

Phylogenetic Biology

Week 2

Biology 1425

Professor: Casey Dunn, dunnlab.org

Brown University

2013

Front matter...

All original content in this document is distributed under the following license:

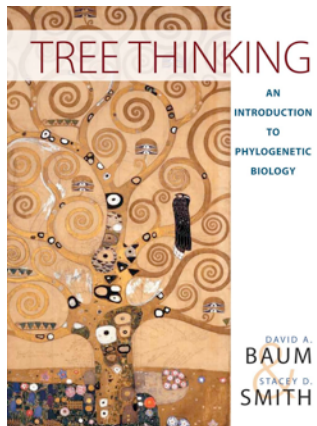


Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License
(http://creativecommons.org/licenses/by-nc-sa/3.0/deed.en_US)

See sources for copyright of non-original content

Sources

Some non-original content is drawn from:



Baum, D and S. Smith (2012) Tree Thinking: and Introduction to Phylogenetic Biology. Roberts and Company Publishers. ISBN 9781936221165

Other non-original content is referenced by url.

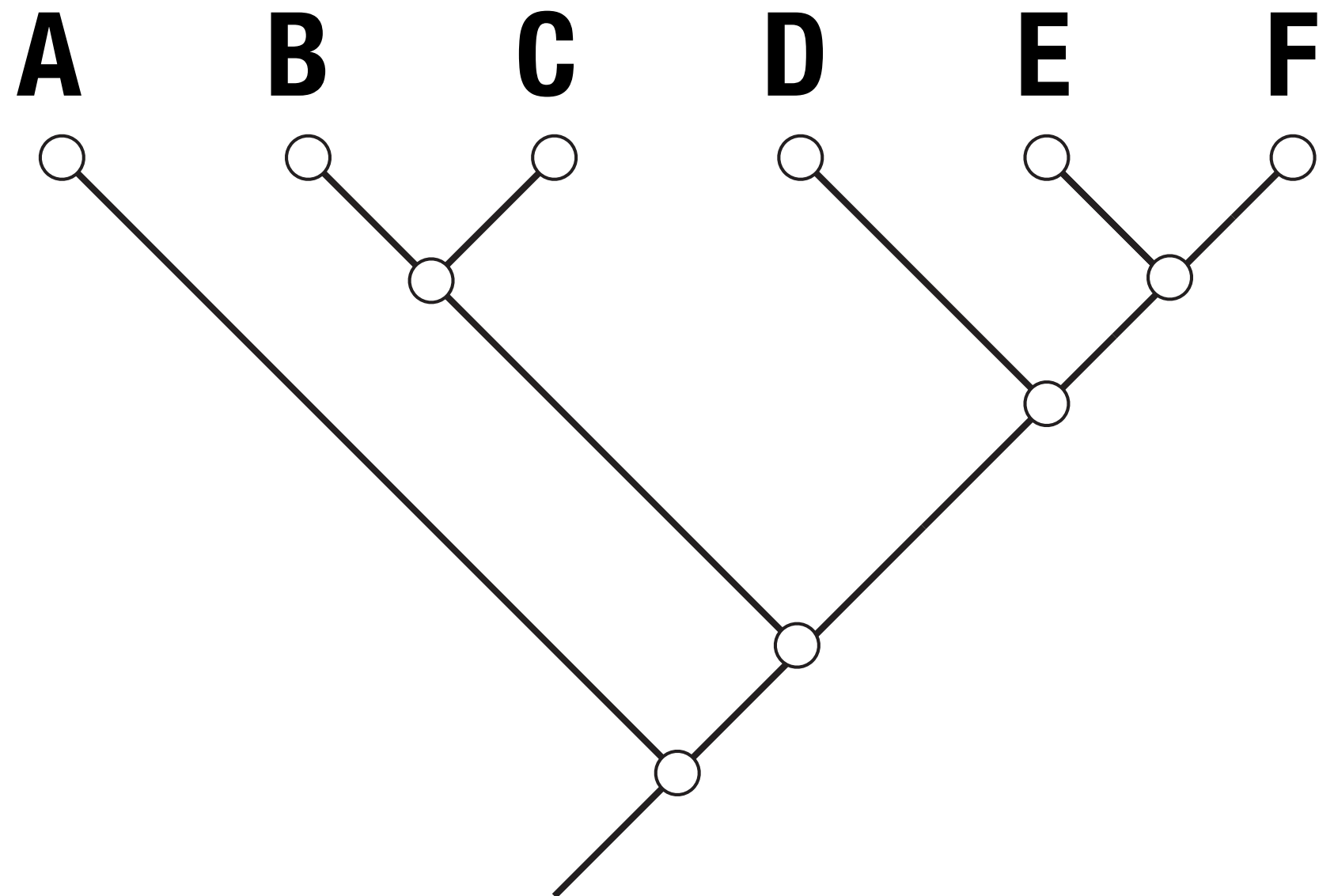
Some definitions...

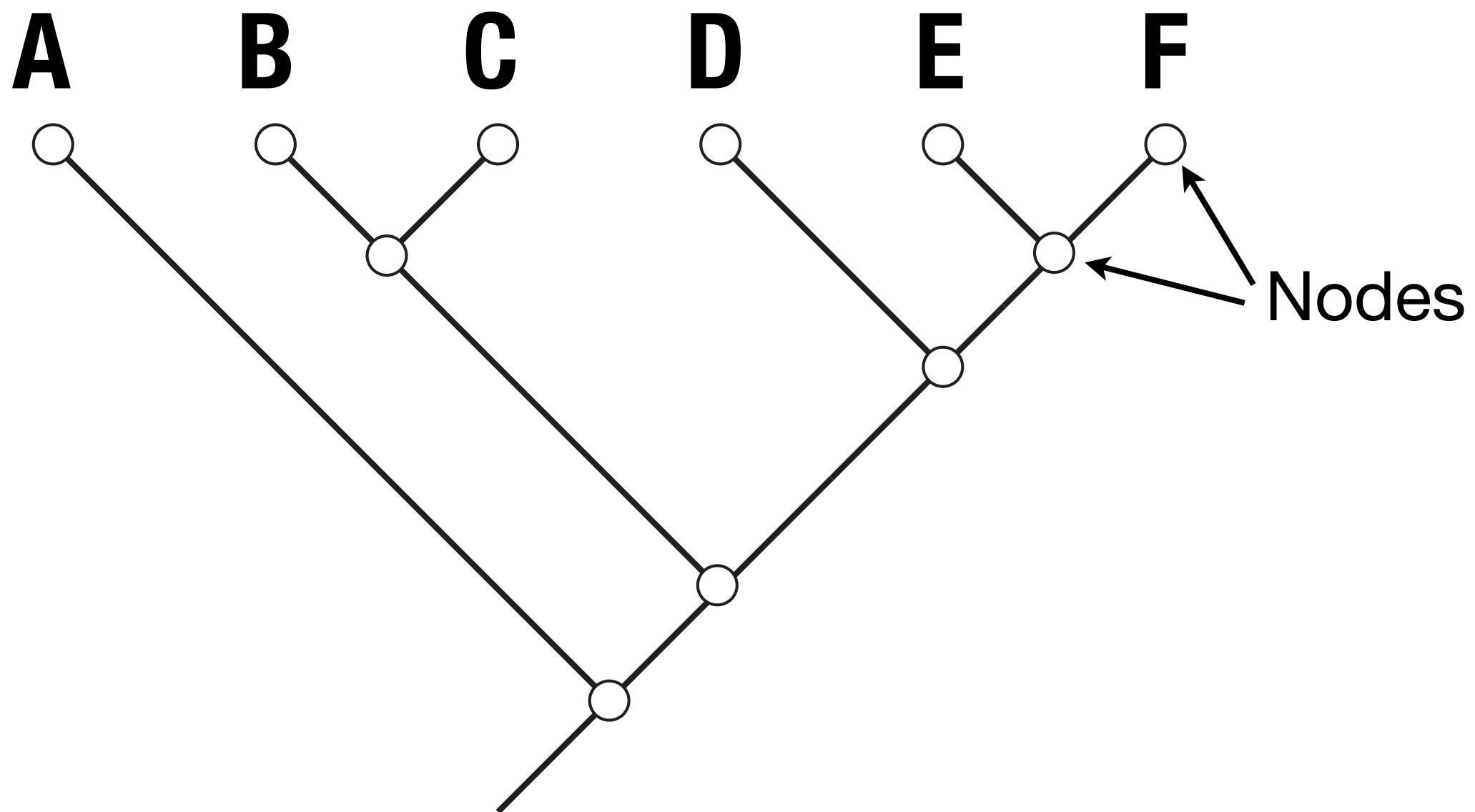
Taxa

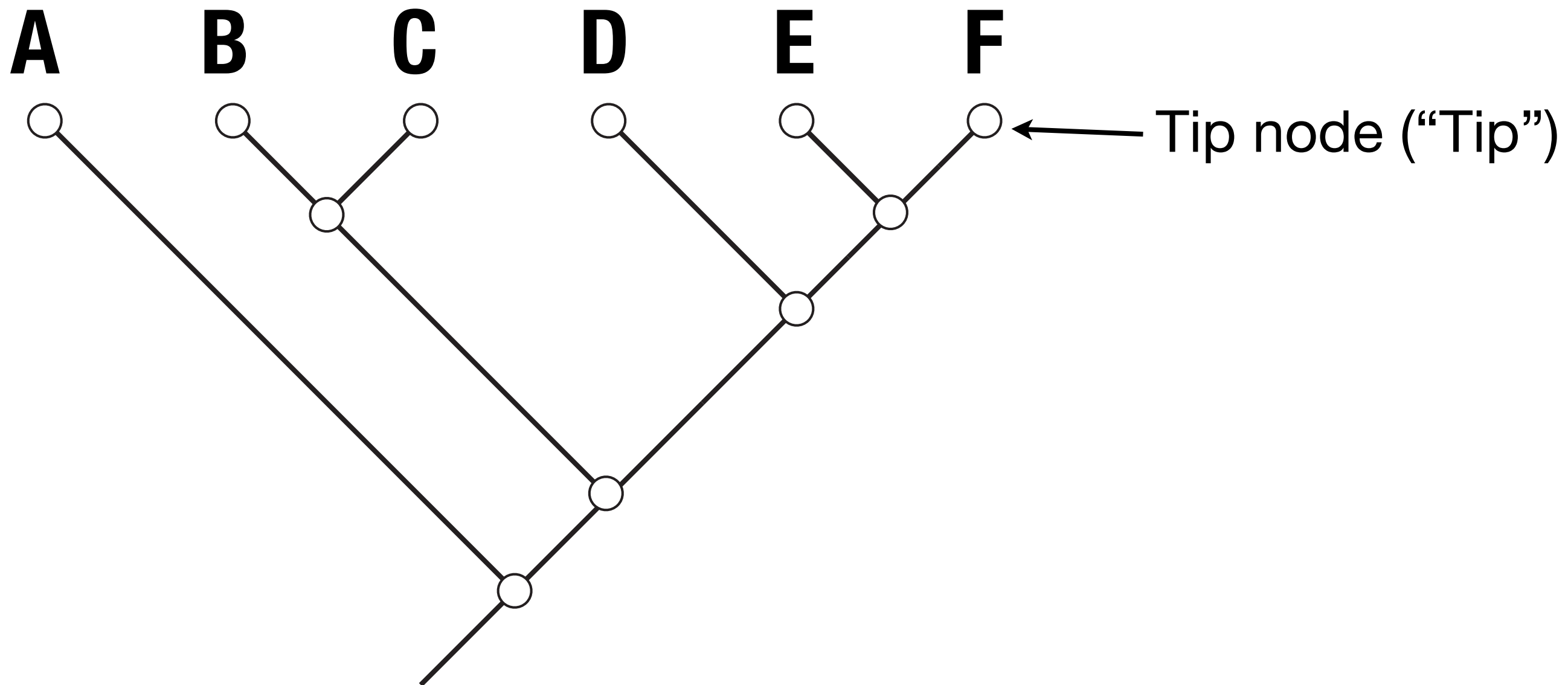
A named group of evolving entities (*e.g.* genes, species, languages)

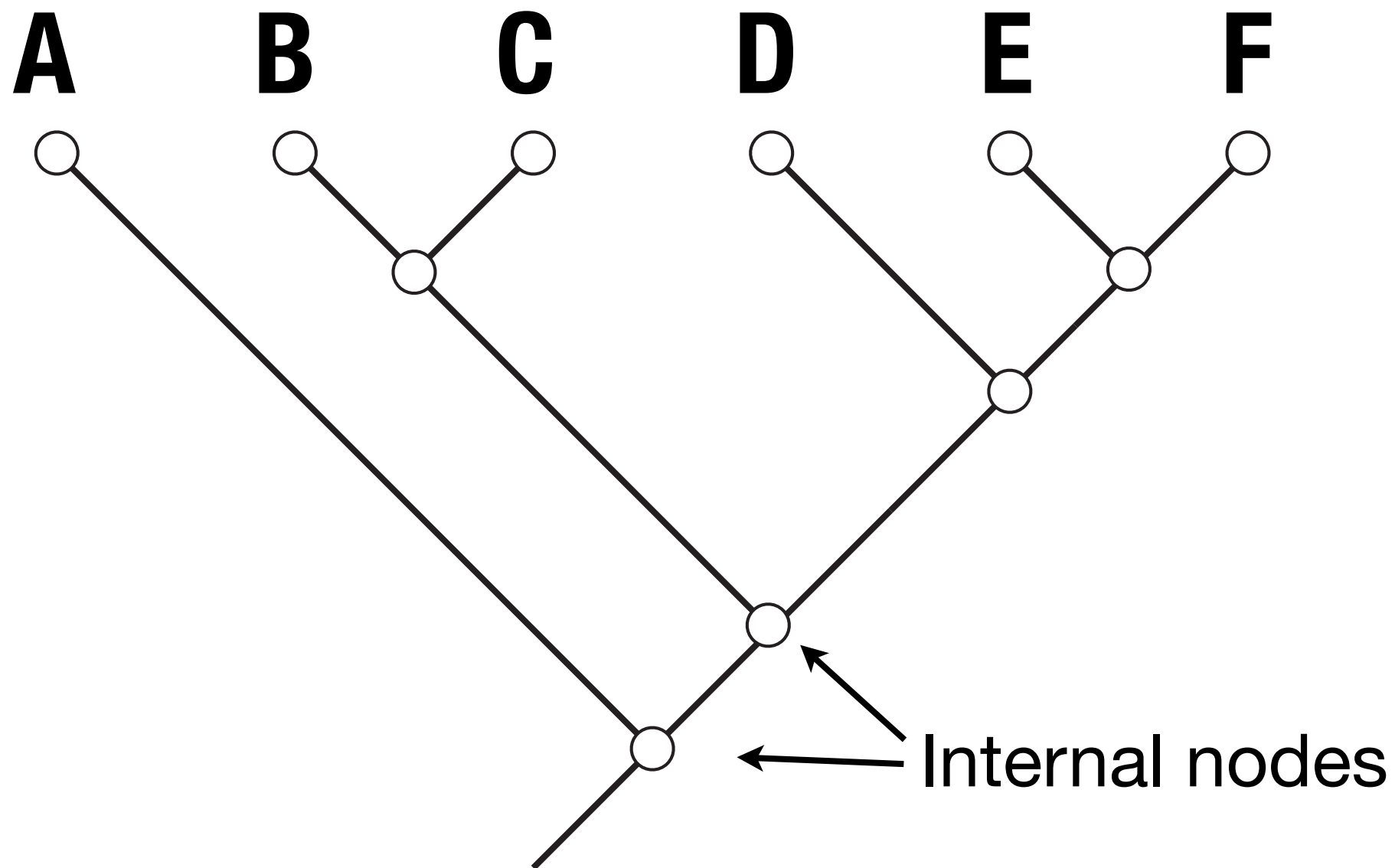
Phylogeny

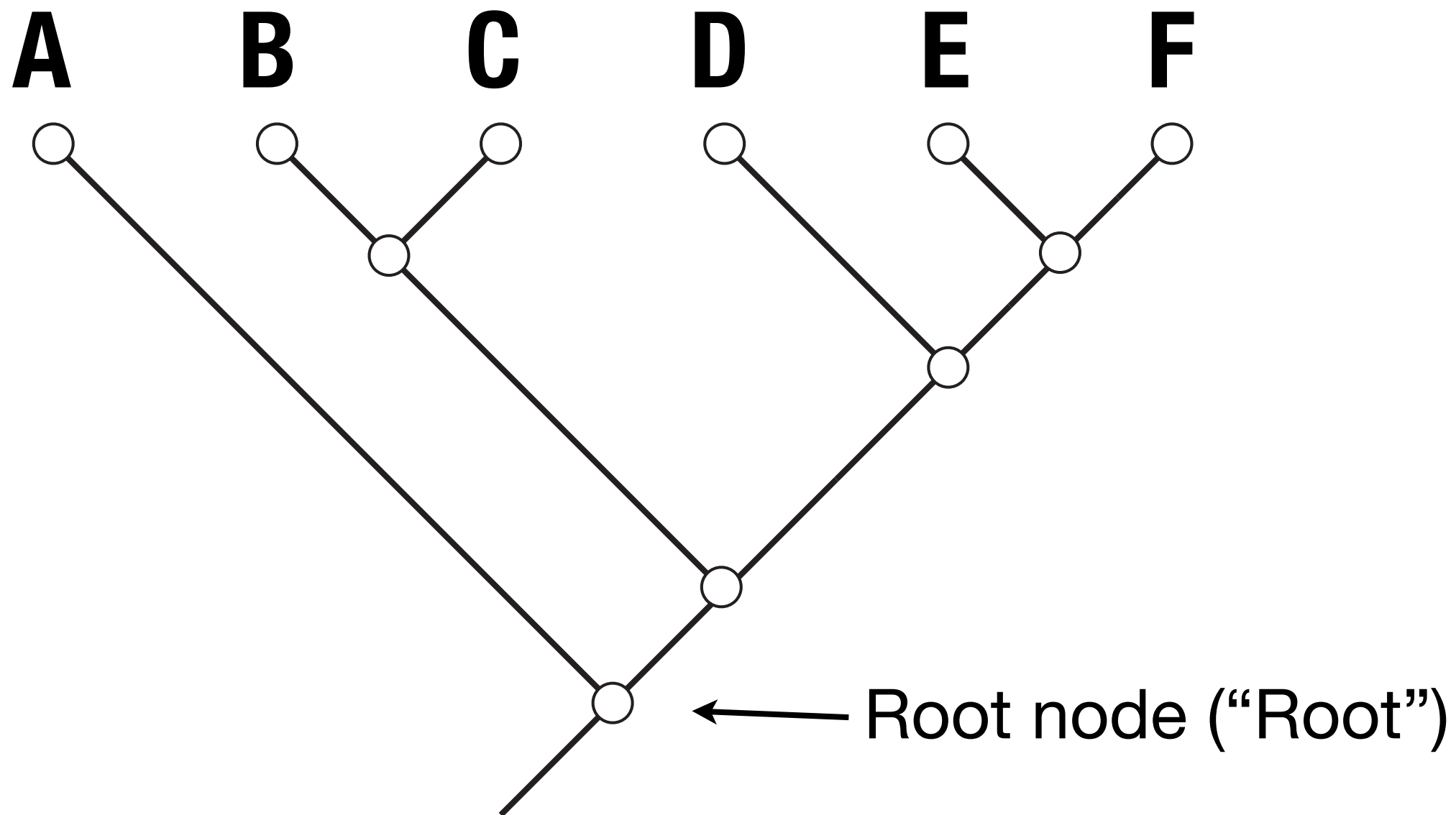
A tree depicting the evolutionary relationships between taxa

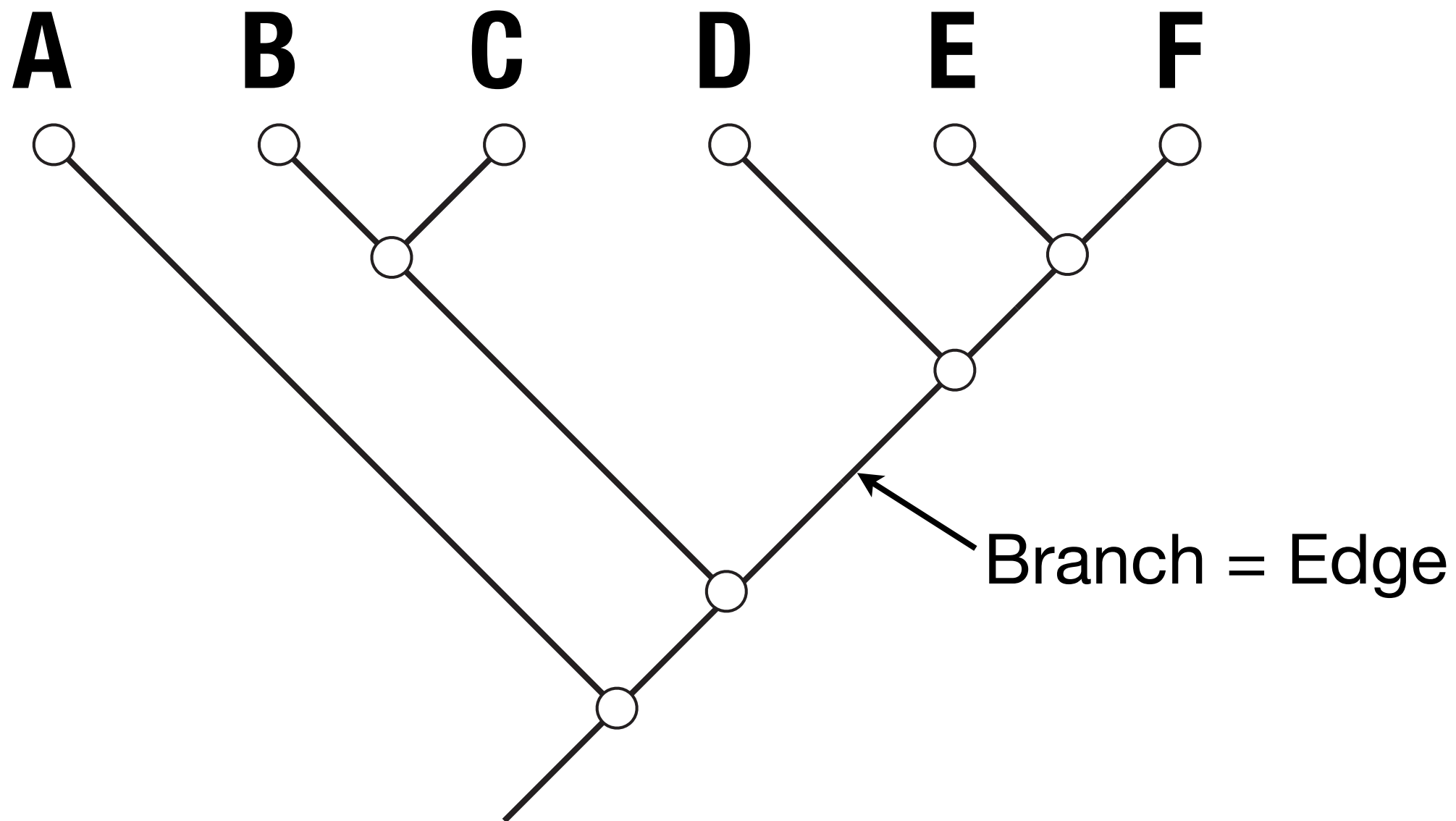






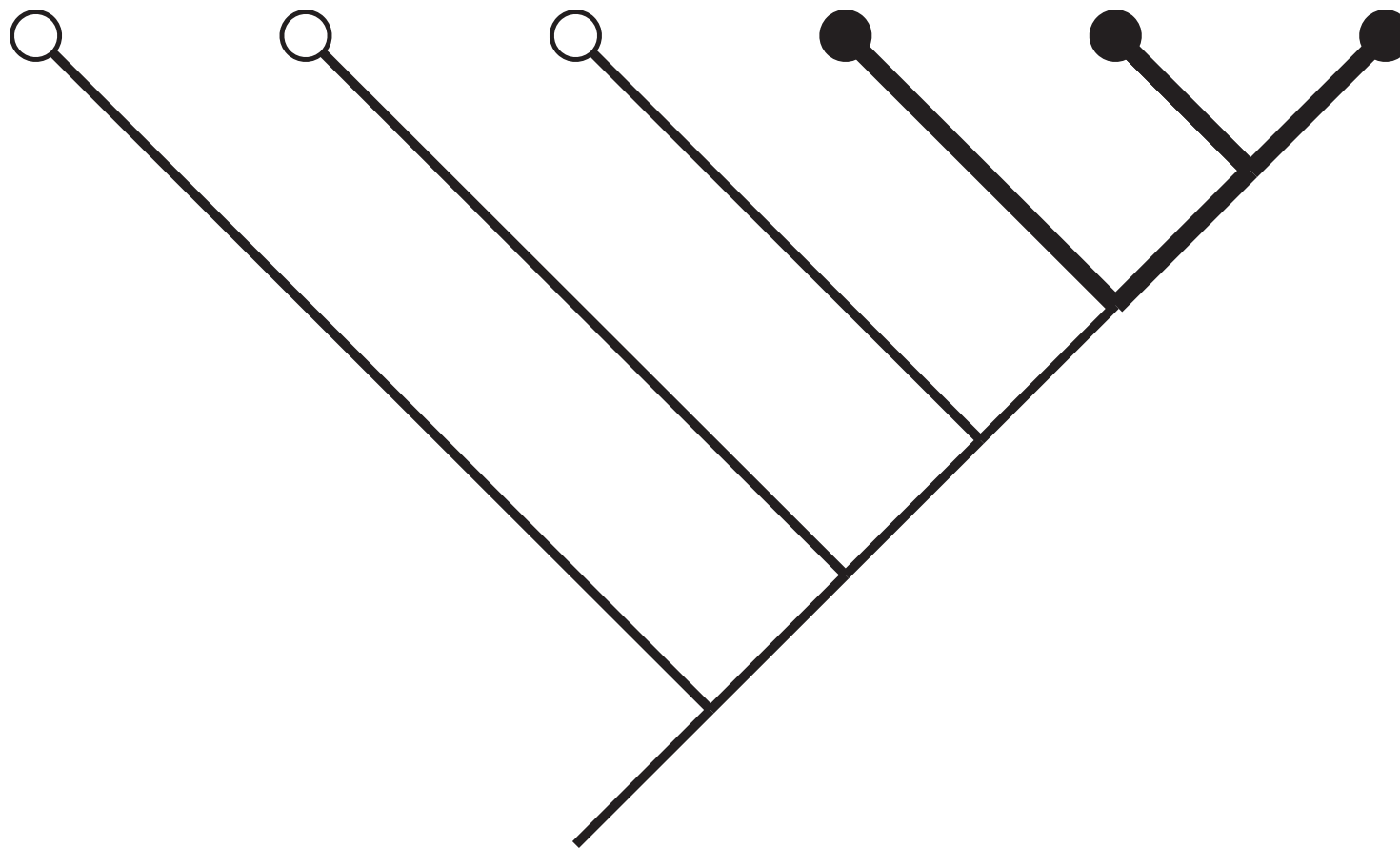






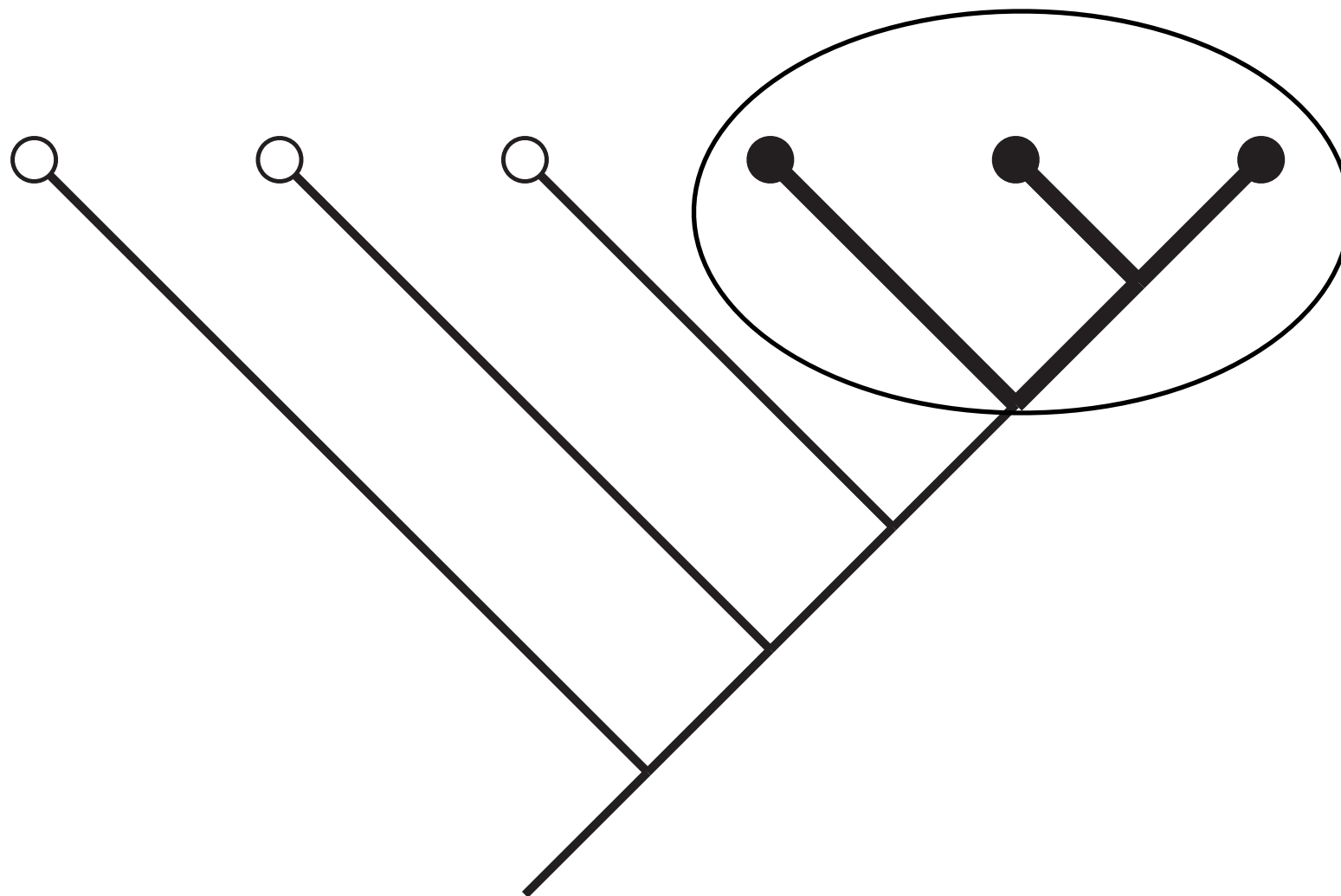
Monophyletic

To include an ancestor, all of its descendants, and nothing that isn't a descendent



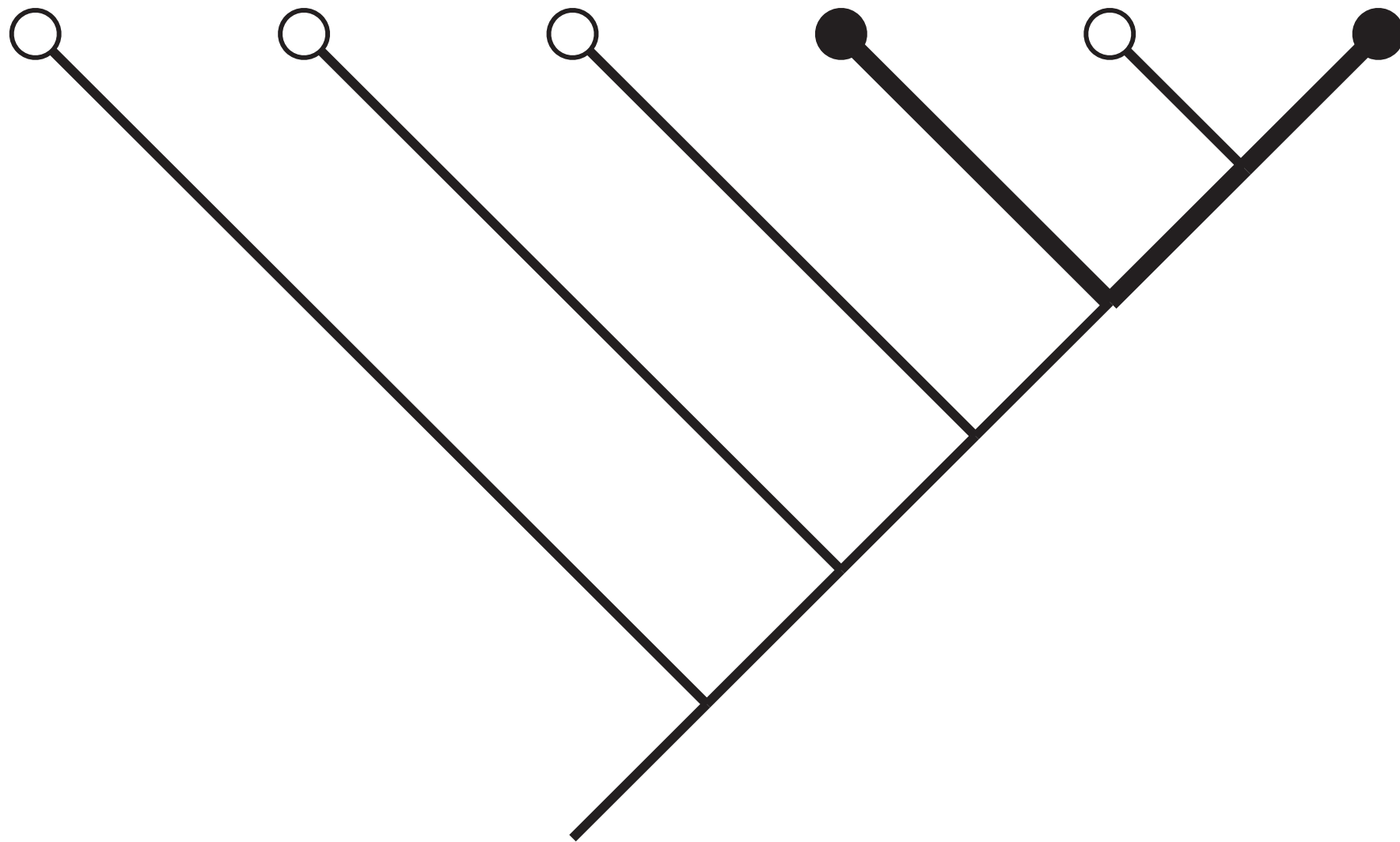
Clade

A monophyletic group



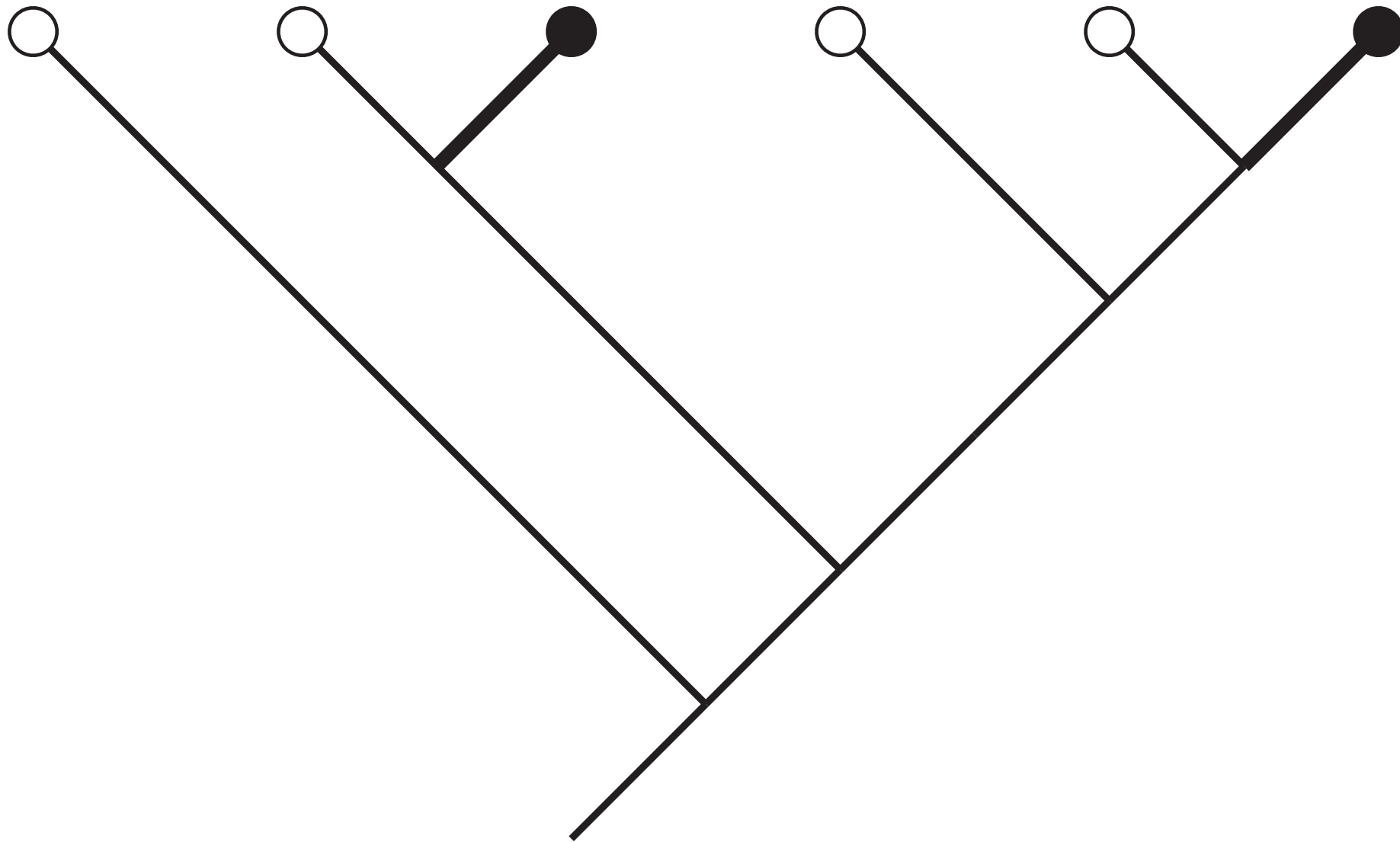
Paraphyletic

To include an ancestor and only some of its descendants



Polyphyletic

To not include an ancestor



Character

Any heritable attribute of a taxon

Homologous characters

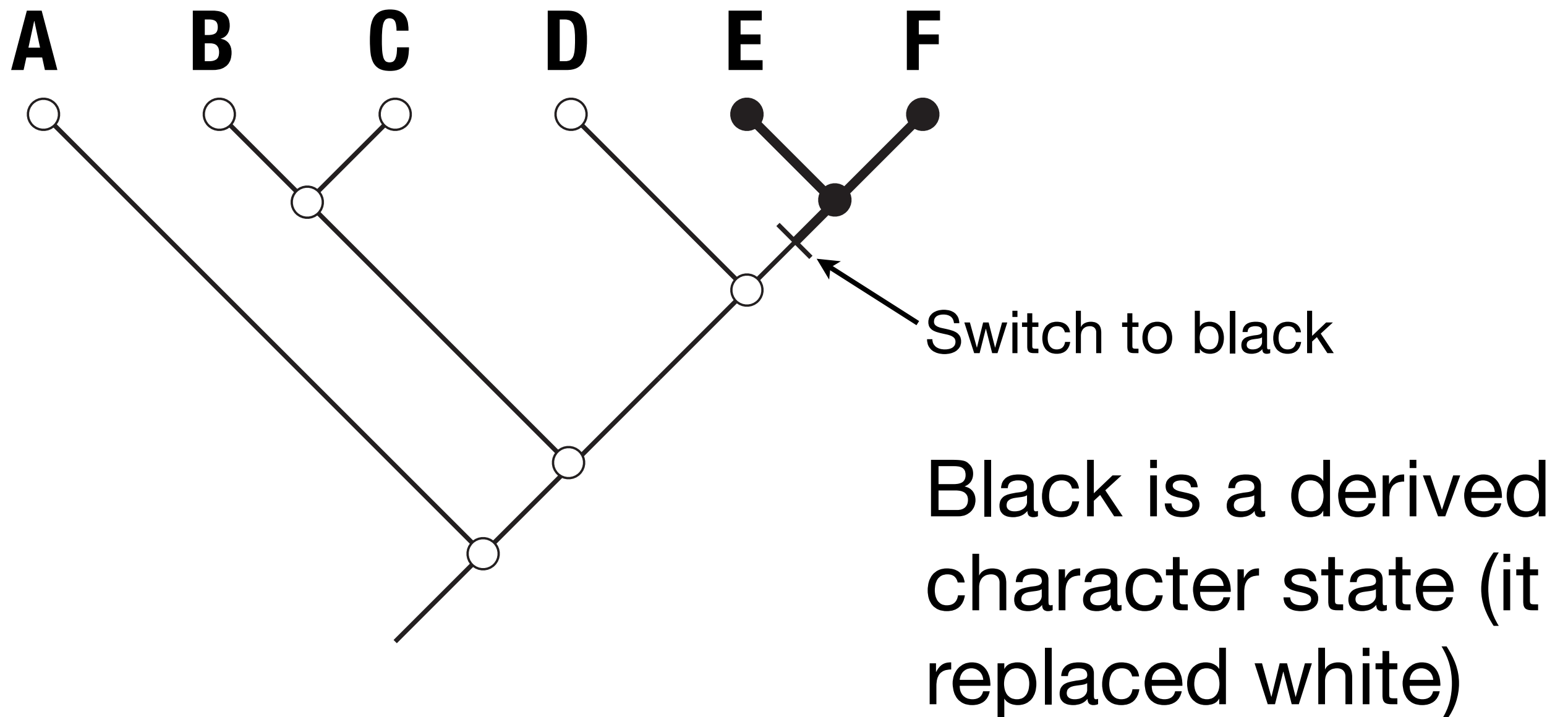
Characters that are present in taxa because they were also present in the most recent common ancestor of those taxa

Character state

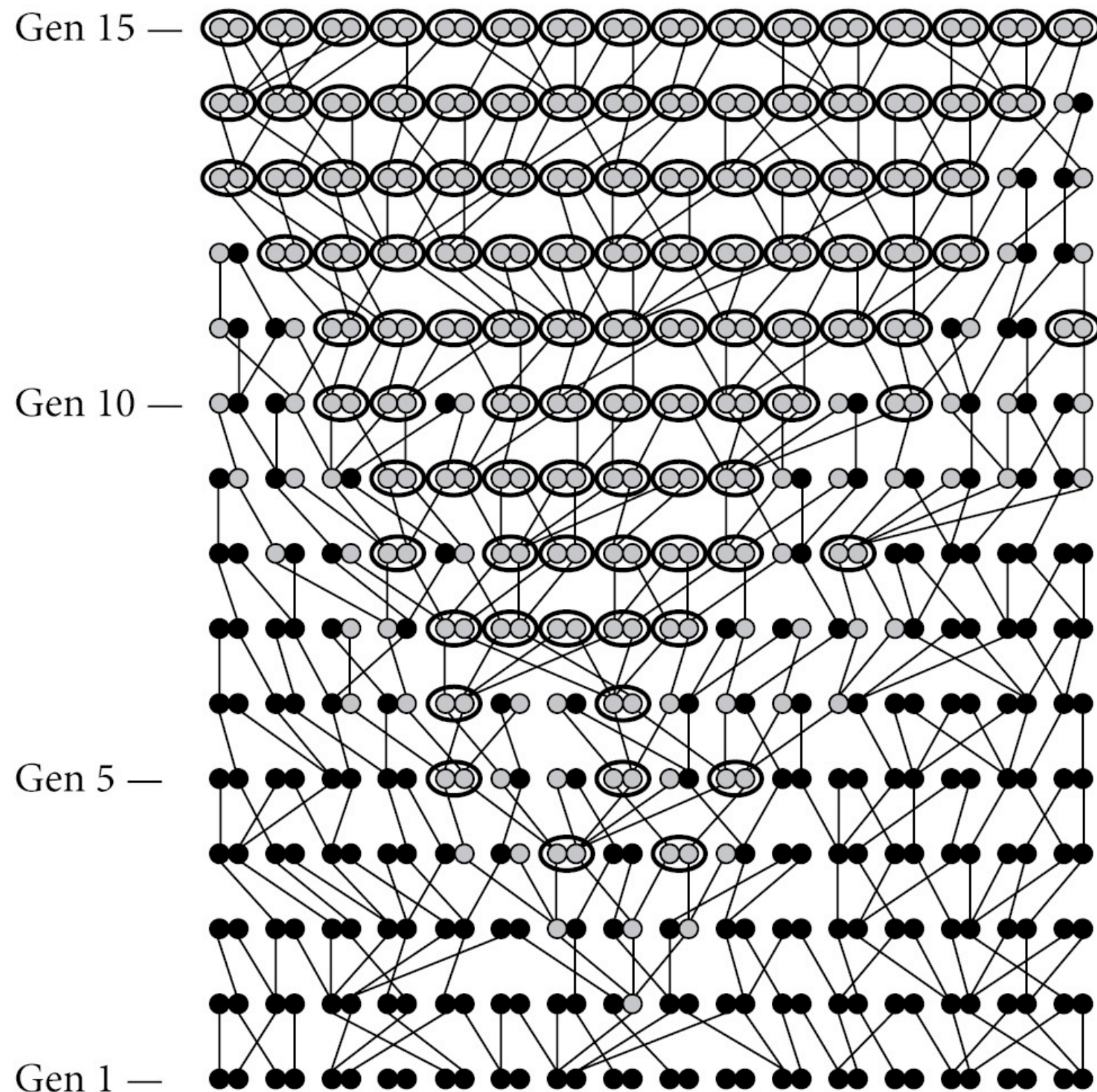
A value that a character can have. (*e.g.*, A, C, G, or T for a DNA nucleotide)

Apomorphy (Derived character state)

A character state that is different from the character state of an ancestor



Apomorphy (Derived character state)

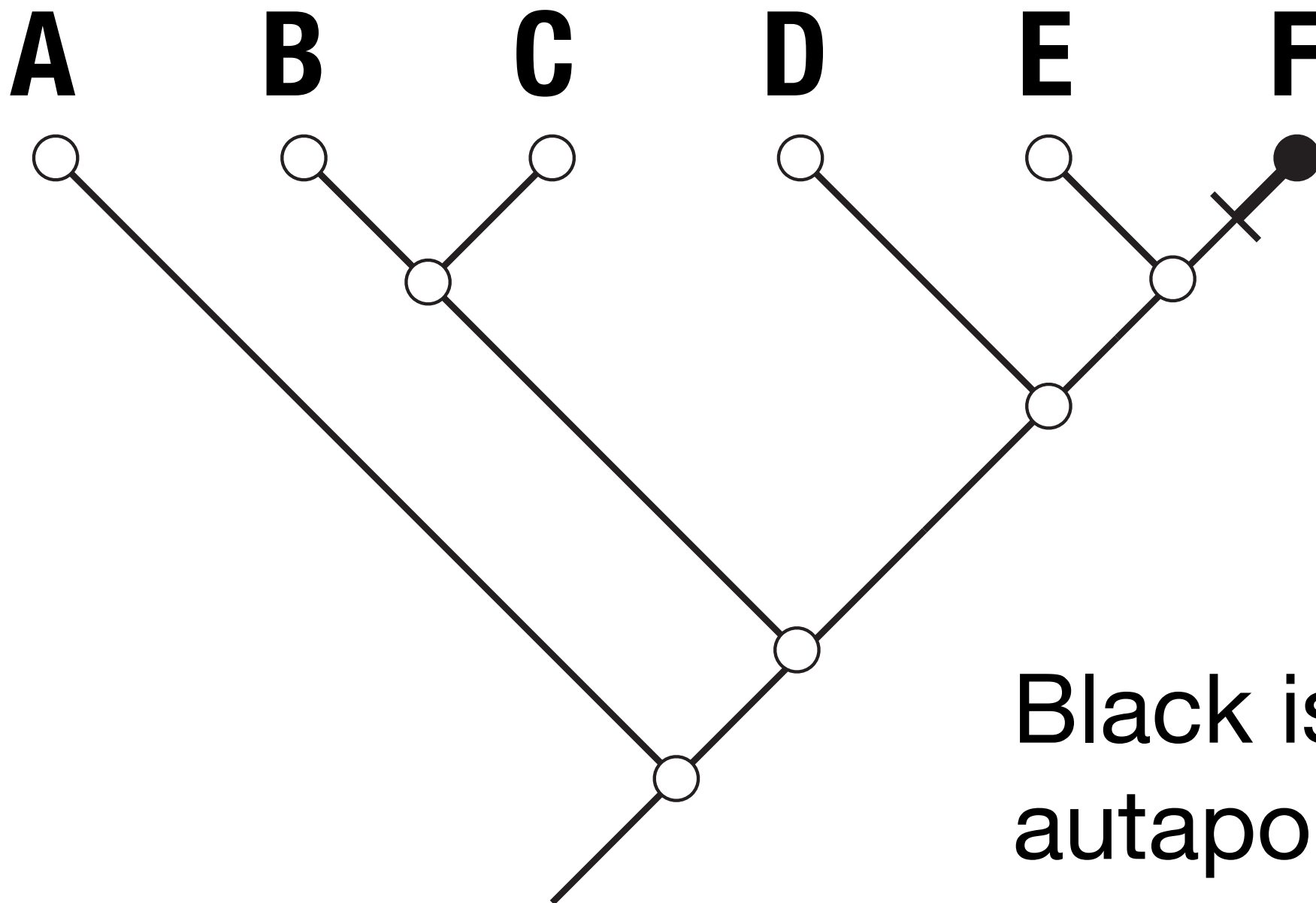


Fixation of derived
character states

Baum and Smith 2012,
Figure 4.1

Autapomorphy

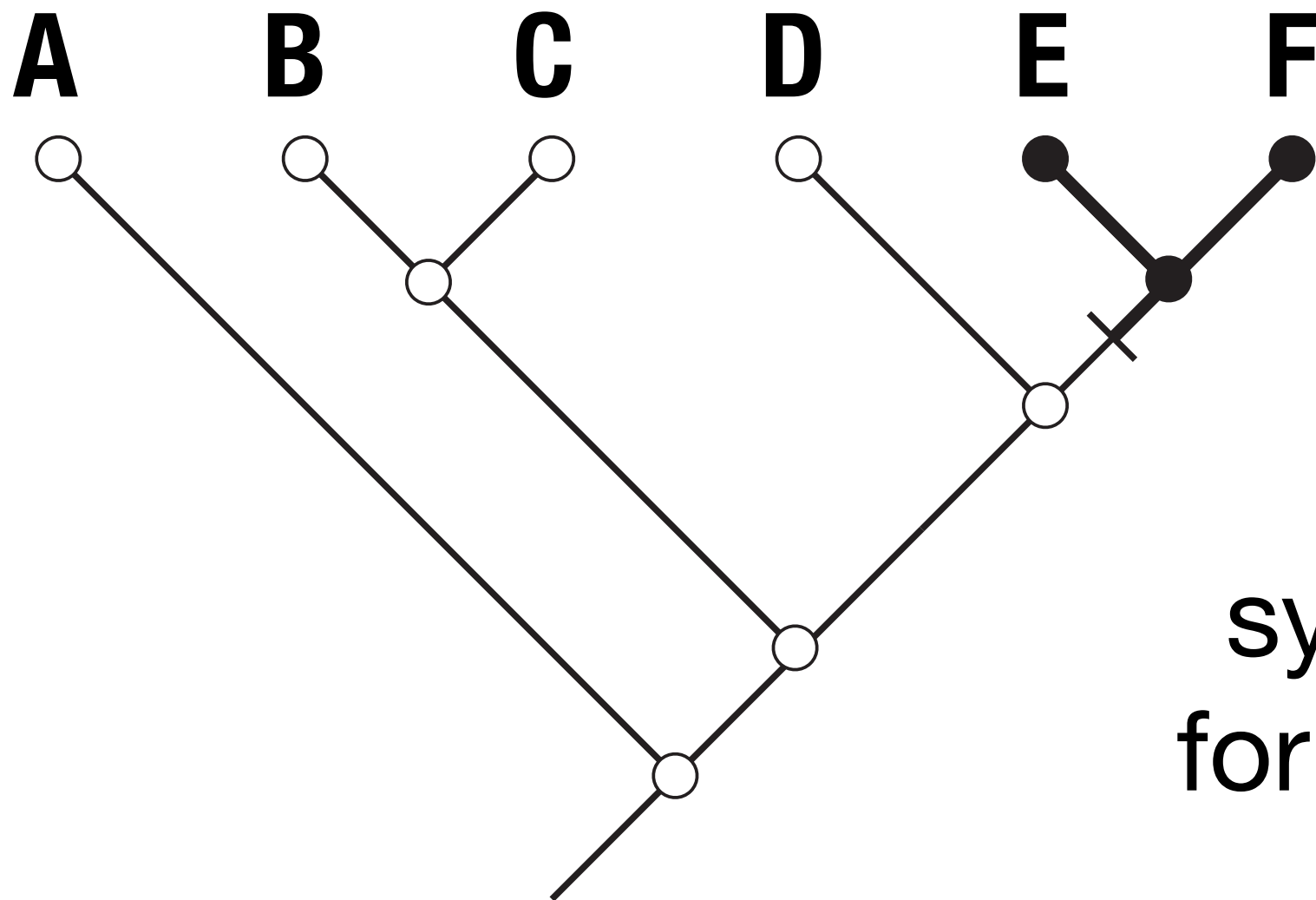
An apomorphy that is unique to a single taxon



Black is an
autapomorphy of **F**

Synapomorphy

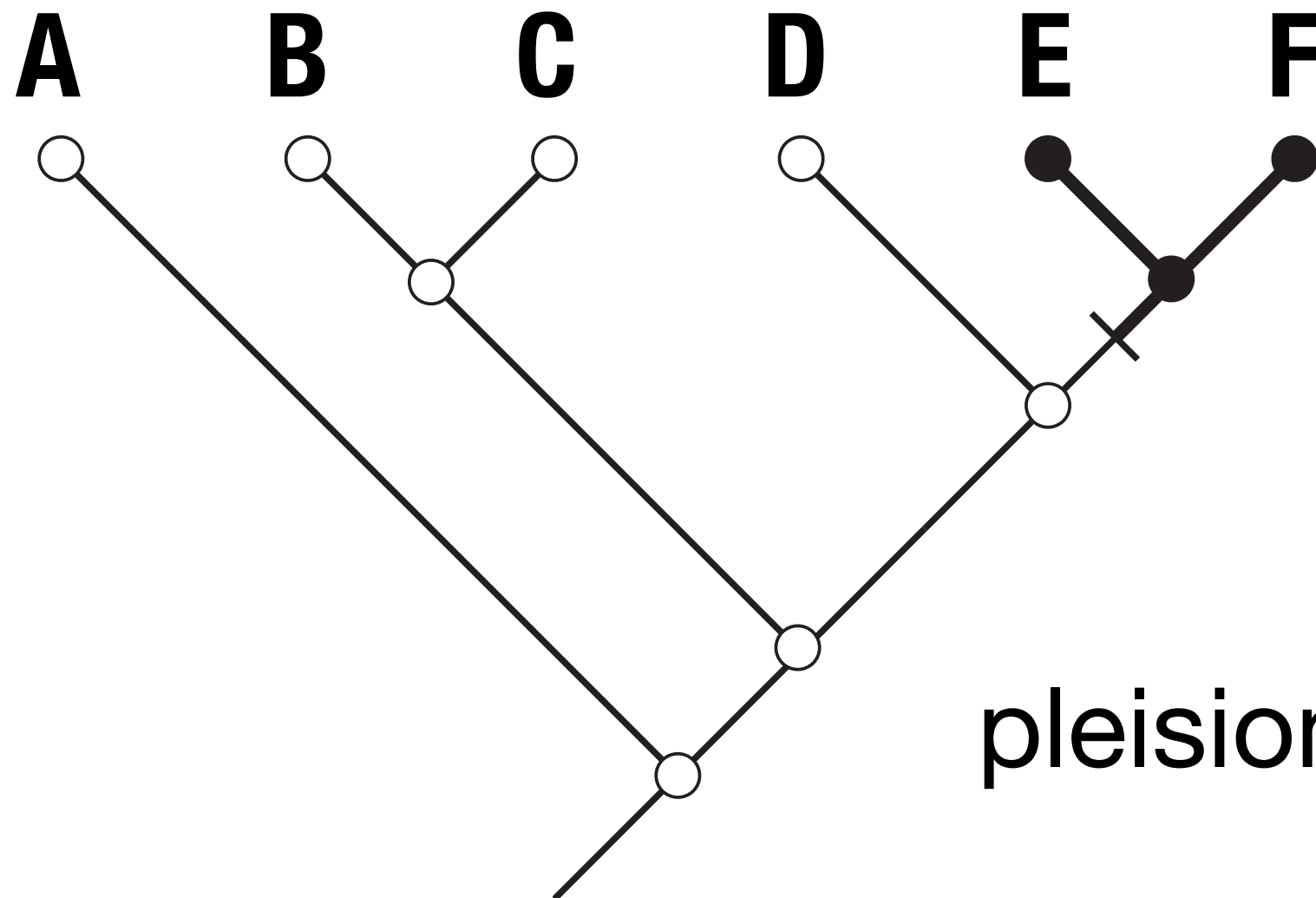
A shared apomorphy



Black is a
synapomorphy
for the clade **EF**

Pleisiomorphy

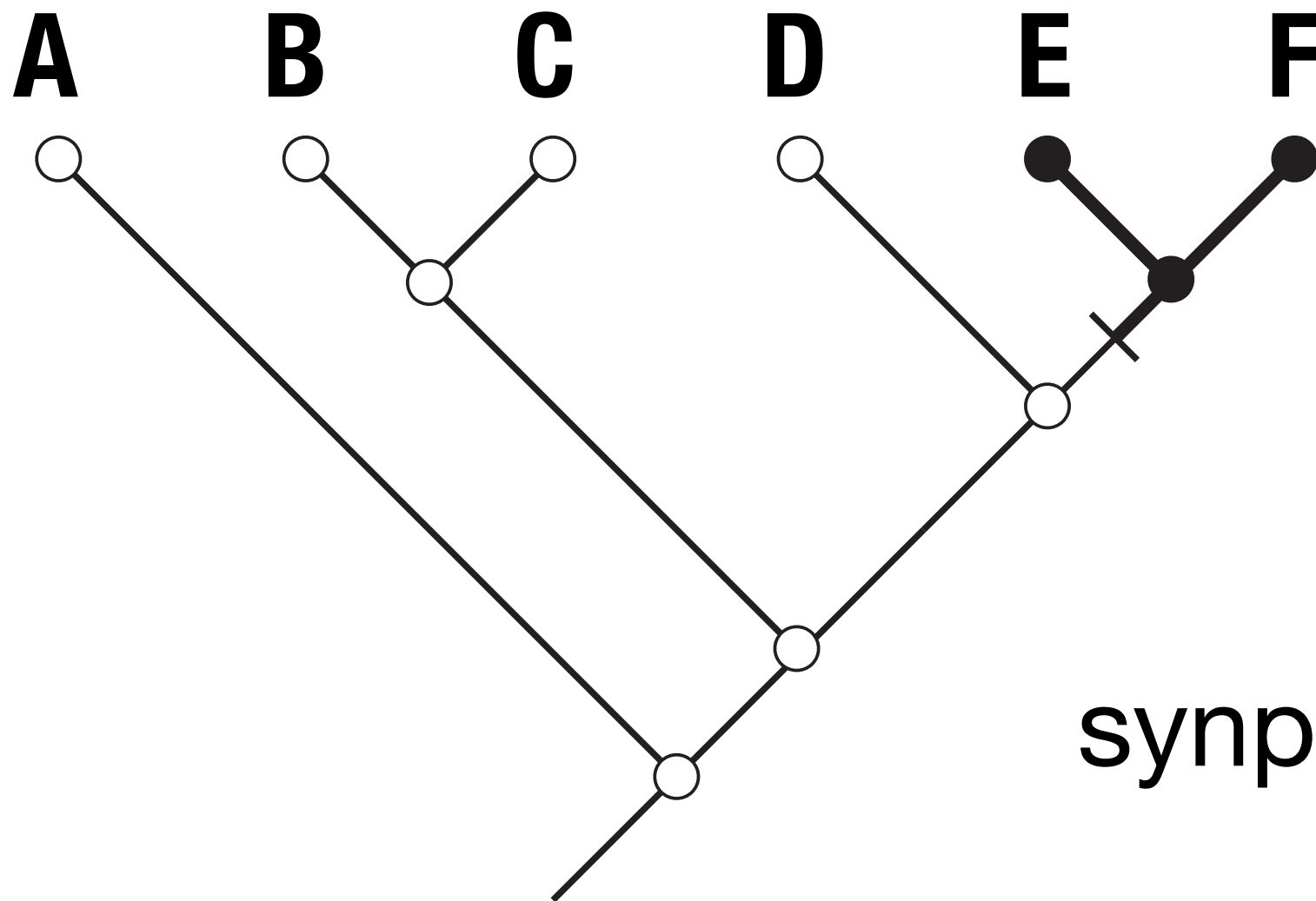
A character state that is the same as that of an ancestor



White is a
pleisiomorphy for **B**

Synpleisiomorphy

A shared pleisiomorphy



White is a
synpleisiomorphy
for **BC**

Homoplasy

Independent origin of the same character state. Causes include convergence and reversals.

Character matrix

A table of character states. Each row corresponds to a taxon, and each column to a homologous character.

Characters

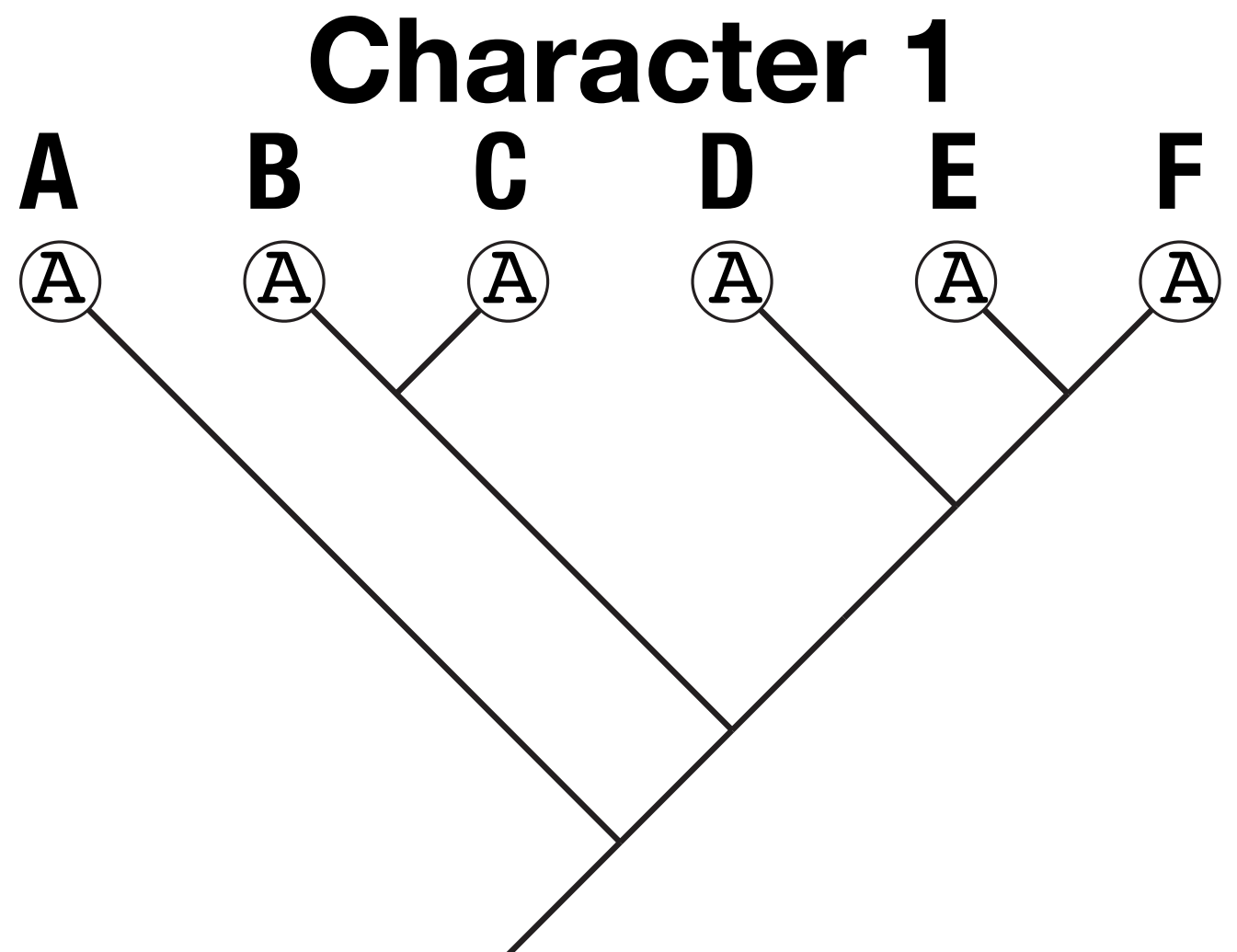
Taxa

	1	2	3
A	A	C	A
B	A	G	A
C	A	G	A
D	A	C	A
E	A	C	T
F	A	C	T

Character matrix

A table of character states. Each row corresponds to a taxon, and each column to a homologous character.

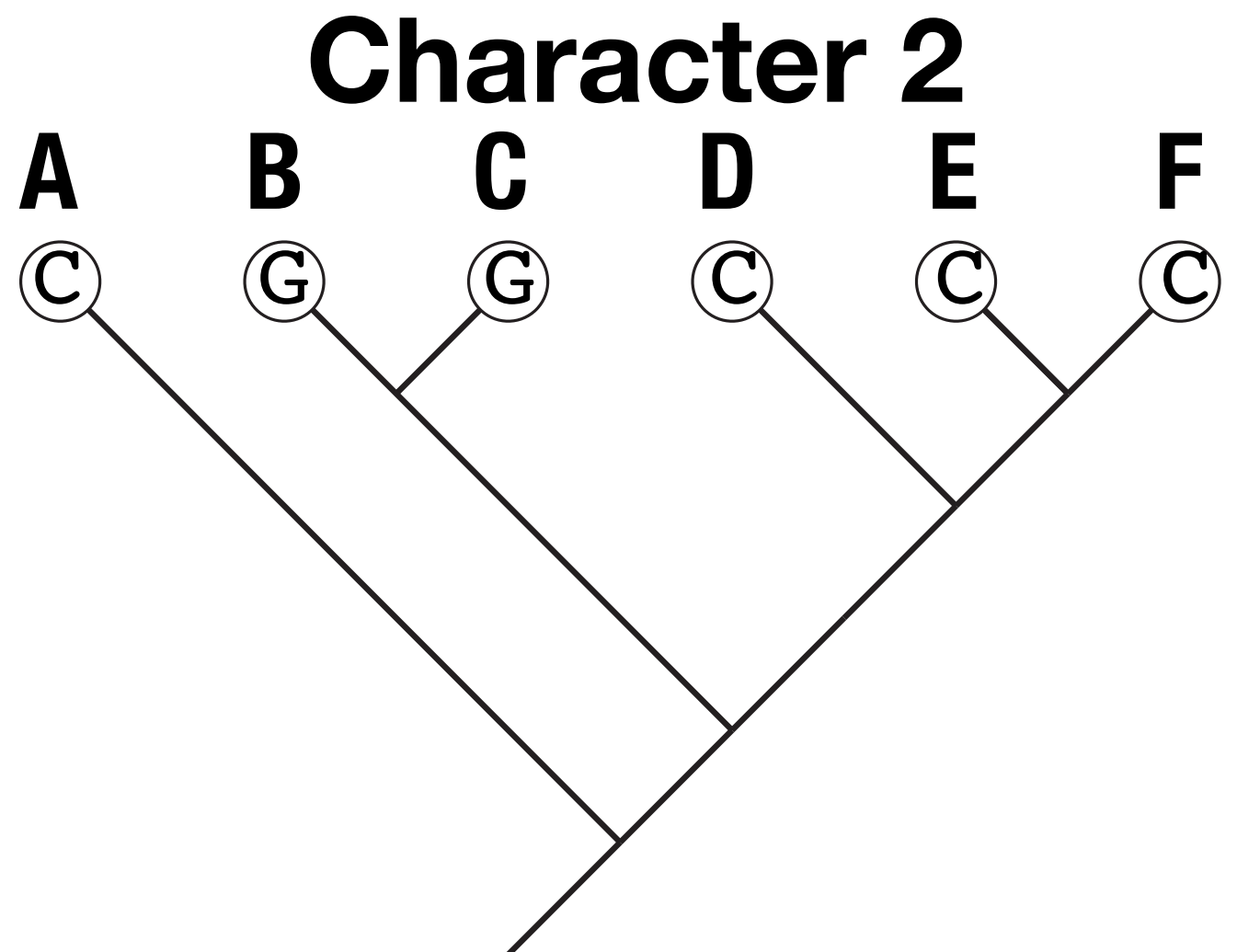
Taxa	Characters		
	1	2	3
	A	C	A
	A	G	A
	A	G	A
	A	C	A
	A	C	T
	A	C	T



Character matrix

A table of character states. Each row corresponds to a taxon, and each column to a homologous character.

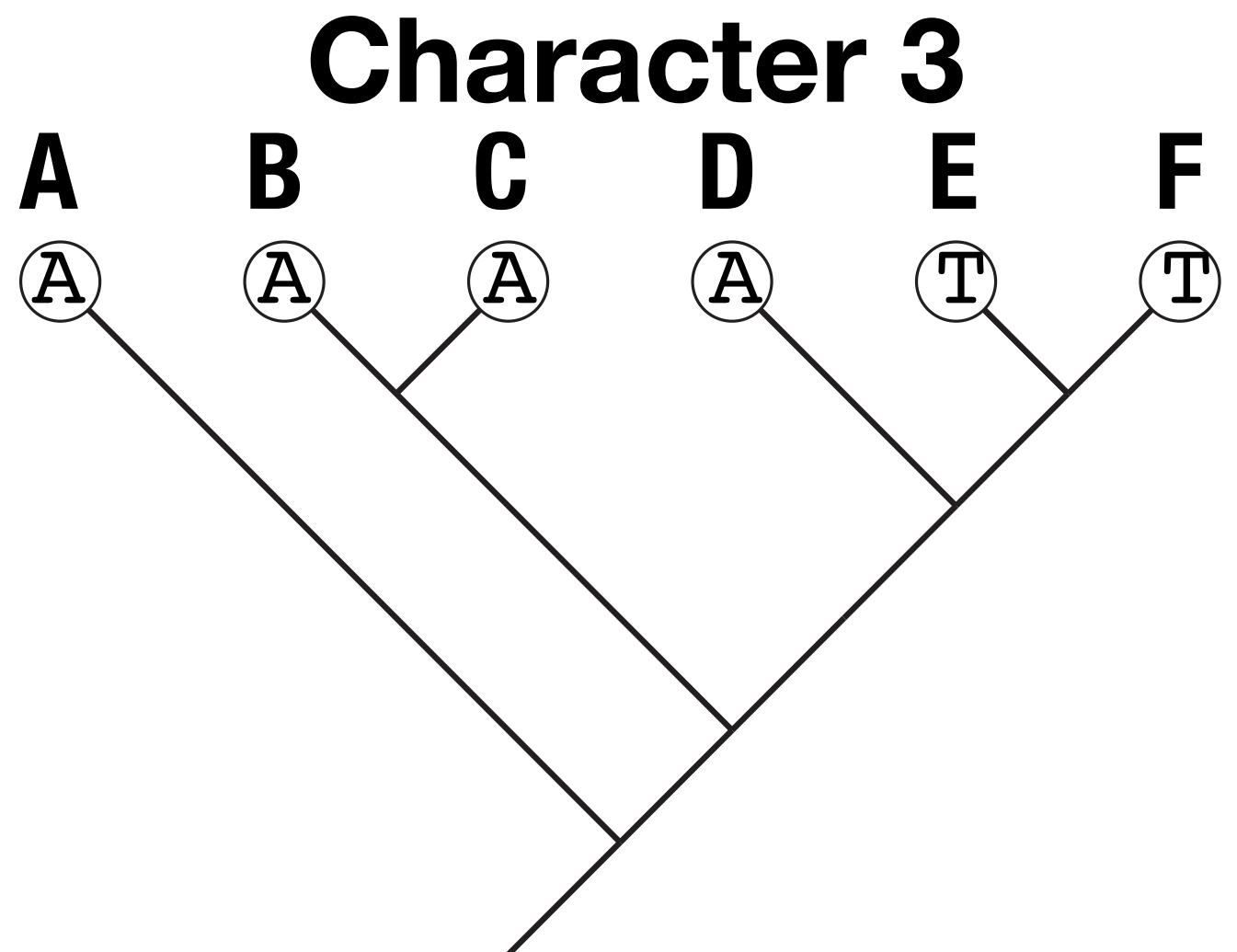
Taxa	Characters		
	1	2	3
	A	C	A
	B	G	A
	C	G	A
	D	C	A
	E	C	T
	F	C	T



Character matrix

A table of character states. Each row corresponds to a taxon, and each column to a homologous character.

Taxa	Characters		
	1	2	3
	A	C	A
	B	G	A
	C	G	A
	D	C	A
	E	C	T
	F	C	T



Phylogenetic Inference

In a typical phylogenetic study, we have a character matrix that describes the character states at the tips of the tree, and we want to know the phylogeny

In a typical phylogenetic study, we have a character matrix that describes the character states at the tips of the tree, and we want to know the phylogeny

Phylogenetic inference is the estimation of the phylogeny based on the character data and a model of character evolution

In a typical phylogenetic study, we have a character matrix that describes the character states at the tips of the tree, and we want to know the phylogeny

Phylogenetic inference is the estimation of the phylogeny based on the character data and a model of character evolution

In the most basic phylogenetic inference projects, we are looking for the single “best” phylogeny (later we’ll expand beyond this)

The model allows us to define an optimality criterion that, for each tree, answers the question, “**How well does this tree account for the observed character?**”

We then measure the optimality criterion for many trees, and pick the tree with the best value

There are different optimality criteria. These include:

Maximum parsimony - The best tree is the tree that minimizes homoplasy, i.e. the simplest (most parsimonious) explanation

Maximum likelihood - The best tree is the tree that maximizes the likelihood of observing the character matrix

How many trees are there?

$$N = \frac{(2t - 5)!}{2^{t-3} (t - 3)!}$$

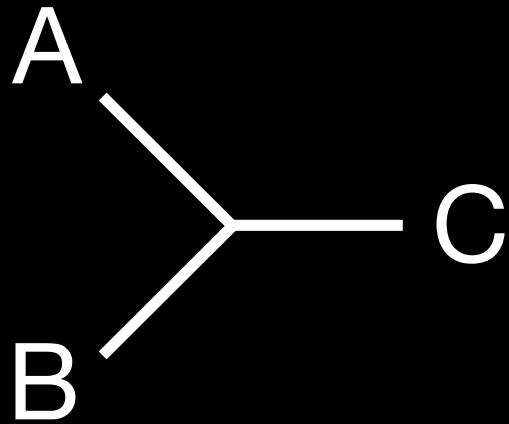
Where:

N is the number of trees

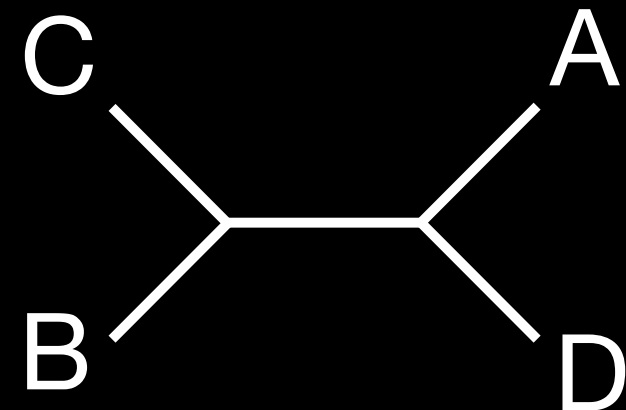
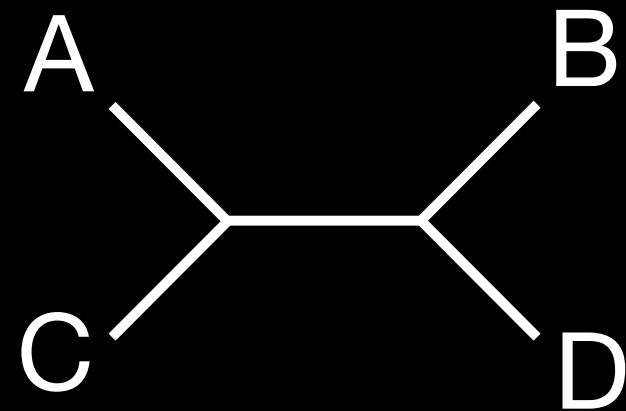
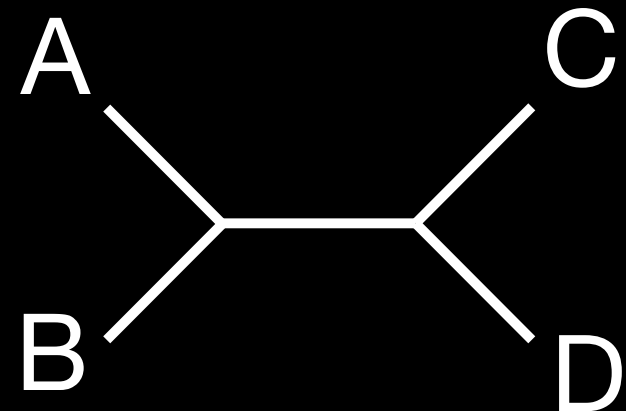
t is the number of taxa

How many trees are there?

$t = 3, N = 1:$



$t = 4, N = 3:$



How many trees are there?

t	N
3	1
4	3
5	15

How many trees are there?

t	N
3	1
4	3
5	15
8	10,395

How many trees are there?

t	N
3	1
4	3
5	15
8	10,395

← 1,470,000- The number of pixels on this slide

How many trees are there?

t	N
3	1
4	3
5	15
8	10,395
10	2,027,025
20	2.21×10^{20}

← 1,470,000- The number of pixels on this slide

How many trees are there?

t	N	
3	1	
4	3	
5	15	
8	10,395	← 1,470,000- The number of pixels on this slide
10	2,027,025	
20	2.21×10^{20}	← 6.02×10^{23} - Avagadro's number (1 mole)

How many trees are there?

t	N	
3	1	
4	3	
5	15	
8	10,395	
10	2,027,025	← 1,470,000- The number of pixels on this slide
20	2.21×10^{20}	
23	1.31×10^{25}	← 6.02×10^{23} - Avagadro's number (1 mole)

How many trees are there?

t	N	
3	1	
4	3	
5	15	
8	10,395	← 1,470,000- The number of pixels on this slide
10	2,027,025	
20	2.21×10^{20}	← 6.02×10^{23} - Avagadro's number (1 mole)
23	1.31×10^{25}	← 5×10^{27} - Molecules of gas in this room

How many trees are there?

t	N	
3	1	
4	3	
5	15	
8	10,395	
10	2,027,025	← 1,470,000- The number of pixels on this slide
20	2.21×10^{20}	
23	1.31×10^{25}	← 6.02×10^{23} - Avagadro's number (1 mole)
30	8.68×10^{36}	← 5×10^{27} - Molecules of gas in this room
40	1.31×10^{55}	
50	2.83×10^{74}	

How many trees are there?

t	N	
3	1	
4	3	
5	15	
8	10,395	
10	2,027,025	← 1,470,000- The number of pixels on this slide
20	2.21×10^{20}	
23	1.31×10^{25}	← 6.02×10^{23} - Avagadro's number (1 mole)
30	8.68×10^{36}	← 5×10^{27} - Molecules of gas in this room
40	1.31×10^{55}	
50	2.83×10^{74}	← 7×10^{79} - Number of atoms in the observable universe

How many trees are there?

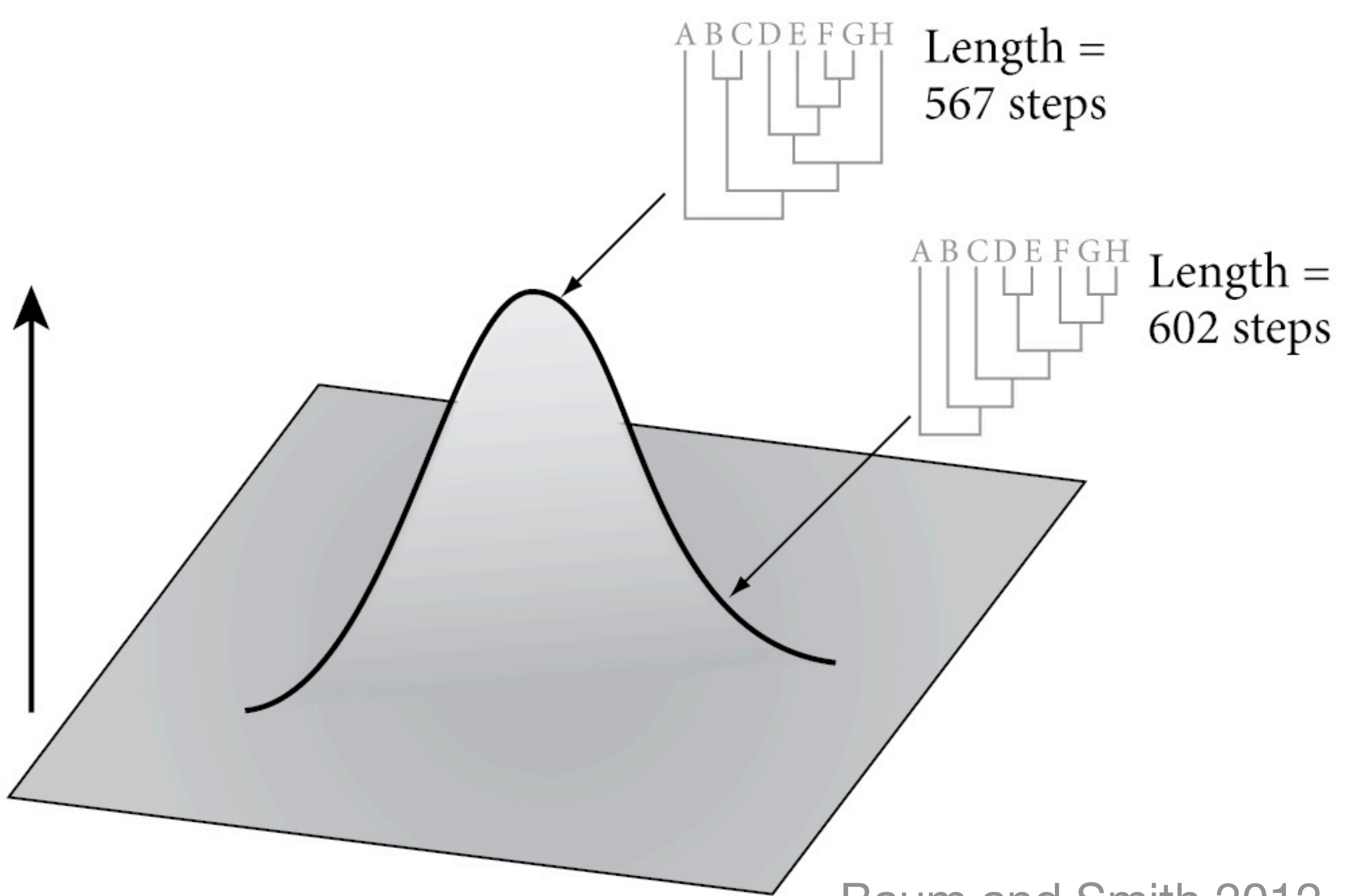
<i>t</i>	<i>N</i>	
3	1	
4	3	
5	15	
8	10,395	
10	2,027,025	← 1,470,000- The number of pixels on this slide
20	2.21×10^{20}	
23	1.31×10^{25}	← 6.02×10^{23} - Avagadro's number (1 mole)
30	8.68×10^{36}	← 5×10^{27} - Molecules of gas in this room
40	1.31×10^{55}	
50	2.83×10^{74}	
60	5.00×10^{94}	← 7×10^{79} - Number of atoms in the observable universe

How many trees are there?

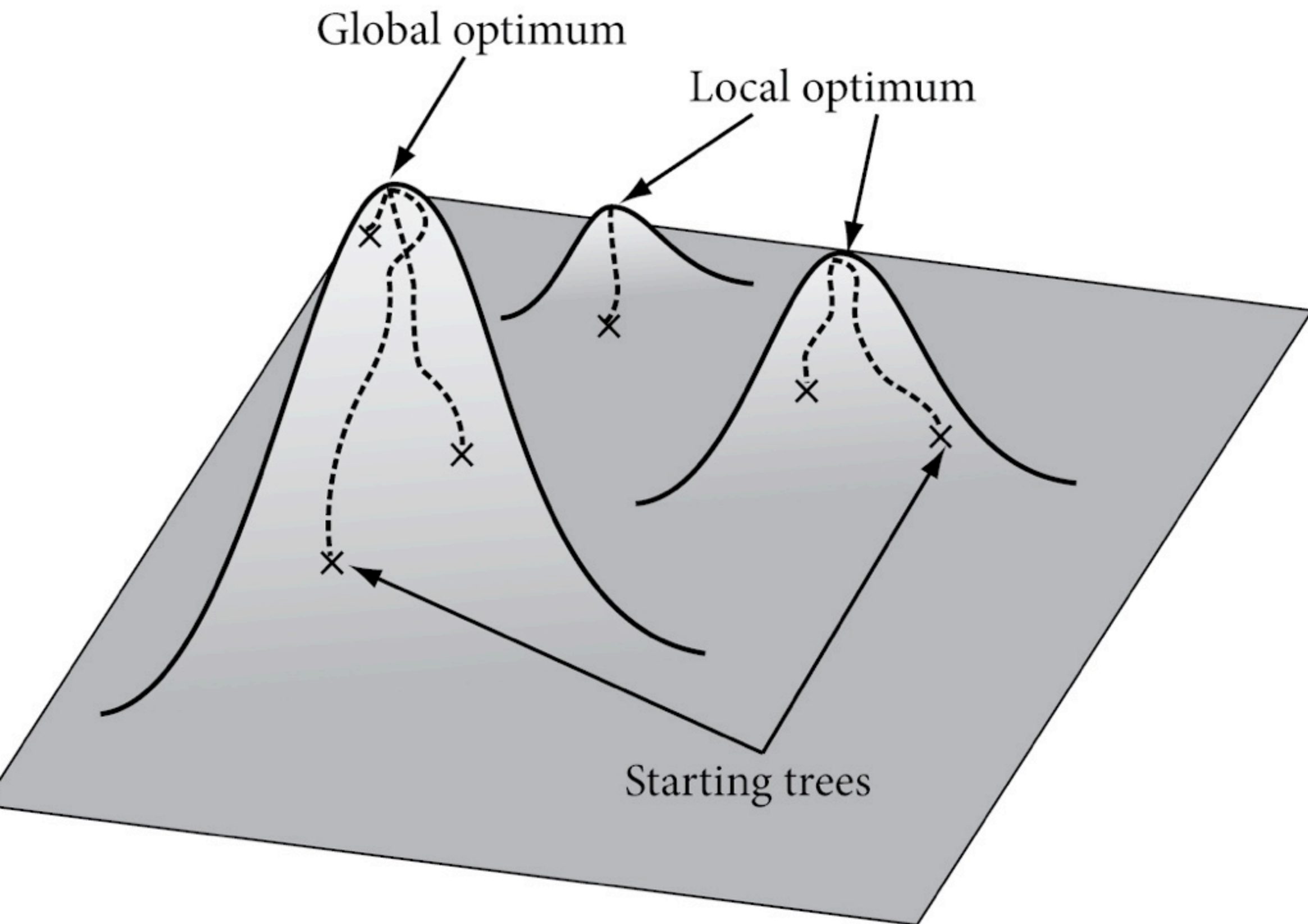
t	N	
3	1	
4	3	
5	15	
8	10,395	
10	2,027,025	← 1,470,000- The number of pixels on this slide
20	2.21×10^{20}	
23	1.31×10^{25}	← 6.02×10^{23} - Avagadro's number (1 mole)
30	8.68×10^{36}	← 5×10^{27} - Molecules of gas in this room
40	1.31×10^{55}	
50	2.83×10^{74}	
60	5.00×10^{94}	← 7×10^{79} - Number of atoms in the observable universe
		← 1×10^{100} - Googol

How many trees are there?

t	N	
3	1	
4	3	
5	15	
8	10,395	
10	2,027,025	← 1,470,000- The number of pixels on this slide
20	2.21×10^{20}	
23	1.31×10^{25}	← 6.02×10^{23} - Avagadro's number (1 mole)
30	8.68×10^{36}	← 5×10^{27} - Molecules of gas in this room
40	1.31×10^{55}	
50	2.83×10^{74}	
60	5.00×10^{94}	← 7×10^{79} - Number of atoms in the observable universe
		← 1×10^{100} - Googol
70	4.99×10^{115}	
	4.00×10^{120}	



Baum and Smith 2012,
Figure 7.8



Baum and Smith 2012,
Figure 7.9

Maximum parsimony

Maximum parsimony

1. For each tree, find the minimum number of steps (character changes) needed to explain the character data.
2. Find the tree with the fewest steps

An example:

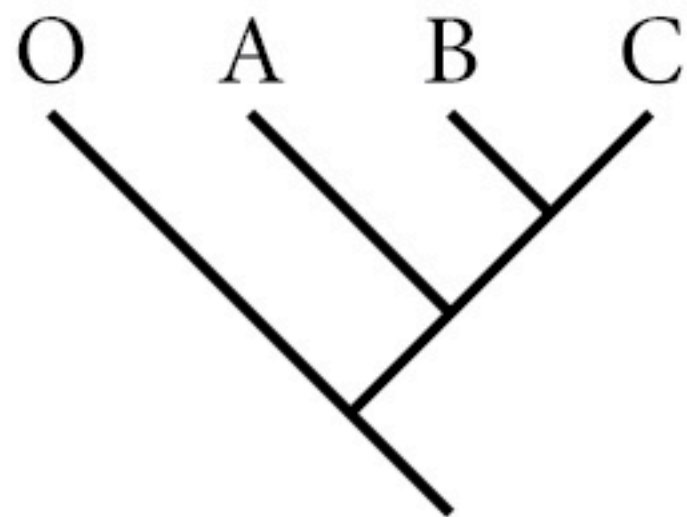
4 Taxa (O, A, B, C)

8 Characters

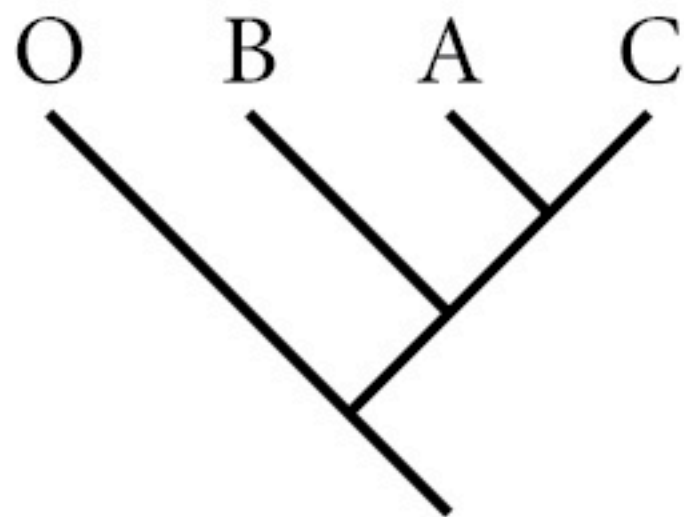
2 Character states for each character

	1	2	3	4	5	6	7	8
O	0	0	0	0	0	0	0	0
A	0	1	0	0	0	1	1	0
B	1	1	0	1	1	1	1	1
C	0	0	1	1	0	0	0	0

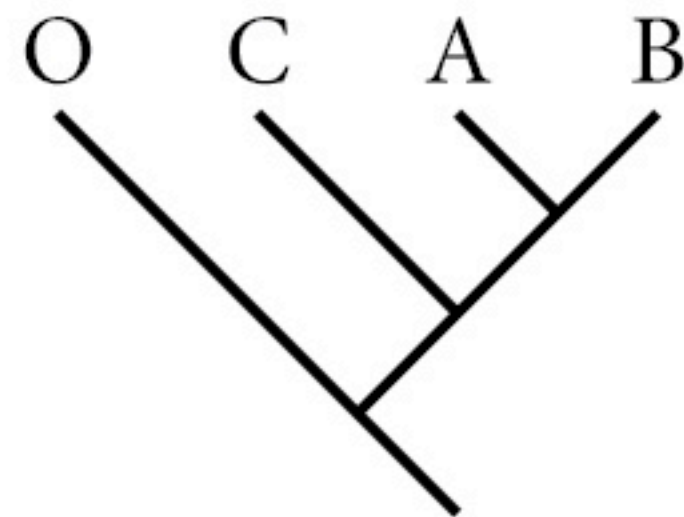
Baum and Smith 2012, Table 7.5



Tree 1



Tree 2



Tree 3

Baum and Smith 2012,
Figure 7.3

Other topics addressed

Rooted vs. unrooted trees

Outgroup/ ingroup

Matrix representation of models