

Mutation and gene editing

Revision notes • IB Biology D1.3: substitutions, insertions and deletions, causes of mutation, germ and somatic cells, and HL knockout, CRISPR and conserved sequences.

For IB Diploma Biology, Standard Level and Higher Level (first assessment 2025). **HL only** is additional higher level content, for HL students only.

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

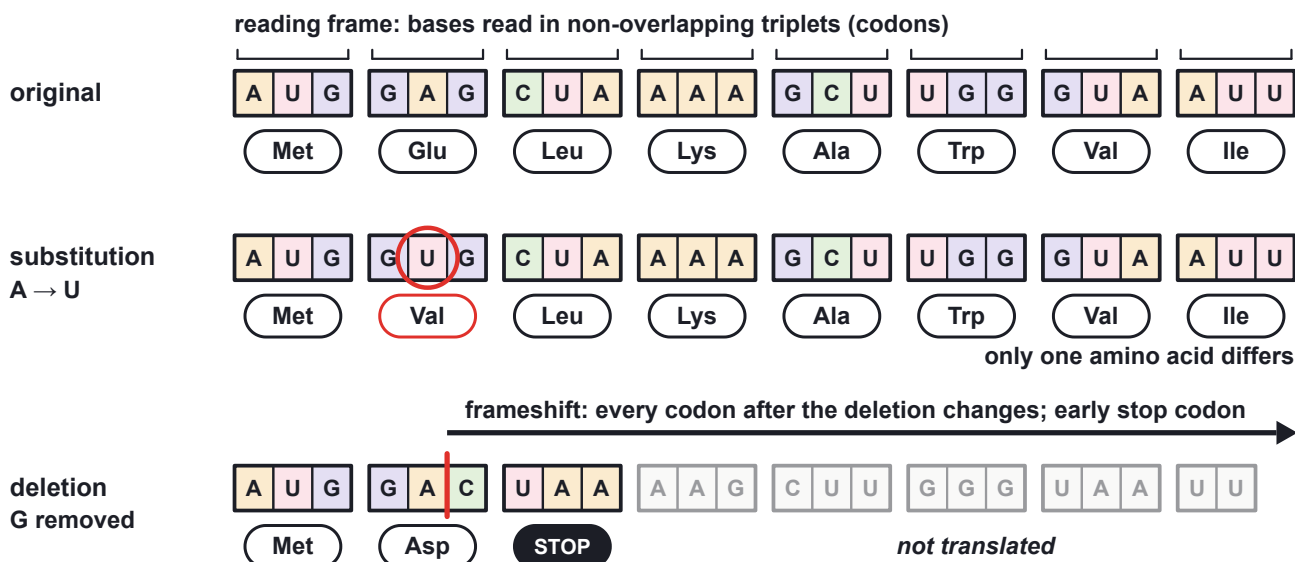
Single-nucleotide polymorphisms (SNPs) are positions in the genome where a single base differs between individuals. They are the result of base substitution mutations. A substitution changes only one codon, and because of the **degeneracy of the genetic code** it may or may not change a single amino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

Gene knockout Spec D1.3.8 HL only

Gene knockout is a technique for investigating the function of a gene by changing it to make it inoperative. The gene's base sequence is altered, or part of it removed, so that no functional product is made.

1. Choose a gene whose function is unknown.
2. Make the gene inoperative in an organism (a knockout organism).
3. Compare the phenotype of the knockout organism with that of a normal (wild-type) organism.
4. Any difference, such as in development, behaviour or metabolism, suggests what the gene normally does.

Details of knockout techniques are not required. Students should know that **libraries of knockout organisms** are available for some species used as models in research. For example, collections of yeast strains exist with nearly every gene knocked out one at a time, and international projects are building a

knockout mouse for each mouse gene. Researchers can obtain an organism with a particular gene knocked out instead of making it themselves.

Practical skill: A knockout experiment needs a control: wild-type organisms of the same strain, raised in the same conditions, so that any difference can be attributed to the missing gene.

CRISPR-Cas9 gene editing Spec D1.3.9 HL only

How CRISPR-Cas9 editing works

1. A **guide RNA** is made with a base sequence complementary to the target sequence in the gene to be edited.
2. The guide RNA is combined with **Cas9**, an enzyme that cuts DNA (an endonuclease).
3. The guide RNA binds by complementary base pairing to the target sequence, bringing Cas9 to exactly that site.
4. Cas9 cuts both strands of the DNA at the target site.
5. The cell repairs the break. Repair often adds or removes a few nucleotides, which disables (knocks out) the gene. If a DNA template with a desired sequence is supplied, the cell can use it during repair, inserting or correcting a sequence.

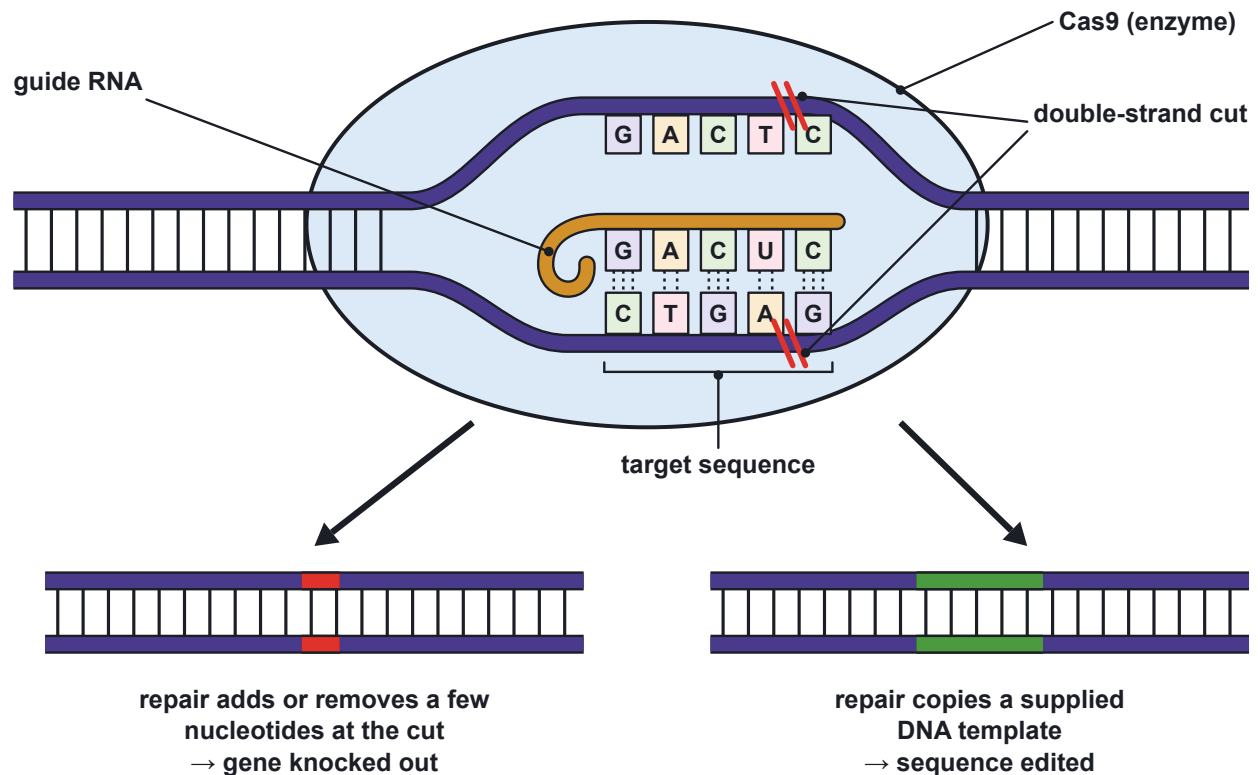
The role of the CRISPR-Cas system in prokaryotes is not required.

Example: treating sickle-cell disease

The first approved CRISPR-Cas9 therapy, exagamglogene autotemcel (exa-cel, sold as Casgevy), treats sickle-cell disease. It was authorised in the UK by the MHRA on 16 November 2023 and in the USA by the FDA on 8 December 2023, for patients aged 12 and over. Blood stem cells are collected from the patient's blood and edited outside the body. Cas9 cuts an enhancer (a regulatory sequence) of the gene *BCL11A*, so red blood cell precursors make much less of the BCL11A protein. BCL11A normally switches off the genes for **fetal haemoglobin** after birth. The patient receives chemotherapy to clear the old bone marrow, then the edited cells are infused. They produce red blood cells containing fetal haemoglobin, which does not form fibres, and most treated patients in the trials had no severe pain crises for at least a year.

Ethics and regulation

Certain potential uses of CRISPR raise ethical issues that must be addressed before implementation, especially editing embryos or germ cells, because changes would be inherited by future generations without their consent, and editing for enhancement rather than treating disease. Scientists across the world are subject to **different regulatory systems**, so there is an **international effort to harmonise regulation** of genome editing technologies such as CRISPR.



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

CRISPR-Cas9: the guide RNA takes Cas9 to the target sequence, Cas9 makes a **double-strand cut** and the cell repairs it.

Conserved and highly conserved sequences Spec D1.3.10 HL only

Conserved sequence	A base sequence (or amino acid sequence) that is identical or similar across a species or a group of species.
Highly conserved sequence	A sequence that has remained identical or similar over long periods of evolution, so it is shared by distantly related groups.

Examples of highly conserved genes include those for histone proteins, which package DNA, and for ribosomal RNA, which are similar in organisms as different as yeast and humans.

Hypotheses to account for conservation

Two hypotheses	
Hypothesis	Explanation
Functional requirements for the gene products	The protein or RNA must have a very precise structure to work, for example because it binds many other molecules. Almost any mutation makes it less functional, so individuals carrying it are selected against and the mutation is removed from the population. The sequence stays the same because changes are eliminated, not because they do not happen.
Slower rates of mutation	Some sequences may mutate less often than others, for example because of their position in the genome or more effective repair, so fewer changes arise in the first place.

For protein-coding genes the two hypotheses make different predictions. If functional requirements are the cause, base substitutions that do not change the amino acid (possible because the code is

degenerate) should still build up at a normal rate, while substitutions that change the amino acid are rare. If the mutation rate is slower, both kinds of change should be rare. In mammals, histone genes show almost no amino acid changes but a normal rate of the silent changes, which supports the functional-requirements hypothesis for these genes.

Exam tip: Name both hypotheses when asked to account for conserved sequences: functional requirements of the gene product (with selection removing changes) and a slower rate of mutation.

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?
5. **HL only** What does the guide RNA do in CRISPR-Cas9 editing?

Quick check answers

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.
- 5 . It binds by complementary base pairing to the target DNA sequence and brings Cas9 to it.