Overview of systematics and phylogenetics

Pakorn Aiewsakun

Department of Microbiology

Faculty of Science, Mahidol University

Outline

- What is systematics?

- What is phylogenetics?

- How to read a phylogeny?

- How to reconstruct a tree?

What is systematics?
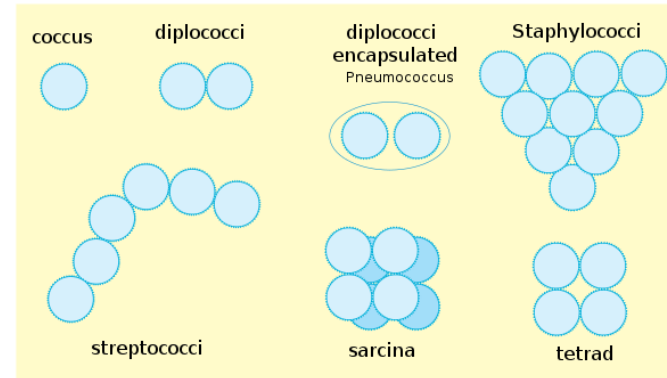
Systematic arrangement of organisms
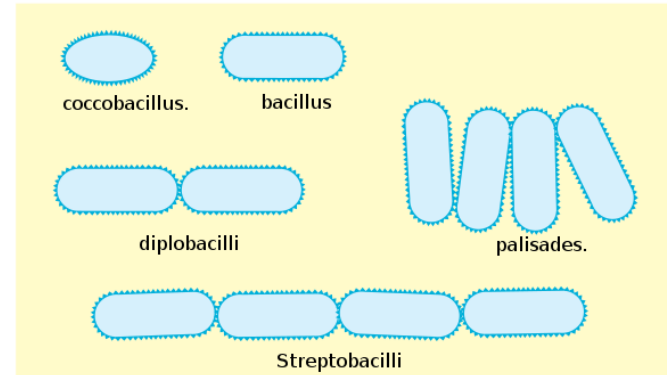
The great diversity of life

Plants
Rkitko, 2009

Animals
Medeis, 2012

Classification of microbes

- Traditionally relying heavily on phenotypic information
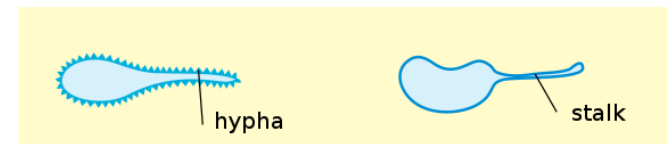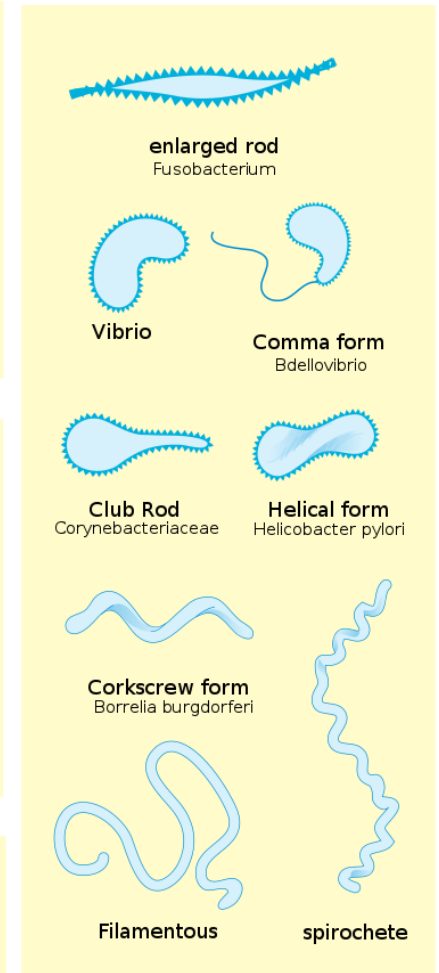

Basic morphological differences between bacteria
Mariana Ruiz LadyofHats, 2006, Wikimedia Commons

Classification of microbes

- Traditionally relying heavily on phenotypic information

A Gram stain of mixed *Staphylococcus aureus* ATCC 25923 (Gram-positive cocci, in purple) and *Escherichia coli* ATCC 11775, (gram-negative bacilli, in pink)

Y tambe, 2010, Wikimedia Commons

10µm
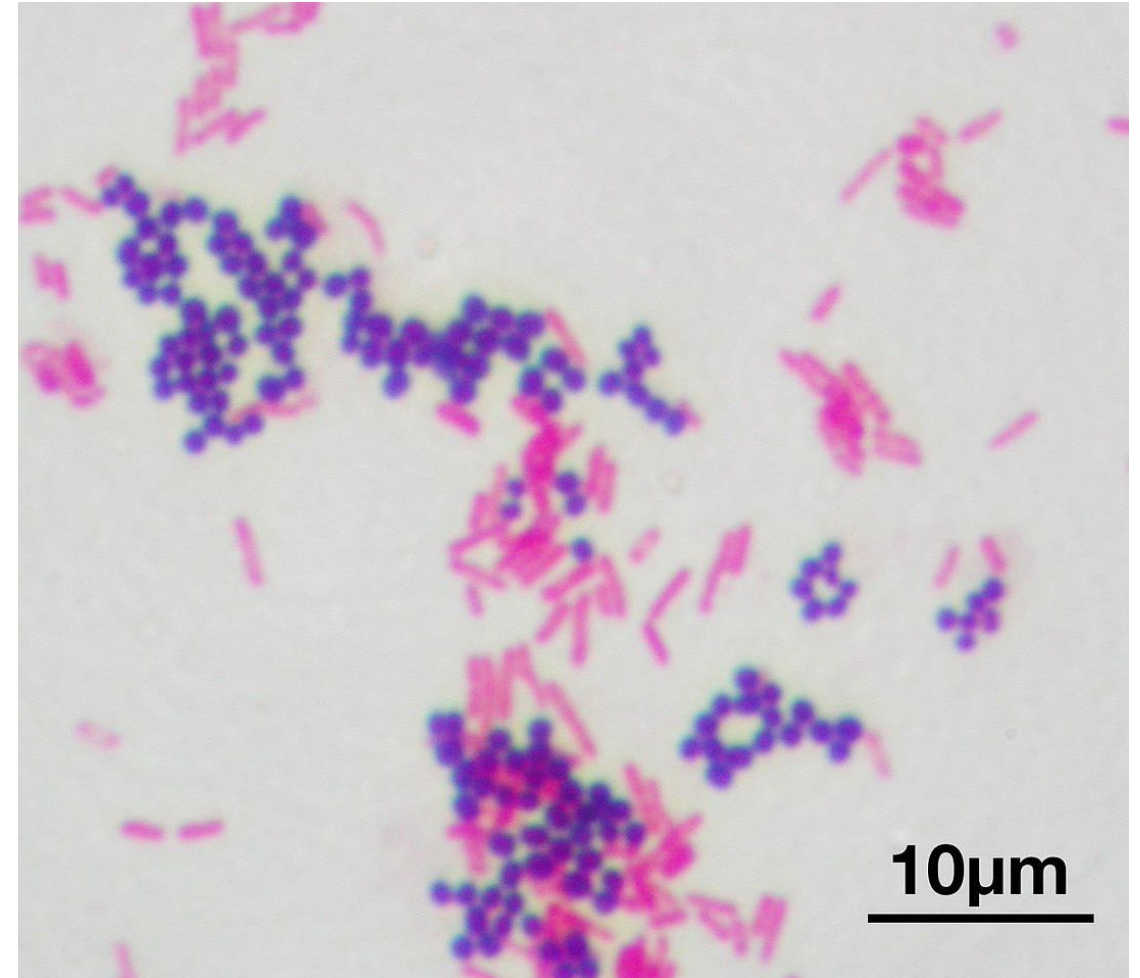
Classification of microbes

- Traditionally relying heavily on phenotypic information

American Society for Microbiology
June 5, 2015 · 🌐

Pic of the Day: Hand print

Hand print on a large TSA plate from my 8 1/2 year old son after playing outside.

This photo was taken from our site www.microbeworld.org. Submitted by: Sturm Thanks to the author(s): Tasha Sturm, Cabrillo College

Submit your pics to #ASM via www.microbeworld.org #potd #science #microbiology #laboratory #microbeworld #handprint #tsa #microbes

Classification of microbes

- Phenotypic features are clearly not quite enough
  - Frankly, there are simply just not enough phenotypic features to distinguish all bacteria on this planet

- It is becoming more popular to classify microbes (and organisms in general) based on their shared evolutionary history = **systematics**

Systematics

- **Biological systematics** is the study of the diversification of living forms, both past and present, and the relationships among living things through time

- Derived from Latin word *systema*, which means **systematic arrangement of organisms**

https://en.wikipedia.org/wiki/Systematics

Systematics

- The field that
  - provides scientific names for organisms
  - describes them
  - preserves collections of them
  - provides **classifications** for the organisms, **keys for their identification**, and data on their distributions

  - investigates their **evolutionary histories**
  - considers their environmental adaptations

Evolutionary taxonomy

Taxonomy

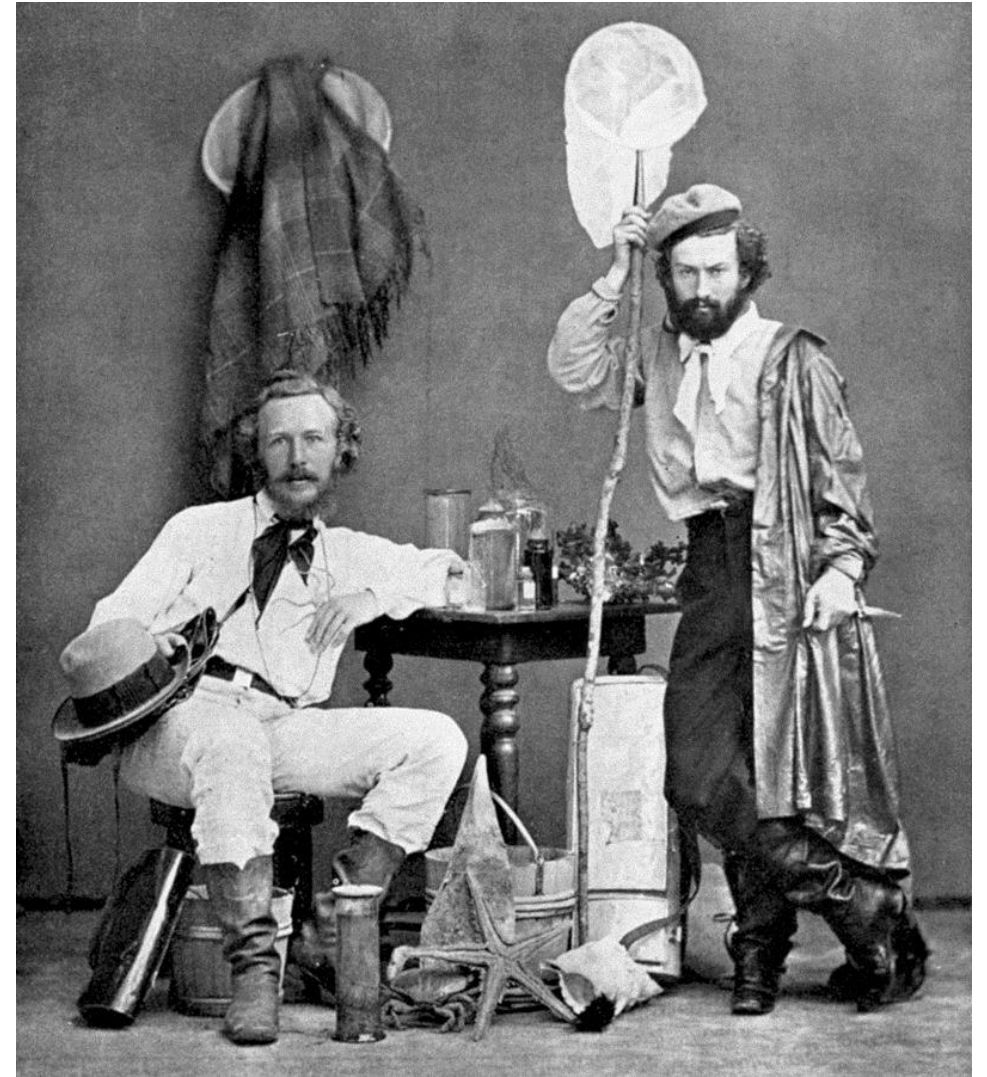Evolutionary study
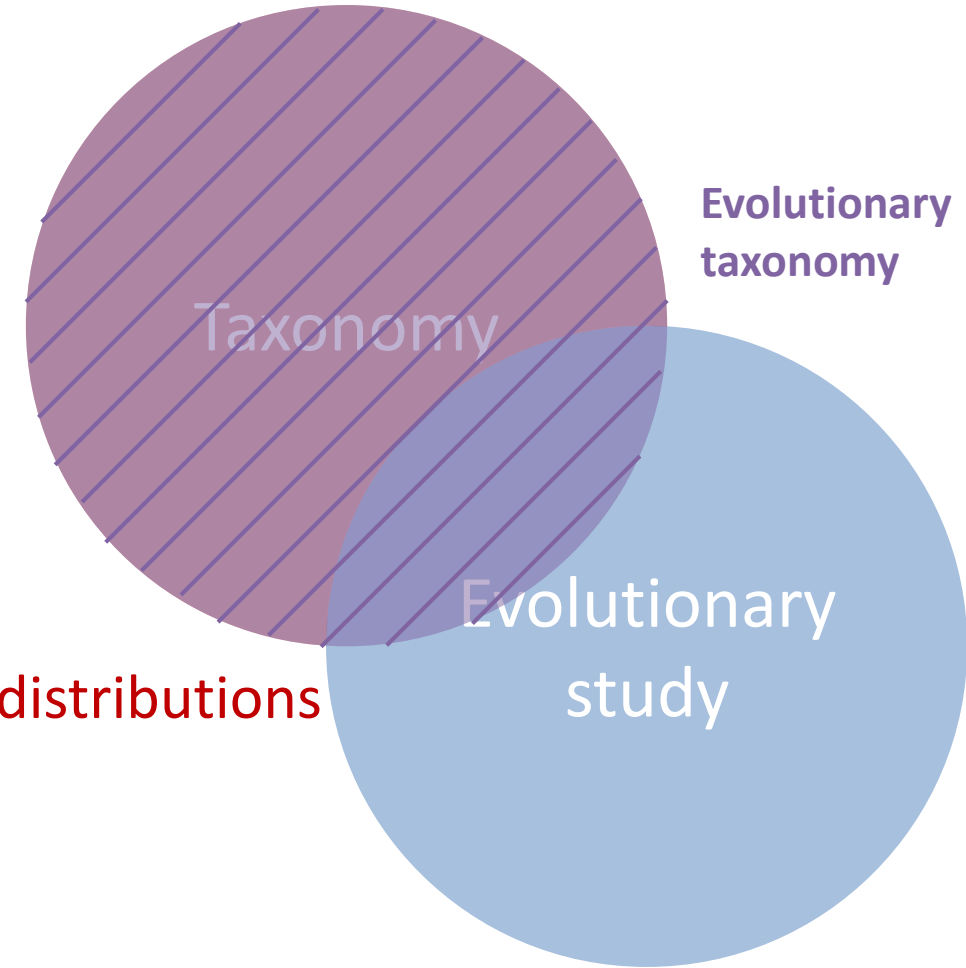
Michener et al., 1970

What is phylogenetics?

The study of the evolutionary history of a group of organisms

Phylogenetics

Phylogenetics (Haeckel, 1866) is the study of the evolutionary history and relationships among individuals or groups of organisms
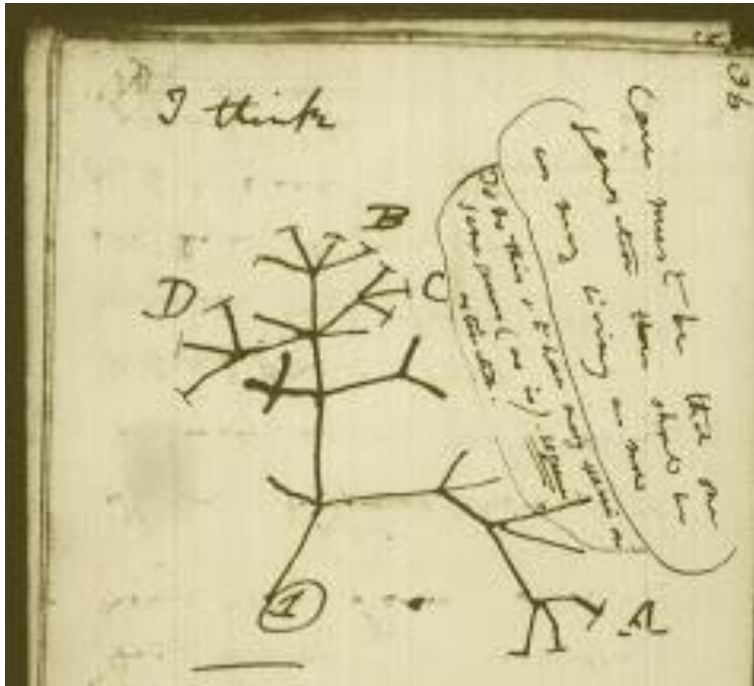
phylé, phylon   = *tribe, clan, race*
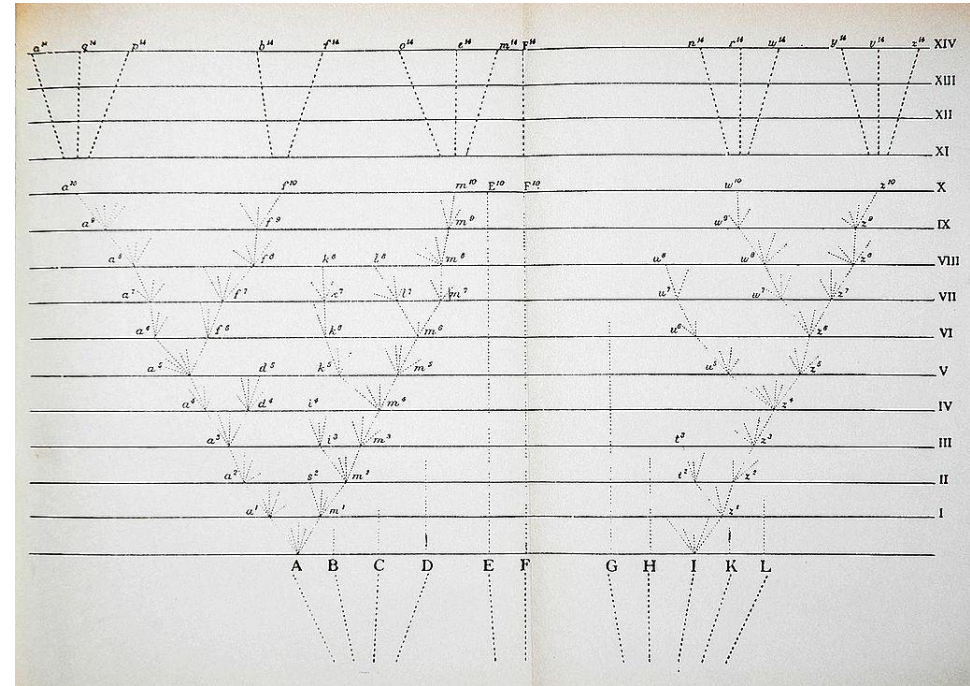*genetikós*   = *origin, source, birth*

Ernst Haeckel with Nicholai Miklukho-Maklai, his assistant, in the Canaries, 1866

Phylogenetic tree

- **Phylogenetic tree** – or **phylogeny** – is a diagrammatic hypothesis representing relationship of a set of evolving entities
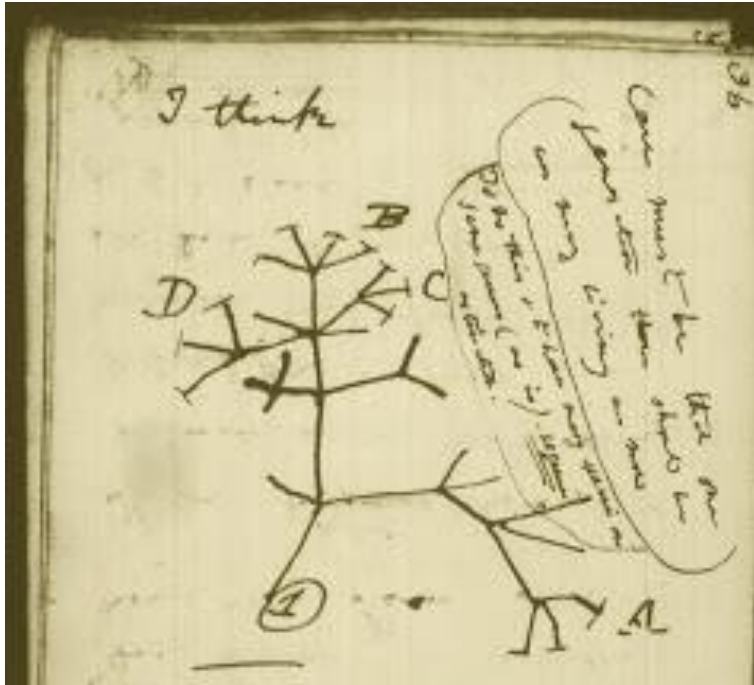

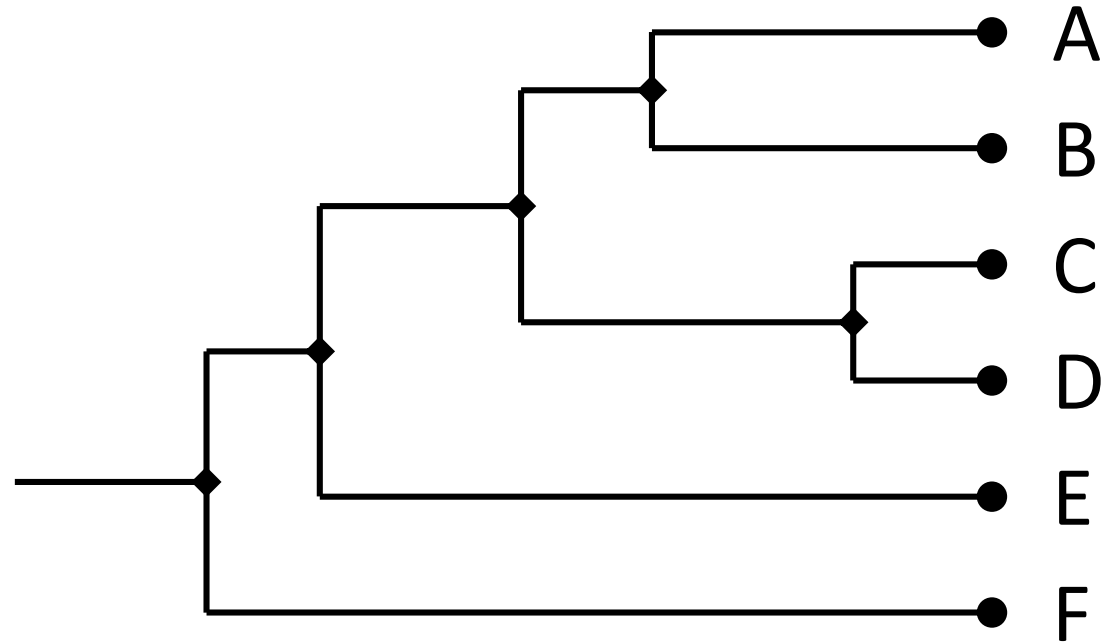Darwin, 1837-1838, Notebook B, on "Transmutation of species", p 36

Darwin, 1859, the Origin of Species
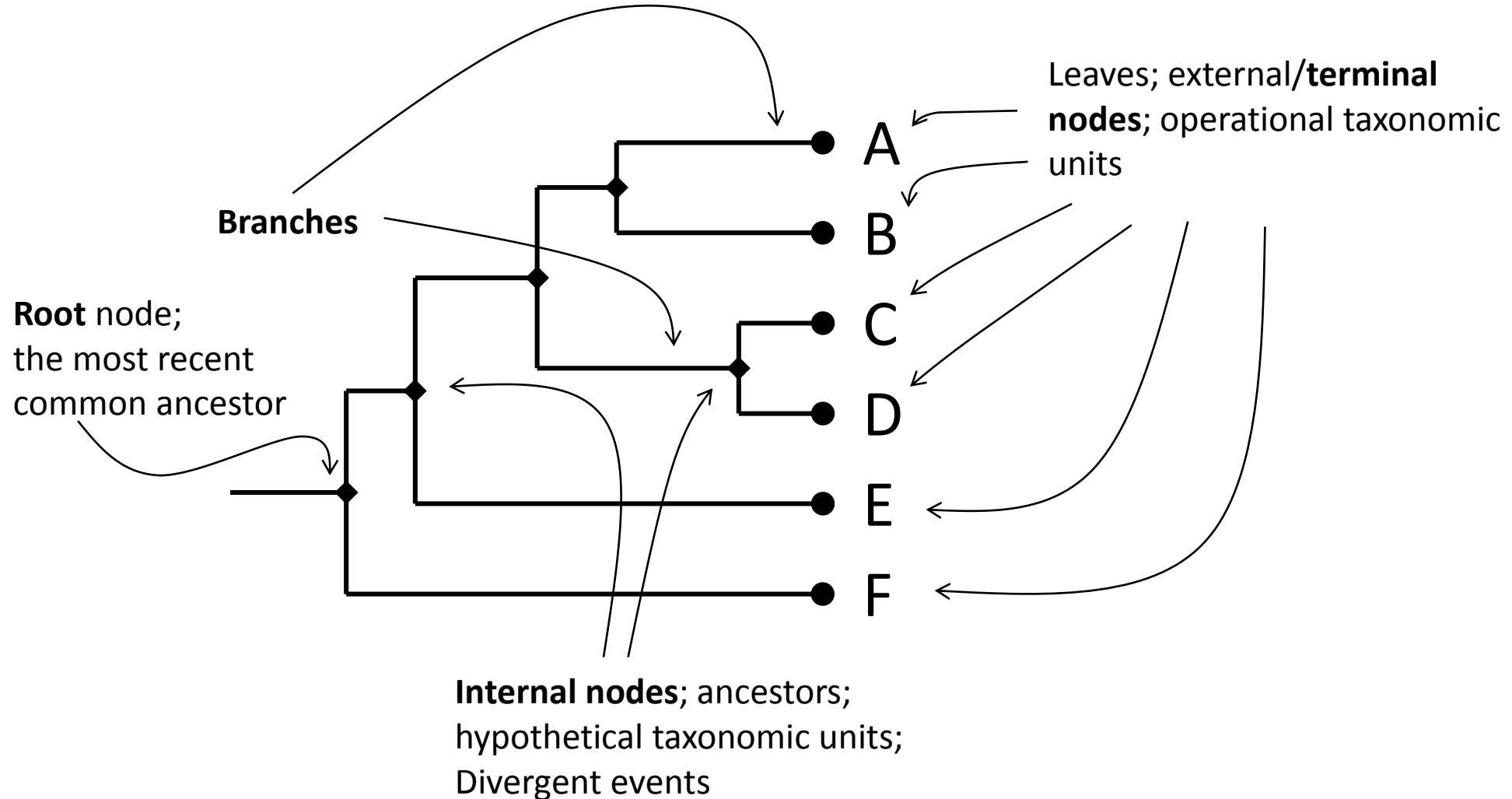
Phylogenetic tree

- **Phylogenetic tree** – or **phylogeny** – is a diagrammatic hypothesis representing relationship of a set of evolving entities
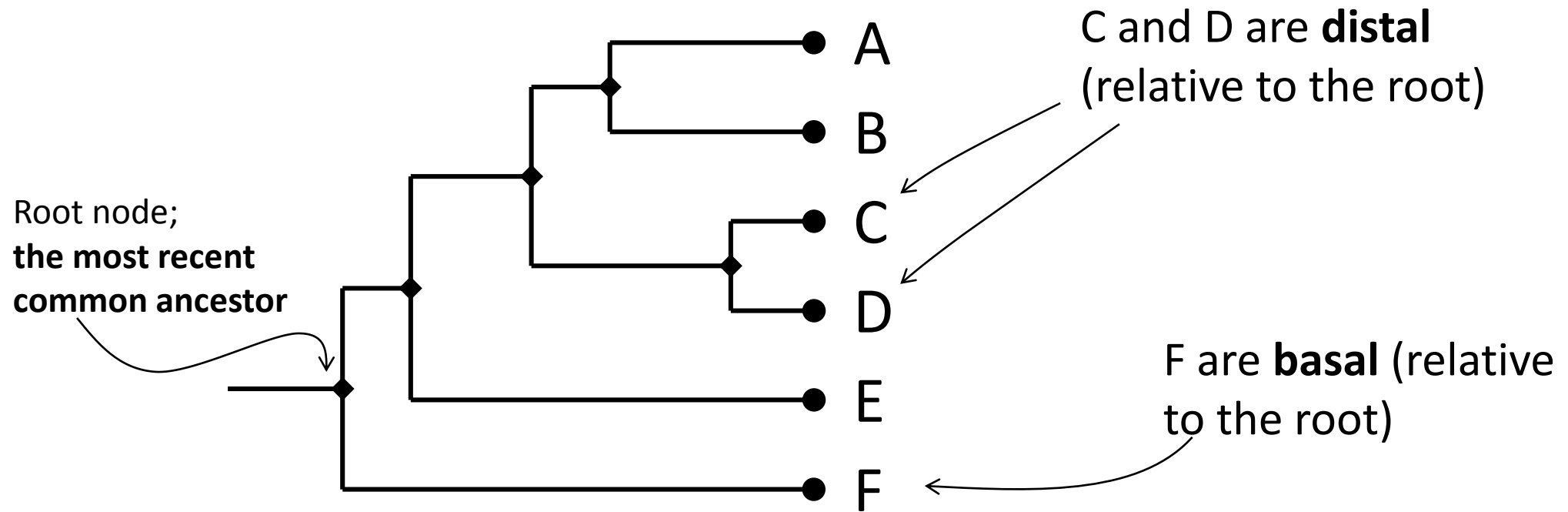

Darwin, 1837-1838, Notebook B, on
"Transmutation of species", p 36
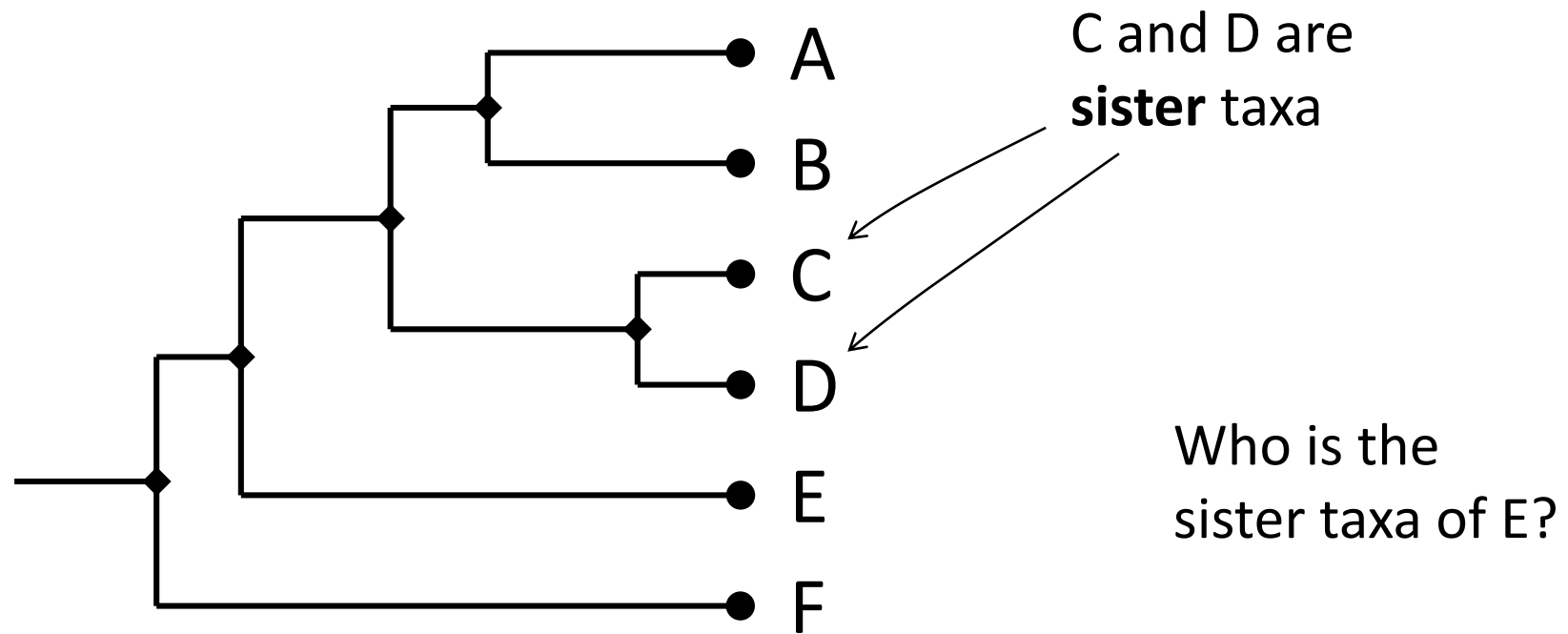
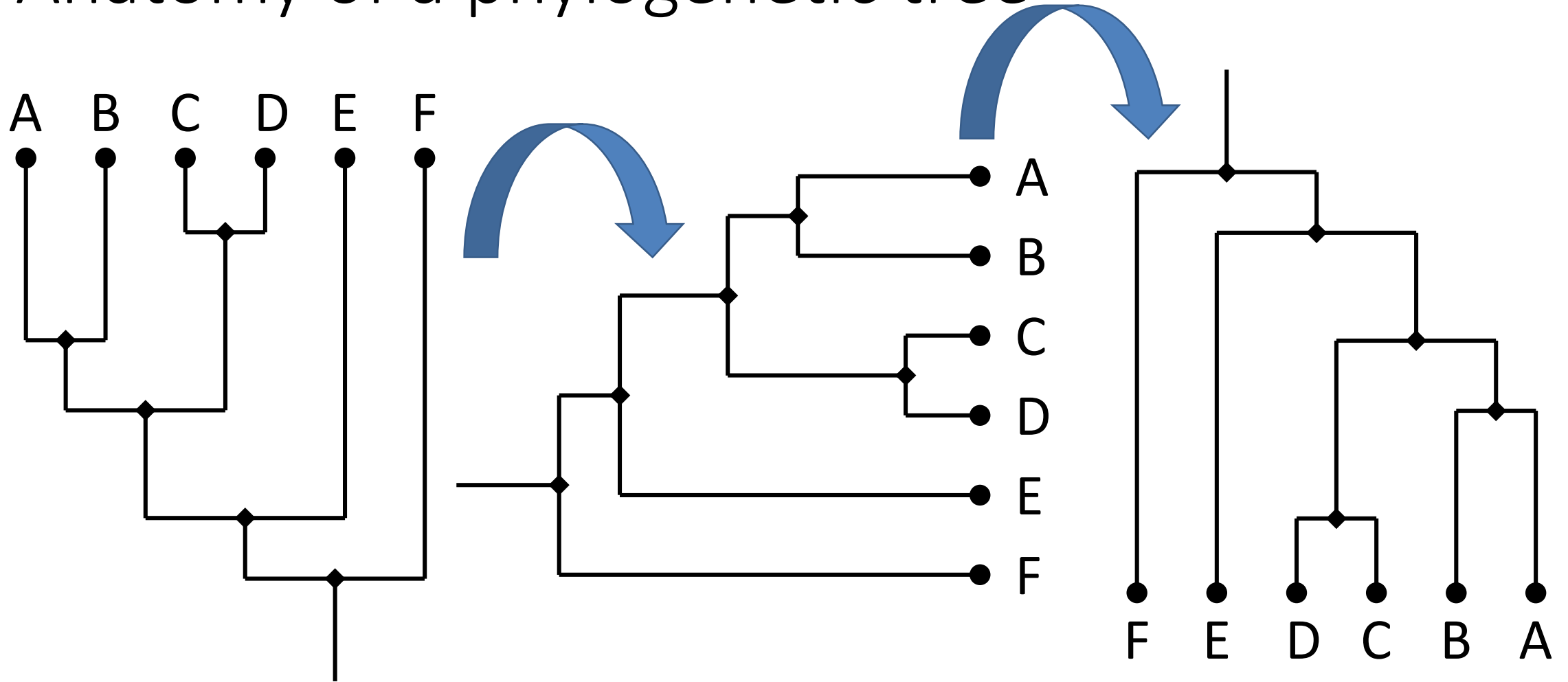# Anatomy of a phylogenetic tree

Leaves; external/**terminal nodes**; operational taxonomic units

Branches

Root node;
the most recent common ancestor

Internal nodes; ancestors;
hypothetical taxonomic units;
Divergent events

A

B

C

D

E

F

Anatomy of a phylogenetic tree

C and D are **distal** (relative to the root)

Root node; **the most recent common ancestor**

F are **basal** (relative to the root)

A
B
C
D
E
F

Which taxon is most like the most recent common ancestor?

Anatomy of a phylogenetic tree

C and D are
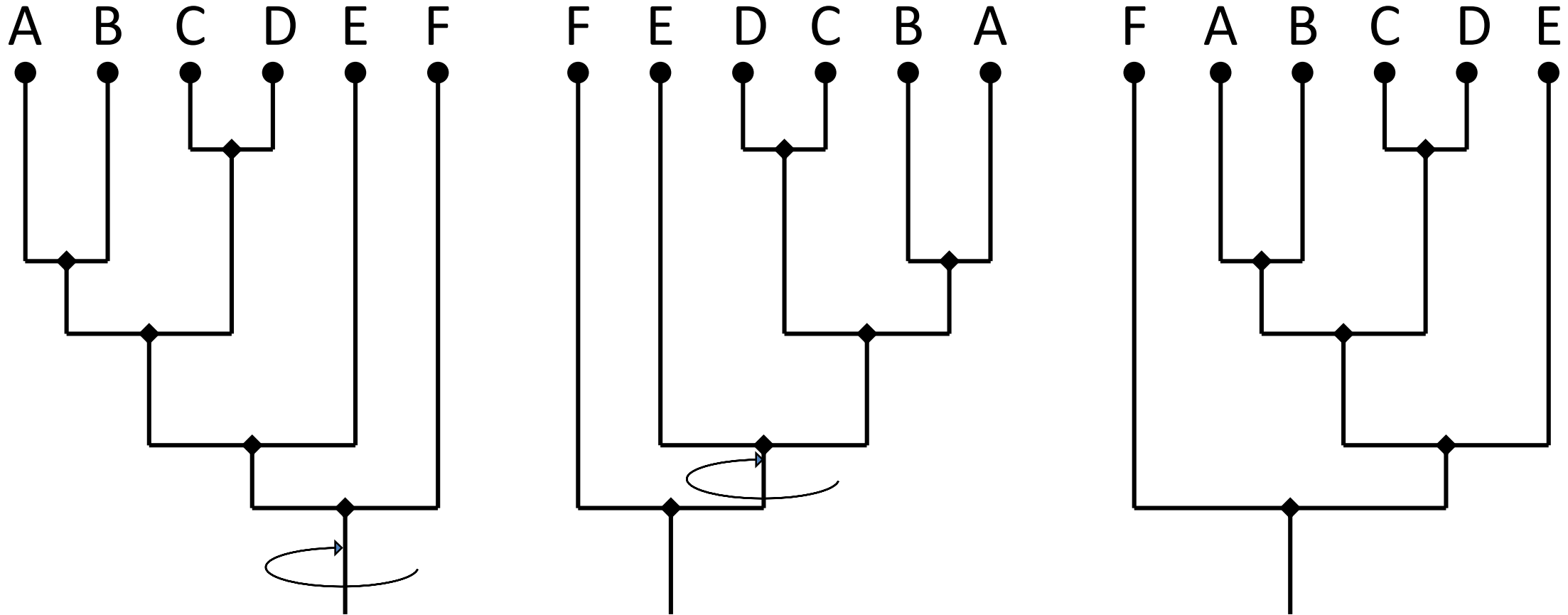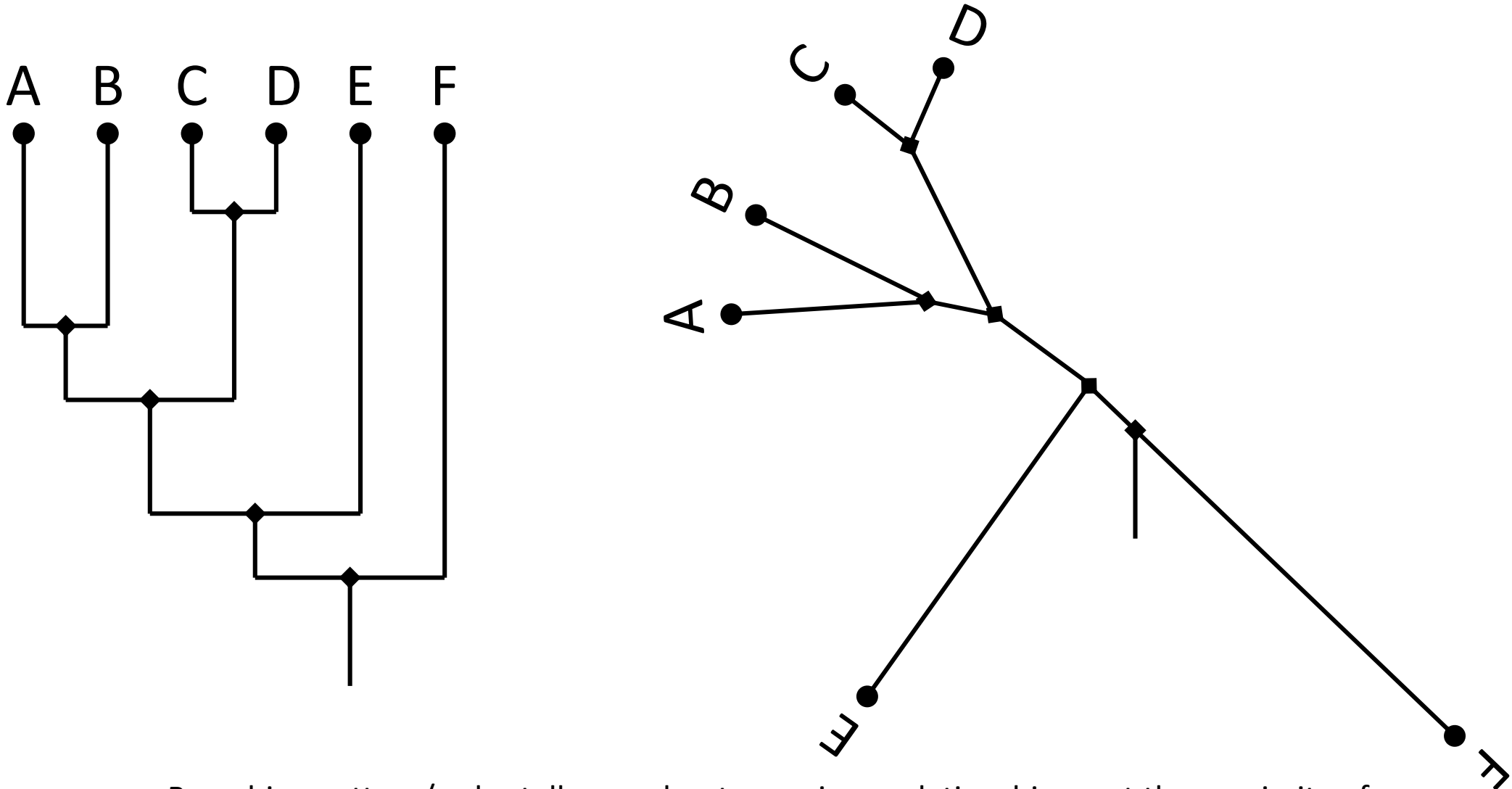**sister** taxa

Who is the
sister taxa of E?

Branching pattern/order tells you about organisms relationships, not the proximity of names

Anatomy of a phylogenetic tree

Branching pattern/order tells you about organisms relationships, not the proximity of names

Anatomy of a phylogenetic tree

Branching pattern/order tells you about organisms relationships, not the proximity of names

Anatomy of a phylogenetic tree

Branching pattern/order tells you about organisms relationships, not the proximity of names
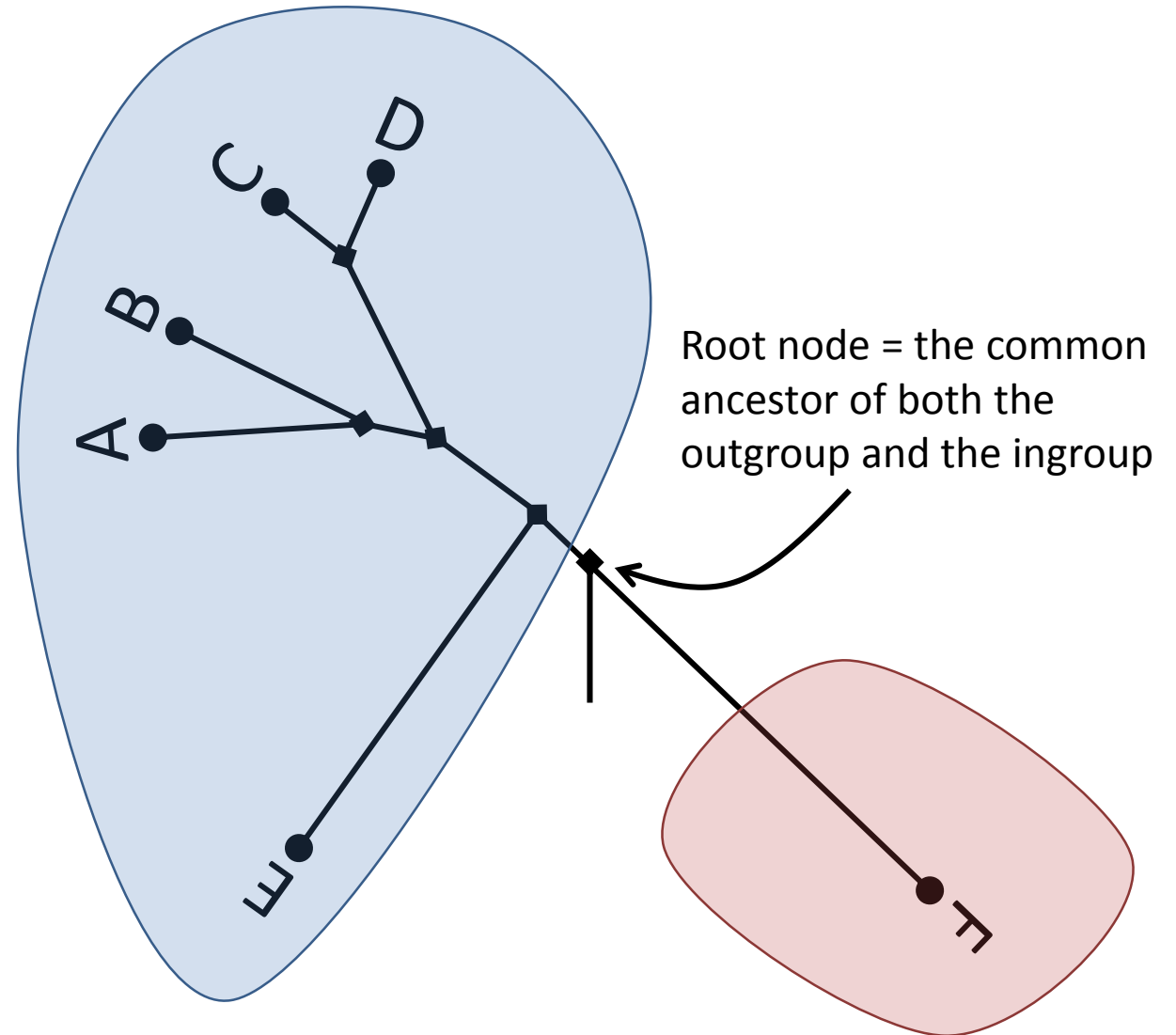
Anatomy of a phylogenetic tree

- An **unrooted tree**

- Only positions the individual taxa relative to each other without indicating the direction of evolutionary processes

- Neither does it tell you which node represents ancestors
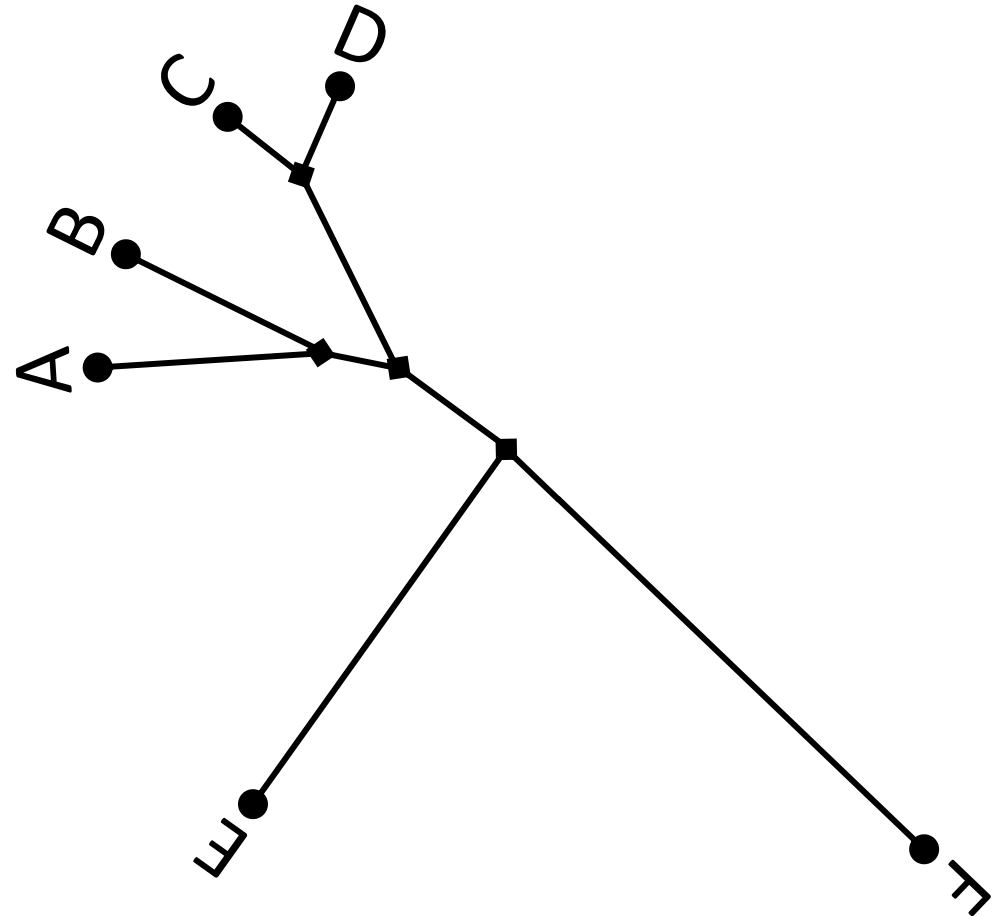
Anatomy of a phylogenetic tree

- A rooted tree

- The tree can be rooted if you have "**an outgroup**"

- The remainder then forms "**an ingroup**".

Root node = the common ancestor of both the outgroup and the ingroup

Anatomy of a phylogenetic tree

- A rooted tree

- The tree can be rooted if you have "**an outgroup**"

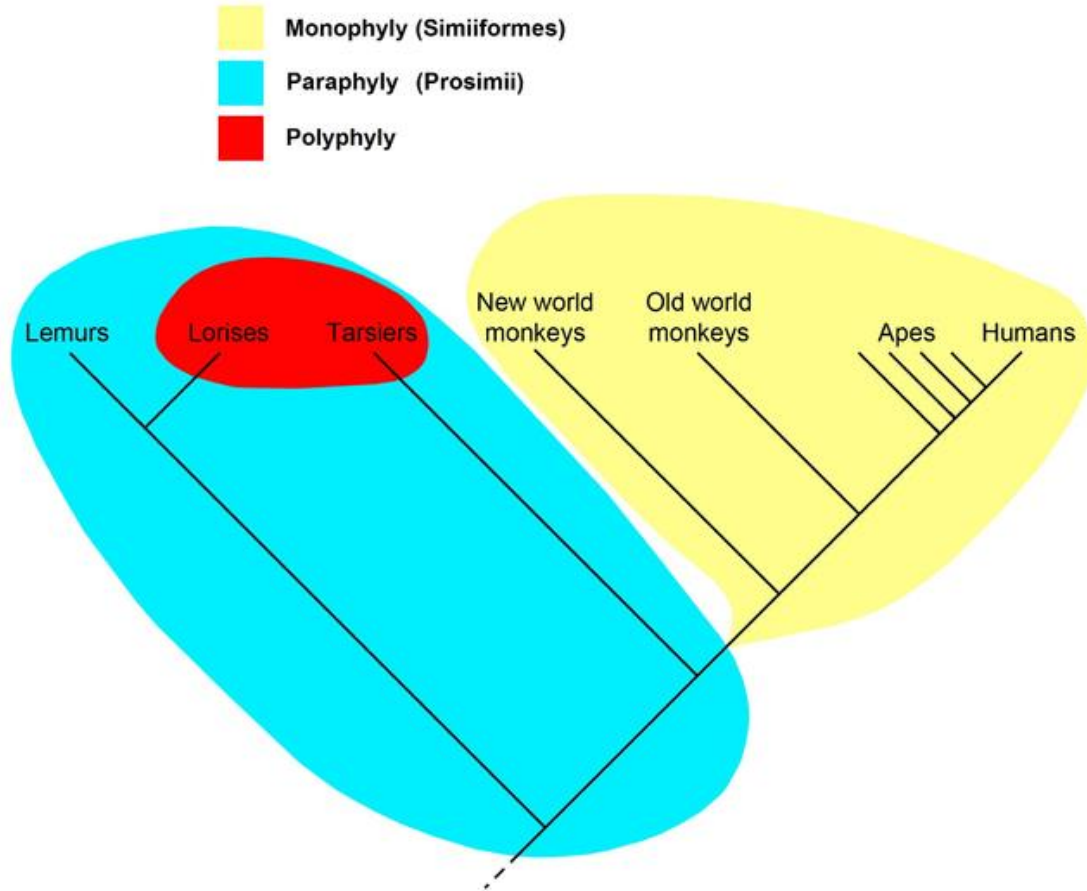- The use of several outgroups can gives you more confidence in rooting
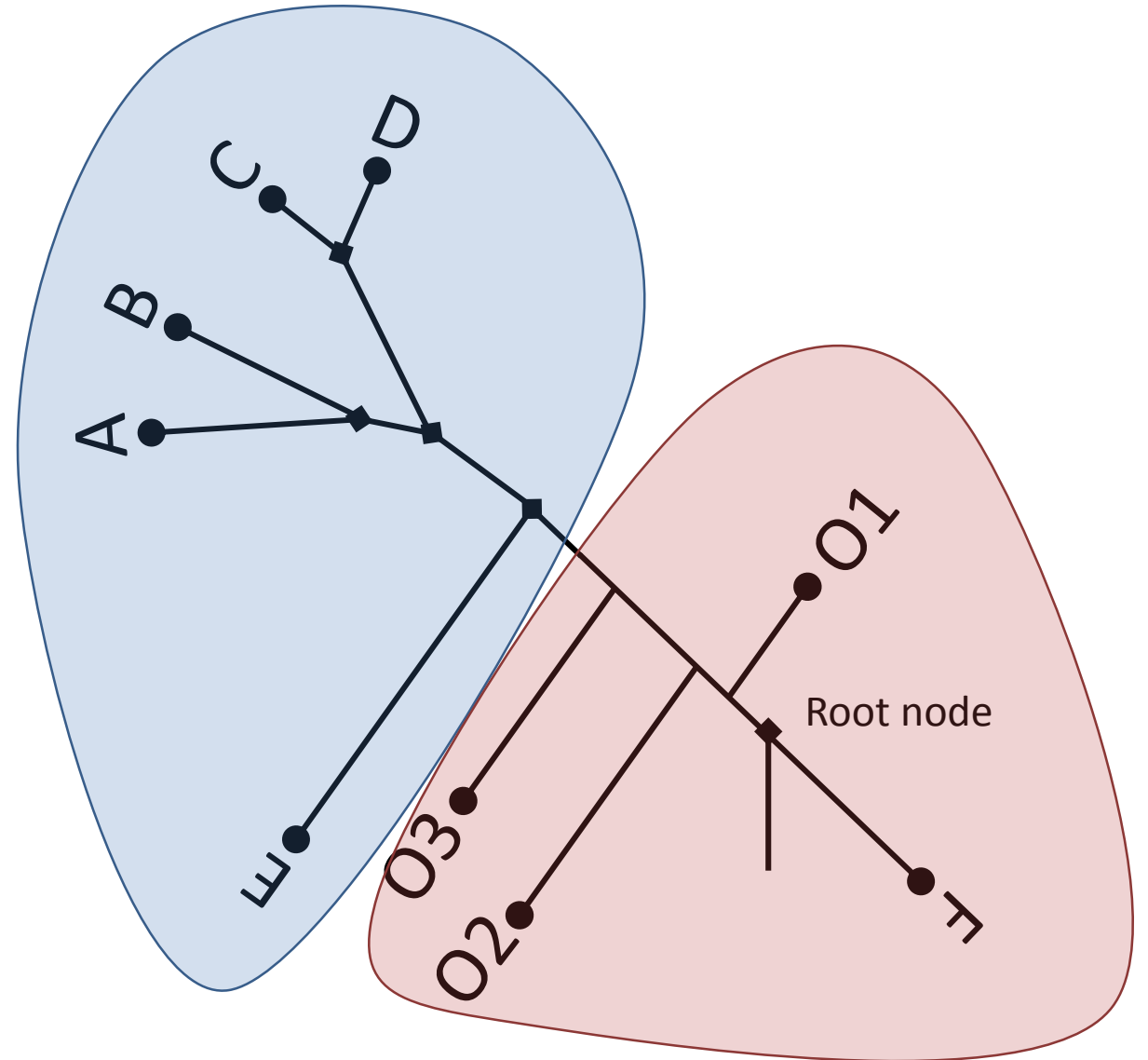
Questions in phylogenetics

- *What are the relationships among a set of organisms?*

- *How did the organisms look like the past, and when did they acquire their traits?*

- *What are the forces under which our organisms are evolving?*

- ***How many groups of organisms do we have here?***

Grouping organisms based on their shared histories

Grouping organisms based on their shared histories

Monophyly (Simiiformes)

Paraphyly (Prosimii)

Polyphyly
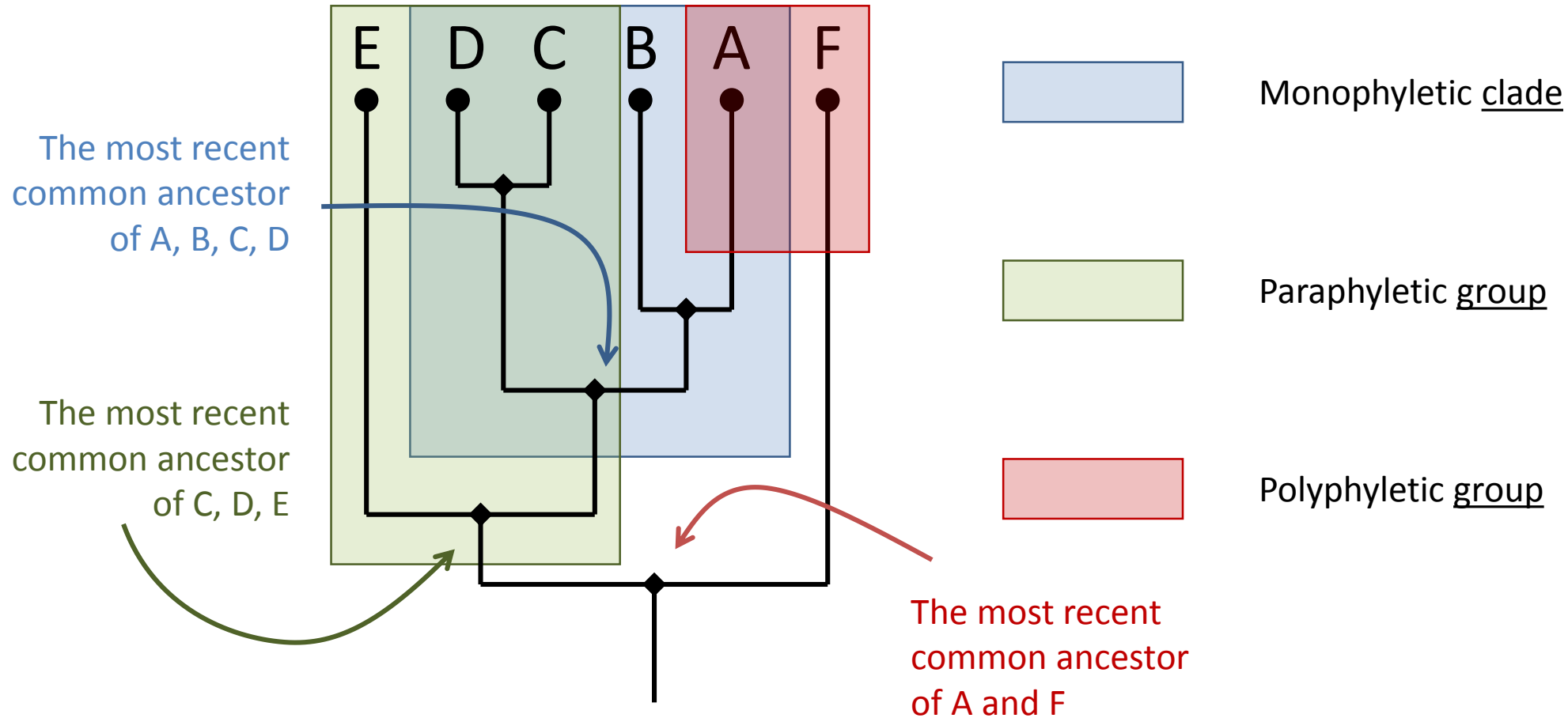
Lemurs Lorises Tarsiers New world monkeys Old world monkeys Apes Humans

A tree of the primates, showing:

a monophyletic taxon: the simians (yellow)

a paraphyletic group: the prosimians (cyan)

a polyphyletic group: the night-active primates (red)

How to reconstruct a tree?

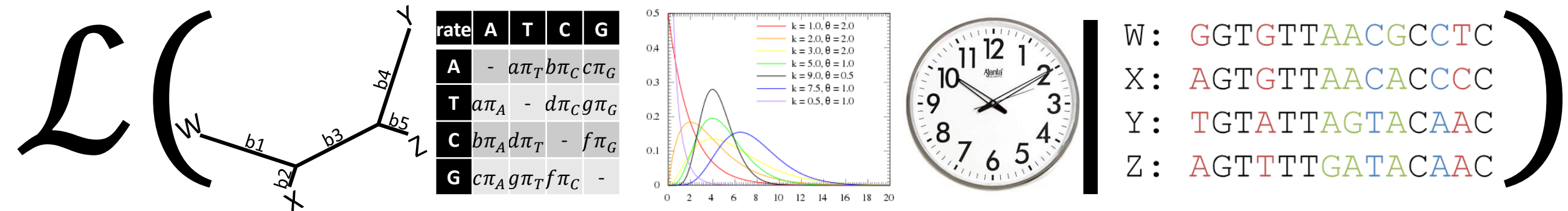Phylogenetic estimation under the maximum likelihood framework

ML phylogenetic estimation

- The fundamental principle of the ML phylogenetic estimation is …

"

*to find the evolutionary hypothesis / scenario that has the highest **likelihood (ℒ)** that can best account the conversion of one sequence into another*

"

ML phylogenetic estimation

- A **maximum likelihood tree** (which is just one of many components of an evolution hypothesis) is a tree that most likely produces the observed data under a specific model of evolution

The Idiot's Guide to the Zen of Likelihood in a Nutshell in Seven Days for Dummies, Unleashed

A gentle introduction, for those of us who are small of brain, to the calculation of the likelihood of molecular sequences

Peter G. Foster*

July 28, 2001

Evaluate tree uncertainty – "Pull Yourself Up By Your Bootstraps"

- ML framework inherently gives you just one BEST tree…

- "Bootstrapping" is a common technique used to assign variability to sample estimates, relying on random sampling with *replacement*

Evaluate the tree uncertainty

Tree distribution

ML tree
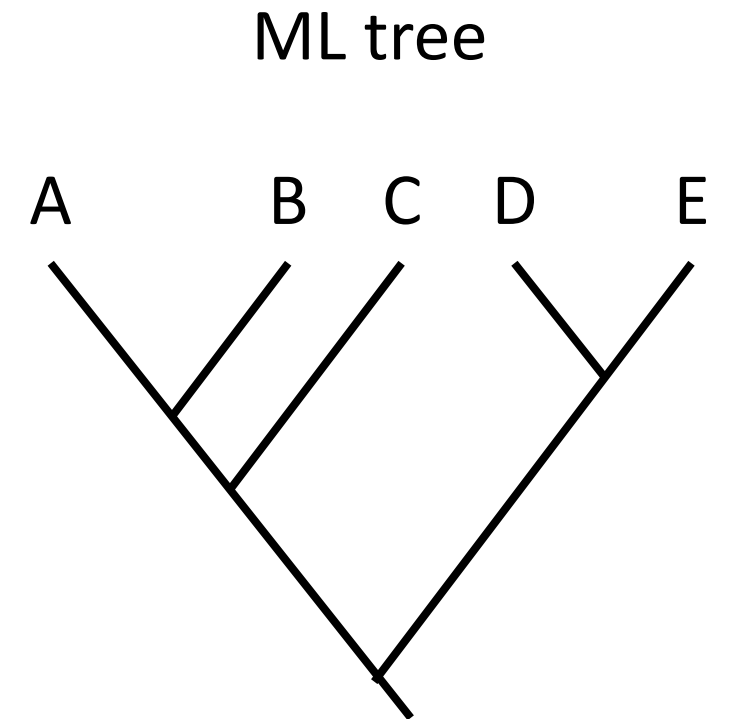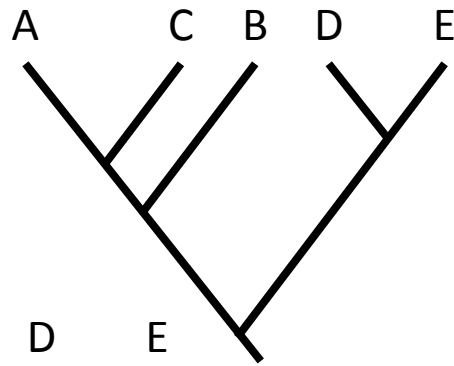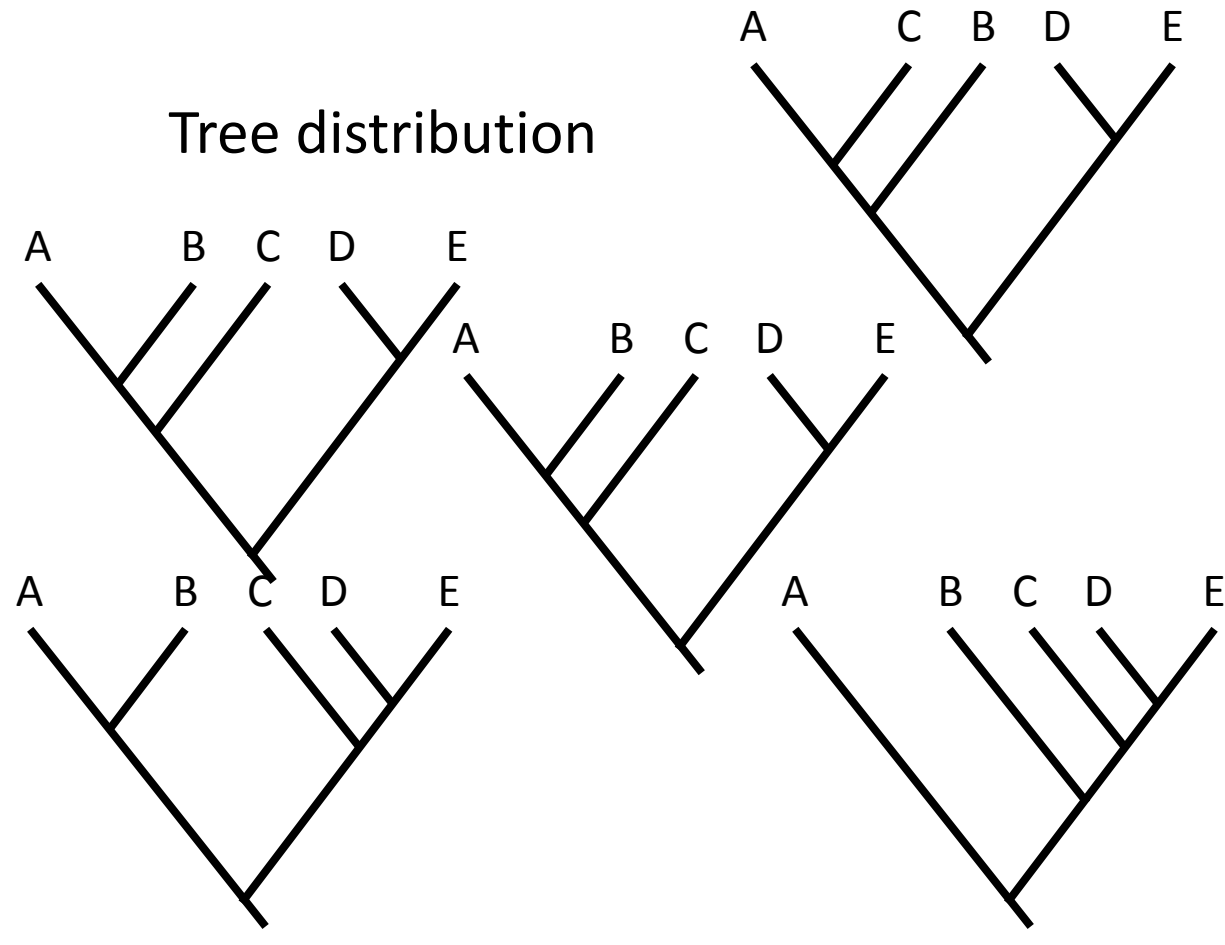
# Evaluate tree uncertainty – "Pull Yourself Up By Your Bootstraps"
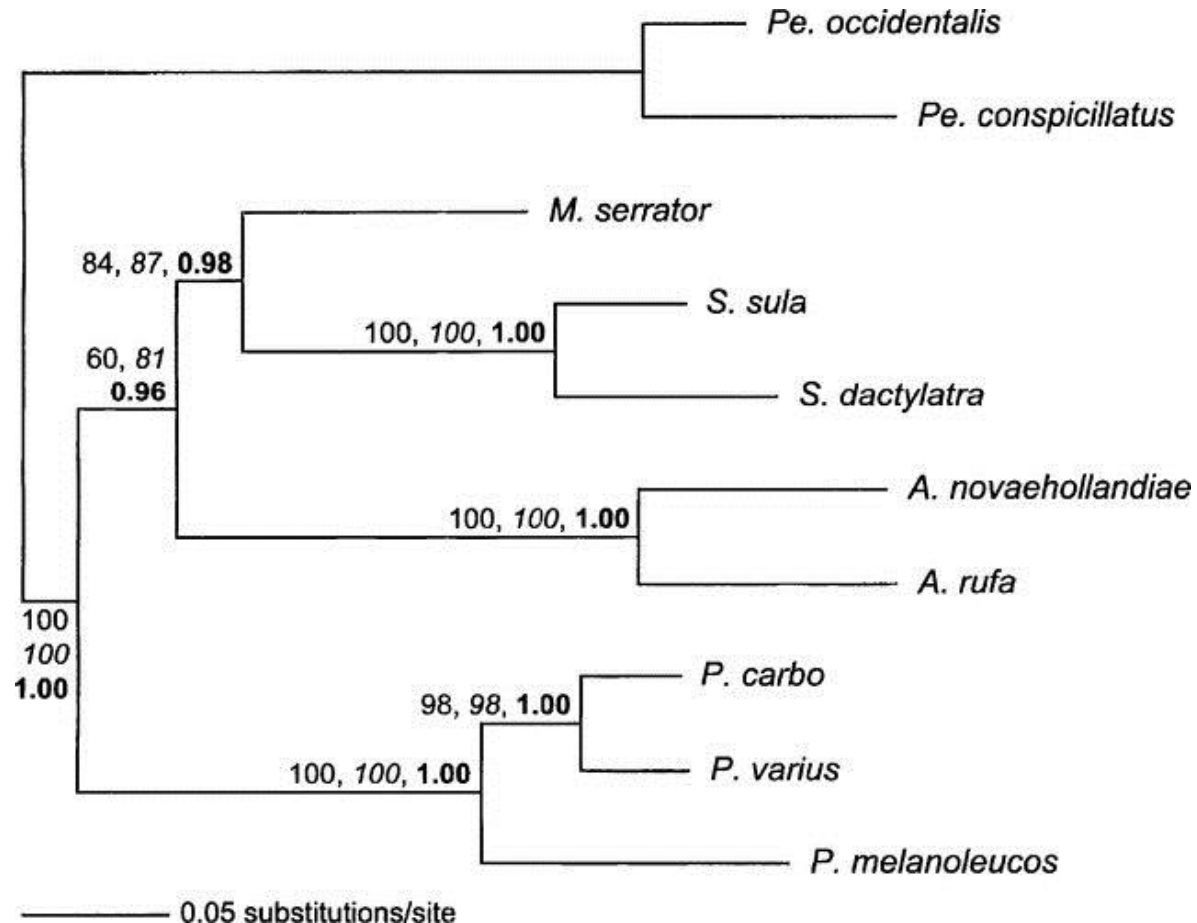
The maximum-likelihood tree (the branch lengths represent the expected number of substitution per site) with A. anhinga excluded. The percentage of bootstrap replicates (out of 1000) that supported each node (for parsimony and, in italics, maximum likelihood) and the Bayesian posterior probabilities (in bold) are shown.

Kennedy et al., 2005 Syst Biol.

> 75% is typically considered good enough

ML phylogenetic estimation

- `IQ-TREE2` will do everything for you!

http://www.iqtree.org/
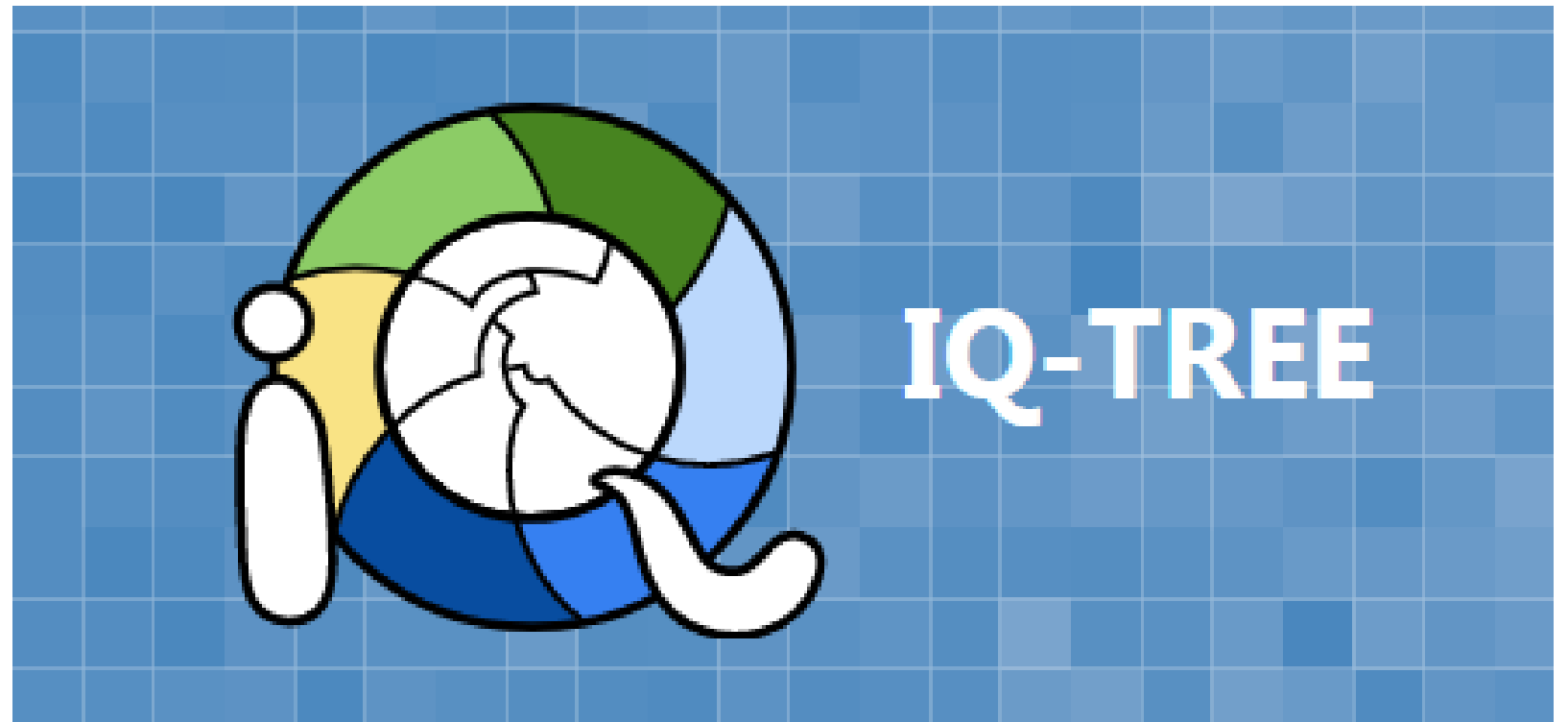
To visualise the tree…

- `Figtree` is a decent program you can use to visualise the tree