# Diversity of organisms

A3.1 · SL

# Similarity and DNA amount need careful comparisons

## What exactly is being grouped or counted?

Organisms can look alike while belonging to different species; more DNA does not automatically mean greater complexity.

A3.1 connects species, names, chromosomes, genomes and sequencing. Identify the evidence and the quantity being compared.

Unreferenced datasets and diagrams are invented. Named NCBI records and research studies are identified in the handout.

# A measured difference need not be a species boundary

## Inherited and environmental effects both matter

Genotype, light, water and age can affect leaf length. Record life stage, method and conditions before comparing samples.

A morphological species concept groups organisms using shared structural traits, useful for fieldwork and fossils.

Use several stable diagnostic characters; one colour or size can mislead.

# Evaluate a morphological grouping

| **Evidence problem** | **Why appearance alone can mislead** |
|---|---|
| Age or sex differences | Members of one species can look different |
| Environmental effects | A trait difference need not be a genetic difference |
| Similar-looking species | Distinct lineages can share appearance |
| Convergence | Similar functional structures can evolve independently |

# The biological species concept concerns populations

## Reproductive relationships differ from appearance

Members of a potentially interbreeding group can produce viable, fertile offspring and are reproductively isolated from other groups.

The definition does not require every individual to mate with every other individual.

Age, sex or temporary separation of two individuals does not automatically define different species.

# Test a colour-based species claim

1. **Invented observation.** Two adult colour forms interbreed naturally and produce fertile offspring.
2. **Inference.** Colour alone is weak evidence for separate biological species.
3. **Further comparison.** Similar-looking populations with consistent reproductive barriers may nevertheless be distinct.

## Decision

Name the species concept and compare population-level evidence.

# A binomial requires both genus and specific epithet

## Write the complete name before abbreviating

*Homo sapiens*: capitalize Homo, use lowercase sapiens, and italicize both. Handwriting uses separate underlines.

After the full name, H. sapiens is acceptable where unambiguous. The epithet alone is incomplete.

Pan troglodytes and Pan paniscus share a genus but are different species. A common genus is not identity of every feature.

# A name is different from evidence of ancestry

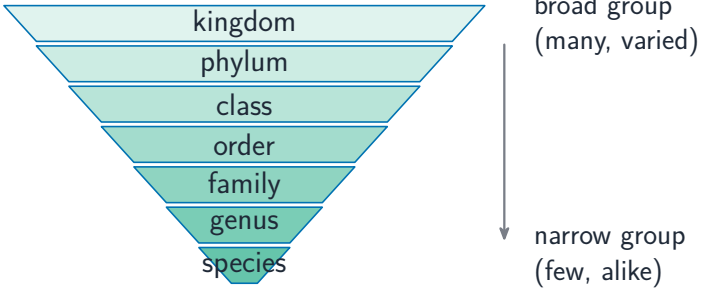
## Taxon and clade describe different ideas

A taxon is a named classification group, such as a species or genus.

A clade contains a common ancestor and all its descendants. A name alone does not establish the branching pattern.

A modern species is not automatically the direct ancestor of a different modern species.

# The taxonomic hierarchy




The binomial name uses genus and species

# Species boundaries can develop gradually

## Reduced gene flow permits divergence

Mutation, selection and other population processes can generate differences after gene flow is reduced.

Mating behaviour, breeding time and compatibility can produce partial isolation before a clear boundary is established.

Compare morphology, fertility, gene flow and molecular relationships. One consistent difference does not answer every species question.

# Calculate a sequence difference without inventing a clock

1. **Invented alignment.** Six differences among 40 aligned homologous sites.
2. **Calculate.** $d = 6/40 = 0.15 = 15\%$.
3. **Interpret.** This describes the selected alignment, not a universal species cutoff or a calibrated date.

## Decision

Test independent genes or characters. Rotating branches at an unscaled node changes page order, not ancestry.

# Compare typical complements, not complexity

| Organism | Diploid body cell and normal gamete |
| --- | --- |
| Human | 46 chromosomes / 23 chromosomes |
| Chimpanzee | 48 chromosomes / 24 chromosomes |
| Interpretation | A larger chromosome count does not automatically mean more DNA or genes |

# A copied chromosome has two sister chromatids

1. **Before copying.** A typical human diploid cell contains 46 chromosomes.
2. **After replication.** 46 chromosomes, each with two sister chromatids: $46 \times 2 = 92$ chromatids.
3. **Count convention.** Chromosomes are counted by centromeres before their separation.

## Decision

Homologous partners are different from the two sister chromatids of one copied chromosome.

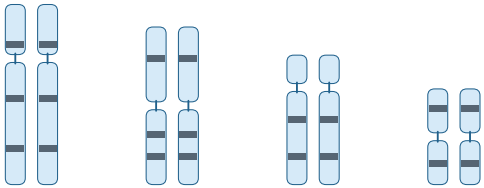
# Nucleus, chromosome, gene, DNA




The chromosome is a long DNA molecule

# Match four invented homologous pairs

Invented diploid complement: 2n = 8




Pair by length, centromere position and bands.
Eight chromosomes; four homologous pairs.

Four pairs = eight chromosomes. Pair by relative length, centromere position and banding pattern.

# A chromosome spread needs analysis

## Arrangement is an interpretation of an image

In a typical preparation, condensed chromosomes are stained or otherwise imaged, then identified and arranged.

Match several features; image orientation may be reversed. Joined sister chromatids are one copied chromosome.

A karyogram can reveal number changes and some structural changes, but usually cannot reveal a single-base substitution.

# Count an invented three-pair complement

1. **Observation.** Two long, two medium and two short chromosomes have matching partner patterns.
2. **Ploidy.** Three homologous pairs: $2n = 6$, so $n = 3$.
3. **After replication.** Six chromosomes and 12 chromatids before separation.

## Decision

An extra copy of one type must be identified explicitly, not forced into normal pairs.

# A historical hypothesis generates present-day predictions

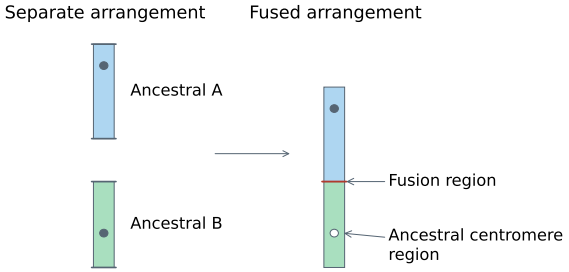
## What should fusion leave behind?

If two ancestral chromosomes fused, corresponding sequence and banding blocks should align along one descendant chromosome.

Internal telomere-like repeats are relevant because telomeres normally lie at chromosome ends.

Evidence for an additional ancestral centromere region adds another prediction to test. A chromosome count alone is insufficient.

# Compare ancestral blocks with a fused arrangement




Ancestral blocks; schematic map.
Internal telomere-like repeats test fusion.

The model conserves chromosome blocks; the junction and ancestral centromere region motivate sequence tests.

# Ancestral chromosomes are not modern chimpanzee parents

## Keep lineage history and nomenclature clear

Human chromosome 2 corresponds to chimpanzee chromosomes 2A and 2B, called 12 and 13 in older nomenclature.

Humans and chimpanzees share ancestors; one modern species is not automatically the direct ancestor of the other.

Two ancestral pairs becoming one fused pair reduces the diploid count by two without implying loss of half the genome.

# A genome includes coding and non-coding information

## Specify nuclear and organellar scope

Genome size is DNA amount, not chromosome or gene count. Nuclear and organellar DNA must be distinguished when a comparison requires it.

A SNP is variation at one nucleotide position in a population. Diploid homologues may carry the same or different alleles there.

Not every SNP changes a protein or produces a visible trait; location and biological context matter.

# A percentage needs its denominator and method

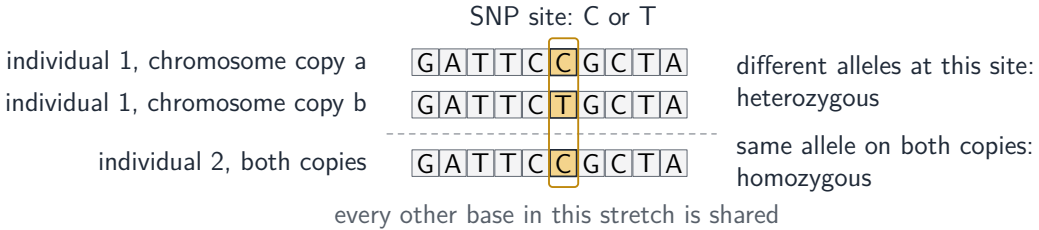
## Compare homologous sequence regions

Within a species, much sequence is shared, but individuals can differ. Between species, sequence and DNA amount can differ substantially.

State whether a comparison uses coding sites, a selected region or an entire assembly.

Similarity does not mean identical organisms, and differences alone do not date divergence.

# A single-nucleotide polymorphism




A SNP is one base that varies

# Record the NCBI assembly version

| **Organism / accession** | **Primary-assembly length, rounded** |
|---|---|
| E. coli / GCF_000005845.2 | 4.64 million bp |
| Yeast / GCF_000146045.2 | 12.1 million bp |
| Human / GCF_000001405.40 | 3,100 million bp |

# An assembly statistic is not a diploid-cell count

## Record what the database actually measured

For human GCF_000001405.40, the primary assembly has 3,099,441,038 bp; all listed units total 3,298,430,636 bp. Total sequence length can differ from ungapped length.

Check included chromosomes and organellar sequences. A human reference can contain both X and Y sequence types.

An assembly's chromosome sequence-type count is not the 46 chromosomes in an ordinary human diploid cell.

# More DNA does not automatically mean more complexity

## Compare compatible quantities and measurement methods

Repeats, other non-coding sequence, duplication and loss affect genome size without a simple complexity ranking.

A published fork-fern study reported about 160.75 billion bp per unreplicated nuclear complement.

A DNA-content estimate and a reference assembly total use different methods and component definitions; do not treat them as exactly interchangeable.

# Correct the DNA basis before dividing

1. **Invented comparable amounts.** Plant: 150 million bp per haploid nucleus. Animal: 3,000 million bp per haploid nucleus.
2. **Ratio.** $R = 3000/150 = 20$.
3. **If plant data were diploid.** 300 million bp becomes $300/2 = 150$ million bp on the same haploid basis.

## Decision

The ratio is about DNA amount; it is not a 20-fold complexity or gene-count result.

# Sequencing produces information that still needs interpretation

## A whole genome is different from one marker

Reads must be checked, assembled or aligned and analysed. Throughput has generally increased and costs fallen compared with early projects.

Research uses include evolutionary comparisons, variant discovery, inherited-condition research, pathogen studies and crop research.

Many genes can test a relationship inferred from one region, but different genomic regions can have different histories.

exercises.lol exercises.com.cn

# A sequence is not a complete prediction of phenotype

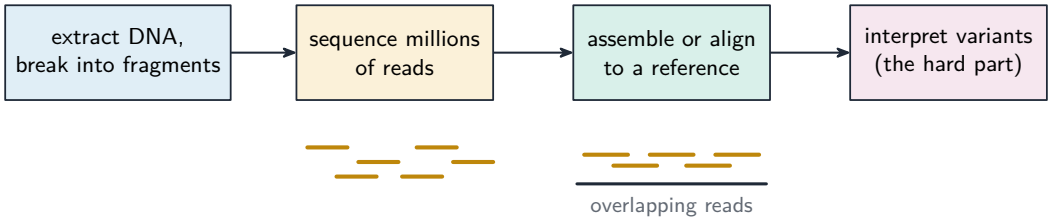
## Potential personalized medicine needs other evidence

Relevant variants may inform prevention or treatment alongside clinical and environmental information.

Some variant effects are uncertain, and environment and development affect outcomes. A sequence is not an automatic diagnosis.

Consent, privacy and access matter when genomic information is collected and interpreted.

# Whole-genome sequencing




cheaper and faster data; accurate interpretation is a separate task

> Then interpret the variants

# Try chromosome counts and an alignment comparison

## Identify the objects counted before calculating

An invented diploid complement has four homologous pairs. Find 2n, n and chromatids after replication before separation.

A selected alignment has eight differences among 50 sites. Find its percentage difference.

Explain why the alignment does not give a universal species threshold or a calibrated divergence date.

# Reveal: keep the count convention and inference separate

1. **Chromosomes.** Four pairs give $2n = 8$; normal haploid set $n = 4$.
2. **Chromatids.** After replication before separation: eight chromosomes, 16 chromatids.
3. **Alignment.** $8/50 \times 100 = 16\%$. A selected difference proportion supplies no universal cutoff or calibrated time model.

# Check yourself

## Concept, quantity and evidence

Name the species concept. Keep chromosomes, chromatids, genes and DNA amount separate.

Pair karyogram chromosomes using several features; test fusion with several predictions. Record database versions and compatible measurement bases.

For identification, state which organisms and conditions support the conclusion.