Phylogenetics (EEB 5349) Spring 2026

A phylogeny in the wild

plewis.github.io/phylogenetics2026/

After this course...

- You should be able to read and understand current **empirical papers** that estimate or make use of a phylogeny
- You should know how to determine the **appropriate model** to use in analyzing your own data
- You should be able to carry out such analyses using modern **phylogenetic software** tools
- You should know enough phylogenetic theory to **correctly interpret** the results

Reading

- **Textbook**:
  - You will receive chapters from a book I am writing. This will be a free textbook when it is done, following Luke Harmon's example (https://lukejharmon.github.io/pcm/)
- **Recommended books**:
  - I've put a list of books on phylogenetics on the course web site. None of these are required, but all are useful for understanding some aspect of phylogenetics

Homeworks (10):

50%

Lab (13):

25%

Lecture (30):

15%

1-on-1 meetings (4):

10%

Grading scale:

$93\% \leq A \leq 100\%$

$90\% \leq A- < 93\%$

$87\% \leq B+ < 90\%$

$83\% \leq B < 87\%$

$80\% \leq B- < 83\%$

$77\% \leq C+ < 80\%$

$73\% \leq C < 77\%$

$70\% \leq C- < 73\%$

etc.

See the grading section of the course web site for details

Today

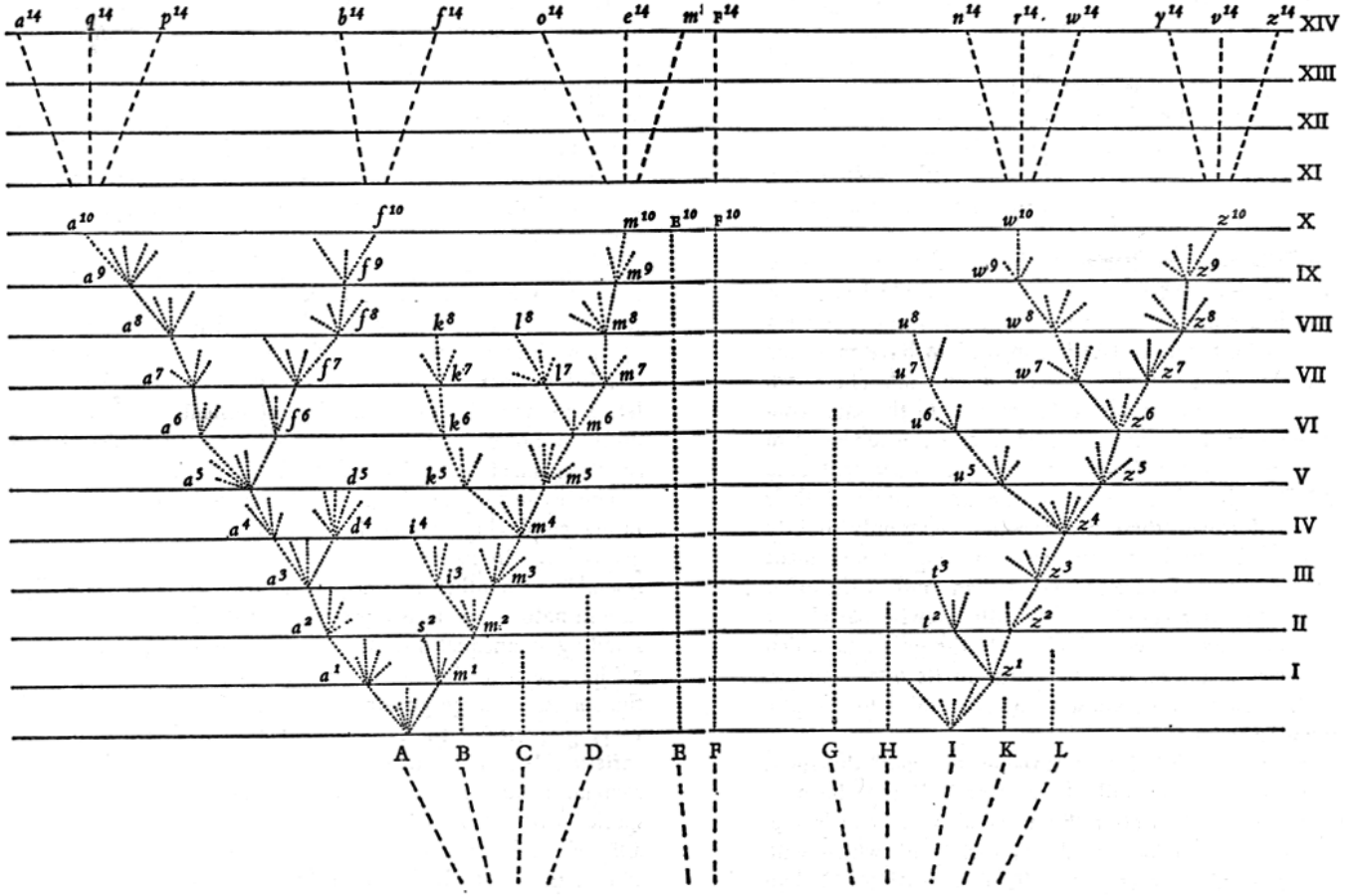
- Of what use is a phylogeny ("tree")?
- What are the parts of a tree?
- How are trees stored and communicated?
- How do you tell if one tree is better than another?

Phylogenetics

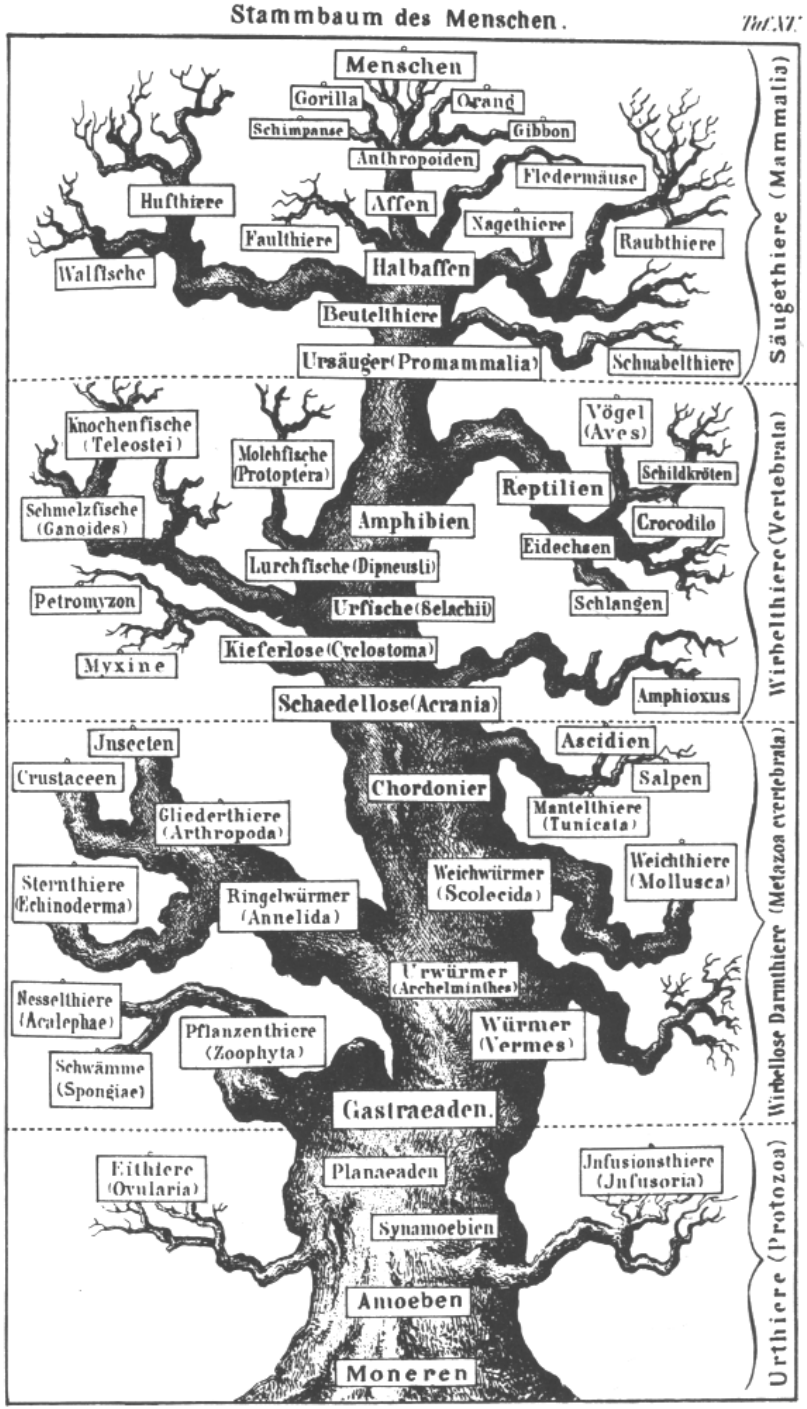
"The affinities of all the beings of the same class have sometimes been represented by a great tree. I believe this simile largely speaks the truth."

- Charles Darwin, O*rigin of Species, 1859*

The only figure in Darwin's 1859 "Origin of Species" was a phylogeny.



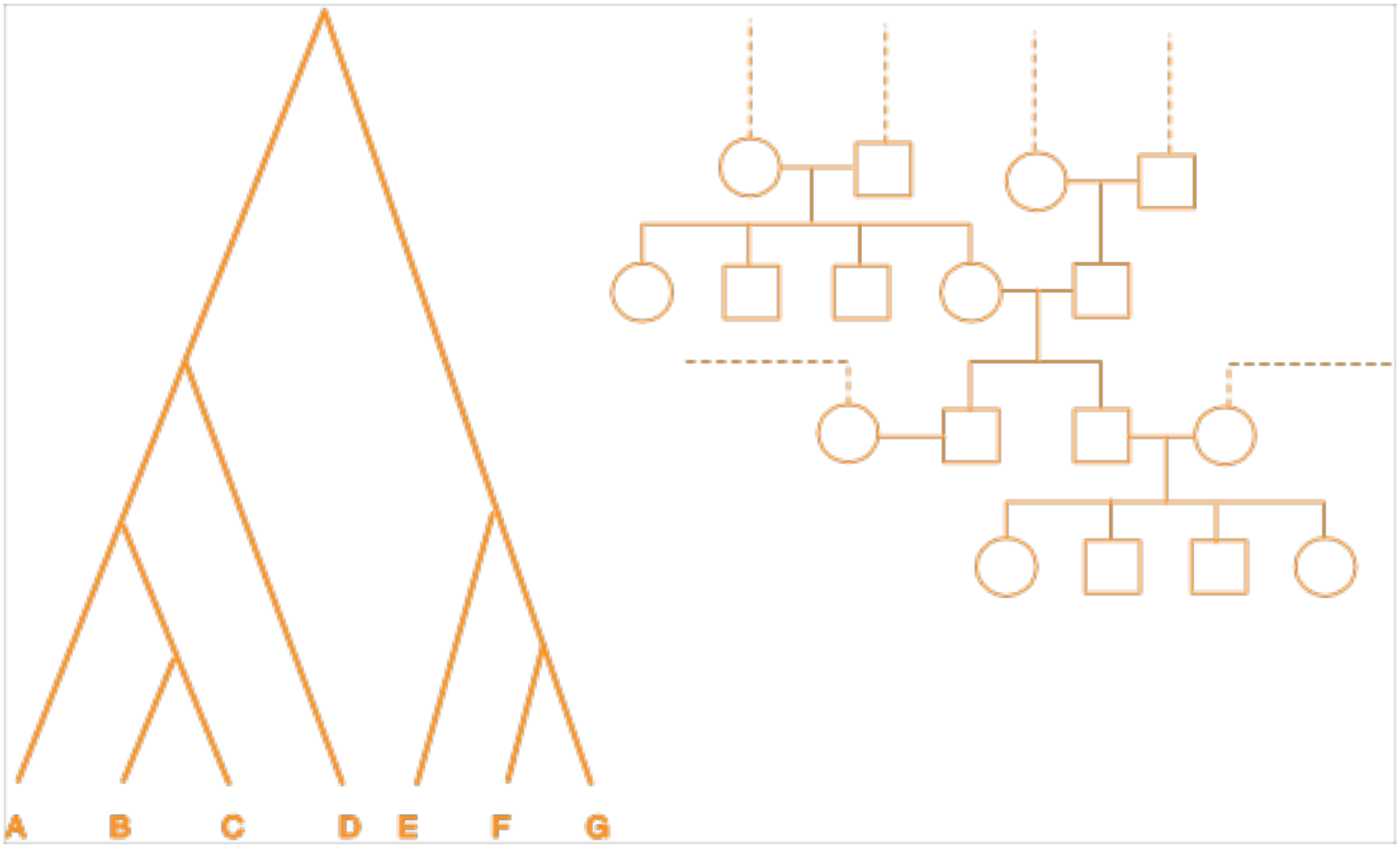
Darwin, C. R. 1859. Origin of species by means of natural selection (or the preservation of favoured races in the struggle for life). Originally published by John Murray. This figure from pp. 160-161 in Penguin Classics edition published 1985 by Penguin Books, London.



Haekel's famous 1874 tree showing hypothesized human ancestors.

Fig. 20, p. 171, in Gould, S. J. 1977. Ontogeny and phylogeny. Harvard University Press, Cambridge, Massachusetts.

Phylogenies: what are they?



Of what use is a phylogeny ("tree")?

Library of Congress

A -- GENERAL WORKS
B -- PHILOSOPHY. PSYCHOLOGY. RELIGION
C -- AUXILIARY SCIENCES OF HISTORY
D -- HISTORY (GENERAL) AND HISTORY OF EUROPE
E -- HISTORY: AMERICA
F -- HISTORY: AMERICA
G -- GEOGRAPHY. ANTHROPOLOGY. RECREATION
H -- SOCIAL SCIENCES
J -- POLITICAL SCIENCE
K -- LAW
L -- EDUCATION
M -- MUSIC AND BOOKS ON MUSIC
N -- FINE ARTS
P -- LANGUAGE AND LITERATURE
Q -- SCIENCE
R -- MEDICINE
S -- AGRICULTURE
T -- TECHNOLOGY
U -- MILITARY SCIENCE
V -- NAVAL SCIENCE
Z -- BIBLIOGRAPHY. LIBRARY SCIENCE. INFORMATION RESOURCES (GENERAL)

Suppose you look up a book:
How to identify flowering plant families, by J. P. Baumgardt
QK 495.A1 B38 1982

http://www.loc.gov/catdir/cpso/lcco/lcco.html

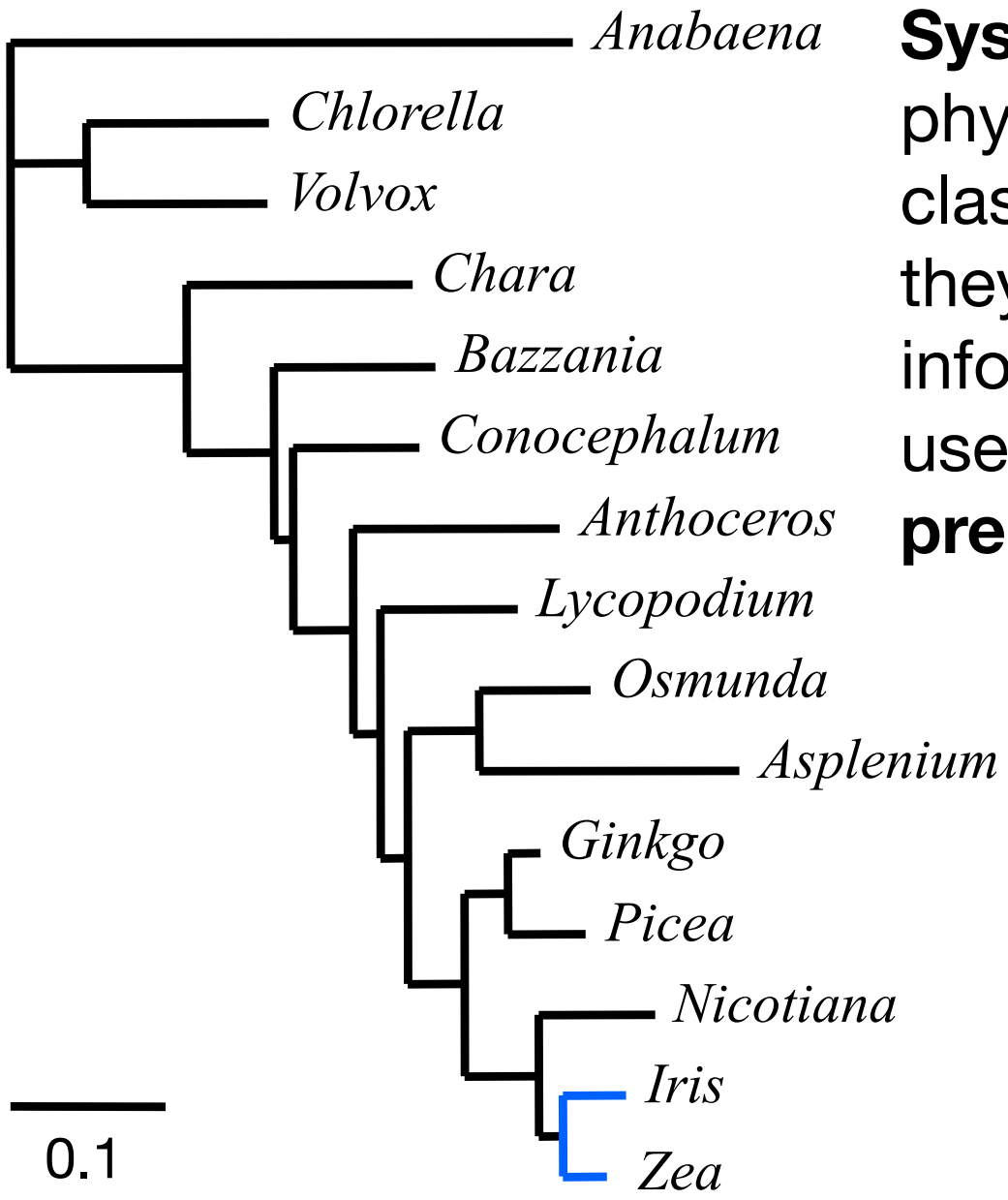
Library of Congress classification

Q	Science (General)
QA	Mathematics
QB	Astronomy
QC	Physics
QD	Chemistry
QE	Geology
QH	Natural history -Biology
QK	Botany
QL	Zoology
QM	Human anatomy
QP	Physiology
QR	Microbiology

QK1-989	Botany
QK1-474.5	General
QK474.8-495	Spermatophyta
QK494-494.5	Gymnosperms
QK495	Angiosperms
QK504-(638)	Cryptogams
QK640-(707)	Plant anatomy
QK710-899	Plant physiology
QK900-989	Plant ecology

Classification allows prediction

- Other books on the same shelf:
  - Identification of flowering plant families, by P. H. Davis
  - Key to the families of flowering plants of the world, by J. Hutchinson
  - Evolution and classification of flowering plants, by A. Cronquist
  - 100 families of flowering plants, by M. Hickey and C. King
  - ...you get the idea...
- Classification enables you to predict features of nearby books given the book in hand
- Imagine if the books were shelved randomly...



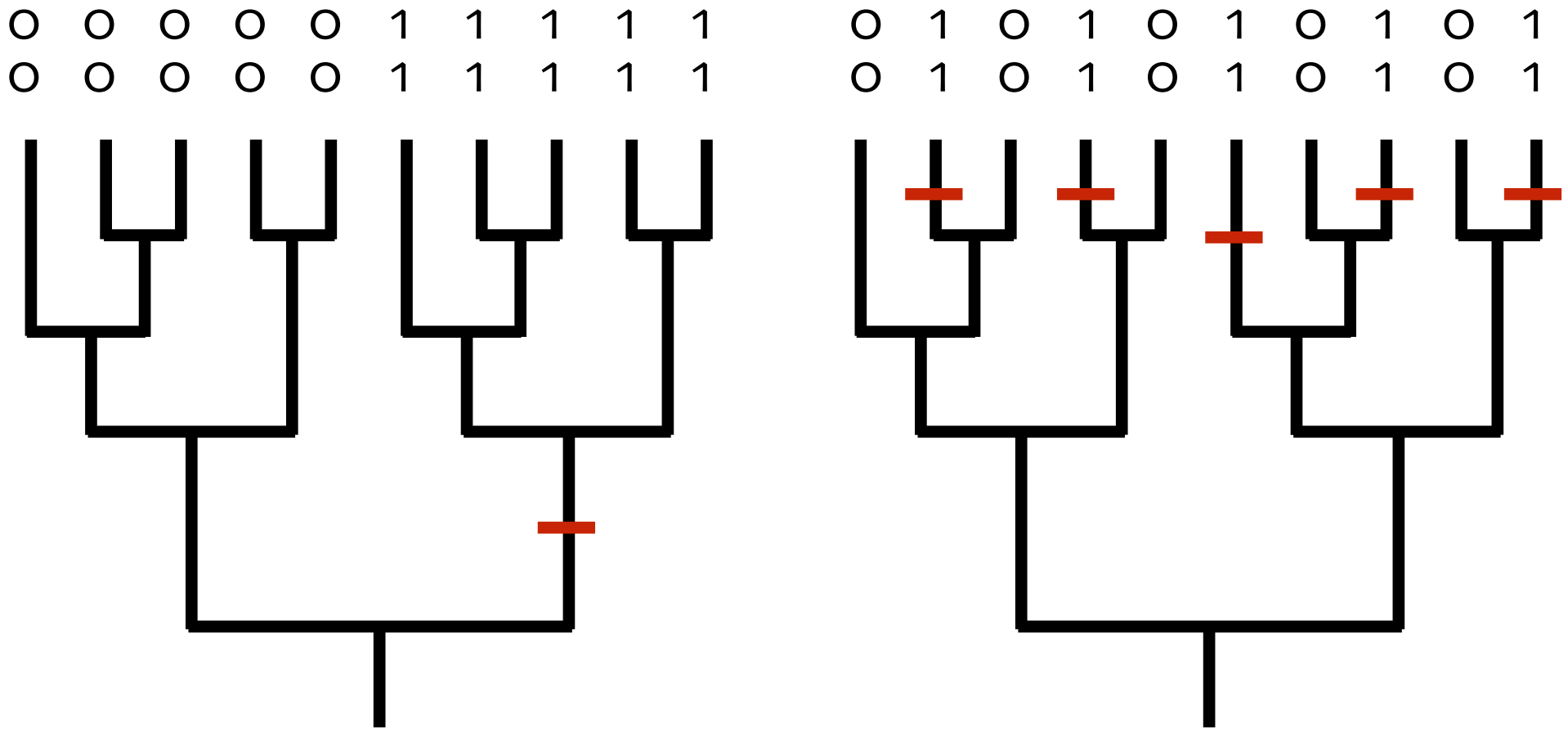
Systematics produces phylogenetically-based classifications because they are rich in information that can be used to make **predictions**

For example, given some knowledge of the features of *Iris*, the phylogeny to the left suggests that you could make more accurate predictions about *Zea* than, say, *Chara*

Correlated Evolution

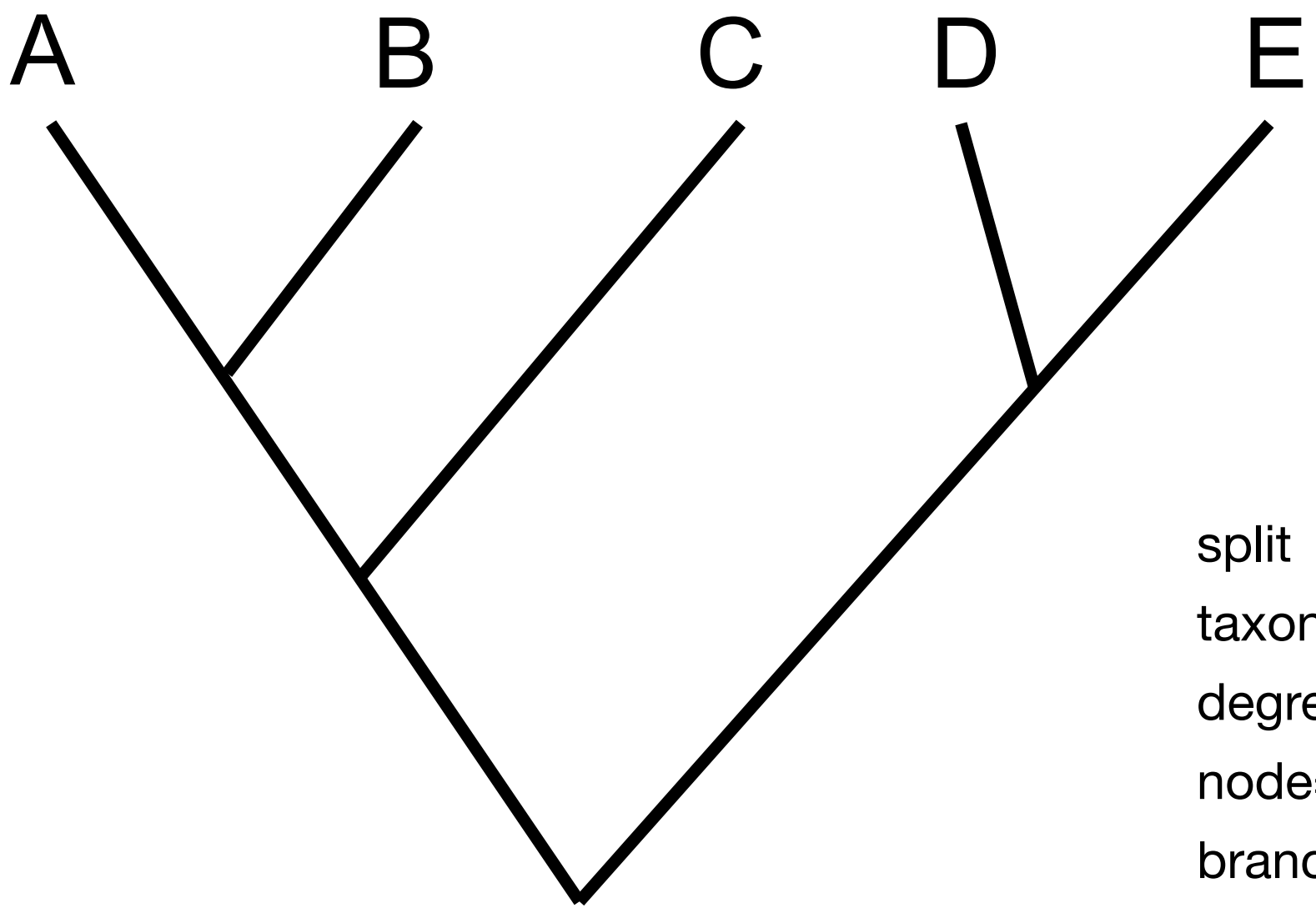
coincidence?

definitely not
independently evolving



What are the parts of a tree?

Tree terminology



split

taxon

degree

node=vertex

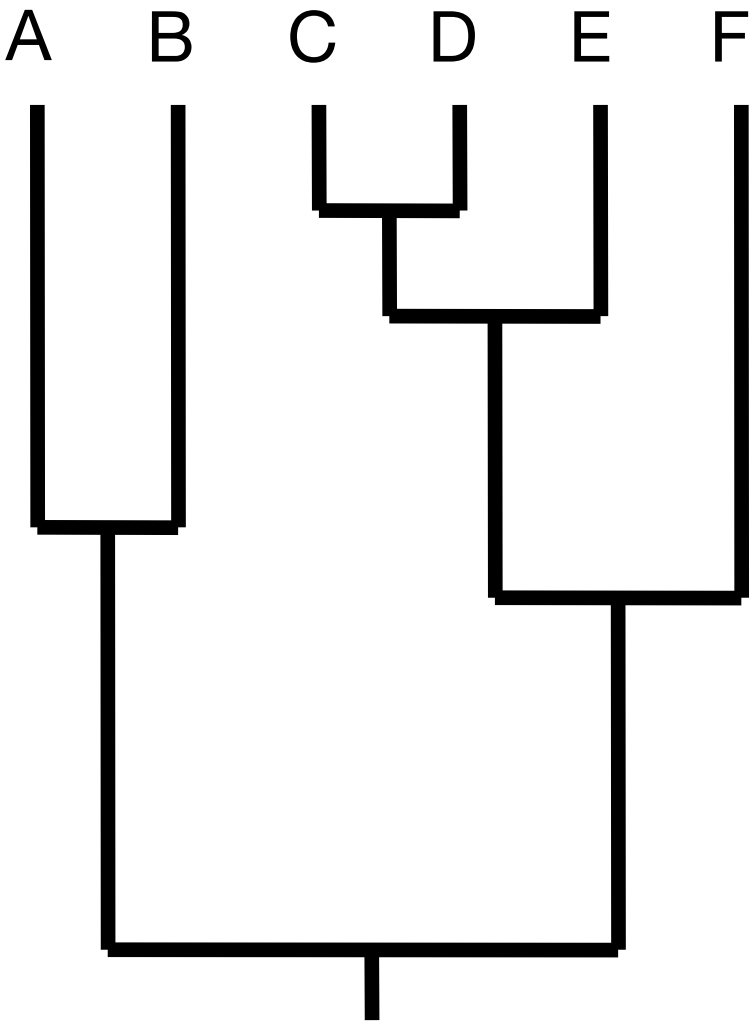
branch=edge

leaf

root

"time tree"

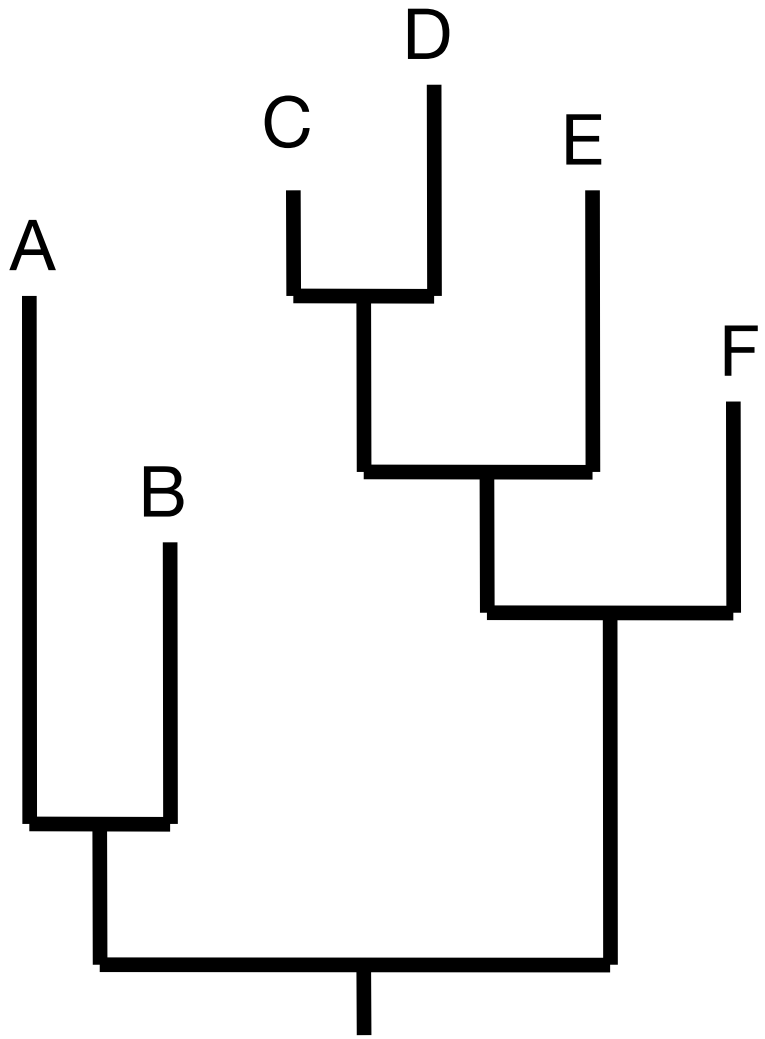
(edge lengths ∝ time)



ultrametric

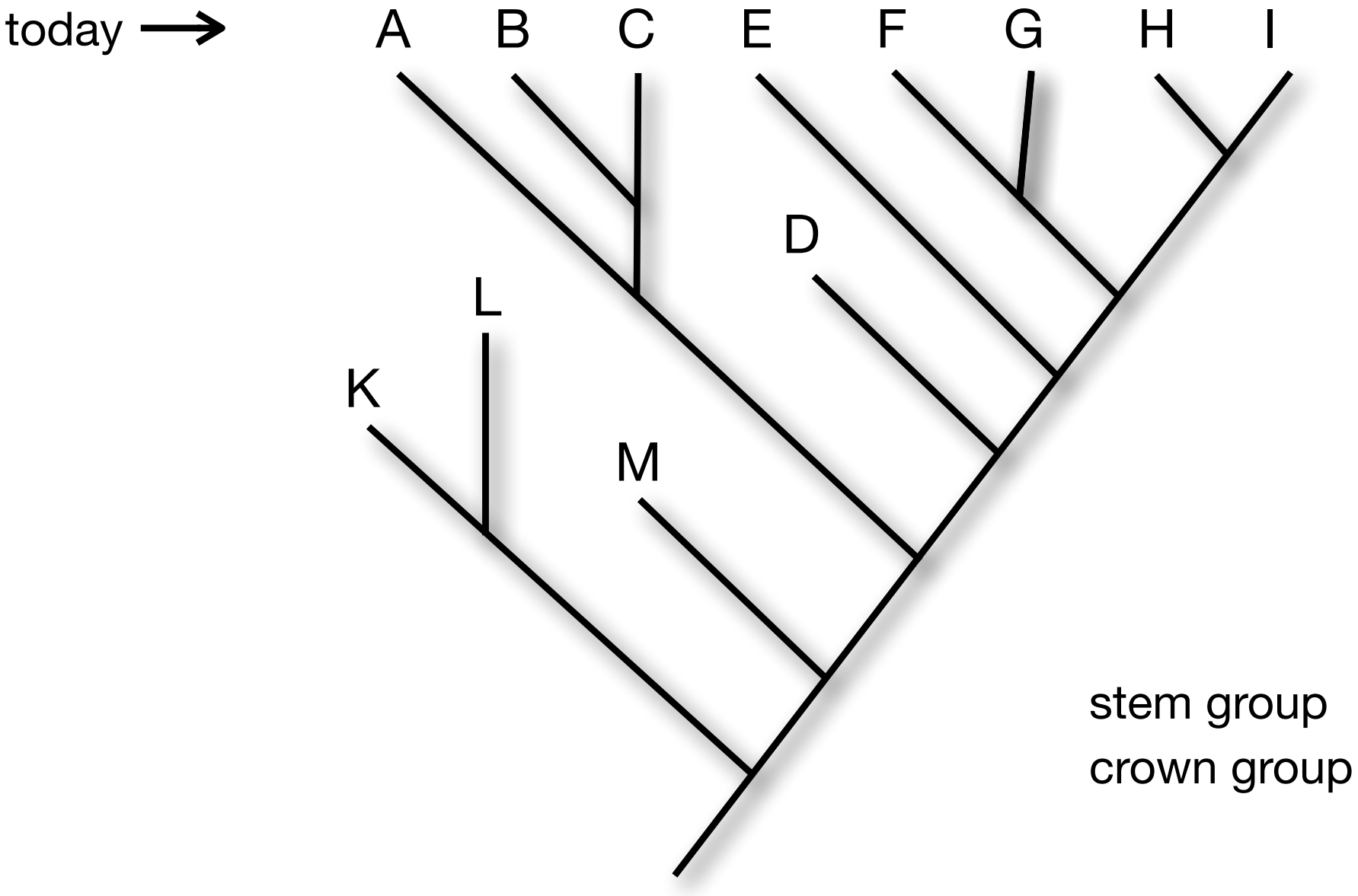
unconstrained tree

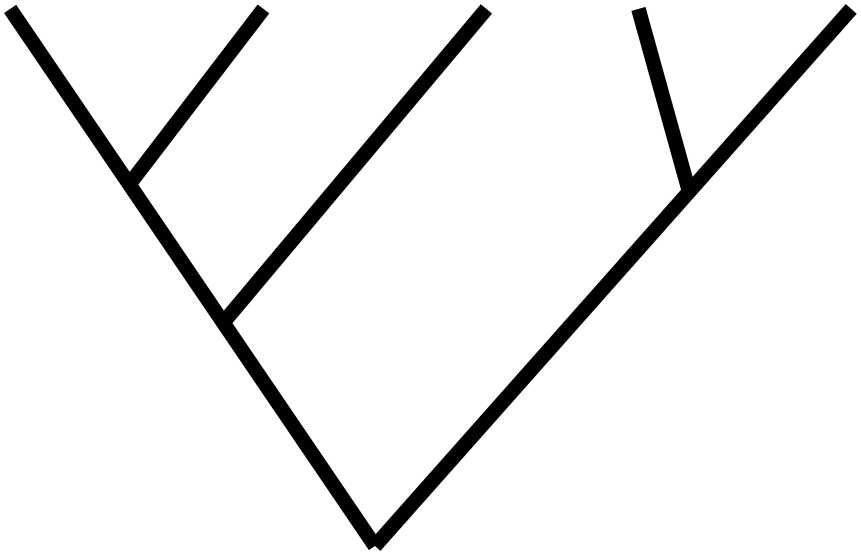
(edge lengths ∝ substitutions)

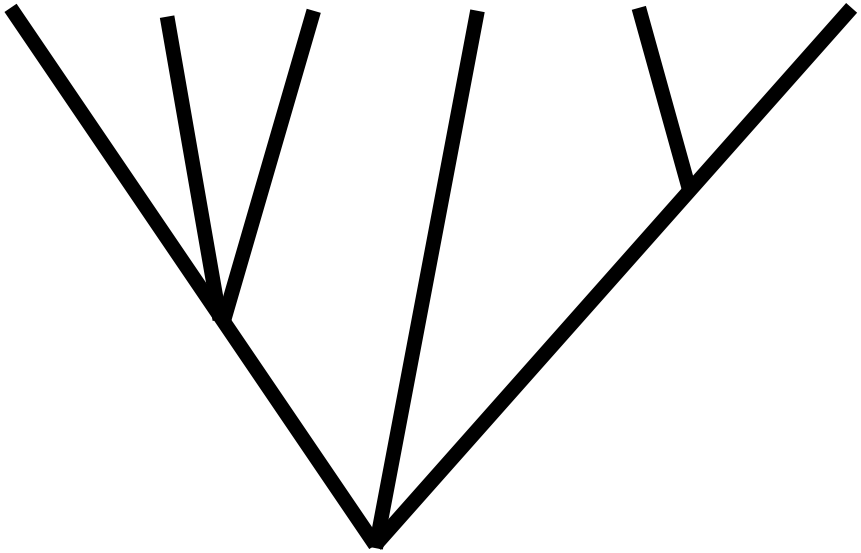


rate x time = substitutions

Tree terminology

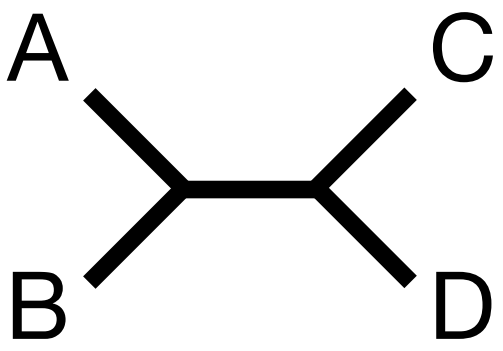


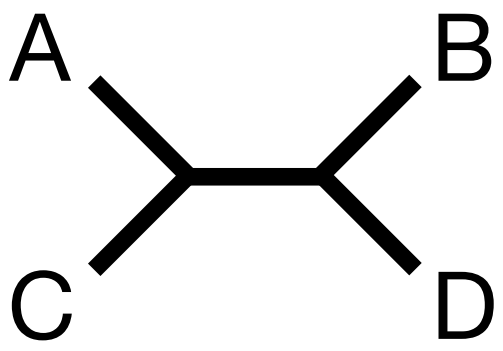


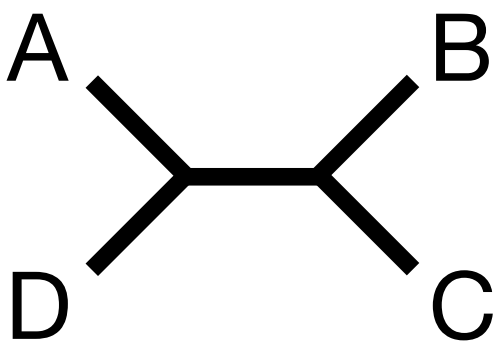


binary

polytomy=multifurcation







topology unrooted rooted

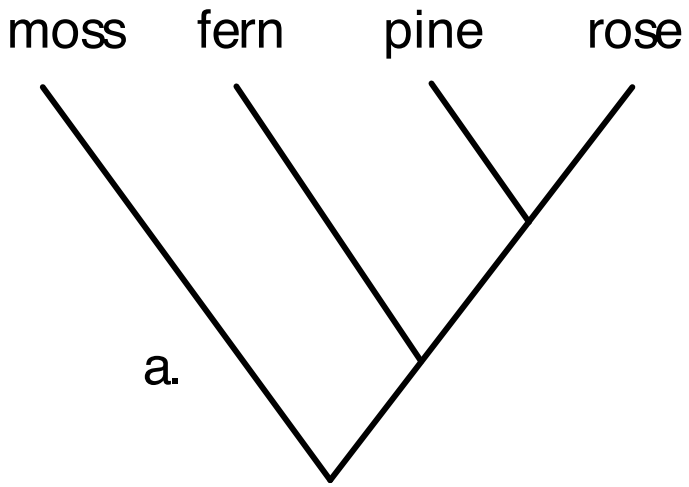
Splits (=bipartitions)

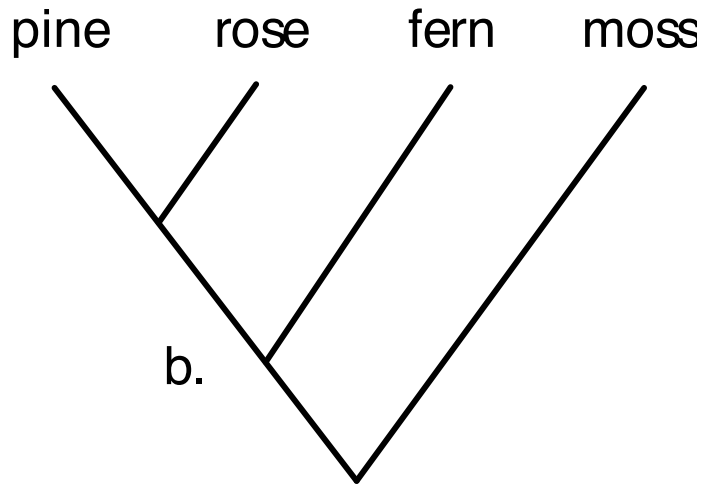
```
Bipartitions found in one or more trees and frequency of
occurrence (bootstrap support values):

          1    1
123456789012345      Freq
-------------------------
..........*.*.*       100
........**.....       100
.**.........*.        100
........***.*.*        94
......*....*...        78
...**********.*        67
.**............        61
......*.*****.*        61
..........*...*        56
...*.*.........        41
..........*.*..        39
..*..........*.        37
.....********.*        33
```

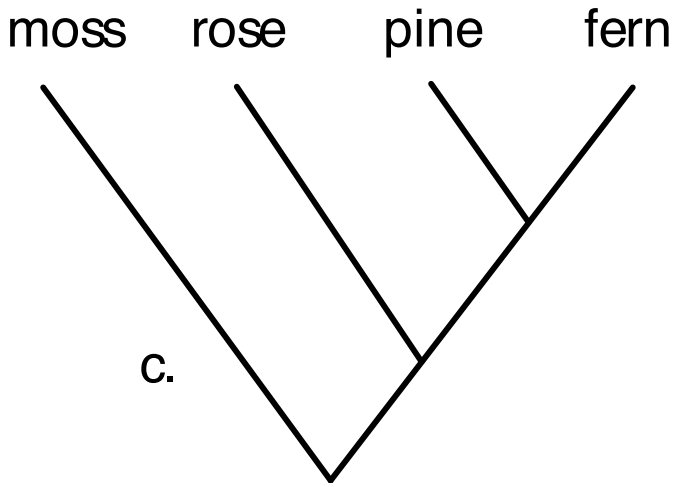
Split on first line defines an edge (branch) that supports a group consisting of taxa 11, 13, and 15.

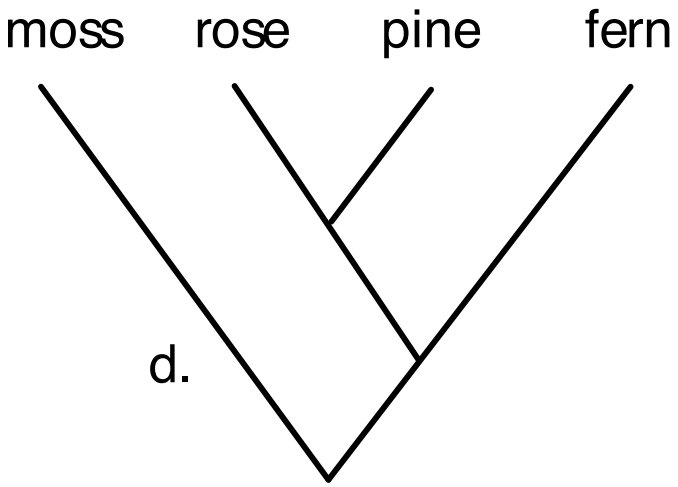
In this analysis, this group is strongly supported (100%) compared to group comprising taxa 3 and 14 (37%)

Which is different?









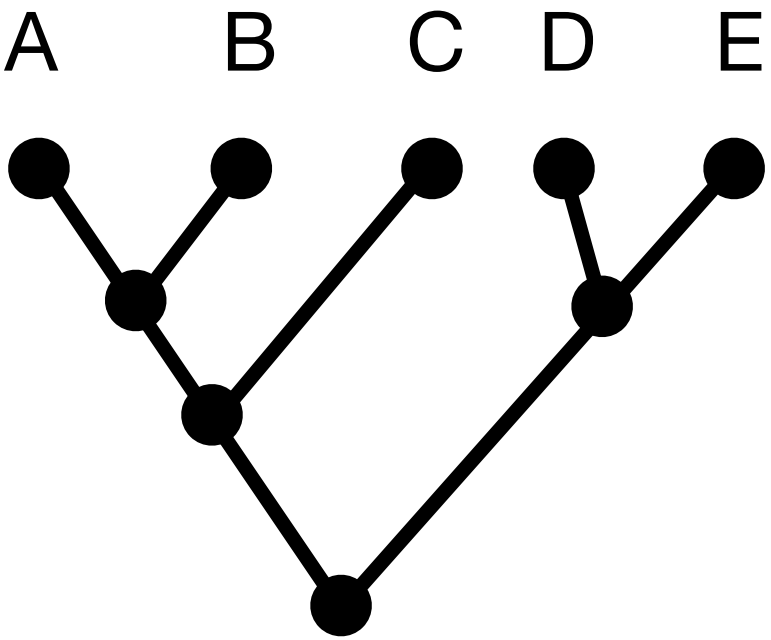
How are trees stored and communicated?

Newick format

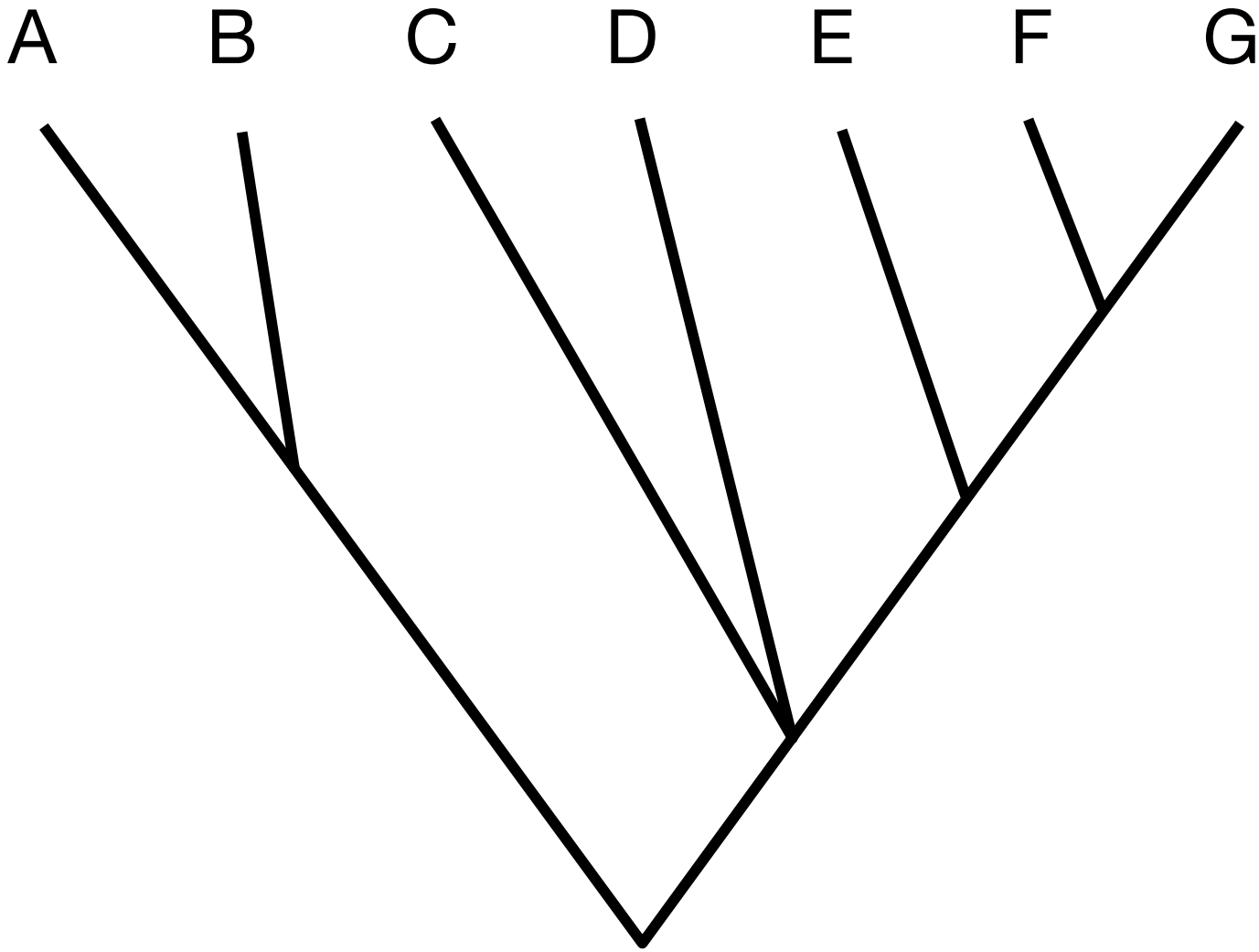
*"**Newick tree format** (or **Newick notation** or **New Hampshire tree format**) is a way of representing graph-theoretical trees with edge lengths using parentheses and commas. It was adopted by James Archie, William H. E. Day, Joseph Felsenstein, Wayne Maddison, Christopher Meacham, F. James Rohlf, and David Swofford, at two meetings in 1986, the second of which was at Newick's restaurant in Dover, New Hampshire, US."*

— https://en.wikipedia.org/wiki/Newick_format

(accessed 17-Jan-2026)



Creating a newick representation



You try it!

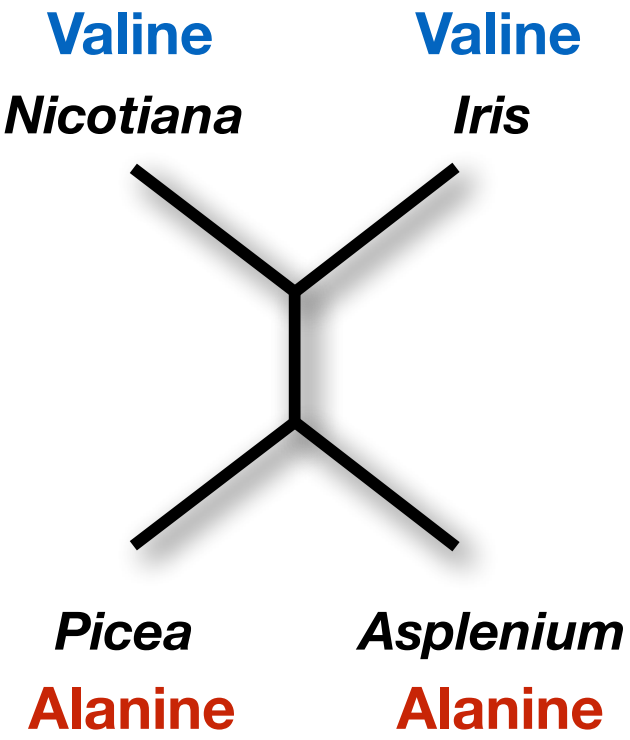
How do you tell if one tree is better than another?

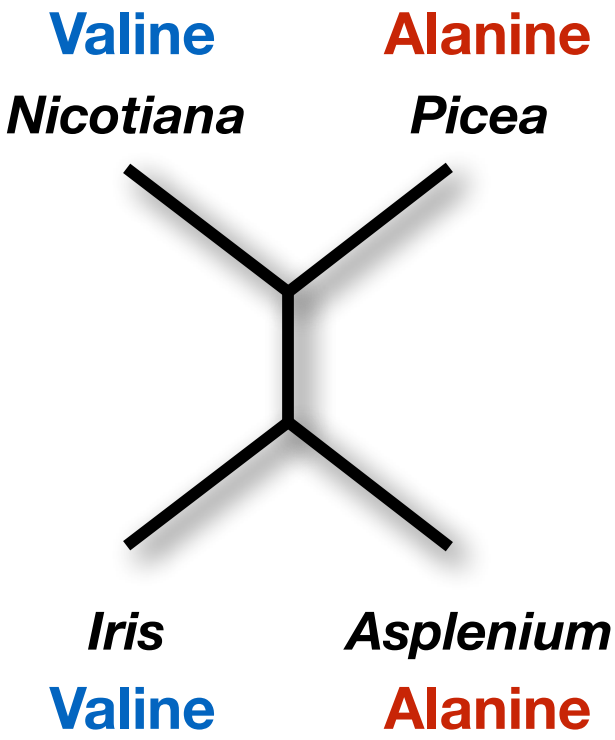
Optimality search combinations

Optimality Criteria	Search Strategies
Maximum Parsimony	Exhaustive Enumeration
Maximum Likelihood	Star Decomposition
Minimum Evolution	Stepwise Addition
Least Squares	Branch-and-Bound
	Branch Swapping

Parsimony criterion

*Given two trees, the one requiring **fewer changes** is best.*





Likelihood Criterion

*Given two trees, the one making us **least surprised** at the data is best.*