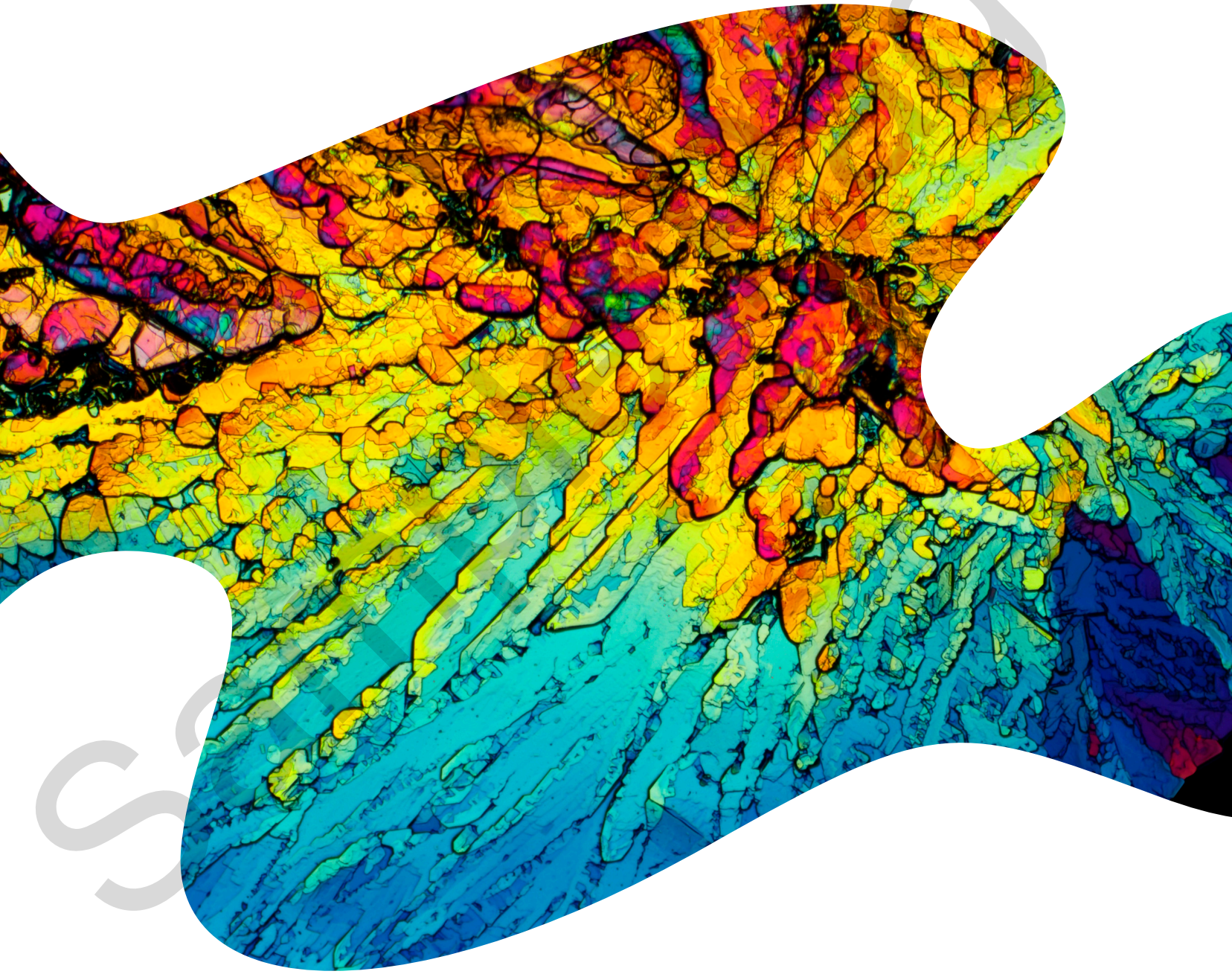
HEINEMANN

BIOLOGY 2

6TH EDITION



VCE UNITS 3 AND 4 • 2022-2026

CHAPTER

04 DNA manipulation techniques and applications

Learning outcomes

By the end of this chapter, you will have developed an understanding of the molecular tools and techniques used to manipulate DNA molecules for particular purposes. You will have learnt how recombinant plasmids are created and then used as vectors in the process of bacterial transformation. You will also have an understanding of the other molecular tools and techniques explored in this chapter, including gel electrophoresis and the polymerase chain reaction (PCR).

Key knowledge

- the use of enzymes to manipulate DNA, including polymerase to synthesise DNA, ligase to join DNA and endonucleases to cut DNA **4.1, 4.2**
- the function of CRISPR-Cas9 in bacteria and the application of this function in editing an organism's genome **4.3**
- amplification of DNA using polymerase chain reaction and the use of gel electrophoresis in sorting DNA fragments, including the interpretation of gel runs for DNA profiling **4.1**
- the use of recombinant plasmids as vectors to transform bacterial cells as demonstrated by the production of human insulin **4.2**
- the use of genetically modified and transgenic organisms in agriculture to increase crop productivity and to provide resistance to disease. **4.3**

4.1 DNA manipulation

FIGURE 4.1.1 A scientist loads DNA samples into a gel electrophoresis chamber.

A microsatellite is a short repeated sequence of nucleotides found at a defined location (locus) on a chromosome.

FIGURE 4.1.2 A scientist extracts fossilised DNA from a Neanderthal (*Homo neanderthalensis*) bone. The DNA will be amplified for further DNA analysis.

DNA sequencing describes any process used to determine the order of the four nucleotide bases—adenine, thymine, guanine and cytosine—in DNA.

To work with DNA, it is necessary to have more than a few DNA molecules. The **polymerase chain reaction (PCR)** is a technique used for DNA amplification—it makes millions of identical copies of a piece of DNA. Using PCR, scientists can amplify the DNA in traces of blood left at the scene of a crime, or a particular gene from a sample of DNA.

Gel electrophoresis is a method used to separate and visualise nucleic acids and proteins according to their size (Figure 4.1.1). This is usually performed after PCR to either confirm that the correct DNA fragment was amplified, or to identify DNA fragments present in the sample (DNA profiling). It is also used to separate DNA fragments to be used in DNA manipulation techniques.

In this section, you will learn about PCR and gel electrophoresis of DNA fragments.

DNA AMPLIFICATION

Many DNA manipulation techniques require a large quantity of DNA to work with. However, sometimes only a very small sample of DNA is available for scientists. For example, only trace samples of DNA may be left at a crime scene or extracted from fossils of extinct species (Figure 4.1.2). Usually only small samples can be removed for medical tests and in embryonic or fetal DNA screening for genetic disorders. In these cases, **DNA amplification** is required to increase the amount of the target DNA sample so that it is large enough to be used or analysed in other techniques and processes.

DNA amplification uses PCR to create a large quantity of DNA that is identical to the initial trace sample. The term **target DNA** is used to describe a particular region of a DNA molecule that a scientist intends to study or manipulate (e.g. a specific gene or a **microsatellite**, which is a variable region of the genome used for DNA profiling).

Polymerases

As the name indicates, polymerase chain reaction (PCR) is based on the action of polymerases. **Polymerases** are enzymes that catalyse the formation of long-chain molecules (polymers), such as DNA and RNA, by linking smaller molecules (nucleotides). These enzymes are found in all living organisms and play an important role in the replication, repair and maintenance of DNA. There are two different groups of polymerases:

- DNA polymerases
- RNA polymerases.

DNA polymerases

Scientists use **DNA polymerase** enzymes in PCR and **DNA sequencing** to synthesise multiple copies of the target DNA. During PCR, the DNA polymerase attaches to the end of the target DNA sequence (called the template) and adds complementary nucleotides to create a new strand of DNA, complementary to the target DNA.

DNA polymerases used in PCR must be stable at high temperatures, have high affinity to the DNA template, and be highly specific to reduce background amplification (i.e. amplification of non-target DNA).

BIOFILE

Thermus aquaticus

A field trip to Yellowstone National Park in the 1960s radically altered the course of molecular genetics research. Thomas Brock, a bacteriologist, found a new species of bacteria in a hot spring, which he named *Thermus aquaticus* (Latin for 'hot water').

This was significant because enzymes are normally denatured if heated to temperatures of 95°C for more than a few seconds. For *T. aquaticus* to survive in the hot springs, its enzymes, including DNA polymerase, need to tolerate these high temperatures. Therefore, the DNA polymerase from *T. aquaticus* (*Taq* polymerase) has proved to be an ideal enzyme for use in PCR.

A scientist obtains a sample of *Thermus aquaticus* from a hot spring.

Taq polymerase

***Taq* polymerase** is the DNA polymerase that is most commonly used in PCR. It was originally extracted from the thermophilic bacterium *Thermus aquaticus*. The heat-resistant properties of *Taq* polymerase make it extremely useful in DNA manipulation techniques such as PCR.

Thermophilic means 'heat-loving'. Thermophilic organisms live in environments with high temperatures.

Reverse transcriptase

Reverse transcriptase is a DNA polymerase that synthesises single-stranded DNA using single-stranded RNA as a template. This is the reverse of the usual transcription process in which DNA is transcribed into RNA. Reverse transcriptase is used in the laboratory to produce DNA molecules that can be amplified by PCR for further analysis. It is also used to make **complementary DNA** (**cDNA**) from modified mRNA that has already had the introns spliced out for some DNA manipulation techniques, as you will learn in Section 4.2.

RNA polymerases

RNA polymerases are enzymes that synthesise RNA from DNA during transcription. Subunits of the RNA polymerase recognise the promoter at the start of a gene. The RNA polymerase attaches to the promoter and unwinds the DNA, allowing it to add nucleotides, one at a time, in the 5' to 3' direction. RNA polymerase adds nucleotides until a 'stop' sequence is reached. A single-stranded piece of RNA is created. RNA polymerase works much more slowly than DNA polymerase. RNA polymerase is used in specialised laboratory techniques to study transcription and RNA amplification.

The polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) is a method of amplifying specific target sequences of DNA. The DNA polymerase known as *Taq* polymerase is used in this process, as it is stable at the high temperatures required during PCR. Each strand of the DNA acts as a template for a new copy of itself.

PCR is carried out in cycles. Each cycle doubles the amount of DNA, resulting in exponential growth. After 30 cycles, there are over one billion copies (Table 4.1.1).

TABLE 4.1.1 Exponential growth in number of target DNA molecules during PCR

Number of PCR cycles (n)	Number of double-stranded copies of original DNA (2^n)
0	1
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024
20	1 048 576
30	1 073 741 824

For PCR to be carried out, specialised equipment and a PCR mixture is required. The PCR mixture contains:

- DNA, including the target DNA to be amplified (the sample) (Figure 4.1.3)
- free nucleotides, to build new DNA strands
- a heat-resistant DNA polymerase (usually *Taq* polymerase), to elongate the new DNA strands by adding the free nucleotides
- two DNA **primers** complementary to the ends of the target DNA, to specify the start and finish of the DNA fragment to be amplified. The two primers are synthetic, single-stranded DNA molecules up to 30 bases in length.

FIGURE 4.1.3 A scientist loads DNA samples into a thermocycler to be amplified by PCR.

Steps in the polymerase chain reaction

The PCR mixture is placed in a **DNA thermocycler**, which alters the temperature in pre-programmed steps. Each PCR cycle involves three steps (Figure 4.1.4):

1. Denaturation: The sample is heated to 95°C to break the hydrogen bonds between the two strands of double-stranded DNA to obtain single strands of DNA.
2. Annealing: The temperature is reduced to 50–60°C. This allows the primers to **anneal** (i.e. bind) to complementary sequences on opposite strands at each end of the target DNA sequence.

Primers determine the start and end points of a nucleotide sequence to be amplified.

3 Extension: The temperature is increased to 72°C. This allows *Taq* polymerase to attach to the primers on the DNA strands. The *Taq* polymerase moves along each strand, adding free nucleotides to form double-stranded DNA.

This three-step cycle of heating and cooling is repeated up to 50 times to ensure there is sufficient target DNA produced to work with.

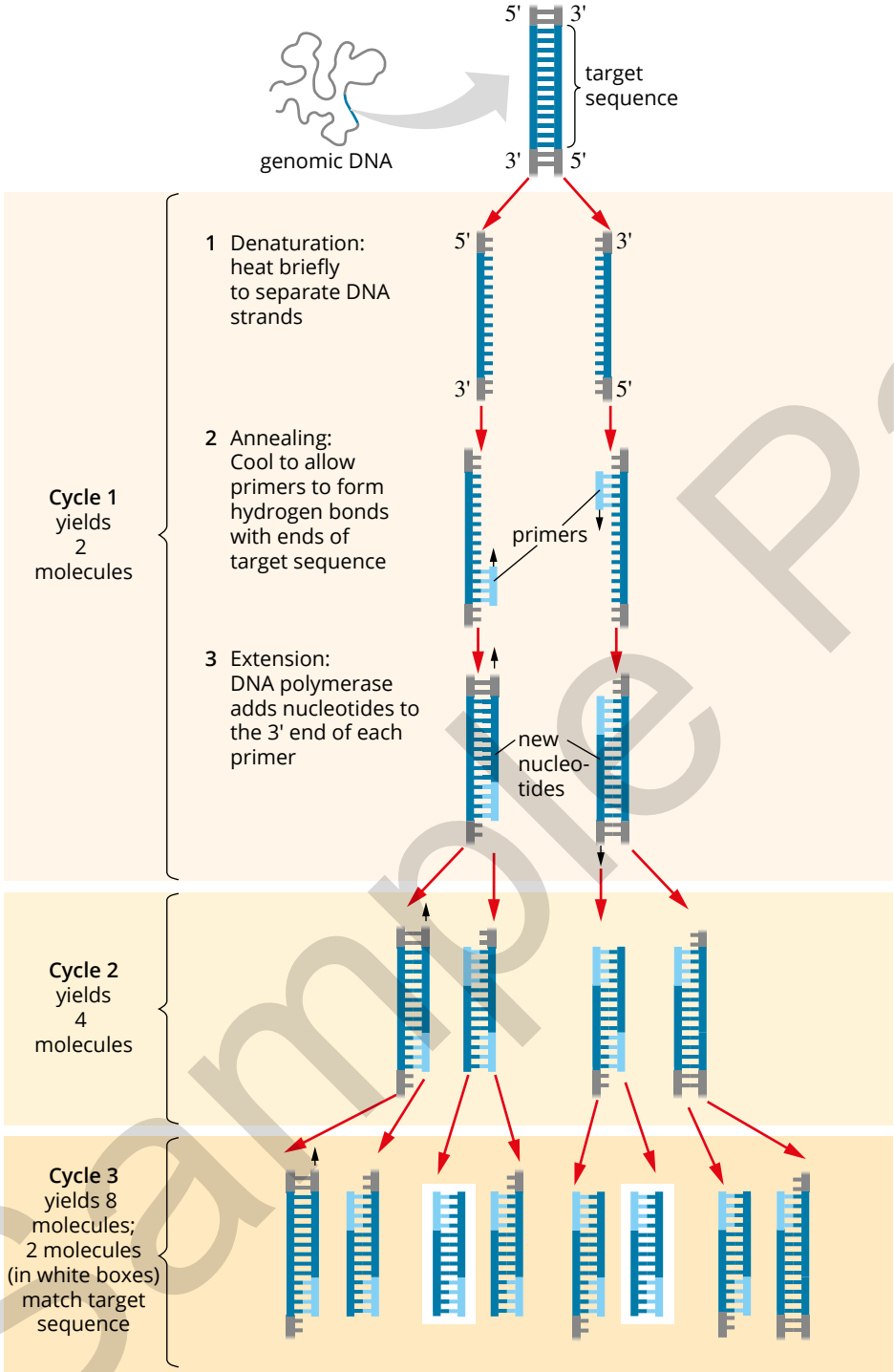


FIGURE 4.1.4 The three steps of PCR: denaturation, annealing and extension. Each PCR cycle increases the amount of target DNA exponentially.

BIOFILE

Trace samples

The process of PCR requires as few as one or two cells for DNA amplification. Scientists at the Victoria Forensic Science Centre have shown that merely touching an object deposits sufficient material for successful DNA amplification. In handling keys, opening a door or driving a car, the cellular material deposited by a criminal provides ample DNA for analysis following PCR.

PA 3

DNA SEPARATION BASED ON FRAGMENT SIZE

Gel electrophoresis is commonly performed after DNA amplification using PCR. Gel electrophoresis allows scientists to separate out the different DNA fragments present in a sample based on their size. This information can be used to match fragments from different samples, as in DNA profiling, or to isolate a particular fragment for further use in another technique, such as DNA recombination and bacterial transformation. You will learn about these techniques in Section 4.2.

Gel electrophoresis

DNA is negatively charged because of the negative charge on the phosphate group in each nucleotide.

Gel electrophoresis is a technique for separating fragments of nucleic acids (DNA and RNA) or to study protein molecules. In this section, only DNA fragments will be considered. When an electric current is applied to the gel, the negatively charged pieces of DNA move through the gel towards the positive terminal. Small DNA molecules move faster than large ones, causing them to separate based on their size.

Gel electrophoresis is used to compare DNA fragments for a number of applications such as:

- DNA screening, such as testing for inherited genetic conditions
- confirming the correct gene has been amplified in PCR
- identifying DNA fragments to be used for genetic engineering.

The following steps and Figure 4.1.5 outline the process of gel electrophoresis:

1. An electrophoresis gel is prepared. It has a jelly-like texture and is usually composed of agarose (a purified form of agar). The gel is rectangular and contains small wells (holes) at one end.
2. The gel is placed into a gel electrophoresis chamber with the wells situated at the negative terminal of the chamber.
3. Each DNA sample is loaded into one of the wells within the gel.
4. A **DNA ladder** containing DNA fragments of known lengths is also run on the gel for comparison with the samples. This allows the length of the sample DNA fragments to be estimated (Figure 4.1.6). A DNA ladder is also known as a molecular weight standard or size standard.
5. The gel is placed in an electrophoresis bath where it is covered with a controlled pH solution that contains ions to conduct an electric current.
6. A power source is attached to the electrophoresis bath and switched on. The electrical current causes the negatively charged DNA fragments to migrate through the gel towards the positive terminal of the chamber.

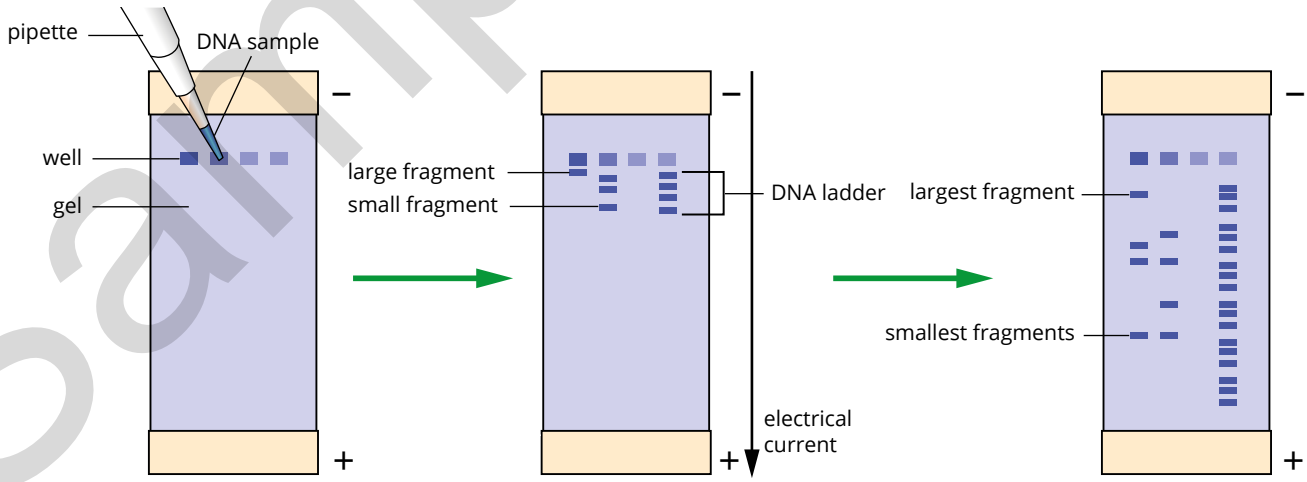


FIGURE 4.1.5 The process of gel electrophoresis, showing two DNA samples being loaded into wells and migration of the DNA fragments through the wells based on their size. The last well contains a DNA ladder for comparing and estimating the fragment length of the two DNA samples.

7 Smaller fragments move faster through the gel, so they migrate further through the gel than larger fragments in a given period of time. This sorts the fragments by length.

8 The DNA fragments are made visible by applying a stain that binds to DNA. This can be done with a fluorescent stain (which may be included in the gel or added after) or with methylene blue stain (which is added after running the gel). Fluorescent stains are viewed with ultraviolet light, while methylene blue staining can be visualised by eye. Bright-coloured bands are observed wherever there is DNA present in the gel (Figure 4.1.7).

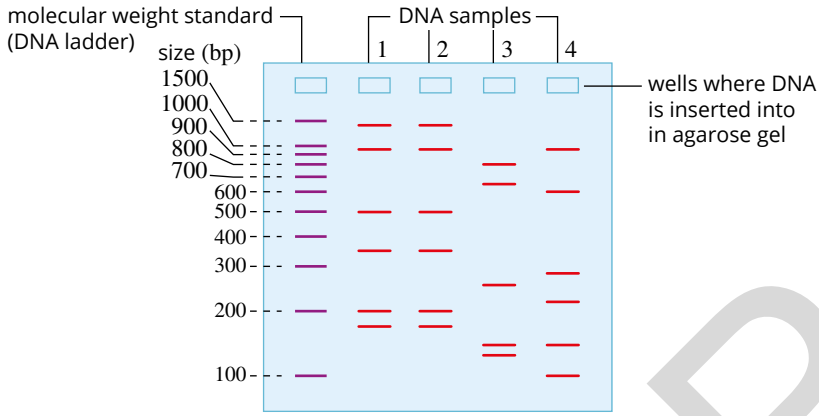


FIGURE 4.1.6 Gel electrophoresis allows a DNA fragment's length to be measured in base pairs (bp) by comparing it with the DNA ladder. Particular fragments can be identified and compared.

FIGURE 4.1.7 A scientist cuts a sample from an electrophoresis gel stained to show DNA fragments as pink fluorescent bands.

CASE STUDY

Surveillance of waterborne RNA viruses

Poor sanitation and faecal contamination of the water supply with viruses, bacteria and protists is a major health problem in many regions globally. The World Health Organization (WHO) provides guidelines on water safety and promotes improved methods of monitoring microbial contamination in drinking water. Outbreaks of poliovirus and hepatitis A and E virus infections may be due to contaminated water. Traditional methods of identifying viruses, by growing them in cells, take several weeks and are not successful for all viruses. Molecular methods such as PCR followed by gel electrophoresis offer faster detection and the potential to detect multiple viruses in one analysis. Viruses such as poliovirus and hepatitis viruses have an RNA genome, so a technique called RT-PCR is used. First, reverse transcriptase (RT) is used to copy the viral RNA into DNA, and then the polymerase chain reaction (PCR) amplifies the DNA.

Primers specific for different viral genes are used in the PCR. Multiple sets of primers can sometimes be used in one PCR, allowing the detection of several viruses at once (Figure 4.1.8). If a virus is present in the water sample, its RNA will be copied to DNA and amplified. The DNA resulting from the reaction can be visualised by the technique of gel electrophoresis.

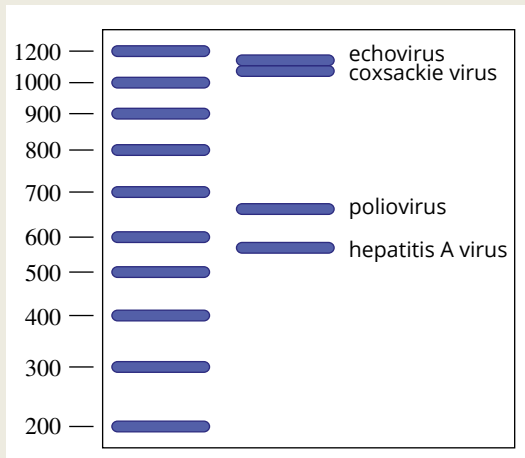


FIGURE 4.1.8 Gel electrophoresis results obtained in water sampling studies. Amplified viral genes were detected as bands of the following size: hepatitis A virus, 589 bp; poliovirus, 671 bp; coxsackie virus, 1084 bp; echovirus, 1128 bp.

COMBINING MOLECULAR TOOLS TO DETECT MUTATIONS

Mutations are usually discovered because of the effect they have on the individual carrying them. If an individual has the symptoms of cystic fibrosis (CF) then the sequence of the gene for CF can be analysed to detect mutations. The gene for CF is the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene. The *CFTR* gene codes for a large membrane protein of the same name, which regulates chloride ion movement across plasma membranes. The *CFTR* gene is very large and many different mutations in this gene can cause disease. The most common mutation, called the ΔF508 mutation, is a deletion of three base pairs, leading to deletion of the amino acid phenylalanine (phe) from position 508 of the protein. Families with a history of CF may wish to undergo screening for this mutation and other common mutations known to cause CF.

PCR and gel electrophoresis can be used to detect the ΔF508 mutant **allele** (gene variant) by:

1. isolating DNA from the individual. The DNA can come from a mouth swab of an adult or from the amniotic fluid surrounding an unborn child.
2. using PCR primers that are complementary to the DNA sequences on either side of the site of the ΔF508 mutation to amplify the DNA.
3. comparing the amplified DNA molecules by gel electrophoresis. For a normal allele the amplified region is 98 base pairs long. In a ΔF508 mutant allele the amplified region is 95 base pairs long (Figure 4.1.9a). A DNA ladder is run next to the samples to enable identification of the normal and mutant alleles based on their size (Figure 4.1.9b).

BIOFILE

Cystic fibrosis

Cystic fibrosis (CF) is an inherited disorder that affects the respiratory and digestive systems. It can significantly shorten the lifespan of people with the condition. In a person with CF, the mucous glands secrete thick, sticky mucus, which clogs the airways, leading to breathing difficulties, respiratory infections and lung damage. The mucus also affects the pancreas, inhibiting the release of important digestive enzymes, which causes a range of nutritional problems. There is currently no cure for CF, but the first treatment that targets the defective CFTR protein, rather than simply treating the symptoms, is now available for some individuals affected with CF in Australia. The medication improves the flow of chloride ions across plasma membranes, which decreases the thickness and stickiness of mucus.

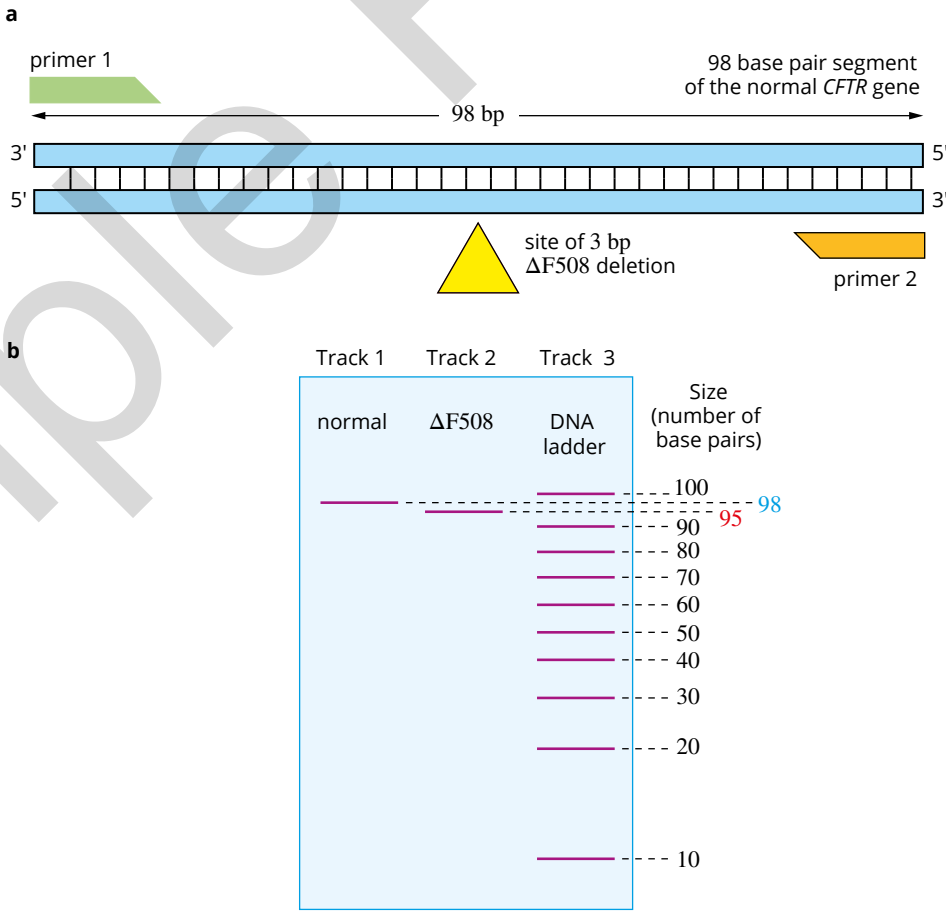


FIGURE 4.1.9 (a) PCR primers that span the site of the ΔF508 mutation are chosen to test for the presence of the CF mutation. (b) Diagram of an electrophoresis gel showing PCR products of an individual carrying two copies of the normal allele of 98 bp (track 1) and an individual carrying two copies of the ΔF508 mutation of 95 bp (track 2). The individual in Track 2 is affected with CF. The DNA ladder in track 3 allows the size of the bands in tracks 1 and 2 to be determined.

CASE STUDY ANALYSIS

Identification of cystic fibrosis carriers

A couple had a healthy daughter but their next child, a boy, failed to grow as expected, had bowel problems, lung congestion and salty-tasting skin. These symptoms led to a diagnosis of cystic fibrosis (CF) (Figure 4.1.10a). The couple were not aware of the incidence of CF in their families, yet a genetic counsellor suggested that they might be carriers of an allele for the disease. To develop CF, you need two copies of the mutant allele (that is, you need to be homozygous for the mutant allele). If you have one copy of the normal allele and one copy of the mutant allele (a heterozygous carrier) you are healthy but able to pass the mutant allele to your children. There is a 25% chance of two carriers having an affected child.

After their son's diagnosis, genetic testing was performed to determine his specific mutant alleles. Testing showed that he carried two ΔF508 mutations (he was homozygous for this mutant allele). Following this the parents underwent carrier testing for the ΔF508 mutation to confirm their carrier status. When the couple fell pregnant with a third child they chose to have genetic tests for the unborn child to determine whether this child could develop CF. A DNA sample from the unborn child was obtained from the amniotic fluid in the uterus and analysed by PCR and gel electrophoresis. The results of the unborn baby's testing, along with the results of the parents, the affected son and the unaffected daughter, are shown in Figure 4.1.10b.

Analysis

1. Using the gel electrophoresis results in Figure 4.1.10b, determine:
   - **a** the size of the mutant allele
   - **b** the size of the normal allele
2. The parents were confirmed as carriers of cystic fibrosis (CF) through genetic testing. Redraw the family pedigree chart, updating the symbols for the parents.
3. Study the gel electrophoresis results in Figure 4.1.10b and determine the CF status of the couple's unborn child.
4. Give your opinion on the social and ethical issues that may arise because of the unborn child's genetic result.

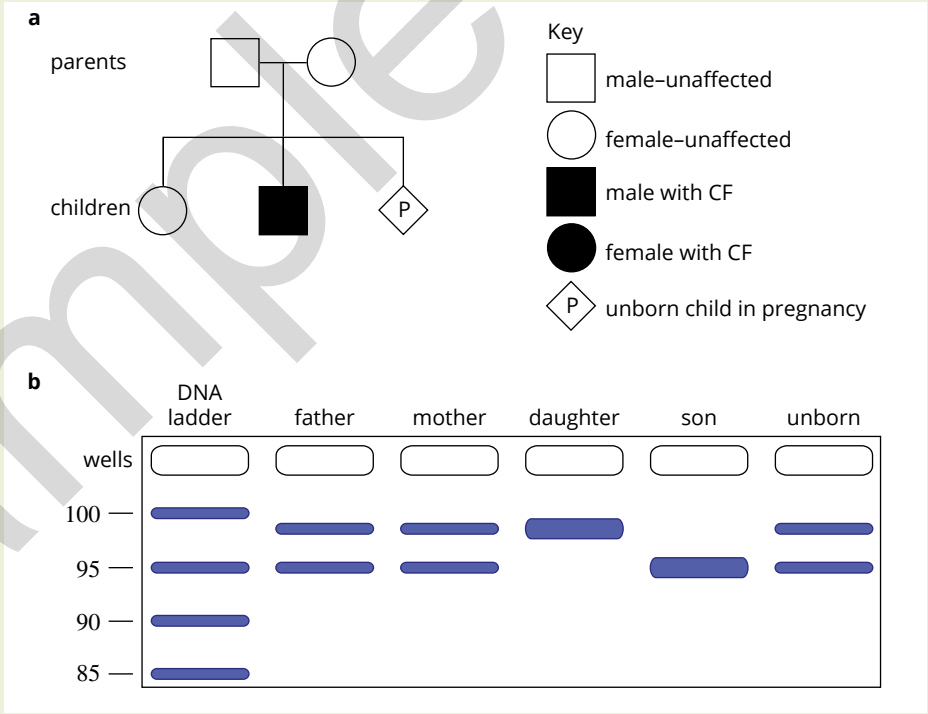


FIGURE 4.1.10 (a) Pedigree chart showing the appearance of CF in the son of two unaffected parents. The first-born daughter is unaffected. The parents want to know whether the unborn child (sex unknown) carries the CFTR mutation. (b) Gel electrophoresis results for the DNA test for the ΔF508 mutation

DNA PROFILING

DNA profiling is a technique that compares and identifies individuals based on their unique DNA sequence. DNA profiling can be used to identify one individual from any other individual. It is often used in forensics to identify the perpetrator of a crime. DNA profiling can also be used to identify bodies after disasters or to confirm if a child is genetically related to a parent (Figure 4.1.11).

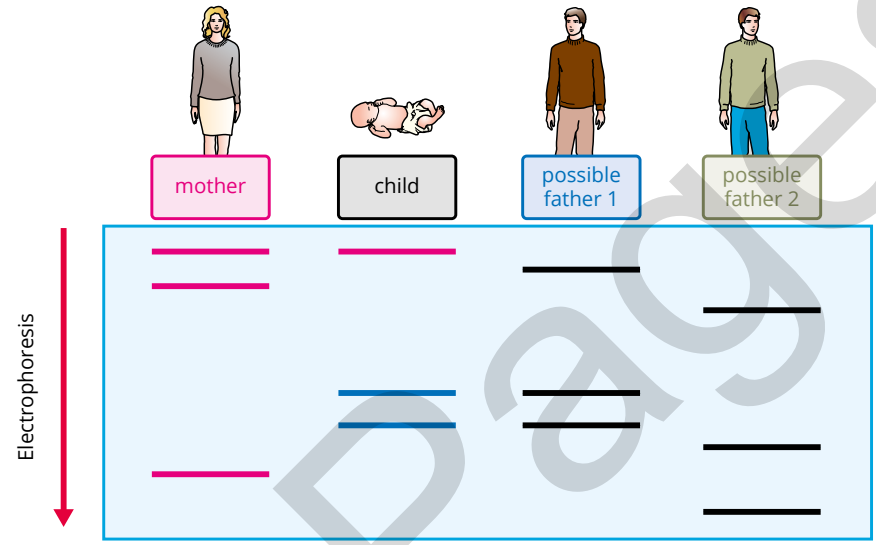


FIGURE 4.1.11 Family members share common bands in DNA profiles, although the combination of bands in each individual is unique.

Short tandem repeats (STRs) are stretches of DNA sequences, usually 2–6 base pairs. that are repeated many times. The number of repetitions varies between individuals.

DNA profiling relies on an individual's unique DNA. The non-coding sections of the DNA (introns) can vary widely between individuals. These inherited variations are called **polymorphisms**. DNA profiling uses the differences between a number of polymorphic sections to identify individuals. In particular, short, repeated sections of between two and six bases, called **short tandem repeats (STRs)**, are examined. STRs are also referred to as microsatellites.

There are thousands of STR loci throughout the human genome and 20 STRs are generally used in DNA profiling. The same STR sequence occurs on each member of a homologous pair of chromosomes, and, by chance, the number of repeats in each STR may be the same or different (Figure 4.1.12). For example, at one location on a chromosome a person might have 25 repeats, while another person might have 45 repeats at the same site.

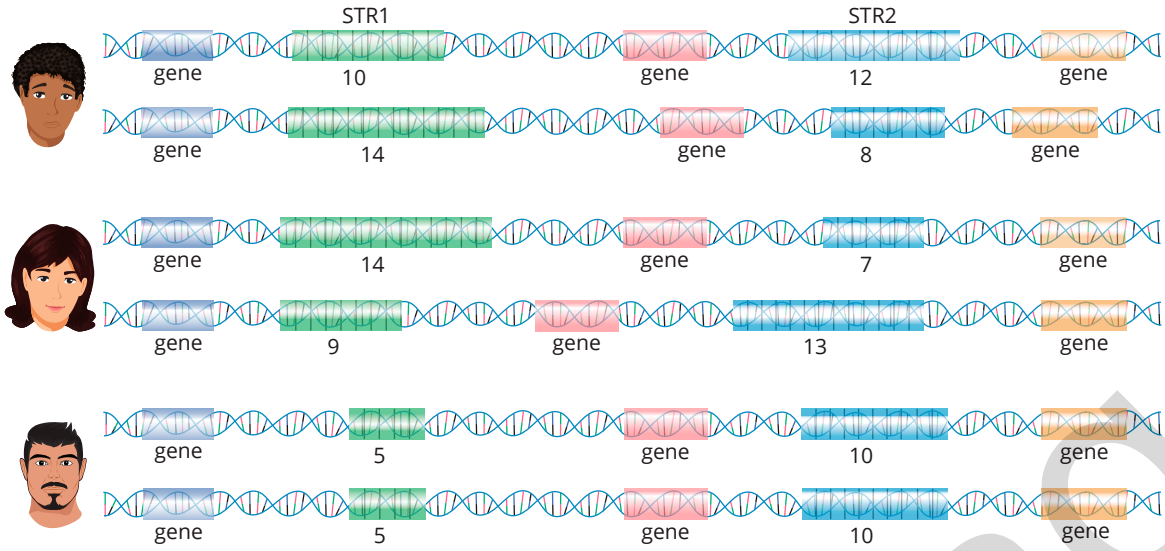
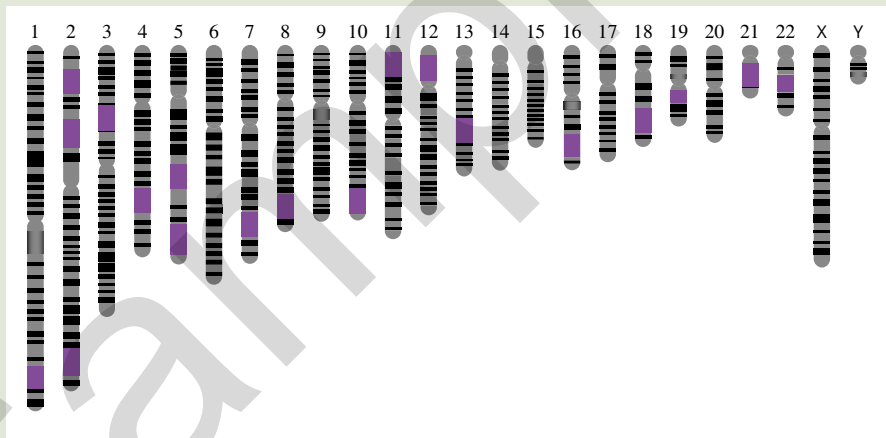


FIGURE 4.1.12 Example of the variation that can be seen in the STRs in three individuals for two STR sites. As chromosomes occur as homologous pairs, each person has two copies of each STR, which may vary in length. The numbers refer to the number of repeats.

BIOFILE

The FBI's CODIS

The 20 STR sites used in DNA profiling are part of CODIS, a United States DNA database developed by the Federal Bureau of Investigation (FBI) in 1997. CODIS stands for Combined DNA Index System. Using 20 STRs gives an extremely high probability that the DNA profile is unique and that the only perfect match of all 20 DNA sites will be with DNA from the same person. DNA profiling also includes regions on the X and Y chromosome for sex determination.



Chromosomal location of the 20 STR CODIS sites used for forensic DNA profiling. Note that there are two STR sites next to each other on chromosome 12.

BIOFILE

Applications of DNA profiling

DNA profiling is used for other applications such as genealogy, biogeographical population comparisons, historical population migration patterns and evolutionary relationships. For these purposes the DNA sequences used for comparisons include STRs (different sites from those used for crime scene analysis), mitochondrial DNA, Y chromosome genes and single nucleotide polymorphisms (SNPs).

Techniques involved in DNA profiling

DNA profiling is a sensitive technique that requires only a small or degraded amount of sample from blood, semen, saliva or hair.

- DNA is extracted from the sample.
- The DNA is digested into smaller fragments by specific restriction enzymes. (You will learn more about restriction enzymes in Section 4.2.)
- The STRs are amplified using PCR, with specific primers for each STR. This produces a much larger sample to test.
- Differences in the size of the STRs can be detected by standard gel electrophoresis or by capillary electrophoresis. Capillary electrophoresis is a rapid, automated system whereby the DNA fragments move in a thin tube under the influence of an electric field. The smaller the size of the fragment, the faster it moves through the capillary tube. As each fragment moves through the tube, a laser detector registers a peak on a graph. As there are two copies of each STR—one on each homologous (paired) chromosome—most STRs appear as a pair of peaks on the graph. If the same number of repeats is on each chromosome, only a single peak is seen (Figure 4.1.13).

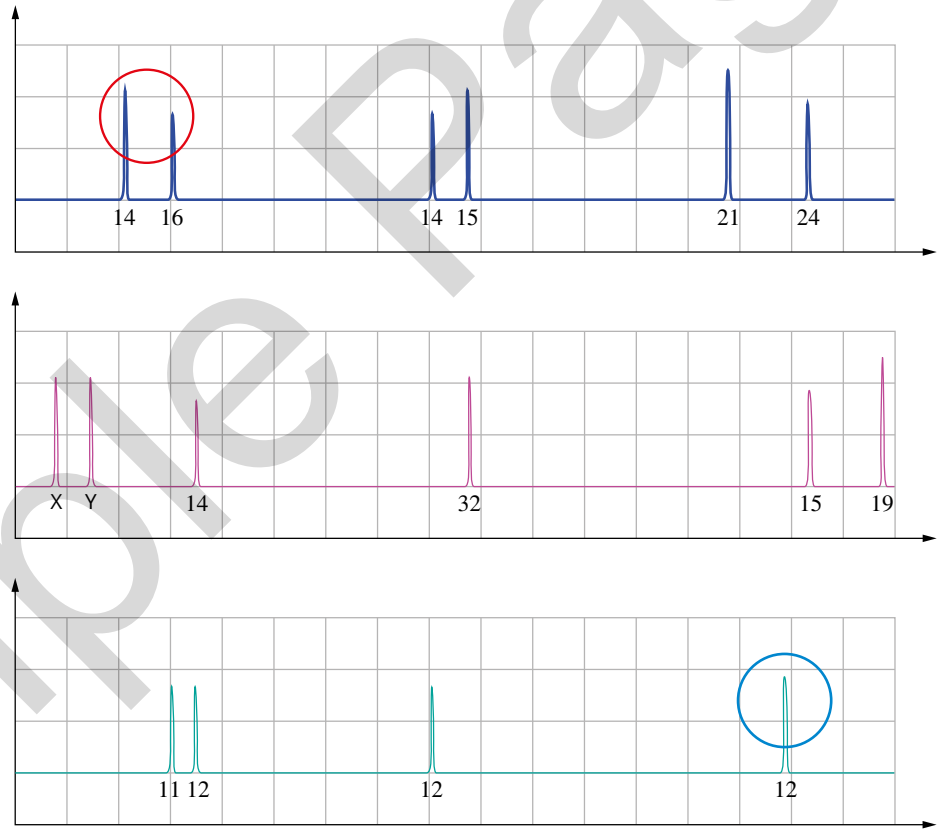


FIGURE 4.1.13 DNA profile of a male. Ten different regions (nine different STRs and the sex chromosome markers) have been analysed. The first two peaks (circled in red) show this person has 14 repeats and 16 repeats of that STR on that homologous pair of chromosomes, while the last peak (circled in blue) indicates 12 repeats of that STR on each homologous chromosome.

- A DNA profile is produced from the STR analysis. DNA fragments from samples collected from the crime scene are compared with the DNA of a suspect or suspects.
- The DNA sample is matched to that of a suspect if the lengths of the particular STRs at all sites are the same.

- If the lengths of 20 STRs from two DNA samples match perfectly, the chance that the two samples are from different people is hundreds of billions to one.

Most recently, DNA profiles have been used to identify key features of an individual's appearance. The presence of alleles for eye, skin and hair colour allows investigators to narrow down the list of suspects. Occasionally the suspect's ancestry can also be determined. This is a developing field of science and the analysis may not yet be reliable.

PA 2

CASE STUDY

Tsar rediscovered using DNA

In July 1918, Tsar Nicholas II of Russia, the Tsarina Alexandra, their five children—Olga, Tatyana, Maria, Anastasia and Alexis (Figure 4.1.14)—three female servants and the royal physician were executed by a Bolshevik firing squad in the town of Ekaterinburg. Historical accounts indicate that two of the children's bodies were burned, although others claim that Anastasia escaped execution. The remaining bodies were thrown into a shallow grave and sulfuric acid was poured over them.

In 1991, two amateur historians, Gely Ryabov and Alexander Avdonin, discovered nine skeletons in a grave near Ekaterinburg. The remains were tested to find out whether they came from the Tsar and his family. DNA extracted from bone tissue samples was amplified by PCR. The first step in the analysis was to identify the sex of the skeletons. This was done using PCR of a gene that is found on the Y chromosome. This indicated that there were two males and seven females.

Using DNA profiling of the bone samples, it was possible to conclude with certainty that five of the skeletons were those of two parents and their three daughters. But these could have been the remains of any family. To establish the identity of the bones, comparison to the DNA from a related person was needed. The evidence came when the DNA profiles of the skeletons were compared with those generated from known relatives of the Tsar (George, brother of Nicholas II, whose remains were exhumed from a crypt in St Petersburg) and Tsarina (Prince Philip, husband of Queen Elizabeth II).

FIGURE 4.1.14 Tsar Nicholas II of Russia with his wife, Tsarina Alexandra, and their five children in 1913

The presence of common bands in the DNA profiles of the five bodies, Prince Philip and George indicated that all of the individuals were relatives. Indeed, it was estimated that the probability that the bones found in Ekaterinburg are the remains of the Tsar and his family is approximately 99 999 out of 100 000.

BIOFILE

Implications and issues related to DNA profiling

Privacy is a contentious issue related to DNA profiling. In Victoria, DNA samples cannot be obtained from a person unless they give their permission. They can, however, be ordered to do so if there is strong evidence that they may have committed the crime under investigation and if the DNA profile could help to confirm or deny their guilt. These DNA samples must be destroyed if the person is not guilty or is not charged. However, in some countries, the DNA may be kept for up to 10 years. This has enabled the identification of criminals who have committed crimes in unsolved cases that occurred before DNA profiling technology was developed. It has also resulted in the exoneration of people who have been wrongly accused.

Some people are in favour of the creation of a 'bank' of DNA samples, provided by everyone in the community, which could be used to solve crimes and perhaps trace the remains of unidentified missing persons. Opponents of a DNA 'bank' argue that there would be potential for these samples to be stolen or used unethically.

CASE STUDY

Freeing the innocent

DNA profiling has been used to incriminate perpetrators, but has also been used to prove the innocence of suspects or the wrongly accused. In 1987, DNA profiling was used in a case for the first time to help free an innocent man and identify the perpetrator.

Alec Jeffreys, who pioneered the technique, was asked to test semen samples found on two 15-year-old victims who were raped and murdered three years apart and compare it to a 17-year-old male, Richard Buckland, who they believed was responsible for both murders. Buckland, who had learning difficulties, had some knowledge of the murder of the second victim, and had later confessed to the second murder. The police had wanted to convict him of the first murder, though he refused to plead guilty to this. Through DNA profiling, Alec Jeffreys confirmed that the two girls were murdered by the same man but it was not Buckland. Buckland was released from custody and another man was convicted of the crimes.

FIGURE 4.1.15 The Idaho Innocence Project uses DNA technology to free people wrongly convicted of crimes.

Greg Hampikian, a biology and criminal justice expert at Boise State University, founded the Idaho Innocence Project. Hampikian works with police and defence lawyers around the world to free innocent people wrongly convicted of crimes by using new DNA profiling technology or exposing errors with existing technology. One such case Hampikian worked on involved an error in police DNA sampling (Figure 4.1.15) that caused two innocent people to be jailed. Hampikian reviewed the DNA sampling procedures and found many errors, including one where an item of clothing containing a DNA sample was not collected until 46 days after the crime, and was passed around to several police investigators before being placed in a different position at the crime scene for photographing, allowing ample time for contamination of the sample. Hampikian also found samples from the same crime scene that were not tested by forensics teams because the quantity of the sample was smaller than what the FBI deemed 'valid for testing'. In fact, as Hampikian discovered, the samples were sufficient in quantity and quality to provide valid DNA evidence.

CASE STUDY

DNA barcodes

A DNA barcode is a short sequence of nucleotides that uniquely identifies a species. It is obtained by using the polymerase chain reaction (PCR) and DNA sequencing (Figure 4.1.16). Sequences are submitted to online databases such as BOLD, the Barcode of Life Database, or other similar databases. There are global DNA barcoding projects underway to catalogue all of life, including bees, butterflies, mosquitoes, fungi, mammals and plants. Scientists and non-scientists alike can access these sequences for research.

To compare and identify species, you need a gene sequence that is present in all organisms but differs slightly between different groups or species. For eukaryotes, the mitochondrial gene *CO1* (cytochrome oxidase subunit 1) is often used. Genes in plant plastids help to further identify plant species. For bacteria, a gene for ribosomal RNA can be used.

Examples of barcoding projects include:

- identifying species and measuring diversity; for example, barcoding organisms of the Great Barrier Reef, CSIRO scientists barcoding Australian fish species
- tracking pathogenic and non-pathogenic bacterial populations
- authenticating food—Is it shark and chips? Is the beef burger really a horse burger?
- monitoring wildlife crime, such as illegal trade in protected and endangered species
- investigating ecology and evolution; for example, seed identification and seed banking of Australian *Acacia* (wattle) genus for conservation and restoration of biodiversity.

FIGURE 4.1.16 Process of obtaining a DNA barcode

WS 7

4.1 Review

SUMMARY

- DNA is often found or extracted in trace amounts and requires amplification to produce a sample that is large enough for scientists to work with.
- DNA amplification uses the polymerase chain reaction (PCR) to rapidly increase identical copies of the target DNA.
- Polymerases are enzymes that catalyse the formation of long-chain molecules (polymers), such as DNA and RNA, by linking smaller molecules (nucleotides).
  - *Taq* polymerase is a heat-resistant DNA polymerase often used in PCR to synthesise multiple copies of the target DNA.
  - RNA polymerases synthesise RNA from DNA during transcription.
- The PCR mixture is added to a test tube and the test tube is placed in a thermocycler. The thermocycler alters the temperature in pre-programmed stages to enable a three-step process to be carried out:
  - denaturation
  - annealing
  - extension.
- Gel electrophoresis is a technique that separates fragments of negatively charged DNA by length.
- DNA profiling compares variable short tandem repeat (STR) regions of the genome for identification of individuals.

KEY QUESTIONS

Knowledge and understanding

1 Which one or more of the following would be a suitable target DNA for PCR?
- **A** a single gene
- **B** a genome
- **C** a short variable region within a chromosome
- **D** a microsatellite
- **E** a whole chromosome

2 **a** Describe the function of a polymerase enzyme in cells and state their function.
b Name two types of polymerase enzyme used in cells and state their function.

3 **a** Identify the function of the enzyme called reverse transcriptase.
b When is reverse transcriptase used in biomolecular techniques?

4 **a** Explain what is meant by amplifying a piece of DNA.
b What do the letters PCR stand for?
c Draw a simple, labelled flow diagram to summarise the key steps in the process of PCR.
d **i** Describe the role of the enzyme *Taq* polymerase.
ii Why does *Taq* polymerase have to be heat resistant?
e If you started a PCR reaction with one DNA molecule, determine how many molecules you would have after:
i 3 cycles of amplification
ii 32 cycles of amplification

5 **a** What is gel electrophoresis?
b Outline how this technique reveals data about the DNA fragments being tested.

6 A DNA ladder is used in gel electrophoresis. Explain the purpose of a DNA ladder.

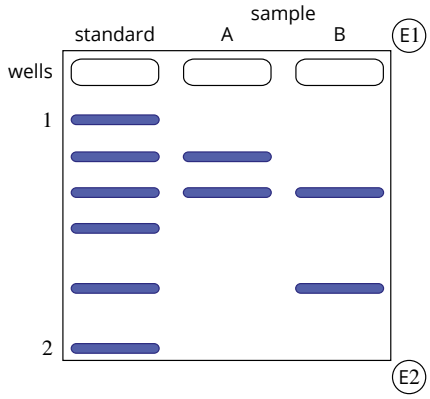
7 **a** Explain how DNA profiling can be used to help determine the guilt or innocence of a murder suspect found with bloodstains on their clothes.
b Explain why a sample may need to be tested in more than one laboratory.

8 Explain if DNA profiling can discriminate between:
a a twin brother and sister
b identical twins.

Analysis

9 You have a single-stranded RNA molecule with the sequence 3'-AAUUGCGCA-5'. If you place it in a test tube with nucleotides and reverse transcriptase, identify the sequence that will be made on the complementary strand.

10 Gel electrophoresis was conducted with a DNA ladder (standard) and two DNA samples (A and B). The diagram illustrates the resulting gel.



a Explain the difference between the DNA fragments labelled 1 and 2.

b E1 and E2 represent the electrodes. Which one is the positive electrode?

c Explain why the DNA migrates through the gel from E1 to E2 when an electric current is applied.

d How many DNA fragments are in:

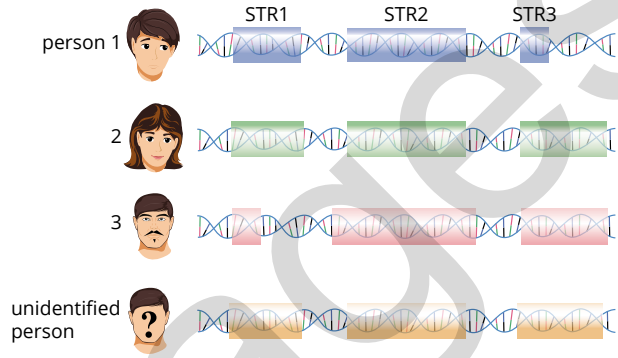
i the standard (DNA ladder)

ii sample A

iii sample B

e You have been given information about the DNA ladder. The sizes of the fragments are 600, 500, 400, 300, 200 and 100 bp. What are the sizes of the fragments in samples A and B?

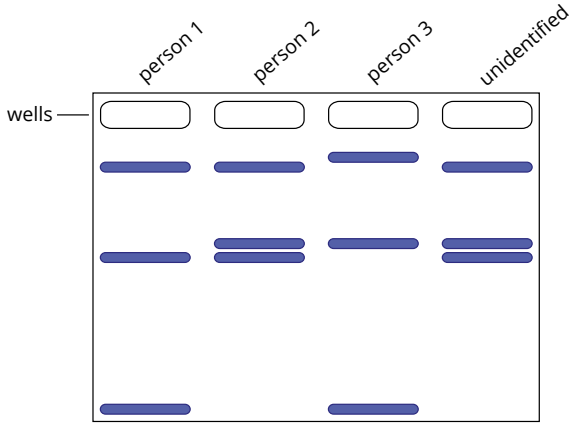
11 The following diagram represents STR regions of DNA used to identify a person who died in a natural disaster. Three people who were looking for a missing sibling submitted DNA for comparison.



a What is an STR?

b List the steps involved in creating a DNA profile using STR analysis.

c The following illustration represents the DNA profile from the STR analysis. Which person is most likely to be the sibling of the unidentified person? Explain your answer.



4.2 Bacterial transformation

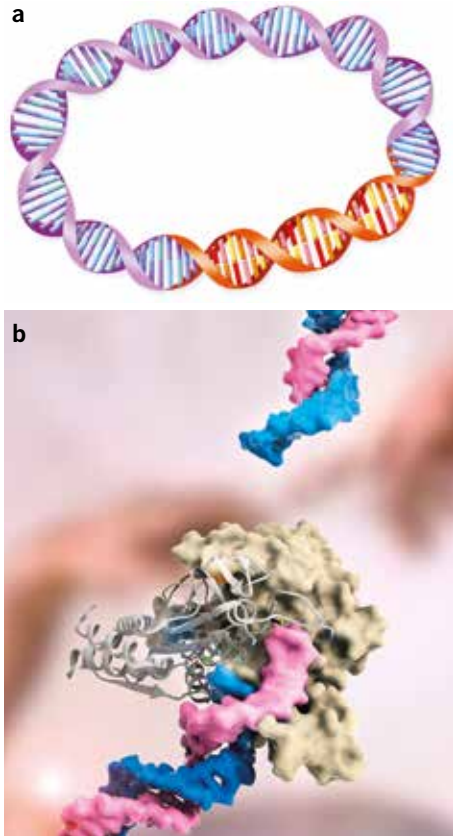


FIGURE 4.2.1 (a) Illustration representing a recombinant plasmid, showing the gene of interest (orange) combined with the plasmid DNA (purple). (b) Computer model showing the restriction enzyme EcoRI (grey) cutting a DNA strand (blue and pink)

Bacteriophages are viruses that infect bacteria.

The base-pairing ability of sticky ends allows DNA from very different species to ligate (join), forming recombinant DNA molecules.

DNA is universal to life on Earth, with the same structured molecule found in all cells. This means DNA from different species can be combined.

Plasmids are small, circular DNA molecules found in bacterial cells. They are often used as **vectors** (carriers) when scientists move target DNA from one organism to another. Genes can be inserted into plasmids, which can then be incorporated into bacterial cells in a process known as **bacterial transformation**. The inserted gene can then be replicated through the self-replicating properties of the plasmid inside the bacterial cells, with the gene expressing the protein for which it codes. In this way target proteins can be mass-produced, such as the human insulin protein for use by diabetics.

In this section, you will learn how restriction enzymes (endonucleases) and ligases are used to create recombinant DNA (Figure 4.2.1), using plasmids as vectors. You will also learn how these plasmids are then incorporated into bacterial cells for replication.

RESTRICTION ENZYMES (ENDONUCLEASES)

DNA molecules are far too long for biologists to work with in their entirety. The discovery and isolation of **restriction enzymes**, also known as **endonucleases**, has enabled scientists to cut DNA into smaller, more usable fragments and isolate particular regions of interest, such as a single gene.

Restriction enzymes are a large group of enzymes that occur naturally in bacteria. They form part of a bacterial cell's defence system, targeting foreign DNA that may enter the cell, such as the DNA of **bacteriophages**. The restriction enzymes cut up foreign DNA into smaller fragments, destroying it and preventing it from replicating. Bacterial cells use a blocking process (methylation) to prevent the restriction enzymes from binding and cutting their own DNA.

Each restriction enzyme targets a specific sequence of nucleotides, usually four to six base pairs in length. This sequence is called a **recognition site**. Every time a restriction enzyme passes its recognition site, it breaks the phosphodiester backbone once on each DNA strand. As a result, the DNA molecule is cut up into fragments of different lengths.

Types of restriction enzymes

There are two types of restriction enzyme, which cut DNA differently:

- sticky-end restriction enzymes
- blunt-end restriction enzymes.

Sticky-end restriction enzymes

Sticky-end restriction enzymes leave DNA fragments with overhanging ends. They cut the DNA backbone at a different location on each strand within the recognition site (Figure 4.2.2). This results in a staggered cut, leaving two fragments with exposed bases or 'sticky ends'. The exposed bases are then able to form complementary base pairs through hydrogen bonding with nucleotides of other DNA molecules that have complementary sticky ends.

EcoRI is an example of a sticky-end restriction enzyme extracted from *E. coli*. It cuts the recognition site GAATTC between the G and A nucleotides on each strand. The sequence on one strand is GAATTC and on the complementary strand it is CTTAAG, the same sequence when read backwards. This is called a **palindromic sequence**. EcoRI cuts at the different G and A location on each strand, so sticky end fragments are produced.

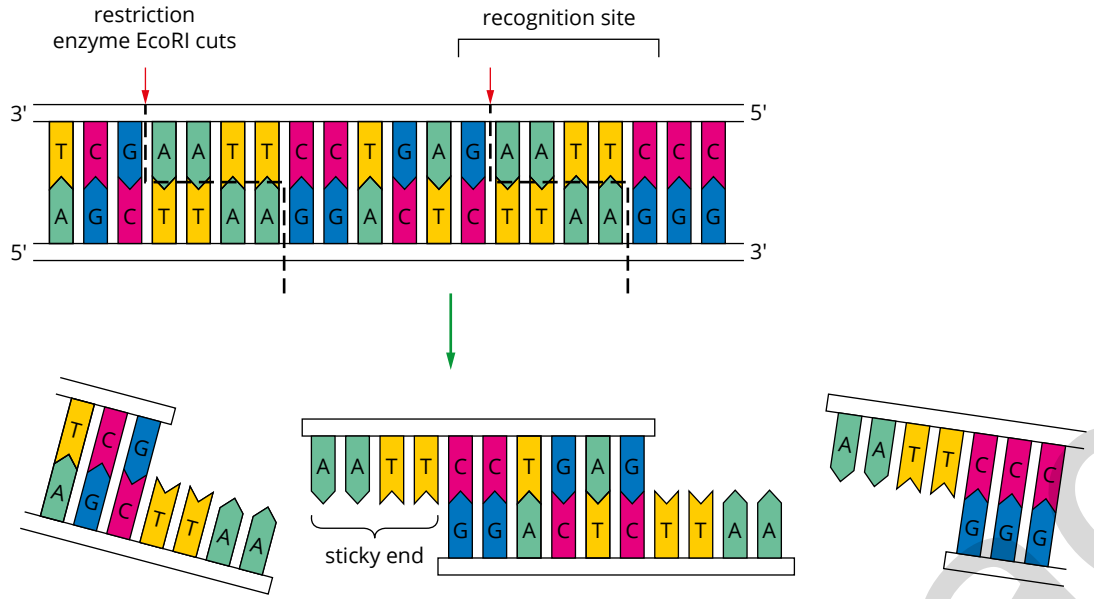


FIGURE 4.2.2 The sticky-end restriction enzyme EcoRI cuts the DNA between the G and the A of its specific recognition site, GAATTC, creating fragments of DNA with sticky ends.

Blunt-end restriction enzymes

Blunt-end restriction enzymes leave clean-cut ends by cutting the sugar–phosphate backbone on both strands of the DNA molecule at the same location within the recognition site (Figure 4.2.3).

HaeIII is an example of a blunt-end restriction enzyme, which is extracted from the bacterium *Haemophilus aegyptius*. It cuts the recognition site GGCC between the G and C nucleotides. These two bases are in the exact same location on either strand of the DNA molecule, resulting in a straight cut through both stands that leaves two fragments with blunt ends.

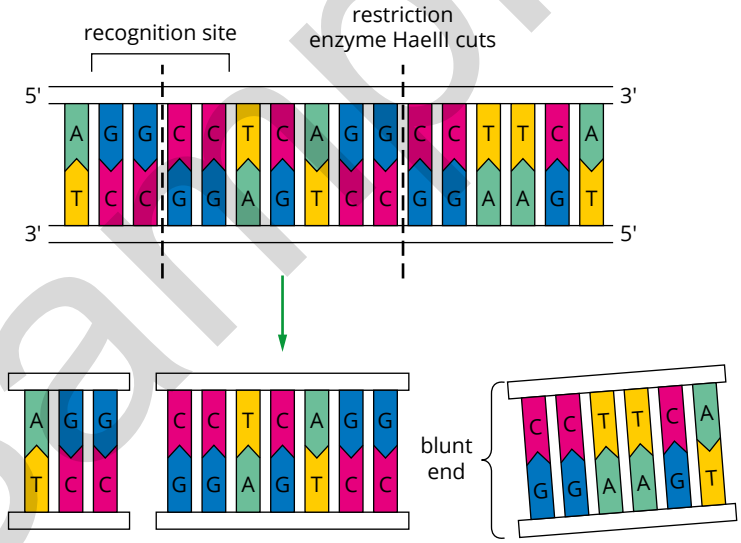


FIGURE 4.2.3 The blunt-end restriction enzyme HaeIII cuts the DNA between the G and the C of its specific recognition site, GGCC, creating fragments of DNA with blunt ends.

BIOFILE

Naming restriction enzymes

The first three letters of the name of a restriction enzyme identify the bacterial species from which they were isolated. The fourth letter refers to the particular strain of bacteria. A roman numeral is also included if more than one restriction enzyme has been isolated from this bacterial strain. For example, the restriction enzyme EcoRI comes from *E. coli* strain RY13, and was the first restriction enzyme isolated from that strain of *E. coli*.

Identifying polymorphisms and mutations

Small variations in DNA sequences (called polymorphisms) occur within a population. To be considered a polymorphism rather than a mutation, the least common allele has to have a frequency of 1% or more in a population.

A mutation may change an allele to a new variant, different from the other alleles in a population. For example, individuals with sickle-cell anaemia have a mutation in the β-globin gene. β-globin is a polypeptide component of haemoglobin, the oxygen carrying protein in your red blood cells.

A common mutation occurring in the β-globin gene is a single base change from A to T, resulting in a single amino acid substitution in the protein—a missense mutation (Figure 4.2.4). This changes the structure of the β-globin polypeptide, causing haemoglobin to clump, and reduces the amount of oxygen carried. The red blood cells take on a sickle shape that tends to clog and rupture the capillaries.

Haemoglobin is a protein made from four polypeptide chains, two alpha (α) globin and two beta (β) globin polypeptides, each with an iron-containing haem group that binds oxygen.

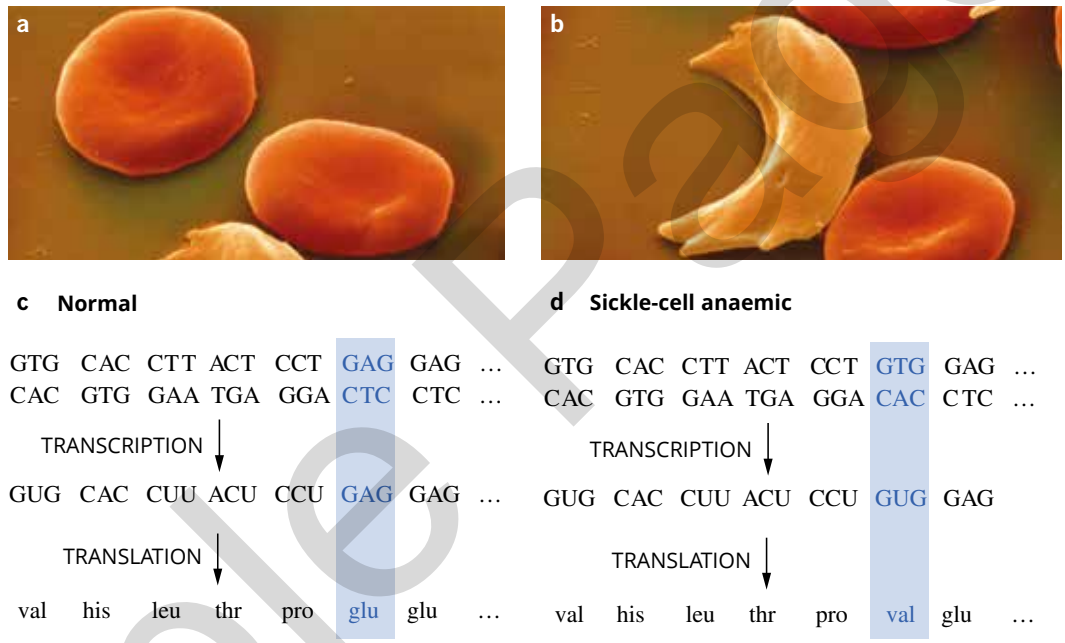


FIGURE 4.2.4 Viewed under the electron microscope there are obvious differences between the (a) normal and (b) sickle-shaped red blood cell. Under each image, expression of the (c) normal and (d) mutated genetic code is compared.

Molecular tools are used to identify individuals carrying the mutation that causes sickle-shaped red blood cells. By chance, the mutation occurs at a restriction enzyme site. The base change eliminates the recognition site for MstII (Figure 4.2.5a). To detect the mutation, DNA is extracted from individuals, the region of DNA containing the recognition site for MstII is amplified by PCR, and then the PCR products are incubated with the MstII restriction enzyme. The normal allele will be cut at the MstII recognition site. The mutant allele will not be cut. The difference in the size of the DNA fragments is identified by gel electrophoresis (Figure 4.2.5b).

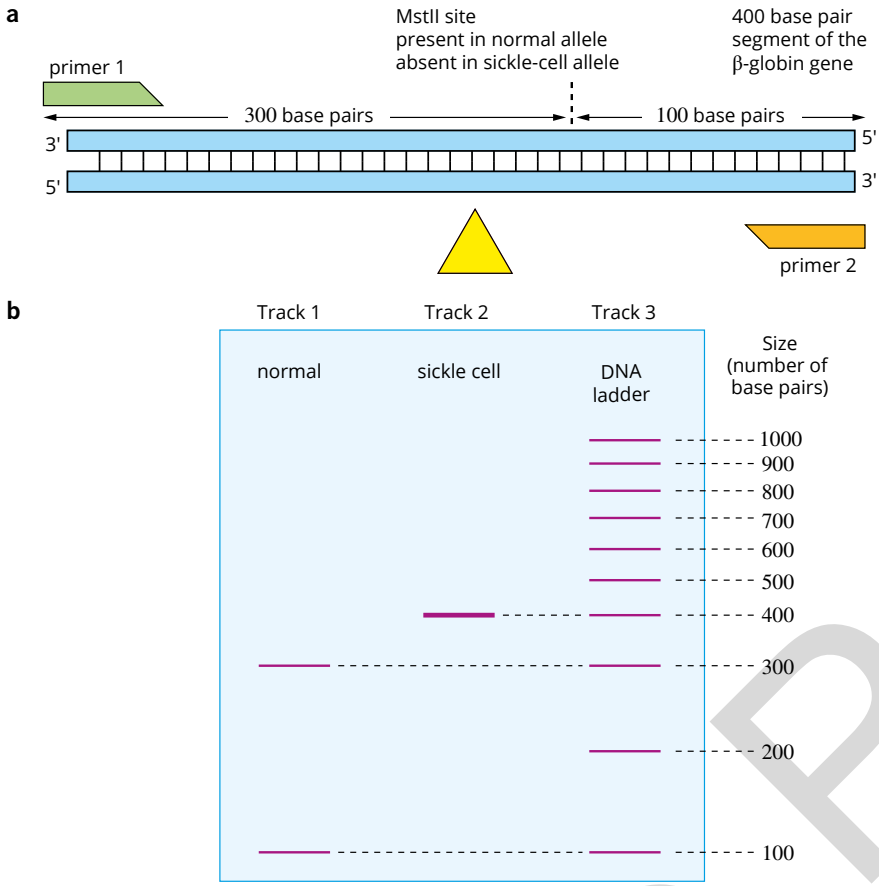


FIGURE 4.2.5 (a) A portion of the β-globin gene with primers either side of a 400 bp region. An MstII recognition site is present in the normal allele, resulting in fragments of 300 and 100 bp when the PCR product is cut with the MstII restriction enzyme. The MstII recognition site is missing in the sickle-cell allele so the PCR product is not cut. (b) Diagram of an electrophoresis gel showing DNA fragments from two individuals. The DNA ladder in Track 3 is used to determine the sizes of the Track 1 and 2 bands and identify which individual has the normal allele (Track 1) and which individual has the sickle-cell allele (Track 2).

LIGASES

Ligases are a group of enzymes that join fragments of DNA or RNA in a process called **ligation**. There are two different groups of ligases:

- **DNA ligases**
- **RNA ligases**

The role of DNA ligase in a cell is to join segments of newly replicated DNA and repair breaks in DNA molecules.

Assuming they were cut with the same restriction enzymes, DNA ligase can join DNA fragments extracted from different organisms or different species, as DNA has a universally consistent molecular structure. Under laboratory conditions, depending on the characteristics of the DNA ends to be joined (sticky or blunt ends), the conditions of the ligation reaction (e.g. incubation time and temperature) need to be adjusted to ensure efficient DNA ligation is achieved.

Ligases join two DNA fragments and create a phosphodiester bond between them.

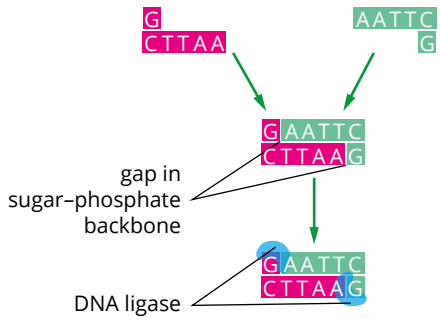


FIGURE 4.2.6 Two sticky-end DNA fragments come together by complementary base pairing, and then DNA ligase (shown as blue circles) permanently links the sugar–phosphate backbone.

Ligation of sticky-end fragments

Ligation to join sticky-end fragments is specific because the exposed bases of sticky-end fragments first bind by complementary base pairing. Complementary bases are attracted by weak hydrogen bonds that hold them together. After this, the ligase joins the fragments (Figure 4.2.6) by creating a phosphodiester bond between the 3' OH end and 5' phosphate end of the adjoining nucleotides. This technique makes recombinant DNA and is used in processes such as **gene cloning**.

Ligation of blunt-end fragments

Unlike ligation of sticky-end fragments, ligation of blunt-end fragments is random. Any two fragments can join if they come in contact and the DNA ligase joins them. For this reason, blunt end fragments are more difficult to use in DNA manipulation processes that require the joining of specific fragments. However, sometimes blunt ends are unavoidable. For instance, a blunt-end enzyme might be the only type available to cut out the target gene without damaging the gene itself. Using DNA ligase, scientists are able to attach short, linking DNA fragments onto blunt-end DNA to create sticky ends.

RECOMBINANT DNA

When DNA from two different species is joined together, the resulting molecule is called **recombinant DNA**. Scientists create recombinant DNA to clone (make multiple copies of) a particular gene. Following this step, they may also produce large quantities of the protein expressed by the cloned gene. For example, an insulin-coding gene may be cloned and then incorporated into bacteria to produce large quantities of the protein insulin to be used as treatment for diabetes.

Insulin from animals

Insulin was previously extracted from the pancreas of other animals, such as pigs and cattle, for the treatment of type 1 diabetes. This was an expensive and time-consuming method that also involved the risk of an allergic reaction to the foreign molecule and potential for contracting diseases. Porcine (pig) and bovine (cattle) insulin are similar, but not identical, to human insulin. Their biological activity is not as effective as human insulin, so it is preferable to use the human hormone. Recombinant human insulin became available for treatment in the 1980s (Figure 4.2.7).

FIGURE 4.2.7 Man giving himself an injection of insulin

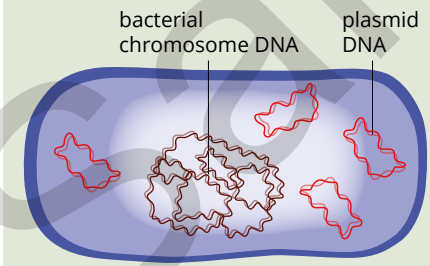
A clone is a genetically identical copy of a gene, cell or organism.

BIOFILE

Plasmids

In Chapter 3 you learnt about the structure of DNA and chromosomes. In eukaryotic cells, DNA forms long, linear chromosomal strands containing thousands of genes. In most prokaryotic cells, DNA forms a single double-stranded circular chromosome.

In addition to the circular chromosome, bacterial cells also contain small circular pieces of double-stranded DNA called plasmids. Plasmids replicate independently of the main chromosome. As bacterial cells do not have a nucleus, both chromosomal and plasmid DNA are located in the cytosol.



Bacterial cells contain a single, circular DNA chromosome (brown) and small rings of double-stranded DNA called plasmids (red).

Using plasmids as vectors

When scientists create recombinant human insulin, they often use a bacterial plasmid as the vector. They insert target DNA, in this case the gene required to produce insulin, the *INS* gene, into the plasmid, producing a **recombinant plasmid**. The plasmid is then placed in a bacterial cell, where the self-replicating system of the plasmid and cell replicates the plasmid genes. Each bacterial cell containing the plasmid will express the protein products of the plasmid genes, including those of the introduced target DNA (in this case, insulin).

Plasmids are used as vectors when creating recombinant DNA for the following reasons:

- Their small size makes them easy to manipulate in a laboratory.
- Plasmids carry a range of restriction enzyme sites. A plasmid containing the appropriate recognition sites can be chosen to suit your needs. For example, if the gene of interest was cut from a chromosome with EcoRI, you would use a plasmid with a single EcoRI site.
- Recombinant plasmids self-replicate independently once they are placed inside a host bacterial cell and at a faster rate than their bacterial host's chromosomal DNA. This is of vital importance in the efficient manufacture of large quantities of insulin and other proteins.

To enable the identification of cells that have incorporated the recombinant plasmid, the plasmids used as vectors must have particular characteristics including:

- an antibiotic-resistance gene
- a **reporter gene** that can be easily identified—such as a gene that produces coloured or fluorescent proteins.

One example is a plasmid containing the ***lacZ* gene** (Figure 4.2.8). Restriction enzyme sites for inserting the gene of interest are located within the *lacZ* gene. If the gene insertion is successful it will disrupt the *lacZ* gene. The use of this process in bacterial selection will be discussed later in this section. The plasmid also contains the antibiotic resistance gene amp^R, which encodes resistance to ampicillin.

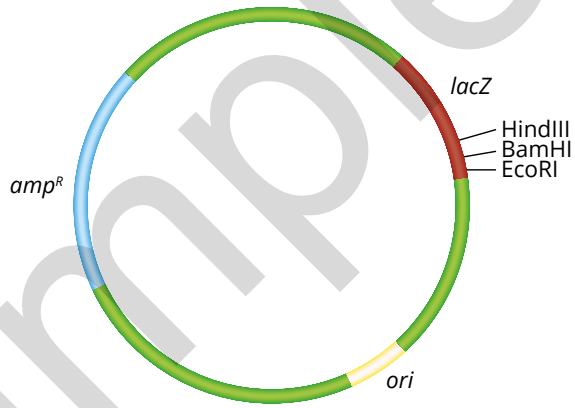


FIGURE 4.2.8 An example of a plasmid used for recombinant DNA and bacterial transformation showing the gene for ampicillin resistance (amp^R), the origin of replication (*ori*) and the *lacZ* gene. Sites for three restriction enzymes lie within the *lacZ* gene.

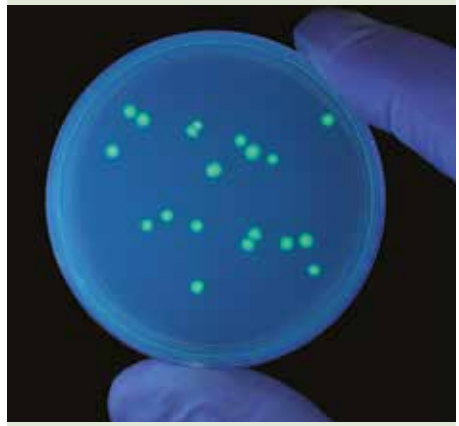
A recombinant plasmid is a plasmid containing a foreign gene that has been inserted using restriction enzymes and DNA ligase.

BIOFILE

Green fluorescent protein (GFP)

Green fluorescent protein (GFP) is a protein from jellyfish. The *GFP* gene can be inserted into plasmids to be used in bacterial transformation. The transformed bacterial cells express the green fluorescent protein and fluoresce green under UV light.

Reporter genes are attached to the gene under investigation and provide visual evidence when the gene of interest is expressed, since they are expressed alongside it. The *GFP* gene is a reporter gene. Other reporter genes express luciferase, an enzyme from fireflies that produces a yellow fluorescent protein, and a red fluorescent protein from a coral.



Transformed bacterial colonies can be identified as they contain the *GFP* gene and fluoresce under UV light.

A reporter gene is a gene that allows detection of gene expression in genetic engineering, such as genes for *lacZ* and fluorescent proteins.

Creating recombinant DNA

The process of creating a recombinant plasmid is outlined below and in Figure 4.2.9:

1. The target DNA, in this case the gene responsible for producing insulin (the *INS* gene), is cut out using a sticky-end restriction enzyme and then isolated.
2. The bacterial plasmid is cut by the same restriction enzyme. The plasmid and the target DNA now have the same sticky ends with exposed bases that are complementary to each other.
3. The *INS* gene and plasmids are placed together. Some plasmids will simply close back up (known as non-recombinant plasmids), while other plasmids will incorporate the *INS* gene by complementary base pairing (known as recombinant plasmids). Reporter genes are necessary to distinguish recombinant plasmids from non-recombinant plasmids.
4. DNA ligase is added to rejoin the sugar–phosphate backbone of the DNA.

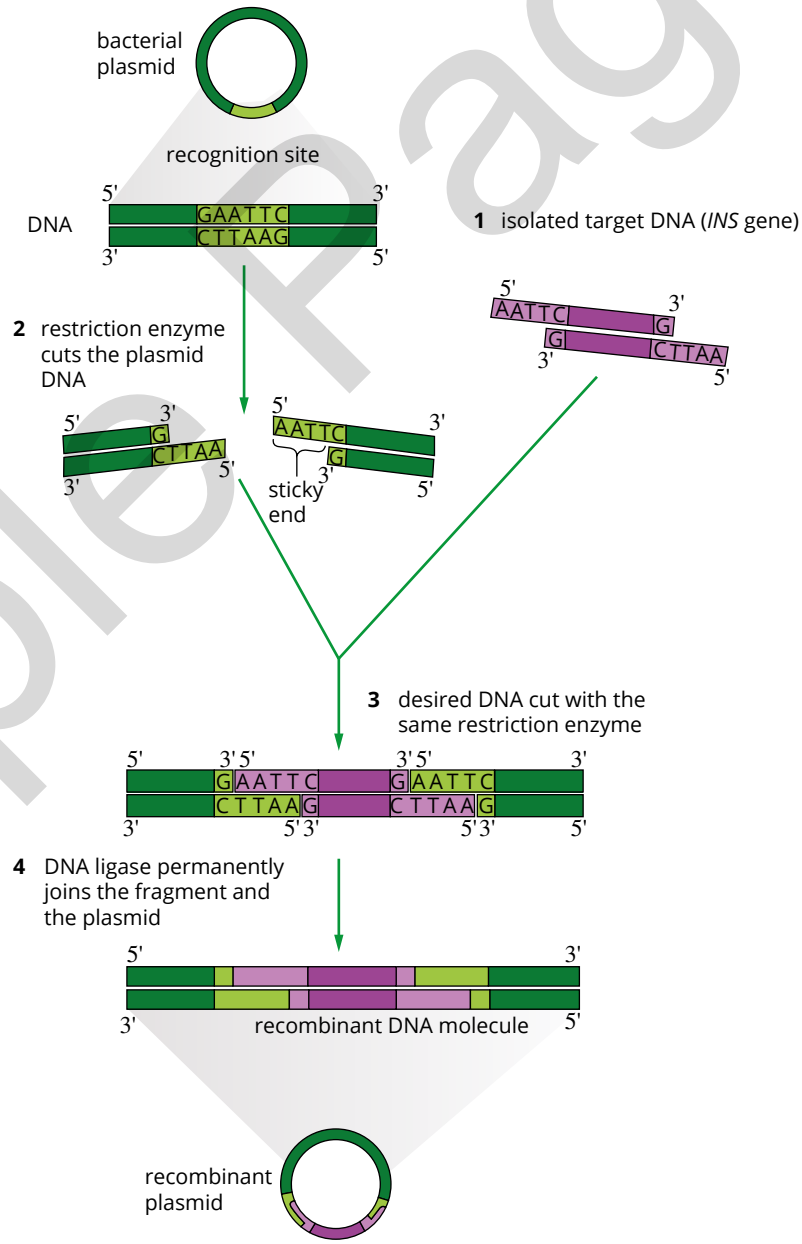


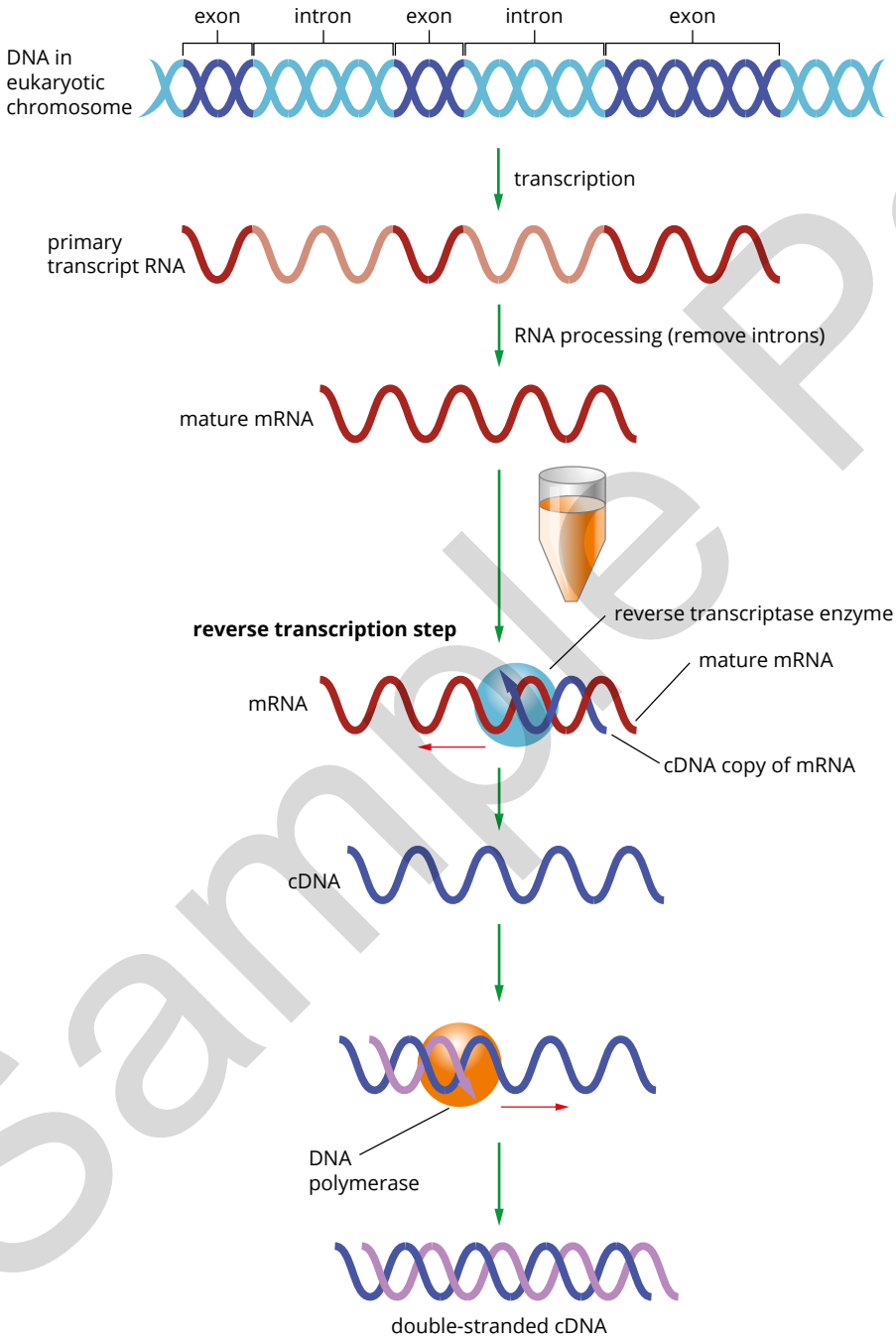
FIGURE 4.2.9 A recombinant plasmid is created by joining a target DNA fragment (the *INS* gene) and a plasmid that have both been cut with the same sticky-end restriction enzyme. They are joined using DNA ligase.

Complementary DNA

To produce a eukaryotic protein, such as insulin, in a prokaryotic cell (i.e. bacterial cell), complementary DNA (cDNA) is used as the target DNA. cDNA is DNA that has been copied from mature mRNA and contains only exons. cDNA is synthesised using the reverse transcriptase enzyme.

Reverse transcriptase

In Section 4.1, you learnt that reverse transcriptase is an enzyme with the ability to make cDNA from mRNA (Figure 4.2.10). This is useful because mature mRNA has already had the introns spliced out. Prokaryotic cells are unable to splice out introns. Reverse transcriptase allows the synthesis of DNA from mature mRNA in a test tube (*in vitro*).



Exons are the coding regions of DNA, while introns are non-coding regions.

FIGURE 4.2.10 Process of creating cDNA from mRNA using a reverse transcriptase enzyme. Once the strand of cDNA is produced, DNA polymerase is used to make the cDNA double stranded.

When cDNA is inserted into a plasmid, which in turn is then incorporated into a bacterial cell, the protein encoded by this cDNA will be expressed (Figure 4.2.11). This method is now commonly used to produce vast quantities of therapeutic proteins including human insulin, growth hormone and cytokines.

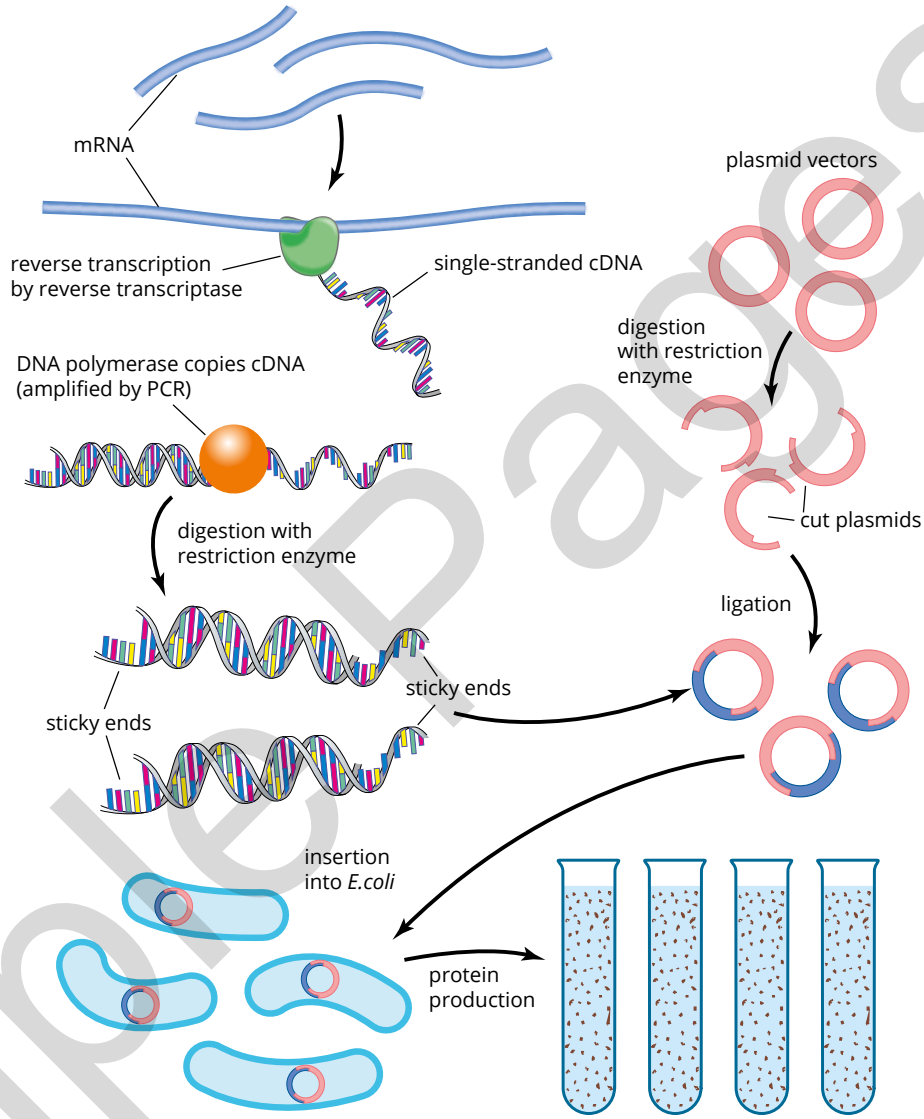


FIGURE 4.2.11 cDNA is used as target DNA to create a eukaryotic protein product within a bacterial cell. The cDNA is treated to create sticky ends so that it can be inserted into a plasmid, creating a recombinant plasmid. This plasmid is then incorporated into a bacterial cell, which expresses the protein product.

Regulatory genes in recombinant DNA

Regulatory genes may be included in plasmids for the purpose of controlling the expression of the target gene that is inserted into the plasmid. The regulatory gene is turned on by an **inducer** molecule, for example a sugar such as lactose or arabinose, or by metal ions such as iron, copper or zinc. Once the regulator is transcribed, the target gene can be transcribed and translated. Inducers are important in regulating gene expression, particularly when the aim of the recombinant DNA technology is protein production, and when a gene is being expressed and studied in plant and animal models.

TRANSFORMING BACTERIAL CELLS

Cells that have had foreign DNA incorporated into them have undergone bacterial transformation and are said to be 'transformed'. For example, when a foreign plasmid is incorporated into a bacterial cell, the bacterial cell is then 'transformed' because it can express a new gene and therefore has a new characteristic. This process is also known as **genetic transformation** and it can be both natural and artificial.

Natural transformation of bacterial cells

Bacterial cells exchange genetic material naturally through several methods. This is how an organism that reproduces asexually (e.g. by binary fission) can evolve and spread antibiotic resistance. **Bacterial competence** refers to the ability of a bacterial cell to alter its genome by taking in DNA from other cells or the environment. Antibiotic resistance genes can be taken in from other bacterial cells. Since antibiotic resistance in disease-causing bacteria is of urgent interest to scientists and the general public, bacterial competence is an important and growing area of research.

Artificial transformation of bacterial cells

Two methods of artificial bacterial transformation are used: heat shock and electroporation.

- Heat shock involves placing bacterial cells and a mixture of recombinant and non-recombinant plasmids in an ice-cold solution containing calcium ions and then rapidly increasing the temperature to disrupt the plasma membrane of the bacterial cells. The plasmids can then penetrate the plasma membrane and enter the bacteria.
- In electroporation, the bacterial cells and a mixture of recombinant and non-recombinant plasmids are subjected to an electrical current that alters the plasma membrane. Again, the plasmids are then able to enter the bacteria.

Very few of the bacterial cells will be transformed with recombinant plasmids. Some will take up the non-recombinant plasmids (plasmids without the target DNA) and others will not be transformed at all or will die from the heat shock or electroporation treatment (Figure 4.2.12). Even a small number of successful transformations is enough to grow a culture or population from.

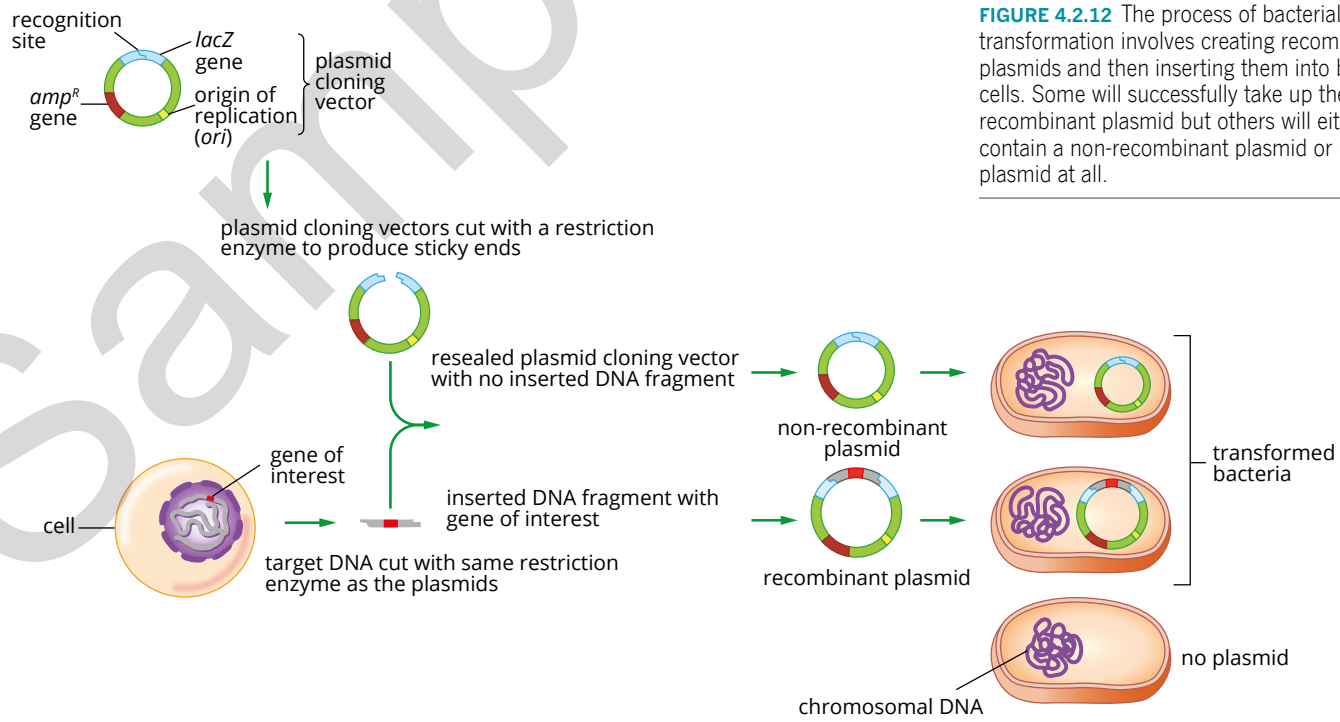


FIGURE 4.2.12 The process of bacterial transformation involves creating recombinant plasmids and then inserting them into bacterial cells. Some will successfully take up the recombinant plasmid but others will either contain a non-recombinant plasmid or no plasmid at all.

Bacteria require nutrients, such as nitrogen, carbohydrates and salt, to grow. Agar is the solid medium on which bacteria are grown (or cultured). The combination of agar and nutrients is called nutrient agar.

BIOFILE

The *lacZ* gene and X-gal

The *lacZ* gene is one of the genes in the bacterial *lac* operon. It codes for the enzyme β-galactosidase, which breaks down lactose into glucose and galactose. The *lacZ* gene is removed from the *lac* operon and inserted into plasmids to act as a reporter gene in recombinant DNA technology.

X-gal is a colourless synthetic compound with a very similar structure to lactose, so it fits in the active site of β-galactosidase and is broken down, releasing a blue reaction product. Using agar plates containing X-gal is an easy way to see whether cells produce β-galactosidase.

Growth of transformed bacteria on agar plates containing ampicillin and X-gal. Bacteria with the non-recombinant plasmid appear blue because the *lacZ* gene is expressed. Bacteria with the recombinant plasmid appear white.

WS 6

Selection and screening of transformed bacteria

When determining which bacterial cells have been transformed with recombinant plasmids containing target DNA, the characteristics of the plasmid vectors described on page 153 become important. Recall from the earlier example (Figure 4.2.8 on page 153) that the plasmid vector contains other genes, including a gene for antibiotic resistance (in this example, ampicillin resistance) and a reporter gene that results in a particular phenotype, such as a coloured product.

Selection of transformed bacteria

To determine which of the bacterial cells have been transformed with the gene for antibiotic resistance, the bacteria are grown on nutrient agar plates that contain an antibiotic (in this case, ampicillin) and are incubated at 37°C, the optimum temperature for the bacteria to reproduce and form colonies. The only bacteria to survive will be those that have taken up the plasmid, whether it is a recombinant or non-recombinant plasmid. These bacteria have the ampicillin resistance (amp^R) gene. All other bacteria will be killed.

Screening for bacteria transformed with recombinant plasmids

In this example, the plasmid also carries the *lacZ* gene, which codes for an enzyme that breaks down an indicator called X-gal, resulting in a blue product. Bacteria carrying the non-recombinant plasmid with an intact and functioning *lacZ* gene produce blue colonies on agar plates. If the target DNA has been successfully inserted within the *lacZ* gene, expression is disrupted and the enzyme coded by this reporter gene is not produced. Therefore, bacteria transformed with recombinant plasmids appear as white colonies (Figure 4.2.13).

Bacteria transformed with the recombinant plasmids are then taken from the agar plate and cultured with nutrients in order for them to replicate and produce the protein (e.g. insulin) encoded by the target DNA.

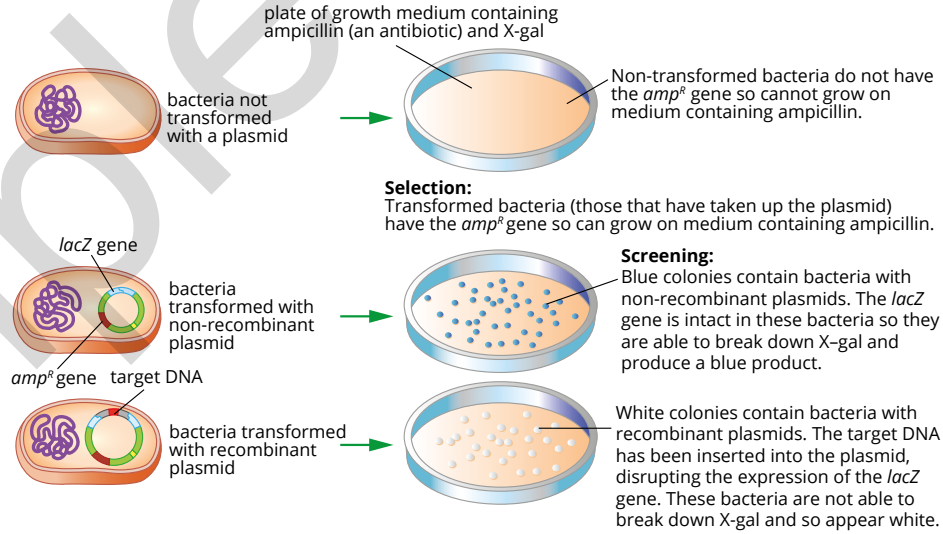


FIGURE 4.2.13 Selection and screening of bacterial cells to identify which cells have been transformed (contain a plasmid), and then which colonies contain recombinant plasmids with the target DNA and which contain non-recombinant plasmids

CASE STUDY ANALYSIS

Recombinant human erythropoietin

Red blood cell production is essential for maintaining oxygen homeostasis. A drop in oxygen supply to tissues (hypoxia) normally triggers the release of the protein erythropoietin (also known as EPO) from the kidneys. EPO promotes red blood cell production in the bone marrow to restore the oxygen-carrying capacity of blood and the delivery of oxygen to tissues (Figure 4.2.14). In chronic kidney disease, not enough EPO is made by the kidneys, resulting in low red blood cell counts and anaemia.

Recombinant human EPO for the medical treatment of this disease is produced in cultured mammalian cells. A copy of the human *EPO* gene is inserted into a plasmid which is introduced into mammalian host cells. EPO is a glycoprotein and must have the correct carbohydrates attached to the protein chain to function properly. Bacteria cannot do this, so mammalian cells must be used for making the recombinant protein (Figure 4.2.15). Recombinant EPO has also been developed for veterinary use, such as recombinant feline EPO for cats with chronic kidney disease.

Because EPO promotes red blood cell production and oxygen-carrying capacity, it has been used by athletes seeking an advantage. EPO has been at the centre of sports doping scandals in recent years, particularly in endurance sports such as cycling, long-distance running and triathlon. It has also been used in horse racing. The World Anti-Doping Agency (WADA) works with drug testing laboratories to develop and validate tests that can distinguish between EPO produced naturally in the athlete and pharmaceutical EPO.

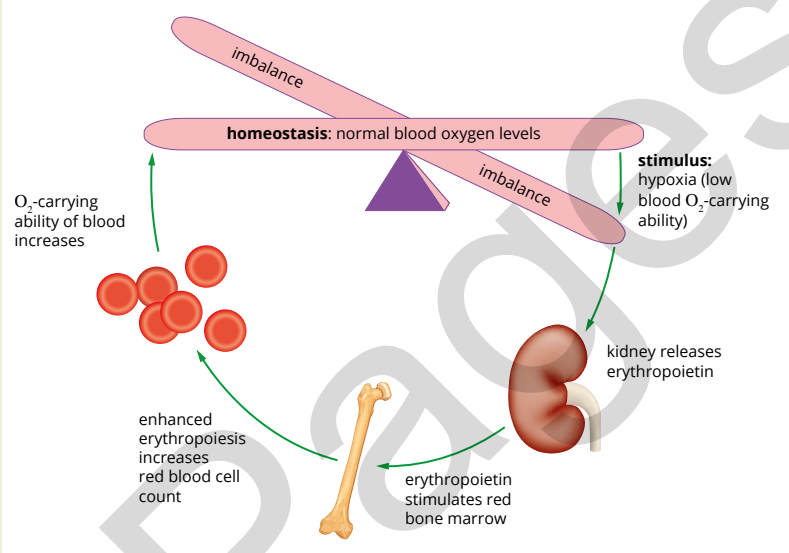


FIGURE 4.2.14 Erythropoietin is a protein released mainly by the kidney to maintain oxygen homeostasis. It promotes red blood cell production (erythropoiesis) in the bone marrow.

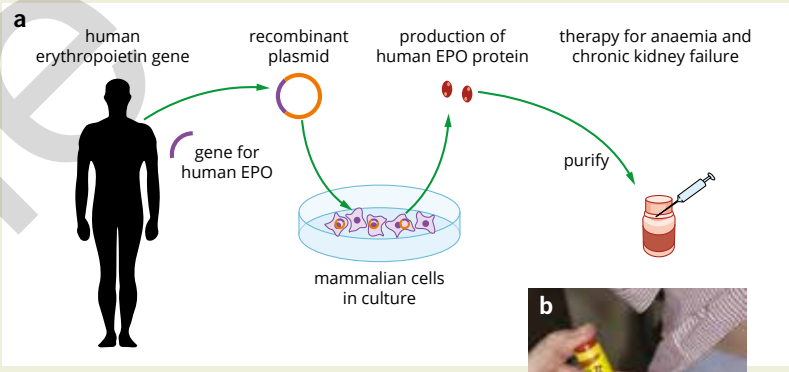


FIGURE 4.2.15 (a) Recombinant human erythropoietin is produced in cultured mammalian cells. It is used to treat anaemia and chronic kidney disease. (b) The brand Aranesp®, for example, is used to treat severe anaemia caused by chemotherapy and chronic kidney failure.

Analysis

1. Referring to Figure 4.2.14, how does normal kidney function compare with kidney function in chronic kidney disease, with respect to erythropoietin and oxygen levels?
2. Examine the text and identify the advantage of using cultured mammalian cells as opposed to cultured bacterial cells in the production of recombinant human erythropoietin.
3. Using Figure 4.2.14 as a reference, deduce what the effect of kidney disease might be on the body.

PROTEIN PRODUCTS OF RECOMBINANT DNA

Recombinant proteins are produced by introducing recombinant DNA into bacteria or eukaryotic cells and allowing them to synthesise the protein. The main types of proteins produced by this technology are hormones, cytokines, enzymes and vaccines for human therapeutic purposes. Therapeutic examples include epidermal growth factor used in the treatment of burns to improve the survival of skin grafts, interleukin-2 used in cancer treatment, antibodies for immunotherapy, vaccines against a number of viruses and to make insulin for the treatment of diabetes. This is much safer and more effective than using proteins purified from other organisms, such as insulin from pigs and growth hormone from human pituitary glands, as was done in the past. Industrial examples include enzymes such as amylase, lipase, protease and cellulase used in food processing, the textile industry and as detergent additives.

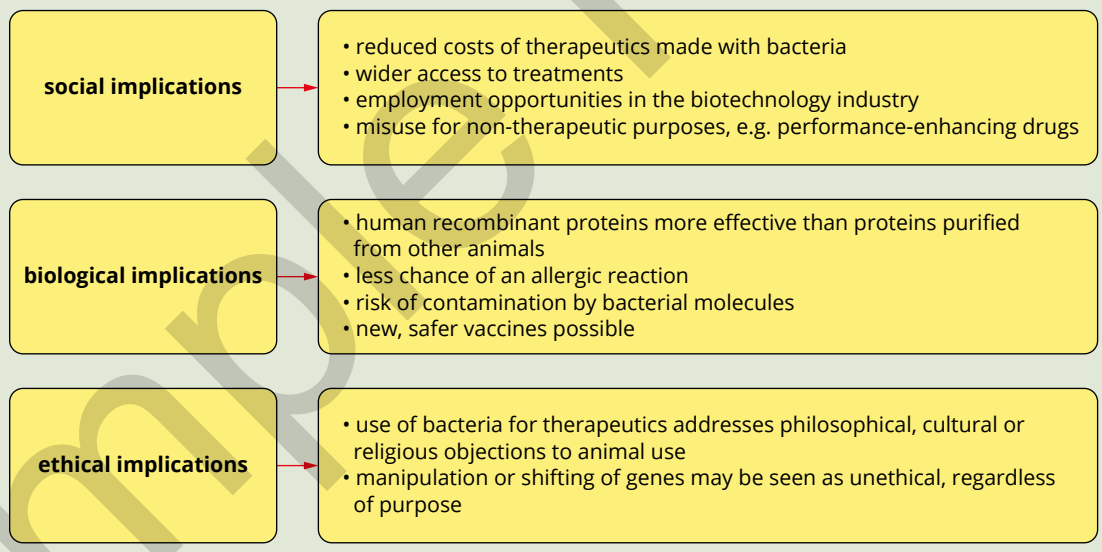
WS 8

BIOFILE

Issues and implications surrounding recombinant DNA technologies

All technologies impact on people's lives and social structures. Social implications may affect people's financial position, lifestyle and reproductive decisions. Biological implications relate to the organisms used in and affected by the technology, its safety and short-term or lasting changes to human biology. Philosophical, moral and religious issues are also part of the impacts of biotechnologies.

Some of the implications, both negative and positive, of gene technologies for making therapeutic recombinant proteins are listed in the figure below.



Implications of recombinant DNA technology and recombinant protein production

CASE STUDY

Natural transformers

Bacterial transformation is used in crop biotechnology to transfer genes into plants and thus introduce a desirable trait. Plants transformed in this way are called transgenic. The bacterium *Agrobacterium tumefaciens* is commonly used as it has the ability to infect plant tissue and incorporate specific parts of its plasmid into the host plant's DNA.

The study and use of *A. tumefaciens* has a long history. This common soil bacterium was identified as the cause of the tumours of crown gall disease over 100 years ago. A plasmid that can move from the bacterium into plant cells was later identified as the key agent. This is a natural transformation of plant cells by a bacterial plasmid. The plasmid in *A. tumefaciens* is called the Ti or tumour-inducing plasmid (Figure 4.2.16). It has genes that direct the movement of the plasmid from bacterium into plant cells and for insertion into the chromosomes of the plant. It also has genes for growth-promoting plant hormones that cause the tumours.

This plasmid is used for genetic engineering. The genes that cause tumour growth are removed, while the genes that ensure transfer from bacterium to plant cell remain. A foreign gene inserted into this plasmid can then be transferred into plant cells. (Figure 4.2.17a). *A. tumefaciens* inserts a section of its plasmid (containing the desired gene) into the plant cell, which is then incorporated into the plant cell DNA. This plant tissue is cultured (Figure 4.2.17b) to create a transgenic plant that expresses the foreign gene. Genes for traits such as insect resistance, herbicide resistance, and tolerance to salinity and frost are all being used in crop biotechnology.

In Australia, transgenic cotton is created using bacterial transformation to confer insect resistance. You will learn more about the production of transgenic plants such as wheat in Section 4.3.

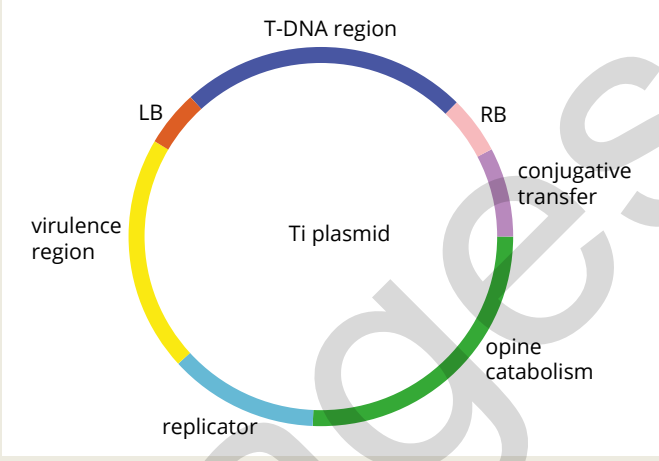


FIGURE 4.2.16 Ti plasmid from *Agrobacterium tumefaciens*. The T-DNA region carries the genes that cause tumours in crown gall disease. This region is removed and replaced with a foreign gene, making a recombinant plasmid.

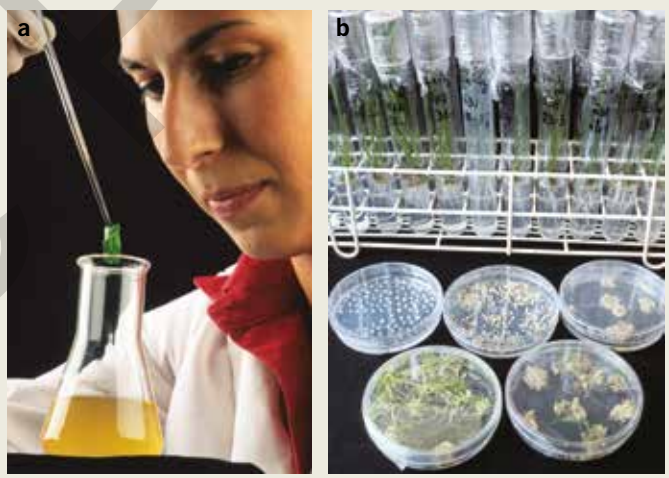


FIGURE 4.2.17 (a) Plant material is exposed to *Agrobacterium* to allow transfer of recombinant plasmid from bacteria to plant cells. (b) The plants carrying the new gene are cultured and selected in the laboratory before release for field testing of the new characteristics acquired by gene transfer.

4.2 Review

SUMMARY

- Restriction enzymes (endonucleases) are enzymes that cut DNA at particular recognition sites.
  - Sticky-end restriction enzymes leave fragments with overhanging ends that have exposed bases.
  - Blunt-end restriction enzymes cut DNA to leave flat-ended fragments.
- DNA and RNA ligase enzymes permanently join fragments of DNA or RNA together in a process called ligation.
- Plasmids are small, circular pieces of double-stranded DNA found in bacterial cells. They replicate independently of the bacteria's chromosomal DNA.
- Recombinant plasmids have had target DNA inserted into them. The same sticky-end or blunt-end restriction enzyme is used to cut both the targeted gene and the plasmid, and then DNA ligase is used to permanently join the two together.
- Plasmids with antibiotic resistance are often used to enable identification of bacterial transformation, as only bacterial cells containing these plasmids will survive when grown in cultures containing the antibiotic.
- A reporter gene produces an identifiable phenotype, such as a coloured product or fluorescence, to identify transformed bacterial cells.
- There are three steps in bacterial transformation:
  - Gene uptake: Bacterial cells are induced to take up the recombinant plasmids either by heat shock or electroporation methods.
  - Selection of transformed bacteria: The bacteria are grown in the presence of an antibiotic. Bacterial cells that have been transformed will survive, as the gene for antibiotic resistance is located in the plasmid.
  - Identification of transformed colonies: Transformed bacteria containing recombinant plasmids are identified by the reporter gene in the recombinant plasmid.

KEY QUESTIONS

Knowledge and understanding

1 **a** Restriction enzymes (endonucleases) are a basic molecular tool in gene technology. What is a restriction enzyme and what can it do?
b Describe the difference between sticky ends and blunt ends produced by restriction enzymes.

2 Define the following terms and give an example of each:
a gene cloning
b recombinant DNA

3 What is a plasmid? Describe the role played by plasmids in gene cloning.

4 Outline the purpose of DNA ligase in recombinant DNA technology.

5 **a** Explain what is meant by genetic transformation.
b List three protein products manufactured using genetic transformation and outline their importance in medicine or agriculture.

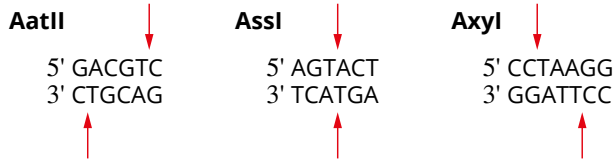
6 Give two features of a plasmid that are important for identifying cells containing a recombinant plasmid.

7 Describe how antibiotics are used to select transformed bacteria.

8 Draw a table listing a positive and a negative aspect of recombinant DNA technology for making human proteins. Include social, ethical and biological issues.

Analysis

9 Three restriction enzymes and their recognition sites are illustrated below.



Consider the DNA sequence below.

5' AGT AGT ACT GAC GCC TAA GGA GTA CTG 3'
TCA TCA TGA CTG CGG ATT CCT CAT GAC

How many pieces would it be cut into if it was mixed with:
a AatII only
b AssI only
c all three restriction enzymes

10 Draw a flow diagram outlining the different techniques and processes involved in insulin production. Use diagrams where possible to assist your explanations.

4.3 Genetically modified and transgenic organisms

Humans have used selective breeding to produce animals and plants with more useful or more attractive characteristics for tens of thousands of years. Animals or plants that expressed the desired characteristics were chosen and selectively bred, in the hope that their offspring would inherit these characteristics. In the past, selective breeding could only utilise characteristics that already existed in the gene pool of a species. We now have the knowledge and skills to use DNA manipulation techniques to alter the genetic material of an organism (Figure 4.3.1). This technology could provide many benefits, but it may also lead to questions about whether its impacts are biologically, socially or ethically acceptable.

FIGURE 4.3.1 Genetically modified cotton on the left shows its insect resistance compared to the non-modified cotton on the right, which has been ravaged by insects.

GENE EDITING WITH CRISPR-Cas9

Organisms may have their genome modified by directed mutation (mutagenesis) or by newer technologies called **gene editing**. An example of gene editing technology is a new technique called **CRISPR-Cas9**, which can cut DNA at specific locations. CRISPR stands for 'clustered regularly interspaced short palindromic repeats', which means segments of DNA with short repetitive sequences that are interspersed with unique DNA sequences.

Function of CRISPR-Cas9 in bacteria

CRISPR arrays consist of fragments of viral DNA that the bacterium has captured from invading viruses. If the same virus attacks again, the bacterium uses the CRISPR arrays to transcribe RNA sequences that are complementary to the virus's DNA sequences. The RNA sequences target the virus's DNA and cut it using endonuclease enzymes, disabling the virus. The enzyme that cuts the viral DNA is a CRISPR-associated protein called Cas9.

Application of CRISPR-Cas9 in editing genomes

The CRISPR-Cas9 system can be used to edit eukaryotic genes by combining the Cas9 enzyme with **guide RNA (gRNA)**. Scientists artificially create the gRNA, which consists of a short length of genetic code that is complementary to the target DNA sequence. The gRNA guides the Cas9 enzyme to the target site and the enzyme cuts the DNA (Figure 4.3.2). Researchers can then use the eukaryotic cell's own DNA repair mechanisms to alter the DNA, either by repairing base pair deletions or insertions that are known mutations, or by inserting new DNA sequences.

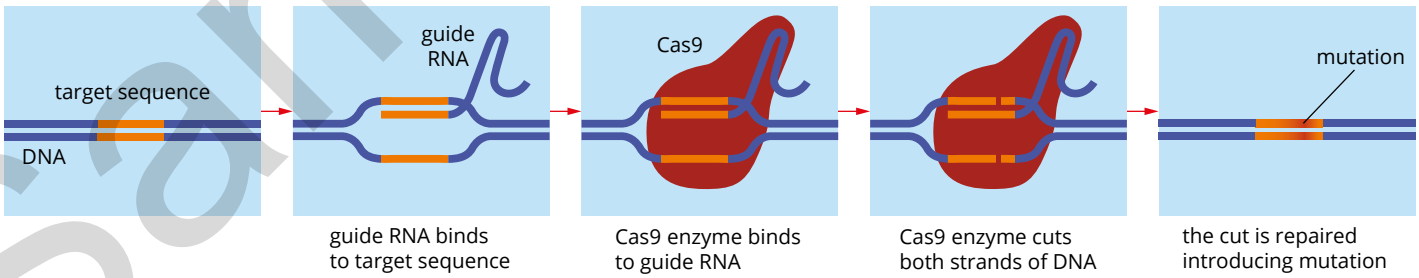


FIGURE 4.3.2 The CRISPR–Cas9 editing tool

Research with CRISPR-Cas9 in humans has already started with cancer genes and editing genes in embryos.

The current understanding of the Cas9–gRNA system is limited. The success rate for the number of genes or cells edited is not well known. Also, recent developments have shown that not every position in the gRNA needs to match the target DNA, resulting in off-target (unintended) sites being edited. Currently, research regarding CRISPR-Cas9 efficiency is inconclusive. There is a wide range of success rates for three different genes (13% to 43%), and success rate in producing transgenic mice embryos ranges from 2% to 88%.

The use of CRISPR-Cas9 to edit the genome in germline cells (eggs or sperm) raises bioethical concerns. If genetic modification occurs in embryos, not all the cells may carry the edited copy of the gene, but any changes made to eggs or sperm will affect all the cells in the embryo that develops after fertilisation. Any changes made will be passed on from generation to generation. Currently, the use of reproductive cells for CRISPR-Cas9 studies is illegal in most countries but patent rights to the technique are being contested by several companies.

Most of the gRNA sequences currently being tested are only around 20 nucleotides long. Such a short sequence is likely to exist somewhere else in the genome that is unrelated to the target site. This means crucial areas of other genes may also be unintentionally edited. The risks currently associated with the technology mean that it will be many years before CRISPR-Cas9 will be routinely used in humans.

PA 4

GENETICALLY MODIFIED AND TRANSGENIC ORGANISMS IN AGRICULTURE

Genetically modified organisms (GMOs) have had their genetic material (DNA) altered in some way.

Transgenic organisms have had genes from another species inserted into their genetic material (DNA).

Over the last few decades, techniques were developed that allowed for the alteration of an organism's genome and for the transfer of genes from one organism to another. Organisms that have had their genetic material altered are known as **genetically modified organisms (GMOs)**. Because the DNA code is universal, almost any gene transferred from one organism to another will express the protein that it expressed in the original organism. This means that a desirable characteristic seen in one animal or plant could be transferred to another organism lacking this characteristic. **Transgenic organisms** are GMOs that have had a gene from another species inserted into their genome. The gene that came from another species is called a **transgene**. Transgenic organisms can be used in agriculture to increase crop productivity and provide resistance to disease.

Genetically modified animals

In agriculture, transgenic cows and sheep are used for improved fertility, meat production, milk quality and yield, and wool quality and yield. The use of genetically modified (GM) farm animals has not expanded to the extent it has for GM plants, perhaps because of detrimental effects of some modifications in animals. For example, genes that promote growth may also cause altered skeletal growth, arthritis, and heart and kidney problems.

FIGURE 4.3.3 Transgenic salmon have been cleared for human consumption in the USA.

To date, GM animals are not approved for human consumption in Australia. In 2015, the United States government allowed genetically modified Atlantic salmon (Figure 4.3.3) to be used for human consumption. A gene from another salmon species, along with a promoter sequence from a fish called a pout, means that the transgenic salmon eat all through the year, not just when the water temperature is warm. This increases the growth rates of these fish dramatically and means they are ready for harvest much sooner than non-modified Atlantic salmon. The eggs of the GM salmon are treated to create infertile adult fish (99% of the adults are reported to be sterile), reducing the chances of interbreeding with wild salmon if they escape from their pens. This is the first genetically modified animal of any type to be cleared for human consumption in the USA.

CASE STUDY ANALYSIS

GM mosquitoes

Genetically modified mosquitoes are being used for disease control. Some mosquito species are vectors of disease-causing viruses or protozoa. Mosquitoes of the *Aedes* genus are vectors for several disease-causing viruses, including the yellow fever virus, the dengue virus and the Zika virus. A biotechnology company has developed genetically modified mosquitoes that carry a dominant lethal gene for the purpose of reducing the population of mosquitoes carrying the viruses.

Males carrying the lethal gene are released into the wild, where they mate with normal 'wild type' females and pass the lethal gene on to their offspring. The offspring die as larvae. The DNA used to make the genetically modified mosquitoes also has a reporter gene for red fluorescent protein, enabling scientists to easily identify the adults and larvae carrying the lethal gene (Figure 4.3.4).

Field trials using genetically modified *Aedes aegypti*, the vector of dengue virus, have been conducted in the Cayman Islands, Brazil and Panama, where dengue fever is a widespread and serious health problem. The recent outbreaks and rapid spread of Zika virus, which is also transmitted by *Aedes* mosquitoes, has prompted trials of these genetically modified mosquitoes in areas of Brazil affected by outbreaks of Zika virus.

Another approach to mosquito control is the release of sterile insects. Mosquitoes of the *Anopheles* genus transmit the malaria parasite *Plasmodium*. Male *Anopheles* mosquitoes have been genetically modified with genes expressed in the testes that cause the males to be unable to make sperm, so they are sterile. Female *Anopheles* mosquitoes mate only once, so mating with a sterile male limits population growth. The genes causing the sterility are linked to a reporter gene for green fluorescent protein for easy identification of the genetically modified insects (Figure 4.3.5). The aim of the research is to reduce the populations of mosquitoes that carry and transmit *Plasmodium* sp. and thus reduce the incidence of malaria, a serious health problem in many developing countries.

FIGURE 4.3.4 The gene for red fluorescent protein has been linked to a lethal gene to enable researchers to identify genetically modified mosquitoes.

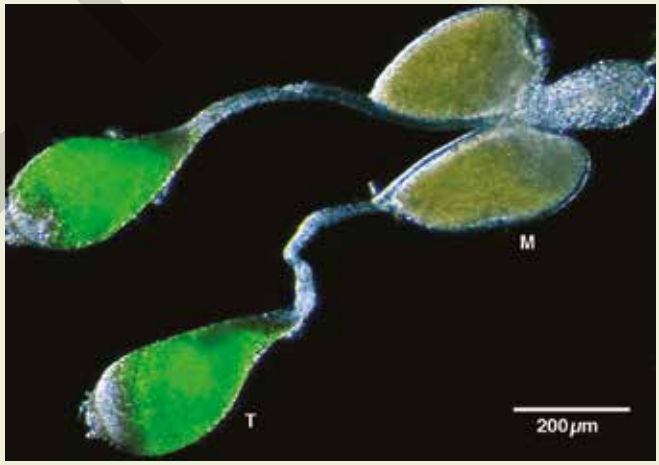


FIGURE 4.3.5 Internal reproductive organs of a genetically modified male *Anopheles gambiae* mosquito. The testes (T), where sperm cells develop, are fluorescent green due to the expression of a green fluorescent protein (GFP) that is linked to the genes causing sterility. The male accessory glands (M) that produce seminal secretions are not expressing GFP.

Analysis

1. What are the advantages and disadvantages of using this genetic approach to controlling insect vector populations?
2. Suggest alternative strategies to the use of transgenic mosquitoes to combat the spread of malaria. Do you think that these would be as effective and safe as transgenic organisms?

Genetically modified plants

GM crops are used in agriculture to increase crop productivity, provide resistance to insect predation and prevent disease. Several GM crops have been developed or are grown in Australia. For example, insect-resistant GM cotton has been grown since 1996, and herbicide-tolerant GM canola was approved for commercial production in Victoria in 2008. In Australia, the Office of the Gene Technology Regulator (OGTR) assesses all GM animals or plants before research, agricultural or commercial use.

Techniques for producing transgenic plants

Transferring a gene into plant cells can be limited by the presence of the cell wall. The introduction of foreign genes into plants is usually done by using a biological vector. One method utilises *Agrobacterium tumefaciens*, a soil bacterium that is able to naturally transfer a plasmid into plant cells (Figure 4.3.6). *Agrobacterium* normally causes crown gall disease because it carries a plasmid with genes that cause the growth of a tumour. A recombinant plasmid (the vector), carrying a desired gene from a different species but lacking the tumour-inducing genes, is introduced into *Agrobacterium* cells. When the transformed *Agrobacterium* is cultured with plant cells, the recombinant plasmid is transferred into the plant cells. These transformed plant cells are then grown in tissue culture into new plants for transplanting into the field as a transgenic crop (Figure 4.3.7).

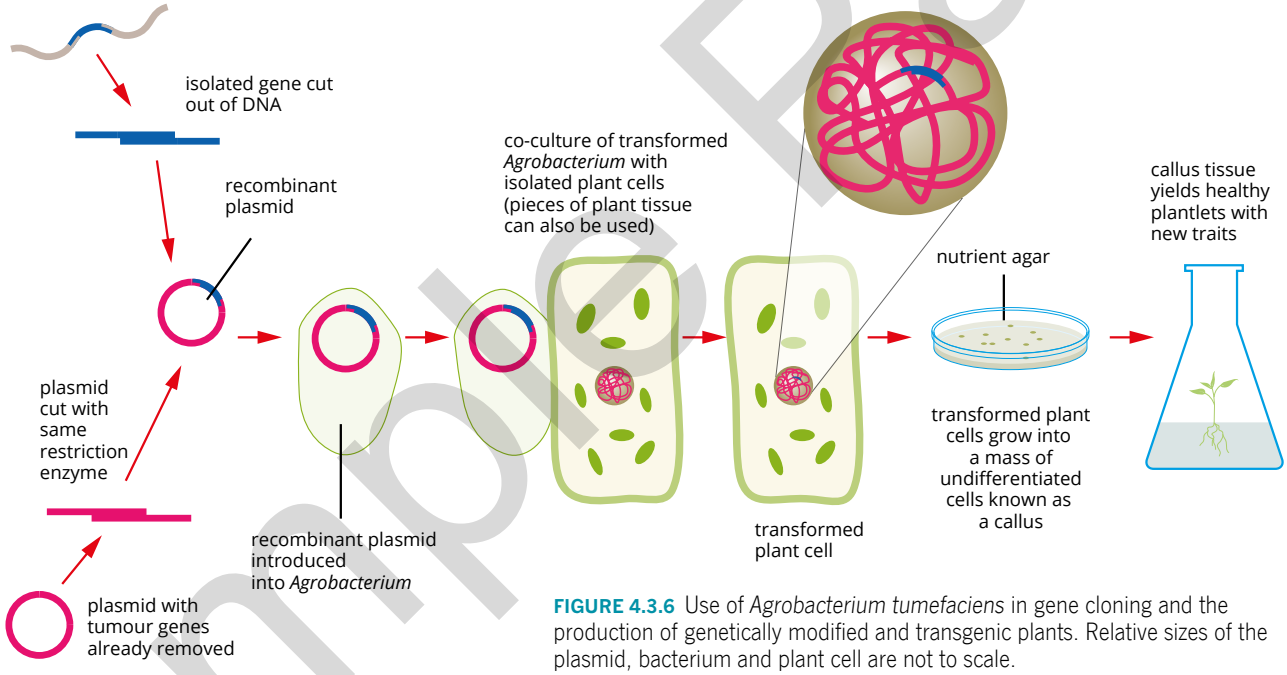


FIGURE 4.3.6 Use of *Agrobacterium tumefaciens* in gene cloning and the production of genetically modified and transgenic plants. Relative sizes of the plasmid, bacterium and plant cell are not to scale.

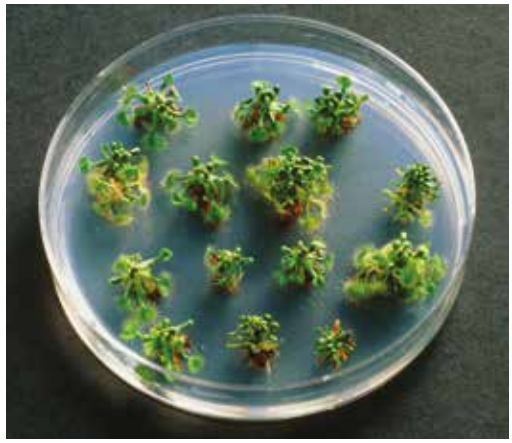


FIGURE 4.3.7 Plant cells carrying the new gene are grown in tissue culture and selected in the laboratory for field testing.

Increased crop productivity—salt tolerant wheat

Soil salinity is a major problem for Australian agriculture. A high level of sodium salts in the soil leads to osmotic water loss from roots and other tissues in which salt accumulates. Cells are stressed due to the altered ratios of sodium and potassium ions in cells. Salt-tolerant plants protect themselves from the effects of salinity by preventing sodium entry into cells, storing the salt in the vacuole, or pumping the sodium out of the cells. Molecular biologists have found the genes that control these features of salt-tolerant plants.

To increase crop productivity, Australian scientists from the University of Adelaide introduced a gene from a salt-tolerant Australian native plant into wheat plants. This greatly improved the grain yield of wheat grown on salty soils without affecting grain yield in normal soil (Figure 4.3.8). The salt-tolerant gene codes for a protein that removes sodium from the leaves, allowing water to move normally from the roots to the leaves. This increases the geographical range that can be used for wheat production in Australia and other countries facing salinity problems, which is becoming increasingly important as the global population grows.

FIGURE 4.3.8 Australian scientists have produced wheat plants that can grow in salty soil.

Disease resistance

In Ireland in the 1840s the introduction of a plant pathogen, *Phytophthora infestans*, resulted in the Irish potato famine, in which 1 million people died and 1.5 million people emigrated. The pathogen causes late blight in potato and tomato crops and is still a major problem, despite first being documented almost 180 years ago. It has been controlled with some success through the use of fungicides, but still causes significant financial loss. Researchers in the United Kingdom identified a gene for resistance to the disease in American black nightshade (a wild relative of the potato). The researchers were able to successfully insert the gene into potatoes, creating resistance to late blight in the genetically modified potato crop without the need for fungicides.

BIOFILE

Pharming for spider silk

Some farm animals are being used for the production of therapeutic proteins, such as antibodies that are difficult to make in bacteria and cultured cells. This process has been referred to as 'pharming' (combining the words 'farming' and 'pharmaceutical'). The products are released into blood or milk and they can be readily extracted from there.

Spider silk is strong and flexible and has many potential applications. Transgenic goats have been bred to produce spider silk protein in their milk.

Spider silk protein (figure at right) is an example of a potentially useful product made in transgenic goats. The gene for spider 'dragline' silk has been put into the genome of goats, along with regulatory genes, so that it is expressed in the milk. Spider silk is of great interest for its extraordinary strength and flexibility. Potential applications include a biopolymer for artificial ligaments and tendons, bandages, biodegradable bottles and tough bulletproof clothing.

CASE STUDY

Bt cotton

Cotton is a plant that attracts many insect pests. To protect the cotton crops, they are sprayed with insecticides up to four times before the crop is harvested. This high use of insecticides impacts the populations of both harmful and beneficial insects, and of the animals that feed on them. Insecticides may also have an impact on human health. In addition, insecticides are expensive.

Bt cotton is a transgenic crop that has been modified to contain two genes from the soil bacterium *Bacillus thuringiensis*. Expression of these genes produces proteins in the cotton plant that kill the main caterpillar pest of cotton by disrupting its digestive system. In Australia, almost all cotton grown is Bt cotton and this has reduced the use of pesticides dramatically. This decreases the environmental impacts of pesticides and reduces costs for farmers. Australian regulators have reported no adverse effects over 15 years of Bt cotton use in Australia. Cotton seed oil extracted from Bt cotton can be sold without GM labelling as the extraction processes separate the oil from the plant's proteins and nucleic acids, therefore the oil does not have any GM components.

BIOFILE

Issues and implications surrounding GMOs

GMOs are controversial for various reasons. Debates surrounding their biological, social and ethical implications are common in the scientific and general media. While the technology has many potential benefits, there is also the uncertainty and risk that comes with any developing technology.

The figure below summarises some common social, biological and ethical implications arising out of the use of genetically modified and transgenic organisms.

social implications	• increased food supply, nutritional content and food quality • expanded range for growth of agricultural species • access to the technology: social equality/inequality • labelling and consumer choice • patents and pricing: control of access by biotechnology companies • costs to farmers
biological implications	• safety of consuming GMOs • cross-pollination between GM plants and wild plants • cross-pollination between GM and non-GM crops • viability of transgenic organisms in the wild • health of GMOs • loss of genetic diversity in species used in agriculture
ethical implications	• violation of animal rights • human self-interest overrides ethical treatment of other organisms • intervention in evolutionary process

Summary of issues arising from the use of genetically modified and transgenic organisms

PA 1 WS 9

4.3 Review

SUMMARY

- CRISPR stands for 'clustered regularly interspaced short palindromic repeats'. In bacteria, CRISPR arrays consist of fragments of viral DNA that the bacterium has captured from invading viruses.
- CRISPR-Cas9 is used by bacteria to target an invading virus's DNA and cut it using endonuclease enzymes, disabling the virus. The enzyme that cuts the viral DNA is a CRISPR-associated protein called Cas9.
- The CRISPR-Cas9 system can be used to edit eukaryotic genes by combining the Cas9 enzyme with guide RNA (gRNA) to cut DNA at specific locations. The eukaryotic cell's own DNA repair mechanisms can then be used to alter the DNA, either by repairing mutations, or by inserting new DNA sequences.
- Organisms that have had their genetic material altered are known as genetically modified organisms (GMOs).
- Transgenic organisms are GMOs that have had a gene from a different species inserted into their genome.
- Genetically modified and transgenic animals are used in research, disease control, medicine and biomolecule production.
- *Agrobacterium tumefaciens* and plasmid transfer is an established method of transferring genes into plant cells.
- Genetically modified and transgenic plants are used in agriculture to provide varieties that resist insect attack, are herbicide resistant or have improved yield or nutritional content.
- A range of biological, social and ethical issues arise from the application of GMOs.

KEY QUESTIONS

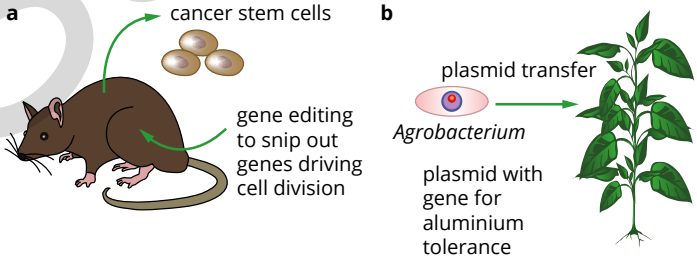
Knowledge and understanding

1 Which Australian regulatory body oversees the development, use and commercial or medical introduction of genetically modified organisms?

2 **a** What does genetic modification of an organism mean?
b Describe a successful application of genetic modification in agriculture in recent years.
c Describe a genetically modified organism and compare it to a transgenic organism.

3 Describe how genetic modification can be useful as a tool to fight vector-borne disease, such as a disease carried and transmitted by an insect vector.

4 The following diagram illustrates two model organisms used in research and the molecular procedures being used to alter a genetic characteristic. State whether the resulting organism is genetically modified, transgenic or both.



5 In your notebook, draw up a table like the one below to identify what you consider to be the key issues surrounding the use of genetically modified organisms. List the points in categories that you consider to be positive or negative aspects of the technology.

	Pros/Positives	Cons/Negatives
social		
biological		
ethical		

6 From the cases described in this section, identify one example of:
a a genetic modification that leads to a reduced environmental impact
b a modification that has potential commercial opportunity for increased production or a new product
c a modification that impacts on both the viability of an animal and the benefit to human health

Analysis

7 Explain why using guide RNA sequences (gRNA) that are 20 bases long for the CRISPR-Cas9 technique is problematic. Consider a solution.

8 Consider the implications of using transgenic animals and suggest why this technology may be opposed by some religious or cultural groups.

Chapter review

KEY TERMS

allele
anneal
bacterial competence
bacterial transformation
bacteriophage
blunt-end restriction enzyme
complementary DNA (cDNA)
CRISPR-Cas9
DNA amplification
DNA ladder
DNA ligase
DNA polymerase
DNA profiling
DNA sequencing
DNA thermocycler
endonuclease
gel electrophoresis
gene cloning
gene editing
genetic transformation
genetically modified organism (GMO)
guide RNA (gRNA)
inducer
lacZ gene
ligase
ligation
microsatellite
palindromic sequence
plasmid
polymerase
polymerase chain reaction (PCR)
polymorphism
primer
recognition site
recombinant DNA
recombinant plasmid
regulatory gene
reporter gene
restriction enzyme
reverse transcriptase
RNA ligase
RNA polymerase
short tandem repeat (STR)
sticky-end restriction enzyme
Taq polymerase
target DNA
transgene
transgenic organism
vector

REVIEW QUESTIONS

Knowledge and understanding

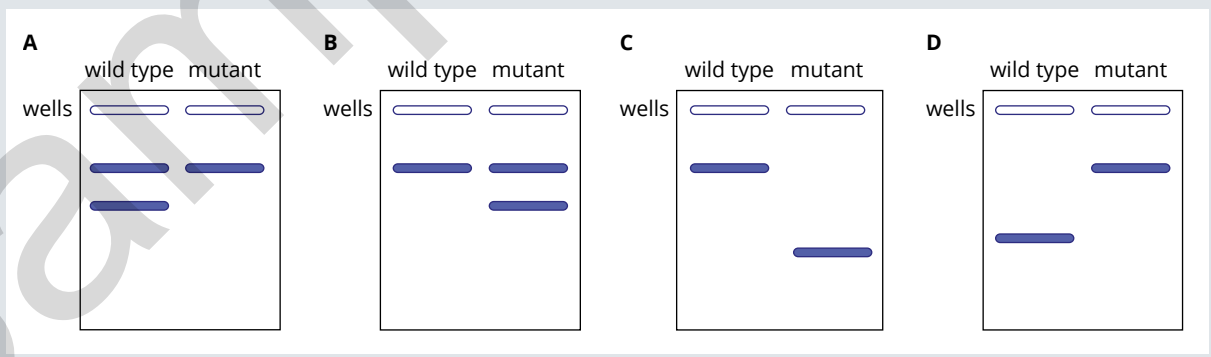
1 Indicate which step in the process of PCR best describes annealing.
- **A** binding the primers
- **B** adding the polymerase
- **C** separating the DNA strands
- **D** building the complementary DNA strands

2 What is reverse transcriptase used to make?

3 A cloned organism:
- **A** increases biodiversity
- **B** is an identical copy of its parent
- **C** can only be a plant
- **D** all of the above

4 The fruit fly, *Drosophila melanogaster*, is commonly used for genetic research. One particular mutation results in the deletion of a section of DNA 200 bp long from one particular gene. The gene was extracted from a fly that is homozygous for the mutant gene and the same gene was extracted from a fly that is homozygous for the wild type (normal) version of the gene. Both versions of the gene were amplified using PCR and then run through gel electrophoresis.

Determine which gel (below) most accurately shows the PCR products.



5 Genes such as the *lacZ* gene can be used as reporter genes. What are reporter genes used to determine?

6 Some students are doing an experiment involving bacterial transformation. Bacteria were incubated with plasmids containing resistance to the antibiotic ampicillin and then grown on agar plates. A plate that will have only transformed bacteria growing will have which of the following?

A nutrient agar only
B nutrient agar and ampicillin
C plain agar with ampicillin
D nutrient agar, ampicillin and penicillin

7 Define 'transgenic organism'.

8 One result of a genetic application was the Flavr Savr tomato. Tomatoes have a short shelf life due the effects of an enzyme called polygalacturonase. This enzyme catalyses the breakdown of the cell walls of the tomato, causing the tomatoes to become soft and unappetising. To slow down this process, the sequence of the polygalacturonase gene was determined and an antisense gene was produced. The antisense gene has a complementary nucleotide sequence to the polygalacturonase gene. The antisense gene was inserted into the tomatoes. When the antisense gene is transcribed, the mRNA produced is complementary to the mRNA for the polygalacturonase gene, so the two mRNAs join to form double-stranded mRNA. Double-stranded mRNA cannot be translated, so the enzyme is not formed and the cell walls are not broken down. The Flavr Savr tomatoes can be considered to be:

A only transgenic
B only genetically modified
C both genetically modified and transgenic
D none of the above

9 Before any gene can be inserted into bacteria to make proteins for human use, the number of copies of the gene must be increased. In order to do this a process called PCR is used.

a i What do the letters PCR stand for?
ii State the role of the 'P' in the process.
iii Identify the source of the 'P' used in this process. Why is that particular source used?

b One particular PCR machine uses the following sequence: heat to 94°C for 1 minute, cool to 56°C for 1.5 minutes, then heat to 72°C for 1.5 minutes.
i Describe what is happening at each stage.
ii How long would it take to obtain 8000 copies of the target DNA?

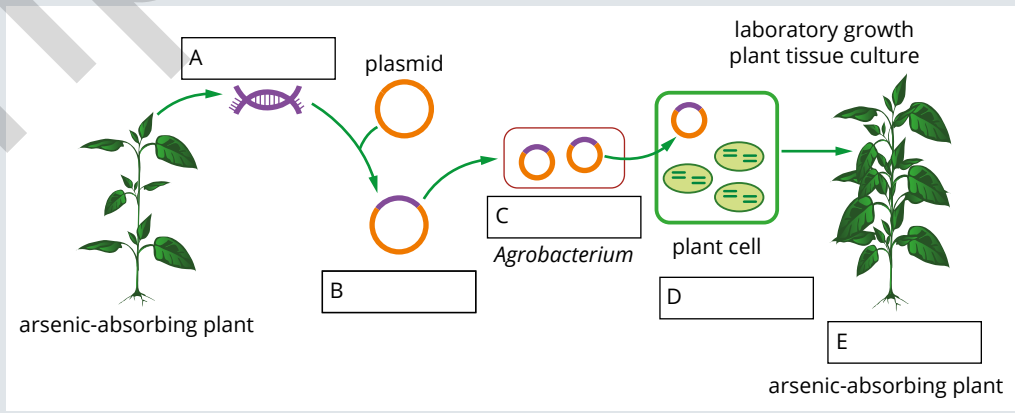
10 What is the purpose of a DNA ladder in gel electrophoresis?

11 How might DNA profiling be used in the legal system?

12 One of the purposes of recombinant DNA technology is to produce large quantities of proteins for therapeutic use. To do this, the gene that codes for protein production is inserted into a plasmid, prior to being placed in a bacterial cell.

a State three reasons why plasmids make good vectors for protein production.
b State three reasons why bacteria make good host organisms for the plasmids.

13 Arsenic contamination of soil is a serious problem in some countries. The arsenic contaminates groundwater and drinking wells. The flow chart below illustrates a process used to insert a gene that enables plants to absorb arsenic from the soil. Name the objects that are labelled A–E.



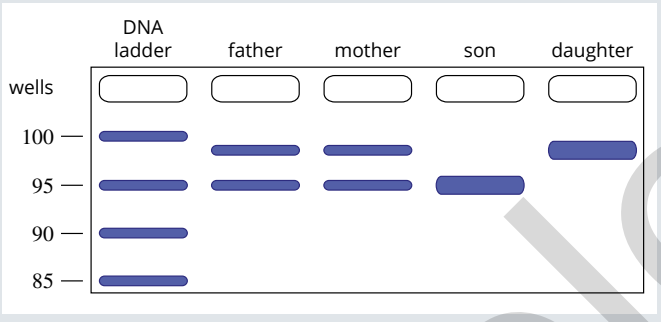
14 **a** What types of human proteins are commonly produced by recombinant DNA technology?
b Suggest an advantage of this method of production compared to a traditional approach.

15 Briefly describe the technical and ethical issues related to DNA profiling.

16 **a** Where do CRISPR arrays come from?
b Outline the steps involved in using CRISPR-Cas9 to edit genes.
c Describe two bioethical issues that arise from the use of CRISPR-Cas9

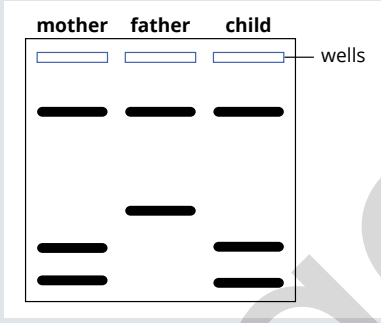
Application and analysis

17 Members of a family with a history of cystic fibrosis (CF) underwent genetic testing to determine whether they carried the common ΔF508 mutation. DNA samples obtained from cheek cells were analysed by PCR using primers specific for the ΔF508 region, followed by gel electrophoresis. The normal allele yields a 98 bp DNA fragment. The mutant allele yields a 95 bp DNA fragment.

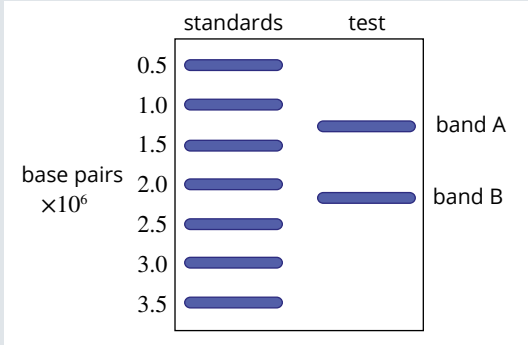


a How is PCR able to identify the allele responsible for cystic fibrosis?
b Describe the purpose of gel electrophoresis in this type of genetic testing.
c The parents are carriers of CF. Explain how the PCR and gel electrophoresis results show this.
d What does the genetic test show about the son?
e The daughter gets a cold and chest infection every winter. Is this likely to be related to the lung congestion seen in CF?

18 A couple wishes to find out if their unborn child has sickle cell anaemia. The figure top right shows the results from the gel electrophoresis of the restriction fragments of the sickle cell gene (located on chromosome 11) for the family. The mother carries the mutation, which results in sickle cell anaemia, while the father does not carry the mutation.
a How does the use of restriction enzymes in the analysis of the alleles of a gene result in different banding patterns?
b Does the child carry the mutation for sickle cell anaemia? Explain your answer.



19 Some genetic disorders are sex-linked, meaning that they present differently depending on whether individuals are male or female. Males have an X and a Y chromosome, while females have two X chromosomes. Duchenne muscular dystrophy (DMD) is a sex-linked genetic disease caused by the deletion of part of the sequence of the dystrophin gene, which is located on the X chromosome. The dystrophin protein is very large. The normal gene is 2 220 390 bp long and contains many exons and introns. The total length of the coding sequences (the exons) is 11 055 nucleotides.
a How many amino acids are in the normal protein?
A young woman who has a family history of DMD is about to start trying to have a family. Her partner's family has no history of the condition.
b The first step the young woman takes is to be tested for the condition. How could she have the genetic change and not know?
c The relevant sections of the X chromosomes of the young woman were cut using an appropriate restriction enzyme and then run on an electrophoresis gel. A set of DNA standards was also run. The gel is shown below.

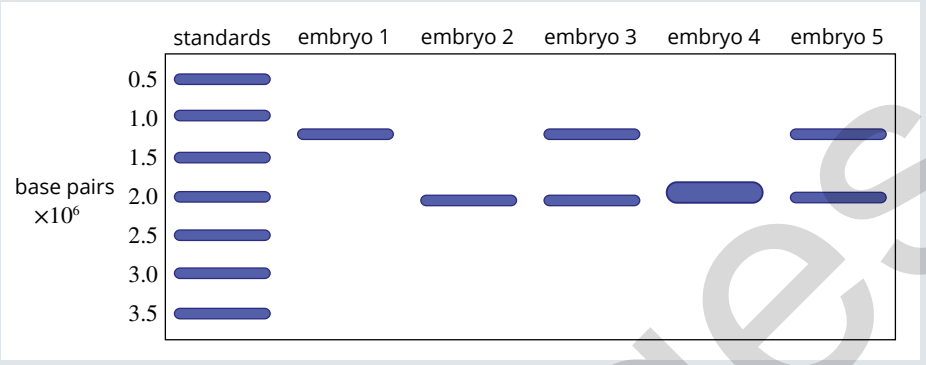


i Explain whether the woman has the DMD mutation.
ii What are the approximate lengths of the pieces of DNA represented by bands B and A?

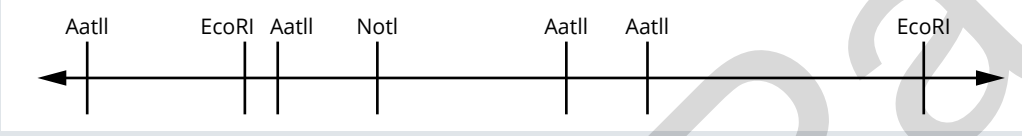
d The woman decides to use IVF and preimplantation testing to become pregnant. The doctors harvest eight eggs, five of which are then fertilised with her partner's sperm. The embryos are then tested for the DMD allele using gel electrophoresis.

i Explain which embryos are male.

ii Explain which embryos are most suitable for implantation.



20 The diagram below represents a linear DNA molecule and shows the position of the recognition sites for a number of restriction enzymes.



a Examine the diagram above and determine how many fragments will be produced when the DNA molecule is cut with EcoRI.

b Examine the diagram above and determine how many fragments will be produced when the DNA molecule is cut with AatII.

c Examine the diagram above and determine how many fragments will be produced when the DNA molecule is cut with NotI.

21 Individuals with haemophilia carry a gene mutation in the *F8* gene, which codes for factor VIII, a clotting factor critical in the formation of blood clots that form after an injury. Blood clots temporarily close off injured blood vessels to stop bleeding, until the vessels are healed by the body. Individuals with haemophilia are at a significantly increased risk of bleeding after injury because they are missing this important clotting factor.

Before the availability of recombinant DNA technology, factor VIII for the management of haemophilia was obtained from blood serum provided by blood donors. Large volumes of blood serum were required in order to obtain sufficient factor VIII for treatment. In addition, the risks of transmitting other blood proteins and viruses (including HIV) was high. Many men with haemophilia contracted HIV as a result of this treatment.

Today, haemophilia is a well-managed disease thanks to the advances in recombinant DNA technology.

Discuss three advantages of using recombinant DNA technology to artificially produce factor VIII.

22 Bacteria are commonly genetically engineered to produce human proteins. The DNA sequence for the gene of one of these proteins is shown below.

```
5' – GCT  TCT  TCC  CGT  GCA  TAT  AGA  TAC  TCT  GAA  ACA  CTG  TGC  GGC  GGT  GAA  CTG
3' – CGA  AGA  AGG  GCA  CGT  ATA  TCT  ATG  AGA  CTT  TGT  GAC  ACG  CCG  CCA  CTT  GAC

... many base pairs ...  CTG  TGC  ACC  TAT  TGT  GCT  ACT  CCC  GCA  AAG  TCC  GAA  TAG  TAG  GCT  TCT
                         GAC  ACG  TGA  ATA  ACA  CGA  TGA  GGG  CGT  TTC  AGG  CTT  ATC  ATC  CGA  AGA

CGC  TGC  TCC  CGT  GCT  TCT  CGC  GTA  TGT  CCG – 3'
GCG  ACG  AGG  GCA  CGA  AGA  GCG  CAT  ACA  GGC – 5'
```

A restriction enzyme is used to cut the gene from the human genome. Four possible enzymes have recognition sequences and cutting sites as shown.

enzyme 1	enzyme 2	enzyme 3	enzyme 4
GC|TTCT CGAA|GA	C|TGTGC GACAC|G	TCC|CGT AGG|GCA	CTG|TGC GAC|ACG

The DNA sequence is quite long, so only the beginning and end are shown along with a section before and after the gene. The start and stop triplets are underlined.

a Explain which of the restriction enzymes would be most suitable to cut out the gene so that it can be inserted into the bacterium that will produce the protein.

b A mutation can occur that changes the base indicated with the arrow from a T to a C. One way to identify individuals who have this mutation is to cut the DNA with a restriction enzyme and run the DNA on an electrophoresis gel. Explain why enzyme 4 is the most appropriate to use for this purpose.

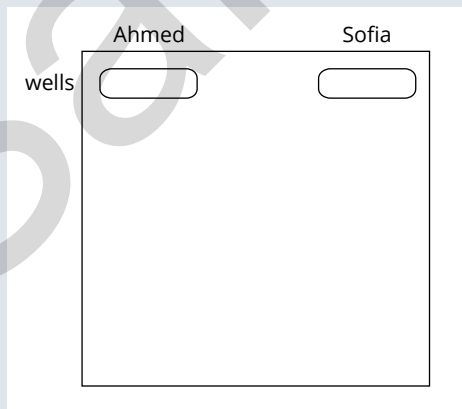
c This mutation runs in one particular family. Ahmed and Sofia are members of the family and decide to be tested. Ahmed turns out to have two normal alleles and Sofia is heterozygous (one normal and one mutant allele).

i If enzyme 4 is used, how many DNA bands will result from the cutting of Ahmed's DNA?

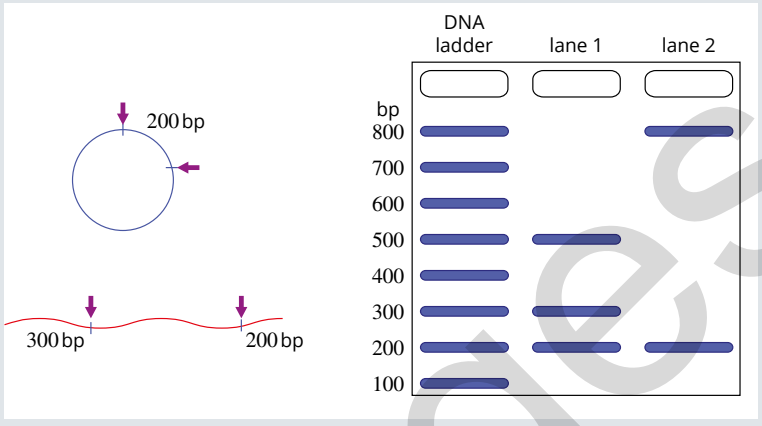
ii If enzyme 4 is used, how many DNA bands will result from the cutting of Sofia's DNA?

iii Redraw and complete the picture of the electrophoresis of the DNA of Ahmed and Sofia after cutting with enzyme 4.

iv Show the positions of the positive and negative terminals on the electrophoresis set below, and explain why you placed them in those positions.

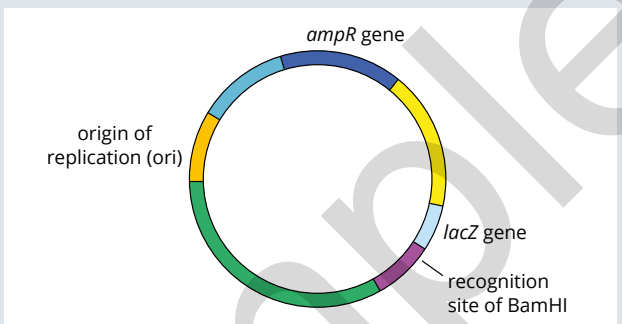


23 A plasmid of total length 1000 bp and a segment of a linear chromosome are being used to make recombinant DNA (diagram to the right). The DNA was cut with the restriction enzyme Tat1, which leaves sticky ends. The cutting sites are indicated by arrows. The resulting fragments were run on a gel (diagram to the right). The purpose of the process is to insert a segment of the chromosome into the plasmid for gene cloning.

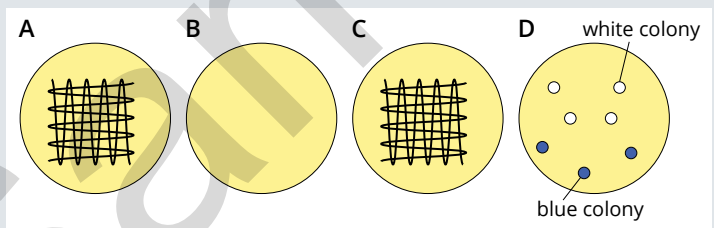


- **a** Which lane on the gel has the fragments from the plasmid digestion? Explain your choice.
- **b** What length is the starting chromosomal DNA? Explain your answer.
- **c** What enzyme will be needed to make a recombinant plasmid using the large chromosome fragment?

24 *E. coli* cells were transformed with a plasmid containing a gene for ampicillin resistance, the *lacZ* gene for blue/white screening and, depending on the success of gene uptake, a gene for a desired protein product. When the production of a recombinant plasmid was successful, the protein-encoding gene was inserted into the *lacZ* gene using the recognition site for the BamHI restriction enzyme. The plasmid is shown below.



The following results were obtained:



	A	B	C	D
added to plate	nutrient agar only	nutrient agar with ampicillin	nutrient agar only	nutrient agar, with ampicillin and X-gal
cultured on plate	untransformed bacteria only	untransformed bacteria only	transformed bacteria	transformed bacteria
description	lawn of bacteria	no growth	lawn of bacteria	blue and white colonies of bacteria

- **a** Identify why plate B is significant.
- **b** Specify the purpose of X-gal in plate D.
- **c** Distinguish between the white and blue colonies on plate D. What evidence can you present to explain the differences?
- **d** Consider the location of the restriction enzyme recognition site within the *lacZ* gene. Do you agree that this location is critical in allowing researchers to determine the success of the bacterial transformation? Explain.

25 At a small country hospital three babies were born on one night. This stretched the resources of the hospital to such an extent that normal procedures failed and the babies were not labelled with their mother's name. In order to ensure the correct babies were taken home by the correct parents, DNA testing was performed. A STR on chromosome 6 that has between 7 and 20 ATTG repeats was investigated in order to match the parents with their babies. The results for the couples and the babies are shown below.

Couple one		Couple two		Couple three	
Mother	Father	Mother	Father	Mother	Father
11, 14	7, 12	14, 20	12, 18	18, 20	11, 18

Baby one	Baby two	Baby three
12, 20	11, 20	12, 14

a Match each baby with its correct parents.

b Explain how you matched the couples with their children.

c Figure 4.1.13 on page 142 shows one way of analysing a series of STRs. It shows the analysis of 10 sites. Some sites have two peaks and others only one. Explain why this is the case.

26 What are some of the effects (positive and negative) of genetically modified and transgenic organisms on biodiversity?

27 Create a table listing the advantages and disadvantages of biotechnology for animal welfare.

28 Does artificial manipulation of DNA have the potential to change populations forever? Consider examples of GMOs that are already in use in Australia. Evaluate the impact these GMOs have had on the populations of these organisms in the short and long term.