Mutations and gene editing

D1.3 · SL and HL

One changed base can have several different outcomes

Follow the sequence, then test the consequence

D1.3 shared: substitutions, insertions/deletions, mutagens and spontaneous errors, randomness, inheritance and genetic variation.

AHL: knockout evidence, CRISPR-Cas9 and conserved sequences. Distinguish a molecular change from a measured biological outcome.

Sequences and numerical examples are invented unless a published study is named.

Distinguish DNA sequence, damage and gene activity

Observation	**Interpretation**
Established sequence change	Mutation; may pass to descendant cells
Damage repaired to original sequence	No lasting mutation at that position
More transcript from unchanged DNA	Changed expression; not automatically mutation
Gene editing	A deliberate change at a chosen DNA sequence

Compare three substitutions in the same coding frame

Coding DNA	Initial polypeptide / signal
Original ATG GAA GCT TAA	→ Met – Glu – Ala; stop
Synonymous ATG GAG GCT TAA	→ Met – Glu – Ala; stop
Replacement ATG GTA GCT TAA	→ Met – Val – Ala; stop
Early stop ATG TAA GCT TAA	→ Met; stop

A substituted base can preserve, replace or stop a codon instruction.

Convert the coding strand before using an RNA codon table

1. **Coding DNA.** ATG GAA GCT TAA, in the stated direction and frame.
2. **mRNA.** AUG GAA GCU UAA: Met–Glu–Ala, then stop.
3. **Interpretation.** Four codons produce three residues here. UAA supplies no amino acid; this is coding DNA, not the complementary template.

Do not infer function from the size of a change alone

New coding triplet	Meaning and limits
GAG → mRNA GAG	Glu remains Glu: synonymous at this codon
GTA → mRNA GUA	Glu becomes Val; function depends on the residue's role
TAA → mRNA UAA	Early stop; observed RNA/product amounts can also change
All three	No length change; no downstream frameshift from the **substitution** itself

A SNP identifies variation at one nucleotide position

Separate sequence variation from a disease claim

An observed nucleotide variant is not automatically harmful. Compare sequence, molecular effect and trait evidence.

Degeneracy allows different codons to specify one amino acid. Synonymous describes that coding outcome.

An unchanged amino acid does not by itself prove every aspect of processing, expression and cell function is unchanged.

Regroup from the same start after deleting base 4

1. **Original mRNA.** AUG GAA GCU UAA: Met–Glu–Ala, then stop.
2. **Remove the first G after AUG.** AUG AAG CUU AA: Met–Lys–Leu and an incomplete group.
3. **Conclusion.** The downstream **reading frame** changes. No complete stop is shown; additional sequence is required to locate termination.

Compare a whole-codon deletion with a frameshift

1. **Remove GAA.** AUG GCU UAA: Met–Ala, then stop; one residue missing.
2. **Frame rule.** A coding insertion/deletion not divisible by three shifts subsequent grouping. Three or six bases preserve the downstream frame; large changes can still disrupt a functional domain.
3. **Limits.** A boundary-crossing deletion can alter the joined codon. A preserved frame does not guarantee normal function; a shifted frame need not change every later residue.

A non-coding change can still alter protein production

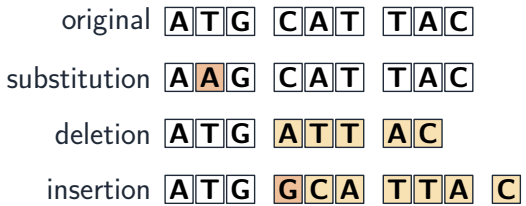
Name the region and test a proposed mechanism

A promoter change can reduce transcription initiation without altering the coding triplets.

A splice-related change can alter which RNA sequence is retained. Non-coding does not mean biologically irrelevant.

Measure DNA sequence, transcript and protein separately. Lower protein abundance can also reflect altered breakdown or later regulation.

Three kinds of gene mutation



red = changed base; shaded = shifted by a frameshift (deletion or insertion)

> Frameshifts usually change much more

Connect a cause to damage, copying and repair

Source	Possible route to a sequence change
Spontaneous errors	Replication or repair can establish a new base sequence
Ultraviolet	Damage involving neighbouring pyrimidines can disrupt normal copying
Ionizing radiation	Direct DNA damage or reactive products can lead to changes
Benzo[a]pyrene metabolites	Can bind DNA and disrupt copying/repair; tobacco smoke is one source

Calculate a final resistant fraction without calling it a rate

1. **Invented sample.** 24 resistant colonies among 12,000 viable colonies.
2. **Fraction.** $24/12000 = 0.002 = 0.20\%$.
3. **Limit.** This is a final phenotype fraction. Survival, selection, generations and expansion of early mutants affect it; new mutation events per copying event were not counted.

Antibiotic exposure can select variants already present

Selection changes representation; it does not specify the needed base

A resistant lineage may arise before exposure. Removal of susceptible competitors increases its representation.

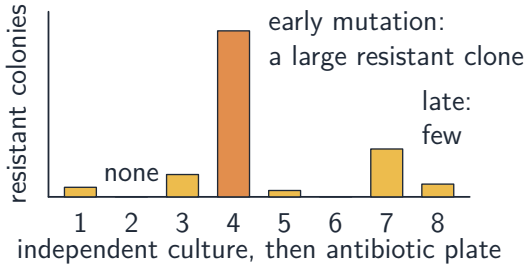
An early mutation can leave many descendants; a late mutation fewer. Independent cultures can therefore vary in resistant counts.

Random with respect to need does not require equal mutation probabilities at all genome sites. Hotspots, repair and exposure can bias rates.

Design the comparison before interpreting the resistant count

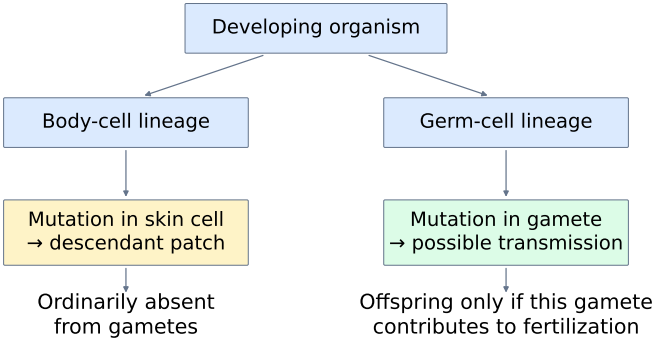
1. **Match.** Starting population, medium, exposure, generations and detection method.
2. **Repeat.** Use independent cultures and measure survival as well as the final resistant fraction.
3. **Infer carefully.** More resistant cells after exposure alone cannot distinguish newly induced mutation from selection and multiplication of earlier variants.

Mutations arise at random



An early mutation, before any antibiotic

Follow mutation through the body or through a gamete



Cell ancestry determines the route of transmission.

Separate cell lineage, transmission and effect

Case	Supported conclusion
Skin-cell mutation	Can pass to descendant skin cells; ordinarily absent from gametes
Germ-cell lineage	May enter a gamete; inheritance requires that gamete to contribute
Later embryonic change	Can create mosaicism across descendant tissues
Growth-control changes	Can contribute to cancer; one mutation is not automatically sufficient

Inherited does not mean beneficial; acquired does not always mean somatic

Identify where the mutation arose

A harmful variant can be somatic or germline. A germline variant can have no detectable effect in the stated conditions.

A change arising during one person's life may occur in a germ-cell lineage, so timing alone does not establish inheritance.

In ordinary sexual reproduction, a mutation restricted to skin cells does not enter an offspring through those skin cells.

Assign each process its actual contribution

Process	Contribution
Mutation	Creates new sequence variants / alleles
Recombination	Rearranges existing alleles into new combinations
Selection	Frequency changes associated with reproductive-success differences
Genetic drift	Chance sampling changes allele frequencies

A useful variant in one environment can carry a cost in another

Fitness describes performance under stated conditions

A resistance variant may help during antibiotic exposure yet have a cost without the antibiotic.

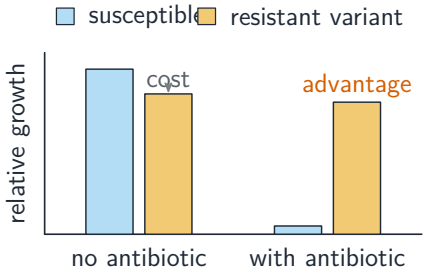
Many changes are neutral in a particular context; changes disrupting important functions are often harmful. No single percentage applies universally.

Heritable variation supplies material for evolution across generations. This does not mean an individual needs a damaging mutation to stay healthy.

Evaluate a commercial genetic-test risk statement

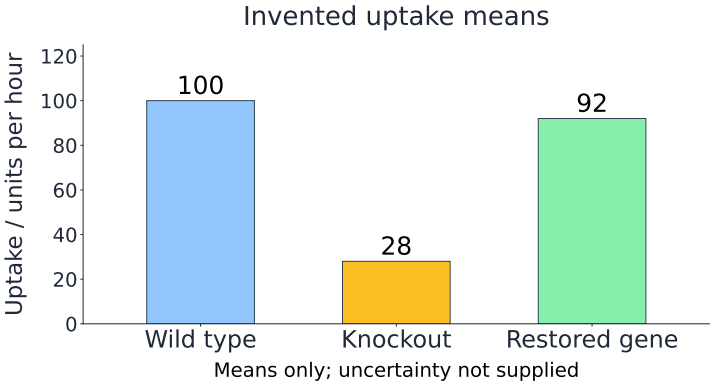
1. **Invented comparison.** Risk rises from 2 in 1,000 to 4 in 1,000.
2. **Two descriptions.** Doubles, a 100% relative increase; absolute increase $= 0.4\% - 0.2\% = 0.2$ percentage points.
3. **Limits.** Association is not a complete individual prediction. Check tested variants, population evidence, effect size and environmental influences; a negative limited panel excludes neither all variants nor all risk.

A mutation's value depends on context



Resistance carries a cost when it is not needed

Compare loss of function with restoration of a functional copy



A knockout and rescue test contribution to a measured process.

Calculate the loss, then limit the mechanistic conclusion

1. **Invented means.** Wild type 100; knockout 28; restored gene 92 units per hour.
2. **Reduction.** $100 - 28 = 72$; $72/100 \times 100 = 72\%$ of wild-type uptake lost.
3. **Conclusion.** The gene contributes to uptake under these conditions. It may act indirectly; the result does not prove its protein binds the nutrient.

A knockout needs controls and a suitable outcome measure

No visible phenotype is not proof of no function

Compare matched genetic backgrounds and conditions. Independent lines and restoration help test unintended-change explanations.

Replicates and variation are needed to assess precision; bars showing means alone do not establish statistical significance.

Another gene may compensate, or a role may appear only under another condition. Knockout libraries support systematic testing in model organisms.

Do not collapse targeting, cutting and outcome into one event

Stage	Meaning and check
Target recognition	Guide RNA complementarity and appropriate PAM help locate DNA
Cas9 cutting	DNA strands cut; intended final sequence not yet guaranteed
Cellular repair	May create insertions/deletions or other changes; verify actual sequence
Functional evaluation	Check relevant cells/alleles, unintended sites and persistent biological effect

Coding disruption and regulatory editing have different targets

Name what was changed before naming the benefit

Small repaired insertions or deletions can disrupt a coding frame, but not every repair disables a gene.

A non-coding target can change gene expression without directly replacing the disease-associated coding base.

An intended-site sequence result does not rule out off-target changes or prove a useful clinical outcome.

Read the denominator and endpoint in a published application

1. **Named example.** Frangoul and colleagues (2024), exa-cel for severe sickle-cell disease: edited blood stem cells returned after conditioning.
2. **Mechanism.** Disrupt an erythroid-specific BCL11A enhancer to increase fetal haemoglobin; not direct correction of the sickle-cell HBB **substitution**.
3. **Evidence and limits.** 44 treated; 30 evaluable at cutoff; $29/30 \approx 97\%$ free of severe vaso-occlusive crises for at least 12 consecutive months. Single-group results are not a guarantee of lifelong cure or zero risk.

Evaluate an application through affected people and consequences

Somatic and potentially inherited edits raise different issues

State likely benefit, uncertainty, consent, fair access and who bears risk. Future descendants cannot consent to inherited changes.

Distinguish treatment from selecting preferred traits. Include animal welfare and alternatives where relevant.

Countries can regulate applications differently; international cooperation seeks to harmonize regulation. Technical possibility alone does not establish permission.

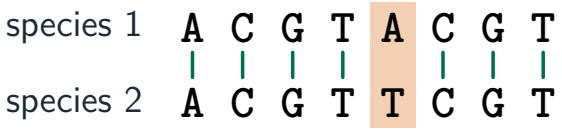
Calculate identity, then compare possible explanations

1. **Invented alignment.** ACGTACGT compared with ACGTTCGT: seven matching positions, no gaps.
2. **Identity.** $7/8 \times 100 = 87.5\%$ DNA identity.
3. **Limit.** Similarity alone does not identify a trait or prove mutation is impossible. Distinguish functional constraint, local mutation rate and divergence time.

Use complementary evidence for a conserved sequence

Evidence	Interpretation
Purifying selection	Disruptive variants leave fewer descendants; sequence retained
Matched neutral comparisons	Help test local-rate and time explanations
Knockout / altered sequence	Tests function; keep appropriate controls
DNA versus protein	Degeneracy can conserve protein despite DNA differences

Percentage identity



7 of 8 match

identity $= \frac{7}{8} \times 100 = 87.5\%$

$7/8 \times 100 = 87.5\%$

Attempt the handout questions before revealing answers

Explain the sequence, lineage and evidence separately

Translate and regroup the stated coding sequences. Explain synonymous, replacement, stop and frame-changing effects.

Calculate resistant fractions and absolute/relative risks; explain mutation, inheritance, selection and drift.

AHL: evaluate knockout/rescue, targeting/cutting/repair, ethics and sequence-conservation evidence.

Check the shared quantitative practice independently

1. **Resistant fraction.** $36/18000 = 0.002 = 0.20\%$; a final phenotype fraction, not automatically a mutation rate.
2. **Risk.** $3/1000 = 0.3\%$ and $6/1000 = 0.6\%$: doubles; 0.3 percentage-point absolute increase.
3. **Interpretation.** A panel tests specified variants, not every possible cause. A sequence association is not a complete individual prediction.

AHL practice: distinguish evidence from an inferred mechanism

1. **Knockout.** $(80 - 20)/80 \times 100 = 75\%$ reduction; restoration to 74 supports contribution, not direct binding.
2. **Controls.** Matched background and conditions, independent lines, replication and measured variation.
3. **Conservation.** Nine matches out of ten give 90% identity. Compare neutral regions, time and experimental function before attributing conservation solely to selection.

Check yourself

State what the evidence supports and what remains uncertain

Name DNA region and codon frame. Non-coding changes can affect expression or processing; a frame-preserving change need not be harmless.

Separate new variants, their inheritance and later changes in frequency. Need does not specify the useful new base.

AHL: separate intended target, cut, repaired sequence and measured function. Conservation and knockout data each need controls and alternative explanations.