MEGA LECTURE

# IB DP BIOLOGY

## Theme D: Continuity and Change

## D1.3 Mutation and Gene Editing

Revision Notes · Standard and Higher Level

**Fahad H. Ahmad**

+92 323 509 4443 | Megalecture.com

Original notes prepared for the IB Diploma Programme Biology course (first assessment 2025)

# What the syllabus requires

D1.3 explains how the base sequence of DNA can change, what those changes do to a protein and an organism, and how biologists now deliberately rewrite DNA. Use this checklist to confirm you can meet every requirement before the exam. Items marked **(HL)** are examined at Higher Level only.

| Understanding | You should be able to... |
|---|---|
| What a mutation is | Define a gene mutation as a change in the base (nucleotide) sequence of DNA, and distinguish spontaneous from induced mutations. |
| Point mutations | Classify a base substitution as silent, missense or nonsense and predict its effect on the polypeptide. |
| Insertions and deletions | Explain how adding or removing bases shifts the reading frame (frameshift) and why this is usually severe. |
| Sickle-cell anaemia | Trace a single base substitution from DNA to phenotype, and account for the link with malaria. |
| Consequences of mutation | Recognise that mutations may be harmful, neutral or beneficial, and are the ultimate source of new variation. |
| Germ-line vs somatic | State that only germ-line (gamete) mutations are inherited, and relate somatic mutation to cancer. |
| Causes of mutation | Describe mutagens (ionising and UV radiation, chemicals) and errors in DNA replication. |
| Mutation rate and neutrality | Explain why mutation rates are low and why most mutations across a genome are neutral. |
| Gene editing (HL) | Outline CRISPR-Cas9 and gene knockout, and discuss their uses and ethical issues. |

> **Exam note:** The molecular understanding of mutation is common to SL and HL. Gene editing with CRISPR-Cas9 and gene knockout as a research tool are Higher Level content, flagged (HL) throughout these notes.

# 1. What a gene mutation is

A **gene mutation** is a change in the base (nucleotide) sequence of a gene. Because the sequence of bases is the information that specifies the order of amino acids in a polypeptide, changing even one base can change the protein a cell makes. Mutations are the original source of all new alleles: every version of a gene that exists today arose, at some point in the past, as a mutation.

Mutation should be kept separate from two other sources of variation. Meiosis and sexual reproduction *shuffle* existing alleles into new combinations but create no new alleles; only mutation supplies genuinely new sequence.

## Spontaneous vs induced

- **Spontaneous mutations** arise without any external cause, mainly from rare, natural errors during DNA replication, or from the chemical instability of the bases themselves (for example a base

occasionally pairing incorrectly).

- **Induced mutations** are caused by an external agent called a **mutagen** → ionising radiation, ultraviolet light or certain chemicals. Mutagens raise the mutation rate above its natural background level.

Both kinds change the same thing (the base sequence); they differ only in what triggered the change. Whatever the cause, the position at which a mutation strikes the genome is **random** → a mutagen cannot direct a change to a gene that would be useful.

> **Key definition:** A gene mutation is a change to the base sequence of DNA. A chromosome mutation (change in chromosome number or structure) is a larger, separate category and is not what D1.3 means by a gene mutation.

# 2. Types of point mutation

A **point mutation** affects a single base. The two broad classes behave very differently, because the genetic code is read in non-overlapping groups of three bases (codons).

## Base substitution

One base is replaced by a different base. Only the single codon containing that base is altered, so the reading frame downstream is undisturbed. A substitution has one of three possible effects on the polypeptide:

- **Silent** → the new codon still codes for the same amino acid, because the genetic code is **degenerate** (most amino acids have several codons). The polypeptide is unchanged.
- **Missense** → the new codon codes for a *different* amino acid. The effect ranges from negligible to severe depending on how important that amino acid is to the protein's shape and function.
- **Nonsense** → the new codon becomes a **stop** codon. Translation ends early, giving a short, truncated polypeptide that is usually non-functional.

## Insertion and deletion (frameshift)

If one or a few bases are **inserted** or **deleted** (and the number is not a multiple of three), every codon from that point onward is regrouped. This **frameshift** changes essentially all downstream amino acids and very often creates a premature stop codon. Frameshifts are therefore usually far more damaging than a single substitution.

## Effects on the polypeptide, at a glance

| Mutation | What changes in DNA/mRNA | Effect on polypeptide | Example |
|---|---|---|---|
| Silent substitution | One base swapped; codon still specifies the same amino acid | None → identical protein | GAA → GAG, both = glutamic acid |
| Missense substitution | One base swapped; codon now specifies a different amino acid | One amino acid changed; may alter folding/function | GAG (Glu) → GUG (Val), sickle cell |
| Nonsense substitution | One base swapped; codon becomes a stop codon | Truncated, usually non-functional protein | CAG (Gln) → UAG (stop) |

| Mutation | What changes in DNA/mRNA | Effect on polypeptide | Example |
|---|---|---|---|
| Insertion (frameshift) | Extra base(s) added; reading frame shifts | Most downstream amino acids changed; often early stop | ...ACU\|GCU... → ...ACA\|UGC\|U... |
| Deletion (frameshift) | Base(s) removed; reading frame shifts | Most downstream amino acids changed; often early stop | ...AUG\|CAU... → ...AUG\|AU... |

**Why degeneracy matters:** Because the code is degenerate, many substitutions (especially at the third base of a codon) are silent. This is a major reason why so many real mutations have no visible effect.

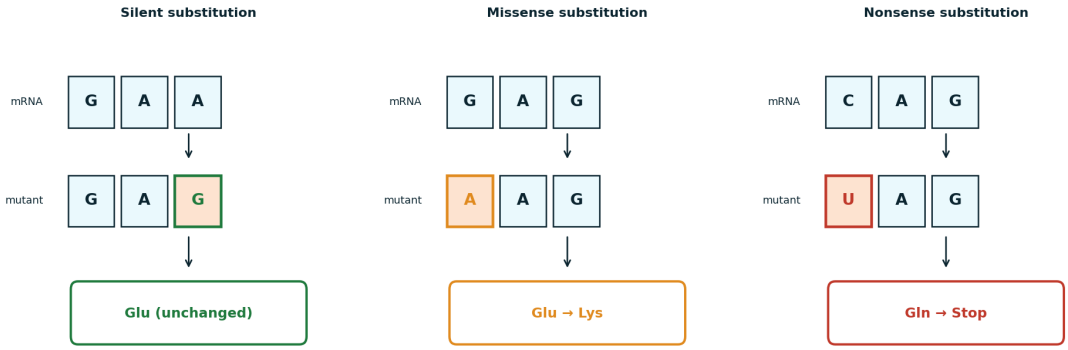


*Figure 1. Three point-mutation outcomes read from the codon table: GAA→GAG is silent (Glu unchanged); GAG→AAG is missense (Glu→Lys); CAG→UAG is nonsense (Gln→STOP). Every translation is looked up from a codon table, not assumed.*

# 3. Sickle-cell anaemia: a worked substitution

Sickle-cell anaemia is the textbook example of a single base substitution whose effect can be followed all the way from DNA to whole-organism phenotype. Haemoglobin in adults (HbA) is built from two α-globin and two β-globin chains. The mutation lies in the gene for the β-globin chain.

## Following the change, step by step

- **DNA:** at the sixth codon of the β-globin gene, the coding sequence changes from GAG to GTG (a single A → T substitution on the coding strand).
- **mRNA:** the transcribed codon changes from GAG to GUG.
- **Amino acid:** GAG codes for glutamic acid; GUG codes for valine. So the sixth amino acid changes from glutamic acid to valine (a missense mutation).
- **Protein:** glutamic acid is polar/negatively charged; valine is nonpolar (hydrophobic). This places a hydrophobic patch on the surface of the altered haemoglobin (HbS).
- **Cell:** when oxygen is low, HbS molecules stick together and polymerise into long fibres that distort red blood cells into a rigid sickle shape.
- **Organism:** sickled cells block small blood vessels (pain, tissue damage) and are destroyed quickly, causing anaemia.

## Genotypes and heterozygote advantage

The sickle-cell allele shows codominance. Individuals who are **homozygous** for the sickle allele ($Hb^SHb^S$) have sickle-cell anaemia. **Heterozygotes** ($Hb^AHb^S$) have sickle-cell *trait*: they make both normal and sickle haemoglobin and are usually healthy.

Crucially, heterozygotes are more resistant to **malaria** → the malarial parasite fares poorly in cells that contain some HbS. In regions where malaria is common, heterozygotes survive and reproduce better than either homozygote (the healthy but malaria-susceptible $Hb^AHb^A$, and the anaemic $Hb^SHb^S$). This is a classic case of **heterozygote advantage**, and it explains why an otherwise harmful allele remains common in malarial parts of the world.

> **Exam link:** Sickle cell ties D1.3 to natural selection: the same allele is harmful, neutral or beneficial depending on genotype and environment. State the environment (malarial region) explicitly when you explain the advantage.

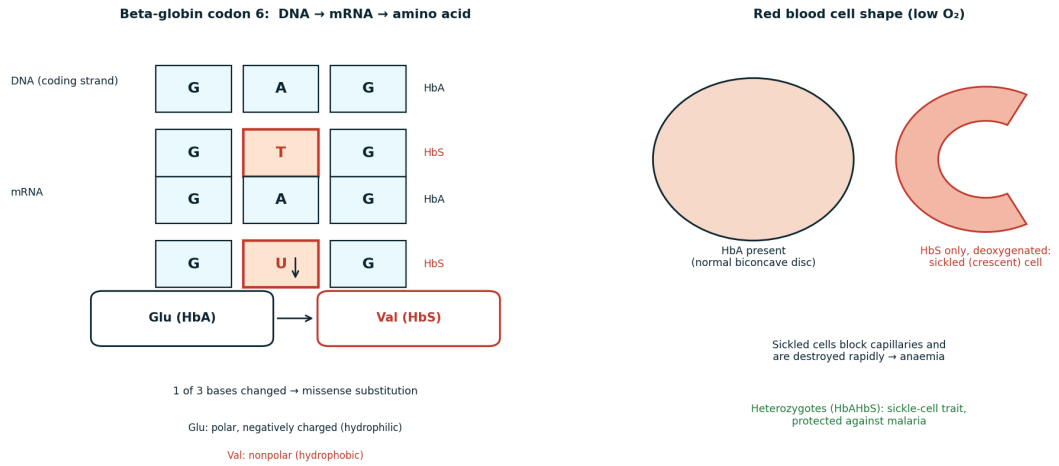


*Figure 2. Sickle-cell substitution audited base by base: coding-strand codon 6 GAG→GTG gives mRNA GAG→GUG and amino acid Glu→Val (missense) - the single base change that produces HbS.*

# 4. Consequences: harmful, neutral or beneficial

Whether a mutation matters depends on where it lands and what the organism needs. The three outcomes are best thought of as a spectrum.

| Outcome | Typical cause | Biological significance |
|---|---|---|
| Harmful | Missense in a critical residue, nonsense, or a frameshift that destroys protein function | Genetic disease; reduced survival or fertility; usually selected against |
| Neutral | Silent substitution, or a change in non-coding DNA / an unimportant residue | No effect on phenotype; accumulates as harmless variation |
| Beneficial | A change that improves function or fit to the environment (rare) | Raises survival/reproduction; favoured by natural selection |

## Mutation as the raw material of evolution

Beneficial mutations are rare, but they are indispensable: mutation is the **only** process that produces genuinely new alleles. Natural selection cannot create variation, it can only act on variation that already exists. Mutation supplies that variation; sexual reproduction then recombines it. Without mutation, evolution

would eventually run out of raw material.

> **Common misconception:** Mutations are not "trying" to help the organism, and they do not happen more often when they would be useful. They are random with respect to need; selection, acting afterwards, is what makes evolution look directed.

# 5. Germ-line vs somatic mutations, and cancer

A mutation is only passed to the next generation if it is present in the cells that form gametes.

| | Germ-line (germ cell) mutation | Somatic (body cell) mutation |
|---|---|---|
| Where it occurs | In gamete-forming cells / gametes | In any non-reproductive body cell |
| Inherited? | Yes → passed to offspring; present in every cell of the child | No → affects only the descendants of that one cell |
| Consequence | Can cause heritable genetic conditions | Local effect; can contribute to cancer |

Only germ-line mutations are inherited. A somatic mutation, however serious for the individual, dies with the body cells it belongs to and is never passed on.

## Mutation and cancer (conceptual)

Cancer arises when somatic mutations disturb the genes that control the cell cycle and cell division. Two categories of gene are involved:

- **Proto-oncogenes** normally *promote* cell division in a controlled way. A mutation can convert one into an **oncogene** that drives division even when it should be switched off → like an accelerator stuck down.
- **Tumour-suppressor genes** normally *restrain* division or trigger repair/cell death. Losing their function removes a brake.

Cancer typically requires **several** such mutations to accumulate in one cell line over time, which is why the risk rises with age and with exposure to mutagens (carcinogens). Because these are somatic mutations, most cancers are not inherited, although an inherited (germ-line) mutation in one of these genes can raise a person's lifetime risk.

> **Precise wording:** Say a mutation "converts a proto-oncogene into an oncogene" or "inactivates a tumour-suppressor gene." Avoid the loose claim that "cancer is caused by a mutation" without naming the type of gene affected.

# 6. Causes of mutation: mutagens and replication errors

Mutations arise both from within (replication errors) and from without (mutagens).

## Replication errors

DNA polymerase occasionally inserts the wrong base. Proofreading and DNA repair correct almost all of these, but the rare error that escapes repair becomes a **spontaneous** mutation. This is why the natural mutation rate is very low but never zero.

## Mutagens (induced mutations)

| Mutagen | How it damages DNA |
|---|---|
| Ionising radiation (X-rays, gamma rays, alpha/beta particles) | Carries enough energy to break DNA strands and remove electrons, causing base changes and breaks in the sugar-phosphate backbone. |
| Ultraviolet (UV) light | Absorbed by DNA bases; causes adjacent pyrimidines (thymines) to bond together (thymine dimers), distorting the helix and causing errors when the DNA is copied. Linked to skin cancer. |
| Chemical mutagens | Molecules such as certain components of tobacco smoke and some industrial chemicals react with bases, altering their pairing or inserting between them and disrupting replication. |

Notice the pattern: mutagens either change a base chemically (leading to a substitution) or damage the strand so that replication introduces an error. Either way they raise the frequency of mutation above the spontaneous background.

> **Radiation note:** Ionising radiation is more penetrating and can reach DNA deep in the body, whereas UV is absorbed near the surface → which is why UV damage shows up mainly as skin cancer.

# 7. Gene editing: CRISPR-Cas9 (HL)

**(HL)** Gene editing lets biologists change a chosen DNA sequence deliberately, rather than waiting for random mutation. The most widely used tool, **CRISPR-Cas9**, was adapted from a natural bacterial defence system against viruses.

## How CRISPR-Cas9 works, conceptually

- A short **guide RNA** is designed so that its bases are complementary to the target DNA sequence to be edited.
- The guide RNA is joined to the enzyme **Cas9**. The guide directs Cas9 to the matching sequence in the genome → precise targeting.
- **Cas9 cuts** both strands of the DNA at that exact location.
- The cell repairs the cut. Repair can **disable** the gene (small errors at the join) or, if a template DNA is supplied, **insert or correct** a chosen sequence → the edit.

## Gene knockout as a research tool

**(HL)** A powerful use of editing is the **gene knockout**: a gene is deliberately disabled so that it no longer produces a functional product. Biologists then observe what goes wrong in the organism and infer the normal **function** of that gene from the phenotype that appears when it is missing. CRISPR-Cas9 makes knockouts fast and precise, which is why it has become a standard investigative method.

## Uses and ethical issues

| Potential uses | Ethical / practical concerns |
|---|---|
| Correcting alleles that cause genetic disease; engineering disease-resistant or higher-yield crops; making gene-knockout organisms for research; developing new therapies | Off-target cuts causing unintended mutations; editing germ-line cells changes future generations who cannot consent; questions of access, consent and equity; the prospect of non-therapeutic "enhancement" |

> **Somatic vs germ-line editing:** Editing somatic cells affects only the treated patient. Editing germ-line cells (embryos, gametes) would be inherited by all descendants → which is the sharpest ethical dividing line in the debate.

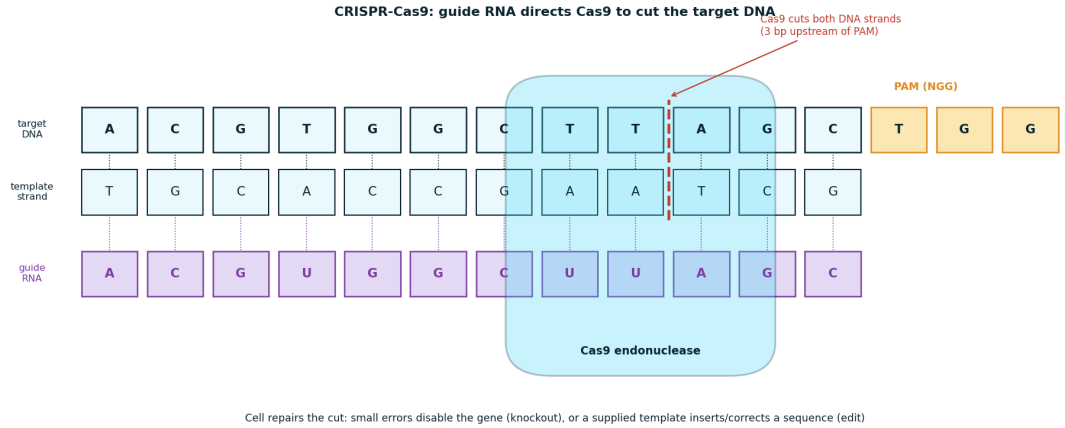


*Figure 3. CRISPR-Cas9: the 12-base guide RNA pairs correctly with the target template strand at 12/12 positions (confirmed by the A−T/G−C pairing rule), and Cas9 cuts 3 bp upstream of the PAM.*

# 8. Mutation rate and why most mutations are neutral

The **mutation rate** is the number of new mutations per gene (or per base) per generation. It is kept very low by the accuracy of DNA replication, by proofreading and by DNA repair. A low rate is itself advantageous: too many mutations would destroy far more useful information than they create.

## Why most mutations are neutral

- **Most of the genome does not code for protein.** A change in non-coding DNA frequently has no effect on any protein and so no effect on phenotype.
- **The genetic code is degenerate.** Many substitutions within coding regions are silent, giving the same amino acid.
- **Many amino-acid changes are tolerated.** A different amino acid in a non-critical part of a protein often leaves its function essentially unchanged.

Because harmful mutations are removed by selection and beneficial ones are rare, the large majority of the mutations that persist in a population are **neutral**. These neutral differences accumulate over time and are the basis of much of the natural variation between individuals.

# 9. Worked examples and exam practice

A short reference table of the codons used below (read from mRNA, 5′ to 3′):

| mRNA codon(s) | Amino acid |
|---|---|
| GAA, GAG | Glutamic acid (Glu) |
| GUU, GUG | Valine (Val) |
| CUU, CUC, CUA, CUG | Leucine (Leu) |

| mRNA codon(s) | Amino acid |
|---|---|
| UGG | Tryptophan (Trp) |
| CAG | Glutamine (Gln) |
| UAA, UAG, UGA | STOP (no amino acid) |

**Worked example 1 - classify a substitution**

A codon in an mRNA changes from CUU to CUC. Classify the mutation and predict the effect on the protein.

Both CUU and CUC code for leucine. The amino acid is unchanged, so this is a **silent substitution** (the code is degenerate). The polypeptide is identical and there is no effect on the phenotype.

Pitfall avoided: a base *has* changed, so it is tempting to say the protein changes. Always translate both codons before deciding.

**Worked example 2 - missense vs nonsense**

An mRNA codon changes from UGG to UGA. Classify the mutation and predict the effect.

UGG codes for tryptophan; UGA is a **stop** codon. Translation stops early, so this is a **nonsense mutation** giving a **truncated** polypeptide that is very likely non-functional.

Compare: if UGG had instead changed to, say, CGG (arginine), a different amino acid would be inserted → that would be a missense mutation, usually less severe than a nonsense one.

**Worked example 3 - explain sickle cell from the DNA**

The sixth codon of the β-globin coding sequence changes from GAG to GTG. Explain, step by step, how this produces sickle-cell anaemia.

GAG → GTG on the coding strand gives mRNA GAG → GUG. This changes the sixth amino acid from glutamic acid (polar) to valine (nonpolar) → a **missense** mutation.

The hydrophobic valine lets haemoglobin molecules (HbS) stick together and polymerise when oxygen is low, distorting red cells into a sickle shape. Sickled cells block capillaries and are destroyed rapidly, causing anaemia.

Only individuals homozygous for the allele ($Hb^SHb^S$) have the disease; heterozygotes are protected against malaria (heterozygote advantage).

**Worked example 4 - distinguish a frameshift**

A coding sequence reads ...AUG CAU GGA... A single base (the C) is deleted. Show why the effect is far greater than a substitution.

Deleting the C regroups every codon that follows: ...AUG CAU GGA... becomes ...AUG AUG GA... The reading frame has shifted, so all downstream codons - and therefore all downstream amino acids - are changed, and a premature stop codon is likely.

This is a **frameshift**: unlike a substitution (which alters one codon), a single-base deletion or insertion can wreck the entire rest of the protein.

## Watch out for these

- **Silent does not mean "no base changed".** A base changed, but the degenerate code gives the same amino acid → no change in protein.
- **Frameshift severity.** Adding or deleting three bases (a whole codon) does *not* cause a frameshift; adding or deleting one or two bases does.
- **Only germ-line mutations are inherited.** A dramatic somatic mutation is never passed to offspring.
- **Mutations are random with respect to need.** A mutagen raises the rate but cannot steer a mutation toward a useful gene.

## Quick reference

| Term | One-line statement |
|---|---|
| Gene mutation | A change in the base sequence of DNA |
| Spontaneous / induced | Arises from replication error / caused by a mutagen |
| Silent | Substitution → same amino acid (degenerate code) |
| Missense | Substitution → different amino acid |
| Nonsense | Substitution → premature stop codon → truncated protein |
| Frameshift | Insertion/deletion (not a multiple of 3) shifts the reading frame |
| Germ-line vs somatic | Only germ-line mutations are inherited |
| Mutagens | Ionising and UV radiation; chemical mutagens |
| CRISPR-Cas9 (HL) | Guide RNA targets the sequence; Cas9 cuts; the cell's repair makes the edit |
| Most mutations | Neutral → non-coding DNA, silent changes, tolerated residues |

## Test yourself

Attempt these without notes; full answers follow.

1. Define a gene mutation and distinguish a spontaneous mutation from an induced one.

2. An mRNA codon changes from GAA to GAG. State the type of mutation and its effect on the protein.

3. Explain why a frameshift mutation is usually more harmful than a single base substitution.

4. In sickle-cell anaemia, why does an allele that causes a serious disease remain common in some parts of the world?

5. Distinguish a germ-line mutation from a somatic mutation, giving one consequence of each.

6. Name two mutagens and state how each damages DNA.

7. Outline how CRISPR-Cas9 edits a chosen gene. (HL)

8. Explain why the majority of mutations across a genome are neutral.

## Answers

**1.** A gene mutation is a change in the base (nucleotide) sequence of DNA. A spontaneous mutation arises naturally (chiefly from a rare, uncorrected replication error) with no external cause; an induced mutation is caused by a mutagen such as radiation or a chemical, which raises the mutation rate.

**2.** GAA and GAG both code for glutamic acid, so this is a **silent** substitution: the amino acid, and hence the protein, is unchanged.

**3.** A substitution alters only the single codon it lies in. An insertion or deletion (not a multiple of three) shifts the reading frame, so every codon downstream is regrouped, changing most of the remaining amino acids and often introducing a premature stop codon → the protein is far more likely to be non-functional.

**4.** Heterozygotes ($Hb^AHb^S$) are protected against malaria while remaining largely healthy. In malarial regions this heterozygote advantage means the allele is favoured by natural selection and so persists, even though homozygotes ($Hb^SHb^S$) suffer sickle-cell anaemia.

**5.** A germ-line mutation occurs in gamete-forming cells and is inherited by offspring (present in every cell of the child, so it can cause a heritable genetic condition). A somatic mutation occurs in a body cell, is not inherited, and affects only that cell's descendants (for example contributing to a cancer).

**6.** Any two, e.g. *ionising radiation* (X-rays/gamma) breaks DNA strands and alters bases by removing electrons; *UV light* makes adjacent thymines bond into dimers that distort the helix and cause copying errors; *chemical mutagens* react with bases and change their pairing during replication.

**7.** (HL) A guide RNA is made complementary to the target DNA sequence and attached to Cas9; the guide directs Cas9 to that sequence; Cas9 cuts both DNA strands at that point; the cell's repair machinery then either disables the gene or inserts a supplied sequence, producing the edit.

**8.** Most of the genome is non-coding, so many changes affect no protein; within genes the degenerate code makes many substitutions silent; and many amino-acid changes fall in non-critical positions that leave function unchanged. Harmful mutations are removed by selection and beneficial ones are rare, so the mutations that remain are predominantly neutral.