



NATURAL  
HISTORY  
MUSEUM  
LOS ANGELES COUNTY



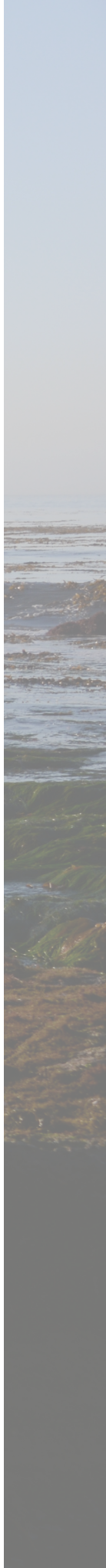
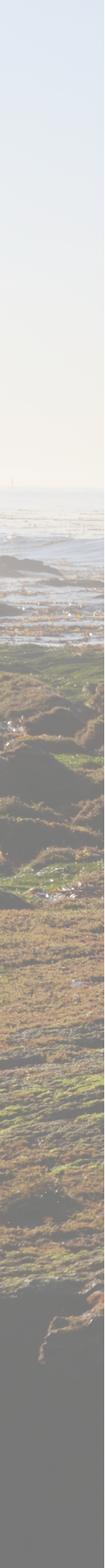
Diversity Initiative for the  
Southern California Ocean

# A Sea of Trees

*Regina Wetzer*

Marine Biodiversity Center & DISCO Project

July 2017



*Where did the diversity we  
have today come from?  
What's so special about it?*

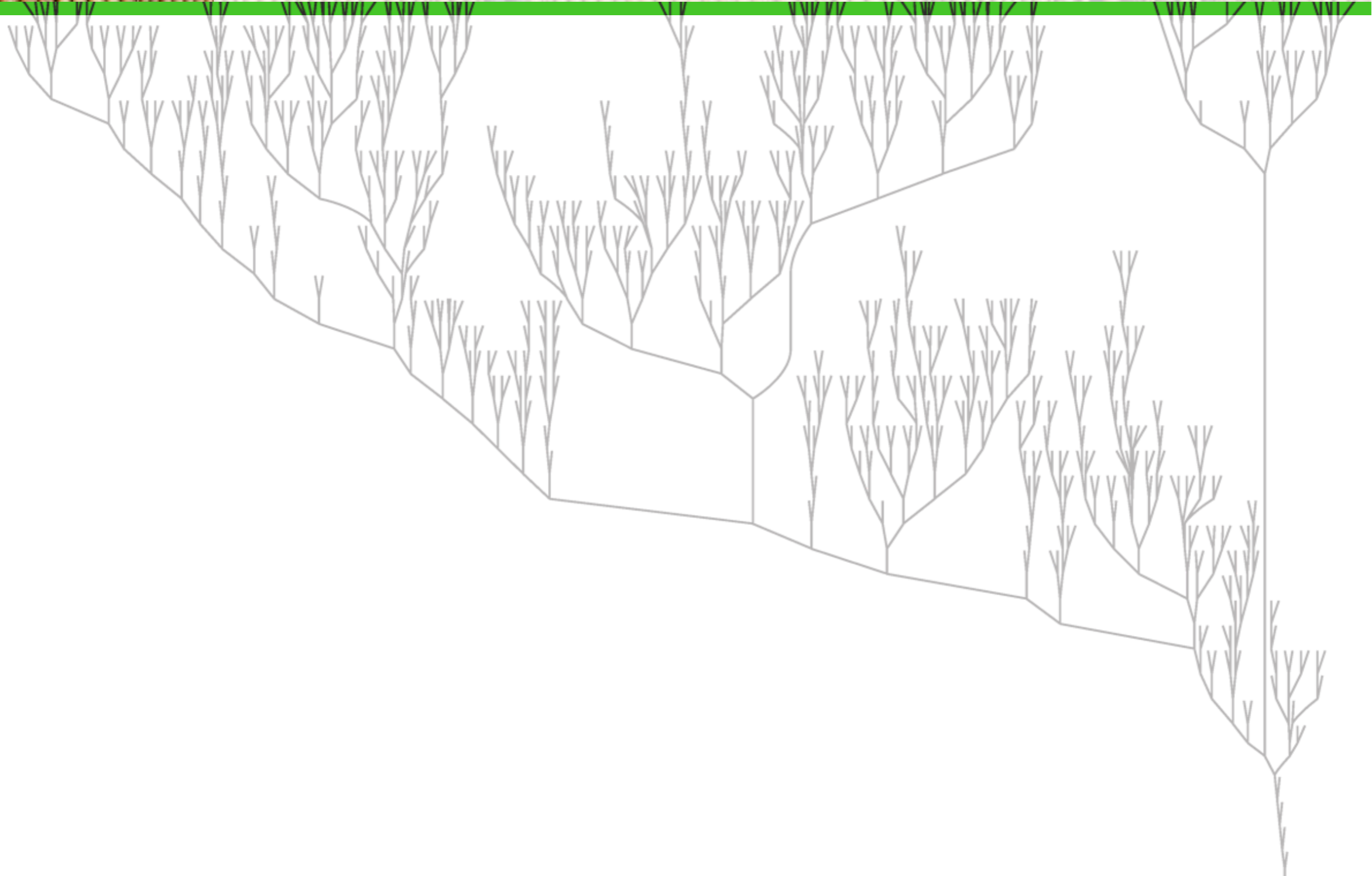
**1,000,000  
described**

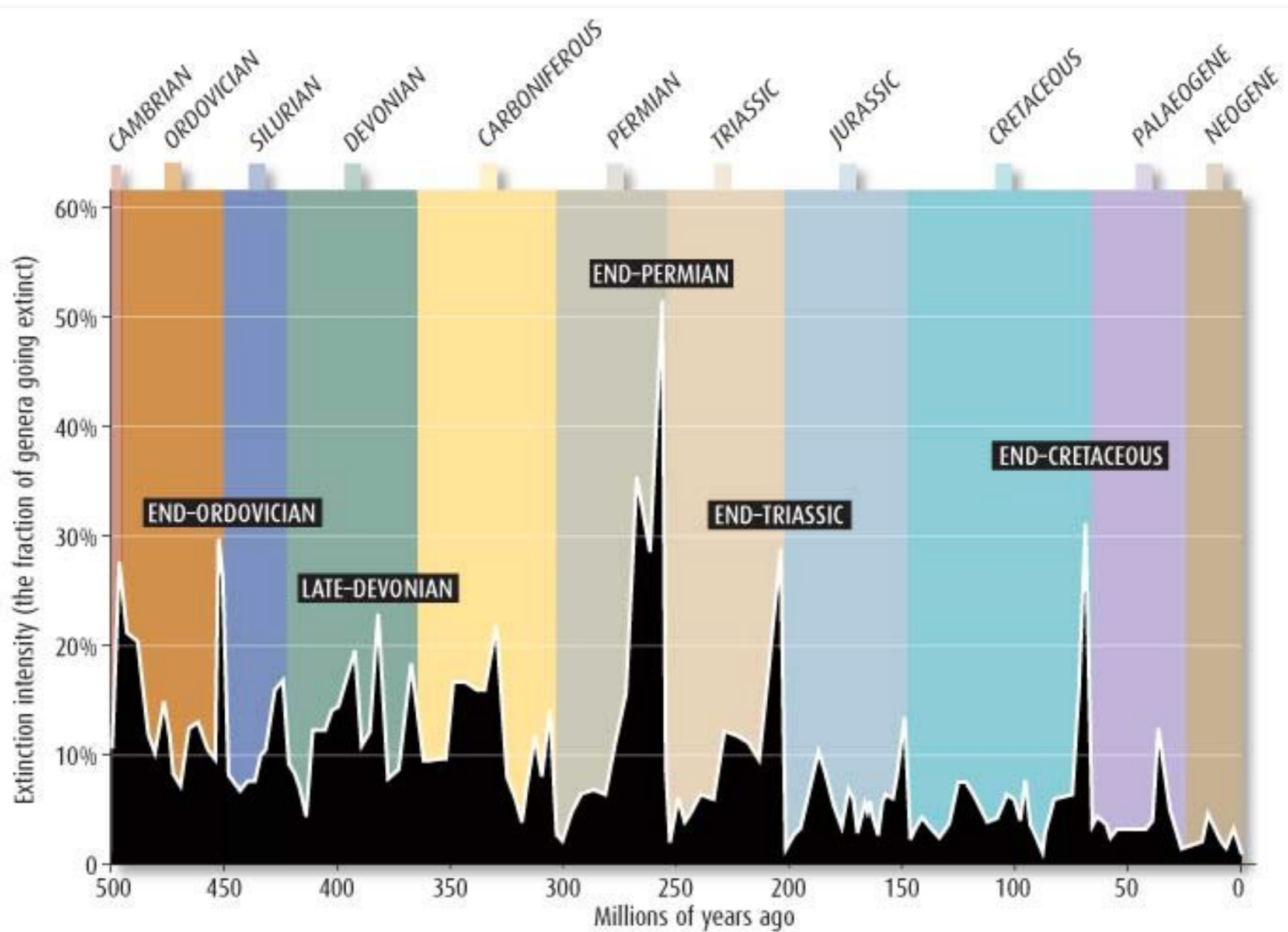
**5,000,000 – 10,000,000  
undescribed**

**Now**

**Time →**

**Origin of Life**





# Investigating the Tree of Life

Phylogeny is the evolutionary history of a species or group of related species

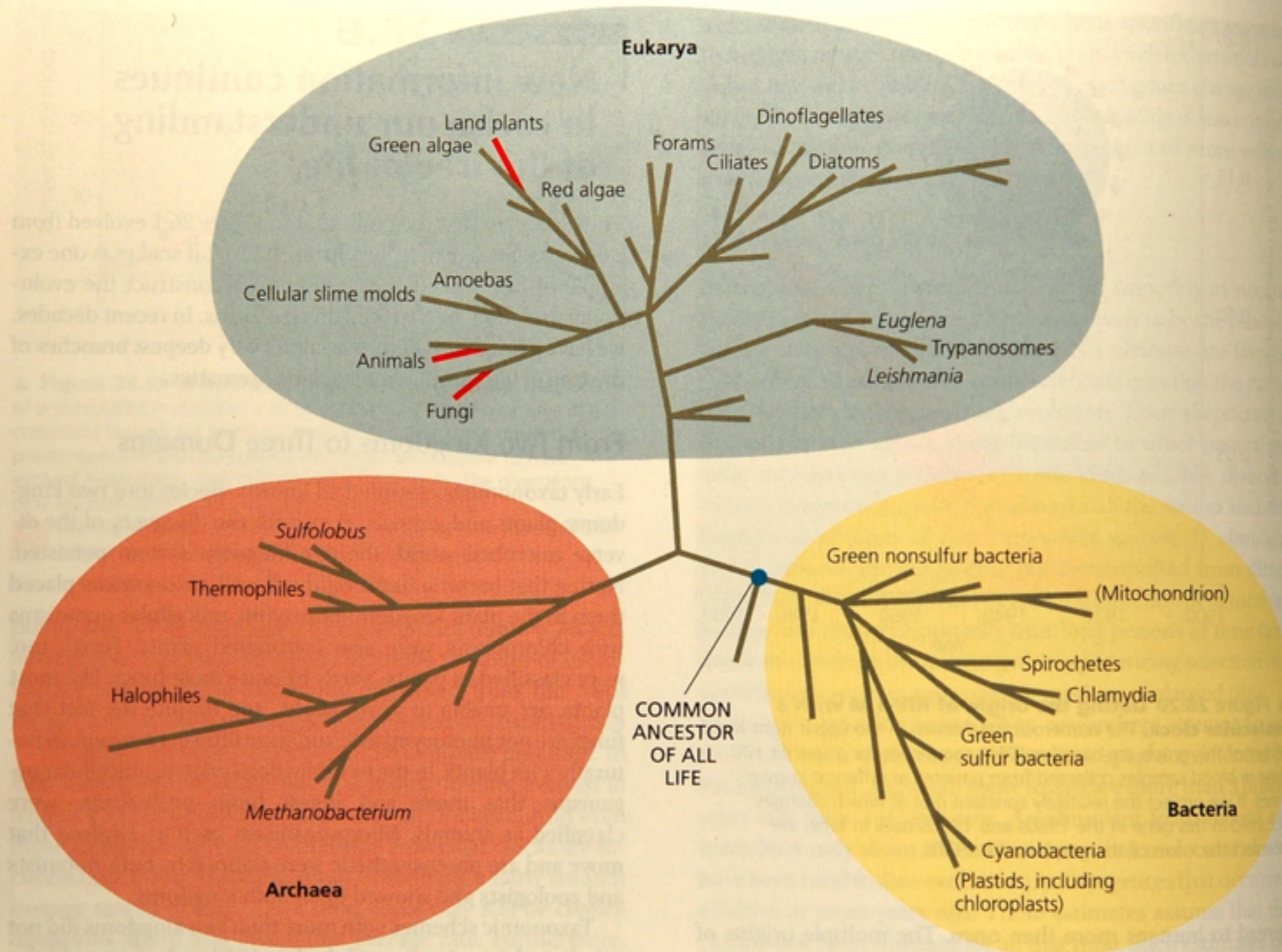
The discipline of systematics classifies organisms and determines their evolutionary relationships

- Systematists use fossil, morphological, developmental, and genetic data to infer evolutionary relationships

# Hierarchical Classification

- Linnaeus introduced a system for grouping species in increasingly broad categories
- The taxonomic groups from broad to narrow are domain, kingdom, phylum, class, order, family, genus, and species
- A taxonomic unit at any level of hierarchy is called a taxon

# Three Domains of life



▲ **Figure 26.21 The three domains of life.** Based on rRNA gene sequences, all of life can be grouped into three domains. Branch lengths are proportional to the amount of genetic change in each lineage. (To simplify the figure, only some branches are labeled.) In this diagram, the lineages within Eukarya that are dominated by multicellular organisms (plants, fungi, and animals) are shown



▲ **Figure 26** history of li gene transfer, more such eve

This reco quence comp parts of ribo workings of t mologies be tected—mal evolutionary of life. How ships. For e genes that i



Fig. 26-3

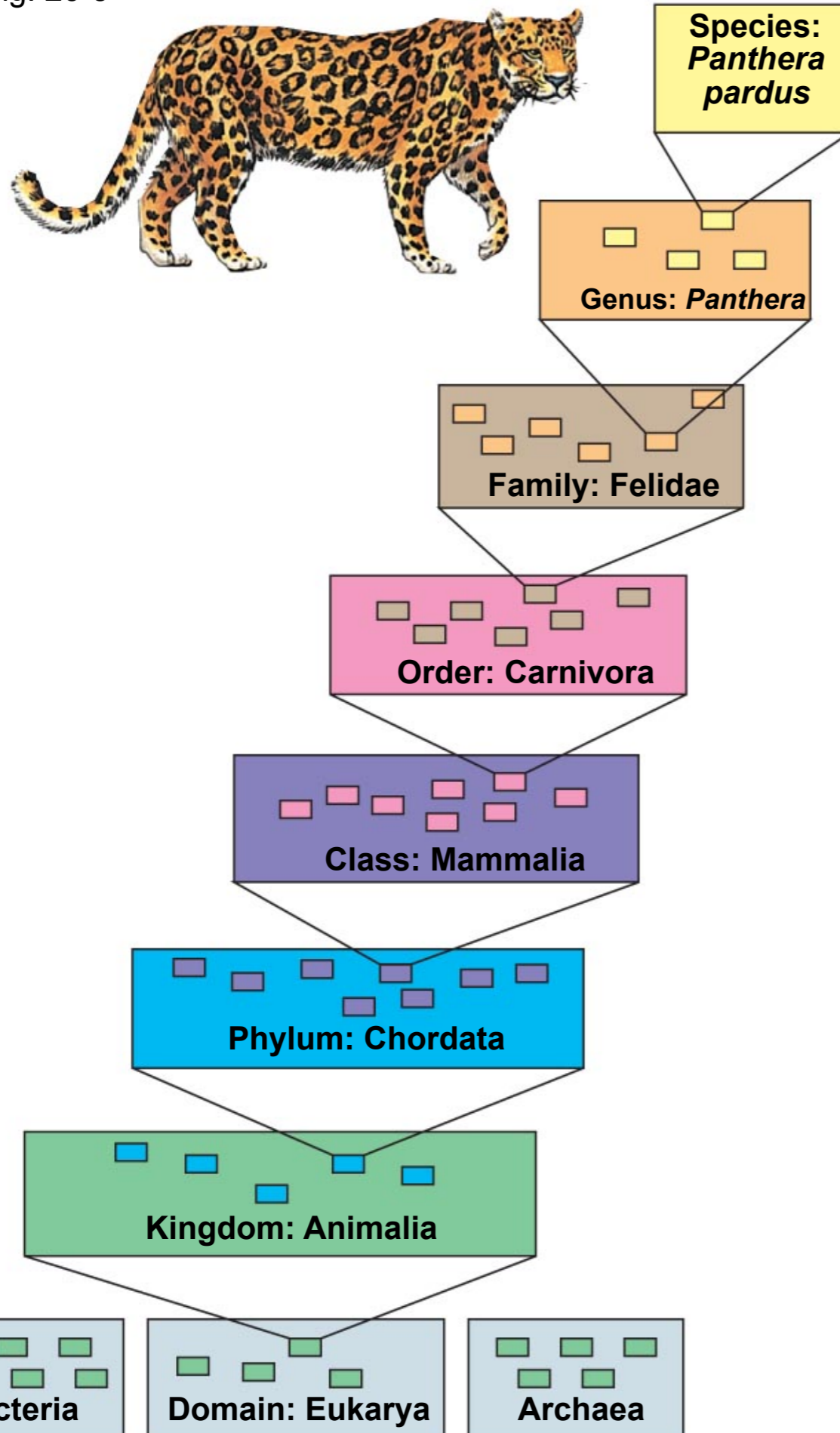
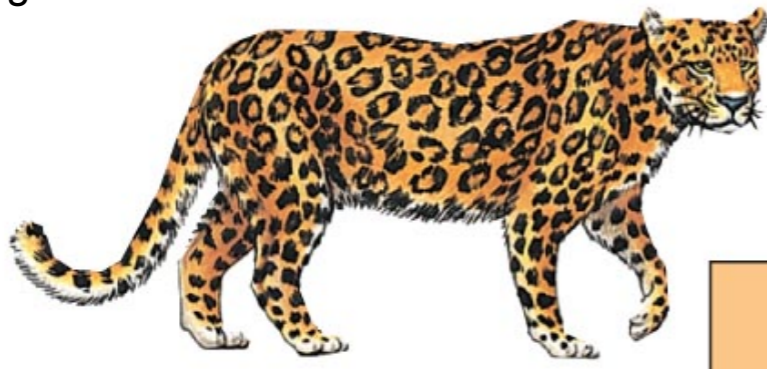


Fig. 26-3



Species:  
*Panthera  
pardus*

Genus: *Panthera*

Family: Felidae

Order: Carnivora

Class: Mammalia

Phylum: Chordata

Kingdom: Animalia

Bacteria

Domain: Eukarya

Archaea

*Bathynomus giganteus*

*Bathynomus*

Cirolanidae

Cymothoidea

Crustacea

Arthropoda

same

same

# Linking Classification and Phylogeny

Systematists depict evolutionary relationships in branching phylogenetic trees

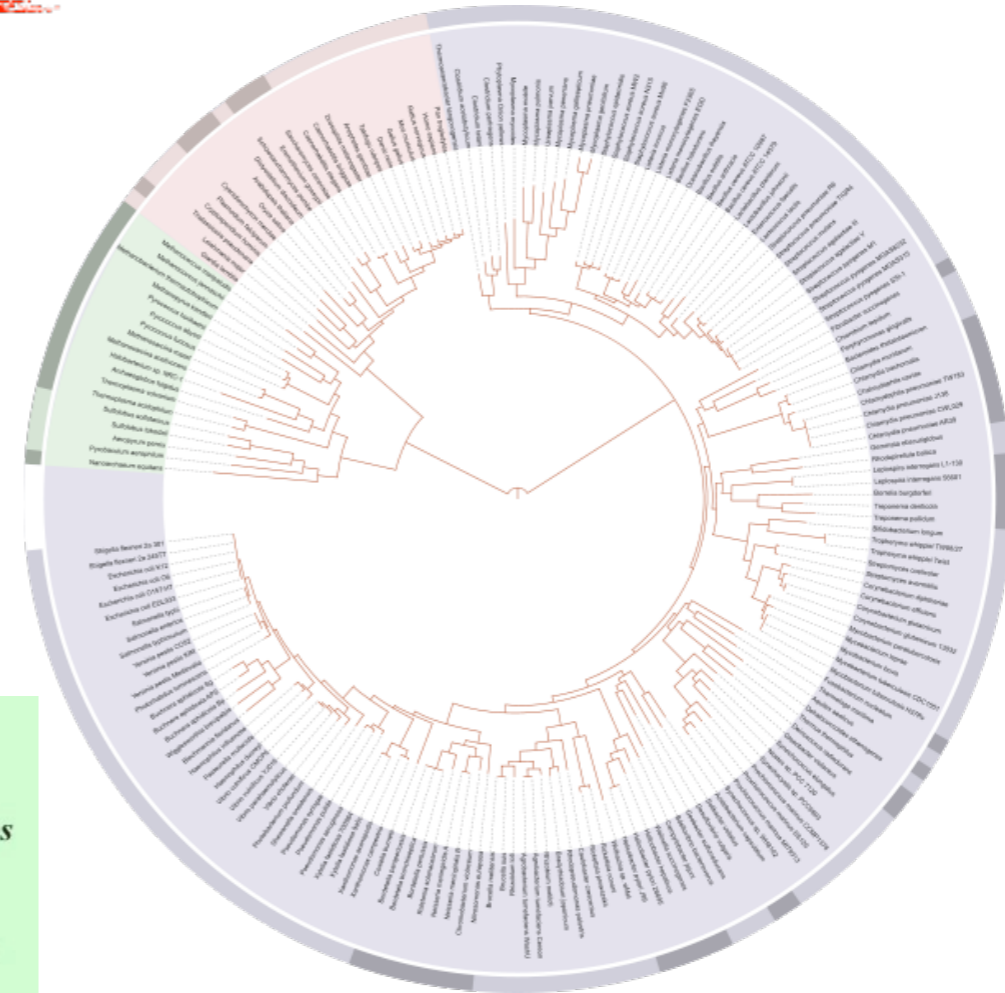
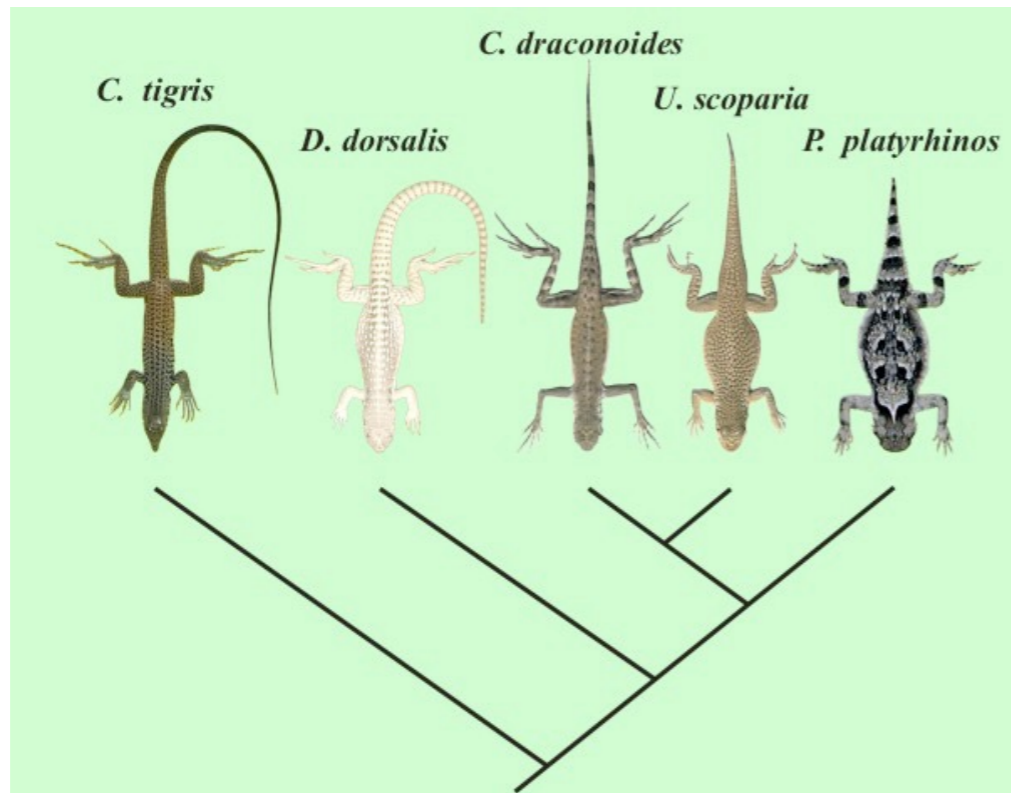
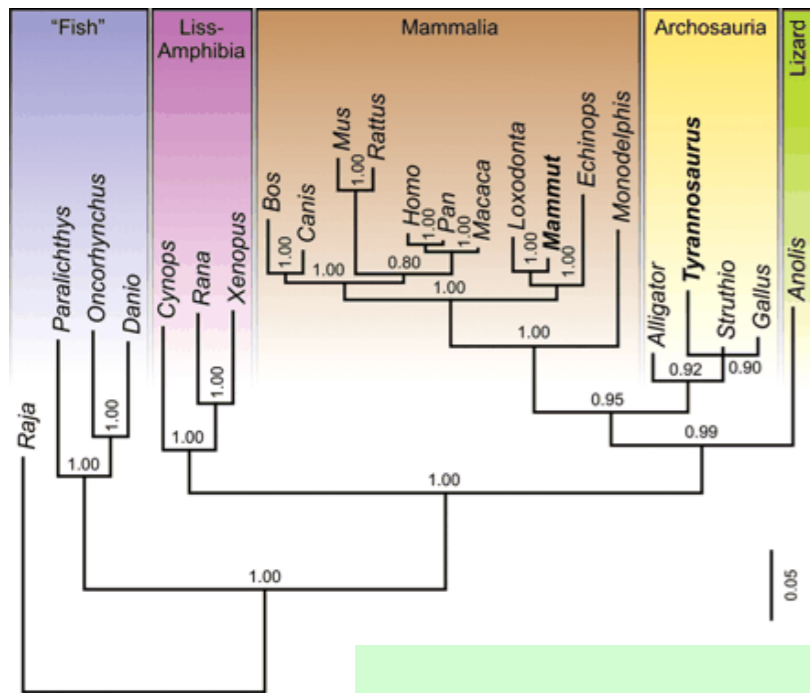


Fig. 26-4

**Order**

**Family**

**Genus**

**Species**

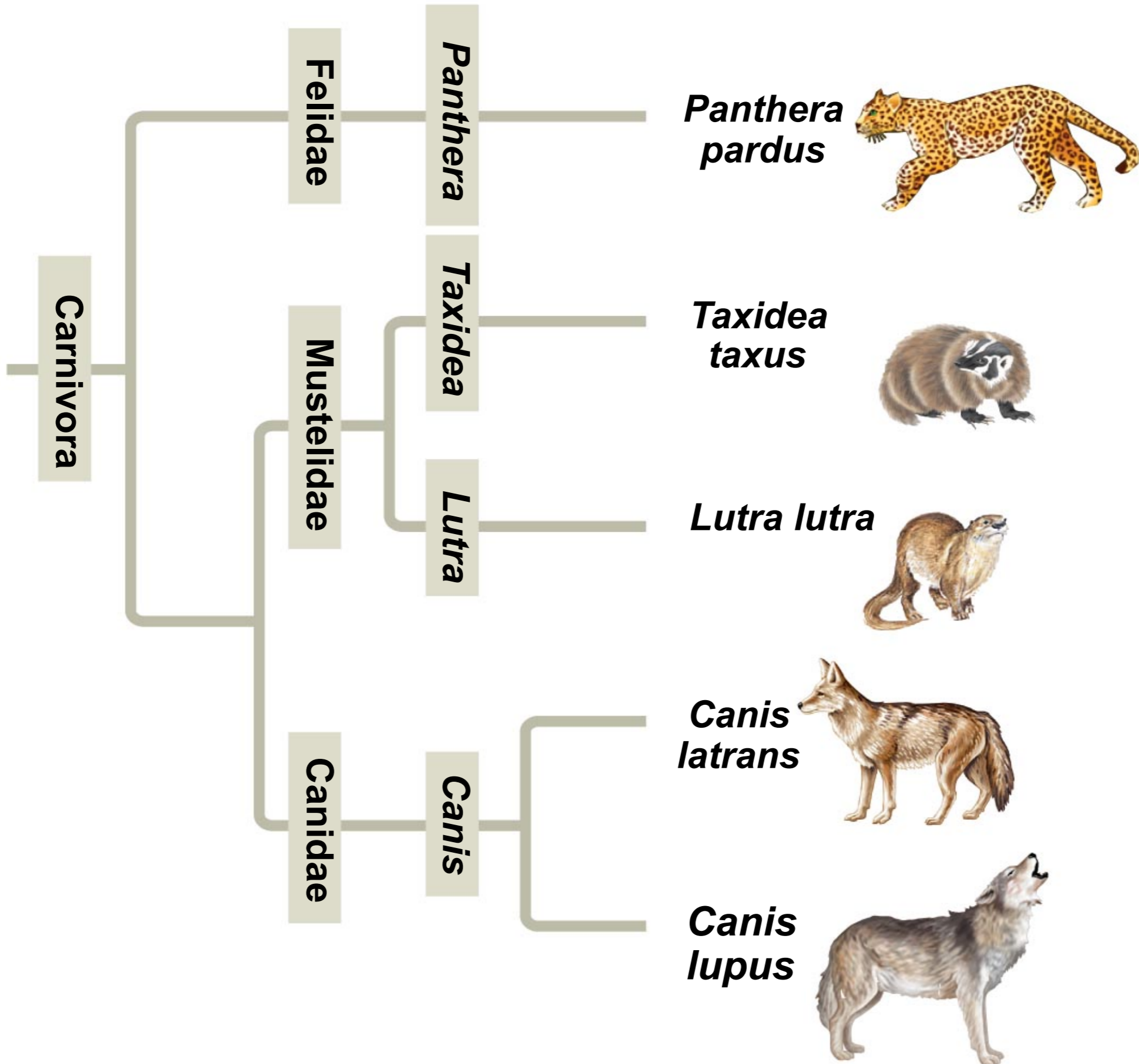
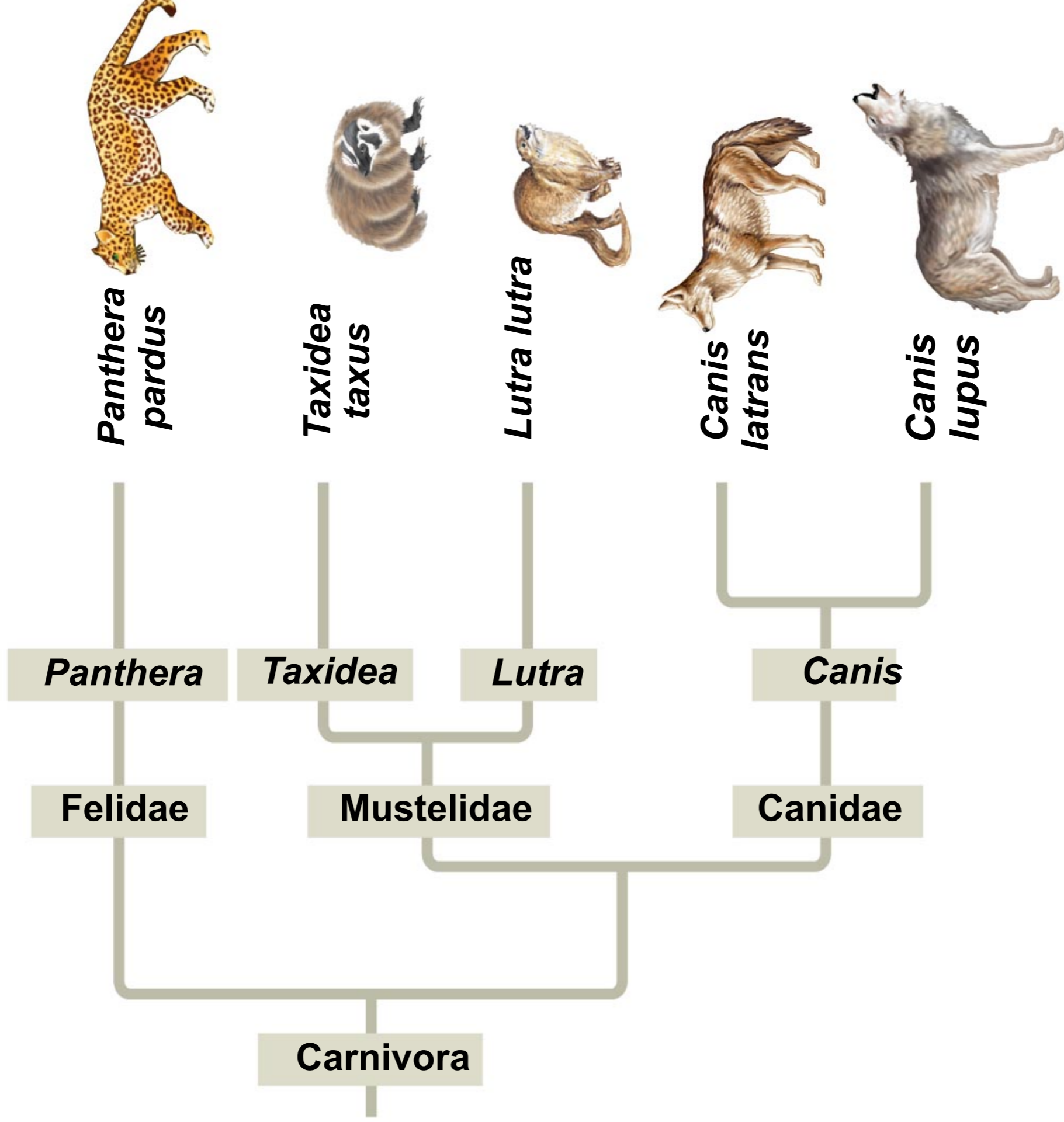


Fig. 26-4

**Order**

**Family Genus**

**Species**

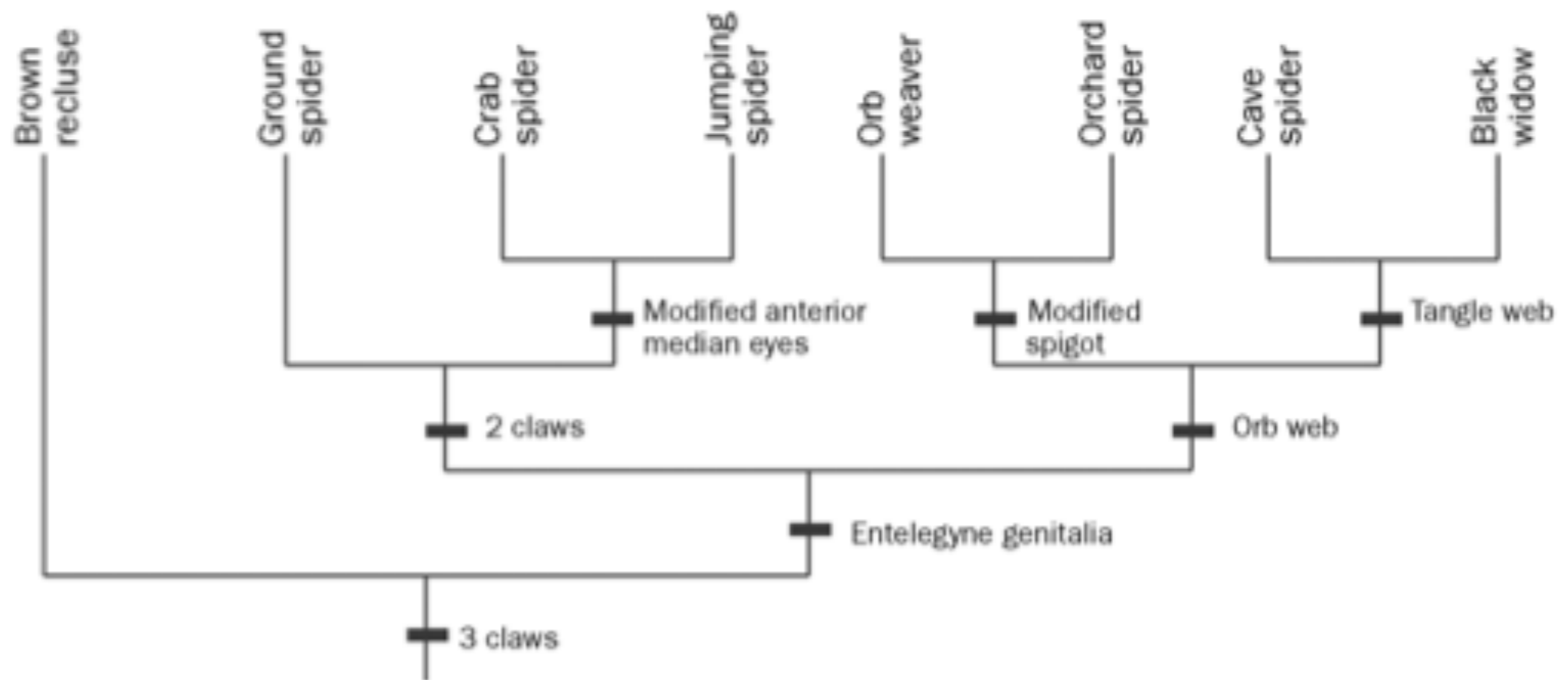


- A phylogenetic tree represents a hypothesis about evolutionary relationships
- Each branch point represents the divergence of two species

Sister taxa are groups that share an immediate common ancestor

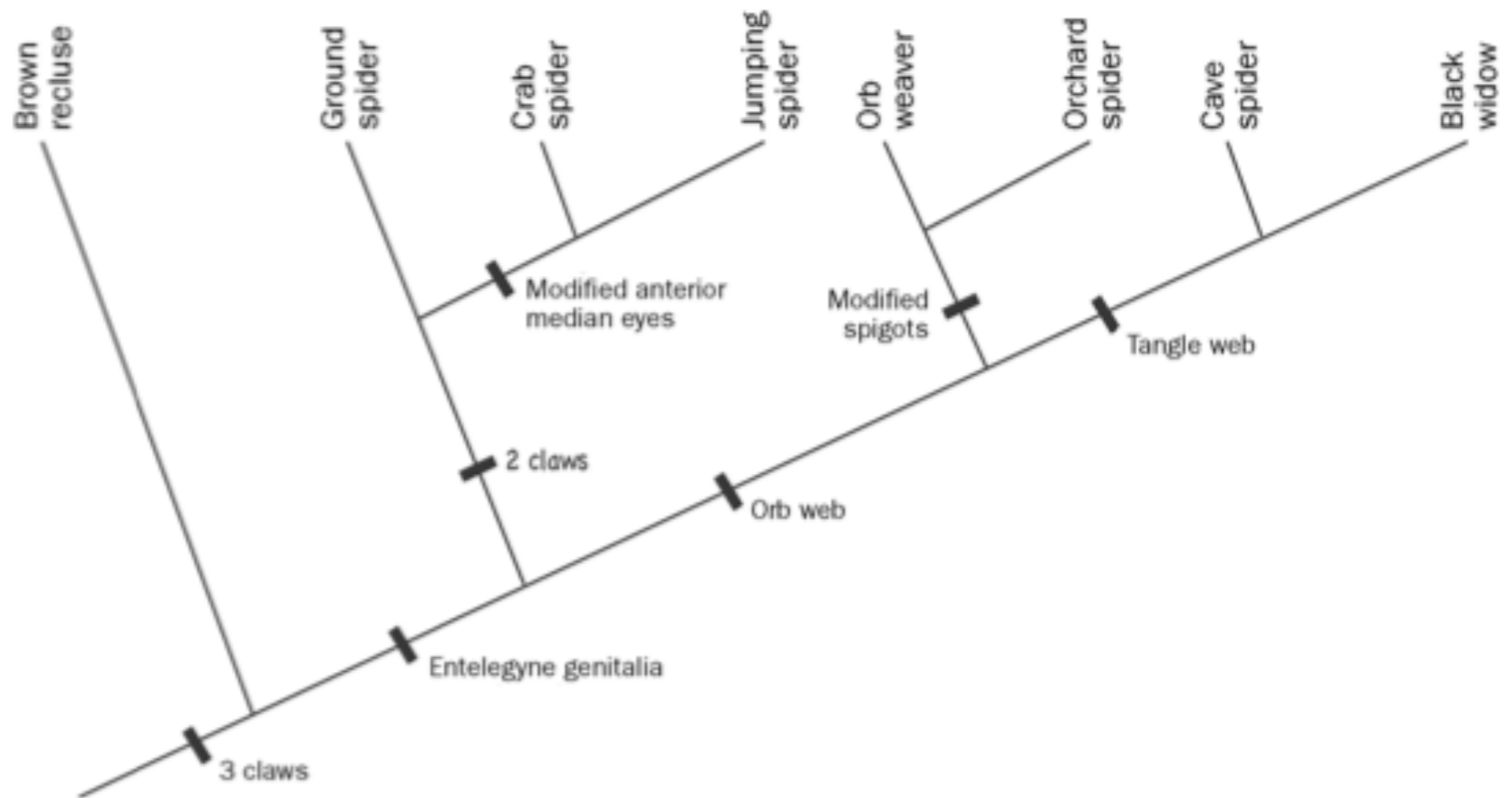
# Different Kinds of Trees

## a. Cladogram in the tree format



# Different Kinds of Trees

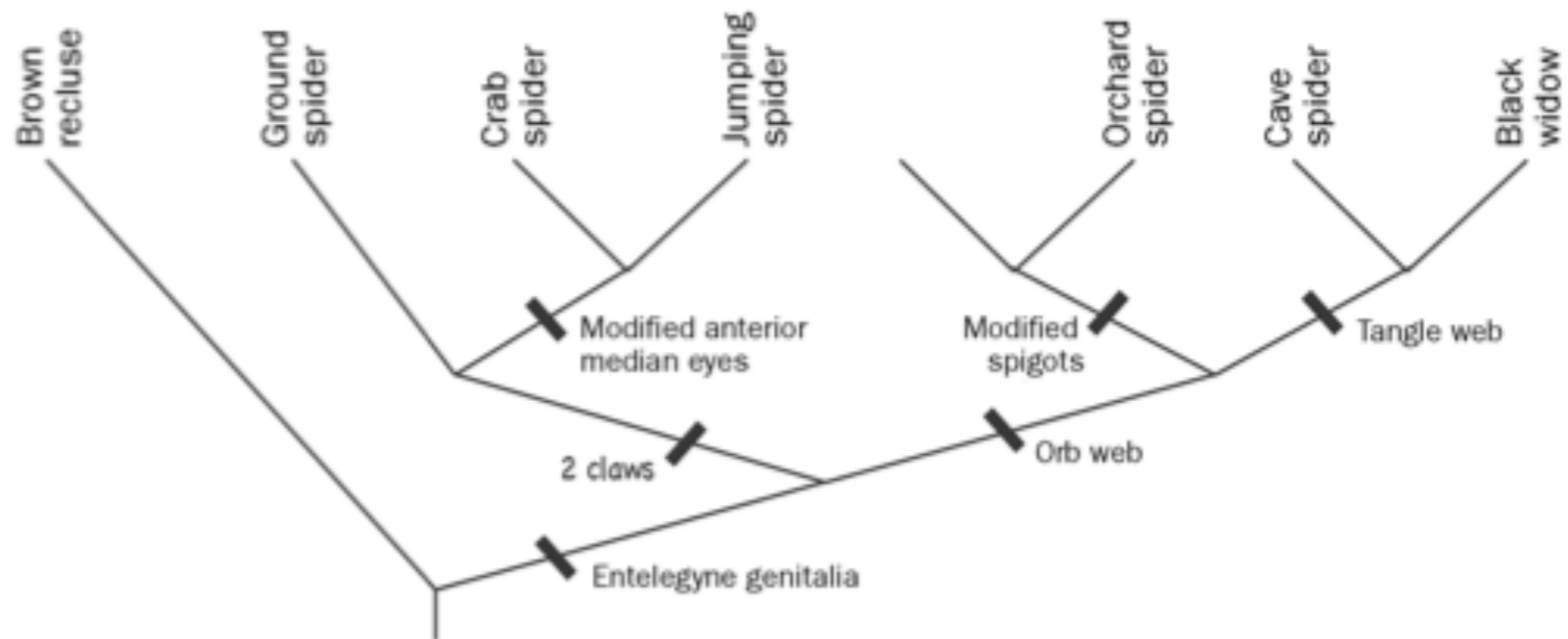
## b. Cladogram in the ladder format





# Different Kinds of Trees

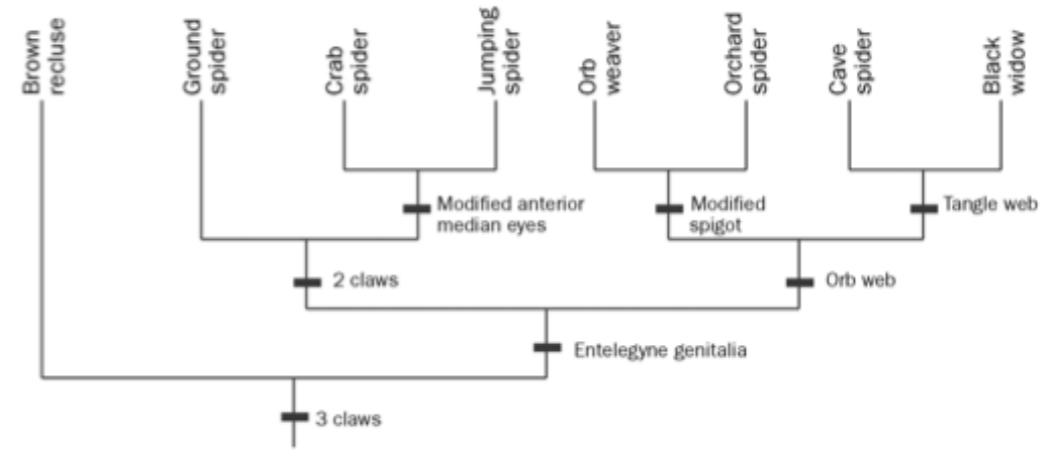
## c. Other cladogram



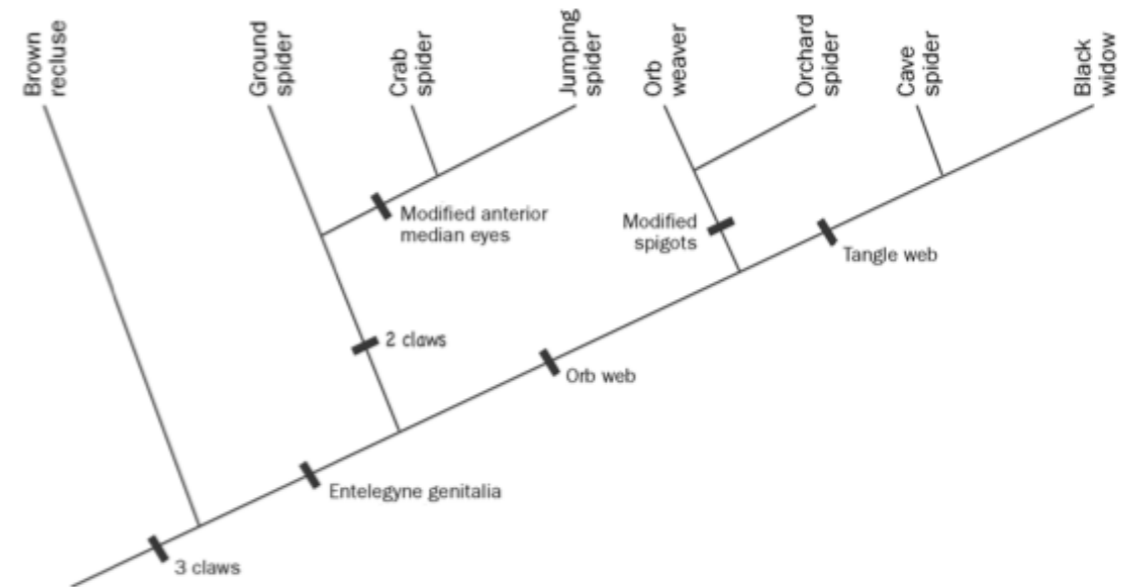
**Figure 1. Three formats for depicting cladograms: (a) tree, (b) ladder, and (c) other.**

# Different Kinds of Trees

a. Cladogram in the tree format



b. Cladogram in the ladder format



c. Other cladogram

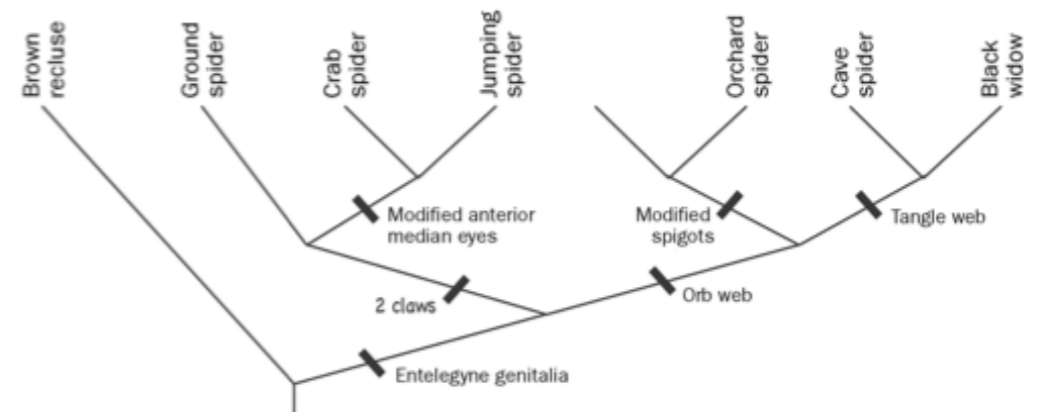
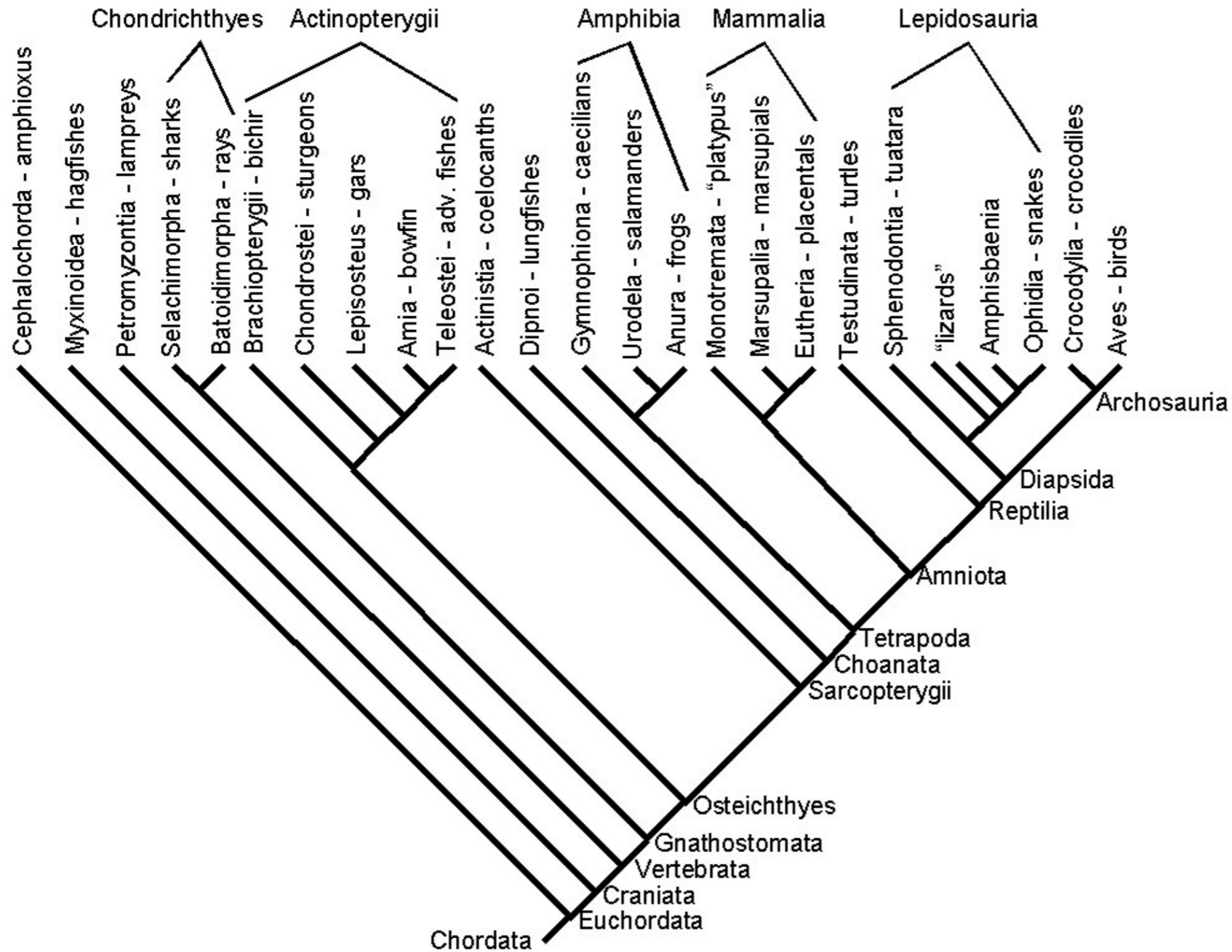


Figure 1. Three formats for depicting cladograms: (a) tree, (b) ladder, and (c) other.

# Different Kinds of Trees



# Different Kinds of Trees

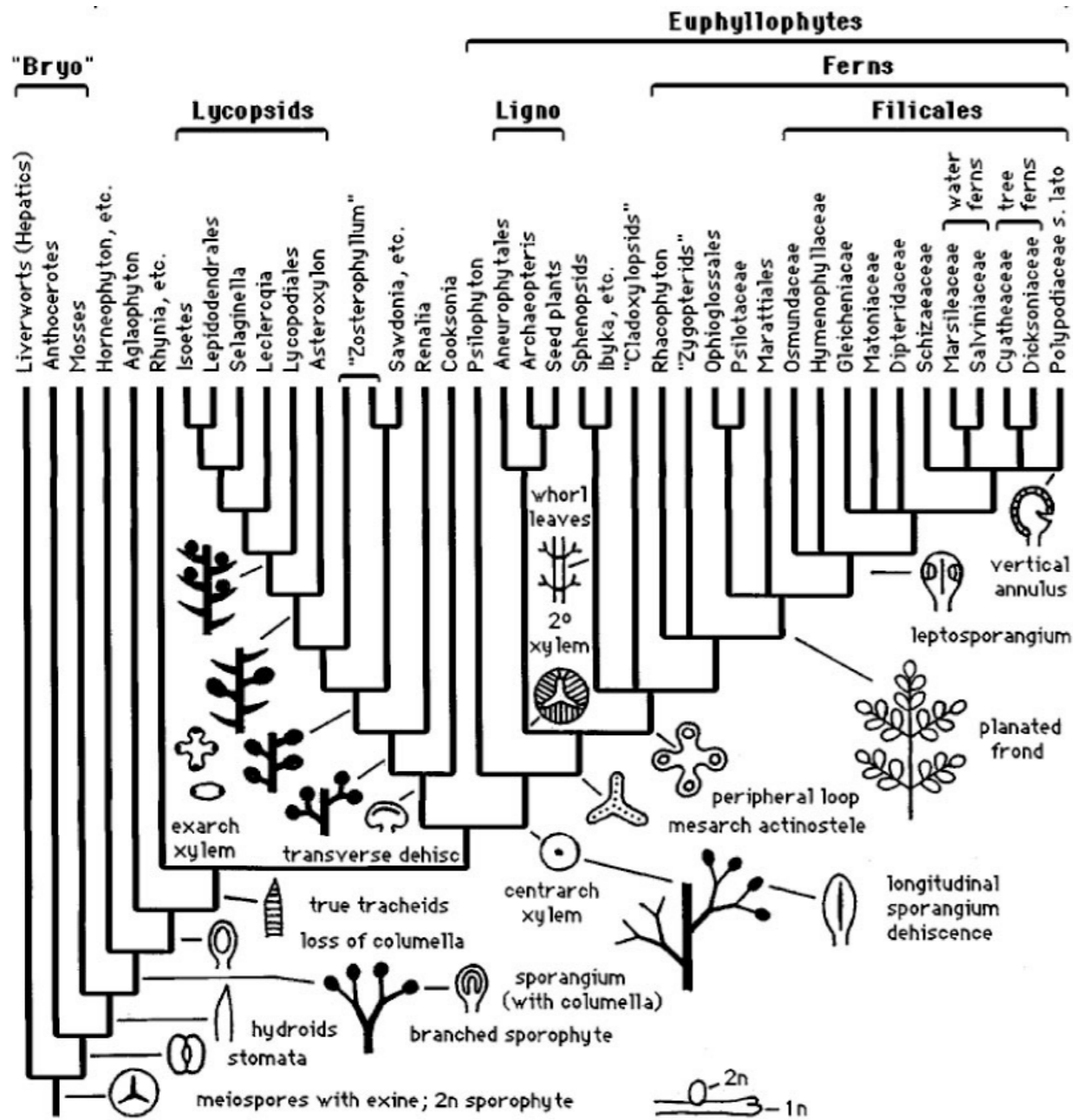
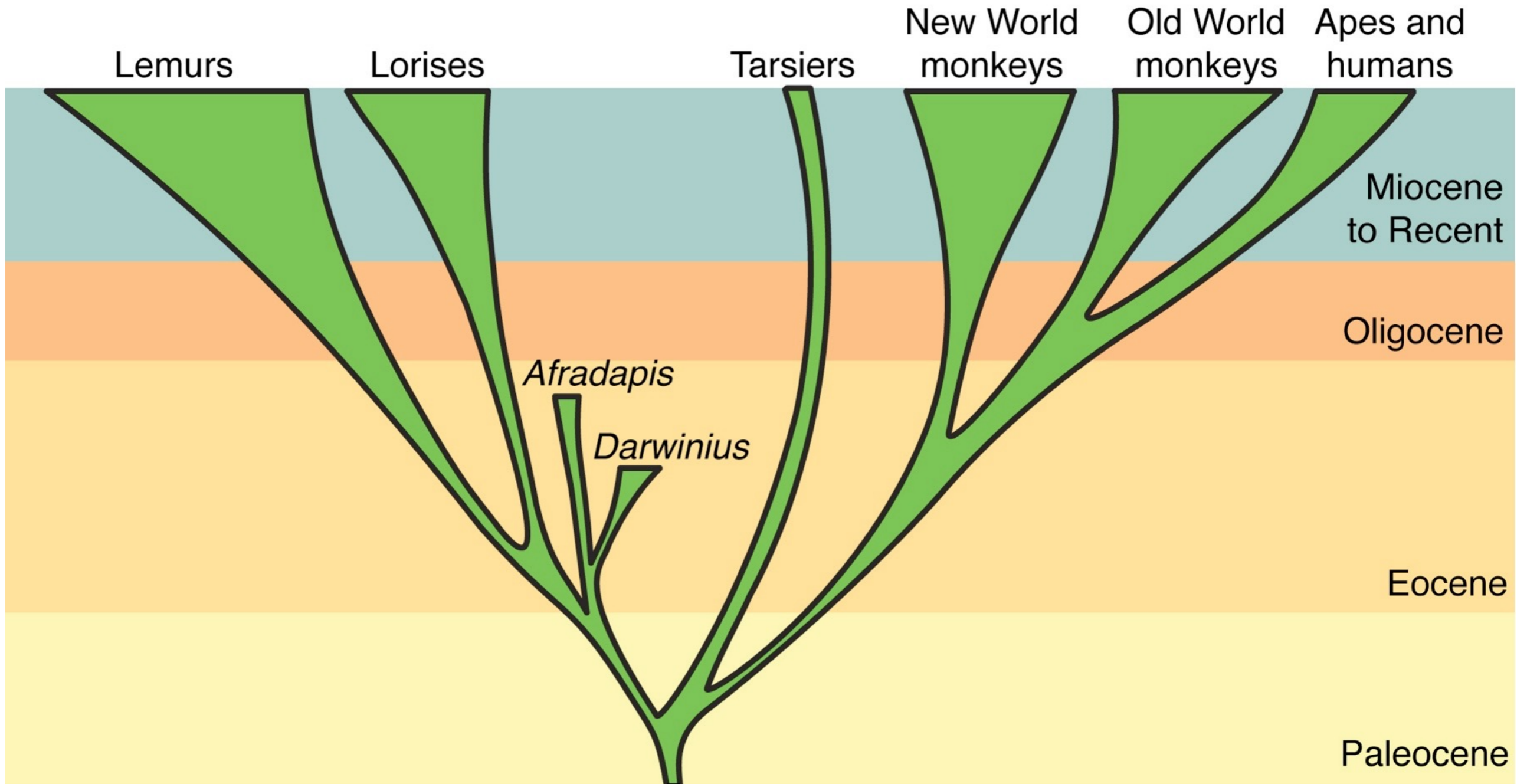
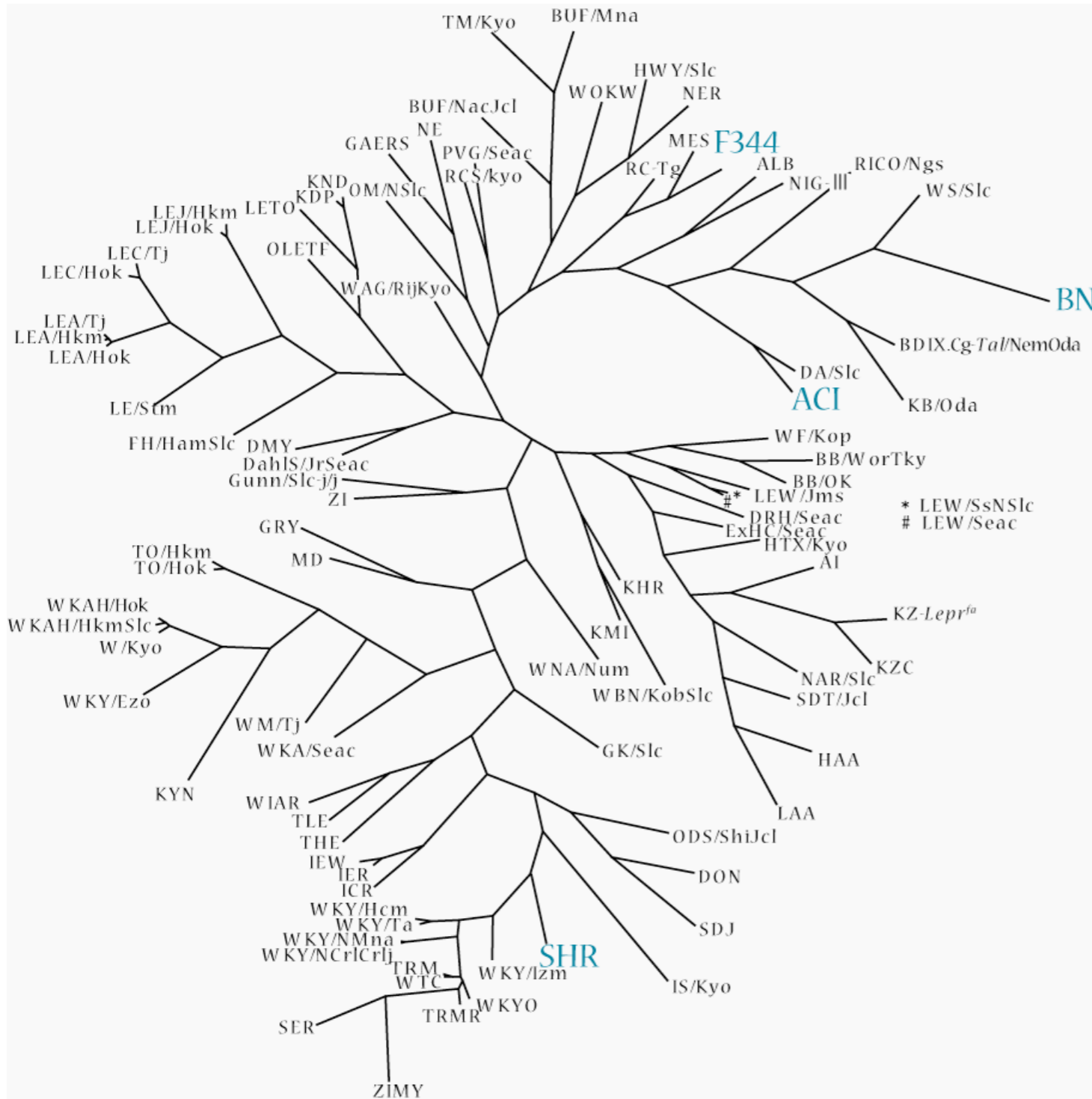


Figure 1 Synopsis of relationships among bryophytes and "lower" vascular plants found in morphological and molecular analyses [based primarily on the work of several others (83, 95, 102)], with sketches illustrating important synapomorphies in vegetative morphology, stem anatomy, and sporangia. "Bryo," bryophytes; Ligno, lignophytes.

# Different Kinds of Trees

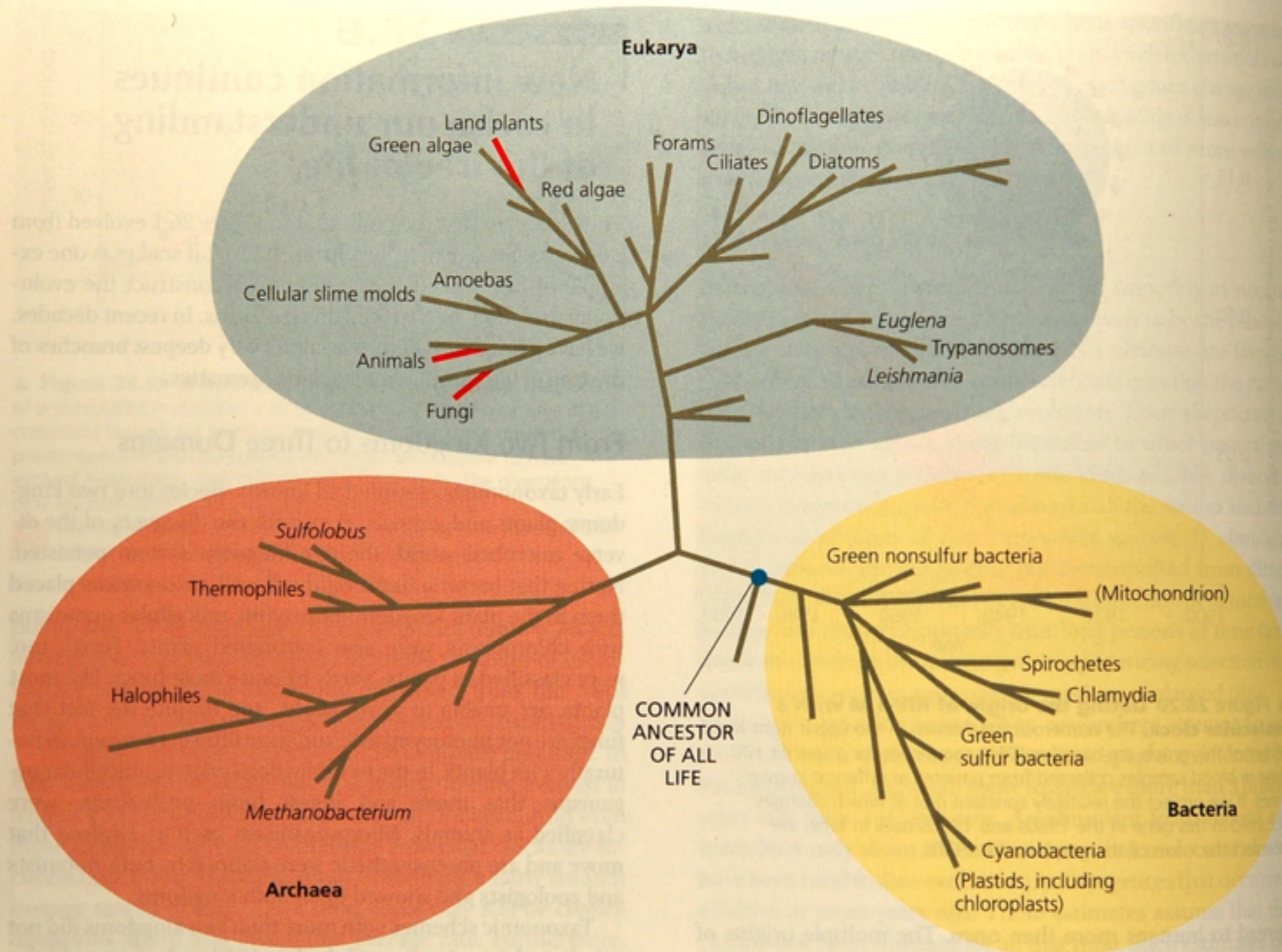


# Different Kinds of Trees





# Three Domains of life



▲ **Figure 26.21 The three domains of life.** Based on rRNA gene sequences, all of life can be grouped into three domains. Branch lengths are proportional to the amount of genetic change in each lineage. (To simplify the figure, only some branches are labeled.) In this diagram, the lineages within Eukarya that are dominated by multicellular organisms (plants, fungi, and animals) are shown

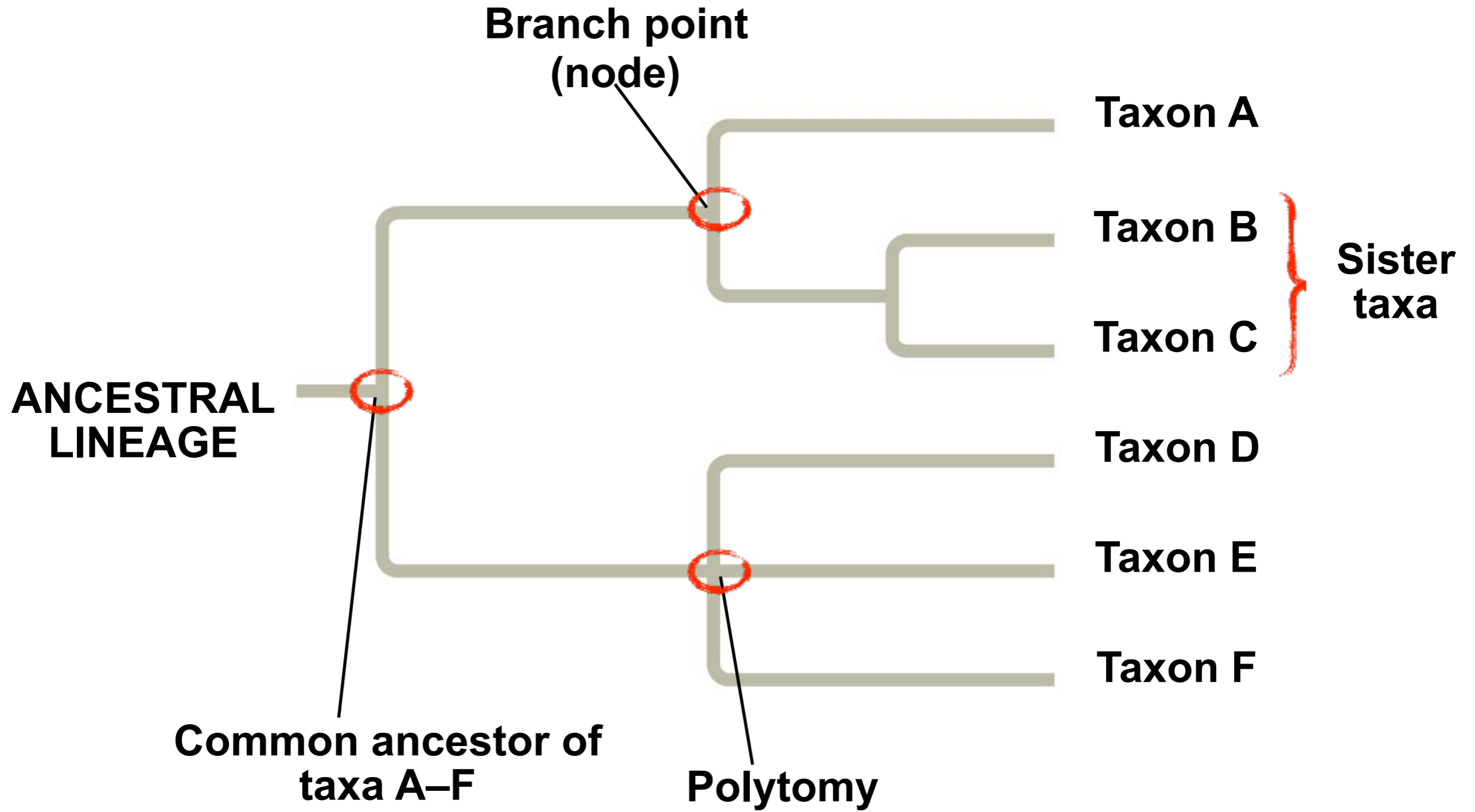


▲ **Figure 26** history of li gene transfer, more such eve

This reco quence comp parts of ribo workings of t mologies be tected—mal evolutionary of life. How ships. For e genes that i

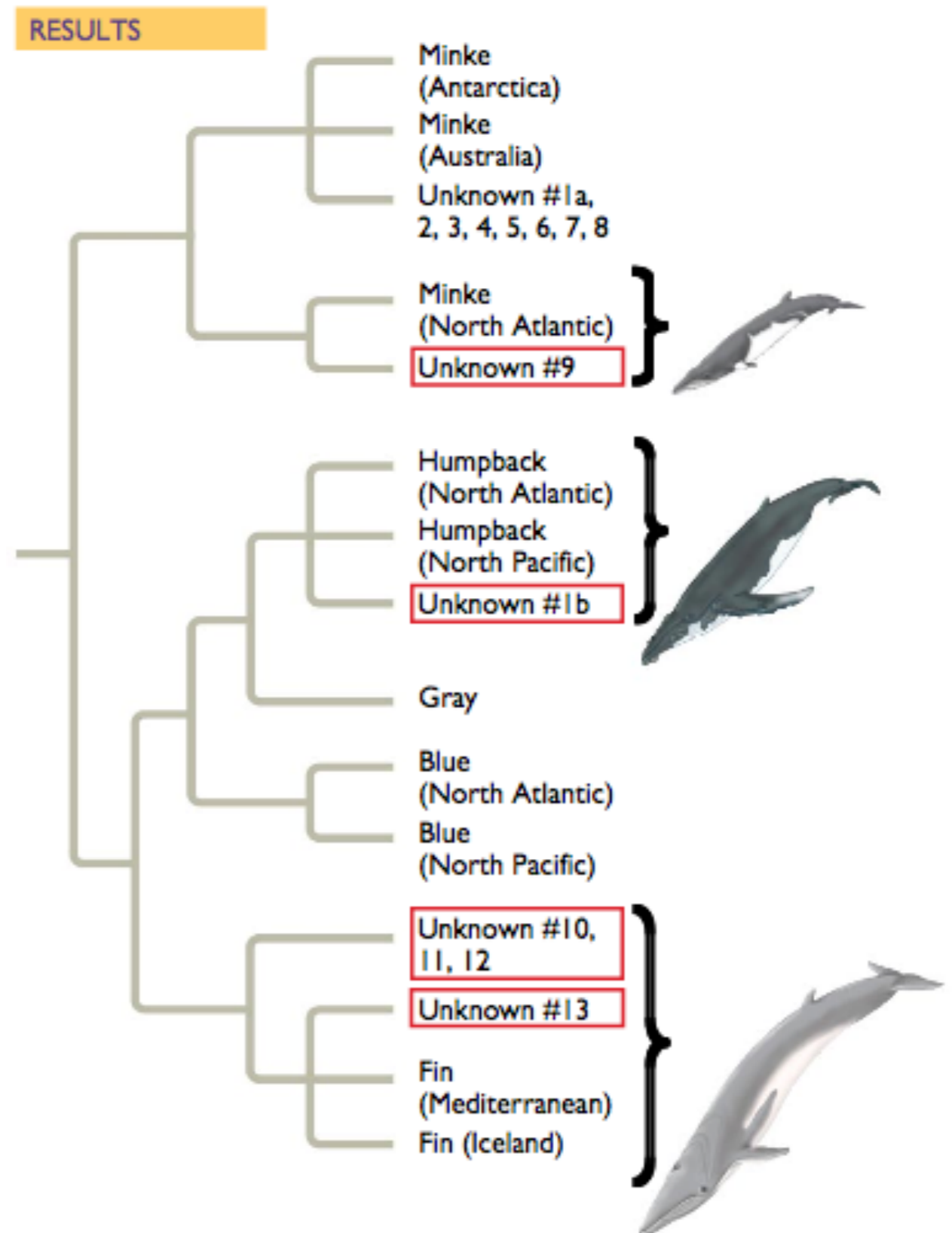


- A rooted tree includes a branch to represent the last common ancestor of all taxa in the tree
- A polytomy is a branch from which more than two groups emerge



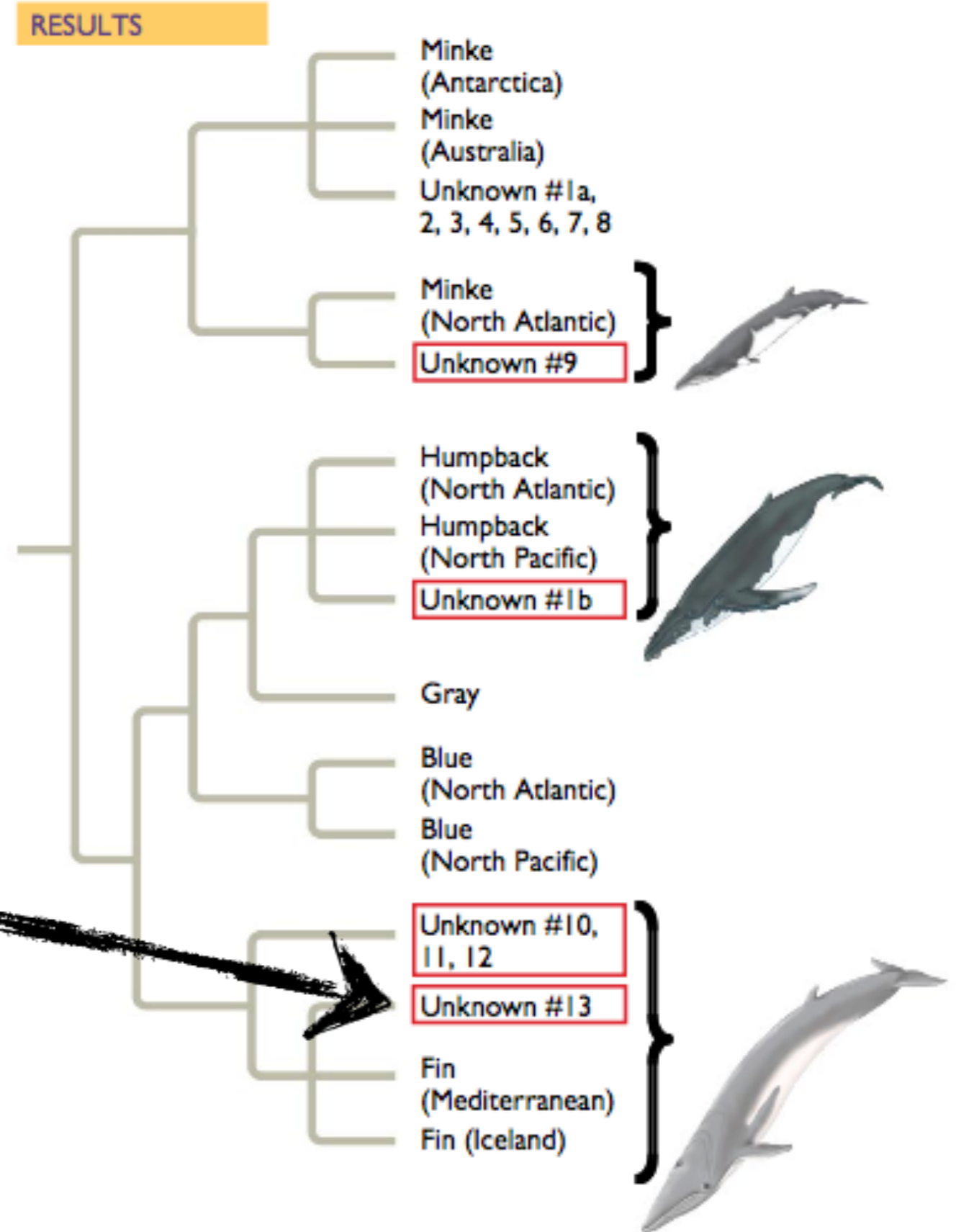
# What We Can and Cannot Learn from Phylogenetic Trees

- Phylogenetic trees do show patterns of descent
- This phylogenetic tree does not indicate when species evolved or how much genetic change occurred in a lineage
- It shouldn't be assumed that a taxon evolved from the taxon next to it



Unknown #13 is most closely related to:

- (a) Gray whale
- (b) Blue whale
- (c) Fin whale
- (d) Minke whale or
- (e) Humpback whale



# Applying Phylogenies

- Phylogeny provides important information about similar characteristics in closely related species
- A phylogeny was used to identify the species of whale from which “whale meat” originated

Where did this fish  
come from?  
.....and what is redfish  
and whitefish anyway?



# Phylogenies are inferred from morphological and molecular data

- To infer phylogenies, systematists gather information about morphologies, genes, and biochemistry of living organisms
- Organisms with similar morphologies or DNA sequences are likely to be more closely related than organisms with different structures or sequences

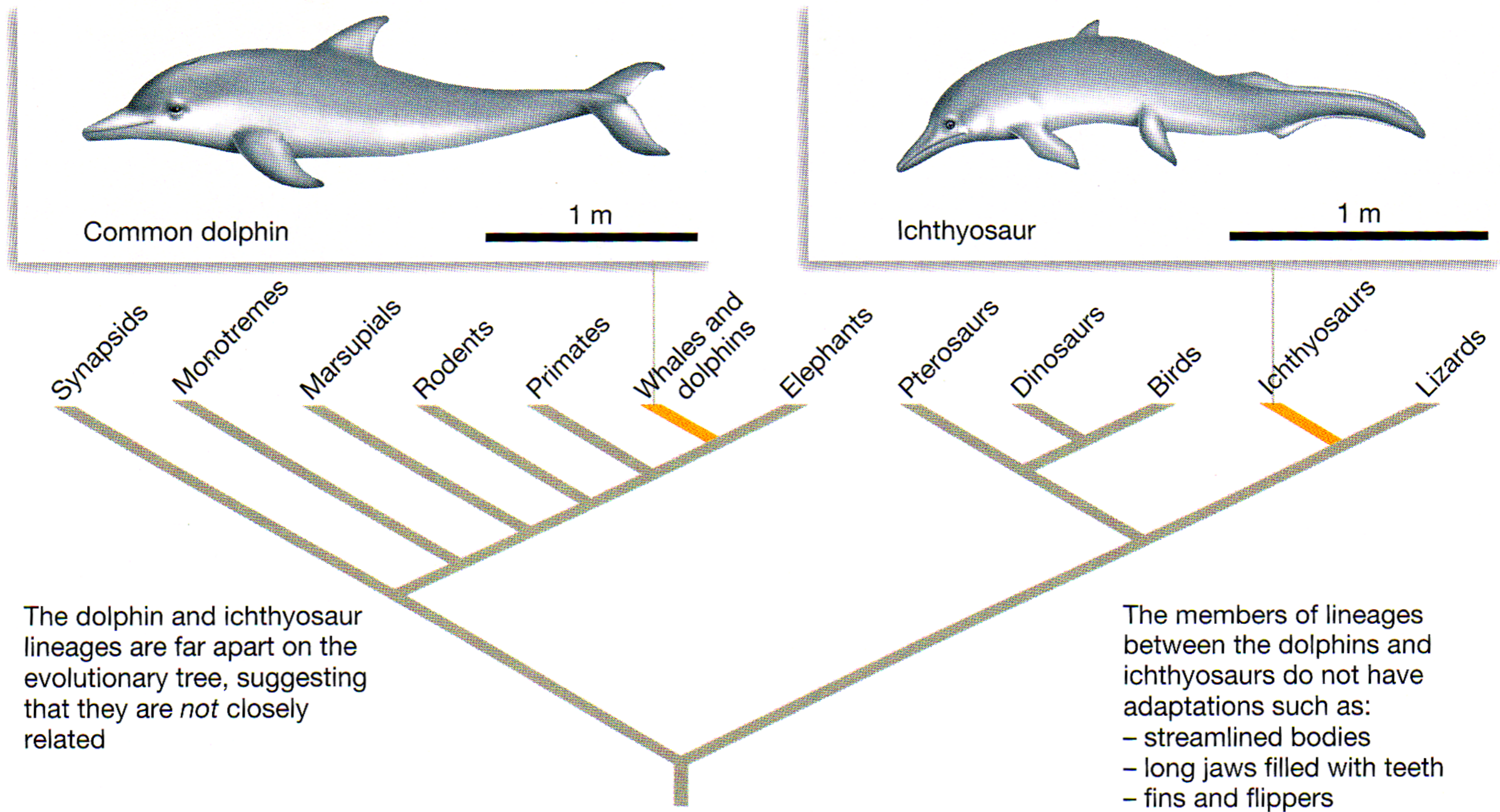
# Identifying Homology and Analogy

- When constructing a phylogeny, systematists need to distinguish whether a similarity is the result of homology or analogy
- Homology is similarity due to shared ancestry
- Analogy is similarity due to convergent evolution



# Convergent evolution occurs when similar environmental pressures and natural selection produce similar (analogous) adaptations in organisms from different evolutionary lineages

(a) Analogous traits: Similarities result from convergent evolution.





- Bat and bird wings are homologous as forelimbs, but analogous as functional wings
- Analogous structures or molecular sequences that evolved independently are also called homoplasies
- Homology can be distinguished from analogy by comparing fossil evidence and the degree of complexity
- The more complex two similar structures are, the more likely it is that they are homologous

# Evaluating Molecular Homologies

Systematists use computer programs and mathematical tools when analyzing comparable DNA segments from different organisms



**Why do systematists need computers to analyze DNA sequences?**

- (a) systematists are lazy**
- (b) there are too many phylogenetic trees to evaluate**
- (c) it is the right thing to do**
- (d) there are too many beetles**

# The number of possible trees can be astronomically large

- for 5 species there are 15 trees possible (fully resolved)
- the number of trees increases explosively as the number of species goes up
- 50 species =  $3 \times 10^{76}$  possible unrooted trees
- 30 million species (the estimated to be alive to day =  $10^{300,000,000}$ )
- no computer today can search through that quantity of trees

$$(2n - 3)!! = \frac{(2n - 3)!}{2^{n-2}(n - 2)!}, \text{ for } n \geq 2$$

total rooted trees and

$$(2n - 5)!! = \frac{(2n - 5)!}{2^{n-3}(n - 3)!}, \text{ for } n \geq 3$$

TABLE 1. THE NUMBERS OF ROOTED TREES WITH *n* LABELLED TIPS AND WITH UNLABELLED INTERIOR NODES. THE LEFT COLUMN COUNTS ALL TREES, THE RIGHT COLUMN ONLY BIFURCATING TREES.

<i>n</i>	All trees	Bifurcating trees
1	1	1
2	1	1
3	4	3
4	26	15
5	236	105
6	2,752	945
7	39,208	10,395
8	660,032	135,135
9	12,818,912	2,027,025
10	282,137,824	34,459,425
11	6,939,897,856	654,729,075
12	188,666,182,784	13,749,310,575
13	5,617,349,020,544	316,234,143,225
14	181,790,703,209,728	7,905,853,580,625
15	6,353,726,042,486,112	213,458,046,676,875
16	238,513,970,965,250,048	6,190,283,353,629,375
17	9,571,020,586,418,569,216	191,898,783,962,510,625
18	408,837,905,660,430,516,224	6,332,659,870,762,850,625
19	18,522,305,410,364,568,764,416	221,643,095,476,699,771,875
20	887,094,711,304,094,583,095,296	8,200,794,532,637,891,559,375
21	44,782,218,857,751,551,087,214,592	319,830,986,772,877,770,815,625
22	2,376,613,641,928,796,906,249,519,104	13,113,070,457,687,988,603,440,625

arise directly from a pre-existing interior node only if that node has only one immediate descendant. It must therefore always be a labelled interior node (although some of the labelled interior nodes will have two immediate descendants). There will be different numbers

of ways of adding species *n* to a tree, depending on how many of these eligible labelled interior nodes there are. Letting *n* = number of labelled nodes, *m* = number of labelled interior nodes with two descendant, and *p* = number of labelled interior nodes with exactly one descendant, we wish to compute *V*(*n*, *m*, *p*), so that by addition over all *m* and *p* we can obtain the total number of bifurcating trees with partially labelled interior nodes (as before, including the case where none is labelled). The recurrence relation is:

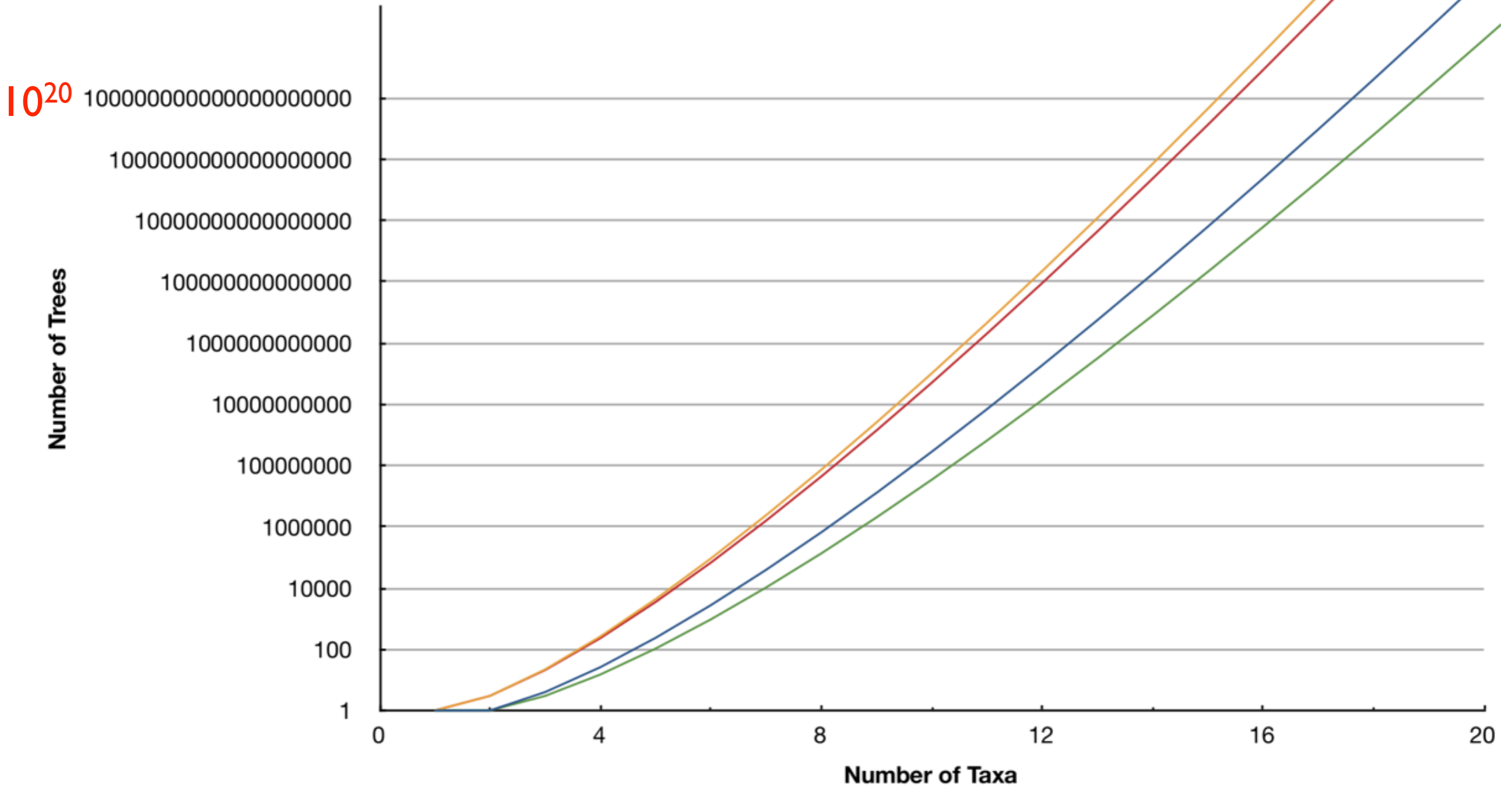
TABLE 2. THE NUMBERS OF ROOTED TREES WITH *n* LABELLED SPECIES, ALLOWING MULTIFURCATIONS AND ALLOWING SOME INTERIOR NODES TO BE LABELLED.

<i>n</i>	Number of trees
1	1
2	3
3	22

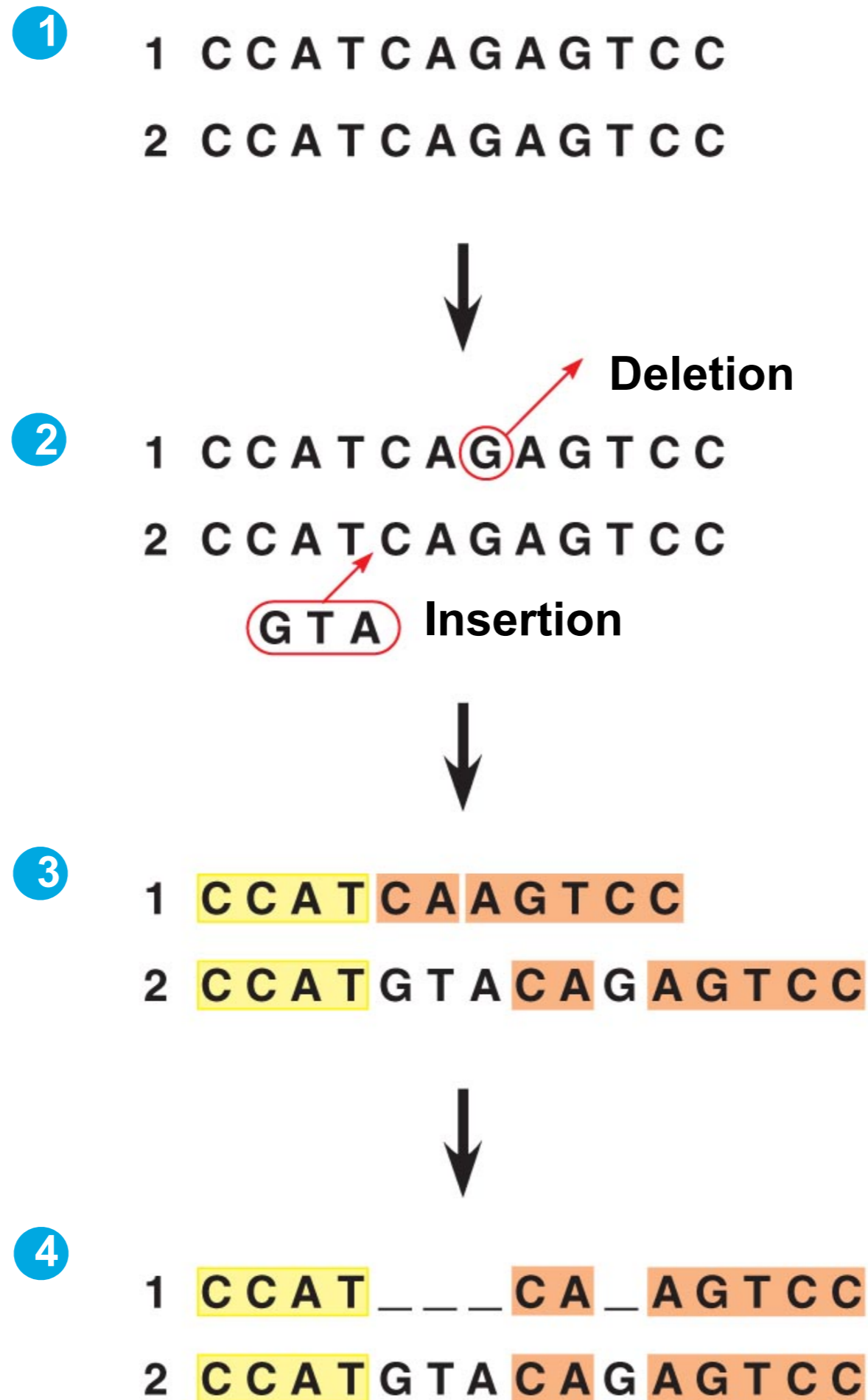
Felsenstein, J. 1978.  
The number of evolutionary trees.  
Systematic Biology.

- 1 taxon = 1 tree
- 2 taxa = 1 tree
- 3 taxa = 4 trees
- 4 taxa = 26 trees....

# How many trees did you say?



- Rooted trees with  $n$  labelled tips, multifurcating, with unlabelled interior nodes
- Rooted trees with  $n$  labelled tips, bifurcating, with unlabelled interior nodes
- Rooted trees with  $n$  labelled tips, multifurcating, allowing some interior nodes to be labelled
- Rooted trees with  $n$  labelled tips, bifurcating, allowing some interior nodes to be labelled



- It is also important to distinguish homology from analogy in molecular similarities
- Mathematical tools help to identify molecular homoplasies, or coincidences

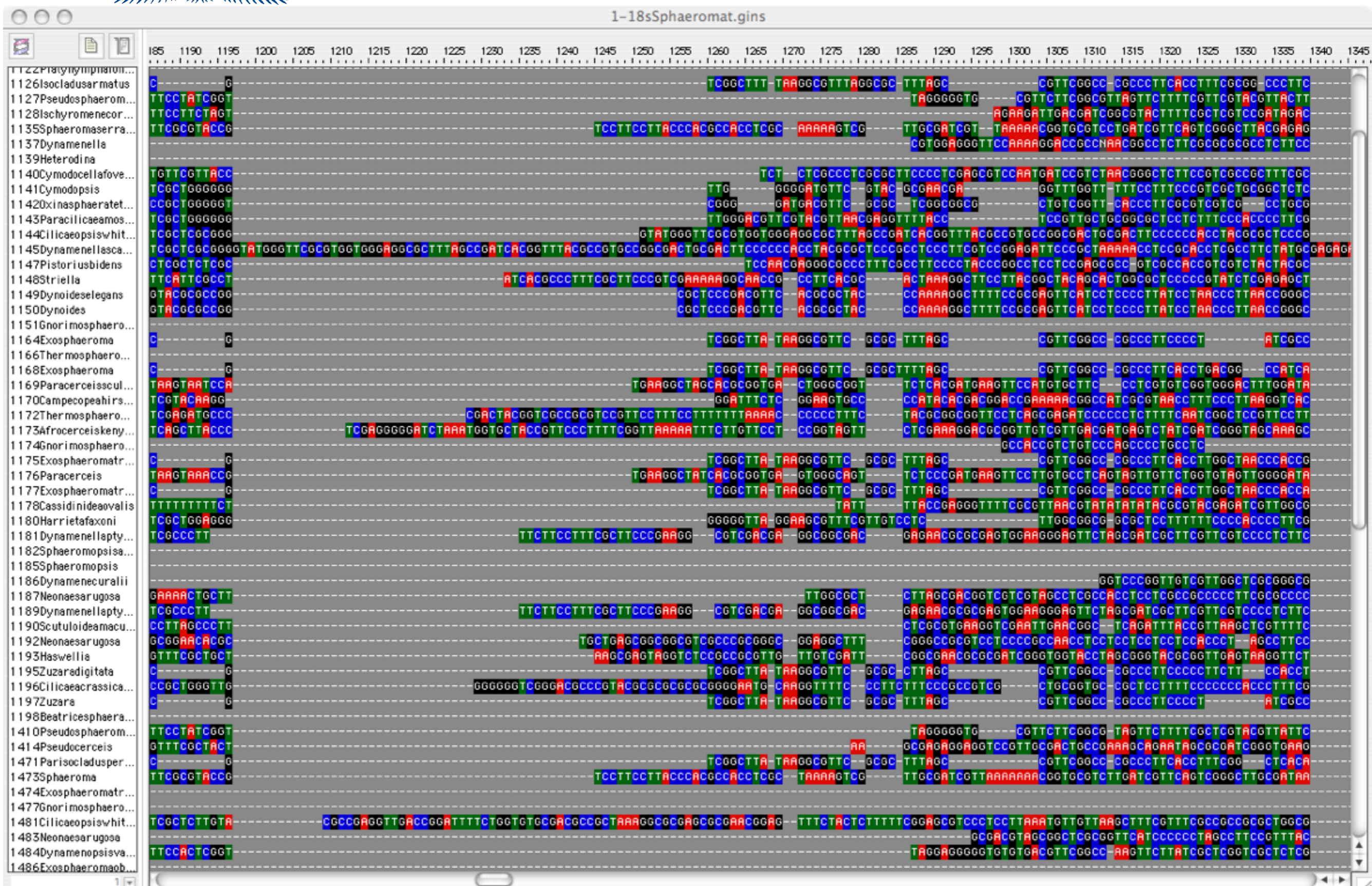
Molecular systematics uses DNA and other molecular data to determine evolutionary relationships







# Hyper-variable regions - 18S rDNA



# Shared characters are used to construct phylogenetic trees

no information



Species 1

A C G G A T A G T C C A C T A G G C A C T A

Species 2

T C A C C G A C A G G T C T T T G A C T A G

Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.



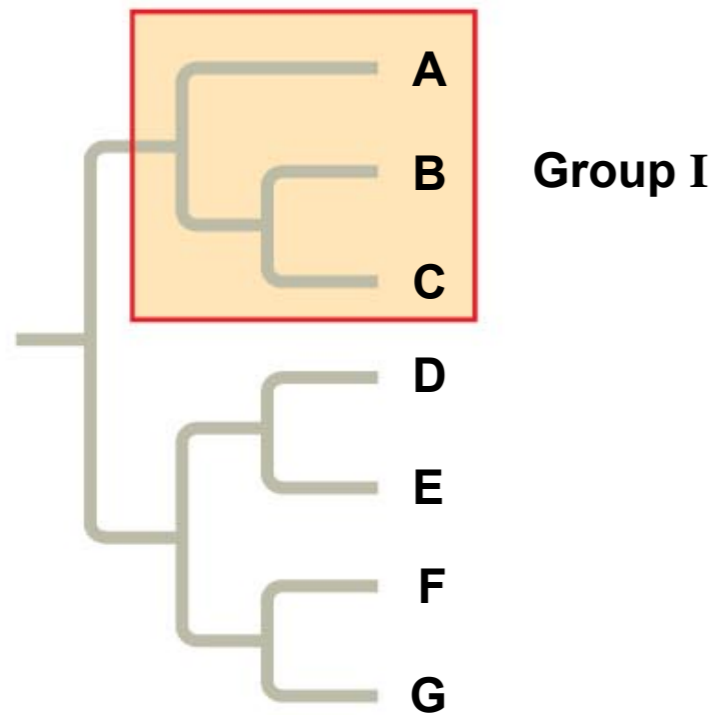
etc.....

Once homologous characters have been identified, they can be used to infer a phylogeny

# Cladistics

Cladistics groups organisms by common descent

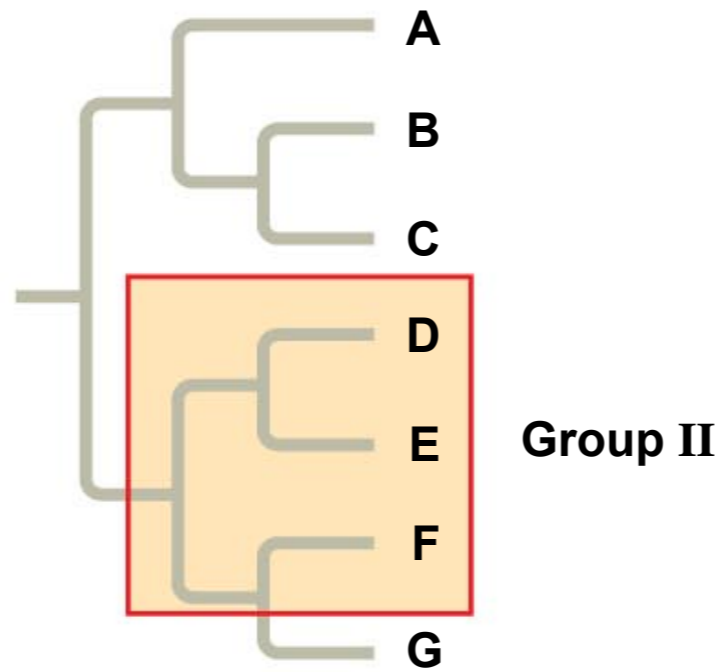
- A clade is a group of species that includes an ancestral species and all its descendants
- Clades can be nested in larger clades, but not all groupings of organisms qualify as clades
- A valid clade is monophyletic, signifying that it consists of the ancestor species and all its descendants



**(a) Monophyletic group (clade)**

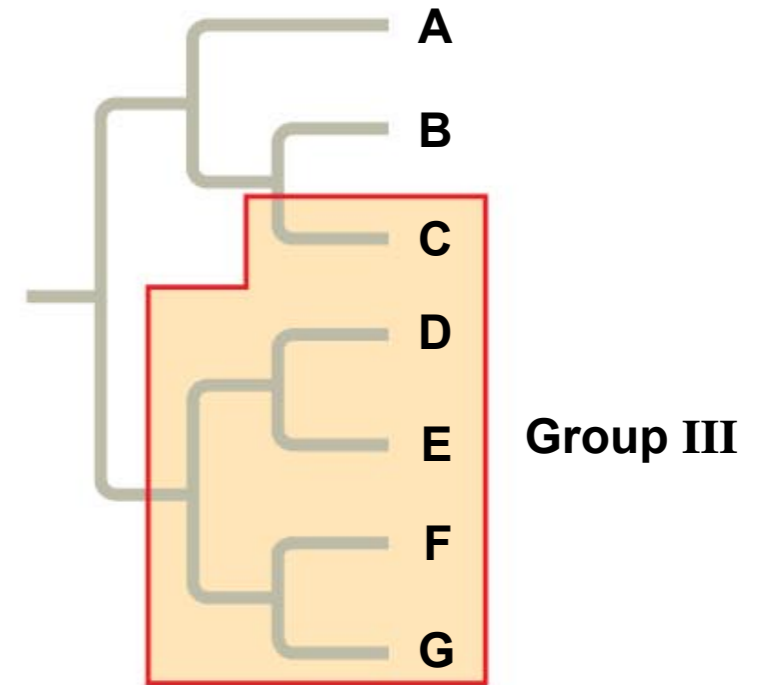
Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

↑  
Ancestor and all  
descendants



**(b) Paraphyletic group**

↑  
A paraphyletic  
grouping consists of  
an ancestral species  
and some, but not  
all, of the  
descendants



**(c) Polyphyletic group**

↑  
A polyphyletic  
grouping consists of  
various species that  
lack a common  
ancestor

# *Shared Ancestral and Shared Derived Characters*

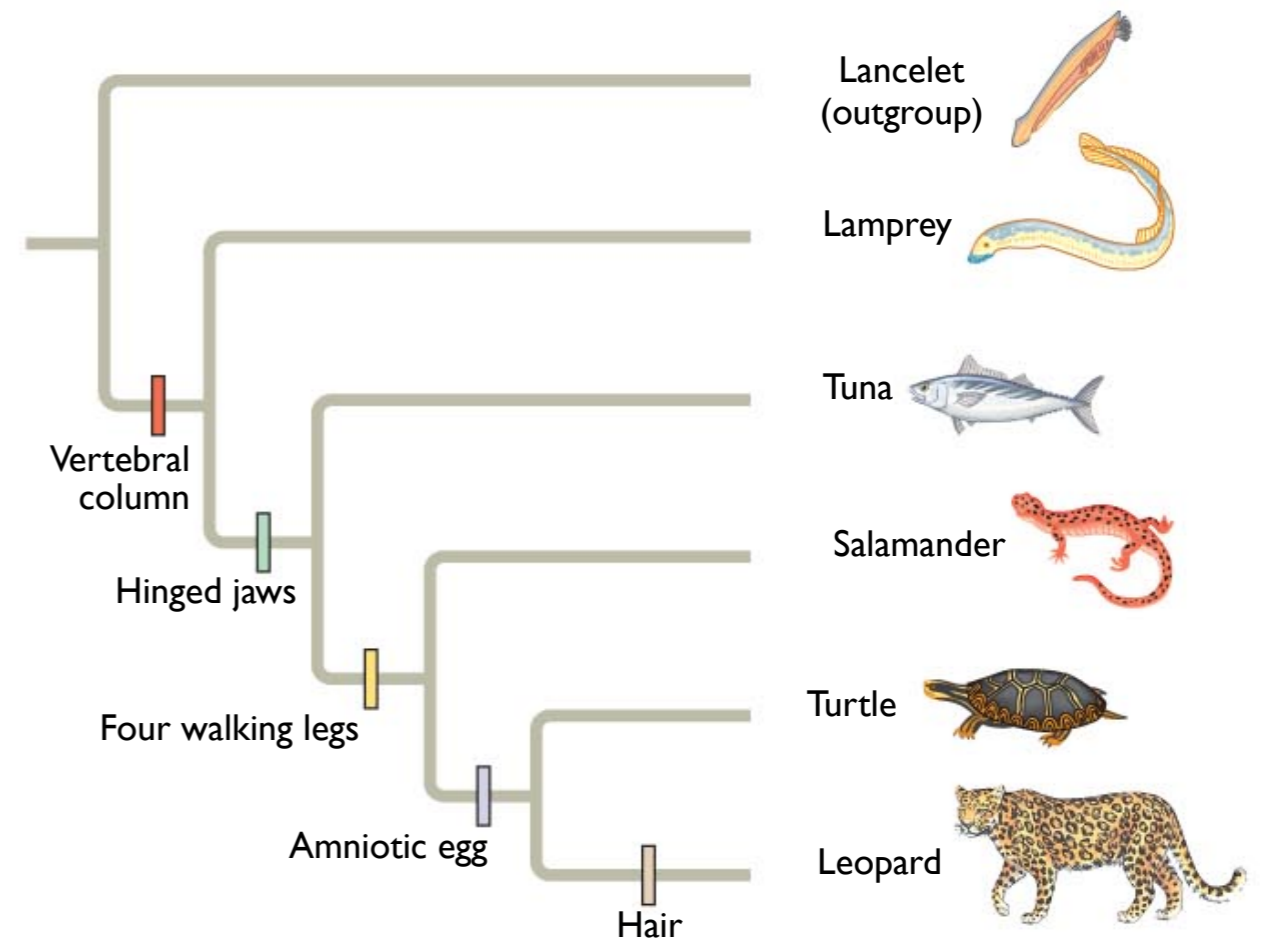
In comparison with its ancestor, an organism has both shared and different characteristics

- A shared ancestral character is a character that originated in an ancestor of the taxon
- A shared derived character is an evolutionary novelty unique to a particular clade
- A character can be both ancestral and derived, depending on the context

# Inferring Phylogenies Using Derived Characters

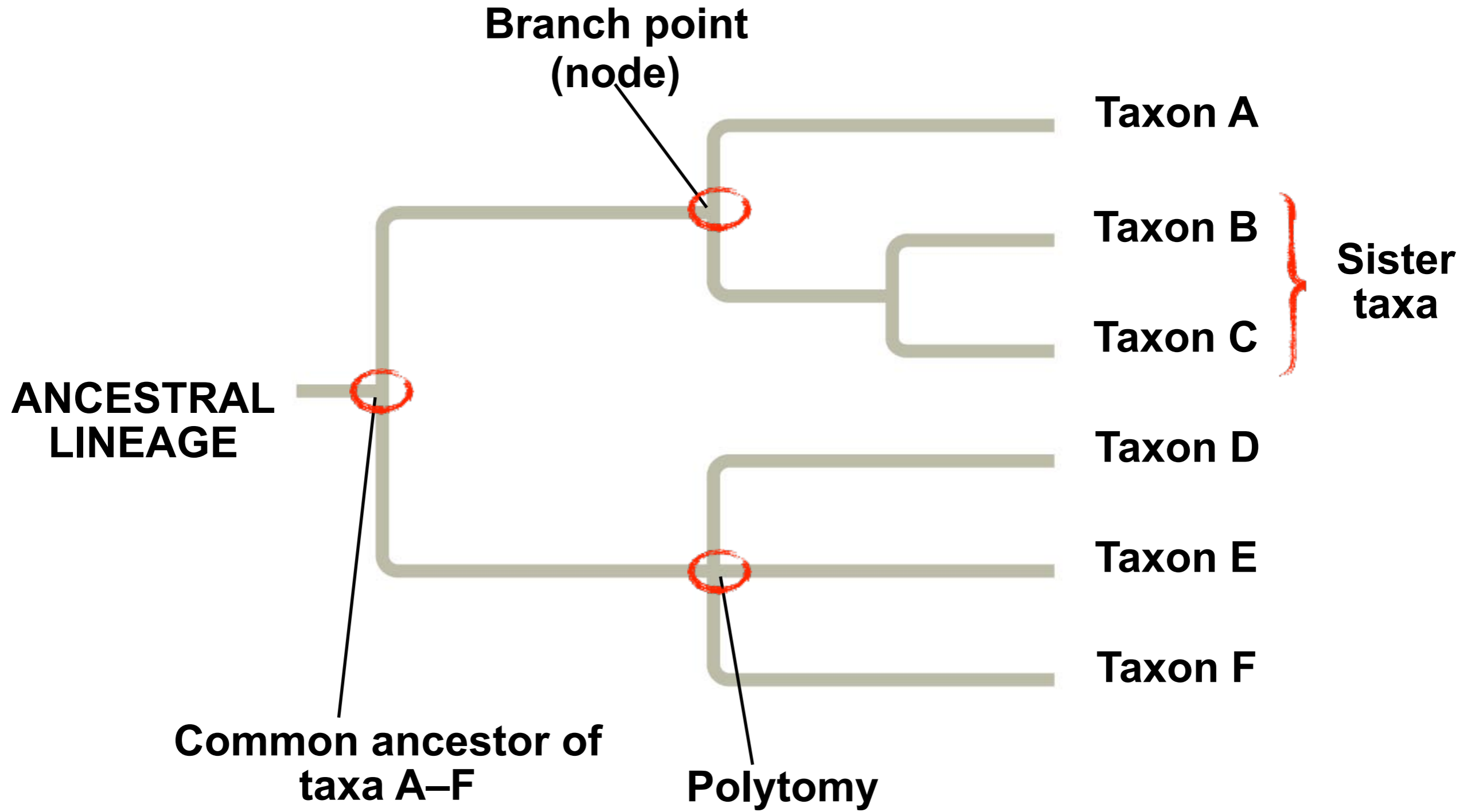
When inferring evolutionary relationships, it is useful to know in which clade a shared derived character first appeared

CHARACTERS	TAXA					
	Lancelet (outgroup)	Lamprey	Tuna	Salamander	Turtle	Leopard
Vertebral column (backbone)	0	1	1	1	1	1
Hinged jaws	0	0	1	1	1	1
Four walking legs	0	0	0	1	1	1
Amniotic (shelled) egg	0	0	0	0	1	1
Hair	0	0	0	0	0	1



(a) Character table

(b) Phylogenetic tree





# Phylogenetic Trees with Proportional Branch Lengths

In some trees, the length of a branch can reflect the number of genetic changes that have taken place in a particular DNA sequence in that lineage

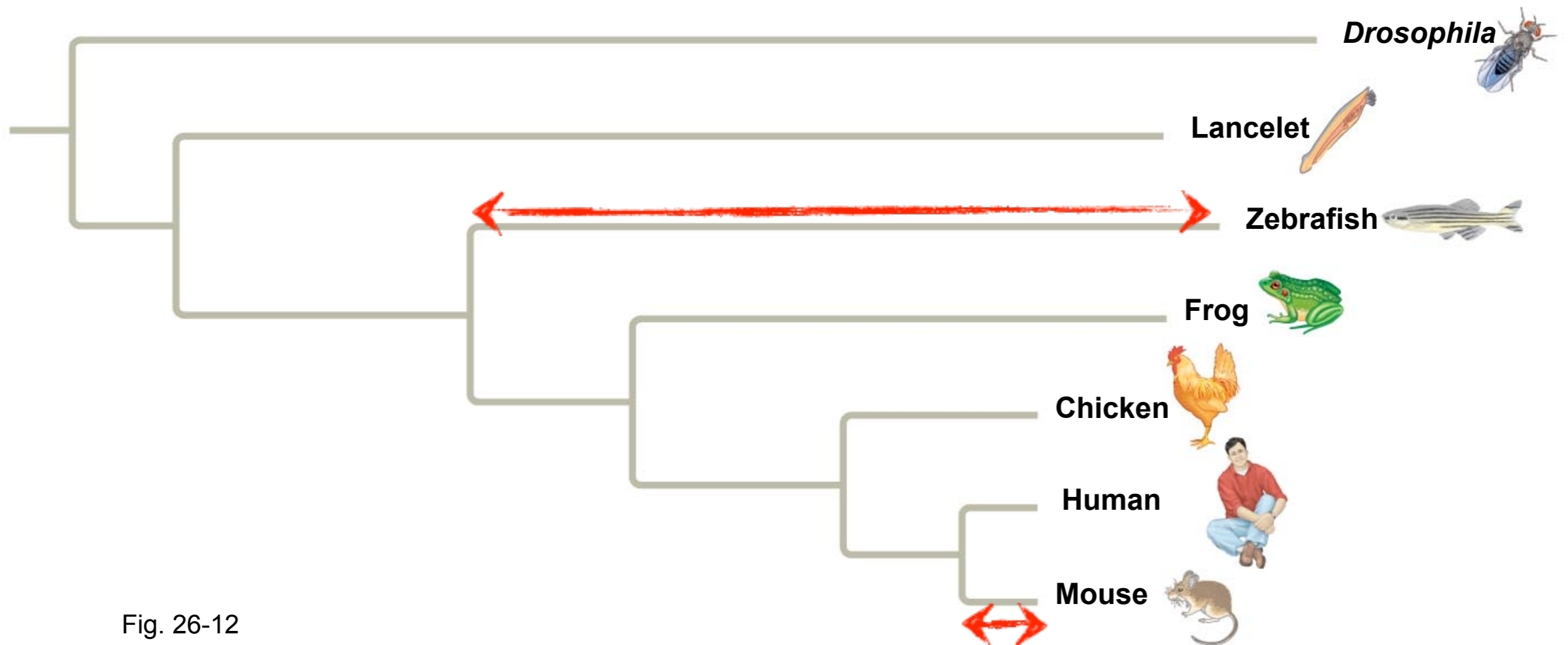
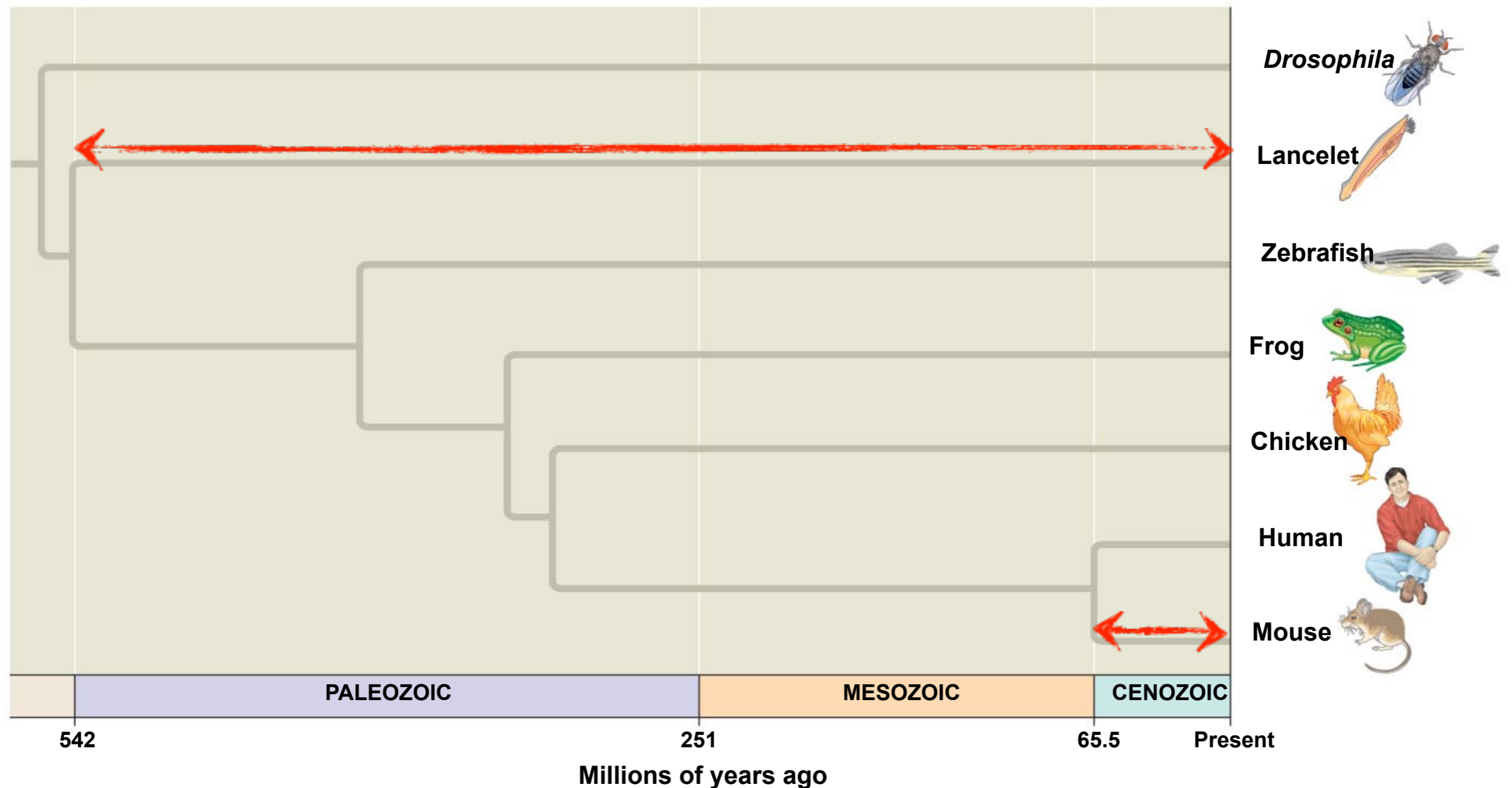


Fig. 26-12

In other trees, branch length can represent chronological time, and branching points can be determined from the fossil record



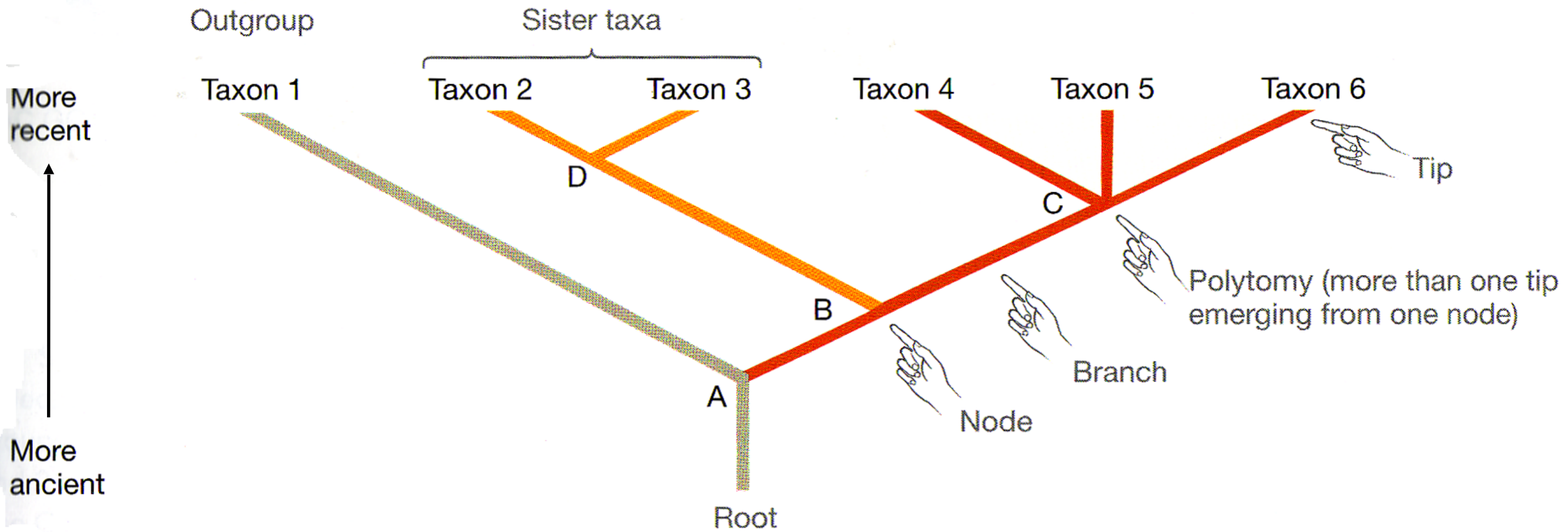
# Maximum Parsimony and Maximum Likelihood

- Systematists can never be sure of finding the best tree in a large data set
- They narrow possibilities by applying the principles of maximum parsimony, maximum likelihood, Bayesian inference, and soon simultaneous analyses

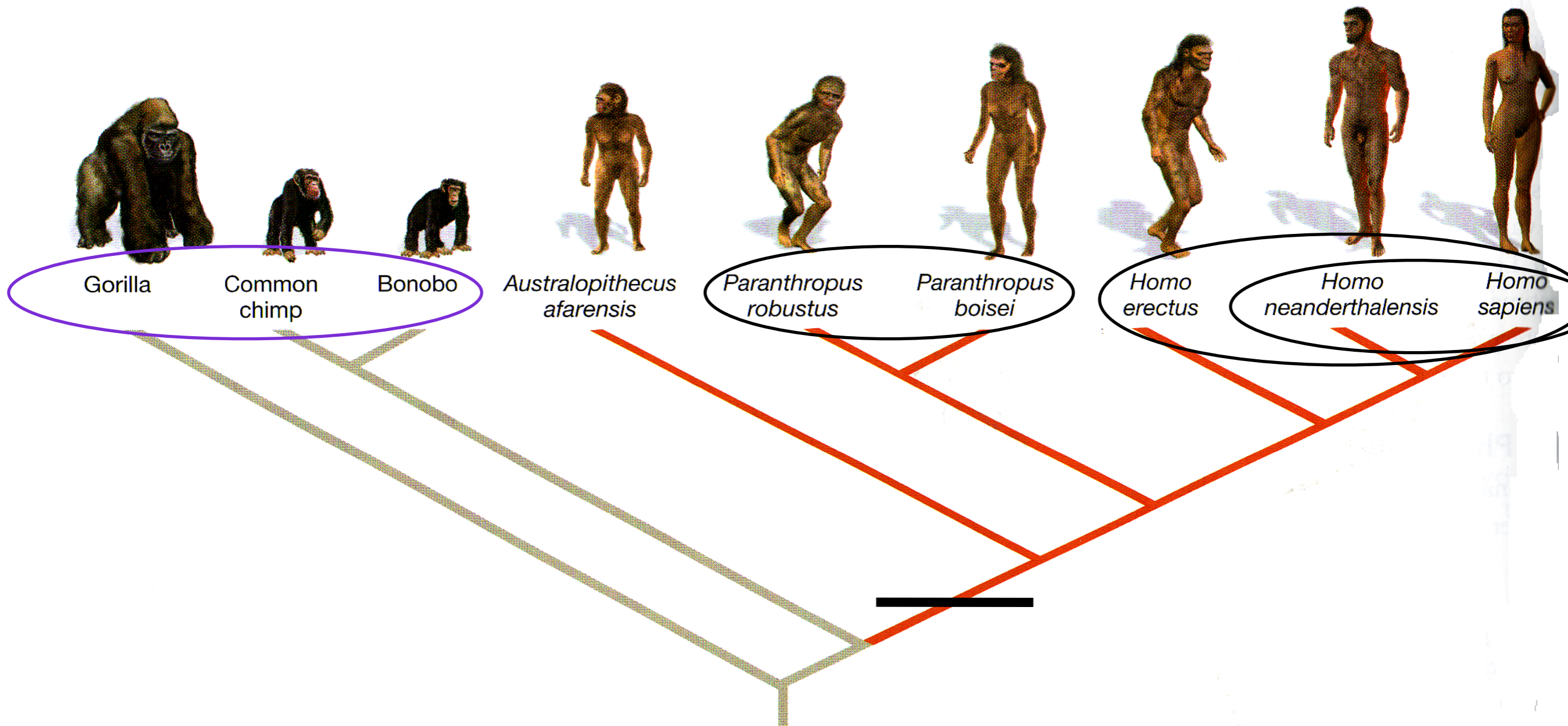
**Maximum parsimony** assumes that the tree that requires the fewest evolutionary events (appearances of shared derived characters) is the most likely

The principle of **maximum likelihood** states that, given certain rules about how DNA changes over time, a tree can be found that reflects the most likely sequence of evolutionary events

# Reviewing Phylogenetic Trees

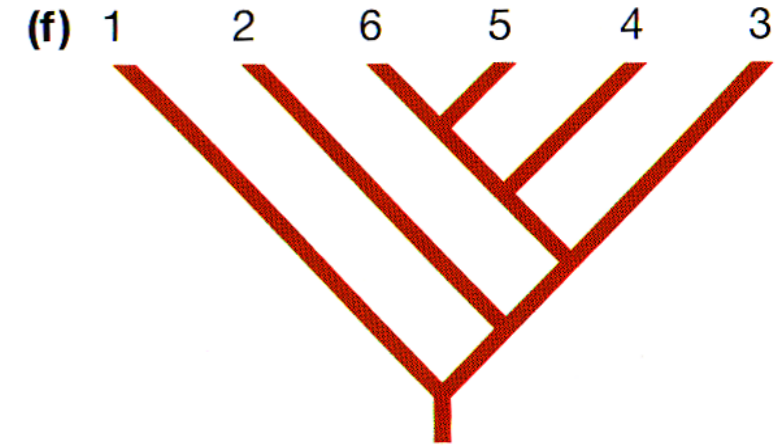
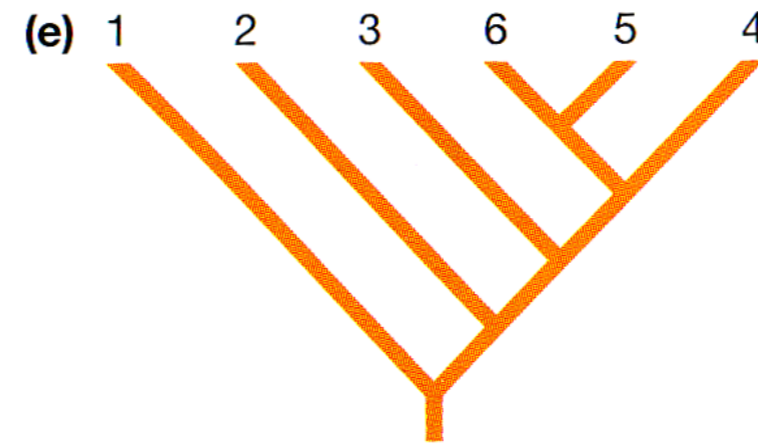
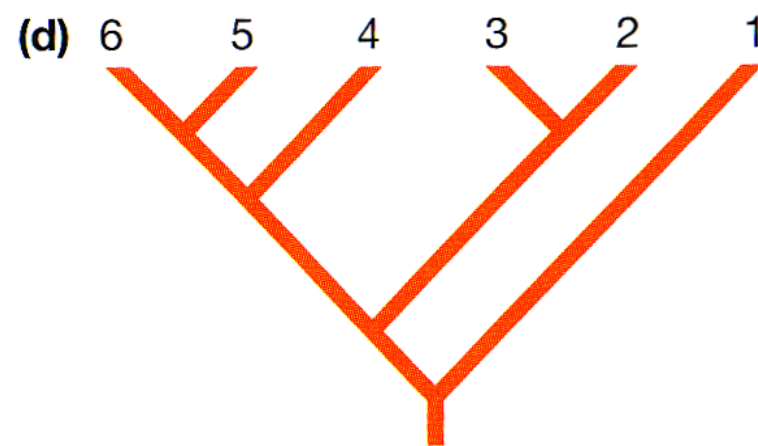
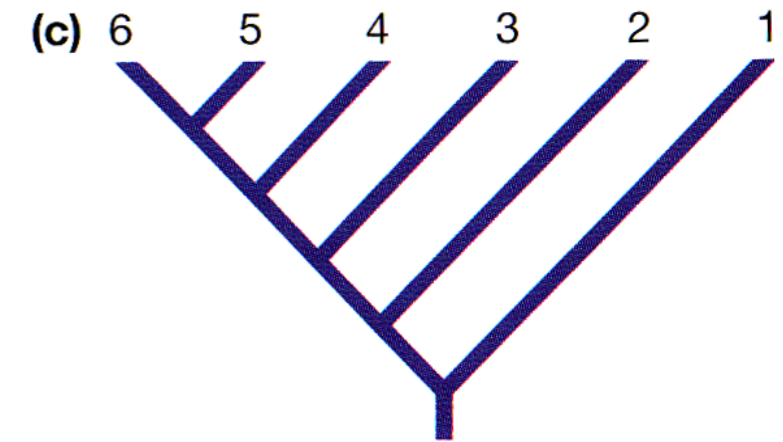
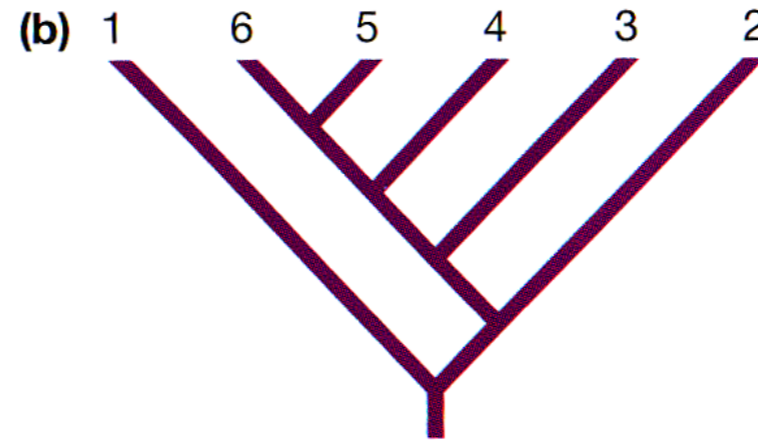
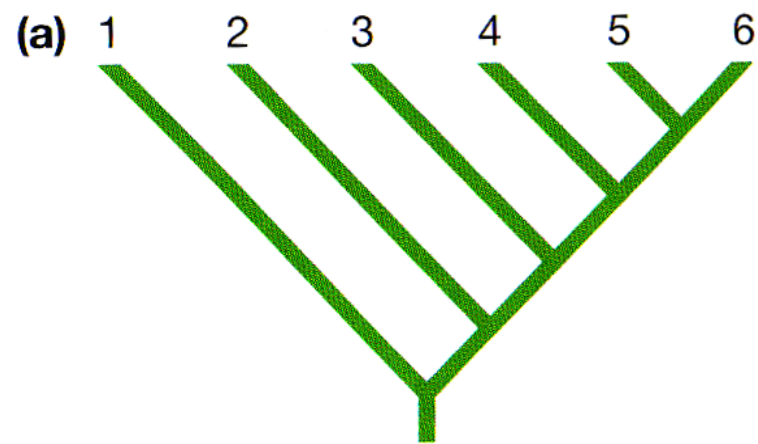


# Example of Phylogenetic Tree



A phylogenetic tree showing the relationships among some of the great apes. Chimps and gorillas walk on all four legs, but all other species on this tree walk on two legs and are considered hominins. **EXERCISE** Add a mark on the phylogeny, and label it "origin of walking on two legs." Circle and label a pair of sister species. Circle and label the monophyletic group called hominins. Label one of the outgroups to the hominins.

# Alternative ways of drawing the same tree

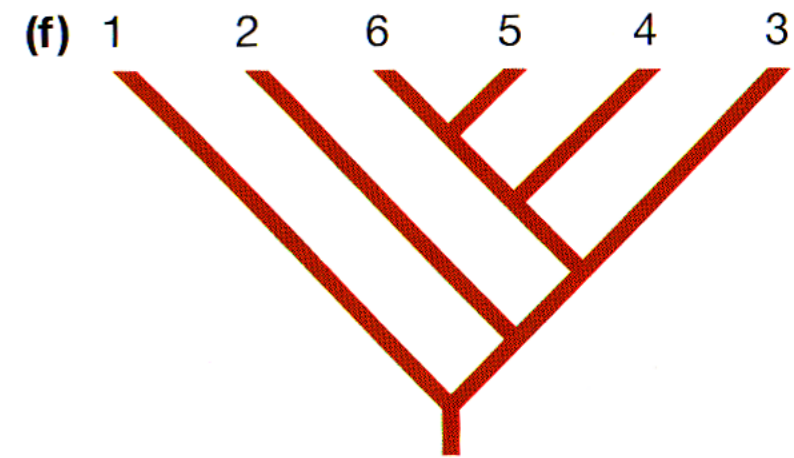
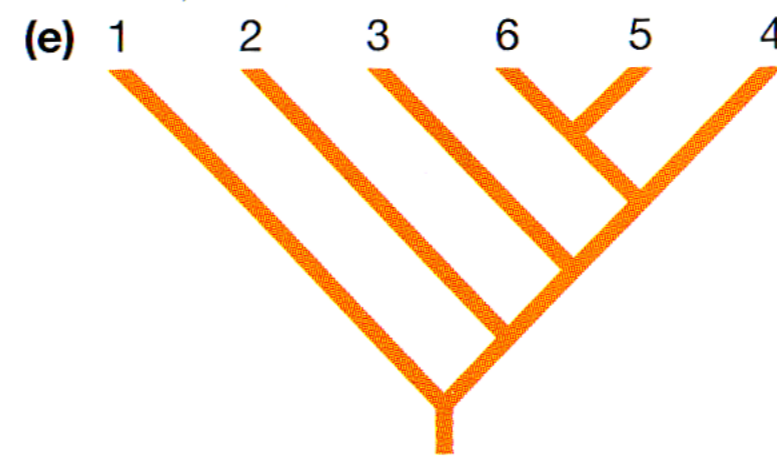
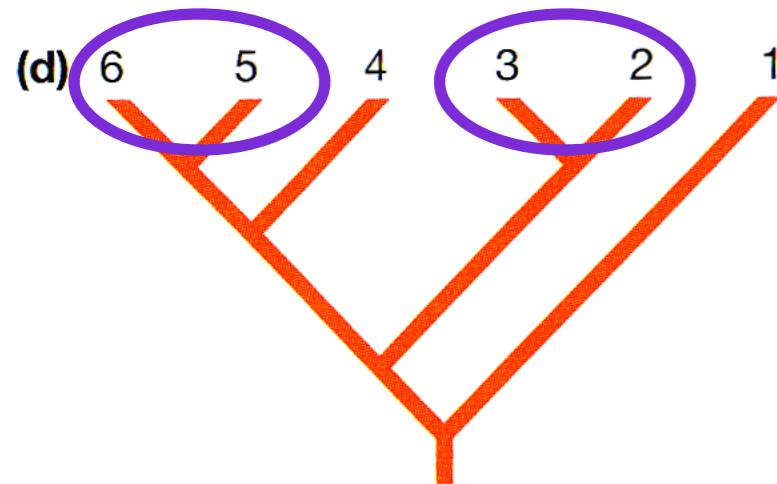
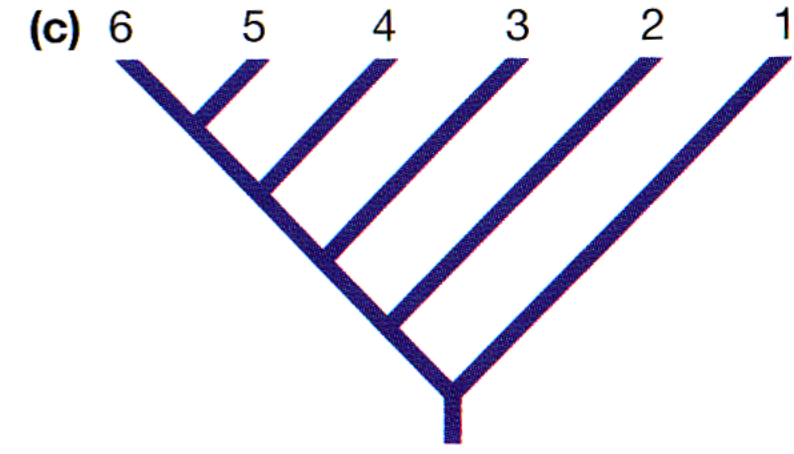
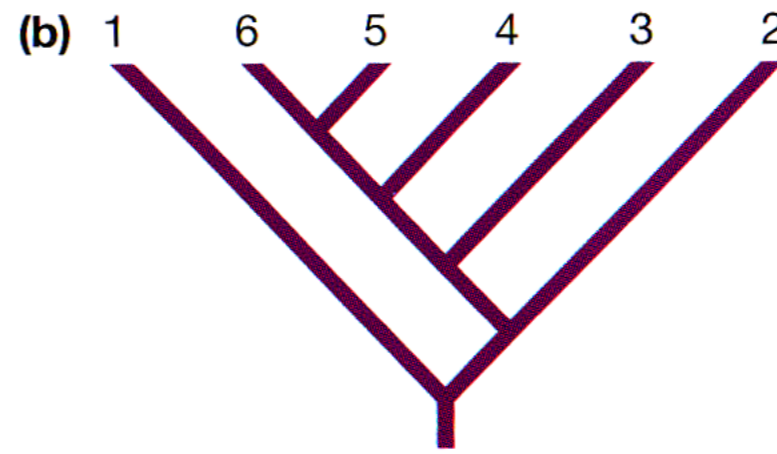
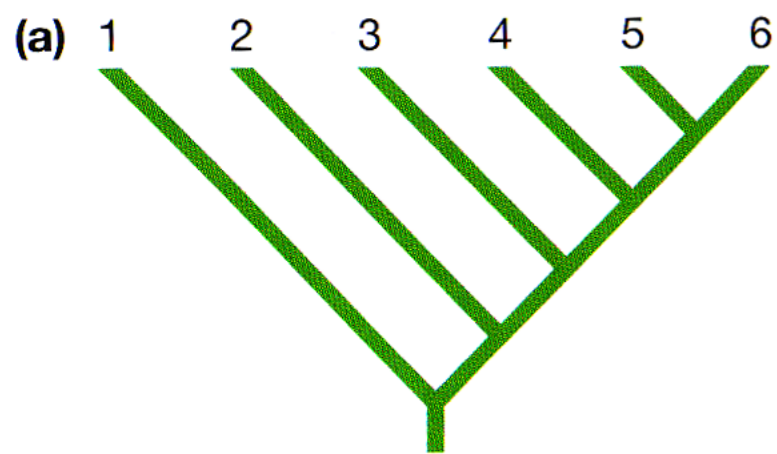


**QUESTION** Five of these six trees describe exactly the same relationships among taxa 1 through 6. Identify the tree that is different from the other five.

Five of these six trees describe  
**EXACTLY** the same relationship  
among taxa 1 through 6.  
Which tree is different from the  
other five?



# Alternative ways of drawing the same tree

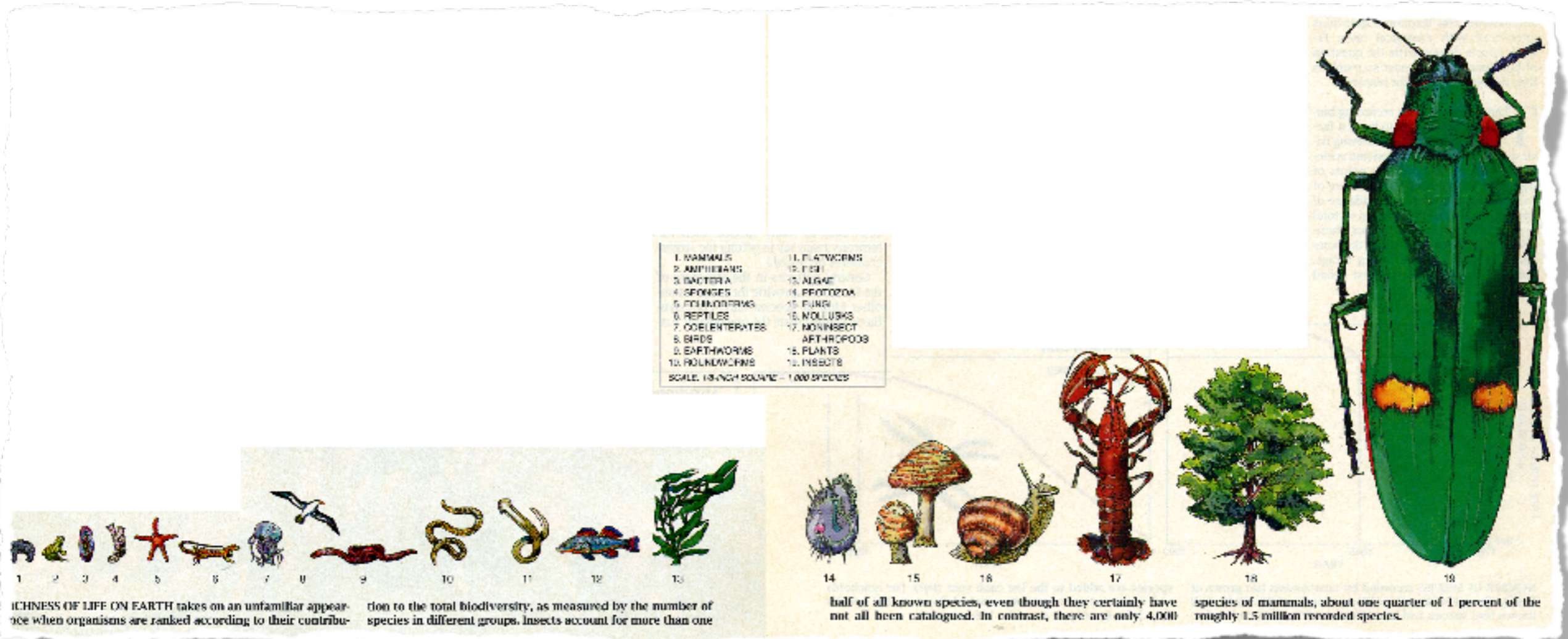


**QUESTION** Five of these six trees describe exactly the same relationships among taxa 1 through 6. Identify the tree that is different from the other five.

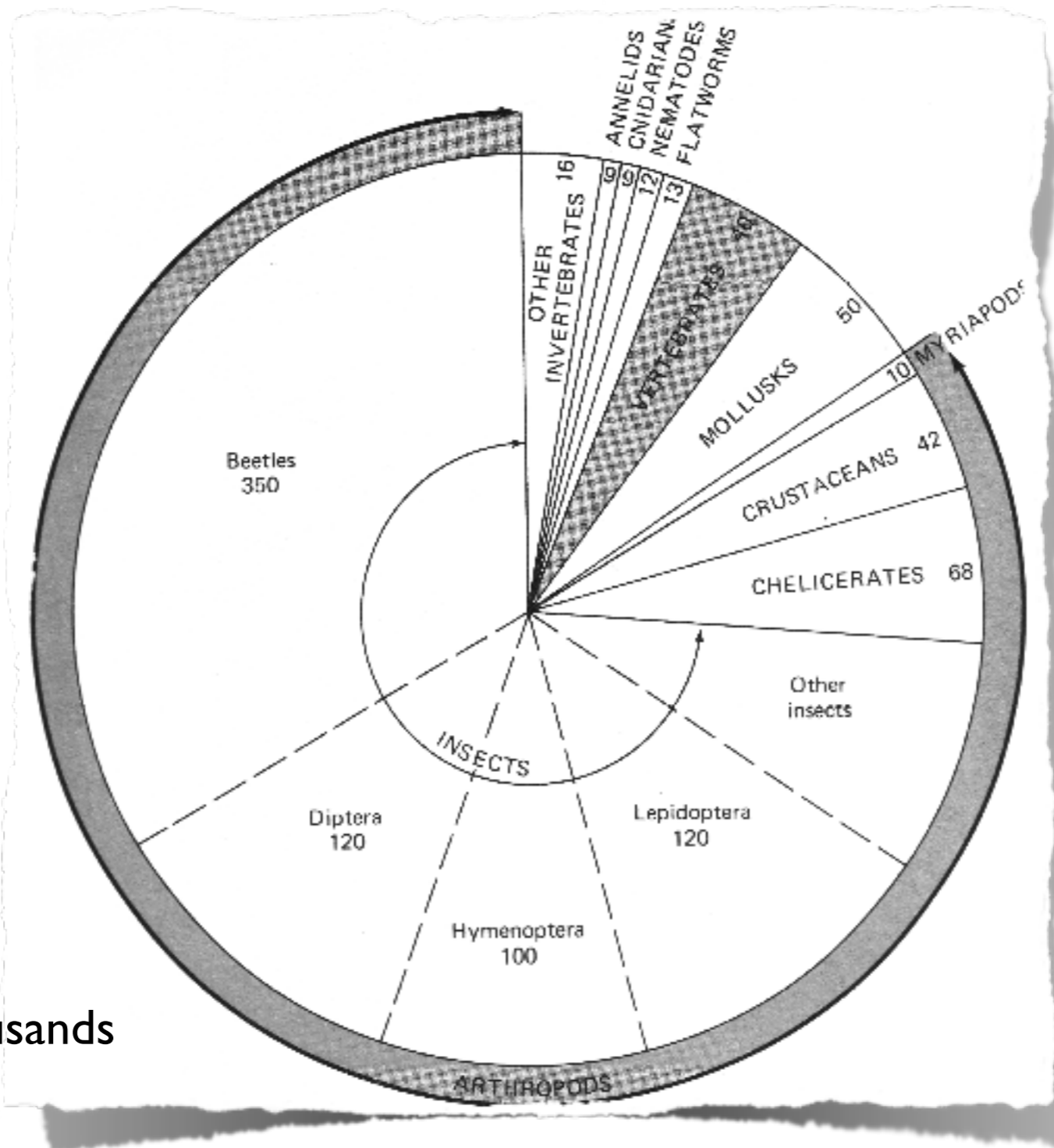


# Biodiversity

- more than 1 million arthropod species described
- zoologists estimate 3 million to 100 million undescribed species



# The Beetles Rule



numbers in thousands





mouse



bat



robin



carp

---

lungs

---

4 bony limbs

---

milk glands

---

hair

---

wings

---

feathers

---

jaws

---

# DATA MATRIX



mouse



bat



robin



carp

---

lungs

+

+

+

-

---

4 bony limbs

+

+

+

-

---

milk glands

+

+

-

-

---

hair

+

+

-

-

---

wings

-

+

+

-

---

feathers

-

-

+

-

---

jaws

+

+

+

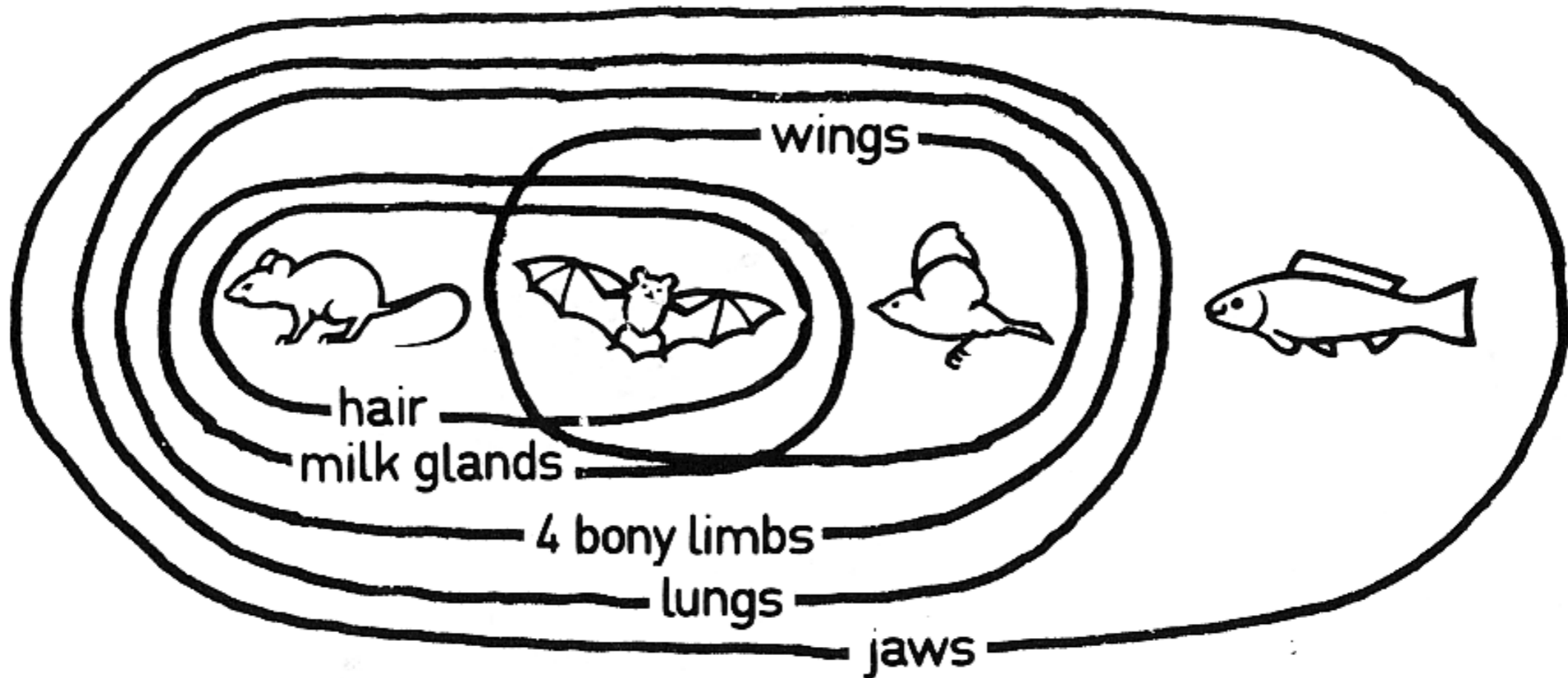
+

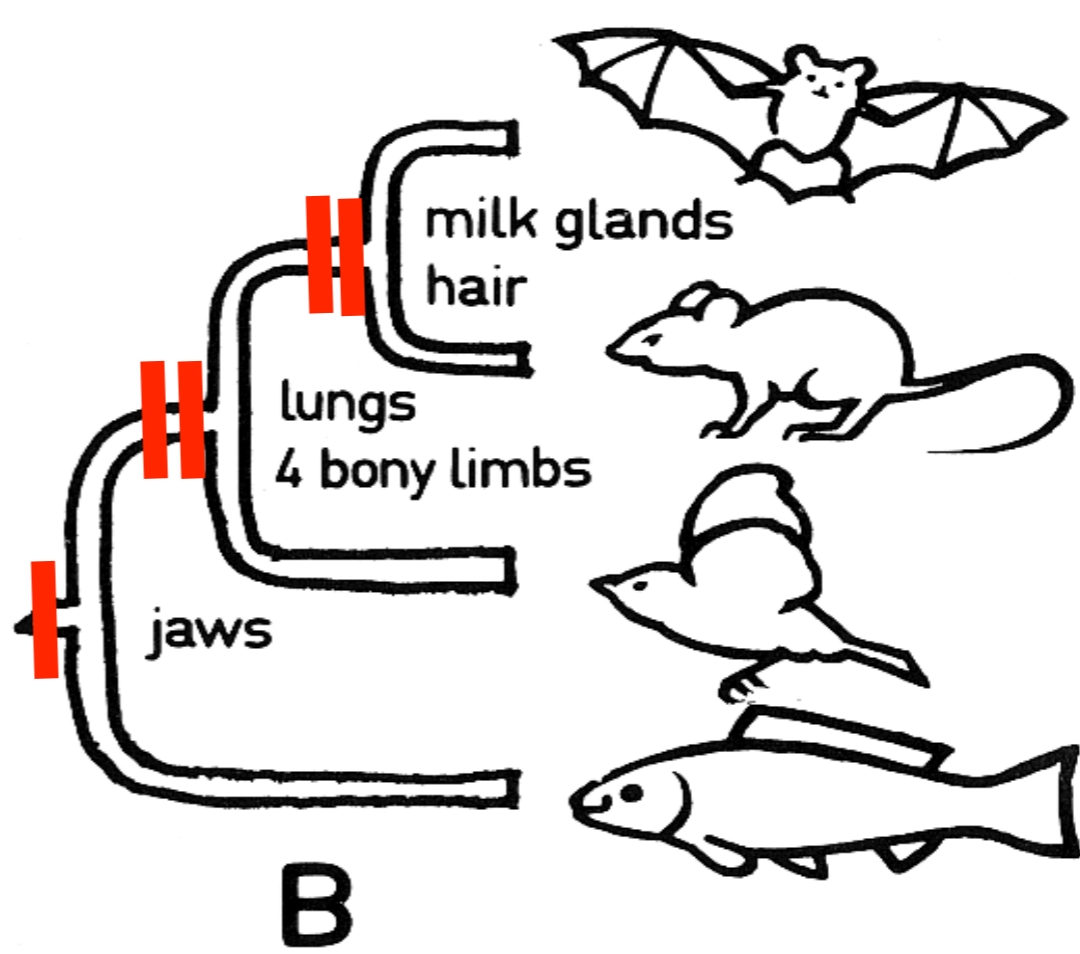
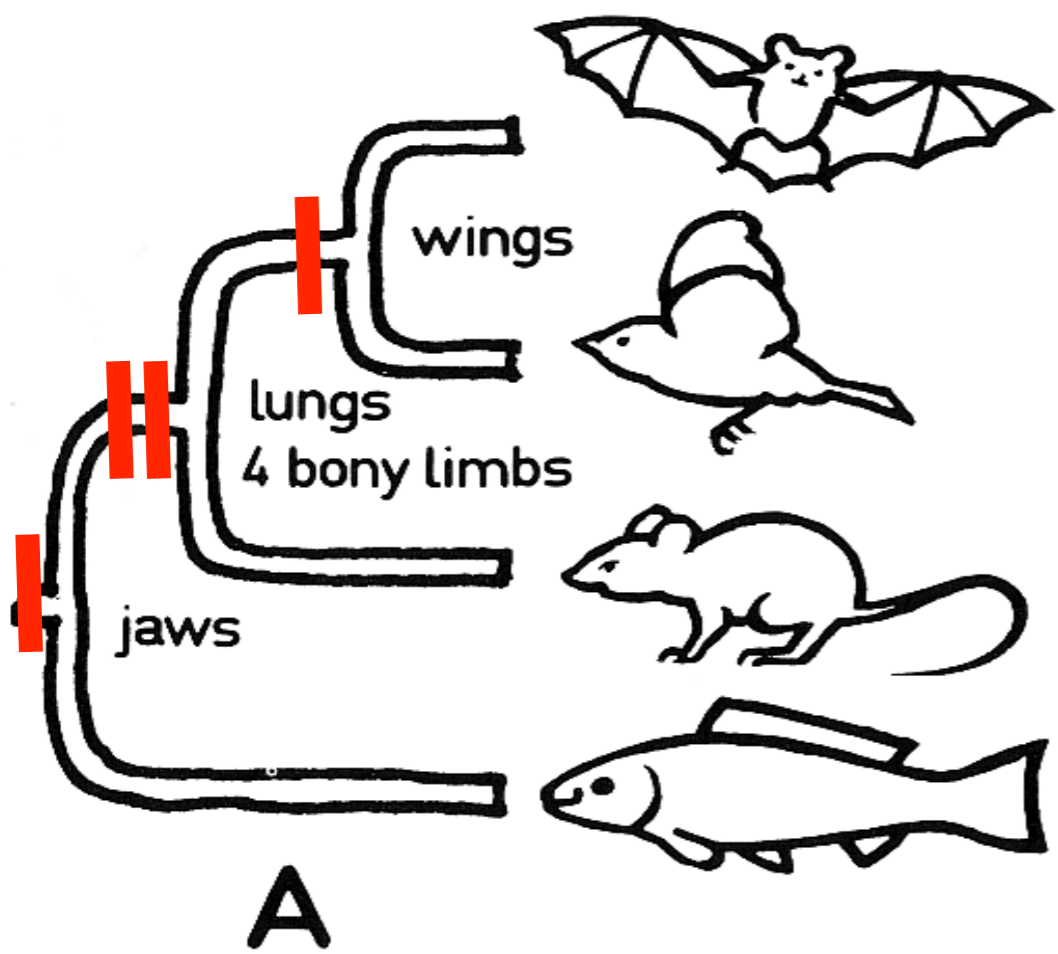
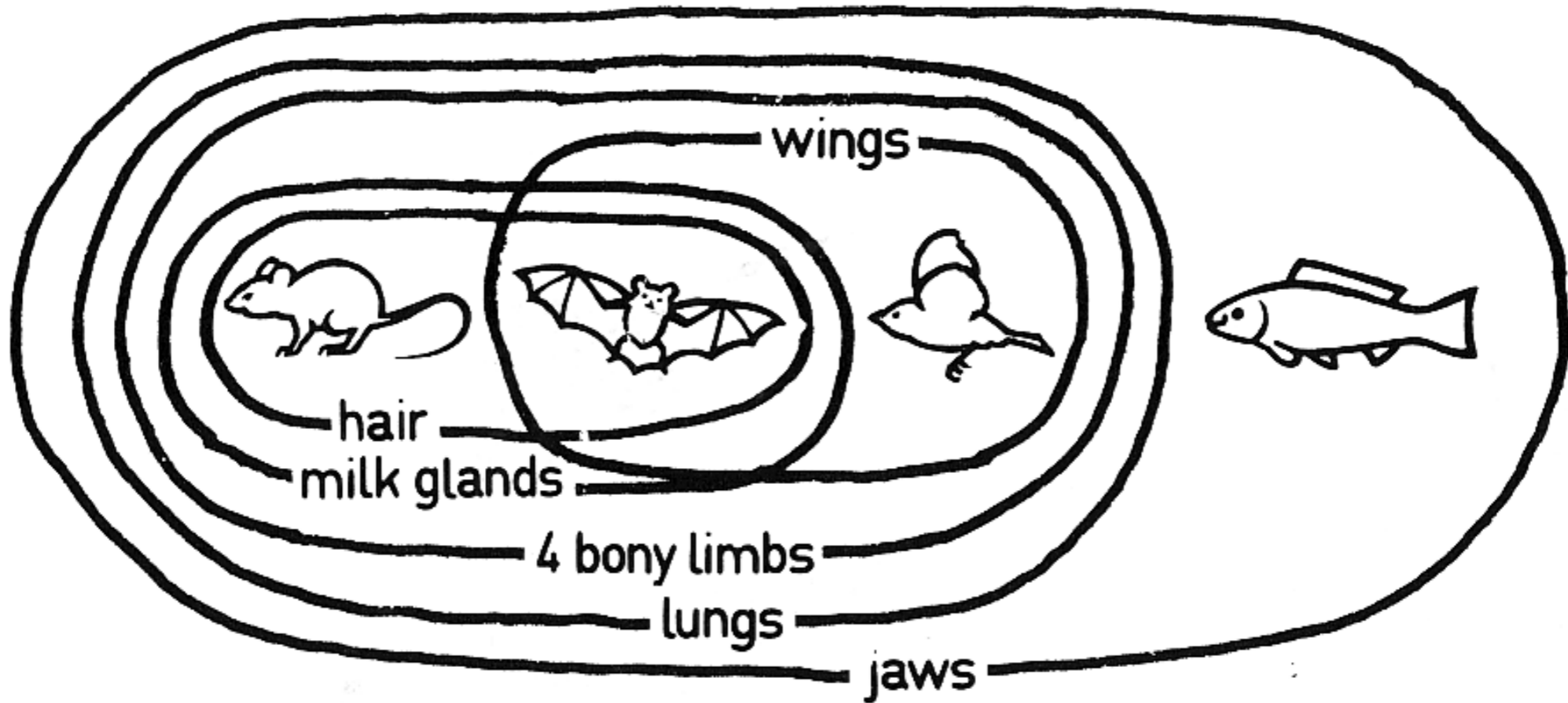
---



lungs	+	+	+	-
4 bony limbs	+	+	+	-
milk glands	+	+	-	-
hair	+	+	-	-
wings	-	+	+	-
feathers	-	-	+	-
jaws	+	+	+	+

# Sort into groups



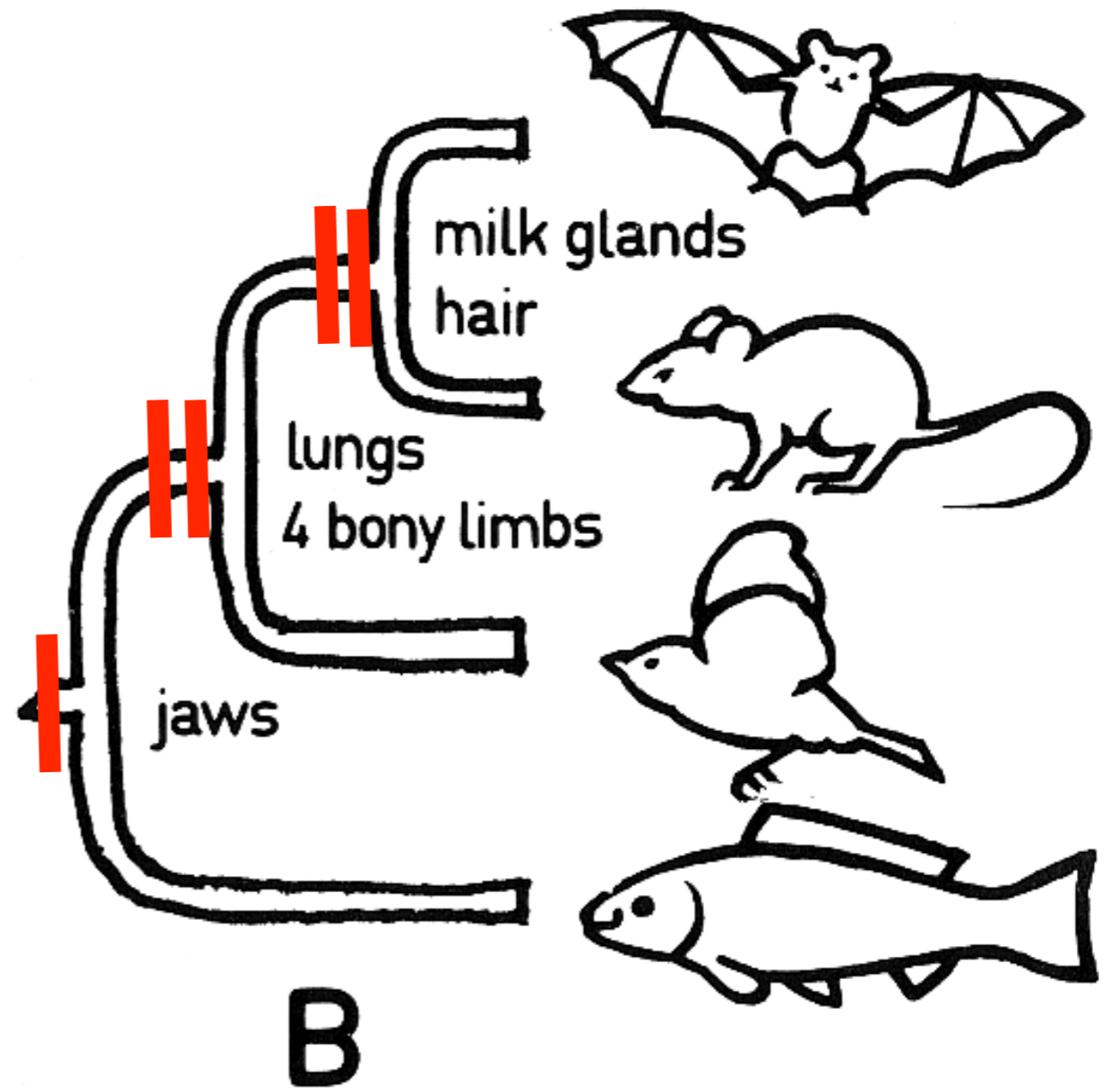
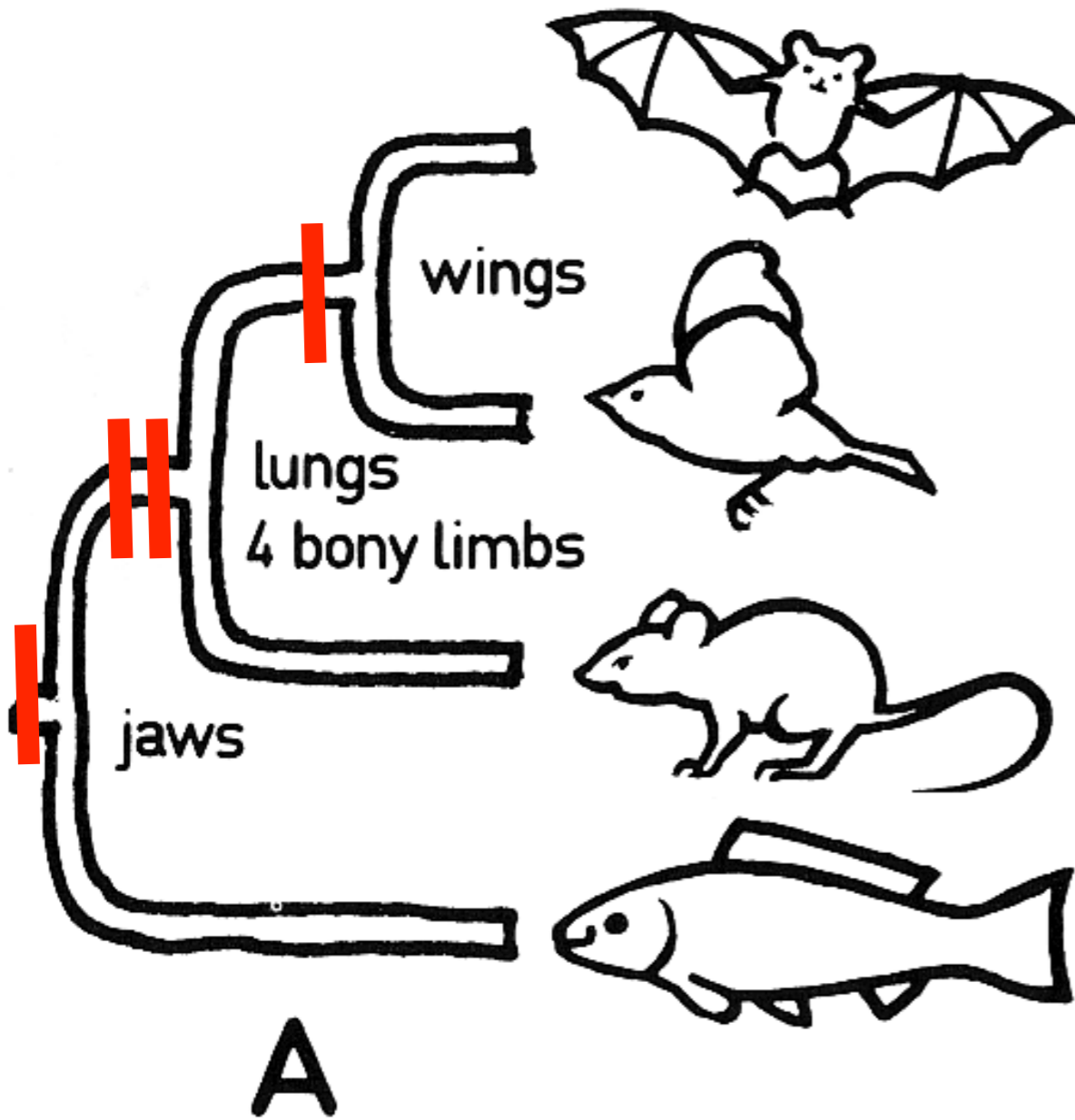


So which of the two  
phylogenies would you  
select as the better  
supported phylogenetic  
hypothesis?

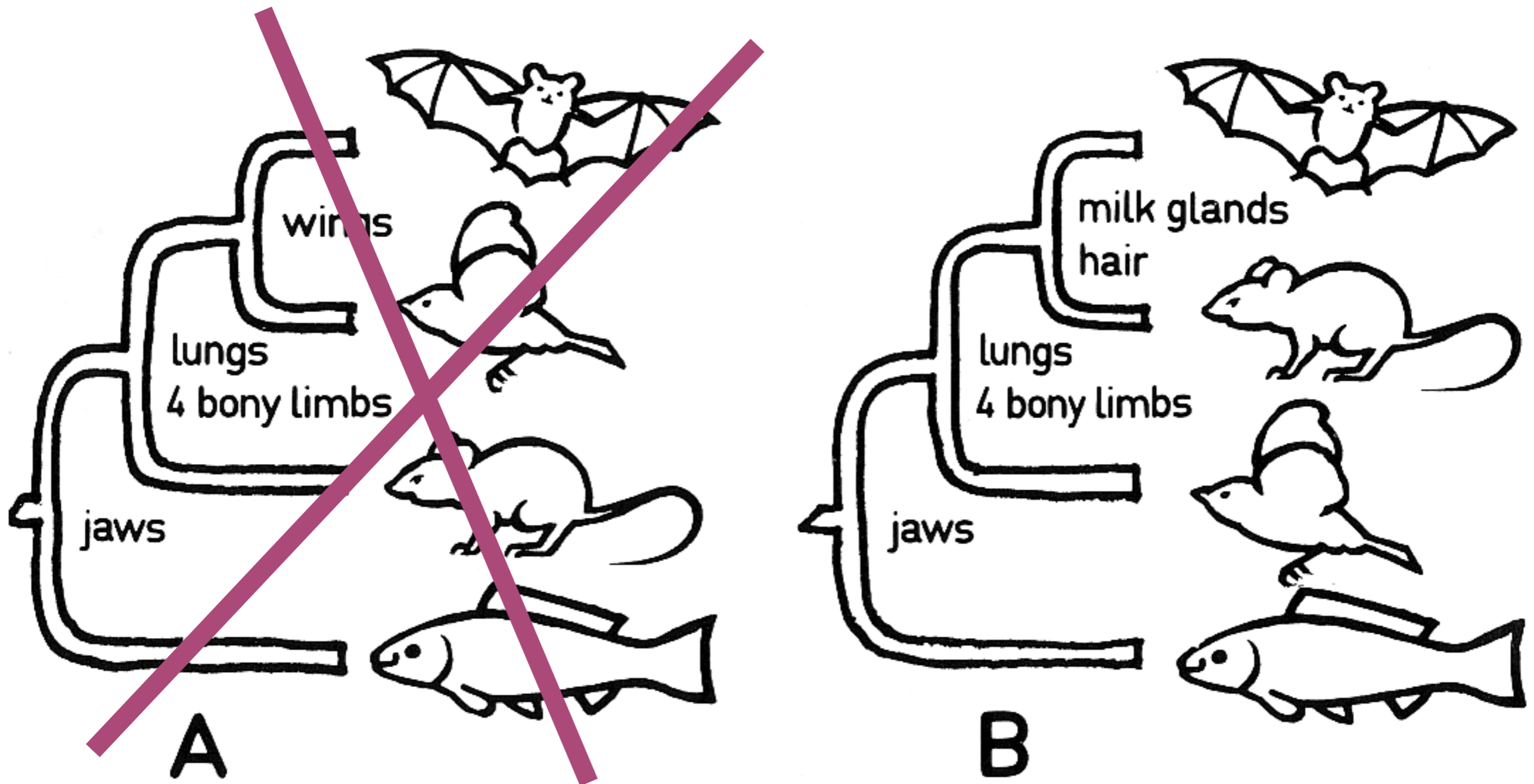




# Multiple trees



# Multiple trees



Choose cladogram B, because it is supported by 5 shared features. Cladogram A is supported by only 4 shared features.