10 BIOTECHNOLOGY PROVIDES EVIDENCE OF EVOLUTION

UNIT 4 CONTENT

SCIENCE INQUIRY SKILLS

- conduct investigations, including the use of virtual or real biotechnological techniques of polymerase chain reaction (PCR), gel electrophoresis for deoxyribonucleic acid (DNA) sequencing, and techniques for relative and absolute dating, safely, competently and methodically for valid and reliable collection of data
- represent data in meaningful and useful ways; organise and analyse data to identify trends, patterns and relationships; discuss the ways in which measurement error, instrumental accuracy, the nature of the procedure and sample size may influence uncertainty and limitations in data; and select, synthesise and use evidence to make and justify conclusions
- select, use and/or construct appropriate representations, including phylogenetic trees, to communicate conceptual understanding, solve problems and make predictions

SCIENCE AS A HUMAN ENDEAVOUR

- developments in biotechnology have increased access to genetic information of species, populations and individuals, existing now or in the past, the interpretation and use of which may be open to ethical considerations
- developments in the fields of comparative genomics, comparative biochemistry and bioinformatics have enabled identification of further evidence for evolutionary relationships, which help refine existing models and theories

SCIENCE UNDERSTANDING

Evidence for evolution

- biotechnological techniques provide evidence for evolution by using PCR (to amplify minute samples of DNA to testable amounts), bacterial enzymes and gel electrophoresis to facilitate DNA sequencing of genomes
- comparative studies of DNA (genomic and mitochondrial), proteins and anatomy provide additional evidence for evolution; genomic information enables the construction of phylogenetic trees showing evolutionary relationships between groups

Source: School Curriculum and Standards Authority, Government of Western Australia

9780170449168

The great majority of scientists accept the general idea of evolution. However, this idea has always been controversial and there are still many people who, for a variety of reasons, do not accept the idea that species evolve. It is important, therefore, to look carefully at the evidence available to support the theory of evolution. Besides the study of fossils, much of the evidence for evolution has come from comparative studies. Traditionally, the focus for comparative studies has been on anatomy and embryology, but the development of technology in more recent times now allows comparative studies to be conducted on both proteins and DNA.

DNA also helps us to understand why evolution happens. When Darwin proposed the theory of evolution, he based it on evidence gained through observation. We now understand the structure of DNA and the inheritance process. This allows us to make sense of the evidence provided and understand the changes that led to them.

10.1 PROCESSING DNA

DNA can be processed by a range of techniques so that it can be analysed and compared for applications such as tracking evolutionary changes. In this section, we will be focusing on:

- making many copies of the sample using polymerase chain reaction (PCR)
- cutting the DNA into smaller lengths with restriction enzymes from bacteria
- separating the lengths of DNA with gel electrophoresis to produce a DNA profile
- determining the sequence of nucleotides, or their bases, using Sanger's method.

Polymerase chain reaction and amplifying DNA

One of the early limitations of DNA analysis was the amount of DNA needed. In many cases there was insufficient DNA, or the organism may have been adversely affected in order to collect enough. The development of the **polymerase chain reaction (PCR)** has enabled small quantities of DNA to be replicated, producing testable amounts to use in analysis techniques. Kary Mullis was awarded the Nobel Prize in Chemistry in 1993 in recognition of his development of the PCR technique.

Nobel prizes
This website provides the press release for the 1993 Nobel Prize in Chemistry.

PCR mimics the natural process of DNA replication that occurs prior to cell division. During the process, the DNA goes through a series of three steps:

1. *Denaturing:* The two strands of DNA are separated.
2. *Annealing:* Short sections of DNA (primers) are bound to the separated strands.
3. *Extension:* The short sections of DNA are extended to produce longer strands.

This sequence is repeated approximately 20–30 times, in a process called **thermocycling**. It takes two to three hours to produce about a billion copies of the DNA.

Denaturation

During natural DNA replication, the enzyme helicase separates the two strands of DNA, allowing each strand to be copied. The polymerase chain reaction uses heat to achieve the same function. Temperatures of approximately 94–96°C are used to break the hydrogen bonds holding the two strands together. This separates the strands without disrupting each individual strand.

Annealing

During the annealing process the temperature is decreased to approximately 50–60°C. This allows short strands of DNA called **primers** to bind to the single DNA strands. The primers are not random sections of DNA. Instead, they are complementary to either end of the section of DNA to be copied.

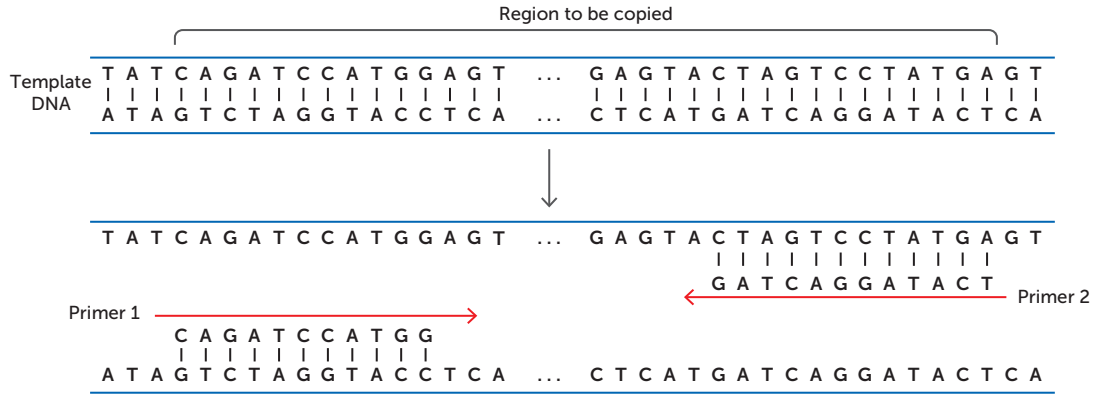


FIGURE 10.1 Primers that are complementary to either end of the region to be copied are used during the annealing process

Extension

The extension step, also known as elongation, also mimics the process of DNA replication. The enzyme DNA polymerase is used to join new, complementary nucleotides to the sections originating with the primers. This extends, or elongates, the nucleotide chain and creates a new strand of DNA. However, it is not the full length of the original DNA, as it starts at the primer and not at the end of the DNA. Eventually, the majority of DNA strands are the length of DNA between the location of the primers.

DNA polymerase attaches to double-stranded DNA. Prior to the extension step, this only occurs where the primers are located. Therefore, the primers act as a starting point and, hence, initiate DNA replication.

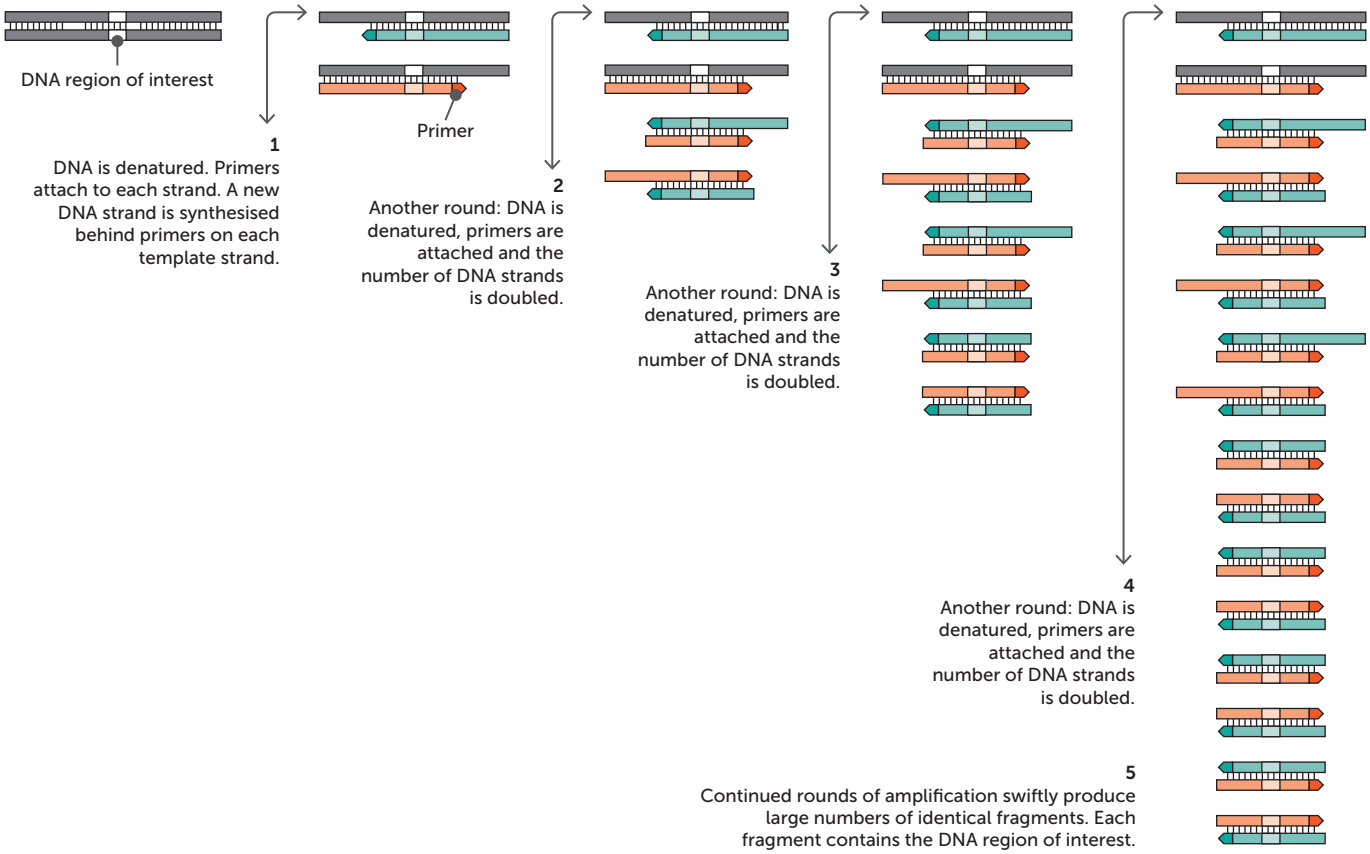


FIGURE 10.2 A diagrammatic representation of the polymerase chain reaction

In the PCR process, the denaturation step is carried out at 94–96°C. However, at this temperature, the DNA polymerase is usually destroyed. This means that more of the enzyme has to be added after the heating stage of each cycle. This is very time consuming and uses large amounts of DNA polymerase.

Therefore, almost all PCR applications now use a heat-stable DNA polymerase. One such enzyme, taken from a heat-loving bacterium called *Thermus aquaticus*, is called Taq polymerase. It does not denature when heated and has allowed the procedure to be simplified and automated, permitting the PCR sample to be alternately heated and cooled. Taq polymerase's optimal temperature is 68–72°C; therefore, the extension phase is carried out at this temperature.

PCR
This website provides an animated sequence on PCR.

PCR virtual laboratory
This website provides a PCR virtual laboratory.

Key concept

The polymerase chain reaction uses heat, primers and DNA polymerase to amplify a section of DNA, producing many copies from a small amount.

Gel electrophoresis and DNA profiling

In Chapter 8, you learnt about **restriction enzymes** and their ability to cut DNA at specific nucleotide sequences. These enzymes can be used in a range of applications, including DNA analysis. When restriction enzymes are added to DNA, it cuts the strands into different lengths depending on the base sequence of the specific DNA sample. The length of these pieces can be analysed and compared with other DNA samples.

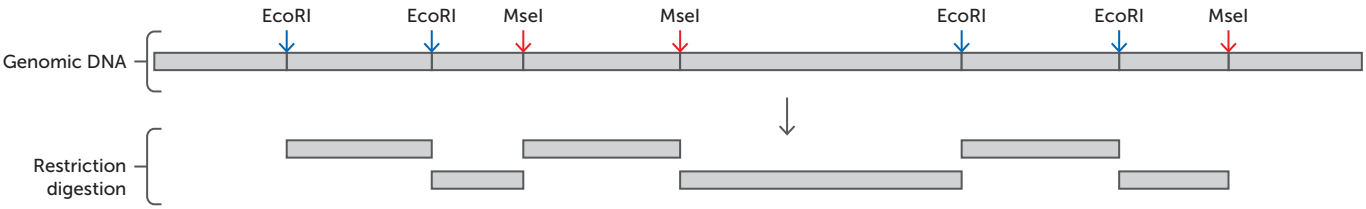


FIGURE 10.3 DNA cut with the restriction enzymes EcoRI and MseI producing pieces of different lengths

Gel electrophoresis is a technique that is able to separate DNA strands based on their lengths. The DNA pieces are placed in **wells** in a semi-solid gel that is immersed in a solution of an electrolyte. There are electrodes at either end of the gel. The negative electrode is closest to the DNA and the positive electrode is at the opposite side. When an electric current is passed through the gel, the negatively charged DNA moves towards the positive electrode. The smaller DNA pieces move faster than the larger ones and so are located further away from the negative electrode when the current is stopped. This results in a pattern of bands that looks similar to the barcodes on products sold in supermarkets. This banding pattern is an individual's **DNA profile**, often called a **DNA fingerprint**.

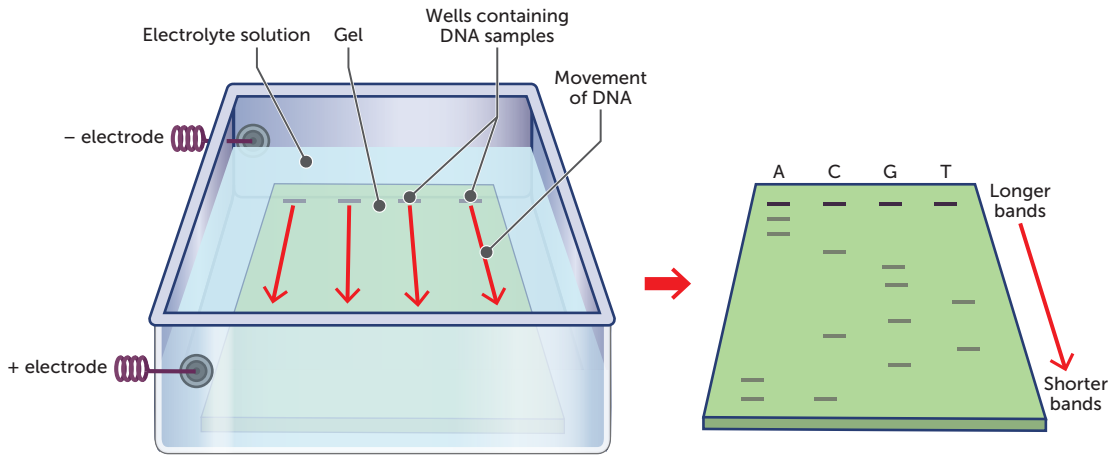


FIGURE 10.4
An illustration of the apparatus commonly used for gel electrophoresis and the resulting profile

Gel electrophoresis
This website contains more information about gel electrophoresis.

Filling the wells

Using a micropipette
This website has more information about how to use a micropipette.

The wells where the DNA is placed are simply depressions in the gel. This means that, when the current is applied, the DNA will move through the gel rather than diffuse through the solution. Therefore, for gel electrophoresis to work effectively the DNA needs to be accurately placed in the wells. This is typically done using a **micropipette**. Micropipettes have disposable tips at the end that can be put on and off the pipette without any contact, reducing any chance of cross-contamination.

iStock.com/Eakkalak Khiwsuwannasri

FIGURE 10.5 Micropipette

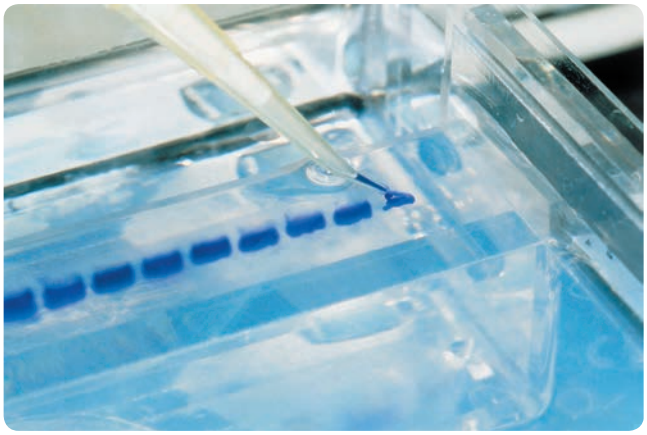


FIGURE 10.6 Micropipettes are used to place the DNA samples into the wells of the gel

DNA ladders

A DNA ladder is often 'run' at the same time as the samples. The ladder contains segments of DNA with known lengths. The results from the unknown sample are compared to the ladder to determine the length of the DNA strands in the sample.

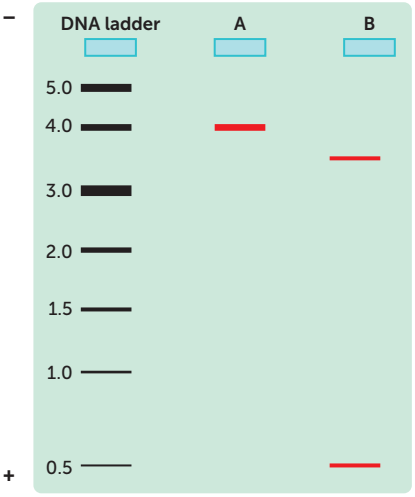


FIGURE 10.7 Samples of DNA can be compared to the DNA ladder

Visualising DNA

Activity 10.1
Investigating electrophoresis simulation

After the DNA strands are separated, scientists use different methods to visualise the DNA. Ethidium bromide can be added to the agar prior to the gel being set. As the DNA moves through the gel it picks up some of the chemical. Upon completion of the 'run', a special ultraviolet light is shone over the gel and the DNA fluoresce. Unfortunately, ethidium bromide is a carcinogen and, therefore, must be handled very carefully.

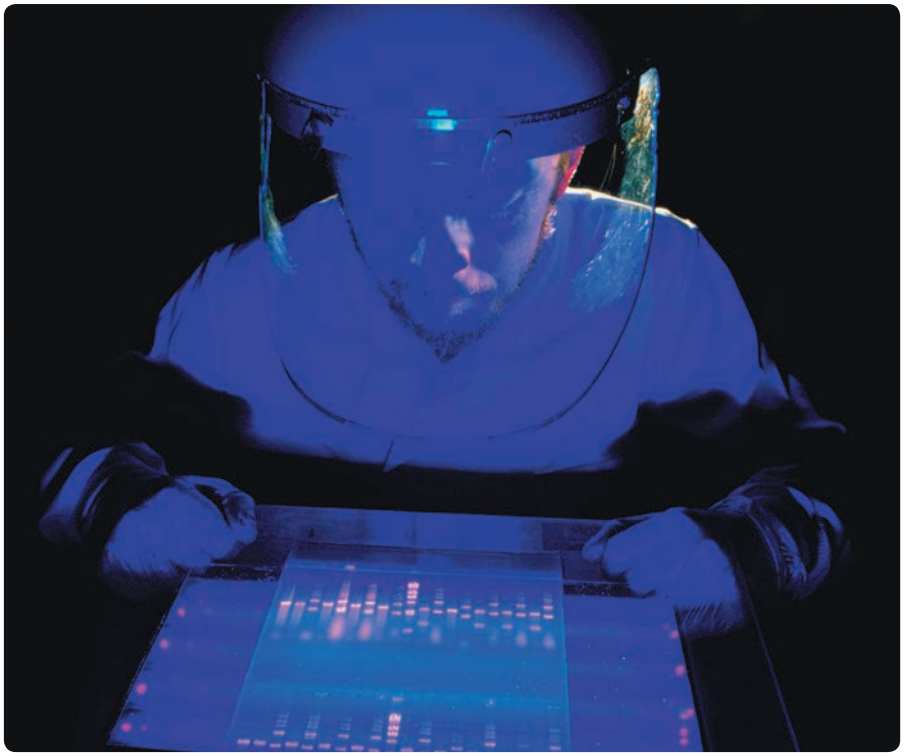
Activity 10.2
Investigating the effect of restriction digestion enzymes on lambda DNA

Methylene blue is a dye that binds to DNA. When the gel is soaked in the dye, the areas containing DNA stain a deeper blue and are therefore visible to the naked eye. Another method is the use of DNA probes. These are short sections of a single strand of DNA with a radioactive or fluorescent molecule that binds to the DNA being tested.

Key concept

A DNA profile is produced by using restriction enzymes to cut DNA into smaller lengths that are separated by gel electrophoresis, which uses an electrical current to move the DNA segments through the gel at a rate proportional to their length.

9780170449168



Alamy Stock Photo/Science Photo Library

FIGURE 10.8
Gel electrophoresis stained with ethidium bromide and placed under UV light

DNA sequencing

DNA sequencing is the determination of the precise order of nucleotides in a sample of DNA. The method most frequently used to determine such a sequence was invented by Frederick Sanger, who was awarded his second Nobel Prize in Chemistry in 1980 for this accomplishment.

DNA is synthesised from four nucleotides, each with a different nitrogenous base – adenine, cytosine, guanine or thymine. Nucleotides are more correctly called deoxynucleotide triphosphates, as they consist of three phosphate groups joined to the sugar deoxyribose with its base.

When DNA forms:

- each nucleotide loses two phosphate groups
- the sugar molecule loses a hydrogen atom from the hydroxy group (OH) when it bonds to the phosphate group of an adjacent nucleotide.

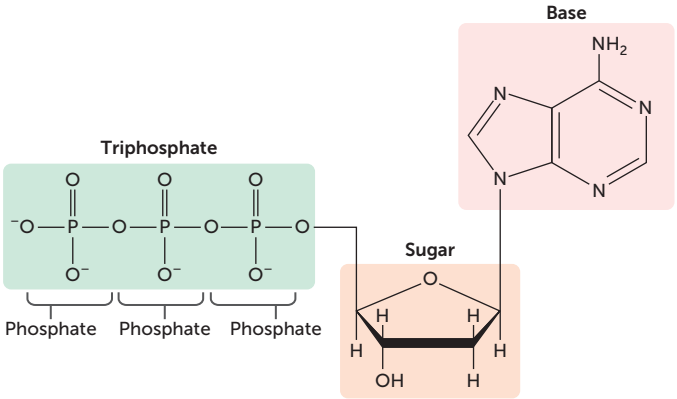
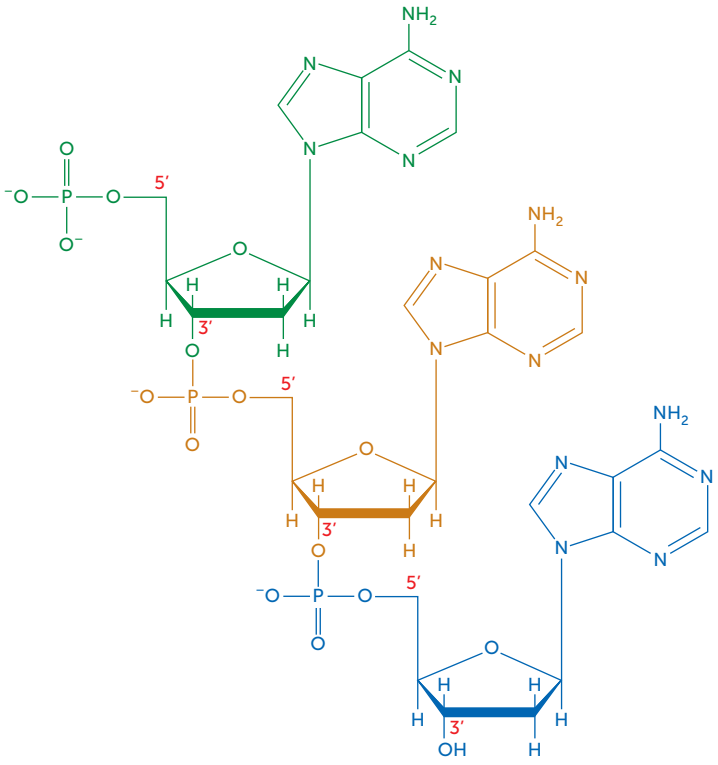


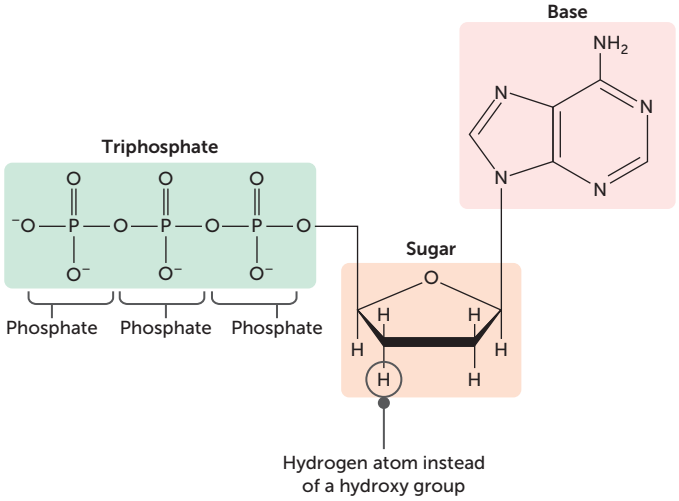
FIGURE 10.9
Structure of a nucleotide

FIGURE 10.10
Three nucleotides joined together in a section of DNA



In Sanger's method of determining a DNA sequence, synthetic nucleotides that lack the OH group are added to the growing strand. These are called dideoxyribonucleotide triphosphates, or **dideoxyribonucleotides** (ddNTPs).

FIGURE 10.11
Structure of a dideoxyribonucleotide triphosphate



The synthetic nucleotide stops the elongation of the sequence because there is no OH group for the next nucleotide to attach to. This happens at each of the nucleotides in the DNA sample, creating different lengths of DNA. These can be separated using gel electrophoresis. Knowing which base was added to create each length allows scientists to determine the order of nucleotides.

9780170449168

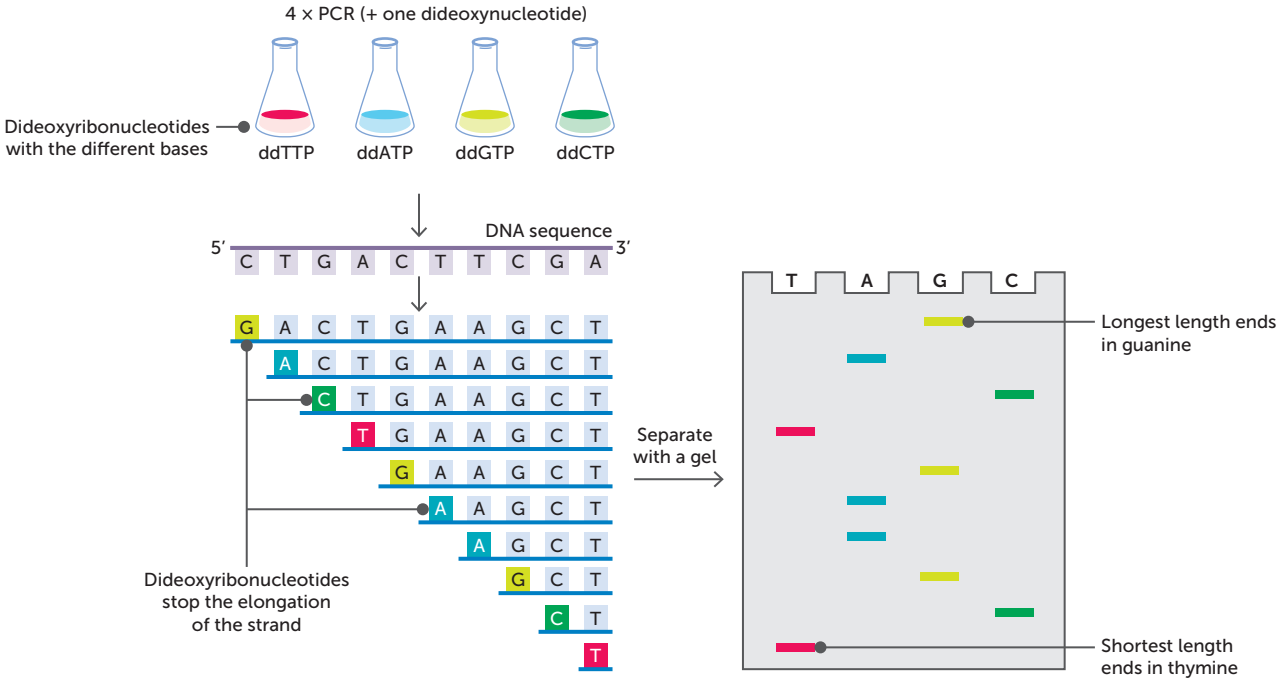
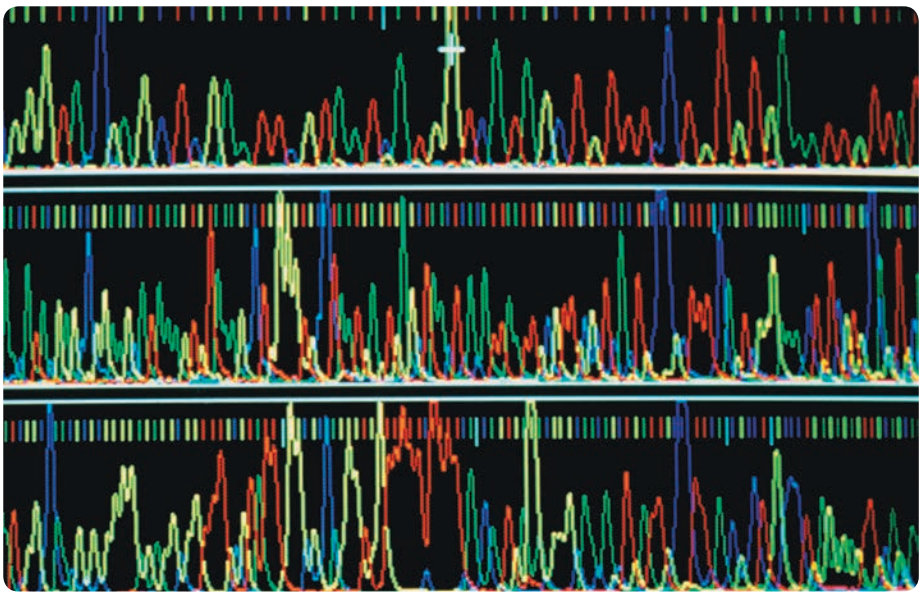


FIGURE 10.12 DNA sequencing using Sanger's method

DNA sequencing can be used to identify mutations or to compare DNA from different organisms. This is useful in identifying inherited disorders such as sickle-cell anaemia, cystic fibrosis and some forms of cancer. It has also been used for maternity and paternity tests, in cases where the identity of the father or mother of a child is in dispute. DNA sequencing can be used by scientists to compare species in order to track evolutionary changes.

Sanger sequencing
This website shows an animation of the Sanger method of DNA sequencing.

DNA sequencing
This website provides a series of annotated slides on DNA sequencing.



Alamy Stock Photo/Science Photo Library

FIGURE 10.13 A DNA sequence displayed on a computer screen

Key concept

The sequence of nucleotides in DNA can be determined by methods such as Sanger's method, which uses dideoxyribonucleotides to stop the lengthening of the DNA strand. This produces segments of different lengths that can be separated by gel electrophoresis.

Activity 10.3
Investigating biotechnological techniques

Ethical considerations of genetic information
This website contains more detailed information about the ethical considerations of genetic information.

Ethical considerations with genetic information

As technologies continue to advance, it is becoming easier and cheaper to obtain genetic information. As with the use of all technologies, the ethical use of genetic information is an important consideration. This includes respecting the following principles.

- *Autonomy*: respect for the right to be self-determining and to choose whether or not to be tested and, if tested, to know and share the information. It also includes the right of an individual to decide their own future, independent of genetic information.
- *Confidentiality*: the use of genetic information is treated sensitively, and is accessed only by those who are authorised to access it.
- *Equity*: the right to fair and equal treatment regardless of genetic information.
- *Privacy*: the right to be 'left alone' and to make decisions regarding genetic testing and the resulting information, independent of others.

key concept

Ethical factors such as autonomy, confidentiality, equity and privacy need to be considered when gaining and using genetic information.

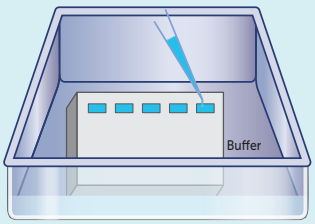
Questions 10.1

RECALL KNOWLEDGE

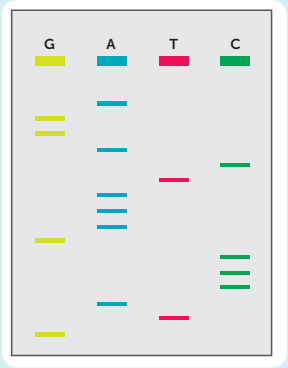
1 What does 'PCR' stand for?
2 List the three steps in PCR.
3 Explain the role of a primer in PCR.
4 What is the advantage of using Taq polymerase over other DNA polymerases?
5 Explain the role of restriction enzymes in DNA profiling.
6 Will the shorter or longer lengths of DNA travel the greatest distance during electrophoresis? Explain your answer.
7 Name the instrument used to place the DNA in the wells for gel electrophoresis.
8 Explain why a DNA ladder is useful in interpreting results from electrophoresis.
9 List three methods of visualising DNA after electrophoresis.
10 Draw a simplified structure of:
   a deoxynucleotide triphosphate
   b dideoxyribonucleotide triphosphate.

APPLY KNOWLEDGE

11 Explain why temperatures of approximately 96°C are sufficient to denature the DNA.
12 Suggest what would happen if the temperature were too high during the annealing step of PCR.
13 'DNA sequencing makes it possible for suitable primers to be chosen for PCR.' Discuss this statement.
14 The image below shows DNA being put into the wells of an agarose gel prior to electrophoresis. Label the positive and negative sides and explain why it must be placed this way.



15 The diagram below shows the bands produced from electrophoresis during DNA sequencing using Sanger's method. Write the base sequence for the section of DNA.



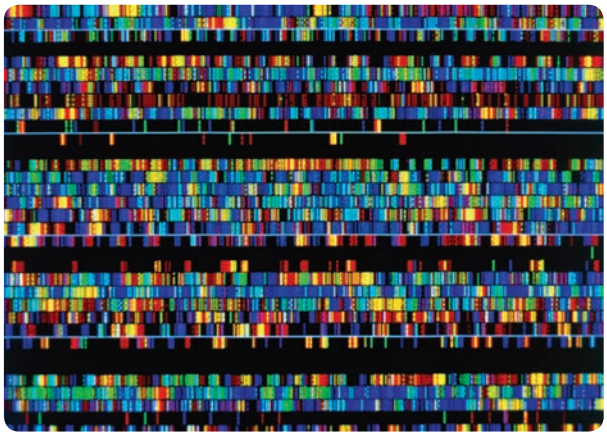
16 Identify two situations when it may be unethical to use genetic information. For each situation, discuss the reasons for and against its use.

10.2 DNA PROVIDES EVIDENCE OF EVOLUTION

Information obtained in DNA profiles and DNA sequencing can be used to compare the DNA of different organisms.

Comparative genomics

Scientists have determined that all living things use the same DNA code, adding weight to the hypothesis that all living things are related to each other and have evolved from a common ancestor. Although all species of organisms have DNA, the sequence of bases in the DNA varies. New genes are gained by mutation; others are lost by natural selection, genetic drift or some other process. Despite the common ancestor, the code in the DNA is different for different species.



Science Photo Library/JAMES KING-HOLMES

FIGURE 10.14 Computer screen display of the human DNA sequence. Each colour represents a specific base. The sequence of bases makes up the genetic code

When speciation occurs, the new species would have very similar DNA. However, as the new species gradually change through mutations, natural selection and genetic drift, they accumulate more and more differences in their DNA. Species that are more distantly related have more differences in their DNA, whereas species that are more closely related share a greater portion of their DNA.

The complete set of DNA in each cell of an organism is called the **genome**. **Comparative genomics** is a relatively new field of biological research in which the genome sequences of different species are compared. By comparing the sequence of the human genome with genomes of other organisms, researchers are able to identify regions of similarity and difference. This procedure provides an effective means of studying evolutionary changes among organisms, helping to identify genes that are preserved among species, as well as genes that give each organism its unique characteristics.

TABLE 10.1 Comparative genome sizes of a number of organisms

ORGANISM	ESTIMATED SIZE (BASE PAIRS)	CHROMOSOME NUMBER	ESTIMATED GENE NUMBER
Human	3.0 billion	46	21 000
Mouse	2.9 billion	40	21 000
Fruit fly	165 million	8	13 000
Roundworm	97 million	12	19 000
Yeast	12 million	31	6 000
Bacteria	4.6 million	1	3 200

The successful completion of the Human Genome Project in 2003 demonstrated that major sequencing projects can generate high-quality data. Consequently, interest in sequencing the genomes of many other species rose significantly. By analysing the genomic features that have been preserved in a number of species over millions of years, researchers are beginning to tease apart the often-subtle differences between animal species. Comparative genomics has revealed a high level of similarity between closely related organisms such as humans and chimpanzees. It has also been used to reveal the diversity of gene composition in different evolutionary lineages. Such research may result in a rearrangement of the way we view some of the evolutionary relationships between primates.

9780170449168

WS

10.1 Biotechnology

Examination of the genome of our closest living relatives, the chimpanzees, shows that they share more than 98% of our DNA (Table 10.2). Scientists quote slightly higher or lower figures depending on what exactly is being compared. Interestingly, humans have 23 pairs of chromosomes, while chimpanzees have 24 pairs. Scientists believe that, at some time in the past, two small chromosomes found in chimpanzees fused to form one of the human chromosomes.

TABLE 10.2 Relationship between humans and great apes using DNA differences

PRIMATES BEING COMPARED	DNA DIFFERENCE (%)
Human–chimpanzee	1.2
Chimpanzee–gorilla	1.2
Human–gorilla	1.6
Chimpanzee–orangutan	1.8
Human–orangutan	2.4
Gorilla–orangutan	2.4

DNA and evolution
This website provides a light-hearted account of using DNA sequencing.

Endogenous retroviruses

In addition to the genes coding for proteins, chromosomes also contain some non-coding sequences of bases in the DNA. These sequences are sometimes referred to as 'junk DNA' as they have no apparent function and appear to serve no purpose. Comparisons of junk DNA provide similar results as those for other parts of the genome: more closely related species have more junk sequences in common. This observation only makes sense if related species have evolved from a common ancestor.

Good examples of stretches of apparently non-functional DNA are **endogenous retroviruses (ERVs)**. An ERV is a viral sequence that has become part of an organism's genome. Retroviruses store their genetic information as RNA, not DNA. Upon entering a cell, a retrovirus copies its RNA genome into DNA in a process known as **reverse transcription**. The DNA then becomes inserted into one of the host cell's chromosomes. A retrovirus only becomes endogenous if it inserts into a cell whose chromosomes will be inherited by the next generation, an ovum or a sperm cell. The offspring of the infected individual will then have a copy of the ERV in the same place, in the same chromosome, in every single one of their cells. All subsequent generations will also have a copy of the ERV at the same location. What scientists have found is that ERVs make up 8% of the human genome, and that other primates also possess some of the same ERVs in exactly the same locations in their genomes.

For example, when comparing the chromosomes of humans and chimpanzees, it has been discovered that the same ERVs are located in the short arm of chromosome 10, the short arm of chromosome 1, the long arm of chromosome 9 and the short arm of chromosome 6 for both species. In all, scientists have found 16 instances of human ERVs matching exactly with chimpanzee ERVs. This is compelling evidence that humans and chimpanzees share a common ancestor. Any retrovirus that became inserted into the genome of a common ancestor would be inherited by both chimpanzees and humans at exactly the same location in the chromosome.

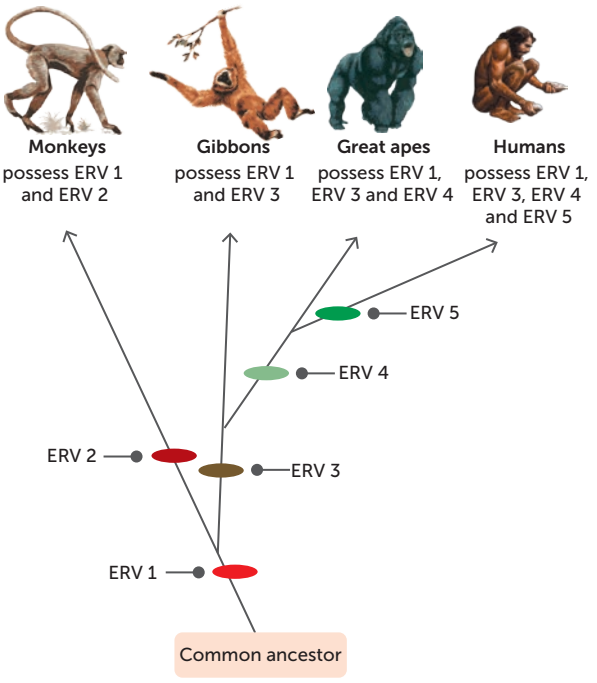


FIGURE 10.15 Simplified example of how endogenous retroviruses could be used to trace common ancestry: the great apes and humans have a more recent common ancestor as they share more endogenous retroviruses

Mitochondrial DNA

Mitochondria are the organelles in the cell where the aerobic phase of respiration occurs to release energy for use by the cell. Most of a cell's DNA is located in the nucleus, but a small amount is in the mitochondria. This is called **mitochondrial DNA**, or **mtDNA**.

Unlike the DNA in the nucleus, which is in the form of very long strands, mitochondrial DNA is in the form of small circular molecules. There are about five to ten of these molecules in each mitochondrion. Mitochondrial DNA has 37 genes, all of which are essential for the mitochondrion to function normally. Twenty-four of the genes contain the code for making transfer RNA molecules, which are involved in protein synthesis. The other 13 genes have instructions for making some of the enzymes necessary for the reactions of cellular respiration.

Most cells contain large numbers of mitochondria and therefore usually have between 500 and 1000 copies of the mtDNA molecule. This makes it a lot easier to find and extract than the DNA in the nucleus, and so smaller samples can be used. In humans, the mtDNA genome consists of about 16 500 base pairs, representing only a fraction of the total amount of DNA in a cell.

Inheritance of mitochondrial DNA

Human eggs and sperm both have mitochondria, but while an egg has many hundreds, a sperm has only about 100, just enough to provide the energy for the sperm to swim to the egg. After a sperm has penetrated the egg at fertilisation, the mitochondria in the sperm are rapidly destroyed. This means that, while our nuclear DNA comes from the nucleus of the egg and the sperm, our mitochondrial DNA comes only from the egg. In other words, we inherit nuclear DNA from both parents, but we inherit mitochondrial DNA only from our mothers. You got your mitochondrial DNA from your mother; she got it from her mother, and so on.

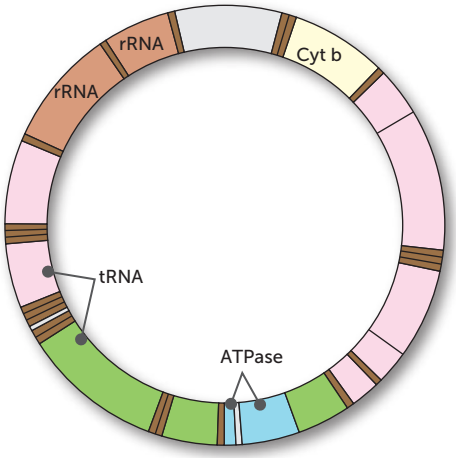


FIGURE 10.16 Model of a molecule of mtDNA showing the location of some of the genes

Evidence from mitochondrial DNA

DNA found in the mitochondria has a higher rate of mutation than nuclear DNA. Because of these mutations, human mtDNA has been slowly diverging from the mtDNA of our original female ancestor, and the amount of mutation is roughly proportional to the amount of time that has passed. Scientists are able to use the similarity between the mtDNA of any two individuals to provide an estimate of the closeness of their relationship through their maternal ancestors. If their mtDNA is identical they will be closely related, perhaps even siblings. On the other hand, if the mtDNA is very different, their last common maternal ancestor lived long ago. The use of mtDNA has been found to be of most value when comparing individuals within a species and for species that are closely related. In this way, it has allowed scientists to track the ancestry of many species back hundreds of generations. For example, through studying mitochondrial DNA it has been possible to trace the migration routes of ancient peoples. Such studies have shown that most Europeans are descended from hunter-gatherers who migrated into Europe during the last Ice Age, rather than from farmers coming from the Middle East. It has also been used to demonstrate the evolutionary relationships between humans and closely related species.

Analysis of mtDNA has become an important tool in mapping the relationships between species. Using such analysis, scientists can verify evidence of evolution gained from other sources. For example, examination of mtDNA has shown that the last common ancestor of modern humans and Neanderthals lived around 600 000 years ago.

mtDNA and Neanderthals
This website provides information about the identification of the genes in mitochondrial DNA of Neanderthals.

Questions 10.2

RECALL KNOWLEDGE

1. Where do scientists gain information from when comparing DNA?
2. Define 'genome'. Describe how sequencing the genome can be used to provide evidence for evolution.
3. Define 'endogenous retrovirus' and 'non-coding DNA'.
4. Describe how endogenous retroviruses are used as evidence for evolution.
5. What do the genes on mtDNA code for?
6. Why do we only inherit mitochondrial DNA from our mothers?

APPLY KNOWLEDGE

7. The DNA of dogs is 85% similar to that of humans, while the DNA of chimpanzees is 98% similar to that of humans. Explain how this information supports the idea that we have a more recent common ancestor with chimpanzees than with dogs.
8. Explain why not all retroviruses are endogenous retroviruses, and why only endogenous retroviruses are useful in providing evidence for evolution.
9. Explain why comparison of structures such as endogenous retroviruses and mitochondrial DNA was not available prior to the development of techniques such as electrophoresis and DNA sequencing.
10. Compare and contrast mitochondrial DNA and nuclear DNA.

10.3 OTHER BIOCHEMICAL EVIDENCE

DNA is not the only chemical substance that can provide evidence for evolution. Other techniques include a comparison of proteins such as cytochrome C and bioinformatics.

Protein sequences

Comparative protein studies also provide evidence for evolution. Proteins consist of long chains of particular amino acids linked together in a precise sequence determined by the DNA. There are tens of thousands of types of proteins in living things and all are fabricated from 20 kinds of amino acids.

Modern biochemical techniques enable the sequence of amino acids in a protein to be determined. By comparing the type and sequence of amino acids in similar proteins from different species, the degree of similarity can be established. Animals of the same species have identical amino acid sequences in their proteins, and those from different species have different amino acids or they are arranged in a different order. Just like DNA analysis, the degree of difference between proteins enables an estimate to be made of the amount of evolution that has taken place since two species developed from a common ancestor. The longer the period of time involved, the greater the number of amino acids that are different.

Amino acids are usually represented by a three-letter code, frequently the first three letters of their name. To make comparison of amino acid chains easier, scientists have also adopted a system of coding whereby one letter is used to represent one particular amino acid. By listing the amino acids for a particular protein in sequence, a comparison can be made with other species. This has been done for a number of proteins that appear to be in all species and which are referred to as **ubiquitous proteins**. Such proteins perform very basic, but essential, tasks that all organisms require for life. Ubiquitous proteins are found in all organisms, from bacteria to humans, and are completely independent of an organism's specific function or the environment in which it lives. Such proteins carry out the same functions no matter where they are found.

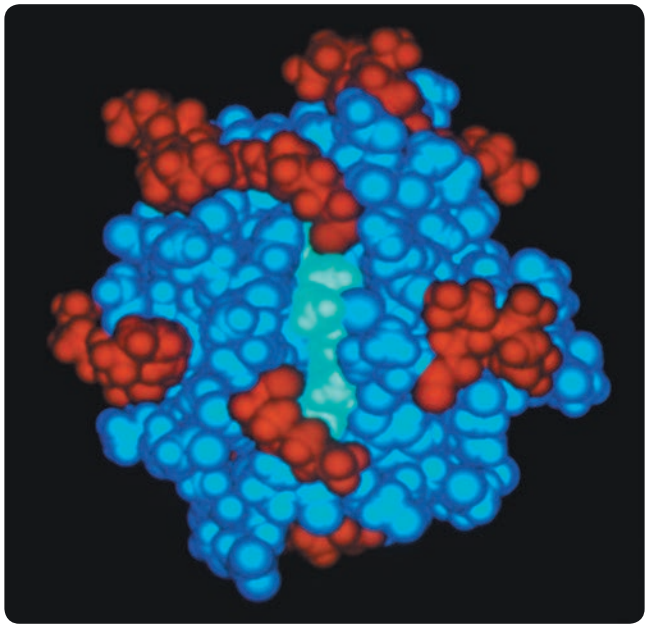
9780170449168

TABLE 10.3 The three-letter and single-letter codes used for amino acids

AMINO ACID	THREE-LETTER CODE	SINGLE-LETTER CODE
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Cytochrome C is a well-researched example of a ubiquitous protein that shows how protein sequences can provide evidence for evolution. This protein performs an essential step in the production of cellular energy. It appears to have changed very little over millions of years of evolution. Human cytochrome C contains 104 amino acids. Regardless of the species tested, 37 of these have been found at the same positions in every sequenced cytochrome C molecule. This strongly suggests that these proteins have descended from an ancestral cytochrome C molecule found in a primitive microbe that existed more than 2000 million years ago.

To compare cytochrome C sequences, they need to be aligned so that the maximum number of positions containing the same amino acids can be determined. The more similarity there is between two molecules, the more recently they have evolved from a common ancestor. By doing such comparisons, scientists have determined that the cytochrome C of chimpanzees and gorillas is the same as that for humans, and for rhesus monkeys it differs by only one amino acid compared with that of humans.



Alamy Stock Photo/RGB Ventures/SuperStock

FIGURE 10.17 A model of cytochrome C

TABLE 10.4 Differences in amino acids in cytochrome C between humans and other species

SPECIES COMPARED WITH HUMANS	NUMBER OF DIFFERENCES FROM HUMAN CYTOCHROME C
Chimpanzee	0
Gorilla	0
Rhesus monkey	1
Patas monkey	1
Rabbit	9
Cow	10
Pigeon	12
Bullfrog	18
Tuna	21
Fruit fly	24
Yeast	44

9780170449168

Activity 10.4
Investigating amino acid sequencing

Other protein sequences have been examined and yielded similar results. The alpha and beta chains of the blood protein haemoglobin are identical in humans and chimpanzees, but the same protein sequences in gorillas differ by one amino acid. When the same chains are examined in gibbons, there are three amino acid differences. A comparison of the delta chain indicates that humans differ from chimpanzees and gorillas by one amino acid, and from gibbons by two. Such protein studies provide more support for the evolutionary relationships between primates that have already been established by DNA comparisons.

Bioinformatics

Bioinformatics has become an important part of many areas of human biological science and is particularly useful in providing evidence for evolutionary relationships. It is a multidisciplinary field that combines all areas of biological science with computer science, engineering, statistics and applied mathematics to help understand biological processes. However, in practical terms, **bioinformatics** is the use of computers to describe the *molecular* components of living things. It uses biochemical analysis to gain information about DNA and proteins, and computer software to store and analyse it. Bioinformatics has been particularly useful in assisting evolutionary biologists to trace the evolution of a large number of organisms by measuring changes in their DNA, rather than through traditional techniques of physical taxonomy or physiological observations. The more similar the genes of two species, the closer their evolutionary relationship.

Comparison of molecules
This website presents evidence to show that comparisons of molecules provide support for evolution.

More recent developments have enabled researchers to compare entire genomes. In doing so, the genes and other biological features in a DNA sequence need to be identified, in a process termed **annotation**. This process needs to be computerised, as most genomes are far too large to be annotated by hand. Annotation is made possible by the fact that genes have recognisable start and stop codons (see Chapter 9 of *Human Perspectives ATAR Units 1 & 2*).

Questions 10.3

RECALL KNOWLEDGE

1. State the relationship between DNA, RNA, amino acids and proteins.
2. How many different amino acids make up proteins?
3. When comparing amino acid sequences, scientists use a single letter rather than the three letters that are usually used to identify them. Why do scientists do this?
4. Ubiquitous proteins are important when comparing proteins.
   - a. Define 'ubiquitous protein'.
   - b. Give an example of a ubiquitous protein.
5. What is bioinformatics?
6. What is annotation, and why is it part of bioinformatics?

APPLY KNOWLEDGE

7. Evolution results from changes in DNA. Given this fact, explain why a comparison of the sequence of amino acids in a particular protein can provide evidence for evolution.
8. Table 10.5 sets out an amino acid sequence from alpha haemoglobin of five different species of animals. Compare each of the amino acid sequences to the one from humans.
   - a. Which species' sequence is the most similar?
   - b. Which species' sequence is the most different?
   - c. Does this correlate with our current understanding of evolution and common ancestors?

TABLE 10.5 The amino acid sequence for alpha haemoglobin of different species

SPECIES	AMINO ACID SEQUENCE FROM ALPHA HAEMOGLOBIN
Human	VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTK TYFPHFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNAL SALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVH ASLDKFLASVSTVLTSKYR
Whale	VLSPTDKSNVKATWAKIGNHGAEYGAEALERMFMNFPSTKT YFPHFDLGHDSAQVKGHGKKVADALTKAVGHMDNLLDALS DLSDLHAHKLRVDPANFKLLSHCLLVTLALHLPAEFTPSVHA SLDKFLASVSTVLTSKYR
Macaw	VLSGSDKTNVKGIFSKIGGQAEDYGAEALERMFATFPQTKTY FPHFDVSPGSAQVKAHGKKVAAALVEAANHIDDIATALSKLS DLHAQKLRVDPVNFKLLGQCFLVVVAIHNPSALTPEVHASLD KFLCAVGNVLTAKYR
Baboon	VLSPDDKKHVKAAWGKVGEHAGEYGAEALERMFLSFPTTKT YFPHFDLSHGSDQVNKHGKKVADALTLAVGHVDDMPQALSK LSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTPAVHASL DKFLASVSTVLTSKYR
Frog	HLTADDKKHIKAIWPSVAAHGDKYGGEALHRMFMCAPKTKT YFPDFDFSEHSKHILAHGKKVSDALNEACNHLDNIAGCLSKLS DLHAYDLRVDPGNFPLLAHQILVVVAIHFPKQFDPATHKALD KFLVSVSNVLTSKYR

CHAPTER 10 ACTIVITIES

Electrophoresis simulation

ACTIVITY 10.1 Investigating electrophoresis simulation

Go to the weblink and work through the electrophoresis simulation. As you go, or after you have finished, answer the following questions.

1. What ingredients are used to make the gel?
2. Describe how the gel is made using the ingredients that you have listed.
3. DNA samples are placed in wells in the gel. Explain how the wells are made.
4. What is the purpose of the DNA size standard?
5. What electrical charge does a DNA molecule have?
6. Which electrical charge is applied to the well end of the gel?
7. Is it possible to tell whether an electric current is running through the gel?
8. What makes the DNA migrate through the gel?
9. Describe the technique that is used to make the DNA visible in the gel.
10. Why do shorter DNA strands move further through the gel than longer strands?

SOUTHERN Biological Developed exclusively by Southern Biological

ACTIVITY 10.2 Investigating the effect of restriction digestion enzymes on lambda DNA

Restriction digestion is the process of cutting DNA molecules into smaller pieces with special enzymes called restriction endonucleases or restriction enzymes. These special enzymes recognise specific sequences in the DNA molecule (e.g. EcoRI GAATTC) wherever that sequence occurs in the DNA.

Aim

To use restriction enzymes to cut DNA into respective fragments
To analyse your restriction digestion using a gel electrophoresis apparatus
Time requirement: 55 minutes

You will need

Restriction digestion materials

Lambda DNA 1 µg (8 uL); restriction digestion buffer (20 µL); EcoRI enzyme (1 µL); HindIII enzyme (1 µL); BamHI enzyme (1 µL); sterile nuclease-free water (200 µL); 4 sterile microtubes 0.5 mL; variable micropipette (2–20 µL); variable micropipette (0.5–10 µL); sterile pipette tips; water bath; microcentrifuge (optional); clock or stopwatch; disposable gloves

Electrophoresis materials

TBE buffer (25 µL); agarose; MIDORI Green safe stain (for pre-staining technique); loading dye 6x (50 µL); variable micropipette (2–20 µL); electrophoresis chamber (blueGel); power supply 100 V (if using an alternative to blueGel); blue light transilluminator (optional)

Note: The above measurements are based on using a blueGel electrophoresis apparatus. If an alternative electrophoreses chamber is being used, increase TBE quantities based on chamber size.

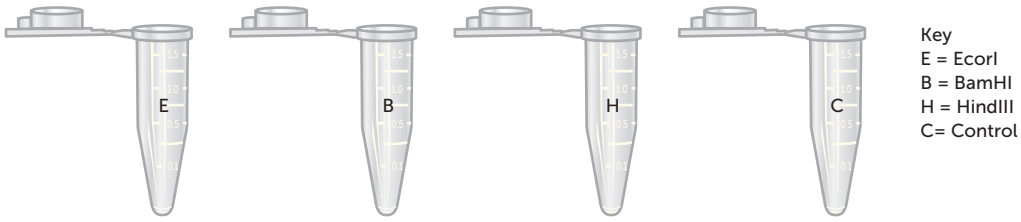
Risks

WHAT ARE THE RISKS IN THIS INVESTIGATION?	HOW CAN YOU MANAGE THESE RISKS TO STAY SAFE?
TBE buffer can cause skin irritation.	Wear appropriate personal protective equipment at all times, including eye protection and gloves. Wash skin immediately if contact does occur.
Disposable gloves may pose an allergy risk.	Use a type of glove that removes allergy risk and is suitable for the chemicals being used.

What to do

Restriction digestion

1 Collect four 500 µL (0.5 mL) microtubes and label them as follows:



2 Using a micropipette, add 42 µL (microlitres) of nuclease-free water to each of the microtubes.
3 Add 2 µL of Lambda DNA to each of the microtubes.
4 Using a fresh micropipette tip, add 5 µL of restriction digestion buffer to each of the microtubes.
5 Using a fresh micropipette tip for each sample, add 1 µL of the EcoRI enzyme to 'E', 1 µL of the BamHI enzyme to 'B', 1 µL of the HindIII enzyme to 'H' and 1 µL of nuclease-free water to 'C'.
6 Mix the samples thoroughly by pipetting up and down a few times using the larger micropipette with fresh tips for each sample. Continue until the samples have an even consistency. To collect the liquid at the base of the tubes, spin with a microcentrifuge, or tap the tubes on a bench.
7 Place the microtubes in a 37°C water bath for 10 minutes.

Analysing your digestion using gel electrophoresis

1 Collect the four tubes from the water bath and add 10 µL of loading dye to each sample.
2 Mix samples thoroughly by pipetting up and down a few times using the larger micropipette with fresh tips for each sample, until the solutions look consistent throughout. To collect the liquid at the base of the tubes, spin with a microcentrifuge. Your samples are now ready to be loaded into the gel.
3 Carefully remove the combs from the gel. Place the prepared 0.8% agarose gel into the gel electrophoresis chamber, ensuring that the wells are at the top or negative electrode section of the chamber.
4 Pour TBE buffer into your gel electrophoresis chamber, ensuring that you completely cover the surface of the gel.
5 Using a fresh pipette tip for each sample, load 10 µL of each of your restriction digest samples into the wells located near the negative electrode and note the specific lanes in which the different samples were loaded.
6 Once complete, carefully place the lid on the gel chamber and press the on button and let it run for 30 minutes. Turn on the built-in blue light to visualise DNA band separation if using a blueGel™ electrophoresis chamber.

9780170449168

Note: If using a gel electrophoresis chamber that requires an external power supply, carefully plug the positive and negative electrodes into the gel box without dislodging the gel. The negative end should be connected to the end closest to the DNA samples. Plug in the power source (set at 100 V), turn it on and let the gel run.

7 After 30 minutes, turn the power supply off and observe the DNA fragments either by turning on blue light or transferring to a blue light transilluminator. Record your results.

Note: If you are using the post-stain method, DNA will not be visible until the gel has been soaked in methylene blue or equivalent for up to 24 hours.

Studying your results

1 How many cuts did each restriction enzyme make?
2 Measure the distance in mm and copy and complete the table below.
3 Graph your results for HindIII digest to determine the sizes of the EcoRI digest and/or BamHI digest. (Try graphing the log – base pairs – vs distance.)
4 Do those fragments add up to the size of lambda DNA? If not, provide possible explanation(s) as to why not.

Analysis of restriction digests of DNA

HindIII		EcoRI			BamHI		
DISTANCE (MM)	SIZE (BP)	DISTANCE (MM)	CALCULATED BP	SIZE (BP)	DISTANCE (MM)	CALCULATED BP	SIZE (BP)
	23 130			21 226			16 841
	9 416						
	6 557						
	4 361						
	2 322						
	2 027						

Discussion

1 Why was 1 µL of nuclease-free water to be added to microtube labelled 'C'?
2 Why do we incubate the restriction digests at 37°C?
3 What is the purpose of the dye?
4 What would occur if the gel electrophoresis chamber were filled with distilled water instead of TBE buffer?
5 Explain why DNA samples must be loaded at the negative end of a gel electrophoresis chamber.
6 What would occur if the electrodes in the electrophoresis chamber were reversed?

Taking it further

Investigate real-world examples of where restriction enzymes are used and how they assist in providing evidence for evolution.

ACTIVITY 10.3 Investigating biotechnological techniques

Throughout this course in human biology, you have had the opportunity to do many activities that have enabled you to inquire scientifically. This activity will allow you to apply some of those skills to investigate a particular biotechnological technique and create your own model of the process.

9780170449168

What to do

Working with a partner, or as part of a small group, select either the polymerase chain reaction or DNA sequencing for further investigation. Both techniques have a number of stages that provide logical steps to allow models to be created.

1. Use a variety of references to establish the exact sequence of steps in the technique being investigated. Collate the information about your sources in a bibliography.
2. Draw a diagram to clearly illustrate all the steps in the process.
3. A scientific model is a simplified representation of an idea or a process. Using different shapes cut out of cardboard to represent the different parts of the process, build a simple model to demonstrate how the technique you are investigating takes place. Depending on the technique being investigated, your shapes may represent various nucleotides or segments of the DNA molecule. There will be many different ways of presenting the model, so do not be surprised if yours is quite different from others' models.
4. Present and explain your model to the other members of the class, or make a video of your model and add an explanation by a voice-over or annotations.

ACTIVITY 10.4 Investigating amino acid sequencing

Haemoglobin is the protein that carries oxygen in the blood. It is found in all mammals and has the same function in each species. You would expect that it would be composed of the same sequences of amino acids. However, this is not the case. The particular protein chain we will study in this activity is composed of 146 amino acids. The numbers in the following table indicate the position of some of the amino acids in that sequence and the letters are abbreviations for the amino acids (see Table 10.3). Six different mammalian species are shown with the amino acids that are present at positions 87 to 116 in the chain.

What to do

1. Examine the table on the following page and count the number of differences in the amino acid sequences for the following pairings of species:
   - human and chimpanzee
   - human and gorilla
   - chimpanzee and gorilla
   - human and rhesus monkey
   - chimpanzee and rhesus monkey
   - gorilla and rhesus monkey
   - human and horse
   - human and kangaroo.
2. Record your data in a table.
3. Using only the data from this section of the haemoglobin molecule, rank the species in order from the one closest to humans to the one most distant.

Studying your data

1. Based on this segment of the haemoglobin molecule, which species of mammal appears to be the most closely related to humans?
2. Which animal appears to be the least closely related to humans?
3. Which of the other pairs of species show close relationships?
4. These sequences of amino acids are generally very similar but not identical. If these species were all descended from a common ancestor, how would the changes in the sequences of the different species have come about?
5. Do you think the differences in the amino acid sequences between the species would affect the function of haemoglobin?

9780170449168

Amino acid sequences in the haemoglobin of six mammalian species

AMINO ACID IN HAEMOGLOBIN	SPECIES					
	HUMAN	CHIMPANZEE	GORILLA	RHESUS	HORSE	KANGAROO
87	T	T	T	Q	A	K
88	L	L	L	L	L	L
89	S	S	S	S	S	S
90	E	E	E	E	E	E
91	L	L	L	L	L	L
92	H	H	H	H	H	H
93	C	C	C	C	C	C
94	D	D	D	D	D	D
95	K	K	K	K	K	K
96	L	L	L	L	L	L
97	H	H	H	H	H	H
98	V	V	V	V	V	V
99	D	D	D	D	D	D
100	P	P	P	P	P	P
101	E	E	E	E	E	E
102	N	N	N	N	N	N
103	F	F	F	F	F	F
104	R	R	K	K	R	K
105	L	L	L	L	L	L
106	L	L	L	L	L	L
107	Q	Q	Q	Q	Q	Q
108	N	N	N	N	N	N
109	V	V	V	V	V	I
110	L	L	L	L	L	I
111	V	V	V	V	A	V
112	C	C	C	C	L	I
113	V	V	V	V	V	C
114	L	L	L	L	V	L
115	A	A	A	A	A	A
116	H	H	H	H	R	E

In summary

Using the information from this sequence of amino acids in haemoglobin, describe the evolutionary relationships between the species in terms of the evolution of humans.

9780170449168

CHAPTER 10 SUMMARY

- Polymerase chain reaction (PCR) is used to make many copies of a section of DNA.
- PCR involves denaturing DNA to separate the strands, annealing primers to the DNA and then extension of the strands.
- Primers are short sections of DNA with a base sequence that is complementary to sections on either side of the DNA that need to be copied.
- DNA polymerase is used to add nucleotides to the growing strand. As the enzyme used needs to be able to withstand the high temperatures during denaturing, Taq polymerase from a heat-loving bacterium is used.
- Gel electrophoresis is used to separate DNA strands based on their length.
- DNA has a negative charge. Therefore, when an electric current is applied to the gel, the DNA will move away from the negative terminal. Smaller lengths of DNA will move more quickly than longer lengths. This will produce a DNA profile.
- Restriction enzymes are used prior to gel electrophoresis to cut the DNA at specific sites.
- DNA ladders are produced by running a sample with known DNA lengths through gel electrophoresis. Results from an unknown sample can be compared to the ladder, giving information about the length of the DNA strands.
- Ethidium bromide, methylene blue or DNA probes are used to visualise the DNA after gel electrophoresis.
- The sequence of nucleotides (or the bases on the nucleotides) can be determined using Sanger's method, which uses dideoxyribonucleotides to stop the extension of the DNA strand. The strands are then separated using gel electrophoresis. Knowing which base was on the dideoxyribonucleotide for each strand length allows the base sequence to be determined.
- Ethical factors such as autonomy, confidentiality, equity and privacy need to be considered when gaining and using genetic information.
- Comparative genomics is the comparison of the DNA of different species. Species with a recent common ancestor will have more similarities than those with a more distant common ancestor.
- Endogenous retroviruses (ERV) are viral sequences that become part of the organism's DNA in a gamete. Comparison of the amount and location of ERV of different species can provide evidence for common ancestors.
- Mitochondrial DNA (mtDNA) is circular DNA found in the mitochondria and is passed down from the mother. The degree of mutation can be used to determine the closeness of the relationship between species.
- The amino acid sequence of ubiquitous proteins such as cytochrome C can be compared and used to infer how recently the common ancestor of two species existed.
- Bioinformatics uses computers to describe the molecular components of a living organism. For example, it can compare the whole genome of species.

9780170449168

CHAPTER 10 GLOSSARY

Annotation Identification of genes in a DNA sequence

Bioinformatics The use of computers to describe the molecular components of living things

Comparative genomics The comparison of genome sequences of different species

Cytochrome C An iron-containing protein that can alternate between a reduced form and an oxidised form; important in the electron transport system in cellular respiration

Dideoxyribonucleotide A modified deoxyribonucleotide that lacks a hydroxyl group of the sugar component

DNA fingerprint A technique that uses the banding patterns of DNA fragments as a means of identification; a DNA fingerprint is unique to a particular individual; also called DNA profile

DNA profile *see* DNA fingerprint

DNA sequencing The determination of the precise order of nucleotides in a sample of DNA

Endogenous retrovirus (ERV) A retrovirus that has become part of an organism's genome and exists in every cell of the body

Gel electrophoresis A process used to separate charged molecules based on their size by pushing them through a gel

Genome The complete set of genetic material in a cell; an organism's complete set of DNA

Micropipette A fine pipette used to measure and transfer very small volumes of liquid

Mitochondria Structures in the cytoplasm of a cell in which the aerobic stage of respiration occurs; singular: mitochondrion

Mitochondrial DNA (mtDNA) DNA found in the mitochondria of the cells, rather than in the nucleus

Polymerase chain reaction (PCR) A technique used in molecular biology for producing multiple copies of DNA from a sample; used in DNA fingerprinting and in identifying diseases

Primer A strand of DNA or RNA that serves as a starting point for DNA replication

Restriction enzyme An enzyme that cuts strands of DNA at a specific sequence of nucleotides

Reverse transcription A process where the base sequence in RNA is copied during the synthesis of DNA

Thermocycling A process of repeated heating and cooling

Ubiquitous protein One of a group of proteins that appears to be in all species, from bacteria to humans; the small protein called ubiquitin was so named because it is present in all types of cells

Well An indentation in the gel used for gel electrophoresis

9780170449168

CHAPTER 10 REVIEW QUESTIONS

Recall

1 a What is DNA sequencing and what is it used for?
b Briefly outline the steps in building a DNA sequence.

2 a What is a 'DNA profile'?
b List two practical applications of DNA profiling.

3 a Outline the steps in the polymerase chain reaction.
b Giving an example, explain what the term 'heat stable DNA polymerase' means.
c What are some of the practical applications of the polymerase chain reaction?

4 a Define 'endogenous retroviruses'.
b How do retroviruses become endogenous?
c What is the value of endogenous retroviruses in a study of evolution?

5 a Define 'mitochondrial DNA (mtDNA)'.
b Describe how mtDNA has been used to provide evidence for evolutionary relationships between species.
c Give an example of where mtDNA has provided information about such a relationship.

6 Describe how the sequence of amino acids in proteins can be used to determine the degree of similarity between species.

7 a Define 'ubiquitous proteins'.
b Why has cytochrome C been so valuable in providing evidence for evolution? Give examples of species that contain cytochrome C.
c Besides cytochrome C, what other proteins have been used to provide evidence about relationships between species?

8 a How has bioinformatics assisted biologists in refining evolutionary relationships?
b What role has comparative genomics played in the study of evolutionary changes among organisms?

9 List the key areas that need to be ethically considered when deciding whether to gain or use genetic information.

Explain

10 One of the most frequently used ways to sequence DNA is to take advantage of the way it replicates. Explain how, if the sequence of bases on one side of a fragment of DNA is known, the sequence on the other side is known as well.

11 The polymerase chain reaction is a method of amplifying a small amount of DNA into a much larger amount. Explain the advantages of being able to do this.

12 Using an example, explain how the study of DNA in different species has added to the evidence for evolution.

13 Describe how each of the following has facilitated DNA sequencing:
a polymerase chain reaction
b gel electrophoresis
c bacterial enzymes.

9780170449168

Apply

14 When ancestral species evolve into two or more separate species, those new species would exhibit considerable similarity in their DNA. What causes the DNA to change over time? How has the information from DNA been used by scientists to speculate on the relationships between species?

15 Modern technology has provided the means to compare DNA and protein sequences. How has this changed the traditional way of looking at the relationships between humans and apes?

16 Explain why mtDNA is only of use when looking at the relationships within a species or between closely related species.

17 Explain why scientists select ubiquitous proteins for their biochemical research on the relationships between species.

18 Refer to Table 10.4, which indicates the degree of difference in the amino acids in cytochrome C between humans and some other species. Using this information, construct a family tree to illustrate a possible relationship between those species.

19 Why would scientists use a comparative study of haemoglobin in different species in a search for data to support their theories of primate evolution?

20 Imagine that you and your sister are identical twins. You have had genetic testing, and now know that you have a genetic predisposition to breast cancer. Discuss whether or not it is ethical to tell your twin about the testing results.

Extend

21 Mitochondrial Eve is a name that has been given to the woman who, when traced through the female line, is the most recent common ancestor for all living humans. The mitochondrial DNA in all humans alive today is derived from her.

- **a** How is the matrilineal line traced back to Mitochondrial Eve?
- **b** How long ago is Mitochondrial Eve believed to have lived?
- **c** In what part of the world did she live?
- **d** Does the fact that the mitochondrial DNA of all humans is derived from Mitochondrial Eve mean that she was the only human female alive at the time?
- **e** How is it possible that one woman could be the matrilineal ancestor of us all?

22 Haemoglobin and cytochrome C have been used to give support to the theory of evolution through natural selection. Scientists have similarly compared the biochemistry of universal blood proteins.

- **a** Have such studies revealed evidence for the relationships between different species?
- **b** Does such evidence imply that some species share a more recent common ancestor than other species?

9780170449168