



Molecules (Continuity and change) revision notes

How DNA is copied, how its base sequence is used to make polypeptides, and how changes to that sequence arise and can be made on purpose. The group covers DNA replication with PCR and gel electrophoresis, protein synthesis, and mutation and gene editing.

For IB Diploma Biology SL undefined. **HL only** is additional higher level content, for HL students only.

1 DNA replication **Spec D1.1**

Exact copies by semi-conservative replication **Spec D1.1.1, D1.1.2**

DNA replication is the production of exact copies of DNA with identical base sequences. It happens before a cell divides, so that each daughter cell receives a full copy of the genome.

- **Reproduction:** DNA must be copied so that it can be passed to offspring, in asexual reproduction and in the production of gametes.
- **Growth:** a multicellular organism grows by mitosis, and every new cell needs a complete copy of the DNA.
- **Tissue replacement:** cells lost through damage or wear (for example skin cells or red blood cell precursors) are replaced by dividing cells, which again need copied DNA.

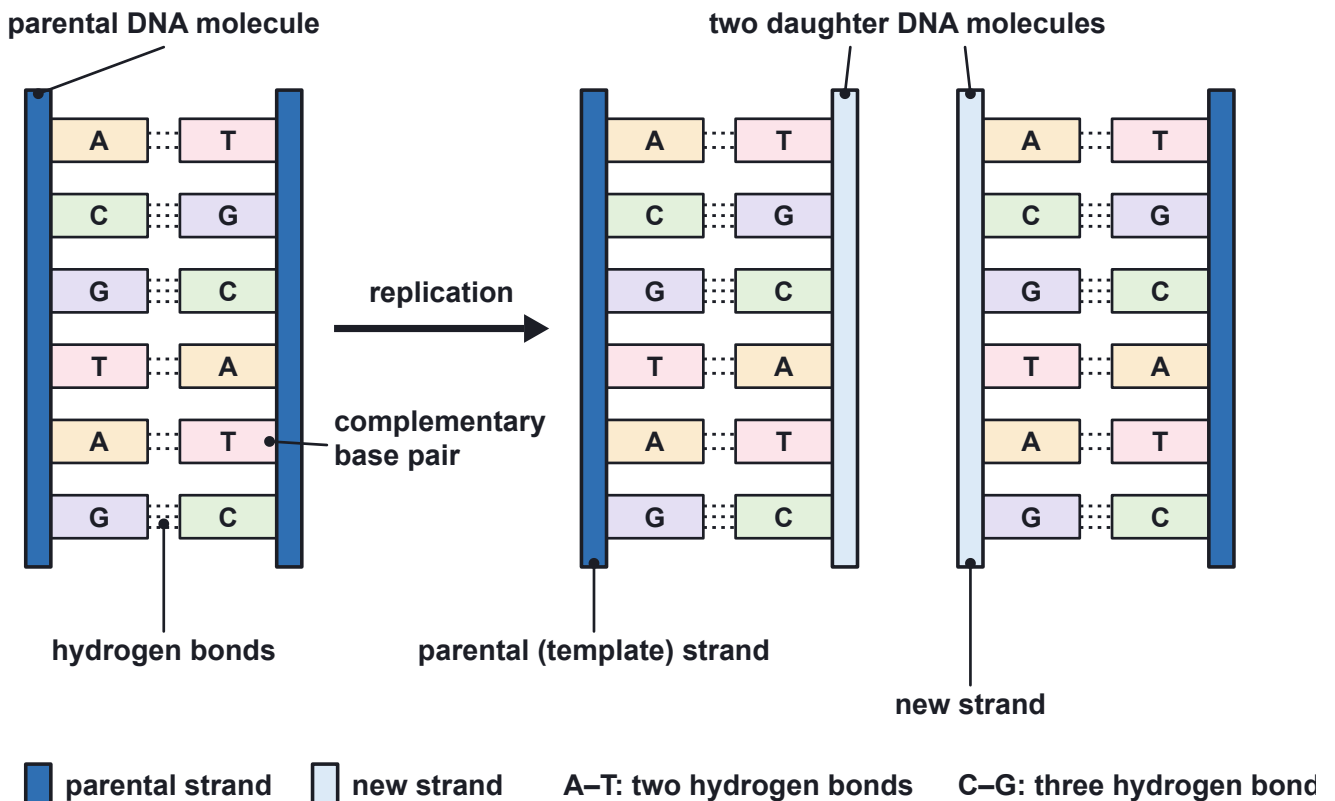
Semi-conservative replication

The two strands of the double helix separate and each one is used as a **template** for a new strand. Each of the two DNA molecules produced therefore has one strand from the original molecule (conserved) and one new strand. This is called **semi-conservative replication**.

Complementary base pairing and accuracy

Free DNA nucleotides are added to the new strand according to **complementary base pairing**: adenine pairs only with thymine (two hydrogen bonds) and cytosine pairs only with guanine (three hydrogen bonds). Because the base on the template strand decides which base is added, the new strand has exactly the sequence complementary to the template, and the new double helix is identical to the original.

A base that does not pair correctly does not fit or form the right hydrogen bonds, so mismatches are rare. This is why replication copies base sequences with a very high degree of accuracy.



[\(opens full size in a new tab\)](#)

Semi-conservative replication: each daughter molecule keeps one parental strand and gains one new strand.

Common mistake: Semi-conservative does not mean half of each strand is new. Each daughter molecule has one complete old strand and one complete new strand.

Roles of helicase and DNA polymerase Spec D1.1.3

Two enzymes are central to replication.

Helicase	Unwinds the double helix and separates the two strands by breaking the hydrogen bonds between complementary bases. Each separated strand can then act as a template.
DNA polymerase	Assembles a new strand: it brings free DNA nucleotides into position opposite their complementary bases on the template strand and links them together with covalent (phosphodiester) bonds to form the sugar-phosphate backbone.

1. Helicase moves along the DNA, unwinding it and breaking hydrogen bonds between base pairs, so the strands separate.
2. Free DNA nucleotides pair with exposed bases on each template strand (A-T, C-G).
3. DNA polymerase links the paired nucleotides into a continuous new strand.
4. Two DNA molecules result, each with one old and one new strand.

Exam tip: Say helicase breaks hydrogen bonds between bases, not covalent bonds. DNA polymerase forms covalent bonds between adjacent nucleotides of the new strand.

PCR and gel electrophoresis Spec D11.4

Polymerase chain reaction (PCR)

The **polymerase chain reaction (PCR)** makes millions of copies of a chosen DNA sequence from a tiny sample. The reaction mixture contains the sample DNA, free DNA nucleotides, two **primers** and **Taq polymerase**, in a buffer. A thermal cycler changes the temperature in repeated cycles.

One cycle of PCR

Stage	Typical temperature	What happens
Denaturation	about 95 °C	Hydrogen bonds between the strands break, so the double helix separates into single strands.
Annealing	about 50–65 °C (often about 55 °C)	Primers bind by complementary base pairing to the ends of the target sequence on each single strand.
Extension	about 72 °C	Taq polymerase adds nucleotides to the primers, building a new complementary strand.

- **Primers** are short single-stranded sequences of DNA. Two different primers are used, one for each strand, and they mark the two ends of the sequence to be copied. DNA polymerase can only add nucleotides to an existing strand, so the primer gives it a starting point.
- **Taq polymerase** is a DNA polymerase from *Thermus aquaticus*, a bacterium of hot springs. It is not denatured at 95 °C and works best at about 72 °C, so it survives every cycle and does not need to be replaced.
- **Temperature changes** drive the cycle: high temperature separates strands, a lower temperature allows primers to bind, and an intermediate temperature suits Taq polymerase.

Each cycle doubles the number of copies of the target sequence, so the amount of DNA increases exponentially. Thirty cycles typically take one to two hours.

$$\text{number of copies} = \text{starting number of copies} \times 2^n \text{ (n = number of cycles)}$$

Maths skill: copies after PCR

A sample contains 50 copies of a target DNA sequence. How many copies are present after 20 cycles of PCR, assuming each cycle doubles the DNA?

1. $2^{20} = 1\,048\,576$

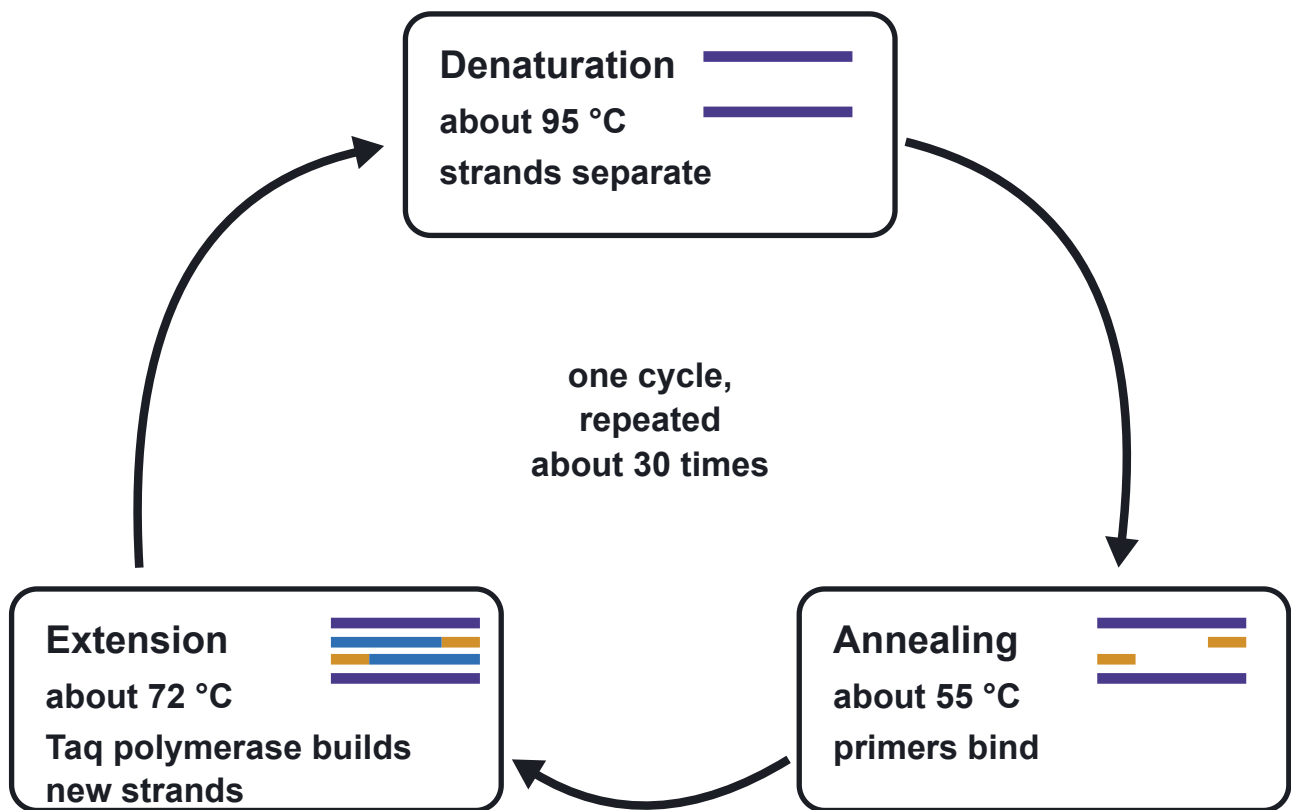
2. $50 \times 1\,048\,576 = 52\,428\,800$

Answer: About 5.2×10^7 copies.

Gel electrophoresis

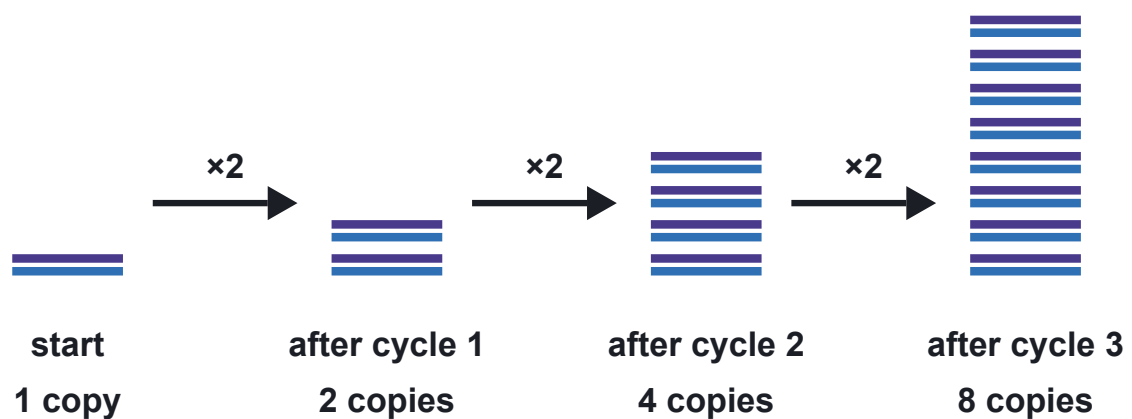
Gel electrophoresis separates DNA fragments according to their length. Samples are placed in wells at one end of an agarose gel and a voltage is applied. DNA has a **negative charge** (from its phosphate groups), so fragments move through the gel towards the positive electrode. The gel acts like a mesh: **smaller fragments move faster and travel further** than larger ones in the same time.

A DNA ladder (a mixture of fragments of known lengths) is run in one lane so that fragment sizes can be estimated. The DNA is stained so that the bands can be seen.



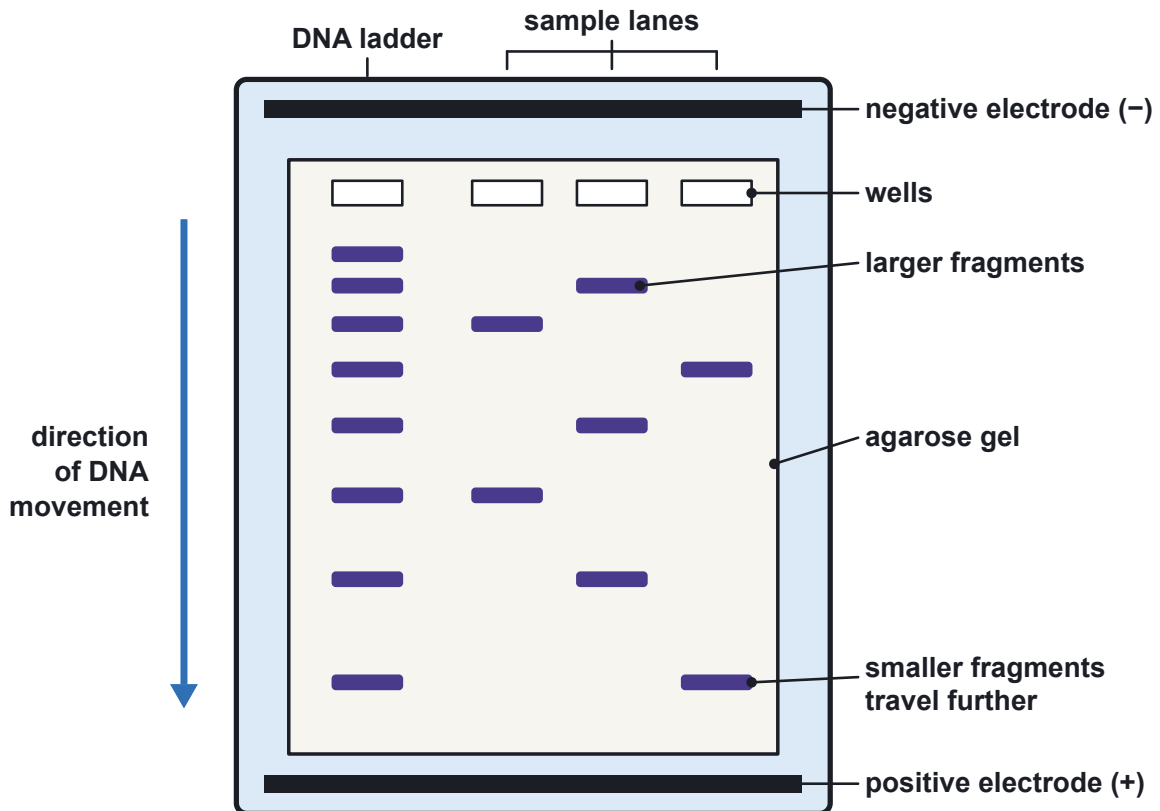
— template DNA
 — primer
 — new strand

copies of the target sequence double every cycle



[\(opens full size in a new tab\)](#)

One cycle of **PCR**. Each cycle doubles the number of copies of the target sequence.



[\(opens full size in a new tab\)](#)

DNA is negatively charged, so it moves towards the **positive electrode**; **smaller fragments travel further**.

Practical skill: When reading a gel, compare each band with the ladder lane. Bands at the same distance from the wells in different lanes are fragments of the same length.

Applications of PCR and gel electrophoresis Spec D1.1.5

Because PCR can amplify a chosen sequence from a tiny sample and gel electrophoresis can separate fragments by length, the two techniques have a broad range of applications.

- **DNA profiling** for paternity and forensic investigations.
- **Detecting pathogens:** PCR detects DNA (or RNA copied into DNA) of a virus or bacterium in a patient sample, as in tests for viral infections.
- **Genetic screening:** testing for alleles that cause genetic conditions.
- **Research:** amplifying genes for sequencing or for comparing species.

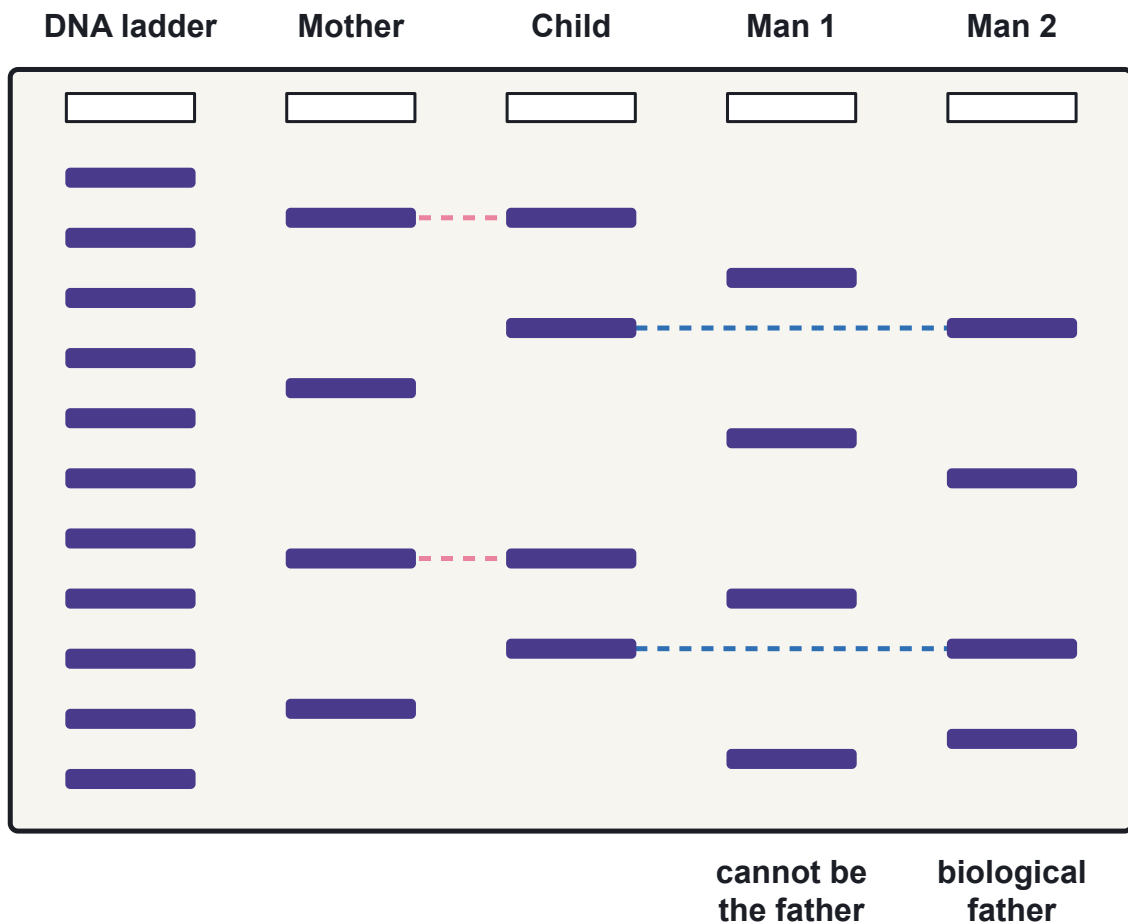
DNA profiling

Some regions of the genome contain **short tandem repeats (STRs)**: a short base sequence repeated many times. The number of repeats at each region varies between individuals, so the length of the region varies. In **DNA profiling**, several of these regions are amplified by PCR and the fragments are separated by gel electrophoresis, giving a pattern of bands (a DNA profile) that is almost unique to each person.

Using DNA profiles

Use	How profiles are compared
Paternity	A child inherits one allele of each region from each parent. Every band in the child's profile must match a band in the mother's or in the father's profile. A man whose profile lacks the non-maternal bands cannot be the father.
Forensic investigation	DNA from blood, hair or skin cells at a crime scene is amplified and profiled, then compared with profiles from suspects. A match links a person to the scene; a mismatch excludes them.

Exam tip: Nature of science: reliability is enhanced by increasing the number of measurements. In DNA profiling, using more markers (more STR regions) reduces the probability of a false match between two different people.



--- child's band matches a band from the mother

--- child's band matches a band from Man 2

[\(opens full size in a new tab\)](#)

Paternity test: each band in the child comes from one parent, so Man 2 is the father and Man 1 is excluded.

Quick check

- 1 . Why is DNA replication described as semi-conservative?
- 2 . What is the role of helicase?
- 3 . Why is Taq polymerase used in PCR?
- 4 . Why do smaller DNA fragments travel further in gel electrophoresis?

2 Protein synthesis Spec D1.2

Transcription and gene expression Spec D1.2.1, D1.2.2, D1.2.3, D1.2.4

Transcription is the synthesis of RNA using a DNA template. Only one of the two DNA strands, the **template strand**, is transcribed. The other strand is the coding strand: the RNA has the same base sequence as the coding strand, except that uracil replaces thymine.

Roles of RNA polymerase

1. RNA polymerase binds to DNA at the start of a gene.
2. It unwinds the double helix and separates the two strands by breaking hydrogen bonds between bases, over a short stretch of DNA.
3. Free RNA nucleotides pair with exposed bases on the template strand by complementary base pairing.
4. RNA polymerase links the RNA nucleotides with covalent bonds to form a single strand of RNA.
5. Behind the polymerase, the RNA separates from the template and the DNA strands re-form their hydrogen bonds and rewind.

Hydrogen bonding and complementary base pairing

Each RNA nucleotide is held in place by **hydrogen bonds** to the complementary base on the template strand. The pairs are: **adenine (A) on the DNA template with uracil (U) on the RNA**, thymine (T) on the template with adenine (A), cytosine (C) with guanine (G) and guanine (G) with cytosine (C). The hydrogen bonds are weak, so the RNA can separate from the template once its nucleotides are linked.

Stability of DNA templates

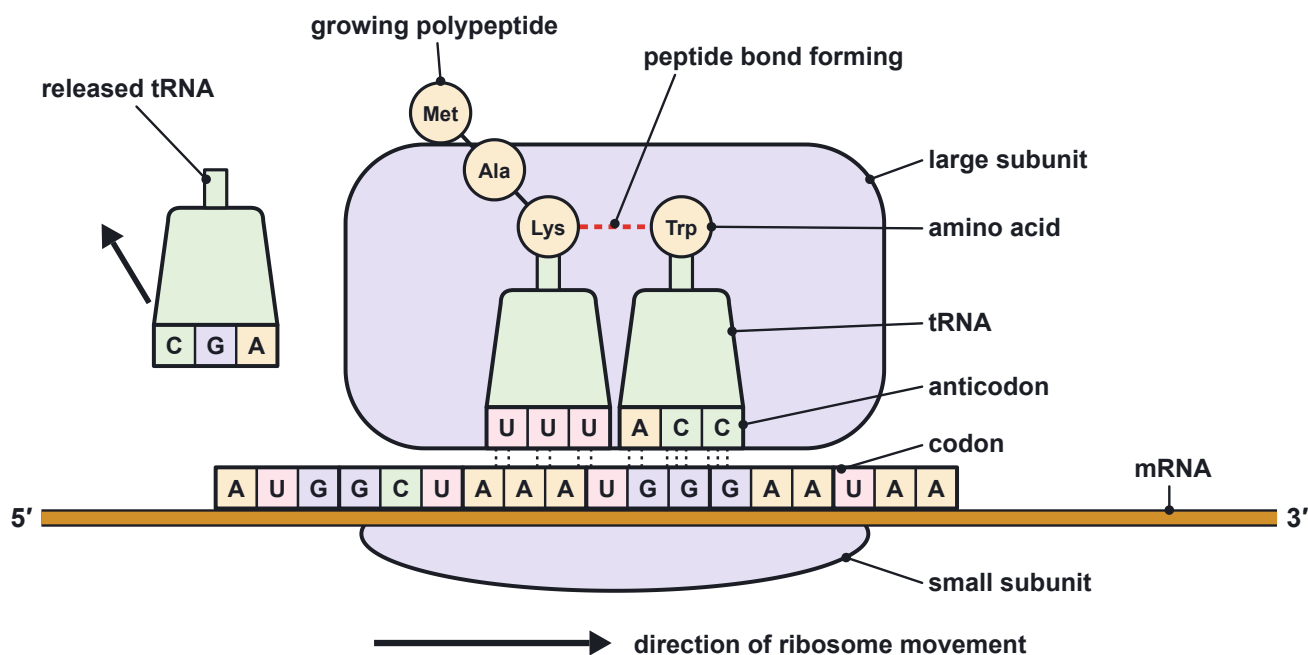
A single DNA strand can be used as a template for transcribing a base sequence without the DNA base sequence changing: only hydrogen bonds are broken and then re-formed. In **somatic cells that do not divide**, such as many neurones, the same DNA must be transcribed again and again, so its sequences must be conserved throughout the life of the cell.

Transcription and gene expression

Not all genes in a cell are expressed at any given time. Transcription is the **first stage of gene expression**, so it is a key stage at which the expression of a gene can be switched on or off. A gene that is not transcribed produces no mRNA and so no polypeptide.

2. A second tRNA, whose anticodon is complementary to the next codon, binds alongside it, bringing its amino acid.
3. A **peptide bond** forms between the amino acid on the second tRNA and the polypeptide, so the polypeptide is transferred to the second tRNA and becomes one amino acid longer.
4. The ribosome moves along the mRNA by one codon (three bases). The first tRNA, now without an amino acid, is released and can collect another amino acid.
5. The next codon is exposed and the cycle repeats, adding one amino acid at a time until a stop codon is reached.

This stepwise movement of the ribosome along the mRNA, one codon at a time, means the amino acids are linked in exactly the order set by the codons.



[\(opens full size in a new tab\)](#)

Elongation: **two tRNAs bind to the large subunit at once**, so a peptide bond can join the polypeptide to the next amino acid.

Common mistake: Codons are on mRNA and anticodons are on tRNA. Do not describe the anticodon as part of the mRNA or the codon as part of the tRNA.

The genetic code Spec D1.2.8, D1.2.9

Why a triplet code

There are four bases in mRNA and 20 amino acids in proteins. A code of single bases would give 4 combinations, and pairs of bases would give $4^2 = 16$, which is not enough. Triplets give $4^3 = 64$ codons, which is more than enough for 20 amino acids plus start and stop signals.

Degeneracy	More than one codon can code for the same amino acid. For example, GAA and GAG both code for glutamic acid. Of the 64 codons, 61 code for amino acids and 3 are stop codons.
Universality	The same codons code for the same amino acids in almost all living organisms, with only minor variations. This is evidence for common ancestry and allows genes to be transferred between species.
Start codon	AUG, which codes for methionine and marks where translation begins.
Stop codons	UAA, UAG and UGA, which do not code for an amino acid and signal the end of translation.

Using a table of mRNA codons

Codon tables give the amino acid for each mRNA codon. Read the mRNA from the start codon in groups of three, without overlapping, and look up each codon. If you are given DNA, first work out the mRNA: it is complementary to the template strand, and the same as the coding strand with U in place of T.

Part of the mRNA codon table (practice extract)			
Codon	Amino acid	Codon	Amino acid
AUG	Met (start)	GAA	Glu
GCU	Ala	GAG	Glu
AAA	Lys	GUG	Val
UGG	Trp	UUU	Phe
CAU	His	UAA	stop

Maths skill: deducing an amino acid sequence

The template strand of a short DNA sequence reads TAC CGA TTT ACC CTT ATT. Deduce the amino acid sequence.

1. Write the complementary mRNA (A→U, T→A, C→G, G→C): AUG GCU AAA UGG GAA UAA.
2. Look up each codon: AUG = Met, GCU = Ala, AAA = Lys, UGG = Trp, GAA = Glu, UAA = stop.

Answer: Met - Ala - Lys - Trp - Glu (then stop).

Exam tip: Check whether a sequence is DNA (contains T) or mRNA (contains U) before using a codon table. Codon tables are always written for mRNA.

Mutations that change protein structure Spec D1.2.11

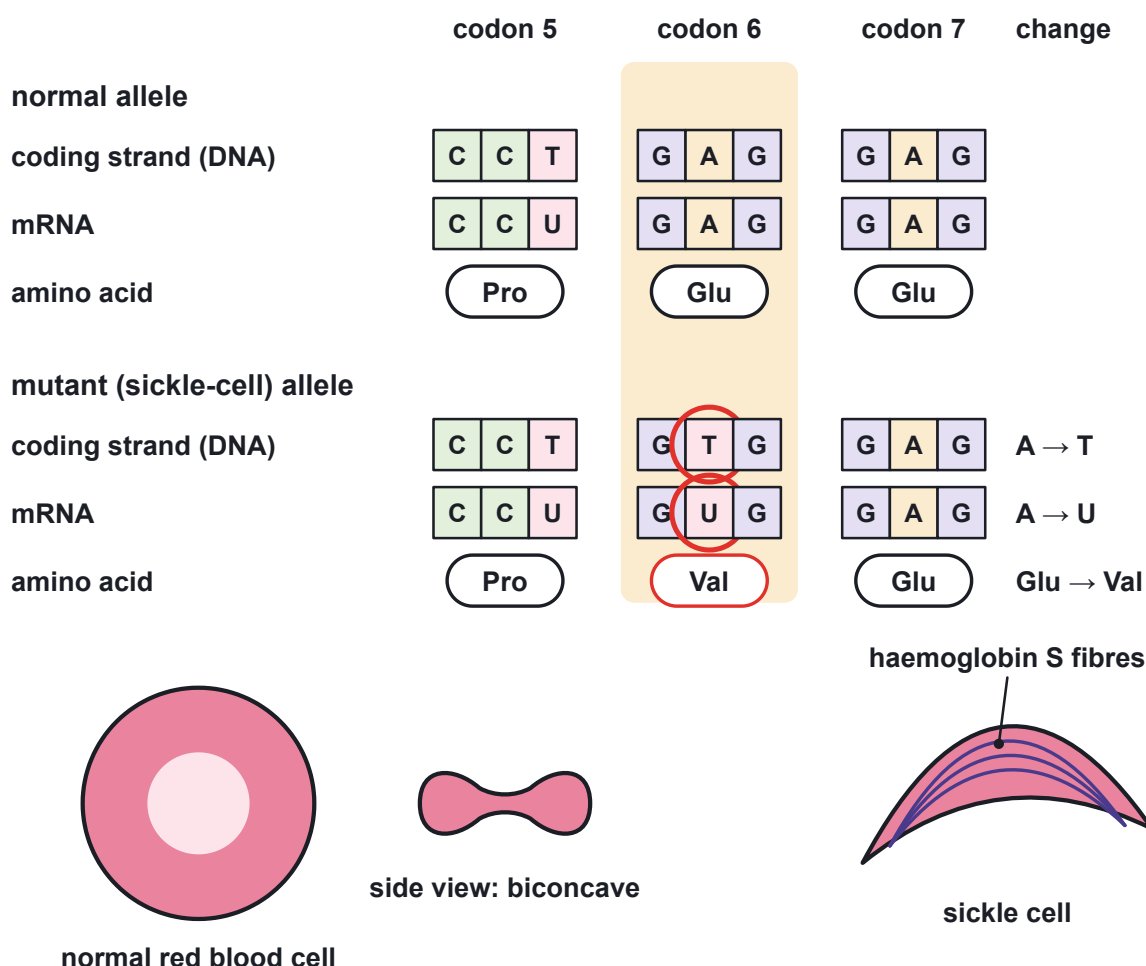
The amino acid sequence of a polypeptide decides how it folds and therefore its three-dimensional structure. A change in the base sequence of a gene can change the codons in the mRNA, and so the amino acid sequence, the folding and the function of the protein.

Example: sickle-cell anaemia

1. The gene for the beta chain of haemoglobin has a **base substitution**: on the coding strand, A is replaced by T (GAG becomes GTG).
2. The mRNA codon becomes **GUG instead of GAG**.

- At the sixth amino acid of the beta chain, **valine replaces glutamic acid**. The resulting form is haemoglobin S.
- Glutamic acid is charged and hydrophilic; valine is non-polar and hydrophobic. At low oxygen concentrations, haemoglobin S molecules stick together into long fibres.
- The fibres distort red blood cells into a rigid sickle shape. Sick cells can block capillaries and carry less oxygen, and they are broken down faster, causing anaemia.

A change to just one base of the gene is enough to alter the protein. This shows how closely protein structure depends on base sequence.



[\(opens full size in a new tab\)](#)

Sickle-cell anaemia: one base substitution (GAG → GTG) changes **glutamic acid to valine** in haemoglobin.

Exam tip: In a sickle-cell answer, give the base change, the codon change (GAG → GUG), the amino acid change (Glu → Val) and the effect on haemoglobin and red blood cells. Each is a separate marking point.

Quick check

- Which base on RNA pairs with adenine on the DNA template strand?
- Why must DNA in non-dividing somatic cells stay unchanged by transcription?
- What is meant by the degeneracy of the genetic code?
- How many tRNAs can bind to the large ribosomal subunit at once?

3 Mutation and gene editing Spec D1.3

Substitutions, insertions and deletions Spec D1.3.1, D1.3.2, D1.3.3

A **gene mutation** is a structural change to a gene at the molecular level: a change to its base sequence. There are three types.

Types of gene mutation		
Type	Change	Example (coding strand)
Substitution	One base is replaced by a different base.	GAG → GTG
Insertion	One or more extra nucleotides are added.	GAG CTA → GAA GCT A...
Deletion	One or more nucleotides are lost.	GAG CTA → GGC TA...

Consequences of base substitutions

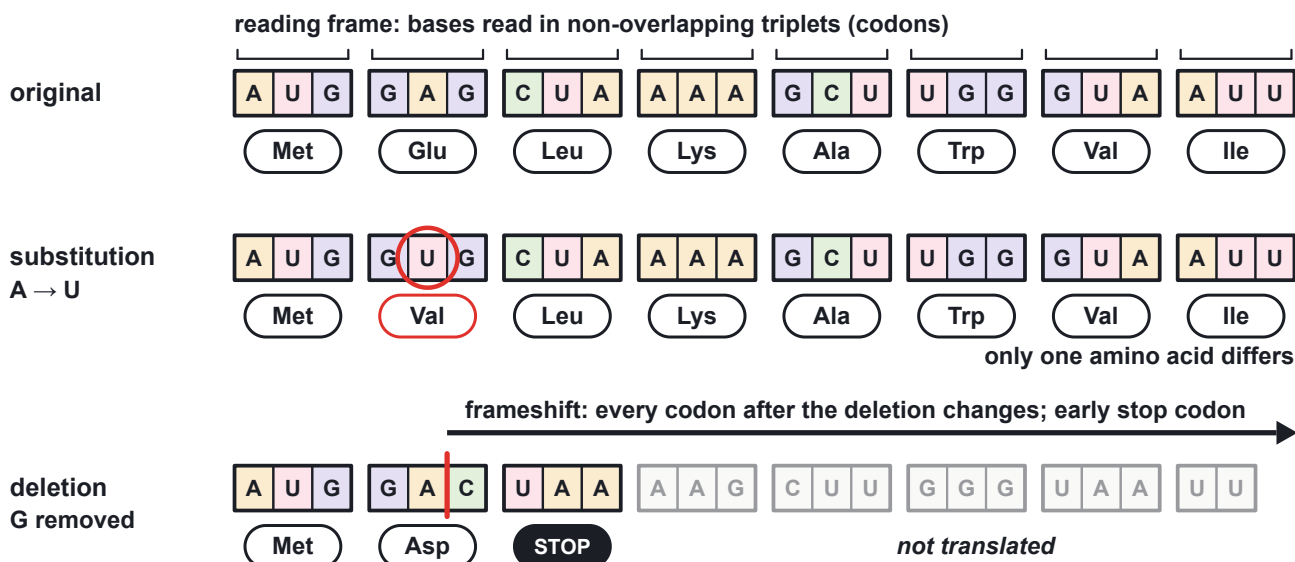
Single-nucleotide polymorphisms (SNPs) are positions in the genome where a single base differs between individuals. They are the result of base substitution mutations. A substitution changes only one codon, and because of the **degeneracy of the genetic code** it may or may not change a single amino acid:

- **Same amino acid:** the new codon codes for the same amino acid (for example GAA → GAG, both glutamic acid). The polypeptide is unchanged.
- **Different amino acid:** one amino acid in the polypeptide is changed (for example GAG → GUG, glutamic acid → valine). The effect may be small or, as in sickle-cell anaemia, large.
- **Stop codon:** the codon becomes a stop codon (for example UAC → UAA), so the polypeptide is cut short and is usually non-functional.

Consequences of insertions and deletions

Because mRNA is read in non-overlapping triplets, inserting or deleting one or two nucleotides shifts the reading frame. This **frameshift** changes every codon after the mutation, so the amino acid sequence from that point is altered and a premature stop codon often appears. The polypeptide is very likely to cease to function.

Inserting or deleting a multiple of three nucleotides does not cause a frameshift, but adds or removes amino acids. Major insertions or deletions of many nucleotides also make it likely that the polypeptide stops functioning.



[\(opens full size in a new tab\)](#)

A substitution changes one codon at most; a deletion causes a **frameshift** that changes every codon after it.

Exam tip: Linking question: how does variation in subunit composition of polymers contribute to function? A single changed nucleotide in DNA can change one amino acid in a protein and alter its function.

Causes of mutation and randomness Spec D1.3.4, D1.3.5

Causes of gene mutation

A **mutagen** is an agent that increases the rate of mutation.

Causes of gene mutation		
Cause	Examples	How it causes mutation
Chemical mutagens	Benzo[a]pyrene and nitrosamines in tobacco smoke; mustard gas	React with bases or bind to DNA, causing bases to pair incorrectly during replication.
Mutagenic radiation	Ultraviolet (UV) light, X-rays, gamma rays, alpha particles from radioactive isotopes such as radon	UV causes adjacent pyrimidine bases (such as two thymines) to bond to each other; high-energy radiation ionises molecules and can break DNA strands.
Errors in DNA replication	A mismatched nucleotide not corrected by proofreading	A wrong base is copied into the new strand and then into every later copy.
Errors in DNA repair	Damaged DNA repaired incorrectly	The repaired region has a changed base sequence.

Randomness in mutation

Mutations are **random**. They can occur anywhere in the base sequences of a genome, although some bases have a higher probability of mutating than others (for example, a cytosine with a methyl group added, often where C is followed by G, can lose an amino group and become thymine, so C → T substitutions are especially common at these sites). There is **no known natural mechanism for making a deliberate change to a particular base** with the purpose of changing a trait. An organism cannot produce a mutation because it would be useful.

Common mistake: Do not write that organisms mutate in order to adapt. Mutations happen at random; natural selection then acts on the variation they produce.

Germ and somatic cells, and mutation as a source of variation Spec D1.3.6, D1.3.7

Germ cells and somatic cells

Consequences of mutation in different cells		
	Germ cells	Somatic cells
What they are	Cells that produce gametes (and the gametes themselves)	All other body cells
Inherited?	Yes: a mutated gene can be passed to offspring	No: not passed to offspring
Cells affected	Every cell of an offspring that inherits it	Only the mutated cell and the cells produced from it by mitosis
Possible consequence	Inherited genetic disease, or new variation in the population	Cancer , if mutations affect genes that control cell division

Cancer develops when mutations in a somatic cell cause it to divide repeatedly without control, forming a tumour. Usually several mutations in the same cell line are needed, which is one reason cancer becomes more common with age and with exposure to mutagens.

Mutation as a source of genetic variation

Gene mutation is the **original source of all genetic variation**: it is the only process that creates new alleles. Meiosis and fertilisation shuffle existing alleles into new combinations, but cannot create new ones. Most mutations are either harmful or neutral for an individual organism, but in a species over the long term they are **essential for evolution by natural selection**, because selection can only act on variation that already exists.

Exam tip: Nature of science: commercial genetic tests can give information about potential future health and disease risk. Without expert interpretation, this information could be misunderstood and cause needless worry or false reassurance.

Exam tip: Linking question: how can natural selection lead to both a reduction in variation and an increase in biological diversity? Mutation supplies the variation; selection removes some alleles within a population but can drive populations apart into new species.

Quick check

1. What is a single-nucleotide polymorphism (SNP)?
2. Why does deleting one base usually have a larger effect than substituting one base?
3. Give one chemical mutagen and one form of mutagenic radiation.
4. Why are somatic mutations not inherited?

Quick check answers

DNA replication

1. Each new DNA molecule contains one strand from the original molecule and one newly synthesised strand.
2. It unwinds the double helix and breaks hydrogen bonds between bases so the strands separate.
3. It comes from a hot-spring bacterium and is not denatured at 95 °C, so it survives the denaturation step of every cycle.
4. They pass more easily through the mesh of the gel, so they move faster towards the positive electrode.

Protein synthesis

1. Uracil.
2. The same sequences must be transcribed throughout the life of the cell, so they must be conserved.
3. More than one codon can code for the same amino acid.
4. Two.

Mutation and gene editing

1. A position in the genome where a single base differs between individuals, resulting from a base substitution mutation.
2. It shifts the reading frame, changing every codon after the mutation.
3. Benzo[a]pyrene (in tobacco smoke); ultraviolet light (or X-rays, gamma rays).
4. They are not in germ cells, so they are not passed into gametes.