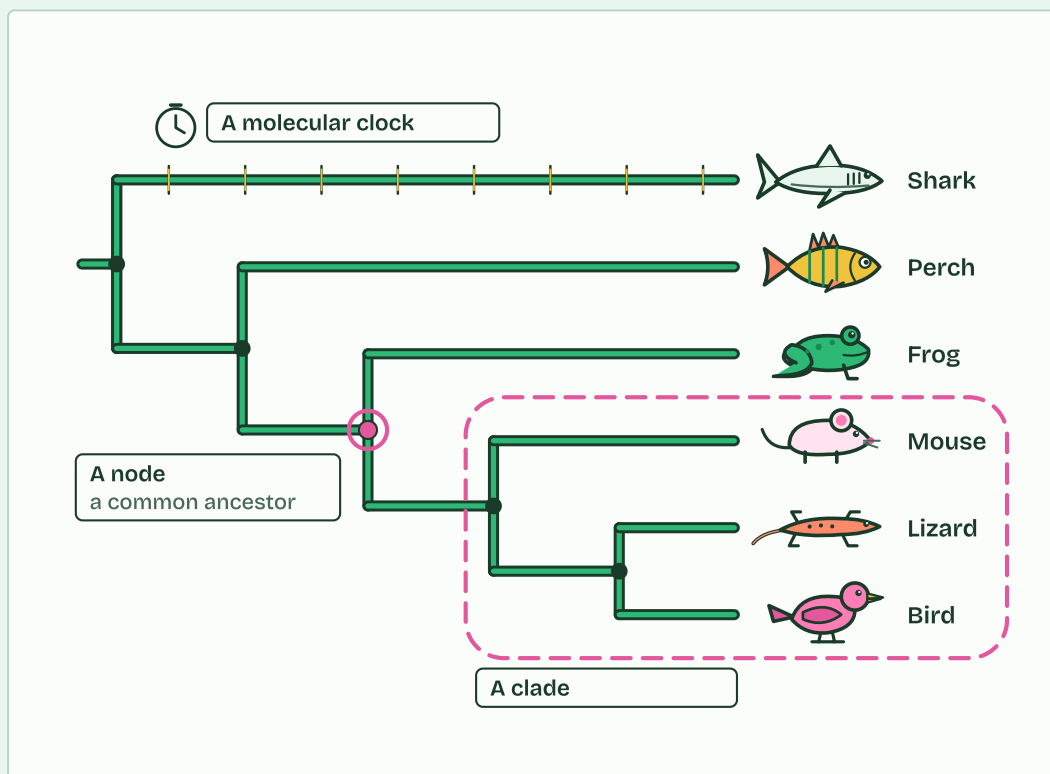


REVISION GUIDE

Sorting life by its *family* tree

Classification and cladistics

Biologists sort millions of species into groups. Here is why the old ranks stopped fitting, how DNA lets you build and read the tree of life, and how it split the prokaryotes into two domains. This guide is higher level only: standard level students can skip it.



9 chapters · a dry lab

The interactive guide, with every animation:

biologybybradford.com/guides/ib-biology-a3-2-classification-and-cladistics



Contents

01	Why sort life at all? HL	3
02	When the boxes do not fit HL	5
03	Groups that are family trees HL	8
04	A clade is an ancestor and all its descendants HL	10
05	Differences that count the years HL	12
06	The tree with the fewest changes HL	16
07	Reading a cladogram HL	20
08	Testing the boxes HL	23
09	A new level above kingdom HL	25
	Key terms	29
	Dry lab: Grow a tree	30
	Check yourself	31
	Answers	32

By the end you can

- ☐ Explain why classification is needed, and what it makes possible A3.2.1
- ☐ Describe why the traditional hierarchy of taxa does not always match evolution, and why a fixed ranking is arbitrary A3.2.2
- ☐ Explain why a classification that follows evolutionary relationships lets you predict the features of its members A3.2.3
- ☐ Define a clade, and say what evidence best places organisms in one A3.2.4
- ☐ Explain how the molecular clock estimates when clades diverged, and why it only gives estimates A3.2.5
- ☐ Build a cladogram from sequence data using parsimony, and say why different criteria can give different trees A3.2.6
- ☐ Read a cladogram: its root, nodes and terminal branches, its clades and common ancestors A3.2.7
- ☐ Use the figwort family to show how cladistics tests a classification, and why a claim can be falsified A3.2.8
- ☐ Explain how rRNA sequences led to the three domains, and what was new about them A3.2.9

Tick each one when you can do it without the book.

01 Why sort life *at all*?

A3.2.1 HL WHY WE CLASSIFY

There are more kinds of living thing than anyone can hold in their head. Biologists have named well over a million species, and many more are still waiting to be found. Nobody can learn them one by one. So we sort them.

Classification turns the diversity of life from a heap into a map.

A good map does two jobs. It helps you **name and find** an organism: once you know which group a beetle or a fern belongs to, you know where to look it up and what it is probably like. And it makes **further study** possible. When a new species is found, its group tells a biologist what to ask about it, such as what its relatives eat, how they breed and how best to protect them, instead of starting from nothing. Medicine, farming and conservation all lean on groups in this way.

KEY POINTS

- Classification is needed because there are too many species to study one at a time.
- Once organisms are classified, a broad range of further study becomes possible.
- Different rules give different groups, so the rule matters.

KEY TERMS

classification

Sorting organisms into groups, and groups into larger groups, by what they have in common.

has a backbone



Fish



Frog



Bird



Bat

no backbone



Butterfly



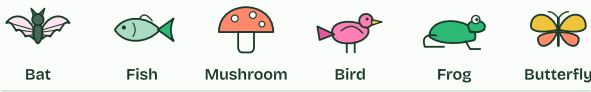
Mushroom

STEP BY STEP

3 STEPS

1

six organisms, in no order

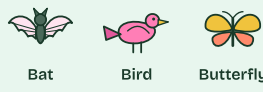


Bat Fish Mushroom Bird Frog Butterfly

Six organisms, to sort **any way you like**.

2

flies



Bat Bird Butterfly

does not fly



Fish Frog Mushroom

Rule one: **flies**. The bat, the bird and the butterfly go together.

3

has a backbone



Fish Frog Bird Bat

no backbone



Butterfly Mushroom

Rule two: **has a backbone**. Now the bat sits with the fish and the frog, and the butterfly has left it. Same six, different groups.

TRY IT FIRST

Before you read on, sort the six organisms in the figure into two groups, using any rule you like. Then compare your rule with a partner's. You will probably both have made sensible groups, and they will not be the same. A rule is worth keeping if the groups it makes tell you something. The next chapters ask which rule tells you the most.

WATCH OUT

Classification is not just naming. A name tells you nothing about a species. A group tells you what to expect from it.

PREDICT

Which is the best reason for classifying organisms?

- (A) So that every species can be given one fixed name, which would stop scientists ever disagreeing about how to group it
- (B) There are too many species to study one by one, and a group lets what we know about one guide the rest
- (C) So that species can be ranked from the simplest up to the most advanced, with humans placed at the very top

Decide before you look. Answer on page 32.

02

When the *boxes* do not fit

A3.2.2 HL THE TRADITIONAL HIERARCHY OF TAXA

Naturalists have long sorted life into a **hierarchy of taxa**. A **taxon** is a named group. Each species sits in a genus, each genus in a family, and so on up through order, class and phylum to kingdom. The system dates from Linnaeus, in the 1700s, and it grouped organisms mostly by features you can see.

It is tidy, and it has a problem. Evolution does not make boxes. It makes branches, and the branches split at different times. The ranks do not always follow that pattern.

Groups made from looks can gather unrelated lines: birds, bats and insects all fly, but each evolved its wings separately. Groups can also leave out relatives. In the old system crocodiles and lizards are in the class Reptilia, and birds have a class of their own, Aves. Yet the figure shows that crocodiles and birds share an ancestor that lizards do not.

Ranks are also **arbitrary**. Variation is gradual, and nothing in nature says how much difference makes a genus or a family. A family of beetles and a family of mammals need not be alike in age or in how much they differ, so a rank measures nothing fixed. **Cladistics** offers another way: unranked groups that follow the branching.

KEY TERMS

hierarchy of taxa

The traditional ranks of classification, one inside the next: kingdom, phylum, class, order, family, genus, species.

taxon

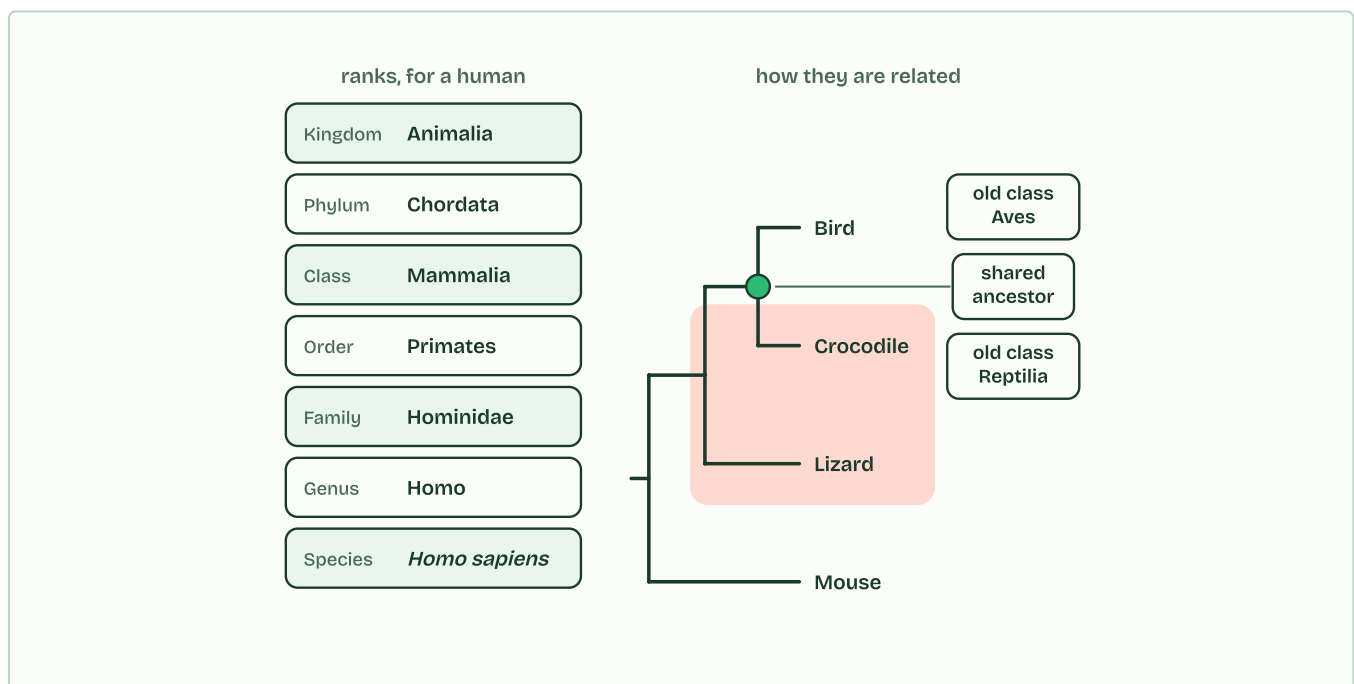
Any named group of organisms, at any rank. The plural is taxa.

cladistics

A method of classifying organisms into clades, groups made of an ancestor and all its descendants, with no fixed ranks.

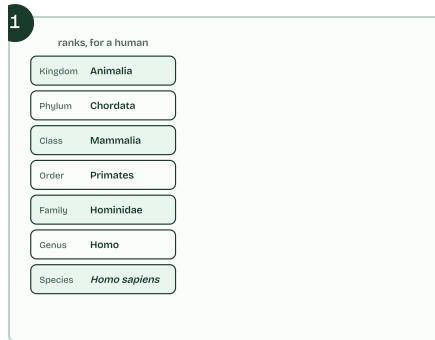
paradigm shift

A big change in the basic ideas and methods of a field of science, when the old ones can no longer cope with the evidence.

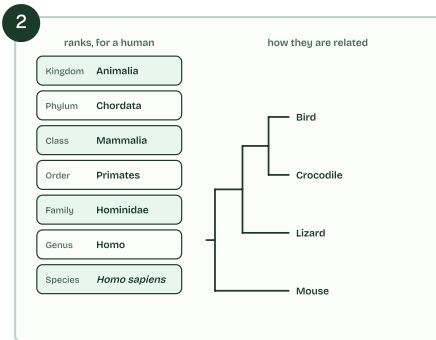


STEP BY STEP

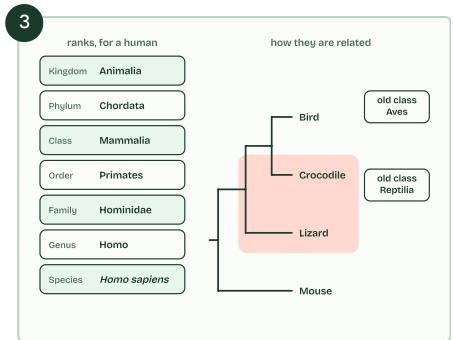
4 STEPS



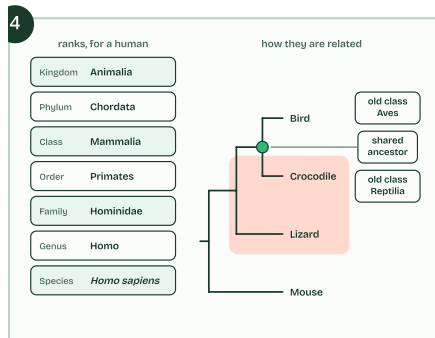
The traditional ranks: each one **neests** inside the next, from kingdom down to species.



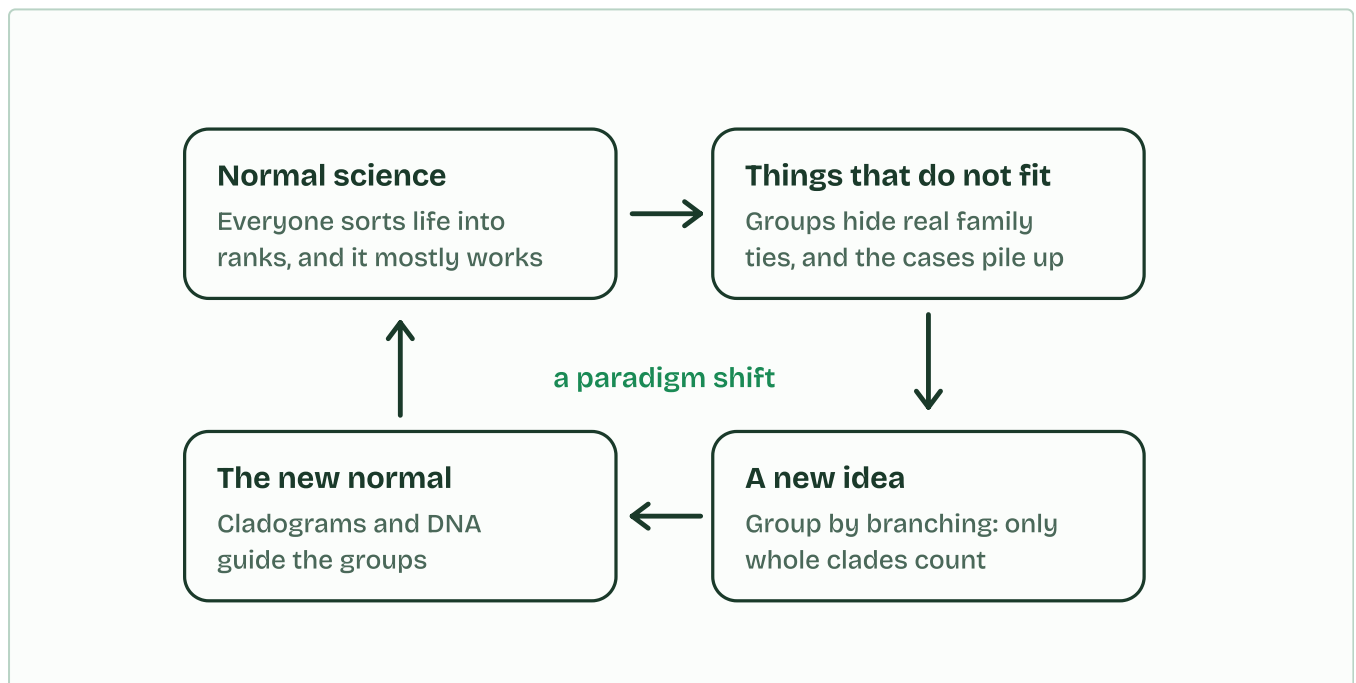
Now the branching tree of descent for four vertebrates, built from DNA (simplified).



The old classes. **Reptilia** holds the crocodile and the lizard, and **Aves** holds the bird on its own.

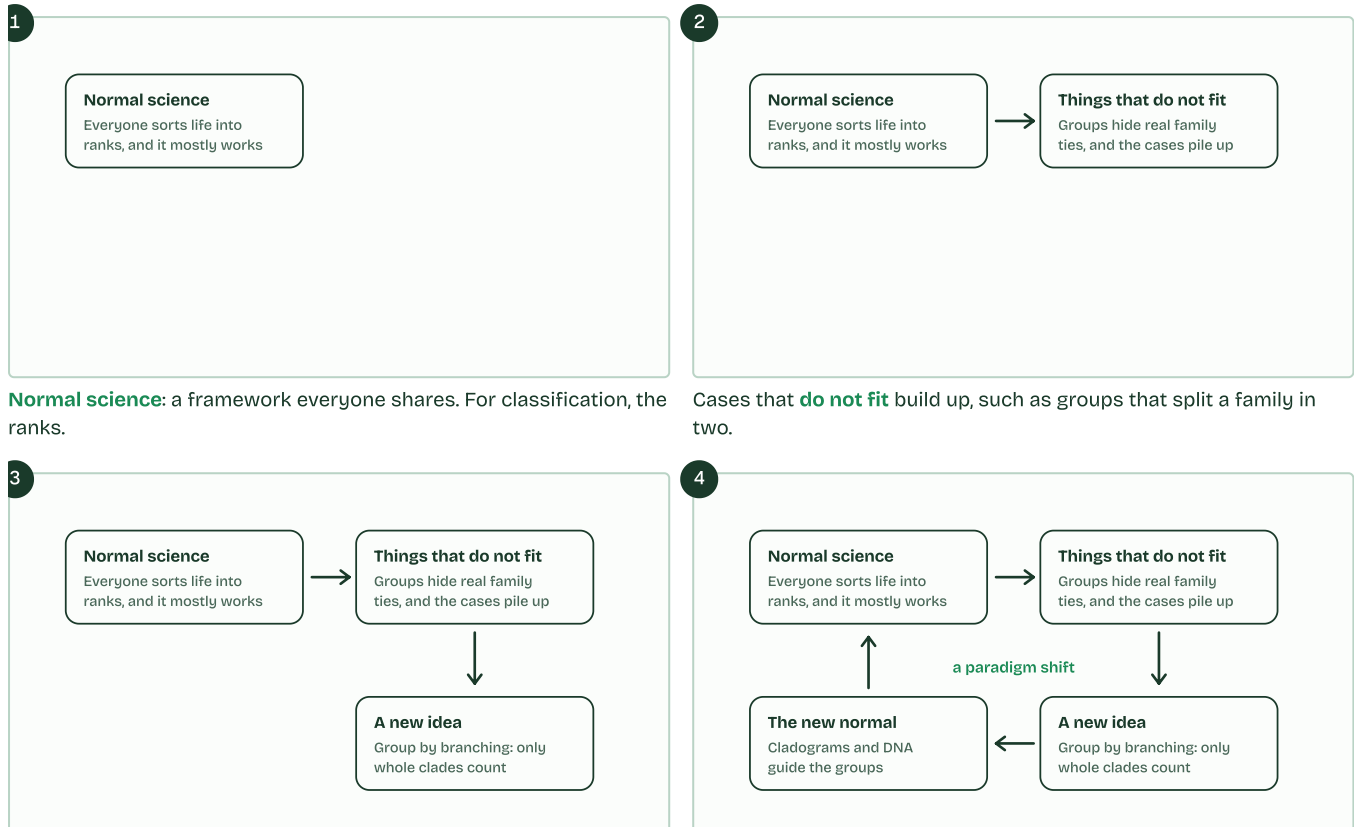


But the crocodile and the bird share an ancestor that the lizard does not. The old class **cuts that family in two**.



STEP BY STEP

4 STEPS



Normal science: a framework everyone shares. For classification, the ranks.

Cases that **do not fit** build up, such as groups that split a family in two.

A **new idea** replaces the rule for grouping: use the branching tree, and group only whole clades.

The new idea becomes the new normal. A change in the basic ideas of a field is a **paradigm shift**.

REMEMBER IT

The ranks, from the top

K King	P Philip	C Came	O Over	F For	G Good	S Soup
Kingdom the biggest group	Phylum a major body plan	Class mammals, birds, insects	Order primates, carnivores	Family cats, dogs, apes	Genus close species together	Species one kind of organism

A PARADIGM SHIFT

A **paradigm shift** is a fundamental change in how a science works. Classification has had one. Everyone used ranks, cases that did not fit piled up, and groups made from branching replaced them. Step through the second figure.

PREDICT

A taxonomist finds that a group in the old system leaves out one species that descends from the same common ancestor as the rest. What does that show?

- (A) The group is a good one, because its members still look alike
- (B) The group is too small, so it needs a higher rank
- (C) The group does not follow the pattern of divergence

Decide before you look. Answer on page 32.

03

A3.2.3 HL CLASSIFICATION THAT FOLLOWS RELATIONSHIPS

Groups that are *family trees*

Think of your own family. Cousins tend to resemble each other more than strangers do, because they share grandparents. Living things are the same, only the family is far bigger and the shared ancestors are further back. The ideal classification follows that family tree, so that every member of a group has evolved from one **common ancestor**, and none of that ancestor's descendants are left out.

The great strength of such groups is prediction. A feature that appeared in the common ancestor is passed to all its descendants, so it is shared across the group. If you know that one member of the group has hair and feeds its young on milk, you can predict the same of a newly found member without testing it. The figure follows one feature through a tree.

The predictions reach beyond looks, to biochemistry, life cycles, behaviour and disease. Plants in one group often make similar chemicals, so a chemical that proves useful in one plant is first looked for in its relatives. A group made from looks alone cannot promise this, because look-alikes with different ancestors need not share anything underneath.

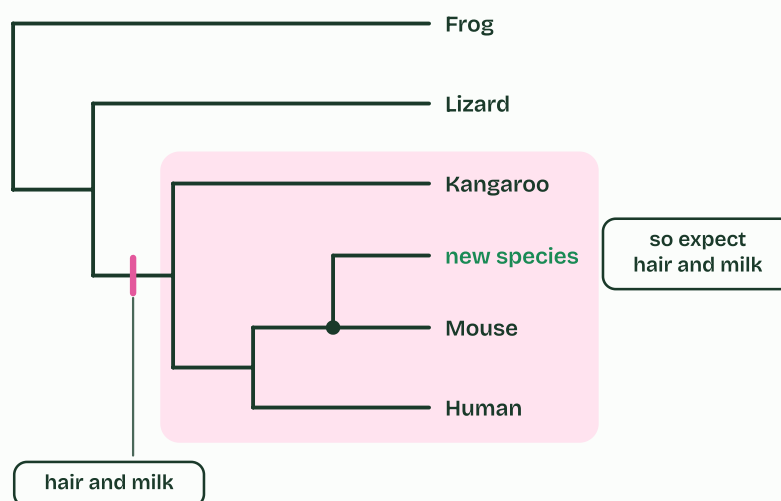
KEY TERMS

common ancestor

An ancestor shared by two or more organisms: the point where their lines split.

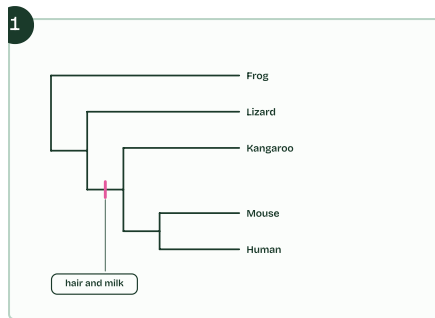
KEY POINTS

- The ideal classification follows evolutionary relationships: all members of a group descend from one common ancestor.
- Features of one member can be predicted for the rest, because they are shared within the group.

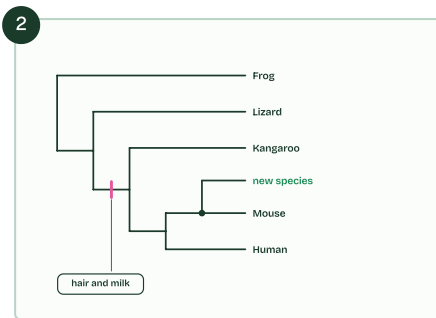


STEP BY STEP

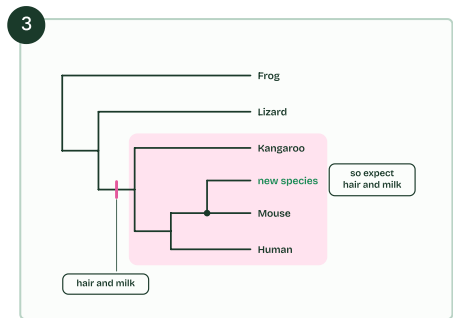
3 STEPS



Hair and milk appeared once, on the branch leading to the **mammals**. Every descendant of that ancestor inherits them, unless a line loses them.



A newly found animal. Its DNA says it is the mouse's closest relative: it **joins the tree** there (a made-up animal).



It sits inside the mammal clade, so you can **predict** its features before you have studied it.

WATCH OUT

A prediction is a good bet, not a promise. A feature can be lost along a branch: snakes belong to a group of four-limbed animals, but their ancestors lost the legs. That is why a prediction is then tested.

PREDICT

A newly found frog falls inside a clade whose other members all make a toxin in their skin. Which statement is best?

- (A) It must make the toxin, because every member of a clade shares every feature of its common ancestor
- (B) Nothing can be said about it until its skin has been tested, because relatives never share a toxin
- (C) It probably makes the toxin, but a feature can be lost, so the prediction needs testing

Decide before you look. Answer on page 32.

04

A3.2.4 HL CLADES

A clade is an ancestor and *all* its descendants

A **clade** is a common ancestor together with every organism that descends from it. Test any group with two questions. Do all its members share one common ancestor? And are all of that ancestor's descendants inside the group? Both answers must be yes. Leave out one descendant, or add one that is not a descendant, and it is not a clade.

Clades nest like boxes inside boxes: birds are a clade, and so is the larger clade that holds birds and crocodiles. Tap the lettered nodes on the tree to ring the clade that starts at each one, then tap the old group called reptiles.

How do you decide which organisms share an ancestor? The most **objective** evidence comes from the base sequences of genes or the amino acid sequences of proteins. A sequence is a long string that can be compared position by position, and anyone counting the differences gets the same answer. **Morphological** traits can also place organisms in clades, but they need a judgement about which features count, and they can mislead when unrelated lines evolve similar structures.

KEY POINTS

- A clade is one common ancestor plus all of its descendants.
- Sequence evidence is the most objective, and morphological traits can also be used.

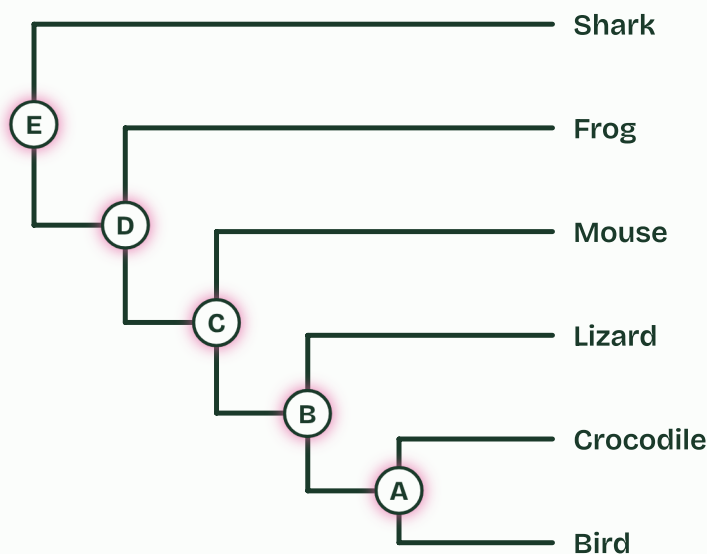
KEY TERMS

clade

A group made of one common ancestor and all of its descendants.

morphological

The shape and structure of an organism's body.



Node E, the root every organism on the tree. The whole tree is always a clade.

Node D add the frog and the ring holds every descendant of node D, so it is a clade too.

Node C add the mouse and the ring still holds everything that descends from this ancestor. Still a clade.

Node A the crocodile and the bird, and nothing else, descend from this ancestor. Together they are a **clade**.

Node B the lizard, the crocodile and the bird all descend from this ancestor, and it has no other descendants. The bird is inside this clade.

The old reptiles the lizard and the crocodile only. They share their ancestor with the bird, which is left out, so this group is **not a clade**. Leave out even one descendant and it fails.

WATCH OUT

A clade is not decided by how organisms look, how big the group is or what rank it once had. Only descent counts. A group of look-alikes that leaves out a relative is not a clade.

PREDICT

A tree shows an ancestor X with four descendants, P, Q, R and S. Someone proposes the group P, Q and R. Is it a clade?

- (A) No, because S also descends from X and is left out
- (B) Yes, because P, Q and R all share the common ancestor X
- (C) Yes, provided P, Q and R look more alike than S does

Decide before you look. Answer on page 32.

05

A3.2.5 HL THE MOLECULAR CLOCK

Differences that *count the years*

Mutations change a gene's sequence at random. When one species splits in two, each line collects its own mutations from then on. So the longer ago two species split, the more differences you find between the same gene in each (you met this with sequences in A4.1). If changes arrive at a roughly steady rate, the differences work like a clock.

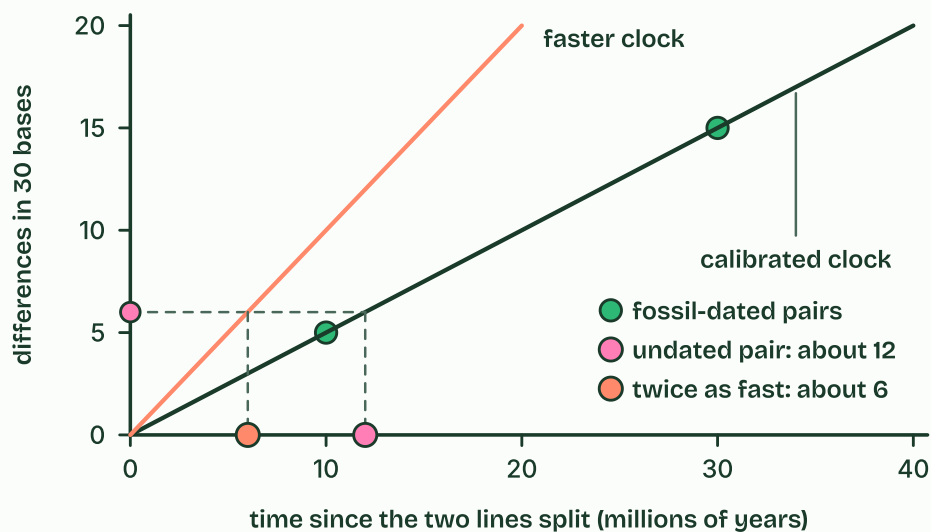
To turn a count into a date you need a calibration. Take pairs of species whose split is dated from fossils, and plot their differences against time. A straight line through the points is the **molecular clock**. Now a pair with no fossil date can be placed: read its number of differences across to the line and down to the time.

The result is an **estimate**, never a date. The rate at which changes build up is not the same everywhere, because it is affected by the length of the generation time, the size of the population, the intensity of selection and other factors. In the figure's last step, a line that collected changes faster gives a quite different answer from the same count.

KEY TERMS

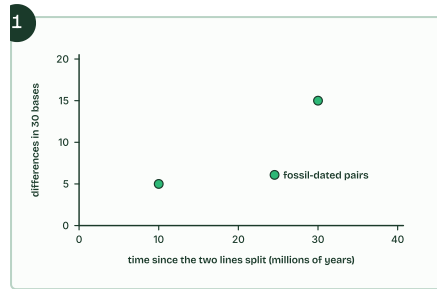
molecular clock

A way of estimating when two clades diverged from a common ancestor, from the number of differences that have built up in their gene or protein sequences.

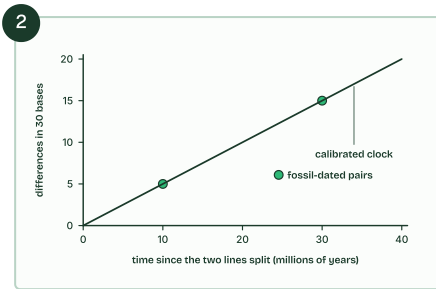


STEP BY STEP

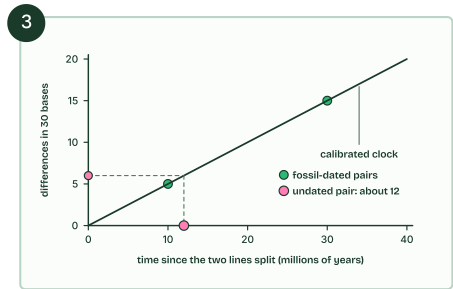
4 STEPS



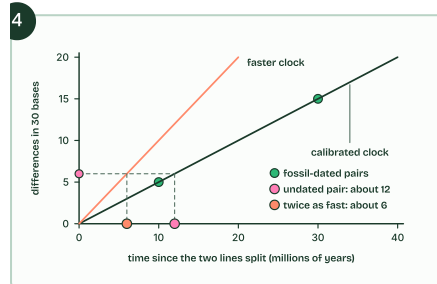
Two pairs of species whose split is **dated from fossils**: 5 differences at 10 million years and 15 at 30 million (made-up numbers).



The differences grow steadily with time, so a straight line through the points is the **molecular clock**.



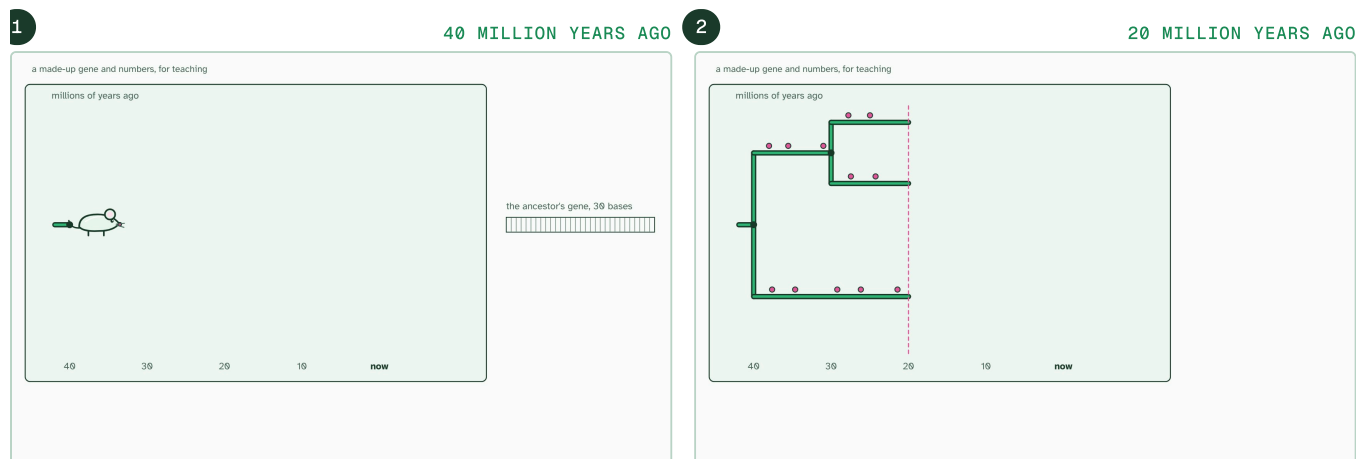
A pair with **no fossil date** differs at 6 places. Read across to the clock, then down: about 12 million years.



If this pair's lines collected changes **twice as fast**, the same 6 differences would mean about 6 million years. The clock gives an estimate, not a date.

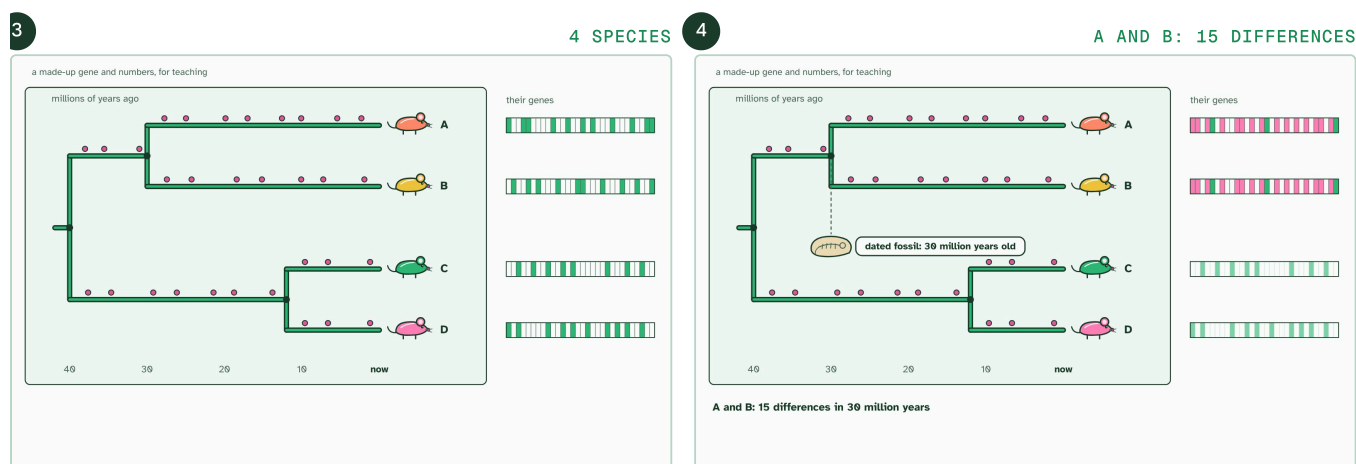
ANIMATION · WATCH IT HAPPEN: TWO LINES COLLECT CHANGES

7 MOMENTS



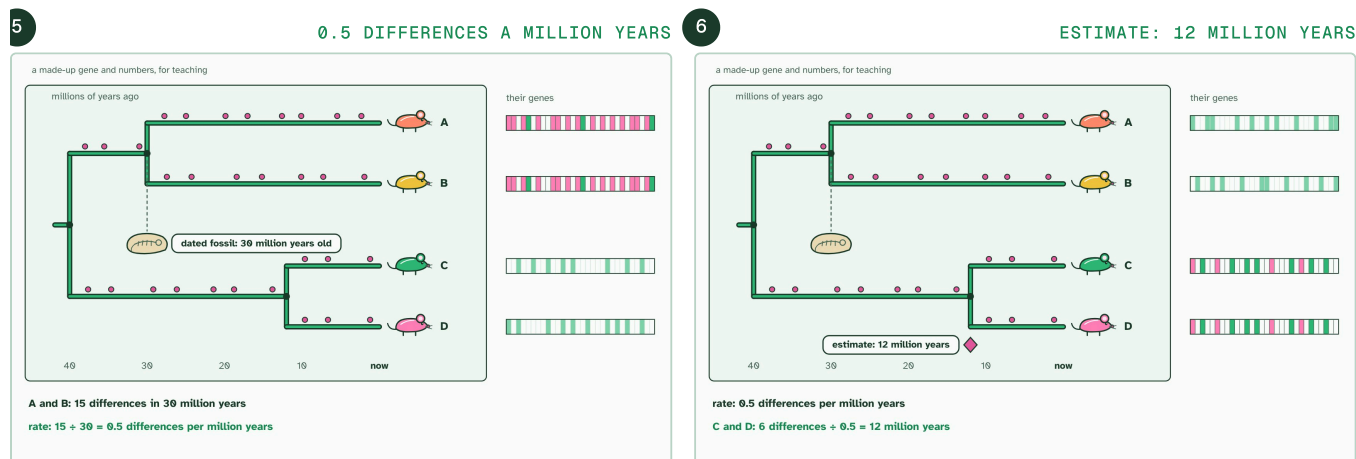
One ancestor species, and one of its genes: a stretch of 30 bases, drawn as boxes.

Its descendants split into new species. In every line, mutations build up in the gene at a fairly steady rate: **the ticks of a clock**.



Today there are four species, A to D. Each gene now differs from the ancestor's, in its own places.

A dated fossil shows that A and B split **30 million years** ago. Their genes differ at **15** bases.

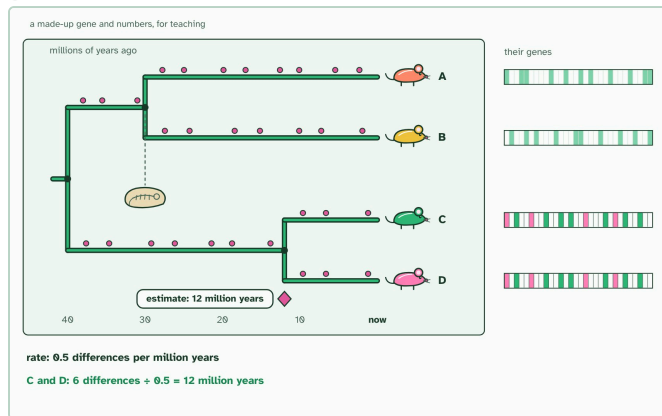


So this clock runs at $15 \div 30 = 0.5$ differences per million years.

C and D differ at **6** bases: $6 \div 0.5 =$ about **12 million years** since they split.

7

ESTIMATE: 12 MILLION YEARS



Only an **estimate**: generation time, population size and selection all change how fast differences build up. Try the settings.

Watch it move, and change it as it runs: biologybybradford.com/guides/ib-biology-a3-2-classification-and-cladistics#clock

REMEMBER IT

What shifts the clock

G Gerbils	P Play	S Sports
Generation time short generations copy DNA more often each year	Population size chance spreads changes differently	Selection pressure selection removes some changes, keeps others

WATCH OUT

A molecular clock is not a watch. It gives an average rate, calibrated against fossils, and different genes and different lines can tick at different rates. That is why scientists quote a range and not a single date.

PREDICT

A clock is calibrated on mammals that breed every few years. It is then used to date a split between two lines of insects that breed several times a year. What will the estimate probably be?

- (A) Too young, because insects gather fewer changes each year than mammals do
- (B) About right, because DNA is copied and mutates in the same way in every animal
- (C) Too old, because the insects probably collected changes faster each year

Decide before you look. Answer on page 32.

06

A3.2.6 HL BUILDING A CLADOGRAM

The tree with the *fewest changes*

A **cladogram** is a hypothesis about the order of splits. To build one from sequences, line up the same gene from each species and compare its position by position. Many trees could join the same species, so you need a rule to choose between them.

The rule used here is **parsimony**: choose the tree that accounts for the sequence differences with the fewest changes. Mutations are rare, so the explanation that needs fewest of them is the most probable. Add an outgroup, a species that split off earlier, to show which base is the older one.

The figure uses made-up data, small enough to count by hand. Each site is a column. A site where one species stands alone costs one change on every tree, so it cannot help. The sites where two species share a change are the ones that decide. Then try the tool: join the beetles yourself, watch the count, and run it again with a DNA column.

A cladogram is not a fact, and the method matters. Different criteria for judgement can lead to different hypotheses. A different gene, a different rule or traits of the body instead of sequences could give another tree, and new data can overturn one.

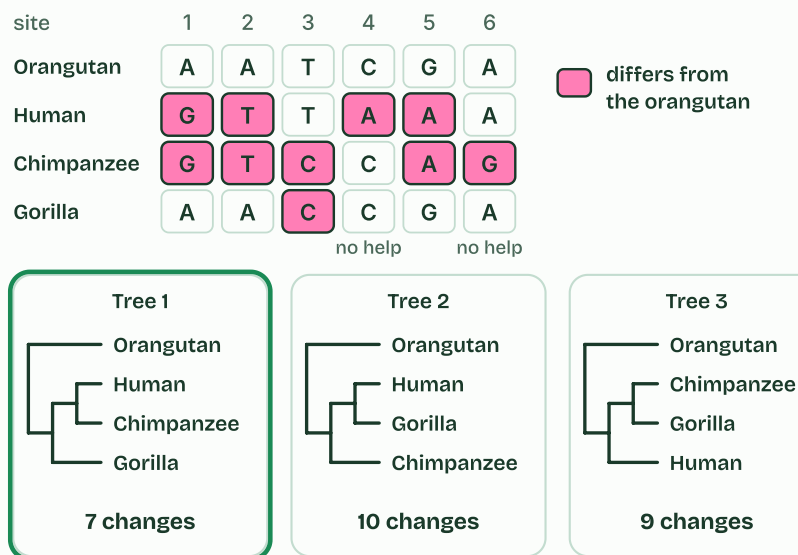
KEY TERMS

cladogram

A branching diagram that shows the proposed order in which groups split from common ancestors.

parsimony

Choosing, from all possible trees, the one that explains the observed differences with the smallest number of changes.



STEP BY STEP

4 STEPS

1

site	1	2	3	4	5	6
Orangutan	A	A	T	C	G	A
Human	G	T	T	A	A	A
Chimpanzee	G	T	C	C	A	G
Gorilla	A	A	C	C	G	A

differs from the orangutan

2

site	1	2	3	4	5	6
Orangutan	A	A	T	C	G	A
Human	G	T	T	A	A	A
Chimpanzee	G	T	C	C	A	G
Gorilla	A	A	C	C	G	A

no help

Tree 1

Tree 2

Tree 3

3

site	1	2	3	4	5	6
Orangutan	A	A	T	C	G	A
Human	G	T	T	A	A	A
Chimpanzee	G	T	C	C	A	G
Gorilla	A	A	C	C	G	A

no help

Tree 1

Tree 2

Tree 3

A made-up stretch of six sites in one gene. Pink marks where a species **differs from the orangutan**, the outgroup.

Three trees could join the human, chimpanzee and gorilla. Sites 4 and 6 differ in only **one species**, so they cost one change on every tree and cannot choose between them.

Count the changes each tree needs to explain all six sites. Tree 1 needs 7, tree 2 needs 10 and tree 3 needs 9.

4

site	1	2	3	4	5	6
Orangutan	A	A	T	C	G	A
Human	G	T	T	A	A	A
Chimpanzee	G	T	C	C	A	G
Gorilla	A	A	C	C	G	A

no help

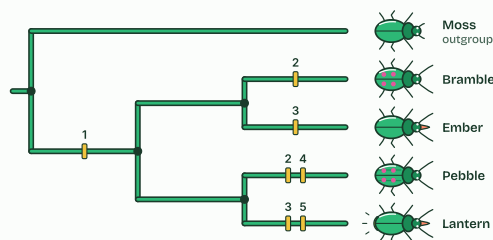
Tree 1

Tree 2

Tree 3

Parsimony picks the tree that explains the data with the fewest changes: tree 1, with the human and chimpanzee as closest relatives.

Traits	1	2	3	4	5
Moss	○	○	○	○	○
Bramble	●	●	○	○	○
Pebble	●	●	○	●	○
Ember	●	○	●	○	○
Lantern	●	○	●	○	●



This tree: 7 changes

Join the beetles into a tree, mark where each trait changes, and watch the count.

STEP BY STEP

2 STEPS

1

Traits	1	2	3	4	5
Moss	0	0	0	0	0
Bramble	●	●	0	0	0
Pebble	●	●	0	●	0
Ember	●	0	0	0	0
Lantern	●	0	●	0	●

Bramble
Pebble
Ember
Lantern

The table gives each beetle's traits (made up for teaching). Moss is the outgroup: its states are taken as the ancestral ones.

2

Traits	1	2	3	4	5
Moss	0	0	0	0	0
Bramble	●	●	0	0	0
Pebble	●	●	0	●	0
Ember	●	0	0	0	0
Lantern	●	0	●	0	●

This tree: 7 changes

One tree, with a bar on each branch where a trait changes: 7 changes. Can you find a tree that needs fewer?

Traits and DNA sites	1	2	3	4	5	a	b	c	d	e	f
Moss	0	0	0	0	0	A	C	G	T	A	C
Bramble	●	●	0	0	0	G	C	G	T	C	C
Pebble	●	●	0	0	0	G	T	A	T	C	C
Cinder	●	●	●	●	0	G	T	A	T	C	C
Ember	●	0	●	0	0	A	C	G	C	A	C
Lantern	●	0	●	●	●	A	C	G	C	A	T

This tree: 16 changes

The same job with a DNA column added. Two of the traits mislead: does the tree with the fewest changes change?

STEP BY STEP

2 STEPS

1

Traits and DNA sites	1	2	3	4	5	a	b	c	d	e	f
Moss	0	0	0	0	0	A	C	G	T	A	C
Bramble	●	●	0	0	0	G	C	G	T	C	C
Pebble	●	●	0	0	0	G	T	A	T	C	C
Cinder	●	●	●	●	0	G	T	A	T	C	C
Ember	●	0	●	0	0	A	C	G	C	A	C
Lantern	●	0	●	●	●	A	C	G	C	A	T

Bramble
Pebble
Cinder
Ember
Lantern

The table gives each beetle's traits (made up for teaching). Moss is the outgroup: its states are taken as the ancestral ones.

2

Traits and DNA sites	1	2	3	4	5	a	b	c	d	e	f
Moss	0	0	0	0	0	A	C	G	T	A	C
Bramble	●	●	0	0	0	G	C	G	T	C	C
Pebble	●	●	0	0	0	G	T	A	T	C	C
Cinder	●	●	●	●	0	G	T	A	T	C	C
Ember	●	0	●	0	0	A	C	G	C	A	C
Lantern	●	0	●	●	●	A	C	G	C	A	T

This tree: 16 changes

One tree, with a bar on each branch where a trait changes: 16 changes. Can you find a tree that needs fewer?

WATCH OUT

The most parsimonious tree is the most probable, not the certain one. Two trees can tie on the count, and then more data, more positions or another gene, is needed to choose between them.

PREDICT

Three possible trees need 14, 11 and 11 changes. What does parsimony tell you?

- Ⓐ The two trees that need 11 are tied, so more data is needed to choose between them
- Ⓑ The tree that needs 14 is best, because it explains the most differences between the species
- Ⓒ Either tree that needs 11 is right, so they must show exactly the same relationships

Decide before you look. Answer on page 32.

07

Reading a *cladogram*

A3.2.7 HL ANALYSING CLADOGRAMS

A cladogram is a map of splits, so you read it from the base up to the tips. The base is the **root**. Each fork is a **node**, which stands for a **hypothetical** common ancestor: it must have existed to explain the split, but it need not be a known fossil. Each line out to a tip is a **terminal branch**. Tap the parts.

To deduce a relationship, follow two tips back until their lines meet. That node is their most recent common ancestor. The nearer it is to the tips, the closer the relatives. Two groups that branch from the same node are **sister groups**. In the figure the human and chimpanzee meet at a node that the gorilla has not yet joined, so a human is closer to a chimpanzee than to a gorilla.

Any node and everything beyond it is a **clade**, so you can read clades straight off the tree. Then use the clade to deduce what its members share.

KEY TERMS

root

The base of a cladogram: the common ancestor of every group on the tree.

node

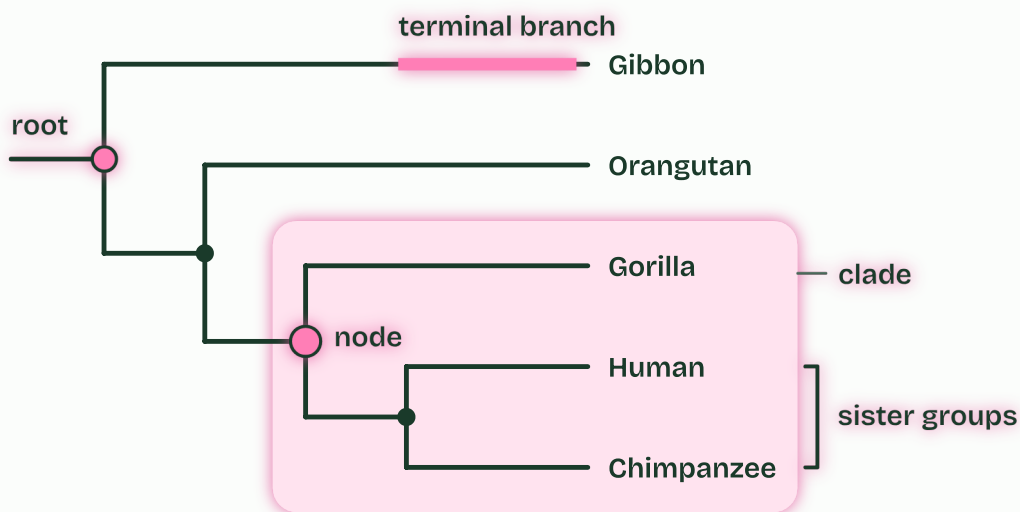
A branching point on a cladogram. It stands for a hypothetical common ancestor of the groups beyond it.

terminal branch

The line from the last node out to a tip, which is a species or group being compared.

sister groups

Two groups that branch from the same node, so each is the other's closest relative.



Clade an ancestor (the node at the left edge of the ring) and all of its descendants. The gorilla, human and chimpanzee make one.

Node a branching point. It stands for a **hypothetical common ancestor**: a species that must have existed to explain the split, though we may never find it.

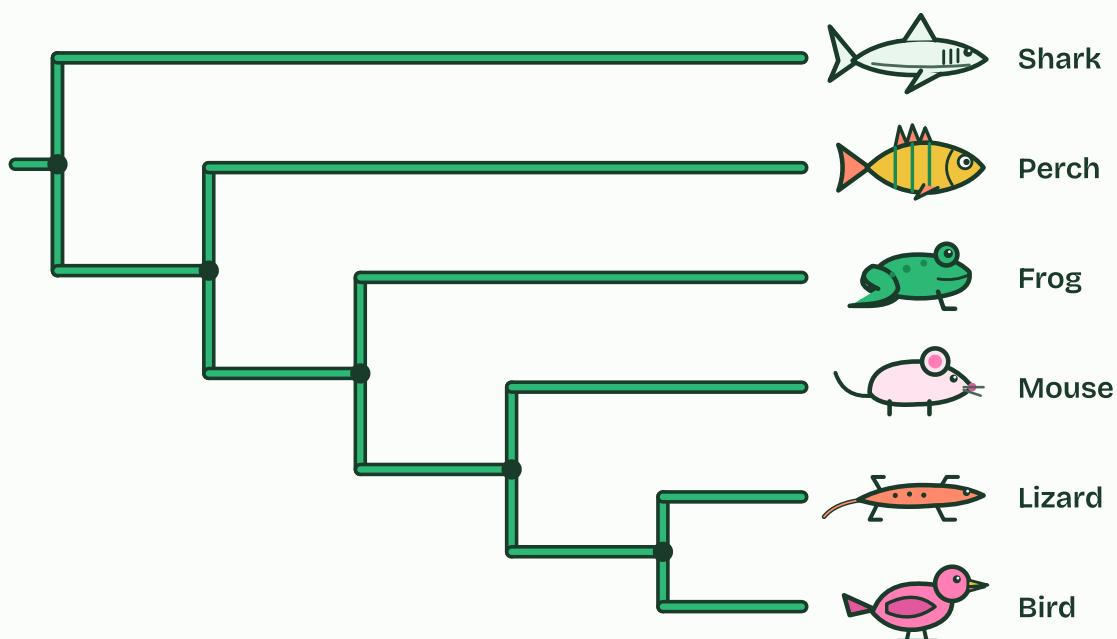
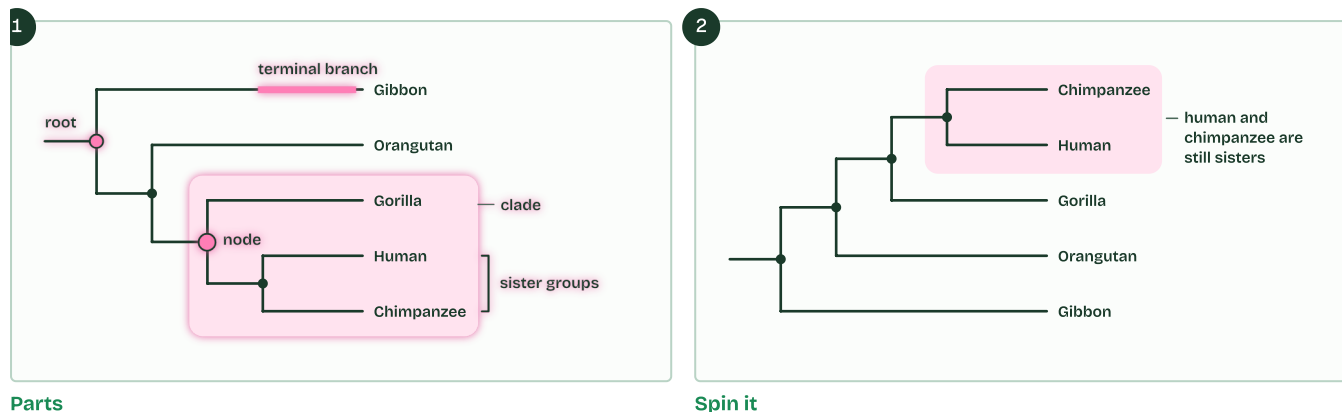
Sister groups the two groups that branch from the same node, so each is the other's closest relative. Here, human and chimpanzee.

Root the common ancestor of every organism on the tree, at its base. Everything else descends from it.

Terminal branch the line from the last node out to a tip. A tip is one species, or one group, being compared.

SWITCH IT

2 SETTINGS



Answer questions about a cladogram: common ancestors, clades and which pairs are closest relatives.

REMEMBER IT

The parts of a cladogram

R Rabbits	N Nibble	T Tasty	C Carrots	S Slowly
Root the base: ancestor of everything	Node a fork: a hypothetical ancestor	Terminal branch the line out to a tip	Clade an ancestor and all its descendants	Sister groups two groups from one node

WATCH OUT

The order of the tips means nothing, because every node can be spun like a mobile. Tips at the top are not more advanced, and a tip is never the ancestor of another tip: humans did not evolve from chimpanzees, they share an ancestor. In this course, only the branching order carries meaning.

PREDICT

In a cladogram, species X and Y branch from node 1, and node 1 and species Z branch from node 2. Which statement is correct?

- Ⓐ Z is the most advanced, because it is drawn at the bottom of the tree, below the others
- Ⓑ X is closer to Y than to Z, because X and Y share a more recent ancestor
- Ⓒ Y is descended from X, because their two lines meet at node 1 on the tree

Decide before you look. Answer on page 32.

08

A3.2.8 HL TESTING A CLASSIFICATION WITH CLADISTICS

Testing the *boxes*

Cladistics can test an old classification. If a group is real, its members should form a clade. A case study is the figwort family, a large family of flowering plants that was built on the structure of its members, mostly their flowers. You do not need to learn the names, only the story.

Researchers then compared the base sequences of a gene from the chloroplasts of these plants and built a cladogram. The figworts were not one clade. Many of the plants belonged with plants that were not figworts, and the family was broken up, its members moved to other families, leaving a much smaller figwort family.

Why had structure misled? Plants in different lines had evolved flowers that look alike, by **convergent evolution**. The likeness came from similar pressures, not from shared ancestry, so a grouping built on it was false.

This is science working, not failing. The old classification was a hypothesis, and it made a testable prediction: its members form a clade. New evidence tested it and it was **falsified**. Scientific knowledge claims may eventually be falsified, which is why they are always open to new tests.

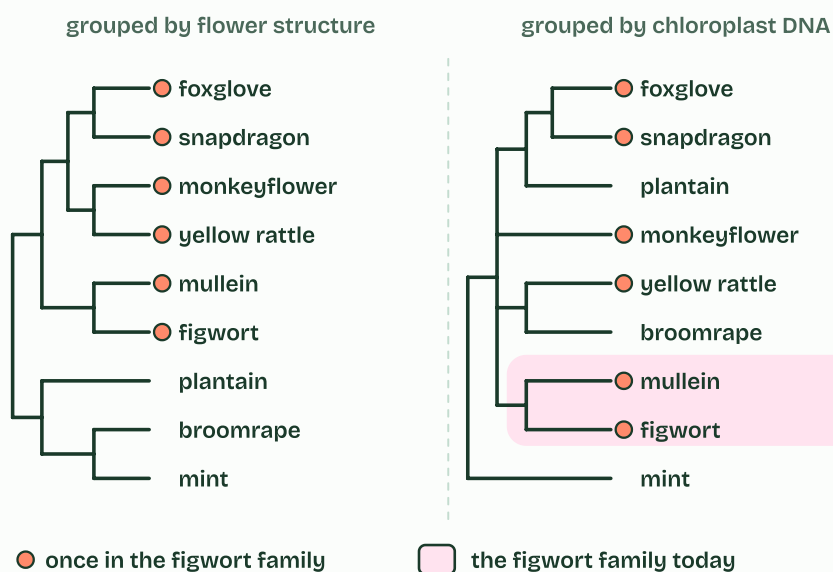
KEY TERMS

convergent evolution

Unrelated lines evolving similar features, often because they face similar demands, so the likeness does not come from a common ancestor.

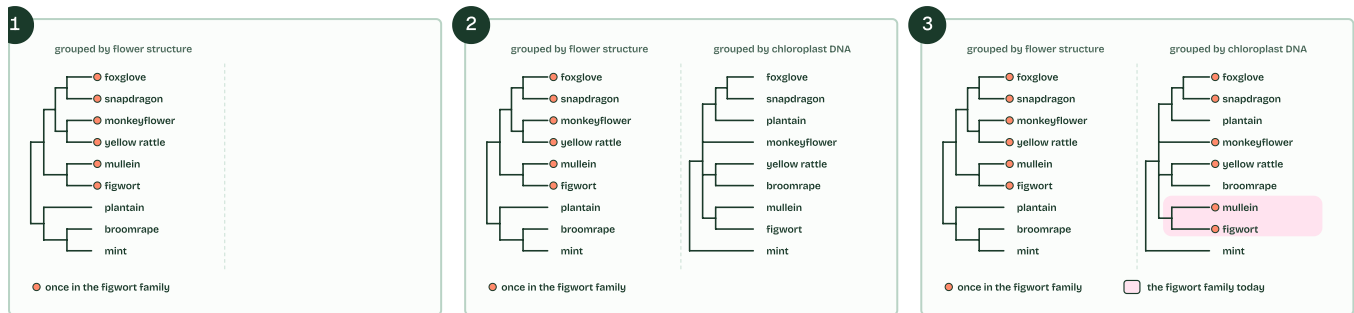
falsified

Shown to be false by evidence.



STEP BY STEP

3 STEPS



Grouped by **flower structure**: six plants with similar flowers sit together as the figwort family (a simplified sketch).

Grouped by **chloroplast DNA** instead. The same plants now fall into several different clades, with other plants between them.

Only a small clade is still the figwort family. The flower resemblance was **convergence**, so the old family was not a real clade.

IN AN EXAM

You can use either the common name or the scientific name of an organism. You are not asked to memorise the details of this case study, but you should be able to explain what it shows.

PREDICT

Chloroplast DNA places two plants with almost identical flowers in different clades. What is the best explanation?

- (A) The DNA must be wrong, because flowers are easier to examine than DNA
- (B) They share a recent ancestor, and the DNA only seems to separate them
- (C) Their flowers probably evolved separately, by convergent evolution

Decide before you look. Answer on page 33.

09

A3.2.9 HL THE THREE DOMAINS

A new level above *kingdom*

For a long time life was sorted into a few kingdoms: animals, plants, fungi, and two kinds of single-celled life, the ones with nuclei and the ones without. The cells without nuclei, the prokaryotes, were treated as one group. Then, in 1977, Carl Woese and George Fox compared **rRNA** base sequences from many organisms, and the picture changed.

Why rRNA? Every living cell has ribosomes, so every organism has rRNA genes to compare. They can be lined up across the whole of life, from a bacterium to a whale. The comparison showed that the prokaryotes were not one group. Some, such as the methane-makers, had rRNA as different from ordinary bacteria's as bacteria's is from ours.

Life falls into three great branches, now called **domains**: Bacteria, Archaea and Eukarya (organisms whose cells have nuclei). Archaea look like bacteria under a microscope, but they share a more recent common ancestor with Eukarya. This was a revolutionary reclassification, with an extra level above the kingdoms.

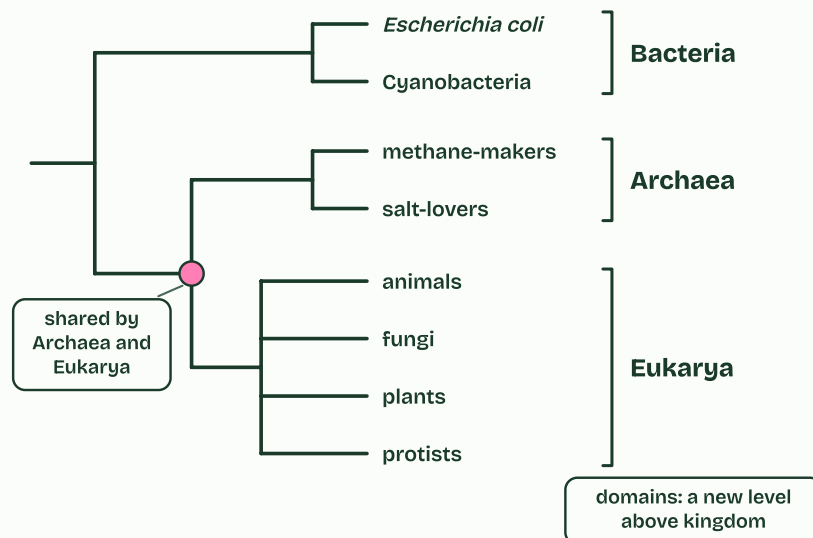
KEY TERMS

rRNA

The RNA that is part of ribosomes. Every organism has it, so its base sequence can be compared across all of life.

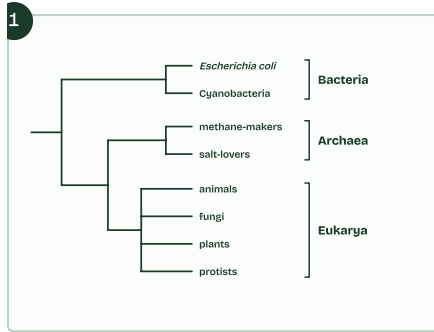
domains

The highest level of classification, above kingdom: Bacteria, Archaea and Eukarya.

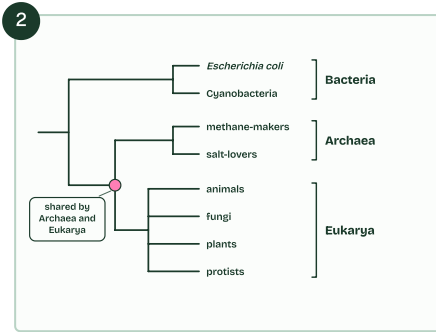


STEP BY STEP

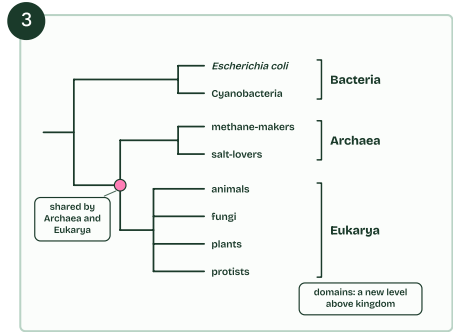
3 STEPS



The tree from **rRNA sequences**: three main branches, the **domains** Bacteria, Archaea and Eukarya. Relationships inside each domain are not shown.



Archaea look like bacteria under a microscope, but their rRNA shows they share a more recent ancestor with **Eukarya**.



A new top level of classification, **above kingdom**, proposed in 1977.

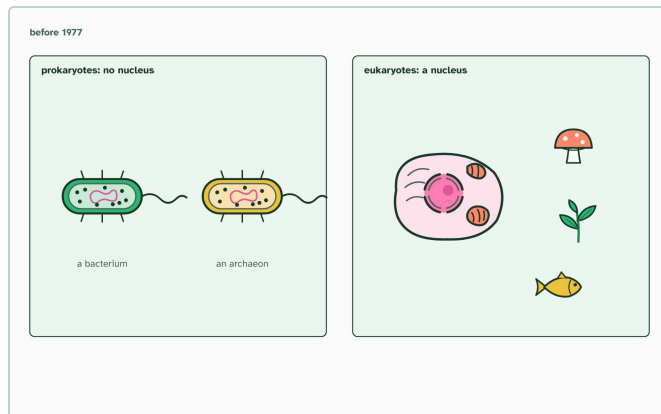
ANIMATION · WATCH IT HAPPEN: THE TREE OF LIFE FROM RRNA

7 MOMENTS

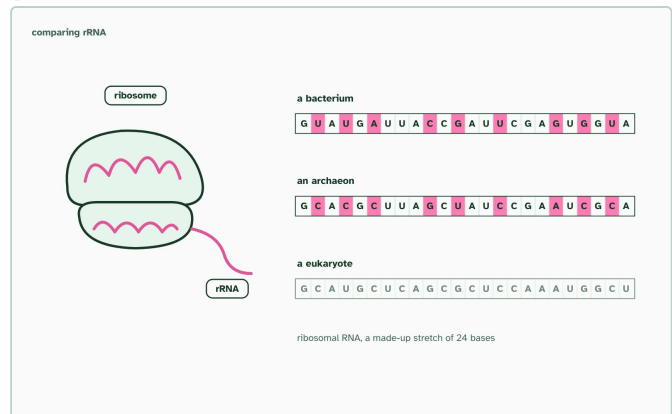
1

TWO GROUPS 2

9 BASES DIFFER



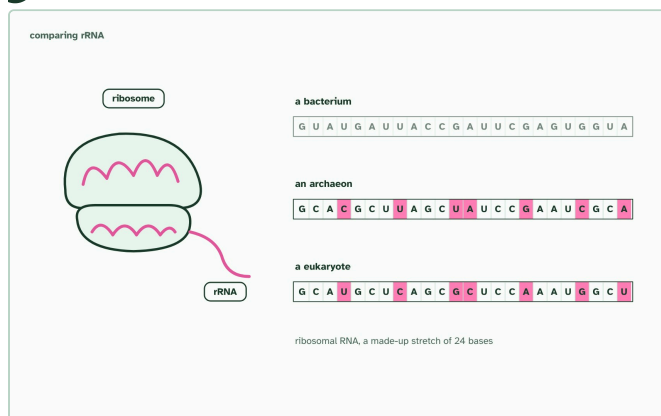
Before 1977, life was sorted by its cells: **prokaryotes**, with no nucleus, and **eukaryotes**, with one. All prokaryotes looked alike, so they were grouped together.



Carl Woese compared the base sequence of **ribosomal RNA**, which every cell has. Some prokaryotes, the archaea, differed from bacteria at many bases.

3

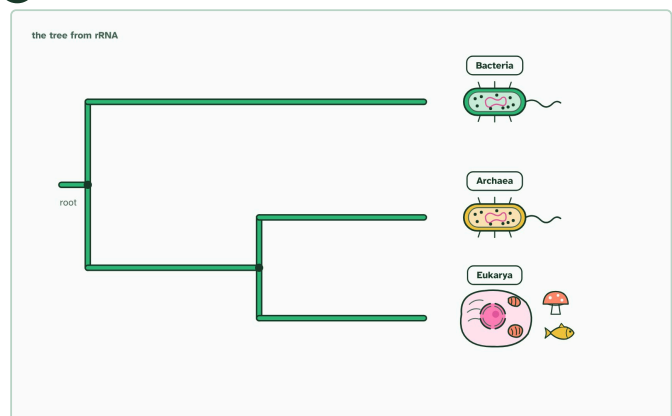
7 BASES DIFFER



In this made-up stretch the archaeon differs from the bacterium at 9 bases but from the eukaryote at 7: **not a bacterium at all**, and a little nearer to us.

4

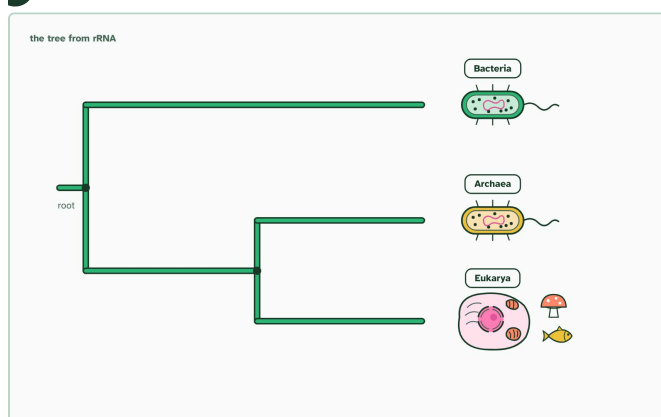
THREE DOMAINS



So the tree of life has three main branches, the **three domains**: Bacteria, Archaea and Eukarya. Archaea and eukaryotes share a more recent common ancestor.

5

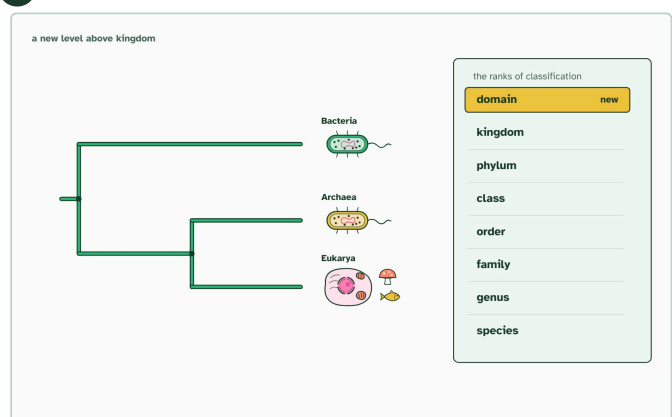
THREE DOMAINS



So the tree of life has three main branches, the **three domains**: Bacteria, Archaea and Eukarya. Archaea and eukaryotes share a more recent common ancestor.

6

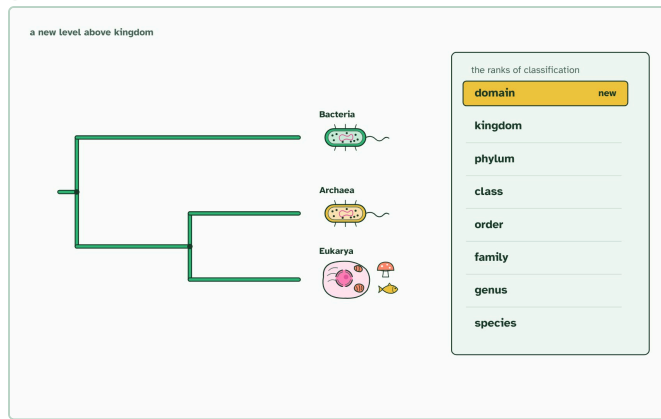
1977



Woese and Fox proposed the three groups in 1977: a new level above kingdom. In 1990 they were named **domains**.

7

1977



Every organism is now classed first by its **domain**, then kingdom, phylum, class, order, family, genus and species.

Watch it move, and change it as it runs: biologybybradford.com/guides/ib-biology-a3-2-classification-and-cladistics#domains

WHERE THE KINGDOMS WENT

Animals, plants and fungi are all in the domain Eukarya. The kingdoms did not vanish: they sit one level below the domains.

WATCH OUT

"Prokaryote" does not name a clade. Bacteria and Archaea are both prokaryotes, but neither is the other's closest relative. Many archaea live in hot springs and salty lakes, but not all do: some live in the guts of animals.

PREDICT

Why was rRNA a good choice for comparing all of life?

- (A) Every organism has ribosomes, so every organism has rRNA that can be compared
- (B) rRNA changes so fast that even close relatives look completely different from each other
- (C) rRNA carries the code for the proteins that every organism uses to copy its DNA

Decide before you look. Answer on page 33.

Key terms

Every term this guide defines, A to Z, with the chapter it is in. Cover the right-hand side and test yourself.

clade 04

A group made of one common ancestor and all of its descendants.

cladistics 02

A method of classifying organisms into clades, groups made of an ancestor and all its descendants, with no fixed ranks.

cladogram 06

A branching diagram that shows the proposed order in which groups split from common ancestors.

classification 01

Sorting organisms into groups, and groups into larger groups, by what they have in common.

common ancestor 03

An ancestor shared by two or more organisms: the point where their lines split.

convergent evolution 08

Unrelated lines evolving similar features, often because they face similar demands, so the likeness does not come from a common ancestor.

domains 09

The highest level of classification, above kingdom: Bacteria, Archaea and Eukarya.

falsified 08

Shown to be false by evidence.

hierarchy of taxa 02

The traditional ranks of classification, one inside the next: kingdom, phylum, class, order, family, genus, species.

molecular clock 05

A way of estimating when two clades diverged from a common ancestor, from the number of differences that have built up in their gene or protein sequences.

morphological 04

The shape and structure of an organism's body.

node 07

A branching point on a cladogram. It stands for a hypothetical common ancestor of the groups beyond it.

paradigm shift 02

A big change in the basic ideas and methods of a field of science, when the old ones can no longer cope with the evidence.

parsimony 06

Choosing, from all possible trees, the one that explains the observed differences with the smallest number of changes.

root 07

The base of a cladogram: the common ancestor of every group on the tree.

rRNA 09

The RNA that is part of ribosomes. Every organism has it, so its base sequence can be compared across all of life.

sister groups 07

Two groups that branch from the same node, so each is the other's closest relative.

taxon 02

Any named group of organisms, at any rank. The plural is taxa.

terminal branch 07

The line from the last node out to a tip, which is a species or group being compared.

Dry lab: *Grow a tree*

Six organisms, a table of features and a stretch of DNA. Choose the data, let the tool build the candidate trees, and see which tree needs the fewest changes.

Six made-up fish, for teaching. River fish is the outgroup.

		traits counted at first				
		1	2	3	4	5
	River fish	☐	☐	☐	☐	☐
	North lake fish	☐	☐	☐	☒	☐
	North cave fish	☒	☒	☒	☒	☐
	South lake fish	☐	☐	☐	☐	☒
	South cave fish	☒	☒	☒	☐	☒
	Marsh fish	☐	☐	☐	☐	☒

1 tiny eyes 2 no pigment 3 taste buds on head
4 barbels 5 fin spine

trees scored

changes needed

The hatch: six made-up fish and their traits. Grow scores every one of the 165 possible trees by the changes it needs, and the graph shows how many trees need each number. Which



Run it online: change the variables, run the trials, and the table and graph fill in as you go. biologybybradford.com/guides/ib-biology-a3-2-classification-and-cladistics#lab

Your question

Does the tree you get depend on the data you choose? When you add more features, or switch from features to sequences, does the winning tree change, and does it ever tie?

What you can change

The data (a table of features or a DNA sequence), the features included, and the number of sites used. The tool counts the changes each candidate tree needs.

What to record

For each run: the data you used, the tree that needed the fewest changes, its score, and the score of the runner-up. Change one thing at a time. The tool gives the same answer for the same data every time, so the experiment is in your choice of data.

Find out

- **Does adding a feature change the tree?** Start with three features and record the winning tree and its score. Add the features one at a time, and record the winner and its score after each. When, if ever, does the winner change?
- **Features or sequence?** Run the six organisms on the table of features, then switch to the DNA sequence. Do both give the same tree? Which gives the clearer winner, and why might that be?
- **Can two trees tie?** Look for a set of features or sites where two trees need the same number of changes. What would you do next, as a scientist, to break the tie?
- **A feature that fits no tree well.** Add a feature that two distant organisms share, as wings are shared by a bat and a bird. What does it do to the scores, and can enough of them change the winner?

Record

- For every run: the data (features or sequence), how many features or sites you used, the tree that needed the fewest changes, its score, and the score of the runner-up.
- Change one thing at a time and keep the rest fixed.
- The tool counts changes by a fixed rule, so the same data always gives the same answer. There is no scatter to average: the experiment is in the data you choose, so repeat a run only to check you recorded it correctly.
- Keep a table of your runs. Compare it with a partner's: did you reach the same winning tree from the same data?

QUESTIONS ON THE LAB

QUESTION 1

You add a feature that only one of the six organisms has. What does it do to the scores?

- (A) It makes the tree that puts that organism next to its closest relative score lower than the rest
- (B) It adds one change to every tree, so it cannot change which tree wins
- (C) It makes that organism the root of the winning tree, because it is different

Answer on page 33.

QUESTION 2

You add several features that two distant organisms share, as a bat and a bird share wings. What is the likely effect on the winning tree?

- (A) The best trees score worse, and with enough such features a different tree can win
- (B) Nothing changes, because the tool ignores features that are shared by distant organisms
- (C) The two organisms always join on the winning tree, because they share features

Answer on page 33.

QUESTION 3

The tool gives exactly the same answer every time you give it the same data. Why is that not a weakness of the model?

- (A) Real classifications are always certain, so a good model should never vary at all
- (B) The tool only works on data it has seen before, so the answer is stored in advance
- (C) Counting follows a fixed rule, so the uncertainty lies in the data chosen

Answer on page 33.



Check yourself

The check for this guide is online: a fresh paper of past-paper style questions from the exam bank on A3.2 each time, marked as you go, and saved to your Exam practice progress when you are signed in. Scan the code, or type the link.

[biologybybradford.com/exam-maker/sit?
topics=A3.2&level=HL&paper=quiz&marks=15&src=ep](https://biologybybradford.com/exam-maker/sit?topics=A3.2&level=HL&paper=quiz&marks=15&src=ep)

Answers

PREDICTIONS IN THE CHAPTERS

- 01** Which is the best reason for classifying organisms?
- B** **There are too many species to study one by one, and a group lets what we know about one guide the rest**
- The second. Sorting makes millions of species manageable, and it lets what is found out about one species guide the study of its group. Names are useful but they do not do that, and ranking species from simple to advanced is not what classification is for: no living species is more advanced than another.
-
- 02** A taxonomist finds that a group in the old system leaves out one species that descends from the same common ancestor as the rest. What does that show?
- C** **The group does not follow the pattern of divergence**
- The group does not follow the pattern of divergence. A group that leaves out a descendant of its own ancestor splits a family in two, however alike its members look. That is the problem with the old class Reptilia, which leaves out the birds.
-
- 03** A newly found frog falls inside a clade whose other members all make a toxin in their skin. Which statement is best?
- C** **It probably makes the toxin, but a feature can be lost, so the prediction needs testing**
- The second. Membership of the clade makes the toxin likely, because it was probably present in the common ancestor. It is not certain, since a line can lose a feature, and that is why a prediction is tested and not simply assumed.
-
- 04** A tree shows an ancestor X with four descendants, P, Q, R and S. Someone proposes the group P, Q and R. Is it a clade?
- A** **No, because S also descends from X and is left out**
- No. A clade holds the ancestor and all of its descendants, and S descends from X as well. P, Q and R do share X, but the group is missing a descendant, so it does not follow the branching.
-
- 05** A clock is calibrated on mammals that breed every few years. It is then used to date a split between two lines of insects that breed several times a year. What will the estimate probably be?
- C** **Too old, because the insects probably collected changes faster each year**
- Too old. With more generations each year there are more chances for changes to build up, so the insects gather differences faster than the mammal clock allows for. The same number of differences stands for fewer years than the clock says.
-
- 06** Three possible trees need 14, 11 and 11 changes. What does parsimony tell you?
- A** **The two trees that need 11 are tied, so more data is needed to choose between them**
- The two 11-change trees tie as most parsimonious. Parsimony cannot choose between them, so a scientist adds data, such as more positions or a second gene, and counts again. A tie is also a reminder that a cladogram is a hypothesis.
-
- 07** In a cladogram, species X and Y branch from node 1, and node 1 and species Z branch from node 2. Which statement is correct?
- B** **X is closer to Y than to Z, because X and Y share a more recent ancestor**
- X is closer to Y than to Z. X and Y share node 1, which is more recent than node 2, the ancestor they share with Z. Neither X nor Y is the ancestor of the other: node 1 is a hypothetical ancestor of both.
-

Q8 Chloroplast DNA places two plants with almost identical flowers in different clades. What is the best explanation?

C **Their flowers probably evolved separately, by convergent evolution**

Their flowers probably evolved separately. Similar structures can arise in unrelated lines, so a classification built on them can be false. DNA is the more objective evidence, and where the two disagree, sequences are used to decide.

Q9 Why was rRNA a good choice for comparing all of life?

A **Every organism has ribosomes, so every organism has rRNA that can be compared**

Every organism has ribosomes, and so has rRNA genes, so the same molecule can be compared across bacteria, archaea and eukaryotes. That is what allowed one tree of all life. rRNA does not code for proteins: it is part of the ribosome that makes them.

QUESTIONS ON THE DRY LAB

Q1 You add a feature that only one of the six organisms has. What does it do to the scores?

B **It adds one change to every tree, so it cannot change which tree wins**

It adds one change to every tree. A feature found in only one organism arose once, on that organism's own branch, whichever way the other five are arranged, so it cannot separate the trees. Only features shared by two or more organisms help, which is why the sites where just one species differs were no help in chapter 6.

Q2 You add several features that two distant organisms share, as a bat and a bird share wings. What is the likely effect on the winning tree?

A **The best trees score worse, and with enough such features a different tree can win**

The best trees score worse, because on a tree where the two are not close each shared feature has to arise twice. With enough of them, a tree that puts the two together can win, even though they are not close relatives. That is convergent evolution misleading a classification, as it did for the figwort family (chapter 8).

Q3 The tool gives exactly the same answer every time you give it the same data. Why is that not a weakness of the model?

C **Counting follows a fixed rule, so the uncertainty lies in the data chosen**

Parsimony is a fixed calculation, so the same data always gives the same count. What varies in real work is the data: which gene, which features and how many sites. That is why different criteria and different data can lead to different hypotheses, and why a cladogram is a hypothesis that new data can change.

How this was made. The course structure, the activities and the teaching choices come from my own classroom, my presentations and my years of teaching IB and AP Biology. I use Claude, an AI, to help write and build the interactive guides, lesson plans, slides and practice questions, and to draw the diagrams and code the interactives. I direct the work and review it against the IB and AP specifications before it goes live.

Written and drawn for Biology by Bradford, biologybybradford.com. Free to print and share with a class.