

# Topic 10: Phylogenetics

Study of  
evolutionary  
relations


### Explore the Tree of Life

#### Browse the Site

- [Root of the Tree](#)
- [Popular Pages](#)
- [Sample Pages](#)
- [Recent Additions](#)
- [Random Page](#)
- [Treehouses](#)
- [Images, Movies,...](#)

#### Learn about ...

##### [Agaricales](#) (a group of fungi)



[image info](#)

The Agaricales, or euagarics clade, is a monophyletic group of approximately 8500 mushroom species...


[read more](#)

[more featured pages](#)

#### News

Darwin 200: the celebration continues...

[read more](#)



[about this picture](#)

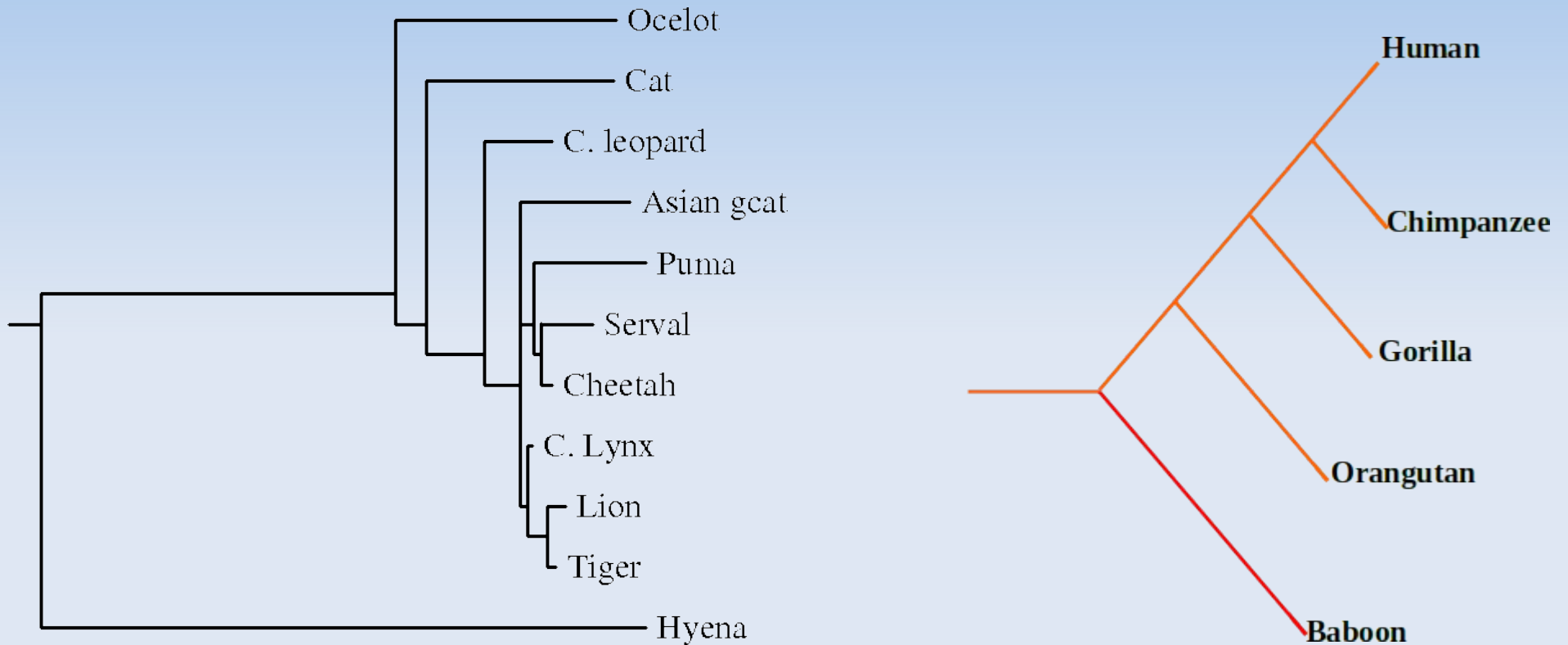
The Tree of Life Web Project (ToL) is a collaborative effort of [biologists and nature enthusiasts from around the world](#). On more than 10,000 World Wide Web pages, the project provides information about biodiversity, the characteristics of different groups of organisms, and their evolutionary history ([phylogeny](#)).

Each page contains information about a particular group, e.g., [salamanders](#), [segmented worms](#), [phlox flowers](#), [tyrannosaurs](#), [euglenids](#), [Heliconius butterflies](#), [club fungi](#), or the [vampire squid](#). ToL pages are linked one to another hierarchically, in the form of the evolutionary tree of life. Starting with the [root of all Life on Earth](#) and moving out along diverging branches to individual species, the [structure of the ToL project](#) thus illustrates the genetic connections between all living things.

[read more about the Tree of Life Web Project...](#)

# BIOT-CIDS 373

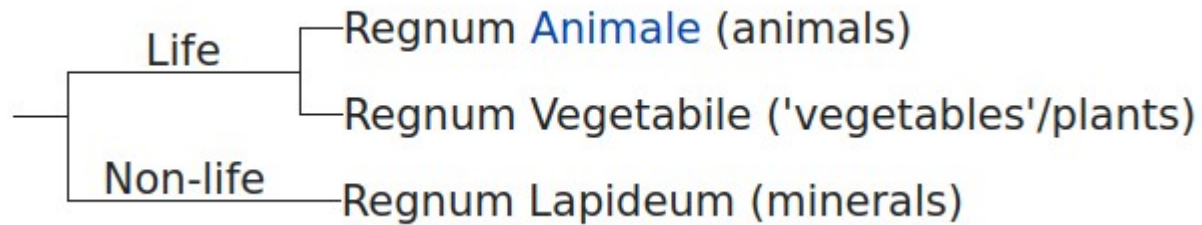
## Phylogenetics



Historically, phylogenetics was based on observable features.  
Now: diagrams like these are created by analyzing sequences!

# The Kingdoms of Life

For 2000 years, from ca. 400 BC to the 1670s, we thought there were just two forms (kingdoms) of life: plants and animals



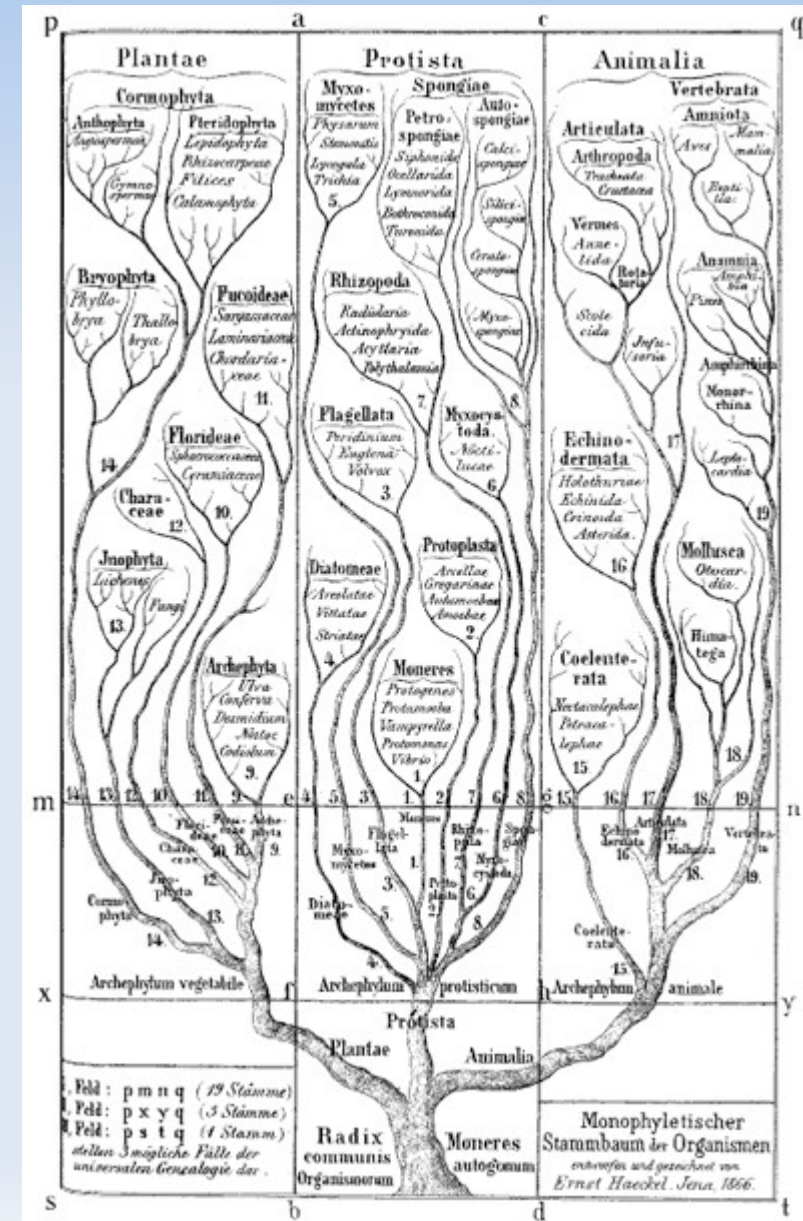
# The Four Kingdoms

Being able to see life forms with **microscopes** changed our thinking.

Ability to observe the cell nucleus → classified organisms into:

- prokaryotes: cells w/o nucleus
- eukaryotes: have cell nuclei

Protists were classified as single celled eukaryotes.



# The Six Kingdoms

It was only in the 20th century that we accepted, based on the **chemistry** of nutrition that Fungi are separate from plants.

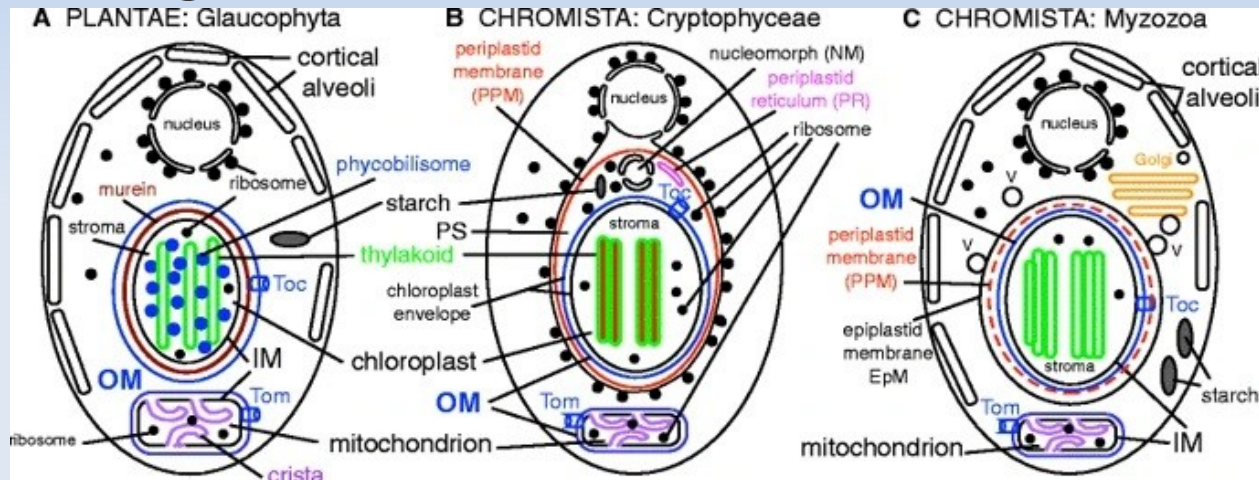
Examining ribosomal RNA "sequences" allowed us to recognize that Archaea were separate from the more common bacteria.





# The Seven Kingdoms

Using a combination of subcellular shapes, biochemistry and sequence analysis, one view (as of 2015) is that there may be 7 "kingdoms":



# Terms

- Phylogenetics
  - Study of evolutionary relationships (sequences or species)
  - Used to infer evolutionary relationship from shared features
- Phylogeny
  - Relationship between organisms & common ancestor
- Phylogenetic tree
  - A graph representing evolutionary history of sequence or species
  - e.g the Living Tree Project from Spain

# New discoveries

- Most life forms have now been placed in 7 to 10 supra kingdom-level groups using *molecular phylogenetics*.
- A recent finding: Hemimastigophora, a predatory protist previously ranked as a phylum.  
In 2018, a phylogenomic analysis placed it outside known groups

## Phylogenomics:

- Burki 2014 The Eukaryotic Tree of Life
- Worden 2015 Rethinking the Marine Carbon Cycle
- Burki 2016 Untangling diversification of eukaryotes

*Also see Domains. There are problems with every classification scheme!*



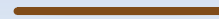
# Tree

The mathematical idea of a "tree" is: a collection of two kinds of things – "nodes" and "edges"

Node



Edge



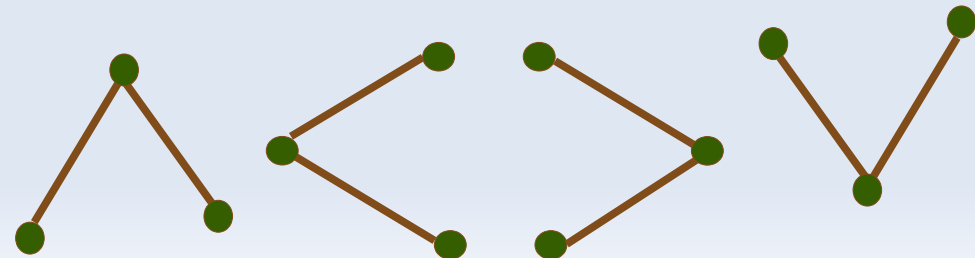
A tree consists of at least one node and zero or more edges with the restriction that there be no loops.

The simplest tree has just a single node



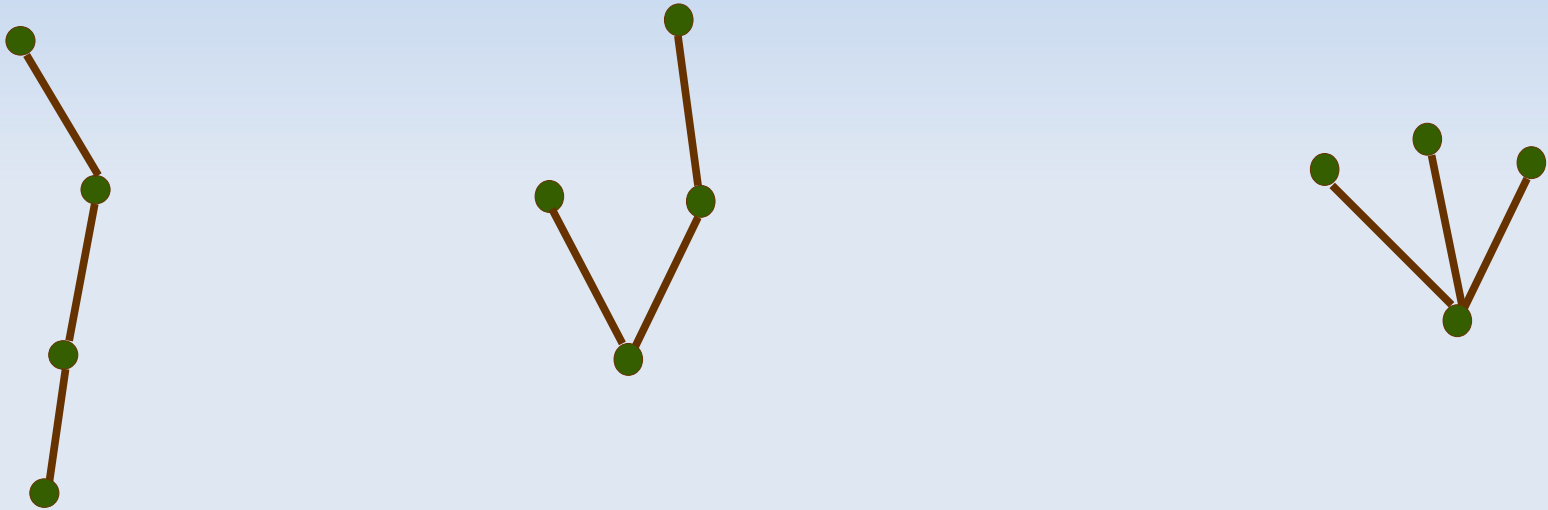
A tree with two nodes has a single edge

Any tree with three nodes must have no more than two edges and no fewer.



# Trees

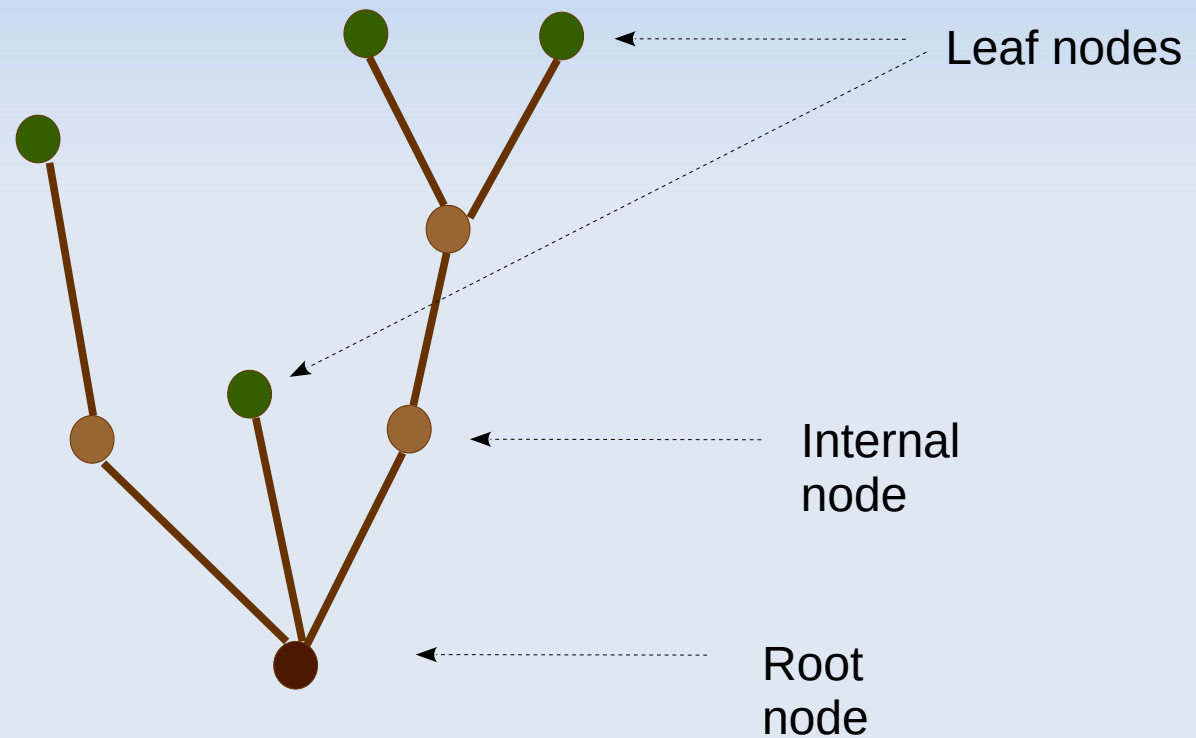
With four nodes, we begin to have more variety in the possible shapes of the trees:



How many edges will a tree with 5 nodes have? How about a 6-node tree? Try drawing these out for yourself.

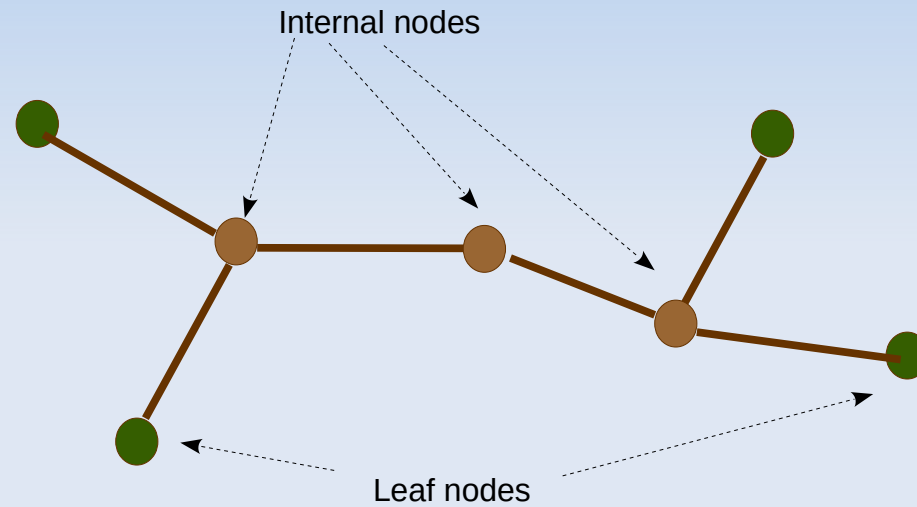
Remember that trees should not have loops

# Types of nodes in trees

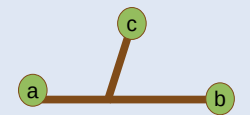
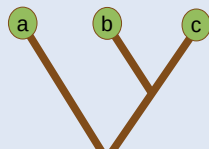
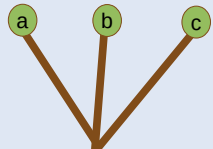
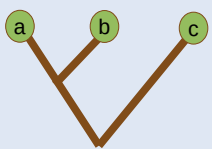


# Unrooted or "free" trees

- Trees without a root node are usually drawn growing outward from a central point:



- 3 rooted trees for 3 leaf nodes! Only one unrooted



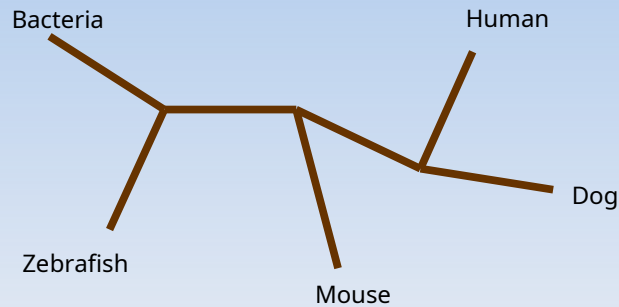
Unrooted tree with  
3 leaf nodes

# Number of possible trees

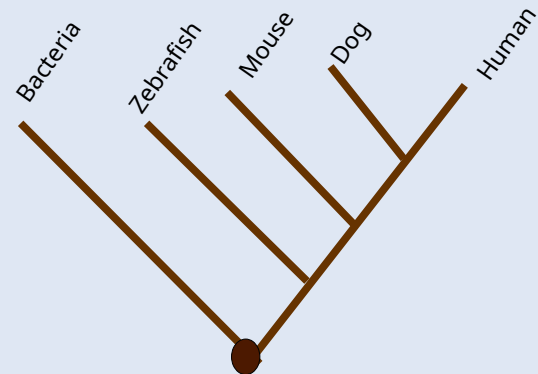
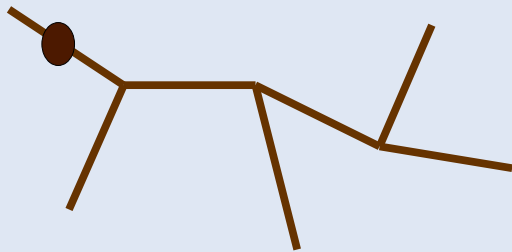
<i>Number of leaf nodes</i>	<i>Number of possible rooted trees</i>	<i>Number of possible unrooted trees</i>
1	1	1
2	1	1
3	3	1
4	15	3
5	105	15
6	945	105
7	10,395	945
8	135,135	10,395
9	2,027,025	135,135
10	34 million	2,027,025



# Converting an unrooted tree



- We can add a root node that splits an edge



# Internal nodes

- Internal nodes represent a hypothetical common ancestor

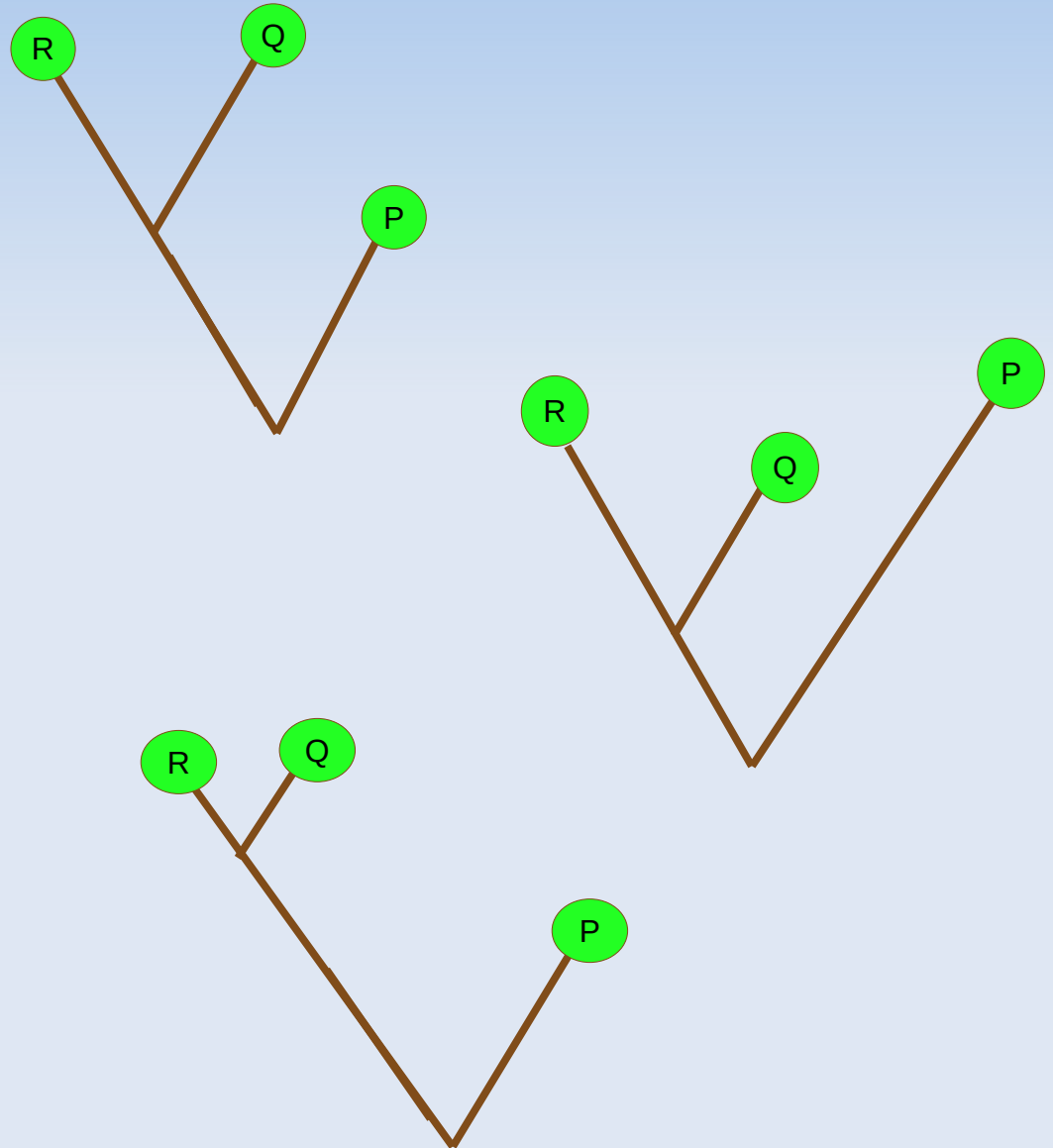


- When we don't have any information about common ancestors, we usually omit them.

# Scaled trees

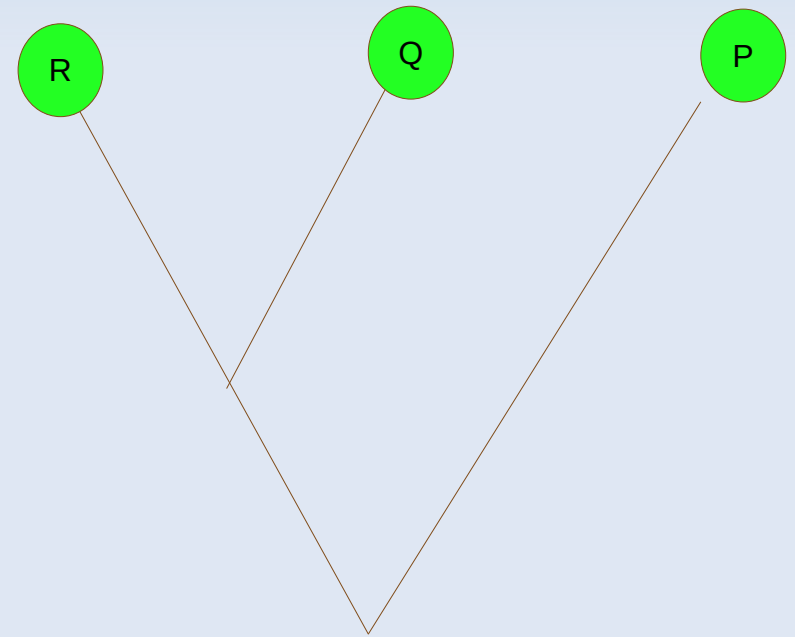
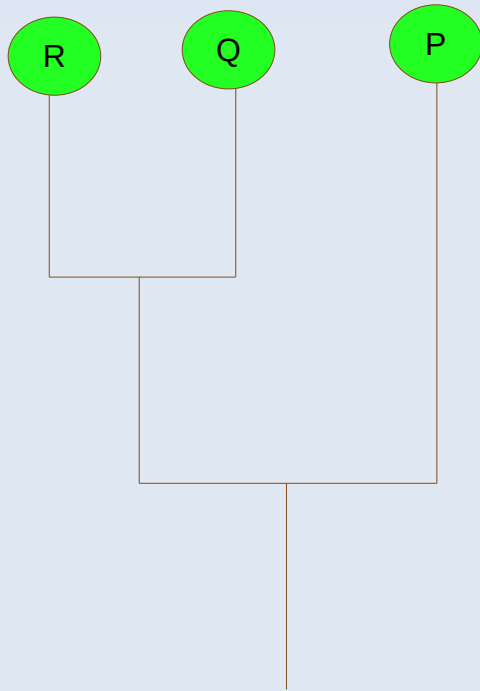
The lengths of edges tell us the "distance" between the nodes – time or sequence differences

Gives us information about how close two species are.



# Cladograms

- Cladograms are trees that show the evolutionary process without indicating time or degree of relatedness



# Phylogenetic Methods

Two main approaches:

- Distance based

- Use the Multiple Sequence Alignment matrix of distances

- Unweighted Pair-Group Method w/ Arith Mean
  - Neighbor Joining

- Best tree

- Try many trees and choose the best one based on:

- Parsimony
  - Maximum Likelihood
  - Bayesian statistics



# Commonly used phylogenetics programs

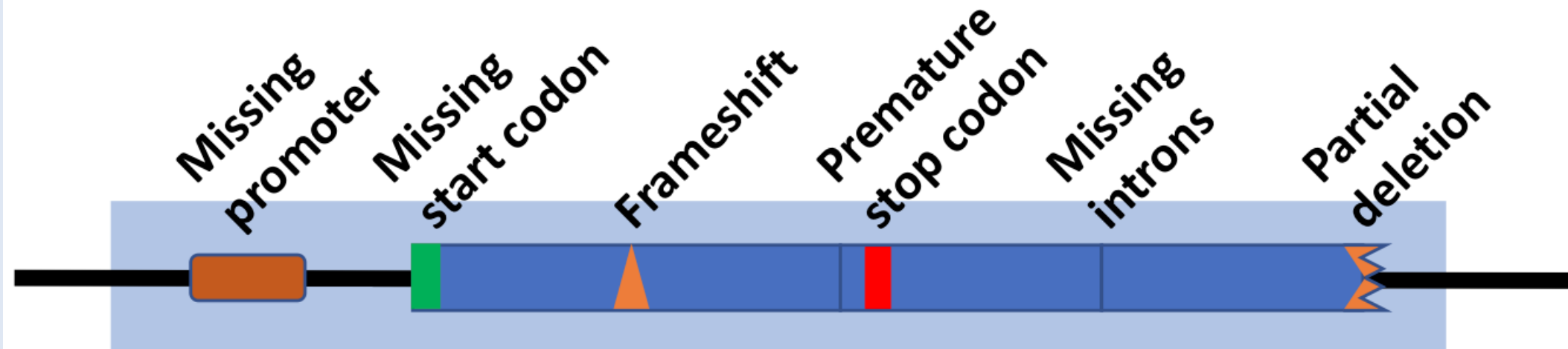
- **PAUP\*** - a commercial package
- **Phylip** – open source
- **Mega** - free
- Many more! See [this list](#)

We could write our own Python program!

# Pseudogene

A pseudogene is segment of genomic DNA that is non-functional but is similar to functional genes in either the same organism or in a related one.

## Common defects of pseudogenes:



Often found by phylogenetic analysis