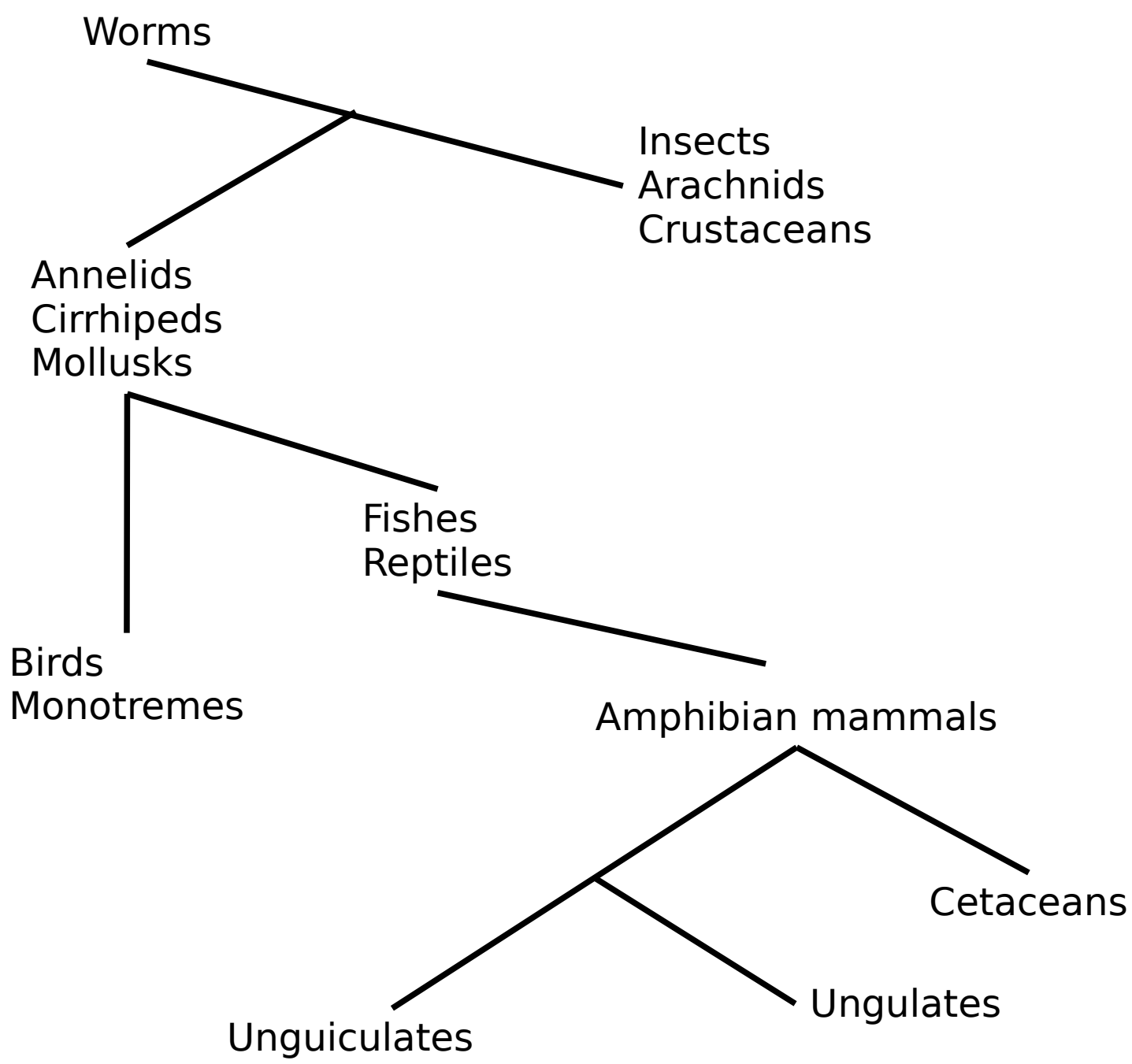


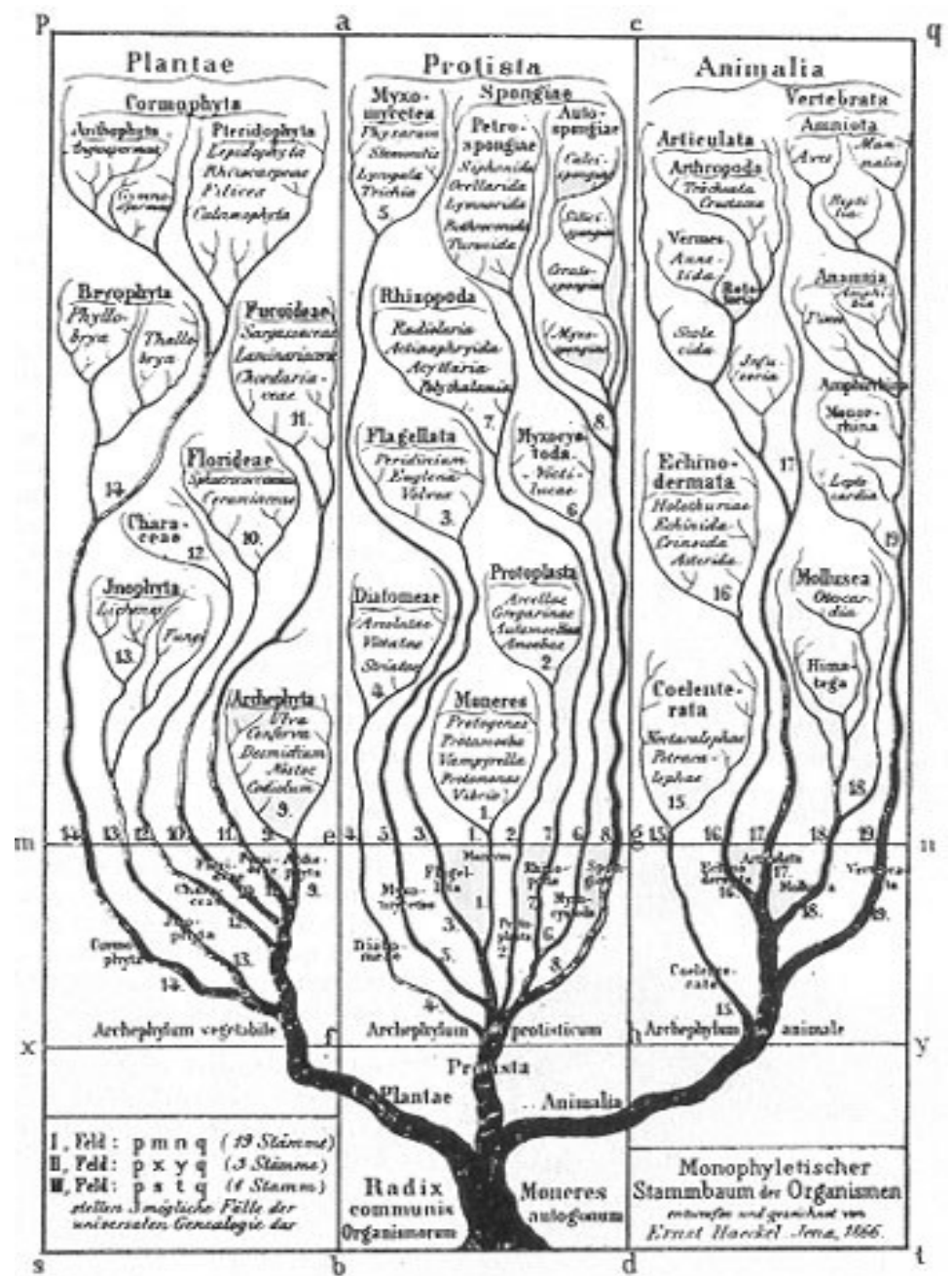
Phylogeny

Term invented by Ernst Haeckel in 1866 in order to define the relationship between animal and plant species across time

Darwin (1872) a phylogeny is the genealogical relationships between all organisms



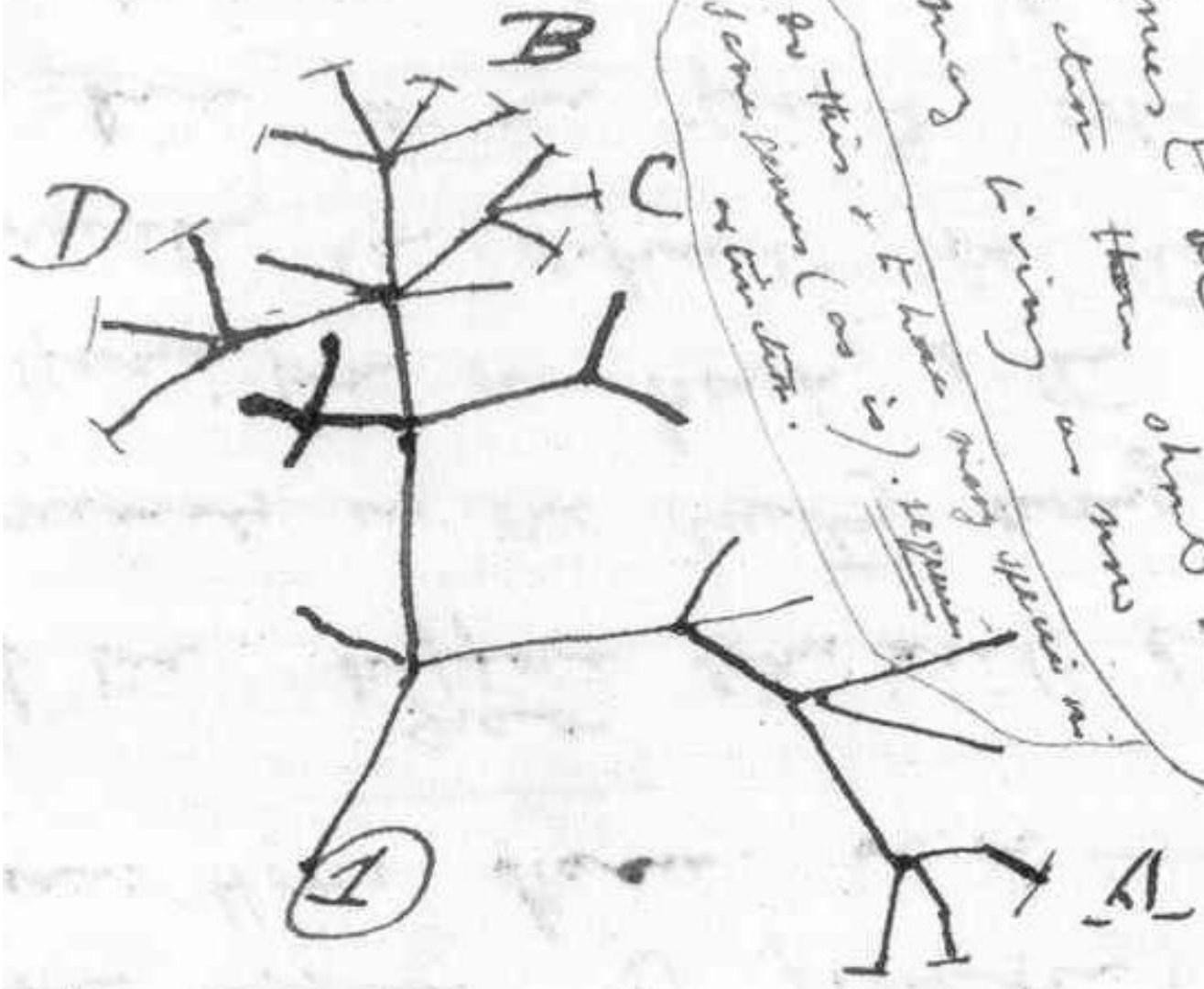
From Lamarck 1809



From Haeckel 1866

I think

There are
obviously
more than
one
species
of
living
monkeys
in
the
Andes
I think
there
are
many
species
of
monkeys
in
the
Andes
I think
there
are
many
species
of
monkeys
in
the
Andes



From Darwin's notebooks

Goals of phylogenetic reconstruction

- Try to explain the evolutionary history of actual characters (morphological traits, genes)
- Show how species are relating to one another
- A phylogenetic tree is a mathematical structure used to model the evolutionary history of a trait

Phenetic - Cladistic

- **Phenetic:** classification of the organisms based on their similarities, trees obtained using a phenetic approach may not reflect evolutionary relationships. A tree based on this method is called a **phenogram**
- **Cladistic (Hennig 1966):** study of the different pathways of evolution, the most parsimonious pathway will be retained to build the best possible tree called a **cladogram**.

Morphological characters

- Morphological
 - easy to access
 - can be influenced by external factors
 - ambiguous
 - qualitative argumentation
 - codification of the characters
 - homology difficult to assess
 - relationship between closely related species
 - sometimes visible morphology can be lacking

Molecular characters

- Molecular
 - heritable
 - only 4 or 20 character states per site (relatively unambiguous)
 - no obvious polarity in their evolution
 - predictable evolution
 - large number of characters (about 17,000 bp in human mitochondrial DNA)
 - independent characters (probably not but better than morphological characters)

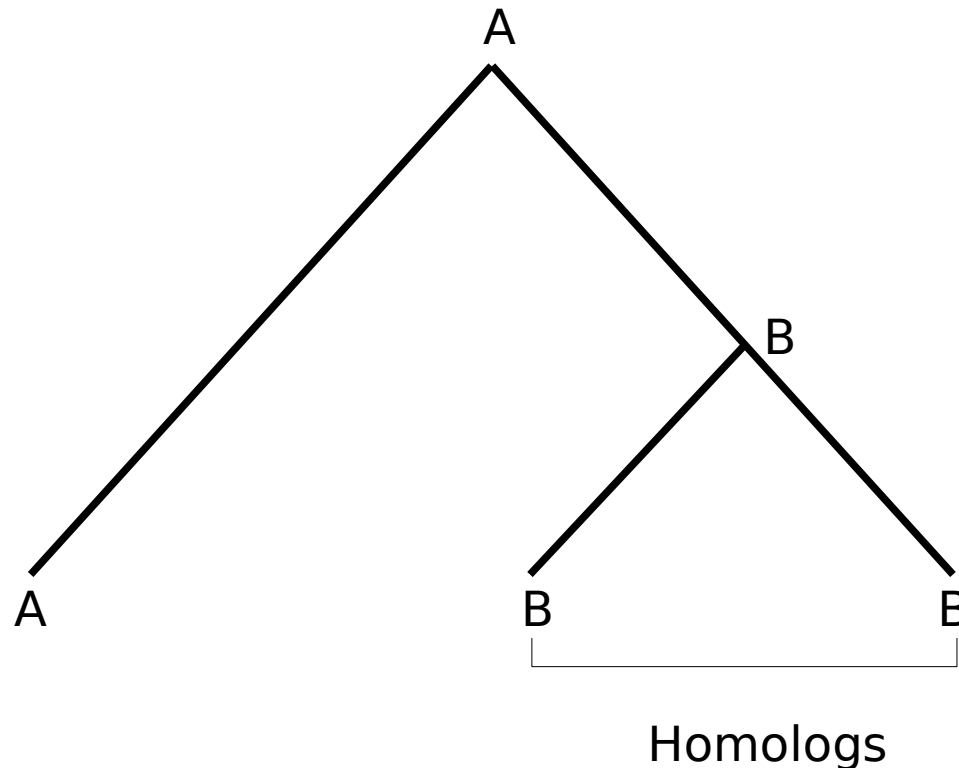
Molecular characters

- Molecular
 - homology can be inferred without bias
 - large body of knowledge (and still accumulating) about how the characters change evolutionarily
 - sequences are known to evolve at different rates
 - rRNA changes very slowly
 - histones changes very slowly
 - immunoglobulin genes changes quickly
 - animal mtDNA changes quickly
 - micro/minisatellites changes very quickly
 - relationship with distantly related species can be inferred
 - easily generated (PCR)

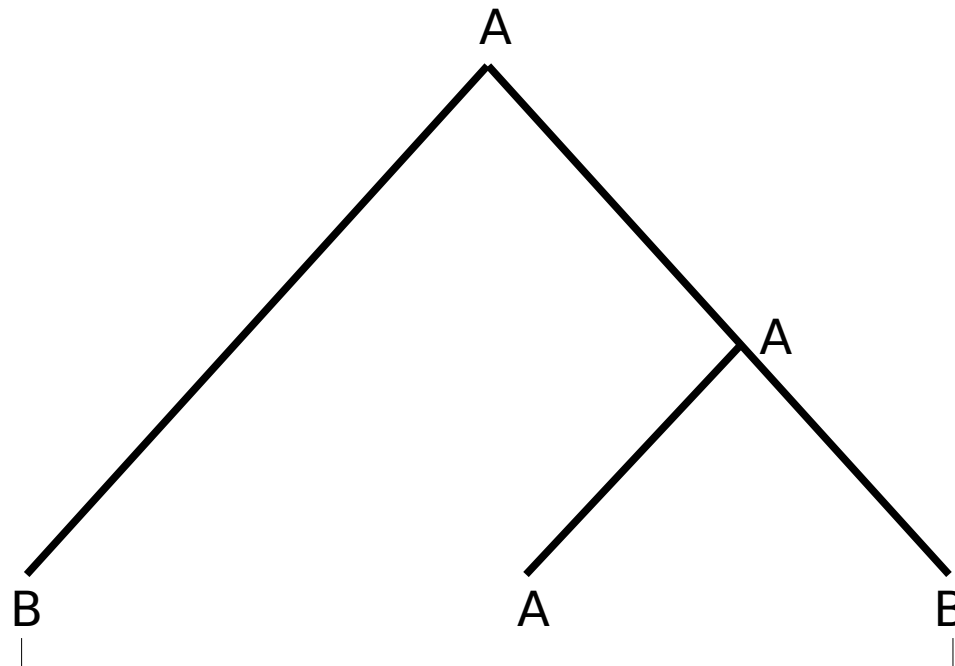
Definitions

- Homology
- Analogy
- Homoplasy
- Convergence
- Reversion
- Orthology / Paralogy
- Character classification

- **Similarity:** resemblance between two characters
- **Analogy:** similar in form or function without sharing an ancestry
- **Homology:** Two traits are homologous if they are derived (with or without modifications) from a common ancestor.



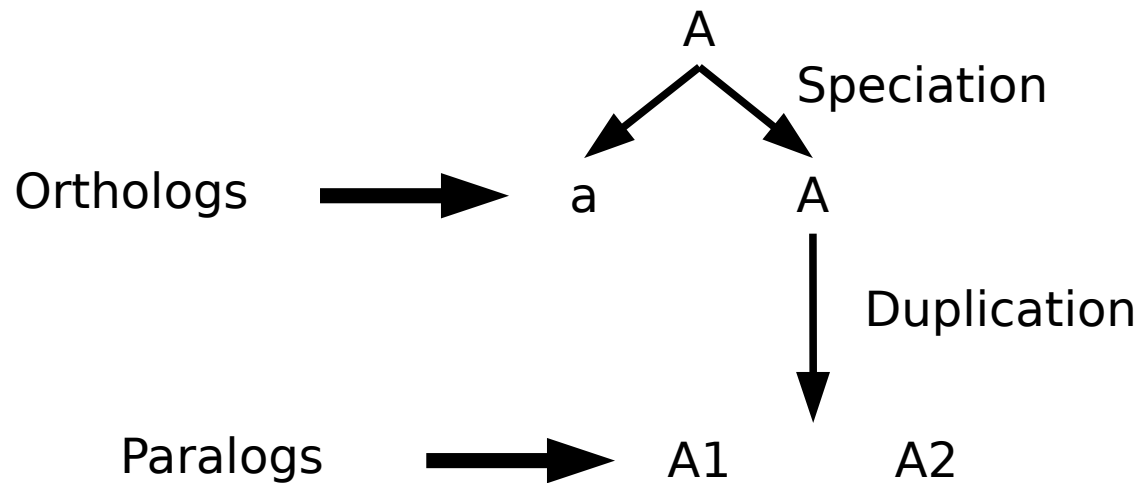
- **Homoplasy:** independent presence of similar characters between species



Homoplasy

- **Convergence:** process whereby non related organisms are evolving similar traits independently due to similar environmental pressures for example.
- **Reversion:** return of a character to one of its ancestral states
- **Parallelism:** identical changes in two or more lineages.

- Homologous traits or sequences can be:
 - orthologs - homologous sequences are orthologous if they are separated by a speciation event
 - paralogs - homology by duplication
 - xenologs - homology through lateral gene transfer

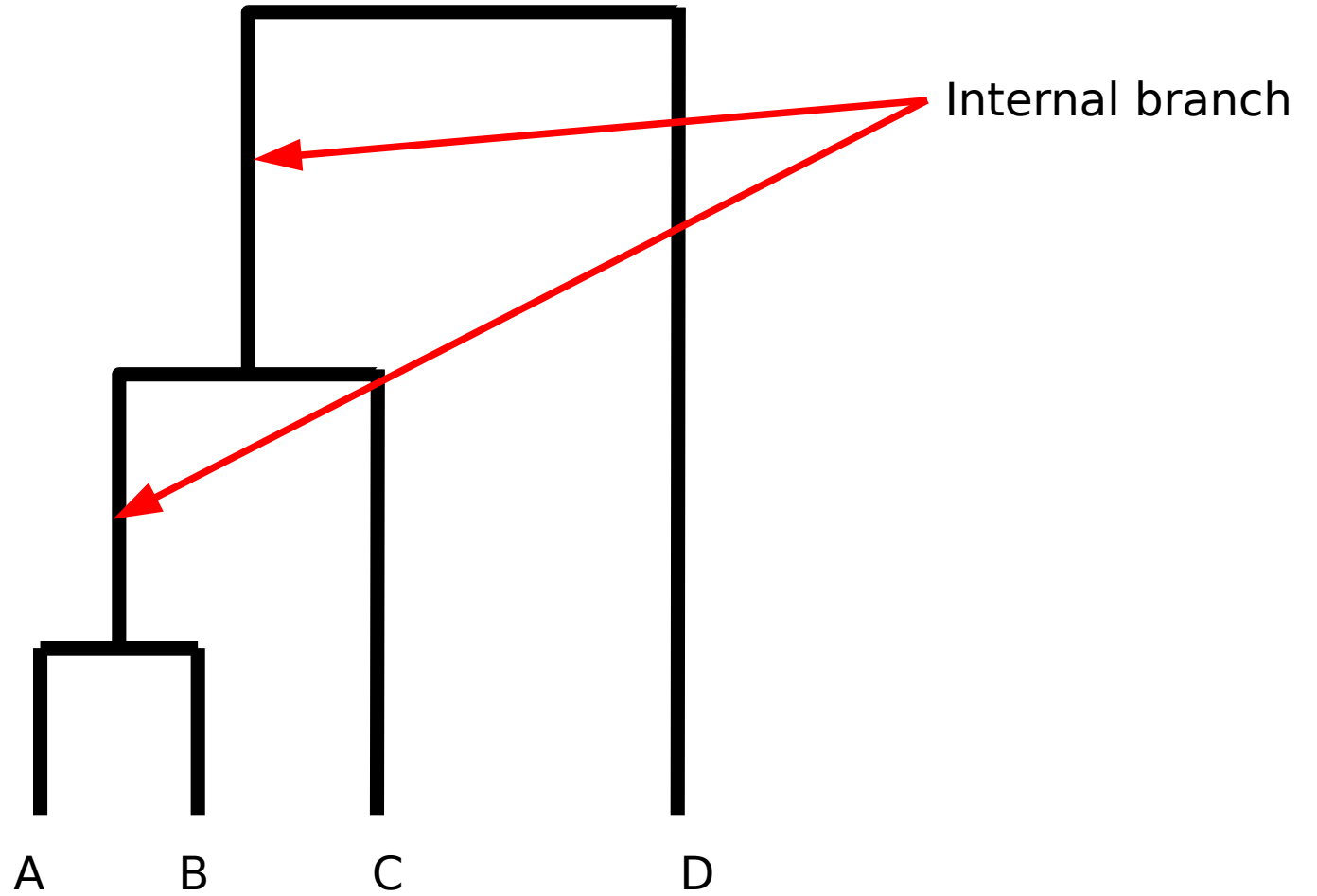


- **Plesiomorphy:** primitive or ancestral character state
 - Primates all have hair.
 - Can we cluster primates as a group based on this characteristic?
- **Apomorphy:** derived state representing an evolutionary novelty
 - Humans are bipedal.
 - Can we phylogenetically place humans based on this characteristic?

- **Symplesiomorphy:** primitive state shared by several taxa
 - Hair is a shared primitive state for all primates
- **Autapomorphy:** derived character state unique to a taxa
 - Humans are uniquely bipedal among primates

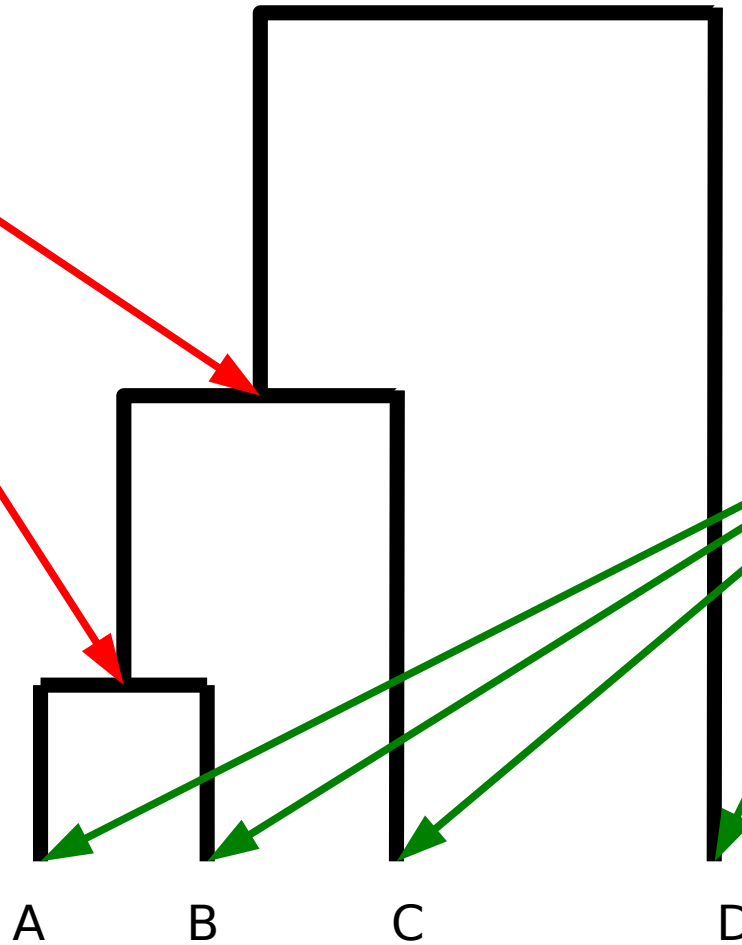
- **Synapomorphy:** derived character state shared by several taxa
 - only these can be used to decipher relationships

Some vocabulary



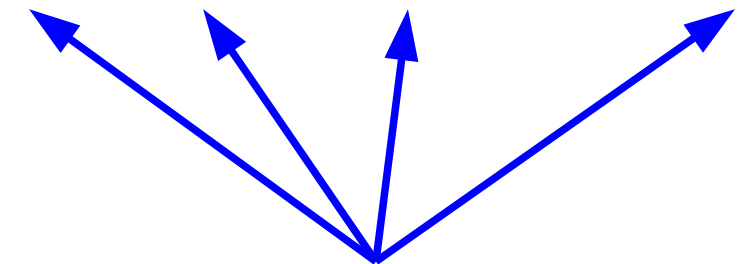
Some vocabulary

Internal node

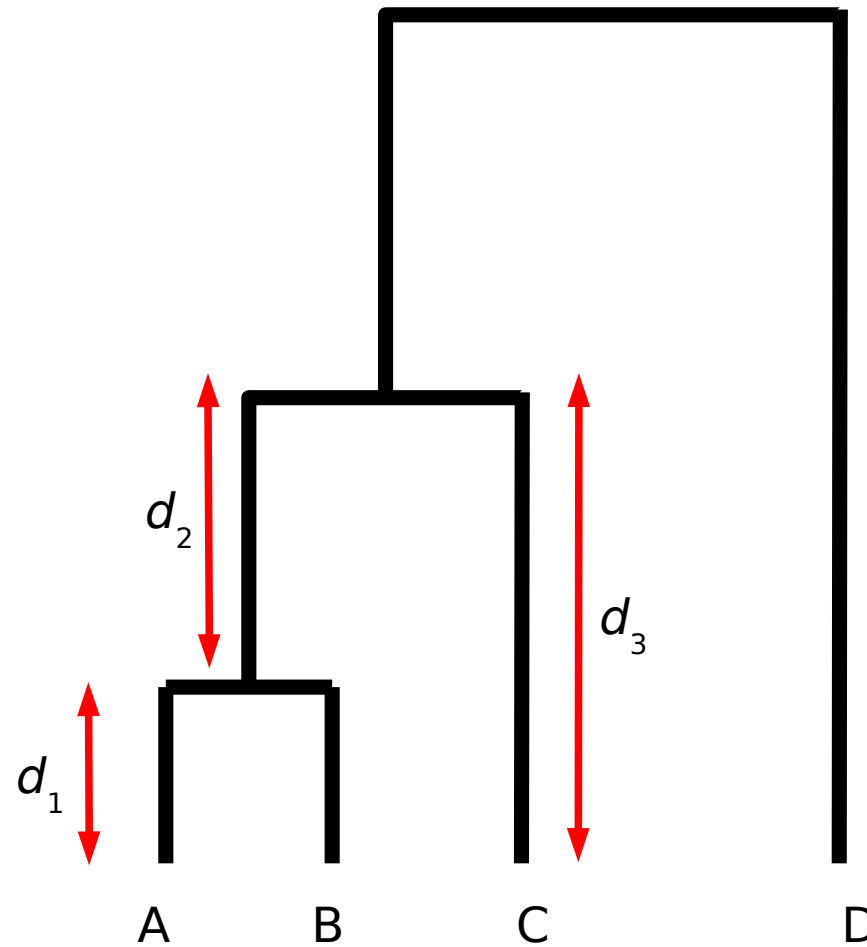


Terminal nodes
Leaf/Leaves

Operational Taxonomic Unit (OTU)



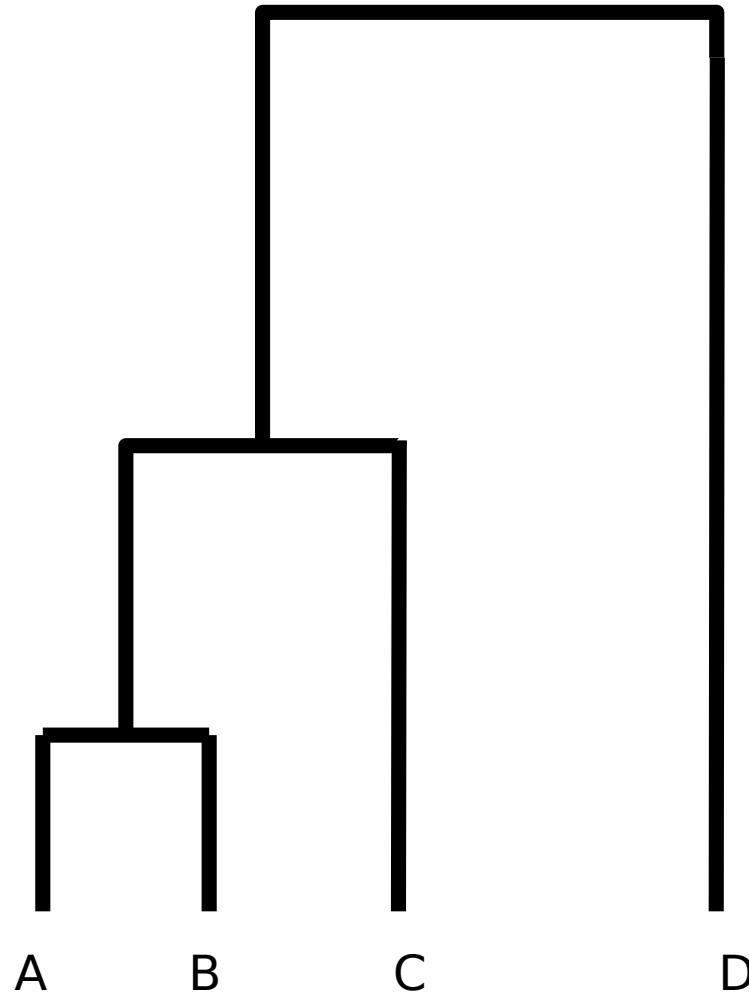
Additive branch length



Sum of the branch lengths between 2 OTUs

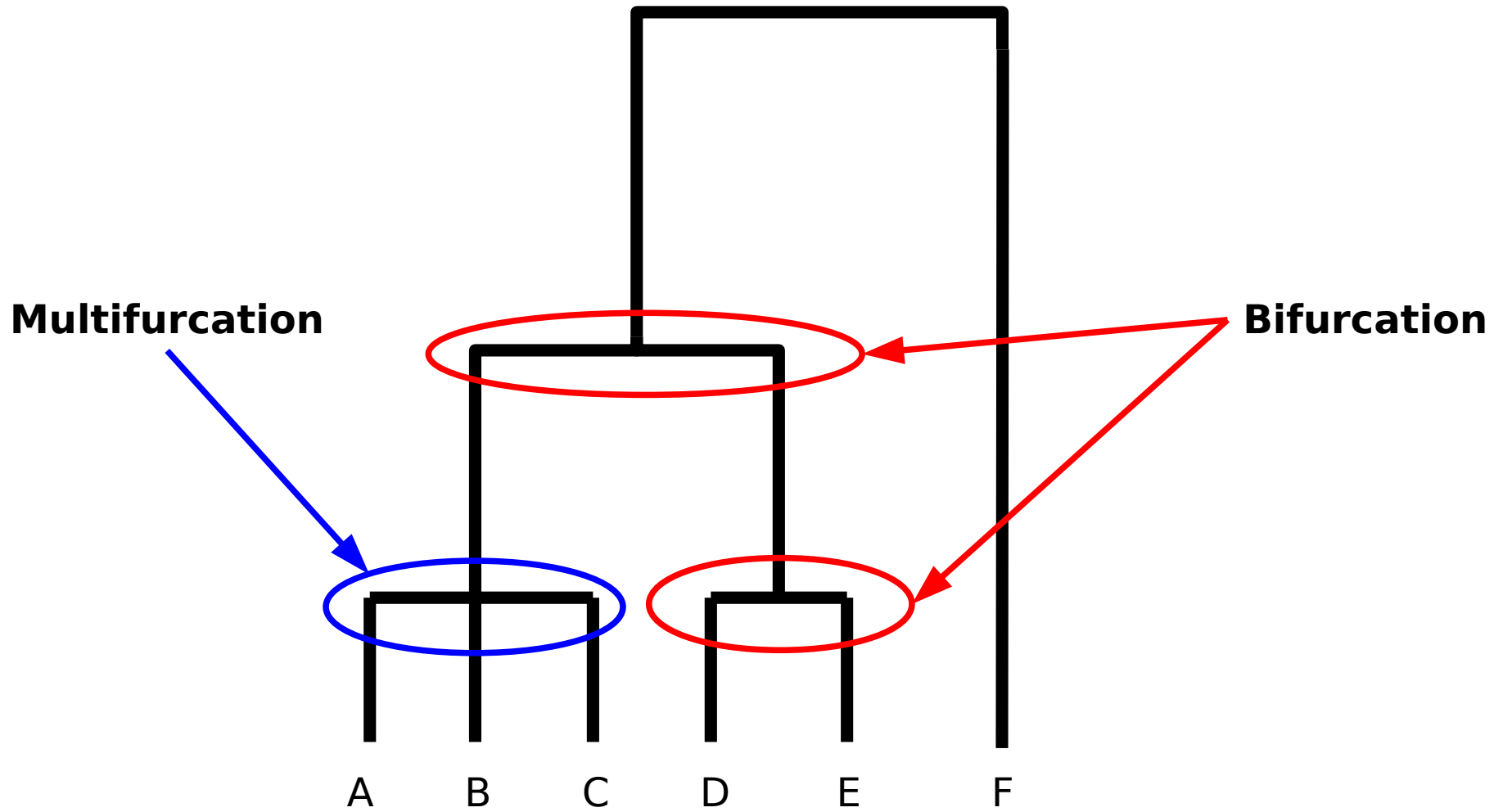
$$D_{AC} = d_1 + d_2 + d_3$$

Ultrametric branch lengths

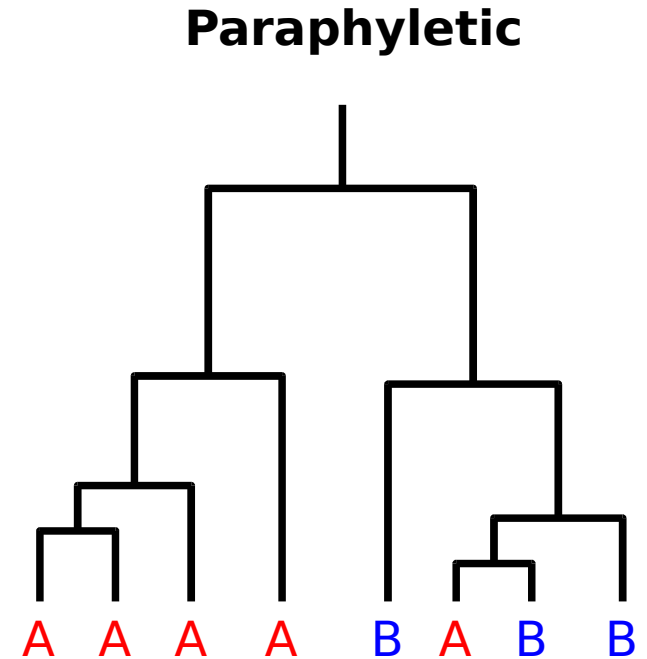
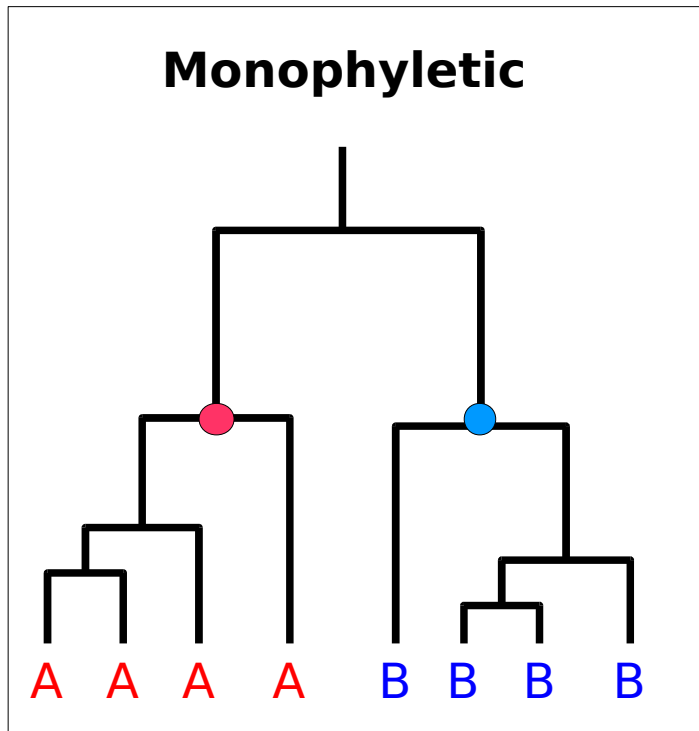


When the distances from any two leaves to their shared common ancestor are equal (as shown here).

Bifurcation / Multifurcation



Monophyletic / Paraphyletic

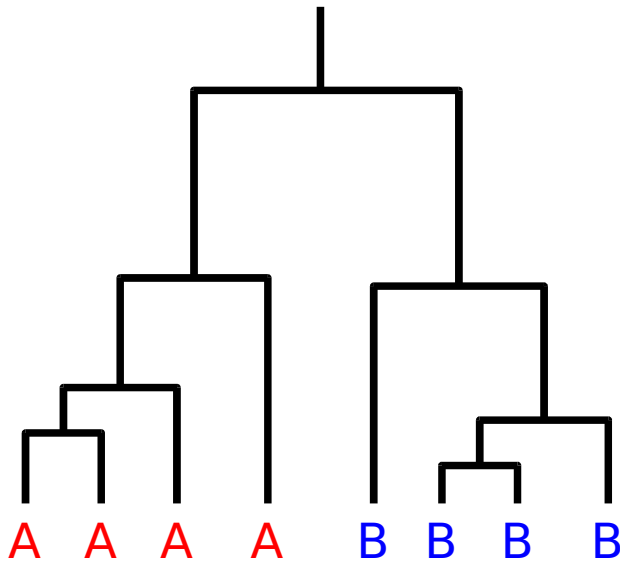


- in the first diagram A and B are **monophyletic groups**; all taxa share a common ancestor and all descendants of that ancestor are members of the group.

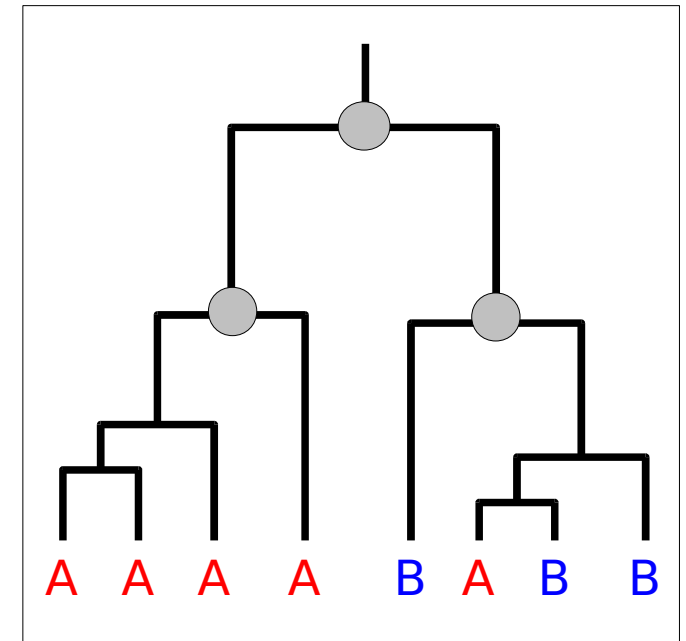
- in the second diagram, B is a **paraphyletic group**; all taxa share a common ancestor but not all descendants of that ancestor are members of the group.

Monophyletic / Paraphyletic

Monophyletic



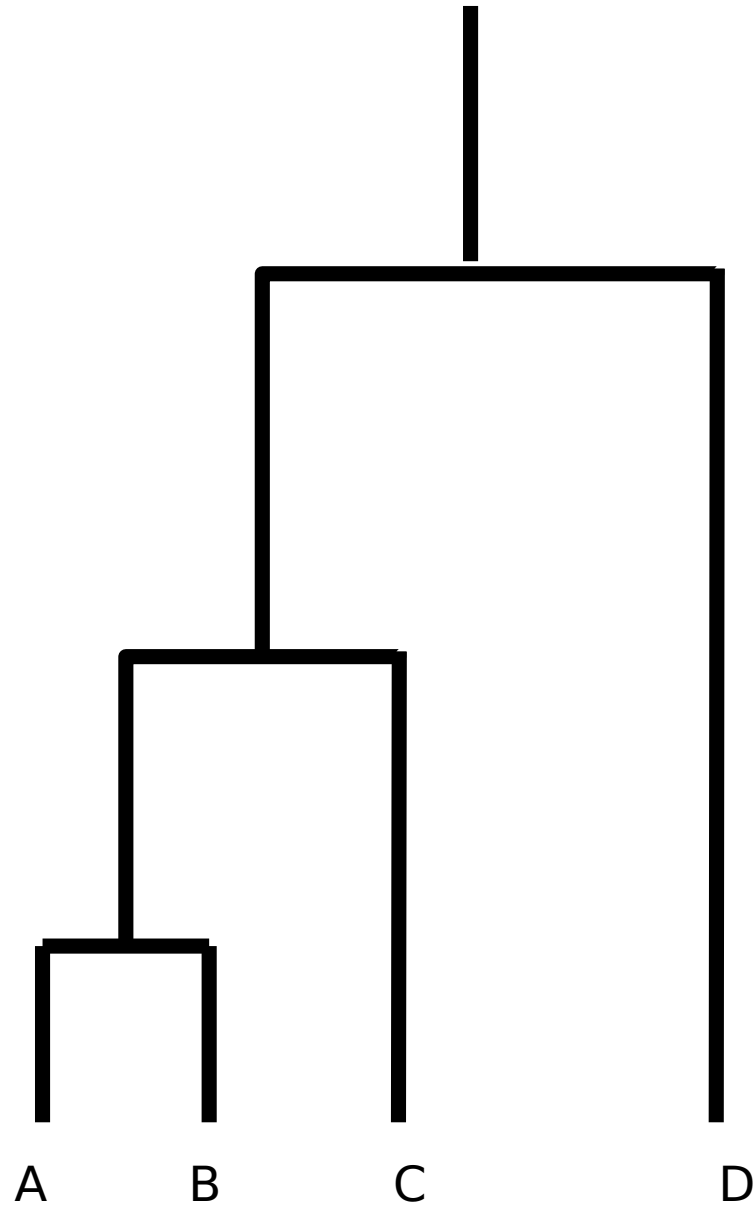
Paraphyletic



- in the first diagram A and B are **monophyletic groups**; all taxa share a common ancestor and all descendants of that ancestor are members of the group.

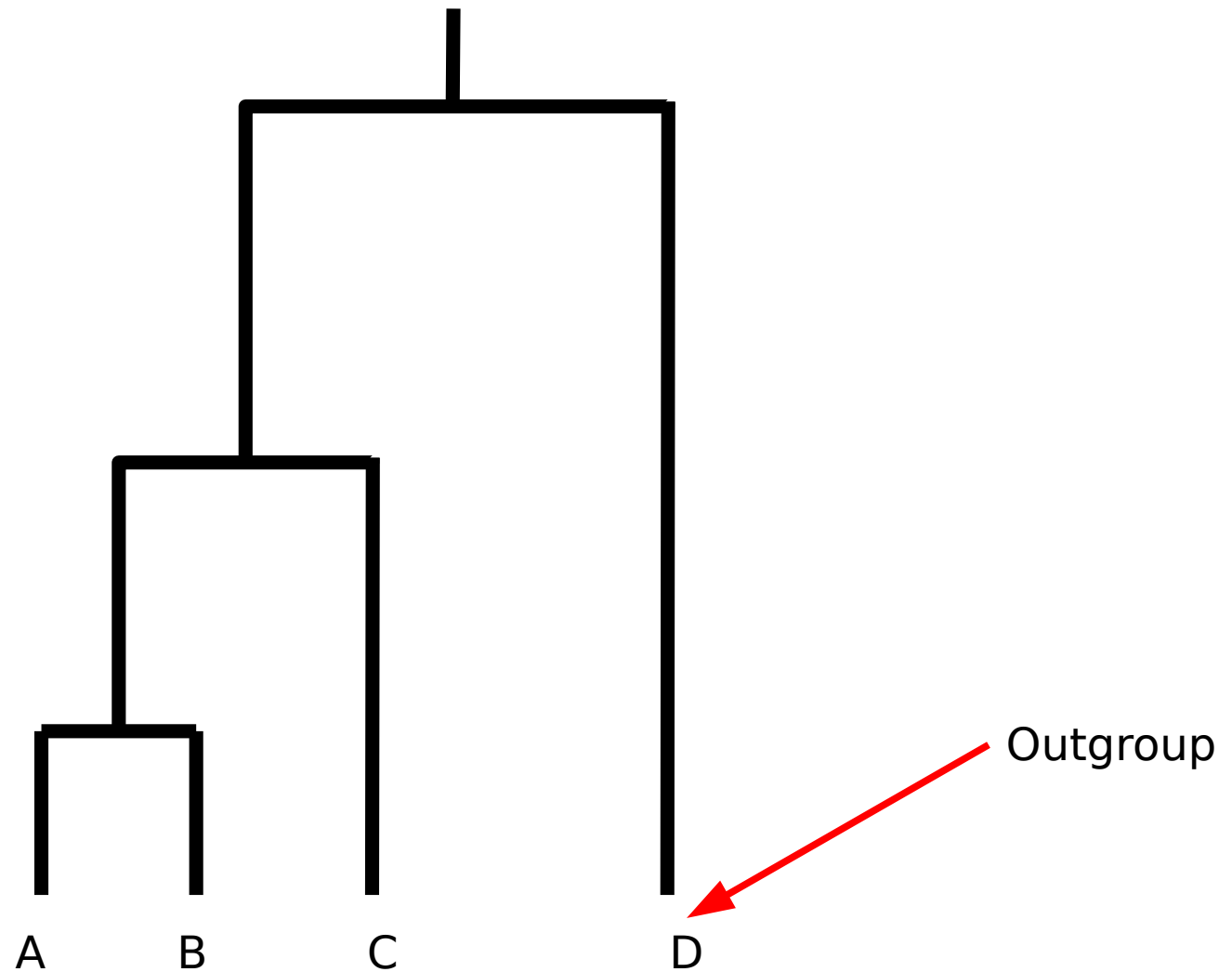
- in the second diagram, B is a **paraphyletic group**; all taxa share a common ancestor but not all descendants of that ancestor are members of the group.

Rooted tree

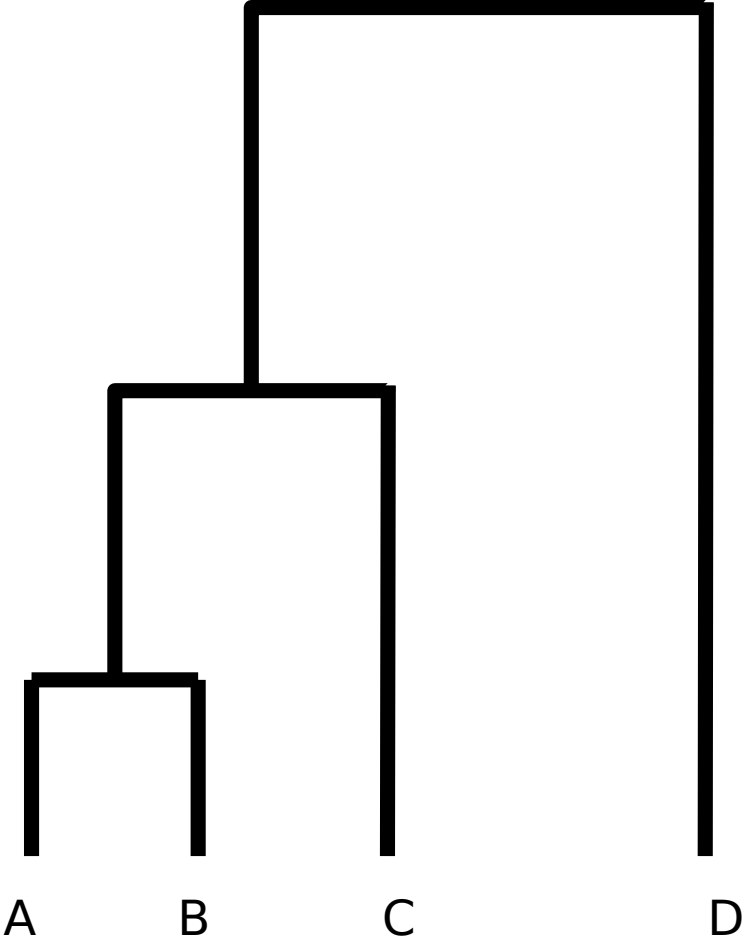


Outgroup

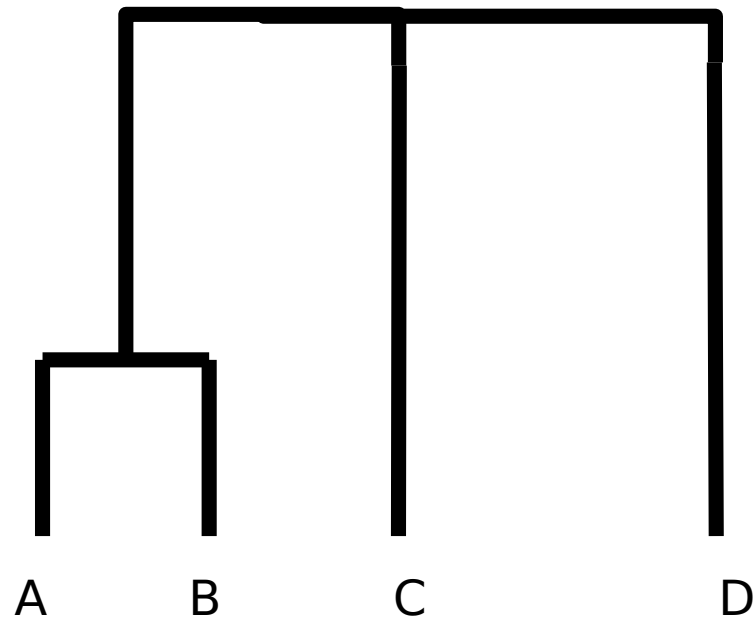
Outgroup: sequences related to the group of sequences compared but not part of this group



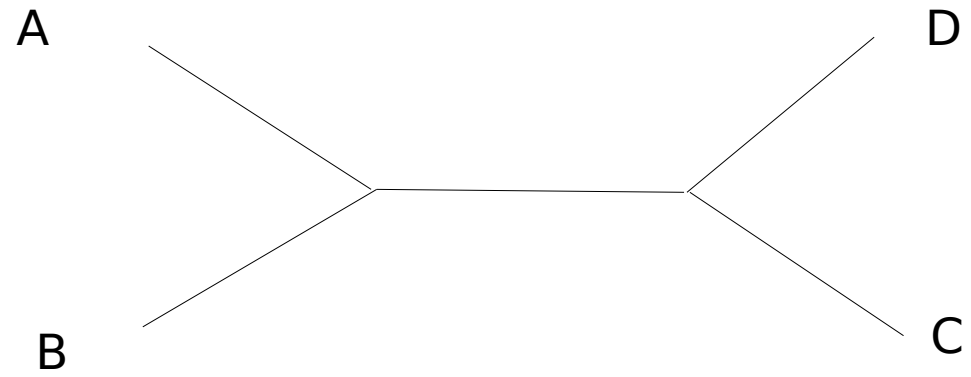
Unrooted tree

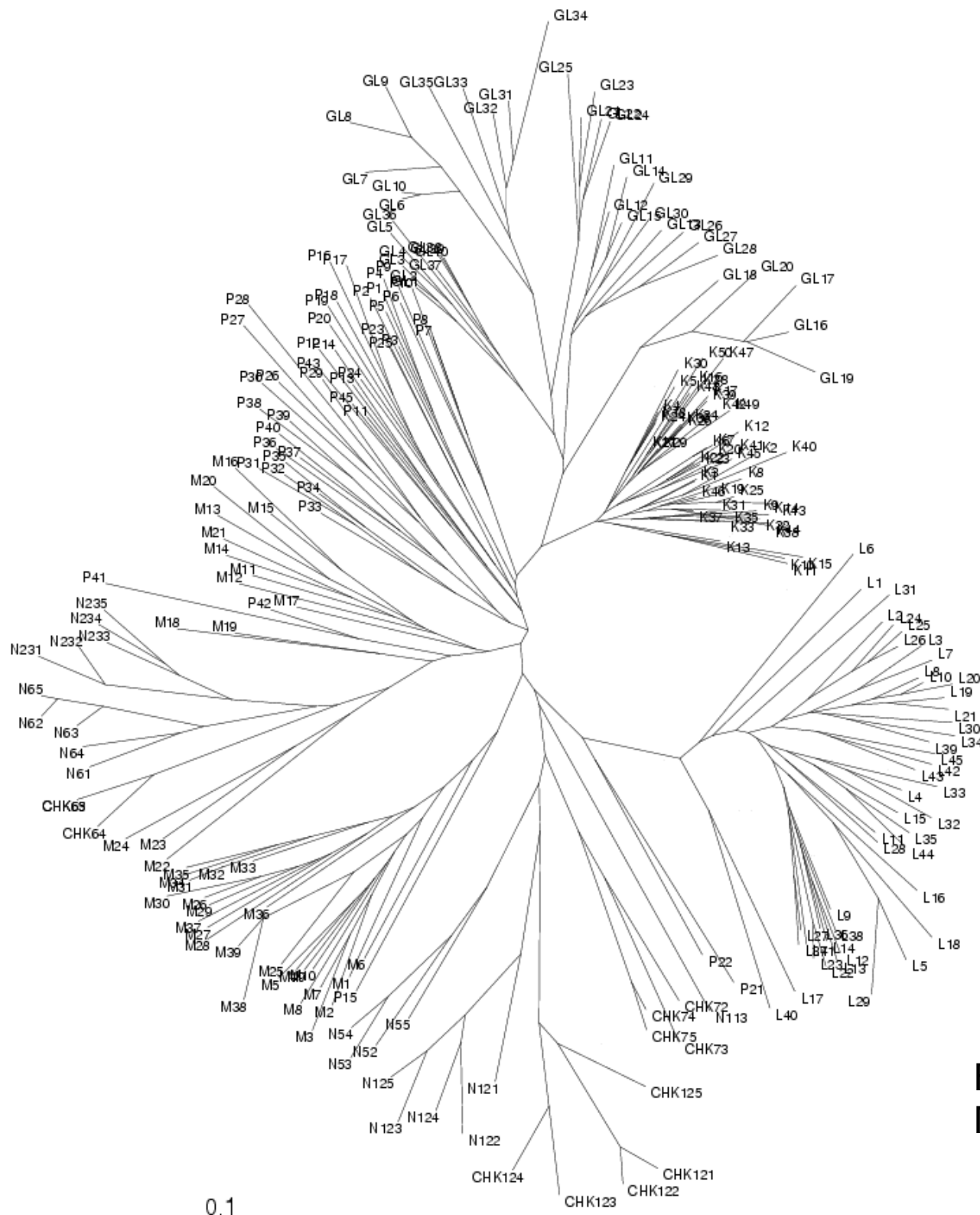


Unrooted tree



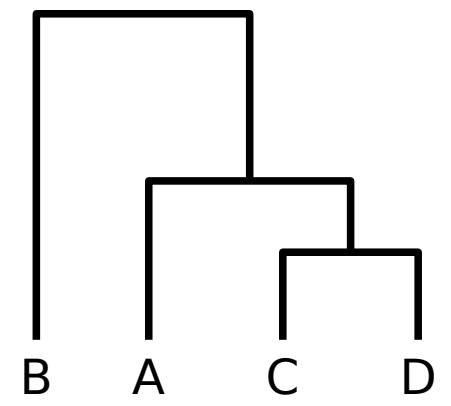
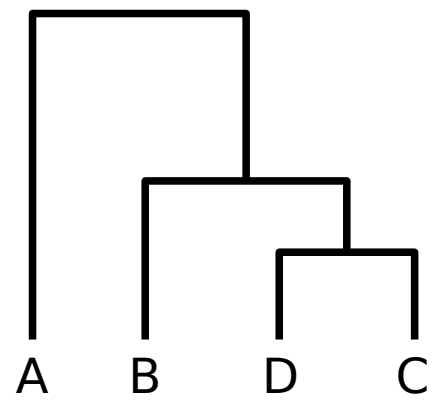
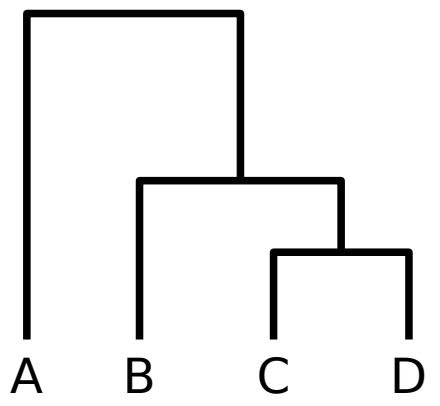
Unrooted tree

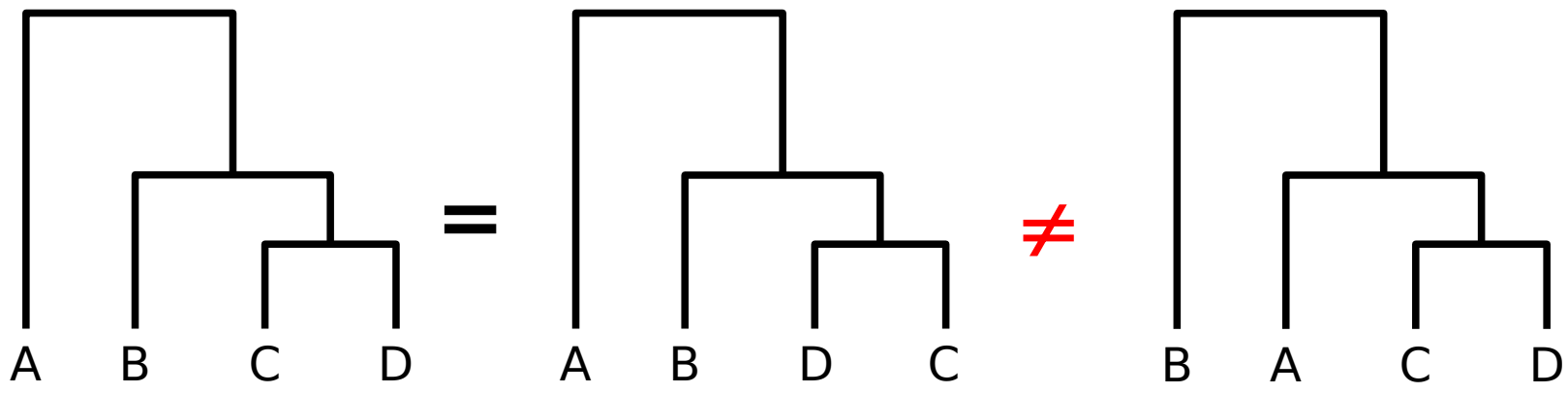


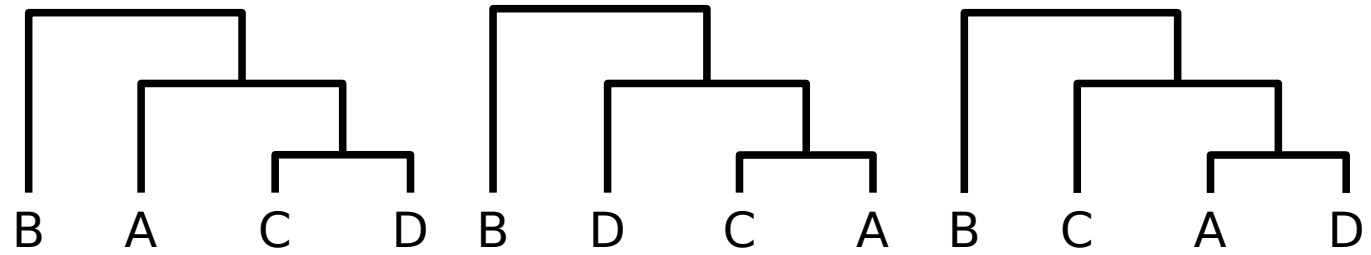
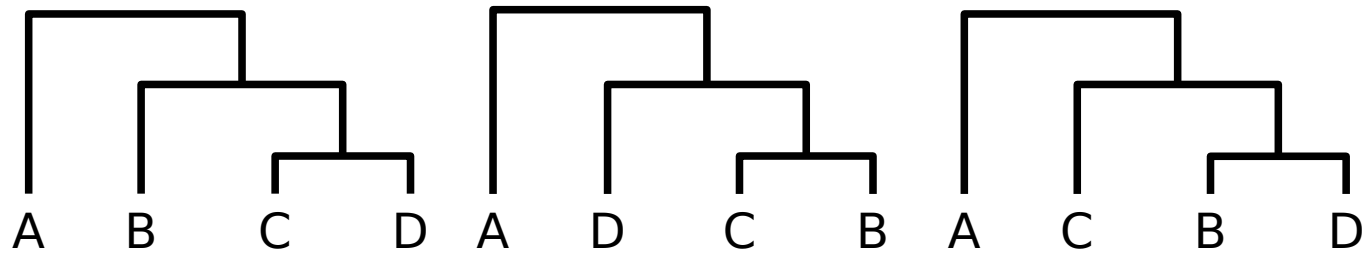


From Haerty et al. 2005.
Mol. Ecol. 14:3801-3807

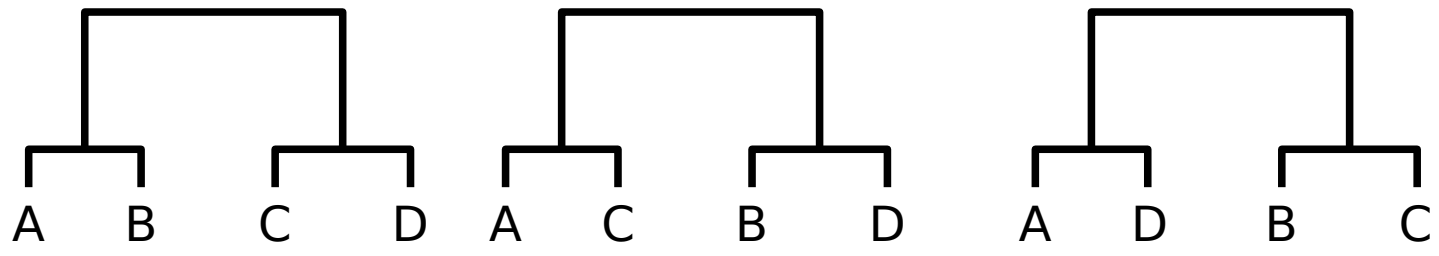
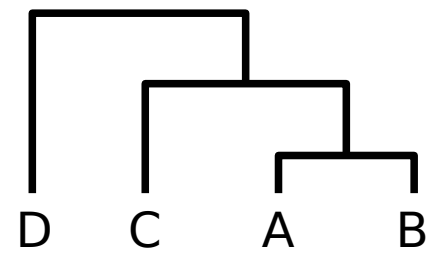
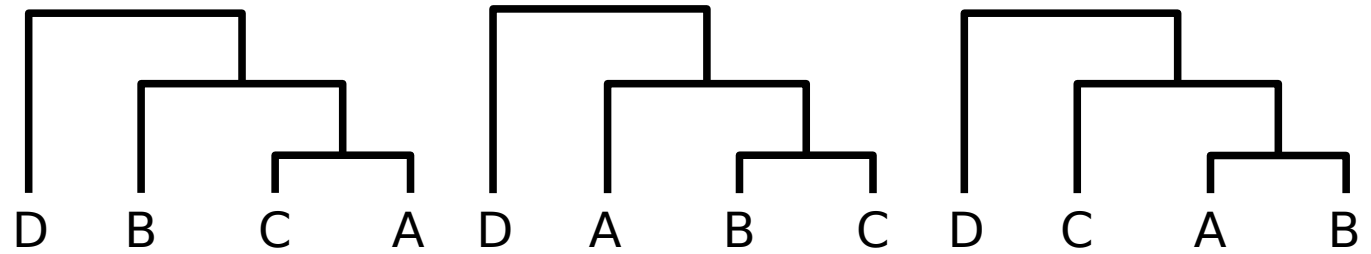
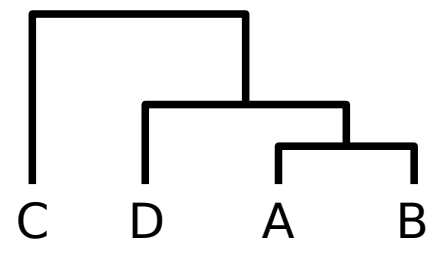
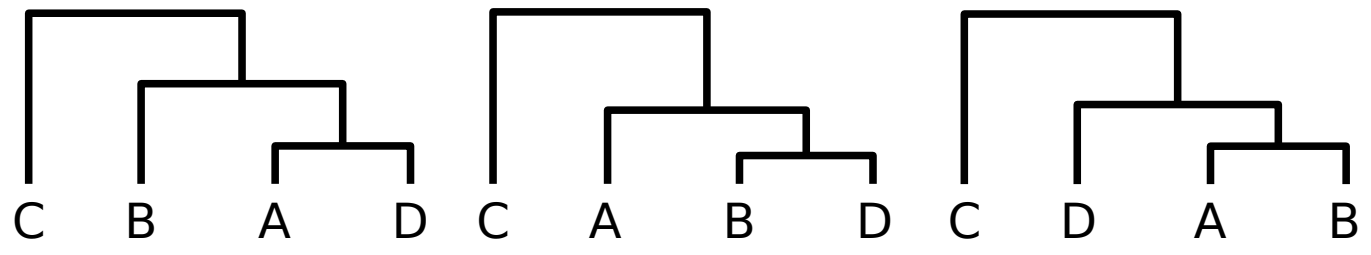
- In a **rooted tree** a single node is designated as a common ancestor giving thereby a unique pathway is connecting this node to any other through evolutionary time
- An **unrooted tree** is only displaying the relationship between the nodes, no information about directionality is given.







4 species
 =
15 trees



Number of nodes

Rooted tree

$$N = (2n-3)! / [2^{n-2}(n-2)!]$$

Unrooted tree

$$N = (2n-5)! / [2^{n-3}(n-3)!]$$

n: number of species

For a rooted tree

→	n= 15	N= 213,458,046,676,875
	n= 20	N= 8,200,794,532,637,891,559,375

→ Only one of these tree is the “true tree”, all the other potential trees must be rejected