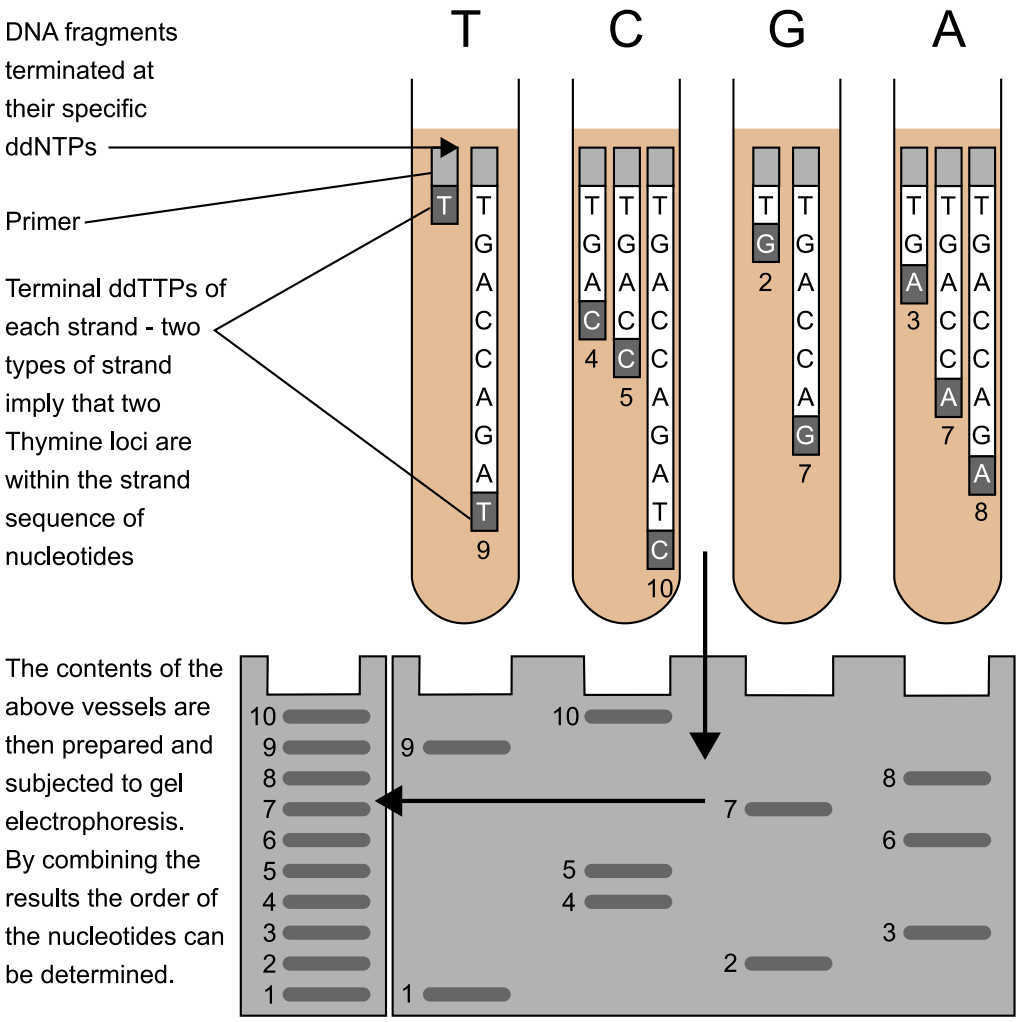
A single strand of DNA is required. This is called the DNA primer. Also needed are DNA polymerase, normal deoxynuclotide triphophates (dNTPs) and modified dideoxynucleotides (ddNTPs). These ddNTPs terminate DNA elongation and may be either radioactively tagged or fluorescently labelled. Labelling in this way allows for detection in automated machines.

The DNA sample is divided into four sequencing reaction vessels, containing normal deoxynucleotides and DNA polymerase. Then, each of the four separate vessels has one of the modified ddNTPs added to it. These ddNTPs stop chain elongation as they are lacking a 3′-OH group and so cannot form a phosphodiester bond to the next nucleotide. As a result, fragments of DNA of different lengths form.

The fragments are denatured and separated by gel electrophoresis. Each reaction is run in a separate lane. The bands of DNA are visualised depending on the marker chosen.

When using autoradiography, X-ray film may be used. If using fluorescent dye, a laser detector is used. Dark bands or specific colours dictate the position of the ddNTP at the end of a DNA strand. As each band appears, it indicates the sequence of nucleotides in the chain.

The use of fluorescent labelled ddNTPs and primers has allowed automated systems to be set up and has greatly increased the speed at which DNA sequencing can be carried out.



Chain termination DNA sequencing method

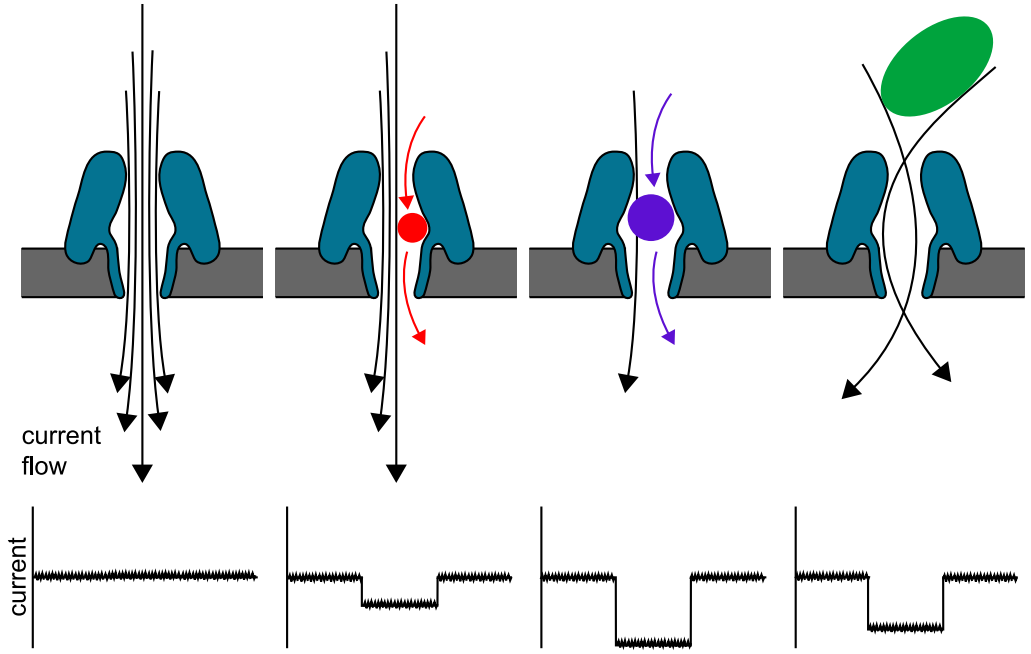
Following the successful completion of the Human Genome Project in 2003, one of the next goals was to develop the $1000 **genome**. That is, an individual's full genome to be described for less than $1000 in a matter of days.

The method has a few drawbacks, typically that some bases at either end of the sequence may be misread, but software can estimate these gaps.

Despite refinement and automation, faster methods have been sought. This is called high-throughput sequencing, or next generation sequencing, and utilises sequencing which runs many parallel processes producing vast numbers of sequences at a time.

Many of these techniques use emulsion PCR and nanopore sensing.

A technology that was announced in February 2012 suggests a significant advance towards the $1000 genome. It is based on nanopore technology and is designed to be capable of reading a single strand of DNA directly. As the DNA passes through the nanopore, it excites an electric current that is particular for each nucleotide. The advantage of this system is that it can measure single molecules directly without the need for amplification, labelling or optical reading instruments.



Nanopore sensing

The above diagram shows a protein nanopore set in an electrically resistant membrane bilayer. An ionic current is passed through the nanopore by setting a voltage across this membrane.

What are the benefits of using nanopores to sequence DNA?

In contrast to current sequencing technologies, nanopores can measure single molecules directly, without the need for nucleic acid amplification, fluorescent/chemical labelling or optical instrumentation.

It is claimed that using a series of these 'chips', or nodes, that a human genome could be sequenced in 20 minutes for approximately $1500 at today's prices. Smaller versions with a more limited capacity can be inserted into the USB of a laptop at a cost of $900.

8.7 End of topic test

End of Topic 8 test Go online

Q2: Determining the order of nucleotide bases is known as _____.

. . . . . .

Q3: What two techniques are combined to perform bioinformatics?

. . . . . .

Q4: As a result of sequencing, phylogenetic trees can be formed. What is their purpose?

. . . . . .

Q5: As well as sequence data, what other information is required to determine the main sequence of events in the evolution of life?

. . . . . .

Q6: The _____ is the sum total of an organism's DNA.

. . . . . .

Q7: Does the quality of DNA in an organism tell you the number of genes?

. . . . . .

Q8: Many genes are found to be highly conserved across species. What does this mean?

. . . . . .

Q9: What advantages might be gained from knowing an individual's genome?

. . . . . .

Q10: What name is given to the field of medicine that aims to use knowledge of patients' genomes to design appropriate courses of medicines?

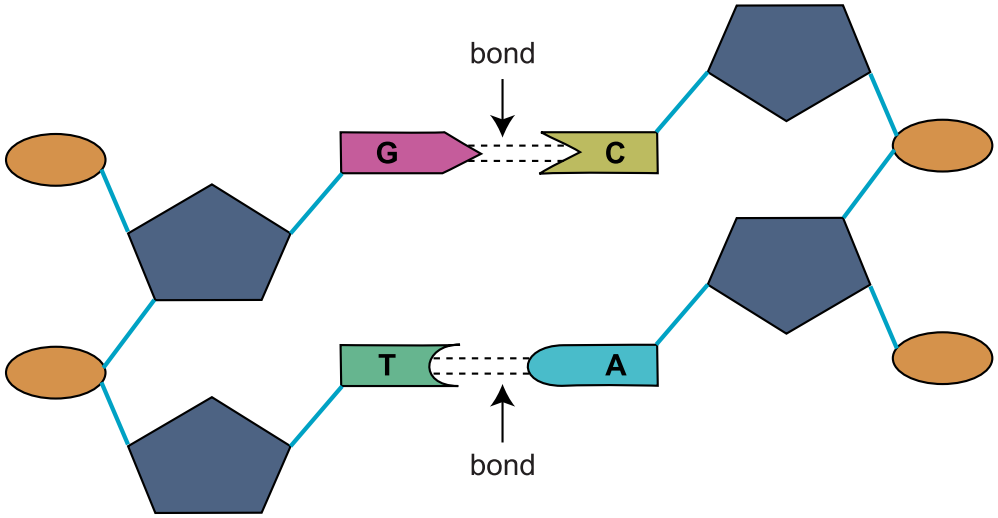
Topic 9

End of unit test

End of Unit 1 test Go online

Structure of DNA

The diagram below represents the four different nucleotides found in DNA.



Q1: Name the type of bond which links the base on one DNA strand to the base on the opposite strand.

. . . . . .

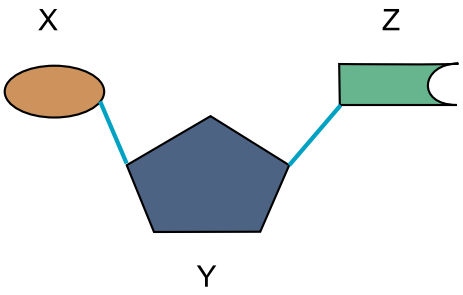
Q2: A section of a DNA molecule contains 8000 bases. Of these, 2800 are thymine. What is the percentage of cytosine bases in the molecule?

. . . . . .

Q3: What is the name of the sugar present in DNA?

. . . . . .

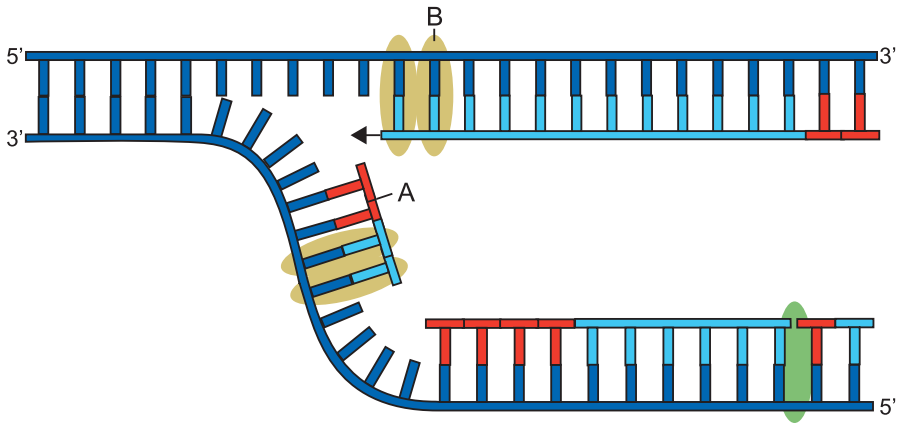
Q4: The diagram below shows a DNA nucleotide.



Identify the parts labelled X, Y and Z in the diagram.

Replication of DNA

The following diagram represents DNA replication.



Q5: Name structure A.

. . . . . .

Q6: Name one essential requirement for DNA replication that is not shown above.

. . . . . .

Q7: Name enzyme B.

. . . . . .

Q8: The polymerase chain (PCR) can be used to obtain many copies of a particular gene. Complete the following sentence.
The _____ bonds of DNA are separated by _____ during PCR.

. . . . . .

Q9: Primers are required for the process of PCR. Which of the following statements is false?

a) Primers mark the start and end of sequence to be copied.
b) Primers prevents the strands re-joining.
c) Primers unwind double helix.
d) Primers allow DNA polymerase to attach.

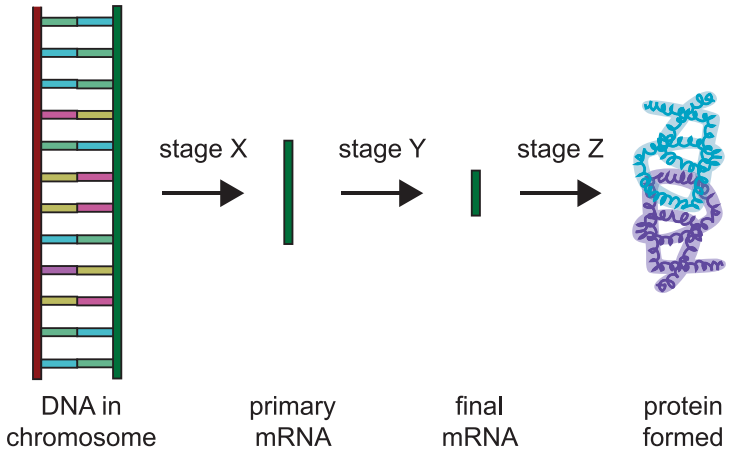
. . . . . .

Q10: Starting with a single molecule of DNA, the polymerase chain reaction goes through three complete cycles. How many molecules of DNA would be produced?

a) 4
b) 8
c) 16
d) 32

Gene expression

The following diagram shows the stages of protein synthesis.



Q11: Name stage X.

. . . . . .

Q12: Name the exact location in the cell where stage X occurs.

. . . . . .

Q13: Name stage Z.

. . . . . .

Q14: Name the exact location in the cell where stage Z occurs.

. . . . . .

Q15: Which enzyme catalyses stage X of this process?

. . . . . .

Q16: Explain why a primary mRNA molecule is longer than a mature mRNA molecule.

. . . . . .

Q17: The mRNA codon for the amino acid methionine is AUG. What is the base sequence of the corresponding anti-codon?

. . . . . .

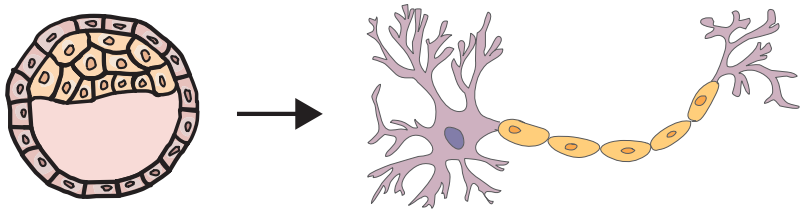
Q18: What name is given to the chemical bonds that join amino acids together in a protein?

. . . . . .

Q19: Name the bonds which hold polypeptide chains together in a protein.

Cellular differentiation

There are hundreds of cell types in the human body which originate from stem cells in the early embryo.



Q20: What are stem cells?

. . . . . .

Q21: A well known drug company has developed a drug for the treatment of the symptoms of a particular inherited disease. However, before they can proceed to clinical trials using volunteers, the company has decided to do additional tests in the laboratory using stem cells. Suggest an ethical consideration which may influence their decision in using stem cells.

. . . . . .

Q22: What is the difference between embryonic and tissue stem cells?

. . . . . .

Q23: Meristems are regions of _____ cells in plants that are capable of cell _____.

Structure of the genome

Q24: What is a gene?

. . . . . .

Q25: Name a non-translated form of RNA.

Mutations

Q26: Listed below are three types of mutation.

a) Deletion
b) Substitution
c) Insertion

Which mutation(s) affect only one amino acid in the protein produced?

. . . . . .

Q27: What word is used to describe a mutation which results in the addition of an incorrect amino acid into a protein?

. . . . . .

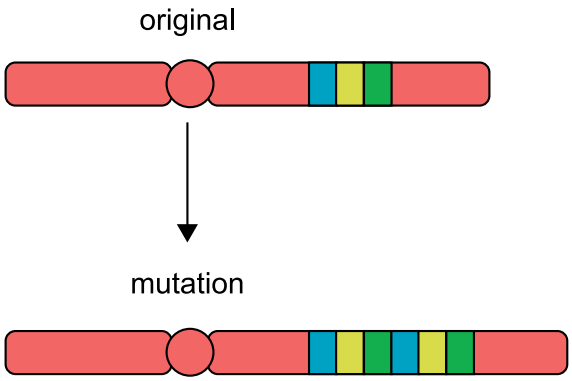
Q28: The sequences below show a type of single gene mutation.

Original sequence: AATTGCTATG
Mutated sequence: AATGGCTATG

Name the type of mutation shown by the sequences.

. . . . . .

Q29: The following diagram shows a type of chromosome mutation.



What type of mutation is shown in the diagram?

1. Deletion
2. Duplication
3. Inversion
4. Translocation

Evolution

Q30: Bacteria can on occasions pass genetic material between themselves. This may be a section of DNA called a _____.

. . . . . .

Q31: This is representative of _____ genetic transfer.

. . . . . .

Q32: Which of the following organisms are capable of vertical gene transfer?

a) Eukaryotes
b) Prokaryotes
c) Prokaryotes and eukaryotes

. . . . . .

Q33: Complete the following sentences using the words from the list.

- _____ selection is where the average phenotype is most successful for a particular habitat.
- _____ selection is characterised by the extreme versions of a phenotype being selected.
- _____ selection is characterised by the selection of one extreme phenotype at the exclusion of all others.

Word list: Directional, Disruptive, Stabilising.

. . . . . .

Q34: What is a species?

. . . . . .

Q35: What type of barrier is involved in allopatric speciation?

. . . . . .

Q36: Many species of cichlid fish are found in Lake Milawi. They are thought to have come from a common ancestor, but cannot interbreed as they fail to recognise each other's courtship patterns. Of the three species, **A** filter feeds on microbes in the water, **B** scrapes algae from rock surfaces and **C** crushes snail shells to extract the meat.

From the information above, why would **A**, **B** and **C** be considered different species?

. . . . . .

Q37: From the information above, what is this type of speciation described as?

Genomics

Q38: Life can be divided into three domains. What process is used to produce evidence to support this claim?

. . . . . .

Q39: Primates, including orangutans, gorilla, chimpanzee and humans, have a common ancestor from 35 million years ago. Humans and chimps diverged from each other 5 million years ago, gorillas 5 million years before that and orangutans 9 million years before that. Complete the sentences using the words from the list. Some of the words may be used more than once each.

Chimps are _____ closely related to gorillas than orangutans.
The common ancestor of chimps and gorillas is _____ recent than the common ancestor of gorillas and orangutans.

Words: less, more.

. . . . . .

Q40: What process relies on the use of computer and statistical analyses to compare genome sequence data?

. . . . . .

Q41: Comparison of genomes reveals that many genes are highly _____ across different organisms.

. . . . . .

Q42: Current research is investigating how people's genetics affects their responses to drugs. What is this field of medicine known as?

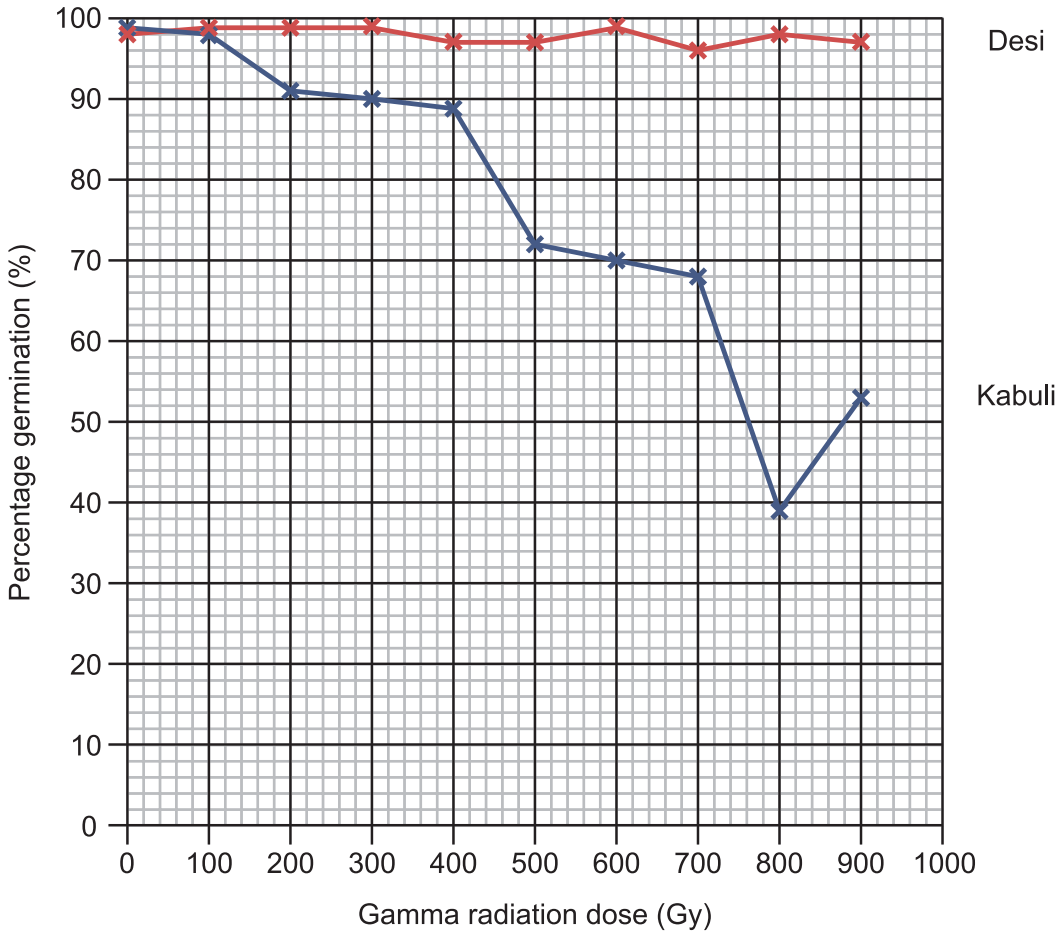
Problem solving

An investigation was carried out to determine the effect of gamma radiation on germination in chickpeas.

Two varieties of chickpeas were exposed to gamma irradiation doses of 0 to 900 Gy (at 100 Gy intervals). Thirty seeds from each group were placed in a petri dish with moist filter paper, and left in an incubator at 25ºC for 7 days. The experiment was repeated three times.

The average results from all three experiments are shown below.

	Percentage germination (%)	
Gamma radiation dose (Gy)	*Kabuli chickpea*	*Desi chickpea*
0	99	98
100	98	99
200	91	99
300	90	99
400	89	97
500	72	97
600	70	99
700	68	96
800	39	98
900	53	97



Q43: What term describes the group of seeds which were not exposed to gamma radiation (0 Gy)?

. . . . . .

Q44: Why were the group of seeds not exposed to gamma radiation included in the study?

. . . . . .

Q45: Name one factor, not already mentioned, which must be kept constant when setting up this experiment.

. . . . . .

Q46: Describe the effect of increasing gamma radiation dose on percentage germination in kabuli chickpeas.

. . . . . .

Q47: Predict the percentage germination of kabuli chickpeas if they were exposed to gamma radiation of 1000 Gy.

. . . . . .

Q48: Draw one conclusion from the results of the experiment.

Glossary

Allele

one of the different forms of a gene

Allele frequency

the prevalence of alternative versions of genes

Anticodon

a triplet of exposed bases on a tRNA molecule

Antiparallel

running in an opposite direction

Apical meristems

growing points (regions of mitosis) found at the tips of plant stems or roots allowing increase in length

Bioinformatics

a process which combines computer science and statistical analysis to study genomes

Blastocyst

an embryo that has developed for 5 to 6 days after fertilisation

Chloroplast

the photosynthetic unit of a plant cell, containing all the chlorophyll

Codon

a triplet of exposed bases on a length of mRNA

Deletion

removal of a length of DNA from a chromosome

Deletion mutation

loss of a section of DNA or a number of nucleotides

Differentiation

the process by which cells or tissues undergo a change towards a more specialised function

DNA ligase

an enzyme that facilitates the process by which fragments of DNA are joined together

DNA polymerase

an enzyme that synthesises DNA strands from individual nucleotides

Double helix

the double helical shape of a DNA molecule

Duplication

repetition of a series of nucleotides within a chromosome

Eukaryote

an organism which possesses a membrane-bound nucleus

Exons

the parts of the initial mRNA which are used to code for proteins

Gene pool

complete set of unique alleles in a species or population

Genetics

the branch of biology that deals with heredity, especially the mechanisms of hereditary transmission and the variation of inherited traits among similar or related organisms

Genome

the entirety of an organism's hereditary information

Genomics

the science of interpreting genes; the study of an organism's genome using information systems, databases and computerised research tools

Induced pluripotent stem cells

somatic (adult) cells reprogrammed to enter an embryonic stem cell-like state

Insertion mutation

the addition of an extra nucleotide

Introns

the parts of the initial mRNA which are removed before translation

Inversion

the inversion (reversal) of a section of DNA within a chromosome

Lagging strand

the strand of DNA that grows in the direction opposite to the movement of the growing fork; it is replicated in fragments

Leading strand

the strand of DNA that is being replicated continuously

Meristem

a growing point in a plant, i.e. a place where mitosis produces new cells

Messenger RNA

(mRNA) is synthesised from a DNA template, resulting in the transfer of genetic information from the DNA molecule to the messenger RNA

Mitochondrion

a structure in the cell responsible for producing energy

Mitosis

nuclear division

Natural selection

the survival of the fittest, whereby only individuals with the most suitable genetic constitution for any set of circumstances pass their genes on

Plasmid

a circular, self-replicating DNA molecule that carries only a few genes

Plastome

the genetic material that is found in plastids in plant cells (for example in the chloroplast). It composes part of the entire genome of photosynthetic organisms

Pluripotent stem cells

these are stem cells, with the potential to make any differentiated cell in the body

Polyploid

organisms which contain more than two set of chromsomes

Primer

a strand of nucleic acid that serves as a starting point for DNA

Prokaryote

an organism which lacks a membrane-bound nucleus

Ribosomal RNA

(rRNA) is the RNA that is a permanent structural part of a ribosome

Ribosomes

structures found in the cytoplasm where protein synthesis occurs

RNA splicing

a process which removes introns from a primary mRNA transcript

Self-renewal

a property displayed by stem cells which allows them to divide to produce more stem cells

Speciation

the formation of a new species

Species

group of organisms which can interbreed to produce fertile, viable offspring

Substitution mutation

the replacement of one nucleotide by another

Transcription

the production of mRNA from a DNA template

Transfer RNA

(tRNA) is a short strand of RNA that is twisted on itself to expose three bases, and which carries a specific amino acid to a ribosome

Translation

the sequencing of amino acids at ribosomes, based on the sequence of nucleotides in mRNA

Translocation

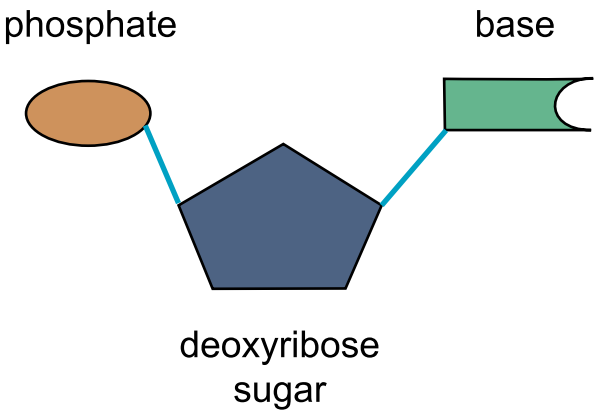
transposition of a length of DNA onto another chromosome

Answers to questions and activities

Topic 1: Structure and organisation of DNA

The structure of DNA: Questions (page 6)

Q1:



Q2:

- The nucleotide guanine pairs with **cytosine**.
- The nucleotide thymine pairs with **adenine**.
- The nucleotide cytosine pairs with **guanine**.
- The nucleotide adenine pairs with **thymine**.

Q3: Double helix

Q4: Hydrogen

Q5: Phosphate, deoxyribose sugar, base

The organisation of DNA in prokaryotes and eukaryotes: Questions (page 8)

Q6: Prokaryotes

Q7: Mitochondria / Chloroplasts

Q8: Plasmid

End of Topic 1 test (page 11)

Q9: Genome

Q10: TACCTGAAATCCA

Q11: 20%

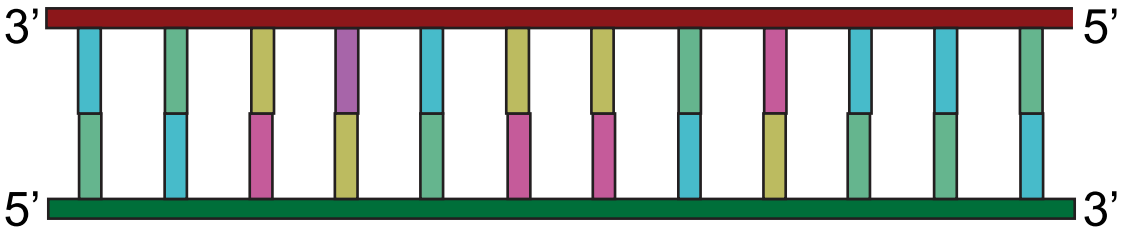
Q12: 4. C, A, B

Q13: Deoxyribose

Q14: Hydrogen bonds

Q15: Sugar-phosphate

Q16:



Q17: Mitochondria / Chloroplasts

Q18: Protein / Histones

Topic 2: Replication of DNA

Extended response question: DNA structure and replication (page 21)

Suggested marking scheme

Each line represents a point worth one mark. The concept may be expressed in other words. Words which are bracketed are not essential. Alternative answers are separated by a solidus (/); if both such answers are given, only a single mark is allocated. In checking the answer, the number of the point being allocated a mark should be written on the answer paper. A maximum of eight marks can be gained.

DNA structure *(maximum of 4 marks)*:

1. DNA can be linear or circular.
2. DNA is made up of two strands of nucleotides.
3. A nucleotide consists of a deoxyribose sugar, a phosphate group and base.
4. Nucleotides are linked together by strong chemical bonds between the deoxyribose sugar of one nucleotide and the phosphate group of another.
5. The bases of DNA always pair up: adenine with thymine, and guanine with cytosine.
6. There are hydrogen bonds between bases.
7. DNA takes the shape of a double helix. . .
8. . . . with antiparallel strands / deoxyribose and phosphate at 3′ and 5′ end of each strand.

DNA replication *(maximum of 4 marks)*:

i. During replication, the hydrogen bonds between the bases in the DNA molecule break and the strands unwind.

ii. A primer binds to each strand of DNA.

iii. DNA polymerase replicates a strand of DNA from free DNA nucleotides.

iv. DNA polymerase adds complementary nucleotides to the 3′ end (of the lead chain) OR in one direction.

v. One strand is replicated continuously.

vi. The other (lag) strand is replicated in fragments.

vii. Fragments are joined by ligase.

End of Topic 2 test (page 23)

Q1: DNA polymerase

Q2: Primer

Q3: DNA ligase

Q4: 2. b, a, d, c - as per:

- The hydrogen bonds between DNA strands break.
- Base pairing occurs between free nucleotides and each of the DNA strands.
- Nucleotides are bonded together by DNA polymerase.
- The DNA molecules coil up to form double helices.

Q5: a) The enzyme is relatively stable at high temperatures.

Q6: 3. d, c, b, f, a, e - as per:

- Temperature of the reaction adjusted to 92-98°C.
- The DNA strands separate.
- Temperature of the reaction adjusted to 50-65°C.
- Annealing of the primers to the single-stranded DNA.
- Temperature of the reaction adjusted to 70-80°C.
- Synthesis of DNA by the enzyme DNA polymerase.

Q7: **Extension** occurs at 70-80°C.

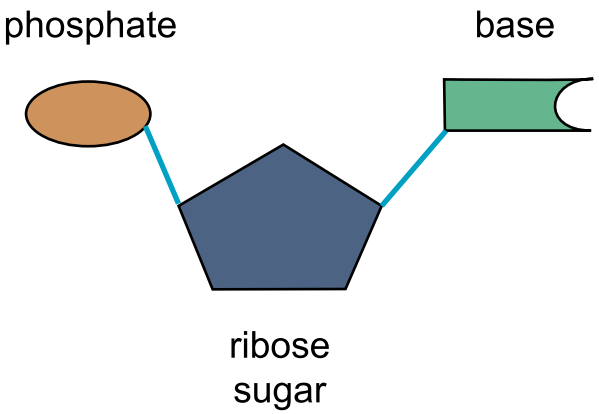
Q8: PCR leads to an **exponential** amplification of desired DNA sequences.

Q9: b) 8

Topic 3: Gene expression

The structure and functions of RNA: Questions (page 29)

Q1:



Q2:

	RNA	DNA
Structure	not a double helix	double helix
Preferred form	single-stranded	double-stranded
Number of types	>1	1
Present in	the cytoplasm and the nucleus	the nucleus
Bases	adenine, uracil, guanine, cytosine	adenine, thymine, guanine, cytosine

Q3: Phosphate, ribose sugar, base

Q4: (A) adenine, (U) uracil, (G) guanine, (C) cytosine

Q5: mRNA, tRNA, rRNA

Q6: **mRNA** copies the code from the DNA molecule and carries it out to the ribosomes where the proteins are synthesised.
tRNAs are found in the cytoplasm, attaching to specific amino acids and bringing them to the ribosomes where the amino acids are joined together.

Q7: c) both DNA and RNA

Transcription: Questions (page 34)

Q8: RNA polymerase

Q9: Nucleus

Q10: Hydrogen bonds

Q11: Primary transcript / Pre-mRNA

Q12: RNA splicing

Translation: Questions (page 37)

Q13:

1. The double-stranded DNA unwinds, hydrogen bonds in the DNA break and the DNA strands separate
2. An RNA nucleotide binds to a complementary nucleotide on one of the DNA strands
3. Hydrogen bonds form between the bases on the RNA and the DNA nucleotides
4. The RNA nucleotides are linked together to form messenger RNA (mRNA)
5. When synthesis of the mRNA is completed, the mRNA separates from the DNA
6. The mRNA leaves the nucleus and enters the cytoplasm
7. A ribosome attaches to the mRNA. Two transfer RNA (tRNA) molecules are also contained within the ribosome
8. Hydrogen bonds are formed between the first codon of the mRNA and the complementary anticodon on a tRNA
9. The second tRNA binds to the mRNA
10. A peptide bond forms between the amino acids carried by the tRNA molecules
11. The first tRNA leaves the ribosome, and another tRNA enters and base-pairs with the mRNA
12. A second peptide bond is then formed. The process continues, with the ribosome moving along the mRNA
13. As each mRNA codon is exposed, incoming tRNA pairs with it and polypeptide synthesis continues until completed

Q14: Anticodon

Q15: Cytoplasm

Q16: Ribosome

Q17: Peptide bond

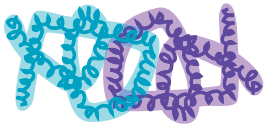
Q18: They cause translation to stop / They are stop codons

Q19:

DNA	C	A	C	A	G	T	G	T	T	T	G	T	C	C	G
mRNA	G	U	G	U	C	A	C	A	A	A	C	A	G	G	C
protein	val			ser			gln			thr			gly		

Protein structure and function: Questions (page 41)

Q20:

Description	Diagram
Chain of amino acids linked by strong peptide bonds	 d)
Polypeptide structure determined by weak hydrogen bonds	 c)
Strong bonds form between special groups of amino acids	 a)
More than one polypeptide makes up the final structure	 b)

Extended response question: Protein synthesis (page 43)

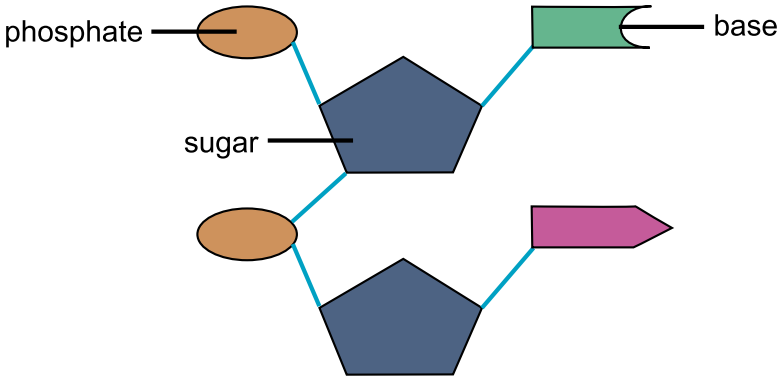
Suggested marking scheme

Each line represents a point worth one mark. The concept may be expressed in other words. Words which are bracketed are not essential. Alternative answers are separated by a solidus (/); if both such answers are given, only a single mark is allocated. In checking the answer, the number of the point being allocated a mark should be written on the answer paper. A maximum of seven marks can be gained.

1. mRNA carries information / code (for proteins) from the nucleus / from DNA.
2. mRNA attaches to ribosome.
3. Three bases on mRNA is a codon.
4. tRNA transports amino acids to ribosome.
5. tRNA transports specific amino acids.
6. Three bases on tRNA is an anticodon.
7. Codons match / pair with their anticodons.
8. Joins / adds correct amino acid onto growing protein/polypeptide.
9. Sequence of bases / codons on mRNA gives sequence of amino acids.

End of Topic 3 test (page 43)

Q21: 3. A, C and B - as per:



Q22: c) It is single-stranded; it has ribose in its backbone; it contains the base uracil

Q23:

a) CGUAAGUAACGU
b) Transcription
c) RNA polymerase

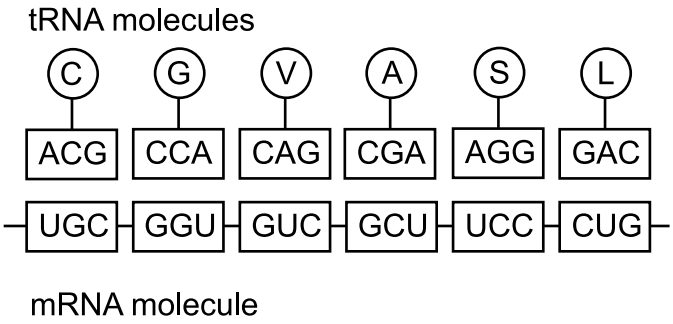
Q24:

The mRNA produced after transcription is called the **primary transcript**; the **introns** are removed, leaving only the **exons** in the final **mature transcript**.

Q25: RNA splicing

Q26: Translation

Q27: CGVASL - as per:



Q28: Anticodon

Q29: Alternative splicing

Q30: Hydrogen

Topic 4: Differentiation in multicellular organisms

Stem cells: Questions (page 54)

Q1: Stem cells have the ability to go through numerous cycles of cell division while maintaining the undifferentiated state.

Q2: Stem cells are undifferentiated cells; this allows them to undergo the process of cell differentiation and gives them the ability to form other types of body cells.

Q3:

- A muscle cell: movement;
- a red blood cell: transport of oxygen;
- a nerve cell: carries impulses.

If you have a different answer, check it with your teacher as it may still be correct.

Embryonic stem cells: Question (page 56)

Q4: The stages of the process of using hESCs to form specialised cells are as follows:

1. Early human embryo Blastocyst
2. Embryo stem cell removed
3. Stem cell cultured in the laboratory
4. Formation of specialised cells: nerve cell, muscle cell, gut cells
5. Undifferentiated stem cells cultured in different culture conditions
6. Formation of mass of undifferentiated stem cells

Extended response question: Stem cells (page 61)

Suggested marking scheme

Each line represents a point worth one mark. The concept may be expressed in other words. Words which are bracketed are not essential. Alternative answers are separated by a solidus (/); if both such answers are given, only a single mark is allocated. In checking the answer, the number of the point being allocated a mark should be written on the answer paper. A maximum of six marks can be gained.

Embryonic stem cells *(maximum of 2 marks)*:

1. Found in developing embryo (blastocyst).
2. Have the capacity to become all cell types.
3. Can be grown relatively easily in culture.

Tissue stem cells *(maximum of 4 marks)*:

i. Found in body tissues.

ii. Are more differentiated than embryonic stem cells.

iii. Can only differentiated into a narrow range of cell types.

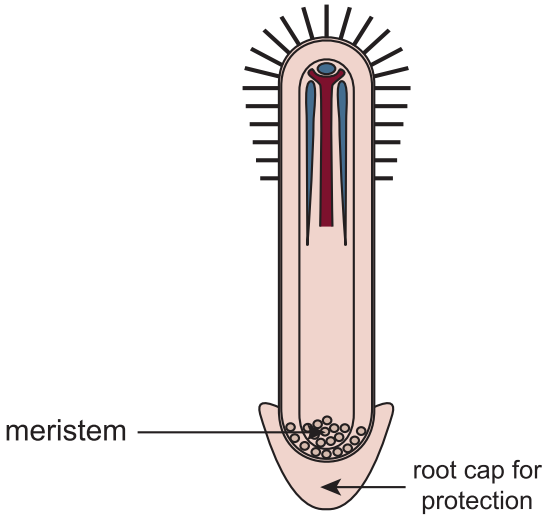
iv. Are rare in mature tissues.

v. Give rise to a limited range of cell types.

vi. Develop into cell types that are closely related to the tissue in which they are found.

End of Topic 4 test (page 63)

Q5: c) It contains specialised cells that differentiate



Q6: **D**, as per

Q7: An unspecialised cell.

Q8: They self-replicate and differentiate.

Q9: Bone marrow, skin epidermis, brain, blood, other suitable organs. *(pick any two)*
If you have a different answer, check it with your teacher as it may still be correct.

Q10: To replace differentiated cells.

Q11:

- Skin grown from stem cells to treat burn victims.
- Bone marrow transplant of stem cells to treat leukaemia.
- Production of replacement organs such as a windpipe for transplants.
- Testing new drugs using stem cells.

Q12: It involves the destruction of embryos.

Topic 5: Structure of the genome

The genome: Question (page 68)

Q1:

Process	Description
Transcription	DNA copied to RNA
Splicing	Introns removed from pre-mRNA
Translating	Exons pass to ribosome where polypeptides are assembled

End of Topic 5 test (page 70)

Q2: The **genome** of an organism is its **hereditary** information encoded in **DNA**.

Q3: c) genes

Q4: False

Q5: mRNA, tRNA, rRNA, RNA fragments *(pick any two)*

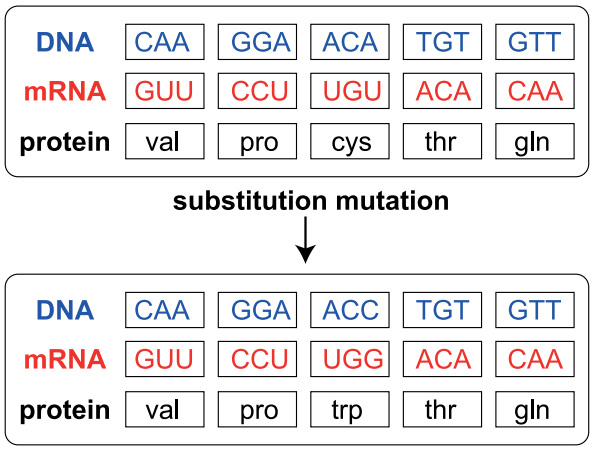
Q6: False

Q7: False

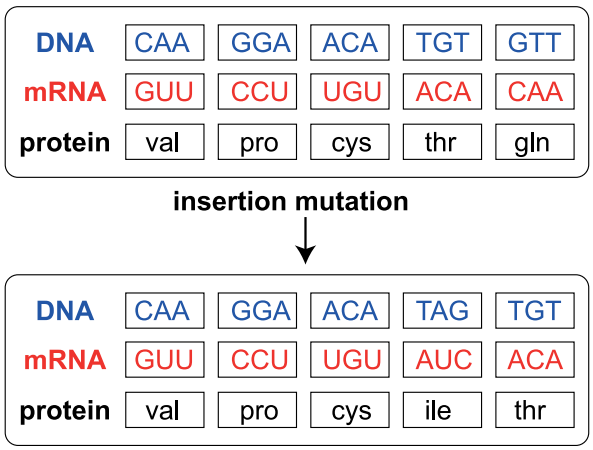
Topic 6: Mutations

Gene mutations: Questions (page 77)

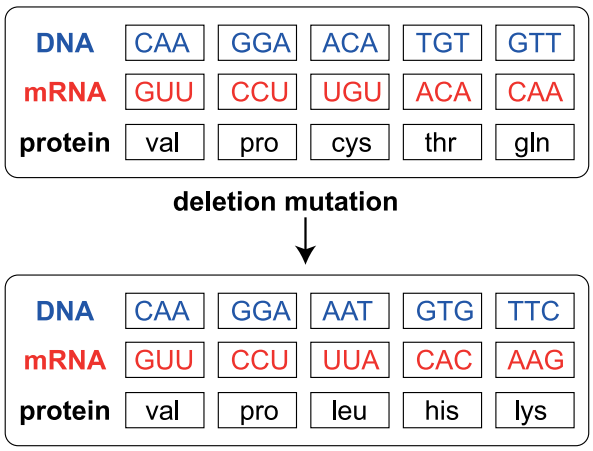
Q1:



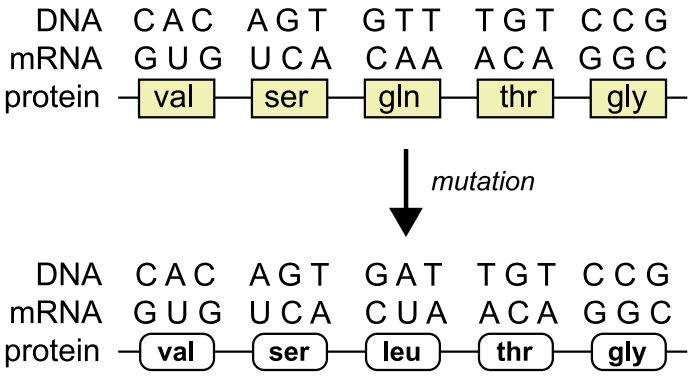
Q2:



Q3:

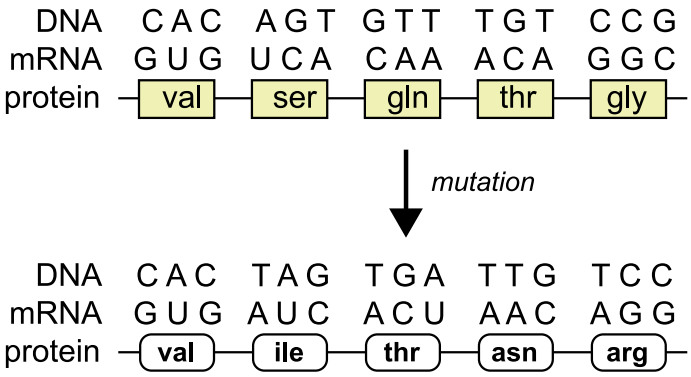


Q4:



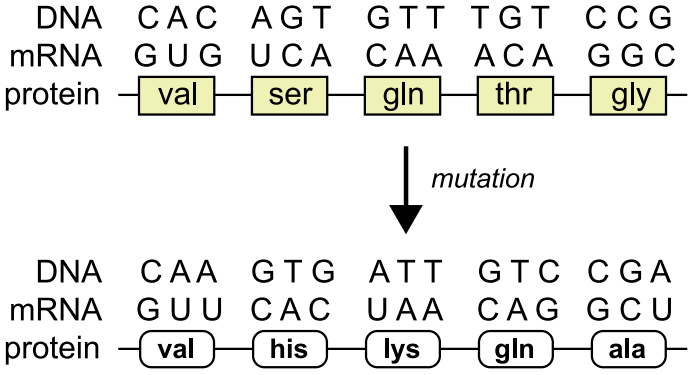
Which kind of mutation is shown? **substitution**
(deletion, insertion or substitution?)

Q5:



Which kind of mutation is shown? **insertion**
(deletion, insertion or substitution?)

Q6:



Which kind of mutation is shown? **deletion**
(deletion, insertion or substitution?)

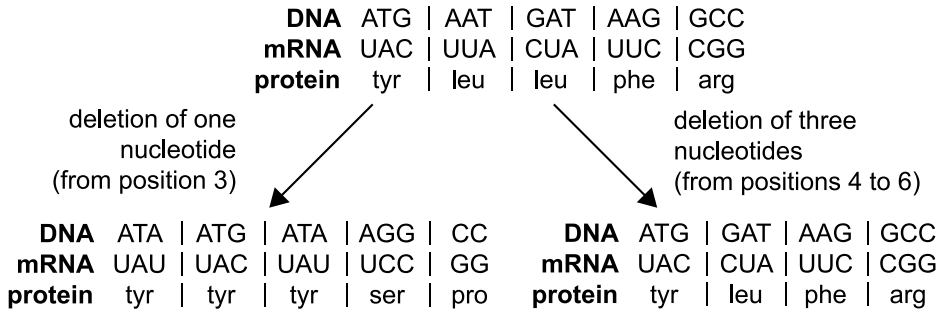
Effects of gene mutations on amino acid sequences: Questions (page 82)

Q7: **4** amino acids have been changed; a deletion mutation is a **frameshift** mutation.

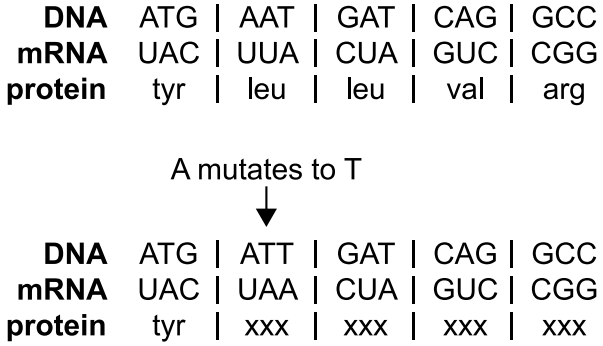
Q8: **4** amino acids have been changed; an insertion mutation is a **frameshift** mutation.

Q9: **1** amino acid has been changed; a substitution mutation is a **point** mutation.

Q10: The deletion of three nucleotides may 'cancel out' the frame shift mutation if the deletion corresponds to a base triplet that is transcribed into a mRNA codon. The pattern of base triplets in the DNA appearing after the mutation will be unaffected. There will be a loss of one amino acid from the protein, but the other amino acids will remain the same. The answer is illustrated by the example given:



Q11: A mutation may introduce a 'stop' codon (UAA, UAG or UGA) into the encoded mRNA. Synthesis of the protein will cease when a stop codon is encountered in the mRNA during translation, and a shorter protein will be produced as a result. Here is an example of how this might occur:



Differences between gene and chromosome mutation: Questions (page 85)

Q12: Chromosome mutation - deletion

Q13: Gene mutation - deletion

Q14: Gene mutation - insertion

Q15: Gene mutation - substitution

Q16: Chromosome mutation - duplication

Q17: Chromosome mutation - translocation

Extended response question: Gene mutations (page 88)

Suggested marking scheme

Each line represents a point worth one mark. The concept may be expressed in other words. Words which are bracketed are not essential. Alternative answers are separated by a solidus (/); if both such answers are given, only a single mark is allocated. In checking the answer, the number of the point being allocated a mark should be written on the answer paper. A maximum of eight marks can be gained.

Gene mutations *(maximum of 3 marks)*:

1. A gene mutation is the replacement or altering of a single nucleotide within a DNA sequence.
2. A substitution mutation means that one nucleotide is substituted for another.
3. An insertion mutation means that one or more nucleotides are inserted into the DNA.
4. A deletion mutation means that one or more nucleotides are removed from the DNA.

Consequences *(maximum of 5 marks)*:

i. The effect of a mutation will depend on its type / location.

ii. Substitution mutations can result in a missense mutation where a single incorrect amino acid is inserted into a protein.

iii. It can also result in a nonsense mutation which results in the code for an amino acid being changed to a stop codon.

iv. Splice-site mutations result in some introns being retained and/or some exons not being included in the mature transcript.

v. Insertions / deletions of 1 or 2 nucleotides mean that all the bases downstream are moved up or down from their place; this means the reading frame is altered.

vi. This type of mutation is known as a frame-shift mutation.

End of Topic 6 test (page 89)

Q18: d) they are random, infrequent occurrences.

Q19: d) Translocation

Q20: a) Deletion

Q21: c) Insertion, deletion, substitution - as per:

Normal sequence		N N **C A C G T A A C G T** N N
Insertion	A	N N **C A C G T A A C C G T** N
Deletion	B	N N **C A G T A A C G T** N N N
Substitution	C	N N **C A C G A A A C G T** N N

Q22: c) Insertion or deletion

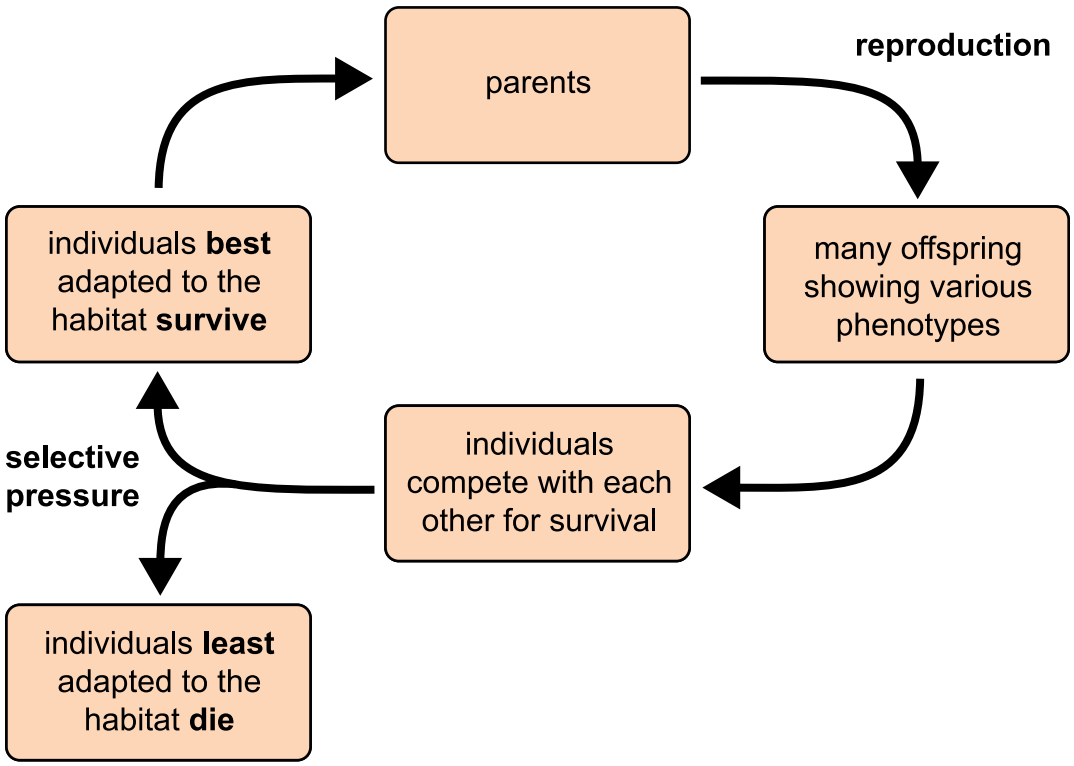
Q23:

1. Nonsense
2. Missense

Topic 7: Evolution

Natural selection: Questions (page 97)

Q1:



Q2: c) All individuals in a population display the same phenotypes.

Q3: b) The reproductive success it experiences during its lifetime.

Species: Questions (page 99)

Q4: A population is a group of individuals that belong to the same species and can interbreed with each other.

Q5: The gene pool is the sum of all the different alleles contained within a population. The allele frequency is the abundance of any given allele in a population.

End of Topic 7 test (page 104)

Q6: Evolution can be described as a change in a **species** over time, and is driven by **natural selection**.

Q7: Inheritance can be described as the passing of **genes / genetic material** between generations.

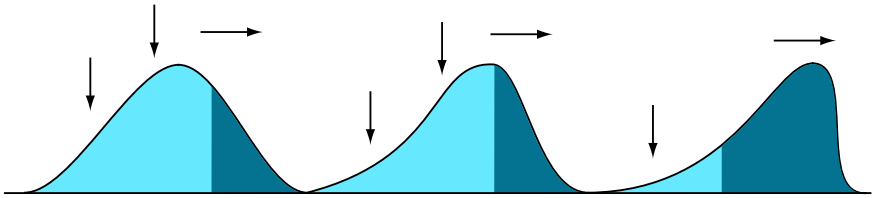
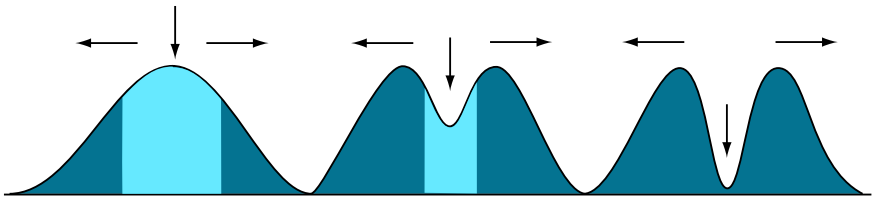
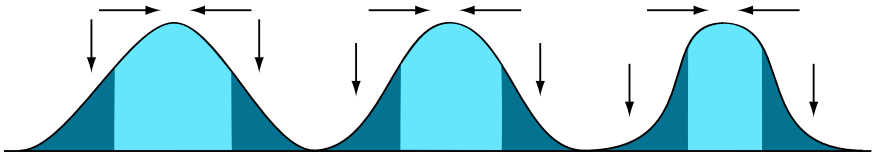
Q8: The exchange of plasmids between bacteria is an example of **horizontal** inheritance.

Q9:

i. 12000

ii. 400%

Q10:

A directional selection	
B disruptive selection	
C stabilising selection	

Q11:

i. Stabilising selection

ii. Disruptive selection

iii. Directional selection

Q12: False

Q13: Sympatric speciation uses behavioural or ecological barriers.

Topic 8: Genomics

Phylogenetics: Question (page 111)

Q1: All life forms are now described as belonging to one of three **domains**. This is largely based on a comparison of **DNA**. The three main groups are **bacteria**, **archaea** and **eukaryota**. Phylogenetic clocks need to be calibrated by using **fossil records**.

End of Topic 8 test (page 118)

Q2: Determining the order of nucleotide bases is known as **sequencing**.

Q3: Computer and statistical analyses.

Q4: To show the relationship between organisms.

Q5: Fossil evidence

Q6: The **genome** is the sum total of an organism's DNA.

Q7: No

Q8: The genes are the same / very similar.

Q9: Screening for genetic defects, better targeted treatment, lower costs of treatment.

Q10: Pharmacogenetics

Topic 9: End of unit test

End of Unit 1 test (page 120)

Q1: Hydrogen bond

Q2: 15%

Q3: Deoxyribose sugar

Q4: X - phosphate; Y - deoxyribose sugar; Z - base.

Q5: Primer

Q6: Nucleotides / ATP.

Q7: DNA polymerase

Q8: The **hydrogen** bonds of DNA are separated by **heating** during PCR.

Q9: c) Primers unwind double helix.

Q10: b) 8

Q11: Transcription

Q12: Nucleus

Q13: Translation

Q14: Ribosome

Q15: RNA polymerase

Q16: Choose an answer from:

- the introns / non-coding regions of genes are removed;
- due to RNA splicing;
- the mature mRNA only contains exons / coding regions of genes.

Q17: UAC

Q18: Peptide bonds

Q19: Hydrogen bonds

Q20: Cells that are undifferentiated / unspecialised.

Q21: Choose an answer from:

- Is it safer than using the drug directly on volunteers?
- Is it right to use embryos to extract stem cells?
- Is it right to deprive sufferers of potential treatment?
- Is it right to use stem cells rather than animals?

Q22: Tissue stem cells are more differentiated than embryonic stem cells.

Q23: Meristems are regions of **unspecialised / undifferentiated** cells in plants that are capable of cell **division**.

Q24: A region of the DNA molecule which codes for a protein.

Q25: tRNA / rRNA / RNA fragments

Q26: b) Substitution

Q27: Missense

Q28: Substitution

Q29: 2. Duplication

Q30: Bacteria can on occasions pass genetic material between themselves. This may be a section of DNA called a **plasmid**.

Q31: This is representative of **horizontal** genetic transfer.

Q32: c) Prokaryotes and eukaryotes

Q33:

- **Stabilising** selection is where the average phenotype is most successful for a particular habitat.
- **Disruptive** selection is characterised by the extreme versions of a phenotype being selected.
- **Directional** selection is characterised by the selection of one extreme phenotype at the exclusion of all others.

Q34: A group of organisms which can interbreed to produce fertile offspring.

Q35: Geographical

Q36: They cannot interbreed.

Q37: Sympatric

Q38: DNA sequencing

Q39: Chimps are **more** closely related to gorillas than orangutans.

The common ancestor of chimps and gorillas is **more** recent than the common ancestor of gorillas and orangutans.

Q40: Bioinformatics

Q41: Comparison of genomes reveals that many genes are highly **conserved** across different organisms.

Q42: Pharmacogenetics

Q43: Control

Q44: To allow a comparison with those which were exposed to gamma radiation.

Q45: Choose an answer from:

- spacing of the chickpeas;
- volume of water;
- length of time exposed to gamma radiation.
- ph;
- oxygen concentration

Q46: As the gamma radiation dose increases from 0 to 800 Gy, the percentage germination decreases from 99% to 39%. As the radiation increases further to 900, Gy percentage germination increases to 53%.

Q47: 68%

Q48: Choose an answer from:

- gamma radiation does not affect germination of desi chickpeas;
- gamma radiation affects kabuli chickpeas more than desi chickpeas.