

# 1.xi.

1. characters on trees
2. parsimony
3. apo- & plesiomorphy
4. outgroup
5. homology & homoplasy
6. character congruence
7. summary
8. home exercise

n                      B(n)  
-----

3

1

4

3

5

15

6

105

7

945

8

10 395

9

135 135

10

$\sim 2 \times 10^6$

15

$\sim 8 \times 10^{12}$

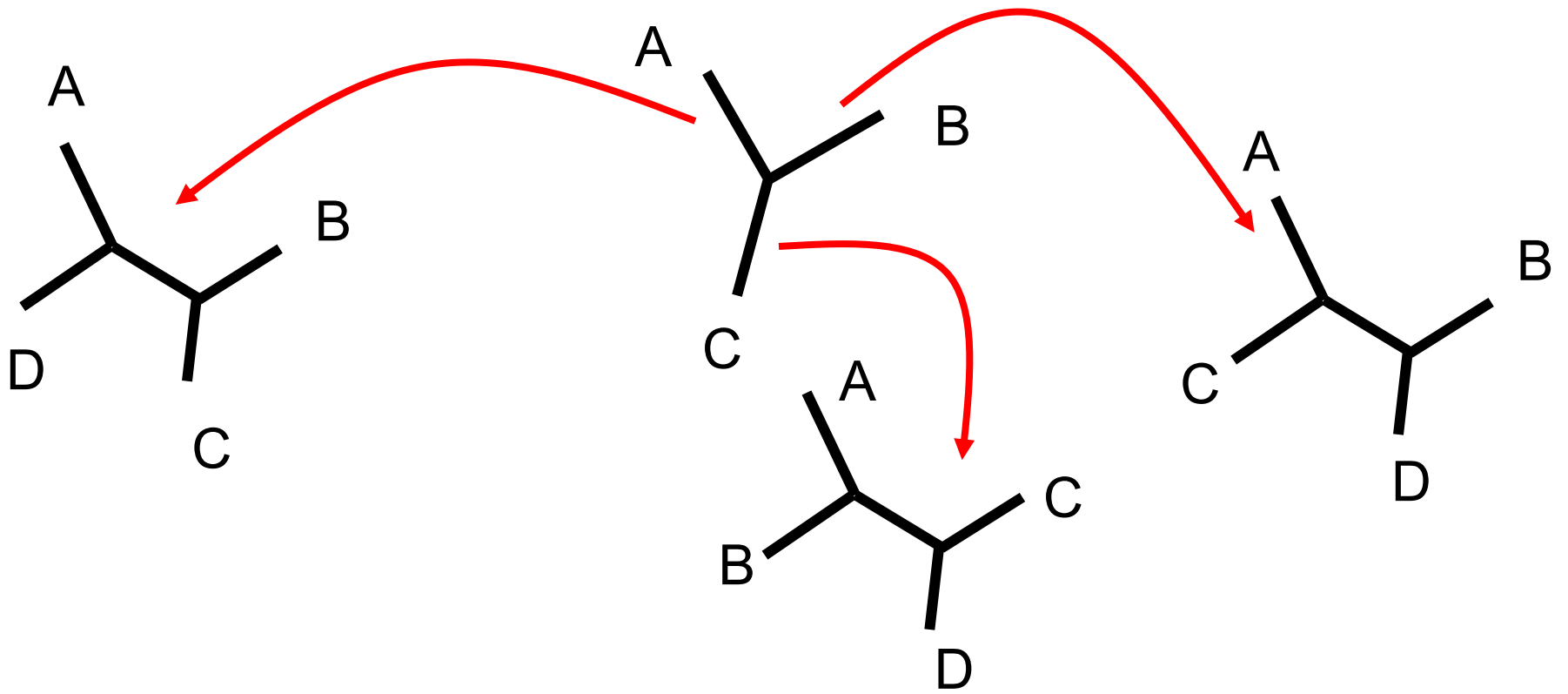
20

$\sim 2 \times 10^{20}$

50

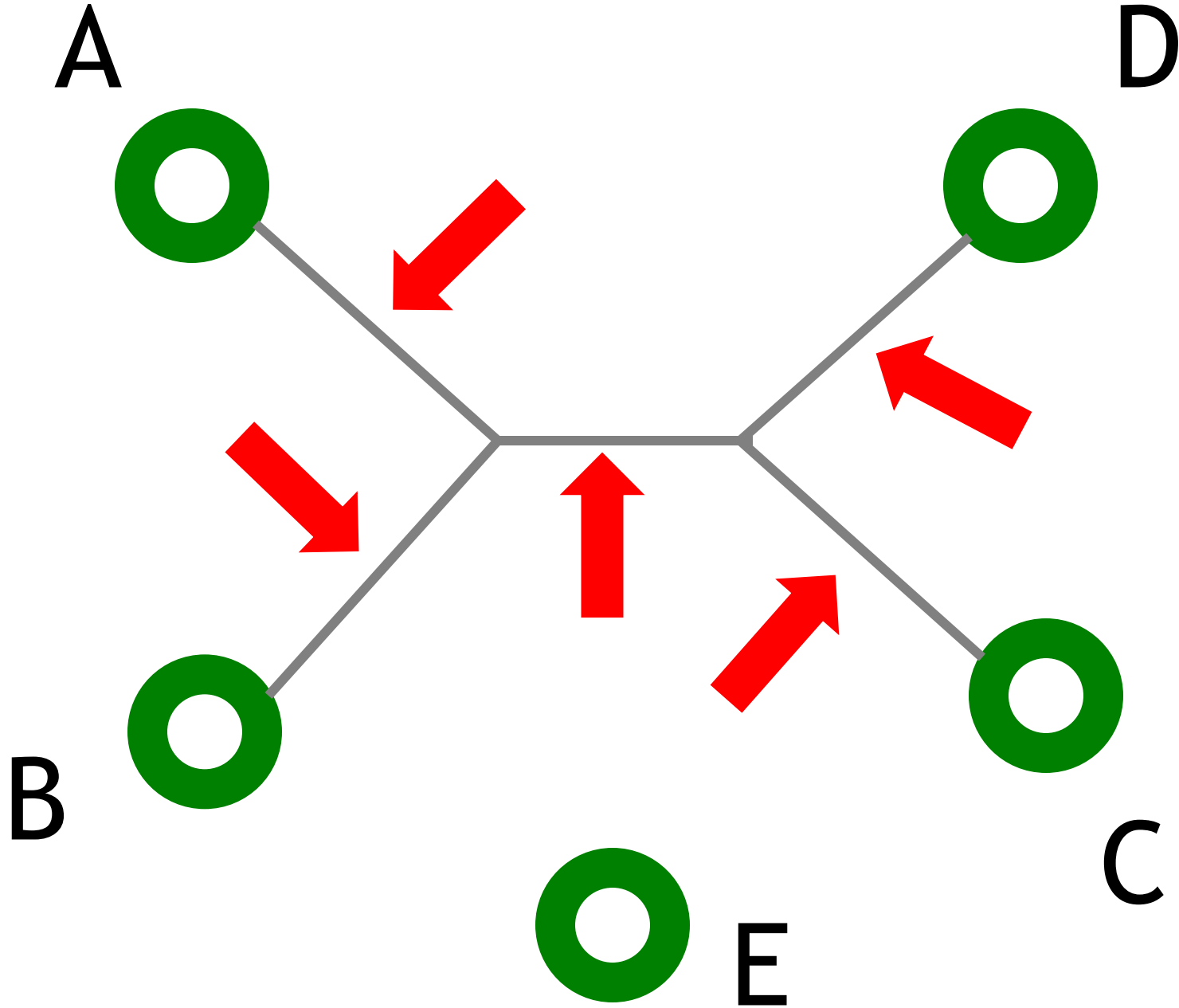
$\sim 3 \times 10^{74}$

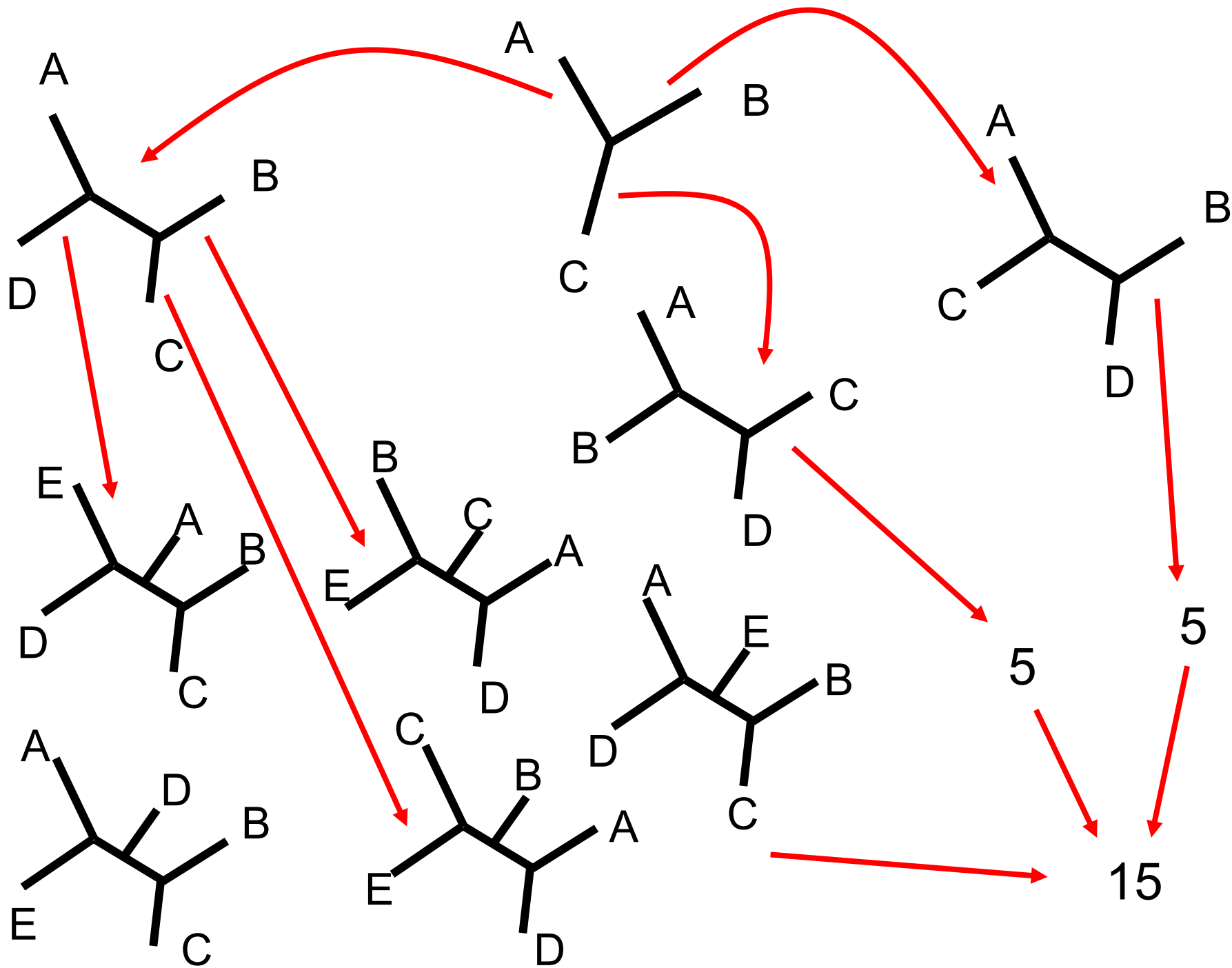
number of possible  
UNdirected trees



3 BRANCHES > 3 DIFFERENT PLACES TO ADD 4<sup>TH</sup> TERMINAL

5 BRANCHES > 5 DIFFERENT PLACES TO ADD 5<sup>TH</sup> TERMINAL












n      branches      B(n)

-----

3	3	x	1
4	5	x	3
5	7	x	15
6	9	x	105
7	11	x	945
8	13	x	10 395
9	15	x	135 135
10			$\sim 2 \times 10^6$
15			$\sim 8 \times 10^{12}$
20			$\sim 2 \times 10^{20}$
50			$\sim 3 \times 10^{74}$

n      branches      B(n)

-----

3	3	x	1		
4	5	x	3		
5	7	x	15		
6	9	x	105		
7	11	x	945		
8	13	x	10 395		
9	15	x	135 135		
10			$\sim 2 \times 10^6$		
15			$\sim 8 \times 10^{12}$		
20			$\sim 2 \times 10^{20}$		
50			$\sim 3 \times 10^{74}$		

characters

0000000001  
1234567890

terminals

<b>A</b>	0010001010
<b>B</b>	0101010000
<b>C</b>	0111000001
<b>D</b>	1010110001
<b>E</b>	0000110101



characters

0000000001

1234567890

terminals

**A**

0010001010

**B**

0101010000

**C**

0111000001



**D**

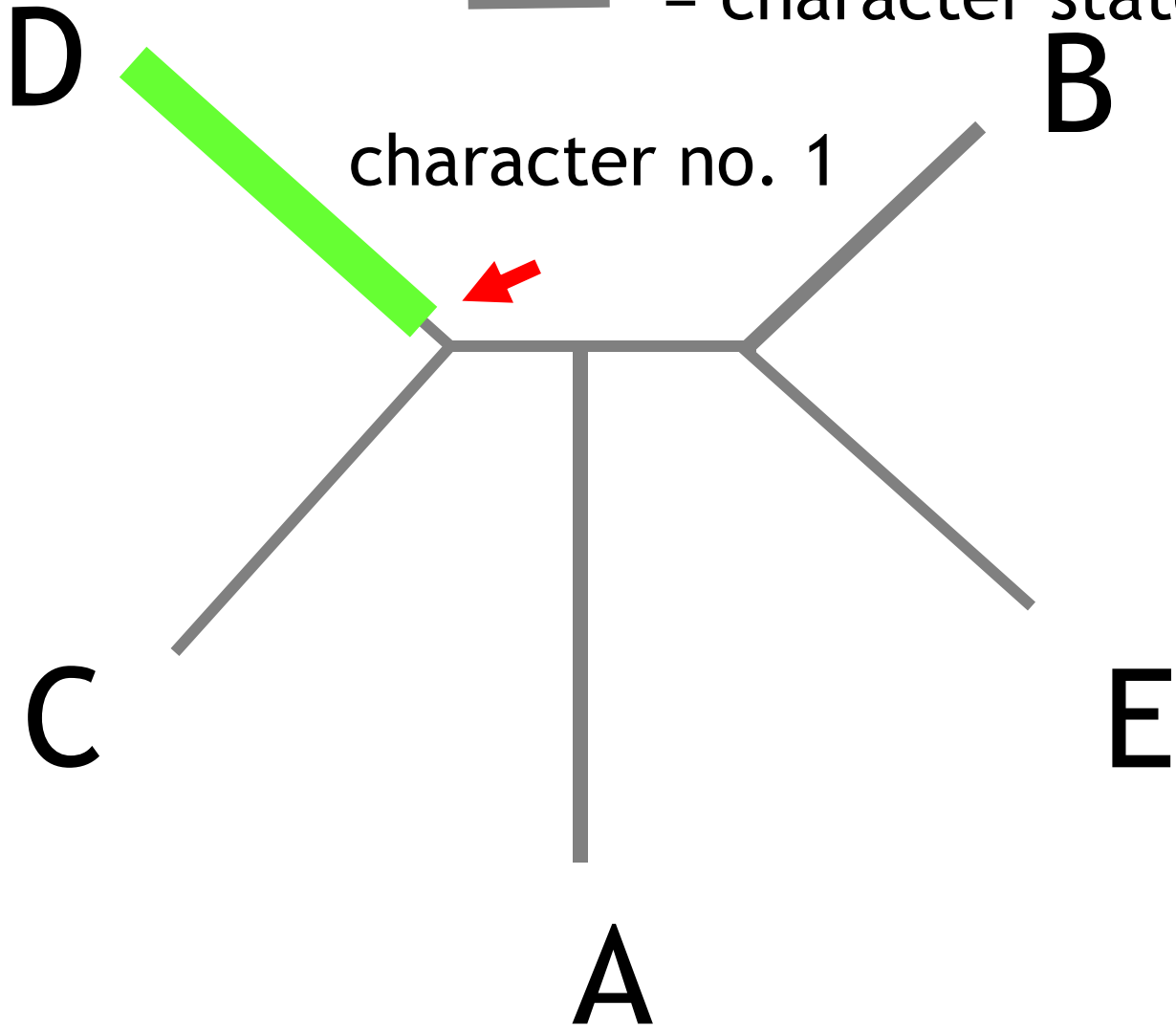




1010110001

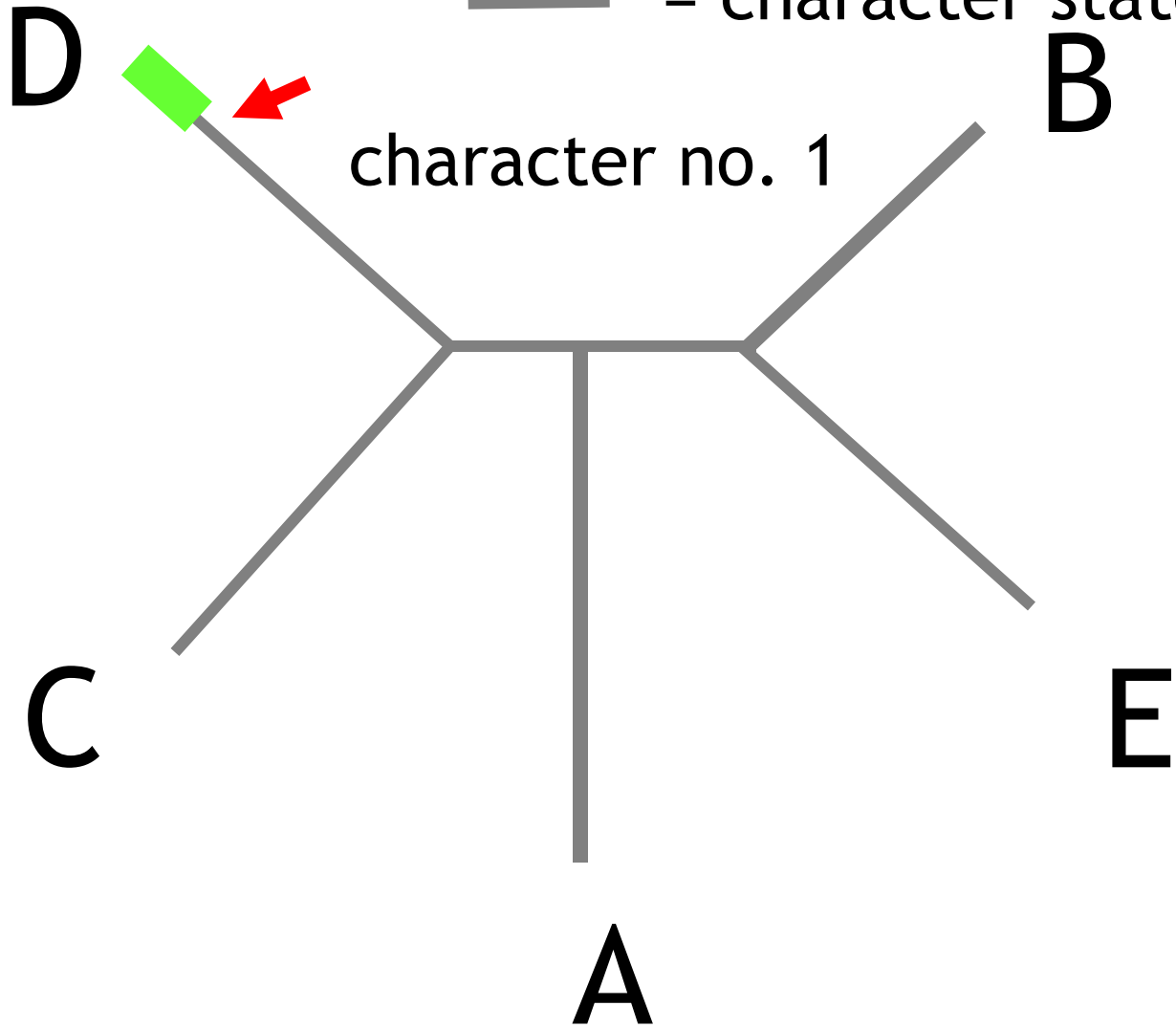
**E**

0000110101

 = character state 1  
 = character state 0



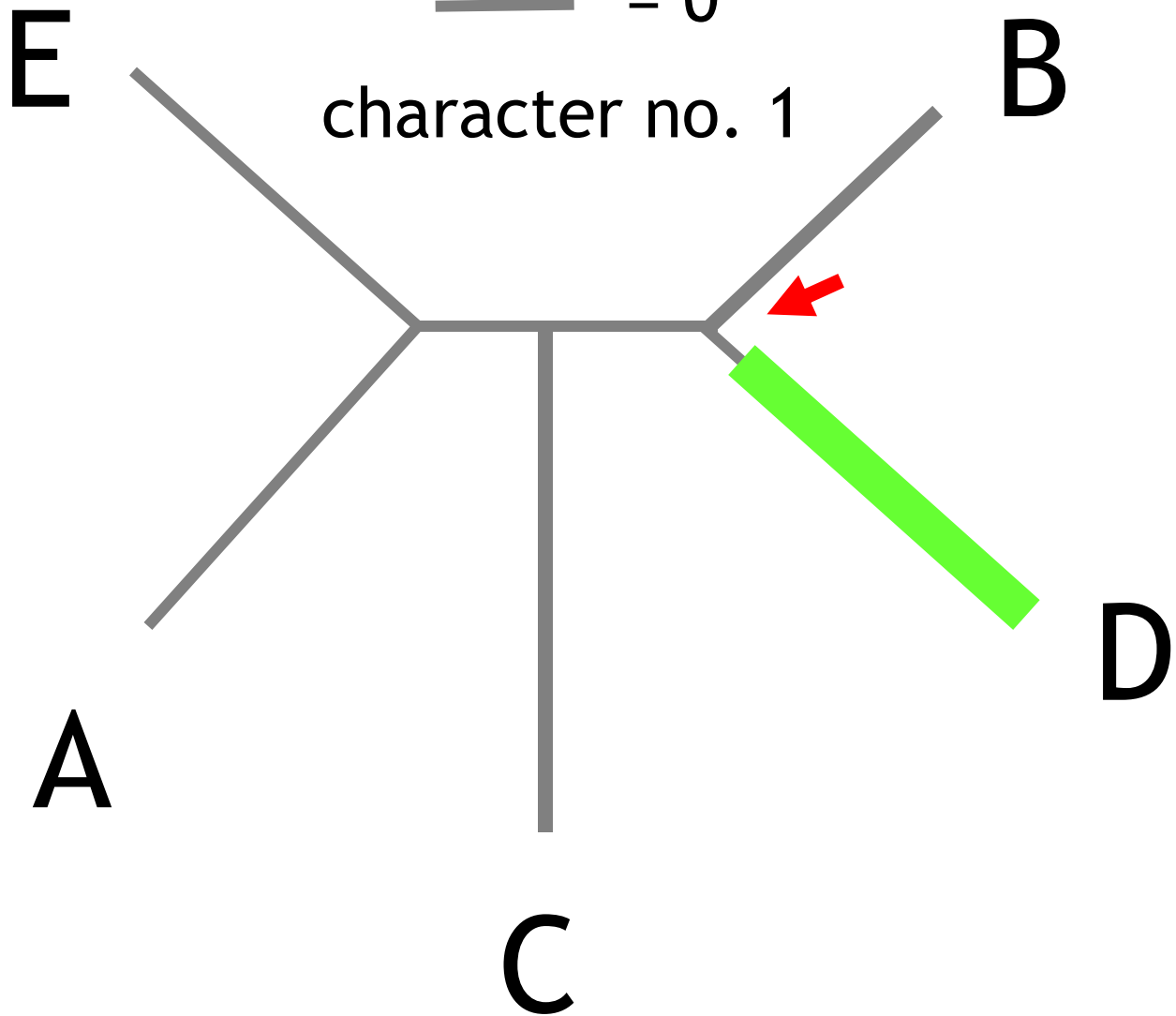
 = character state 1  
 = character state 0

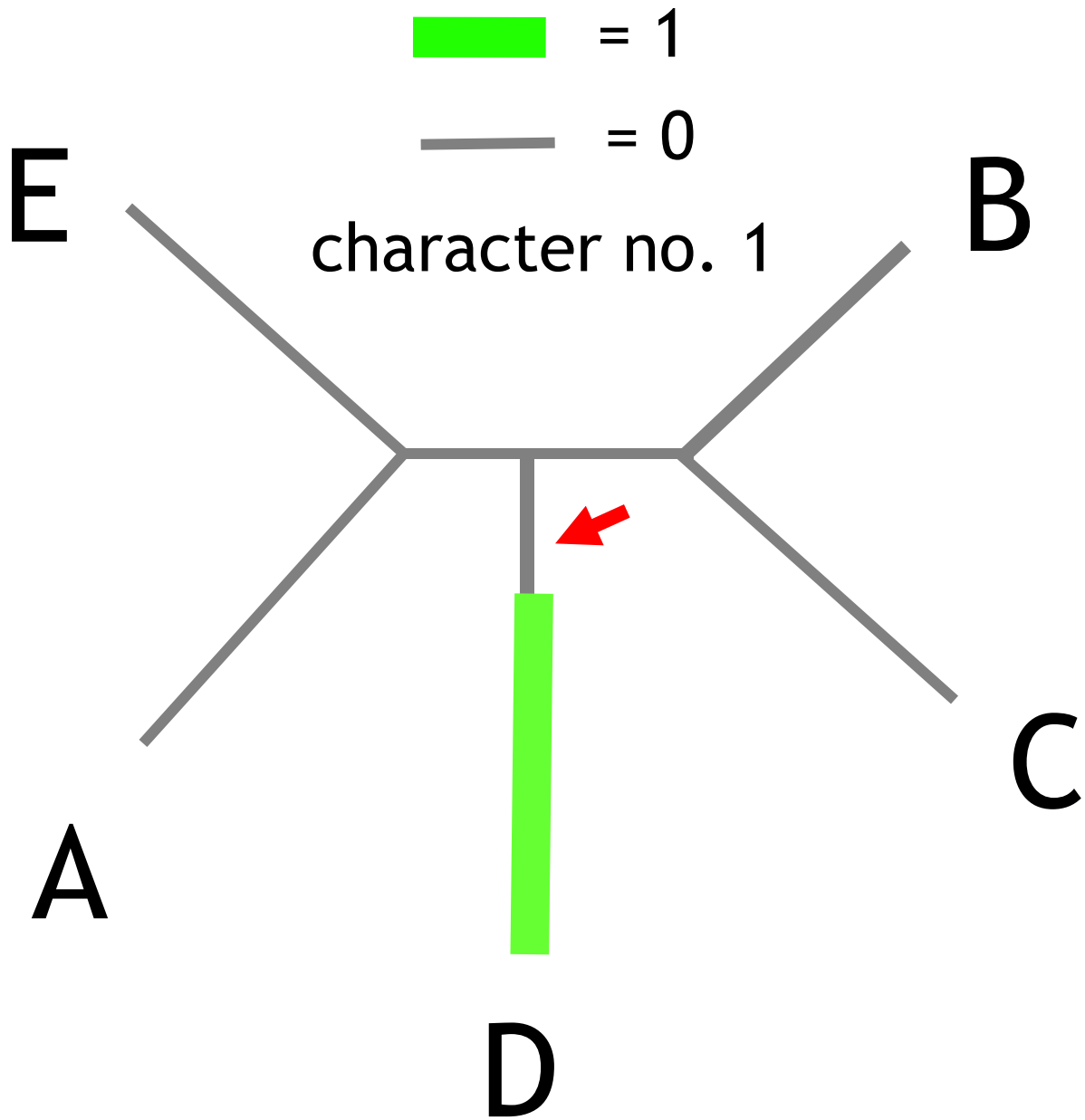


 = 1

 = 0

character no. 1





characters

0000000001

1234567890

characters with only 1 terminal differing from all the others

terminals

**A**

0010001010

**B**

0101010000

these characters can be IGNORED because they are NOT informative about phylogeny

1110000001

0101100001

**E**

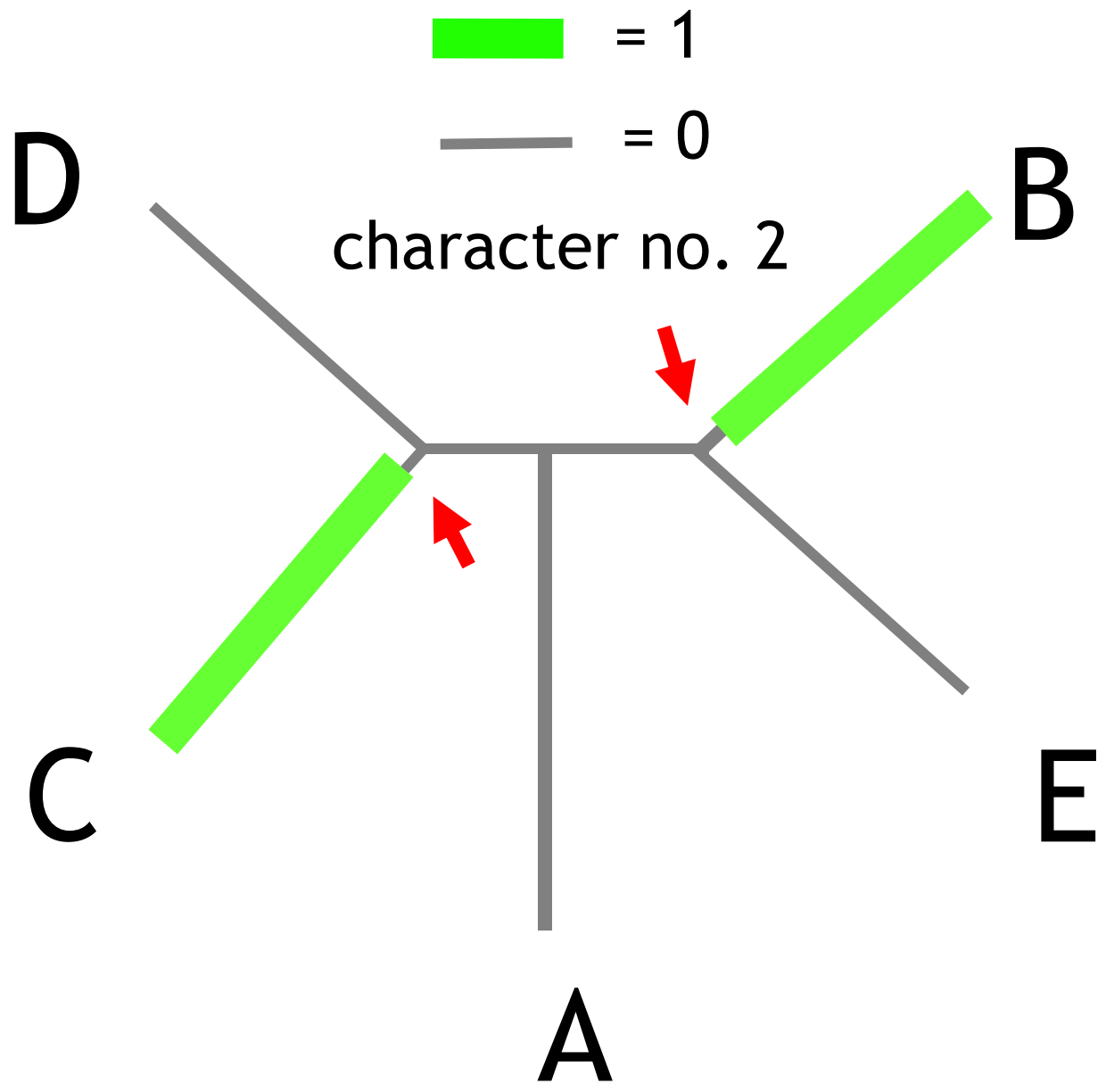
0000110101

characters

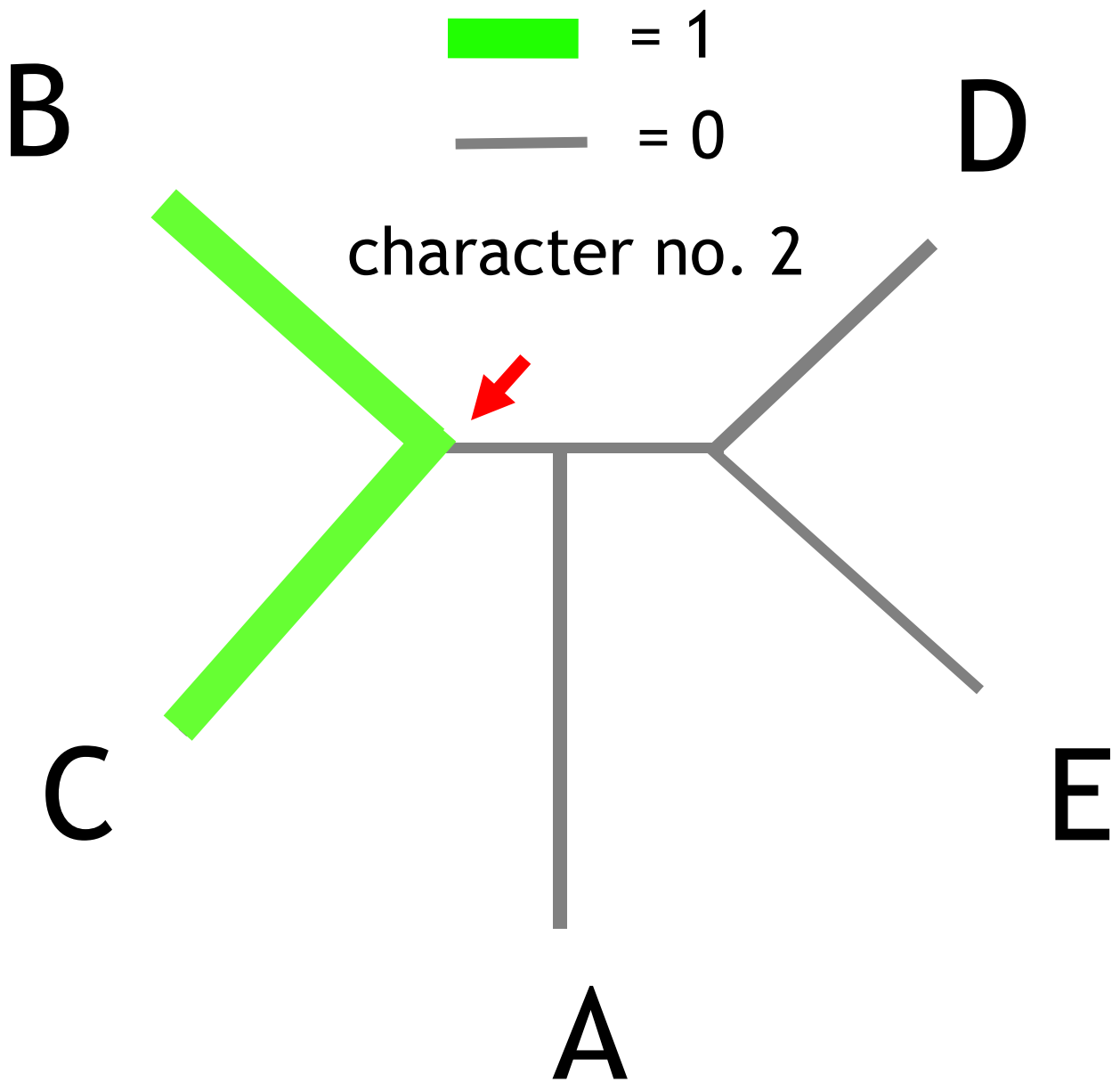
000000001  
1234567890

terminals

<b>A</b>	0010001010
<b>B</b>	0101010000
<b>C</b>	0111000001
<b>D</b>	1010110001
<b>E</b>	0000110101







characters

0000000001

1234567890

terminals

**A**

0010001010

**B**

0101010000

**C**

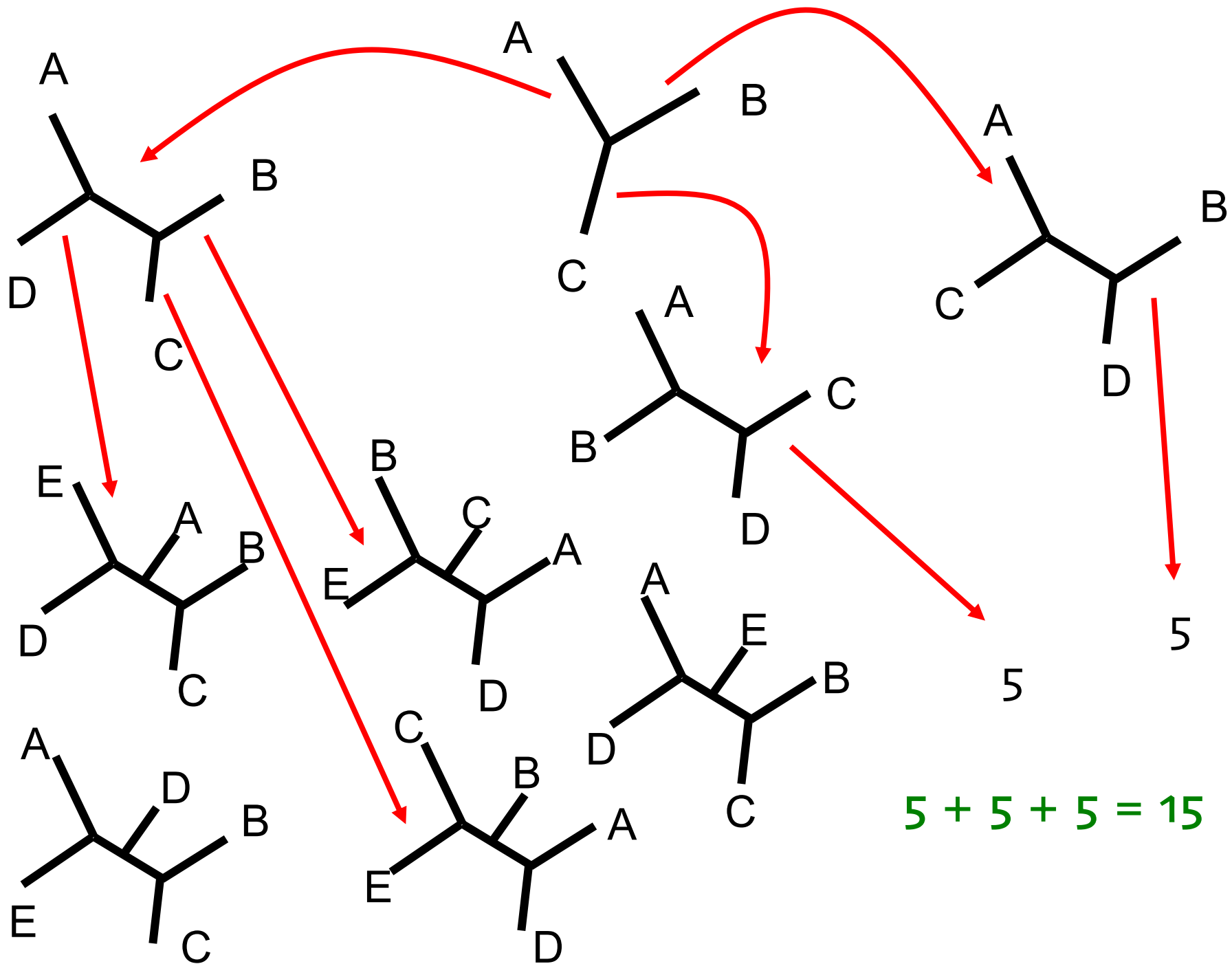
0111000001

**D**

1010110001

**E**

0000110101



characters

0000000001

1234567890

terminals

**A**

0010001010

**B**

0101010000

**C**

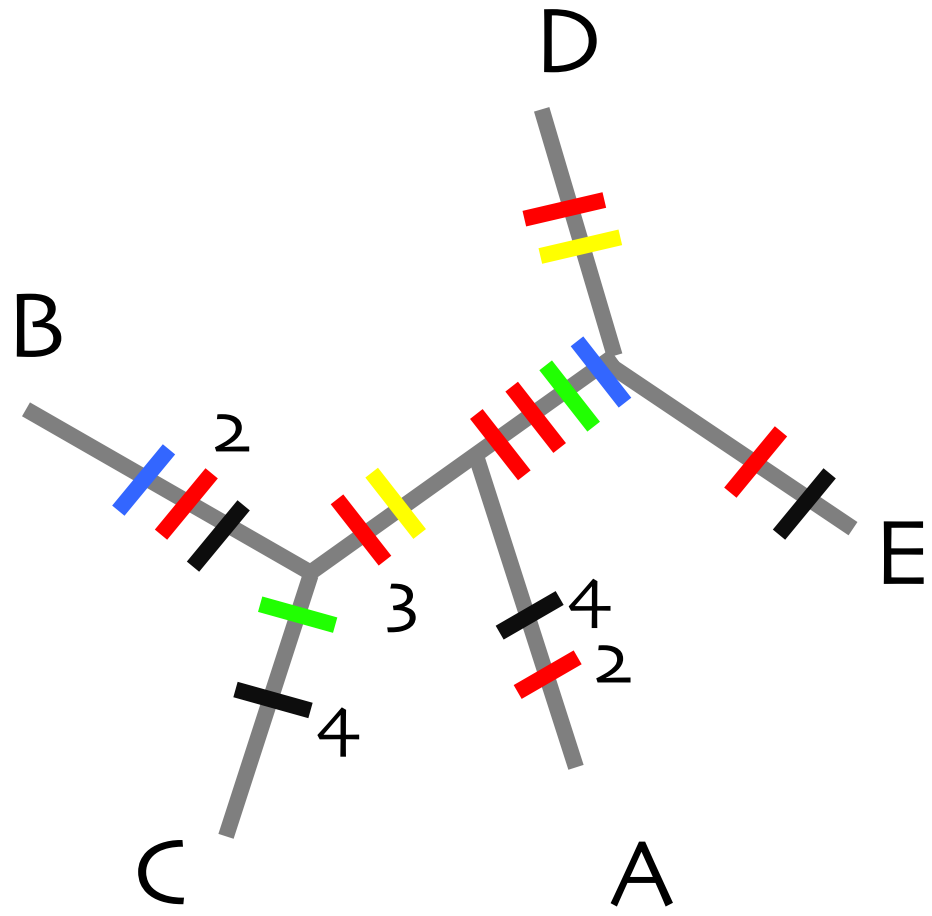
0111000001

**D**

1010110001

**E**

0000110101



characters		0	0	0	0	0	0	0	0	0	1	$\Sigma$
		1	2	3	4	5	6	7	8	9	0	
tree	1	1	1	2	1	2	2	1	1	1	2	14
	2	1	1	2	1	1	2	1	1	1	2	13
	3											
	4											
	5											
	6											
	7											
	8											
	9											
	10											
	11											
	12											
	13											
	14											
	15											

SUM of ch. state changes

EVALUATION of all possible trees feasible only for a SMALL number of trees

characters

trees	0	0	0	0	0	0	0	0	0	1	1	1	$\Sigma$
	1	2	3	4	5	6	7	8	9	0	1	2	
1	1	1	2	2	2	2	2	1	2	2	1	2	20
2	1	1	2	2	2	2	2	1	2	2	2	2	21
3	1	1	2	2	2	2	2	1	1	2	2	2	20
4	1	2	2	1	2	2	2	1	2	2	2	2	21
5	1	2	2	1	2	2	2	1	2	2	2	1	20
6	1	2	2	1	2	2	2	1	2	2	2	2	21
7	1	2	2	2	2	2	2	1	2	1	1	2	20
8	1	2	1	2	2	1	2	1	2	1	2	2	19
9	1	2	2	2	2	2	2	1	2	1	2	1	20
10	1	2	2	2	2	2	2	1	1	2	2	1	20
11	1	2	2	2	2	2	2	1	2	2	1	2	21
12	1	2	1	2	2	1	2	1	2	2	2	2	20
13	1	2	2	2	1	2	1	1	1	2	2	2	19
14	1	2	2	2	1	2	1	1	2	2	2	2	20
15	1	2	1	2	1	1	1	1	2	2	2	2	18

	characters													
	0	0	0	0	0	0	0	0	0	0	1	1	1	
trees	1	2	3	4	5	6	7	8	9	0	1	2	$\Sigma$	
1	1	1	2	2	2	2	2	1	2	2	1	2	20	
2	1	1	2	2	2	2	2	1	2	2	2	2	21	
3	1	1	2	2	2	2	2	1	1	2	2	2	20	
4	1	2	2	1	2	2	2	1	2	2	2	2	21	
5	1	2	2	1	2	2	2	1	2	2	2	1	20	
6	1	2	2	1	2	2	2	1	2	2	2	2	21	
7	1	2	2	2	2	2	2	1	2	1	1	2	20	

**BEST HYPOTHESIS**

- smallest number of changes from one character state to another
- largest part of resemblance between terminals explained by their shared HISTORY (descent from common ancestor)
- EXPLANATORY POWER of the shared history MAXIMIZED

14	1	2	2	2	1	2	1	1	2	2	2	2	20
15	1	2	1	2	1	1	1	1	2	2	2	2	18



characters

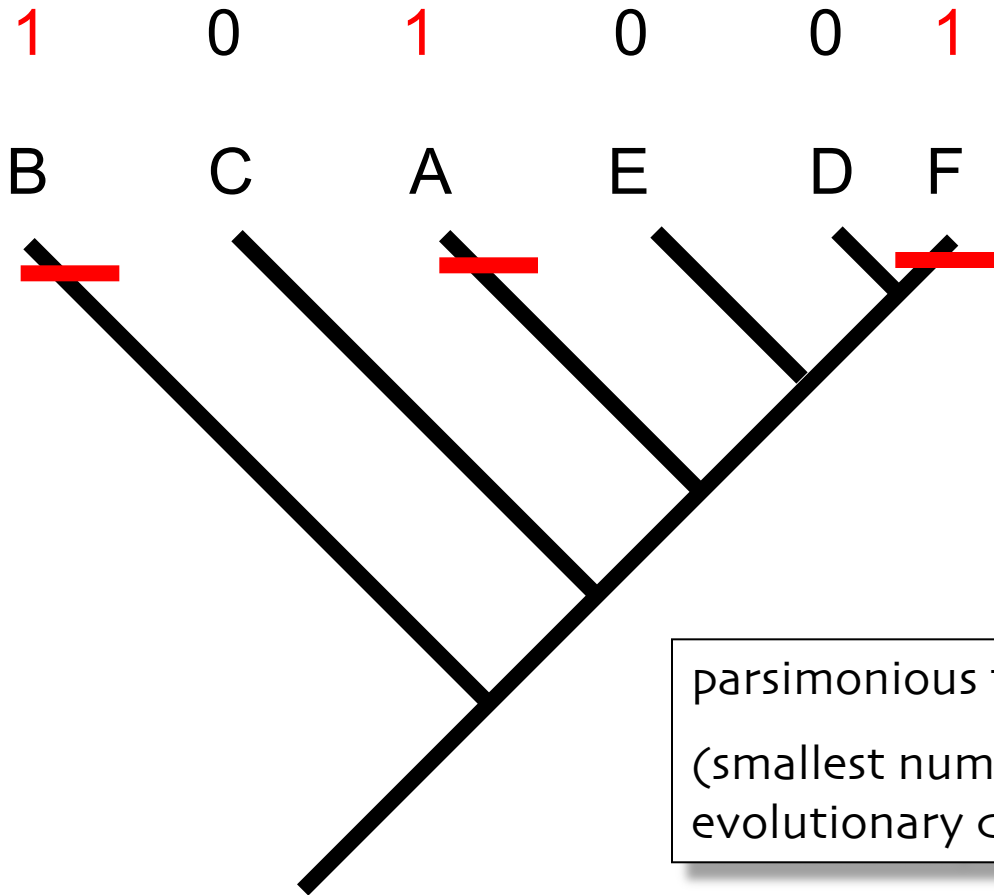
trees	0	0	0	0	0	0	0	0	0	0	1	1	1	$\Sigma$
	1	2	3	4	5	6	7	8	9	0	1	2		
1	1	1	2	2	2	2	2	1	2	2	1	2	20	
2	1	1	2	2	2	2	2	1	2	2	2	2	21	
3	1	1	2	2	2	2	2	1	1	2	2	2	20	
4	1	2	2	1	2	2	2	1	2	2	2	2	21	
5	1	2	2	1	2	2	2	1	2	2	2	1	20	
6	1	2	2	1	2	2	2	1	2	2	2	2	21	
7	1	2	2	2	2	2	2	1	2	2	2	2	21	
8	1	2	1	2	2	2	2	1	2	2	2	2	21	
9	1	2	2	2	2	2	2	1	2	1	2	1	20	
10	1	2	2	2	2	2	2	1	1	2	2	1	20	
11	1	2	2	2	2	2	2	1	2	2	1	2	21	
12	1	2	1	2	2	1	2	1	2	2	2	2	20	
13	1	2	2	2	1	2	1	1	1	2	2	2	19	
14	1	2	2	2	1	2	1	1	2	2	2	2	20	
15	1	2	1	2	1	1	1	1	2	2	2	2	18	

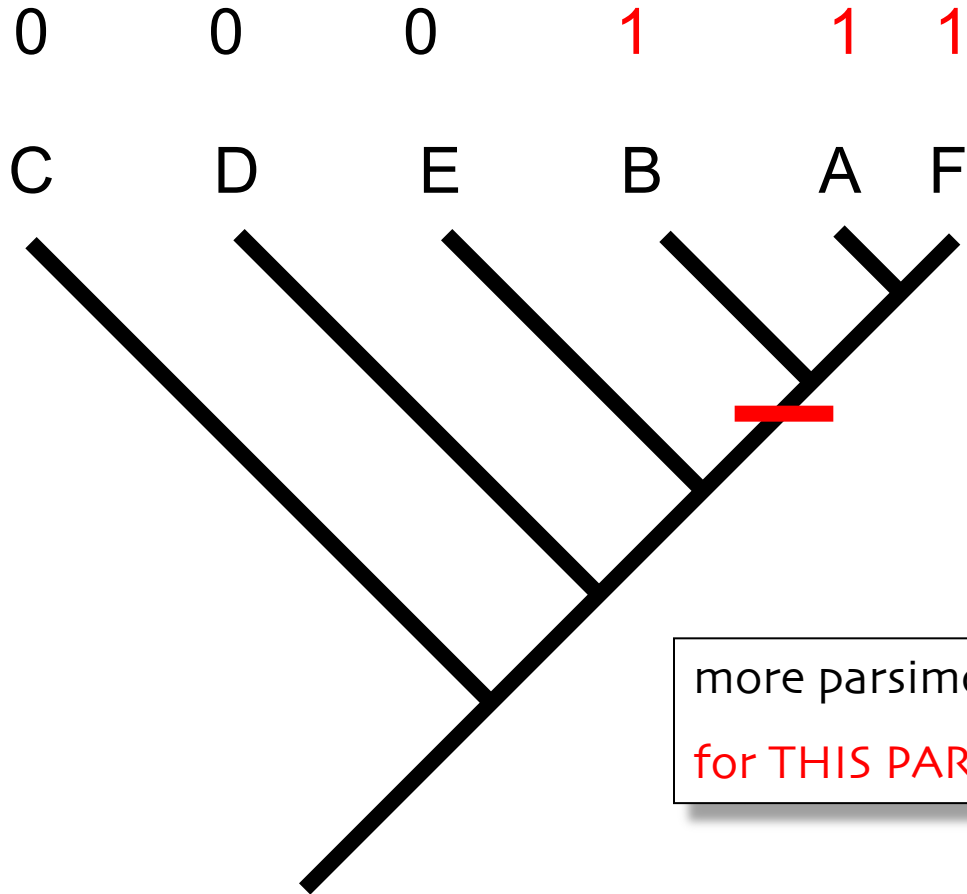
number of character states = n

MINIMUM number of evolutionary changes = n-1

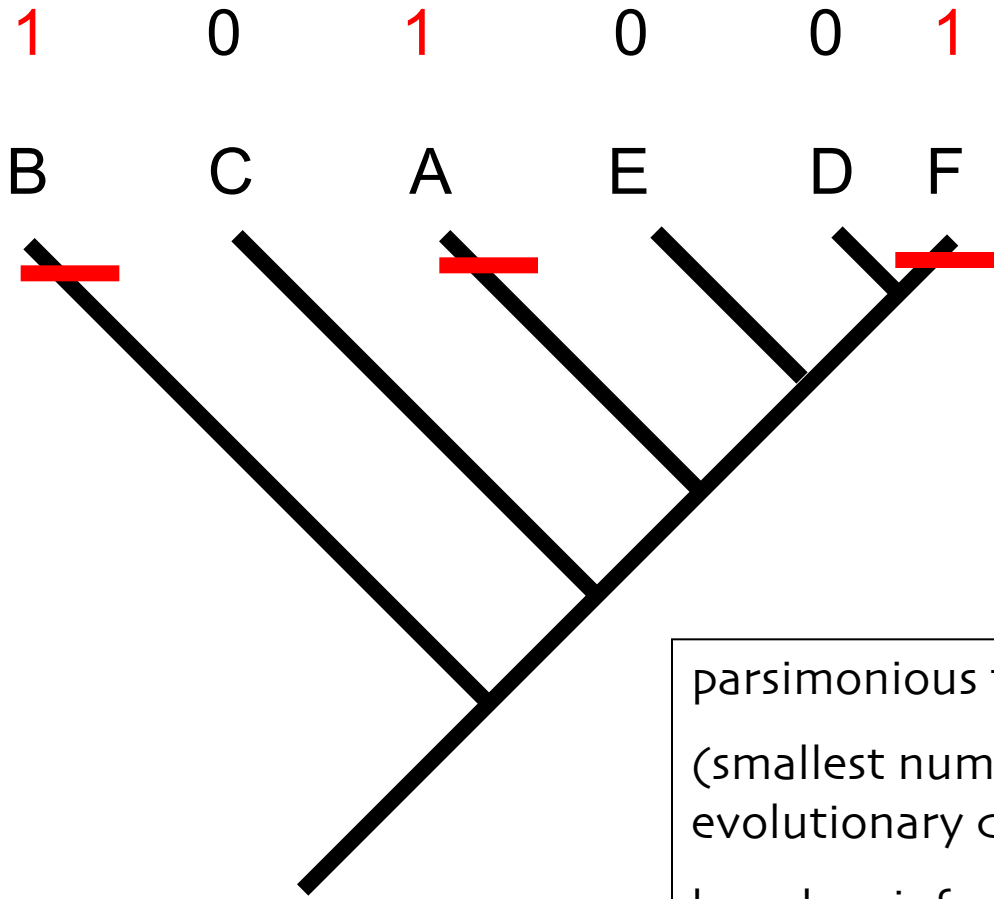
characters

trees	0	0	0	0	0	0	0	0	0	0	1	1	1	$\Sigma$
	1	2	3	4	5	6	7	8	9	0	1	2		
1	1	1	2	2	2	2	2	1	2	2	1	2	20	
2	1	1	2	2	2	2	2	1	2	2	2	2	21	
3	1	1	2	2	2	2	2	1	1	2	2	2	20	
4	1	2	2	1	2	2	2	1	2	2	2	2	21	
5	1	2	2	1	2	2	2	1	2	2	2	1	20	
6	1	2	2	1	2	2	2	1	2	2	2	2	21	
7	1	number of character states = n											20	
8	1	MINIMUM number of evolutionary changes = n-1											19	
9	1	2	2	2	2	2	2	1	2	1	2	1	20	
10	1	2	2	2	2	2	2	1	1	2	2	1	20	
11	1	2	2	2	2	2	2	1	2	2	1	2	21	
12	1	2	1	2	2	1	2	1	2	2	2	2	20	
13	1	2	2	2	1	2	1	1	1	2	2	2	19	
14	1	2	2	2	1	2	1	1	2	2	2	2	20	
15	1	2	1	2	1	1	1	1	2	2	2	2	18	





more parsimonious tree  
for THIS PARTICULAR character



parsimonious tree  
(smallest number of evolutionary changes)  
based on info from **ALL** characters

# PARSIMONY

PLURALITAS NON EST PONENDA  
SINE NECESSITATE

PLURALITY should NOT be assumed BEYOND NECESSITY

William of Ockham

SIMPLE hypotheses are desirable

# PARSIMONY

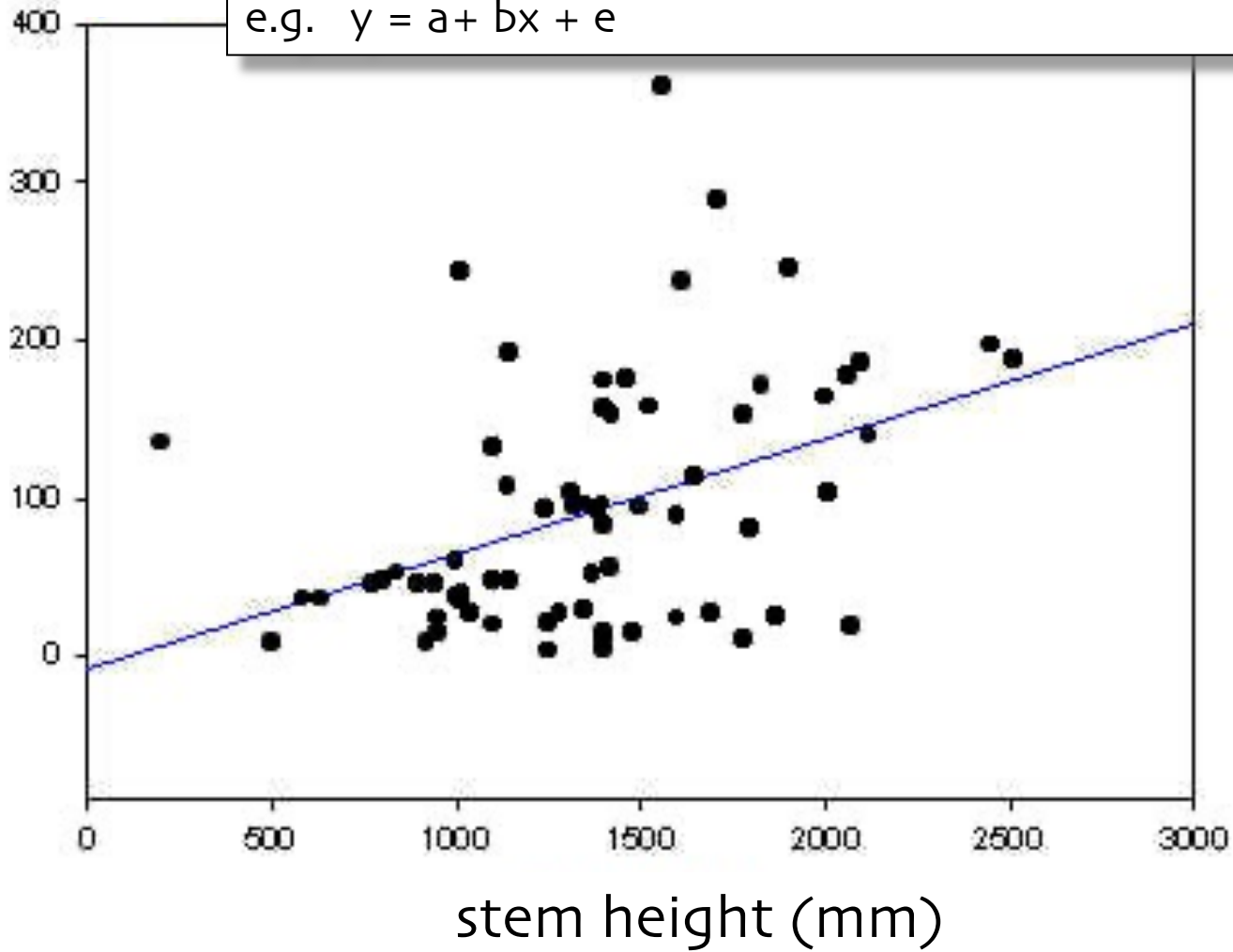


it has been erroneously claimed that use of parsimony as optimality criterion requires unrealistic assumption that evolution & evolutionary history are simple

line represents dependence between two variables

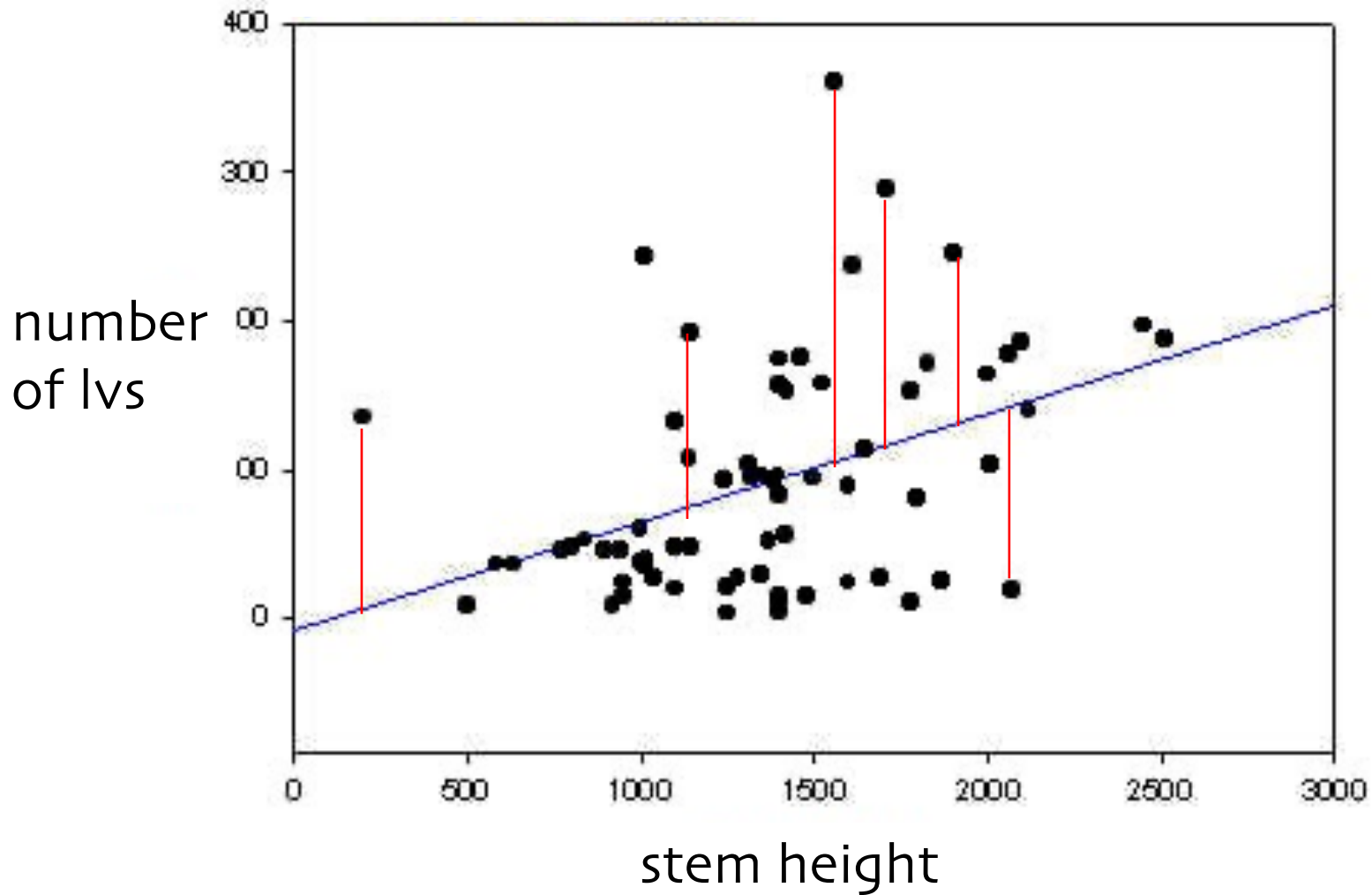
e.g.  $y = a + bx + e$

number  
of lvs



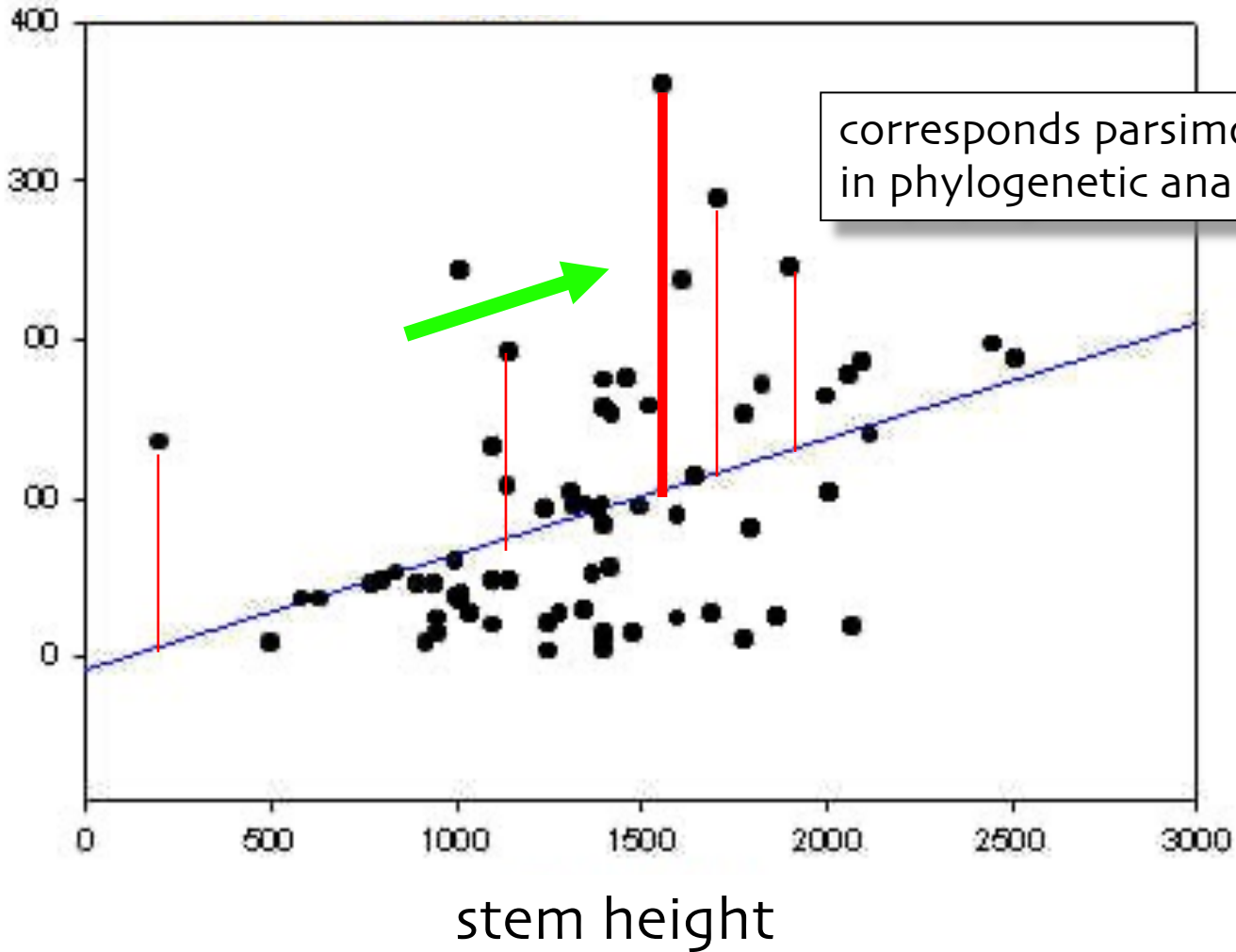


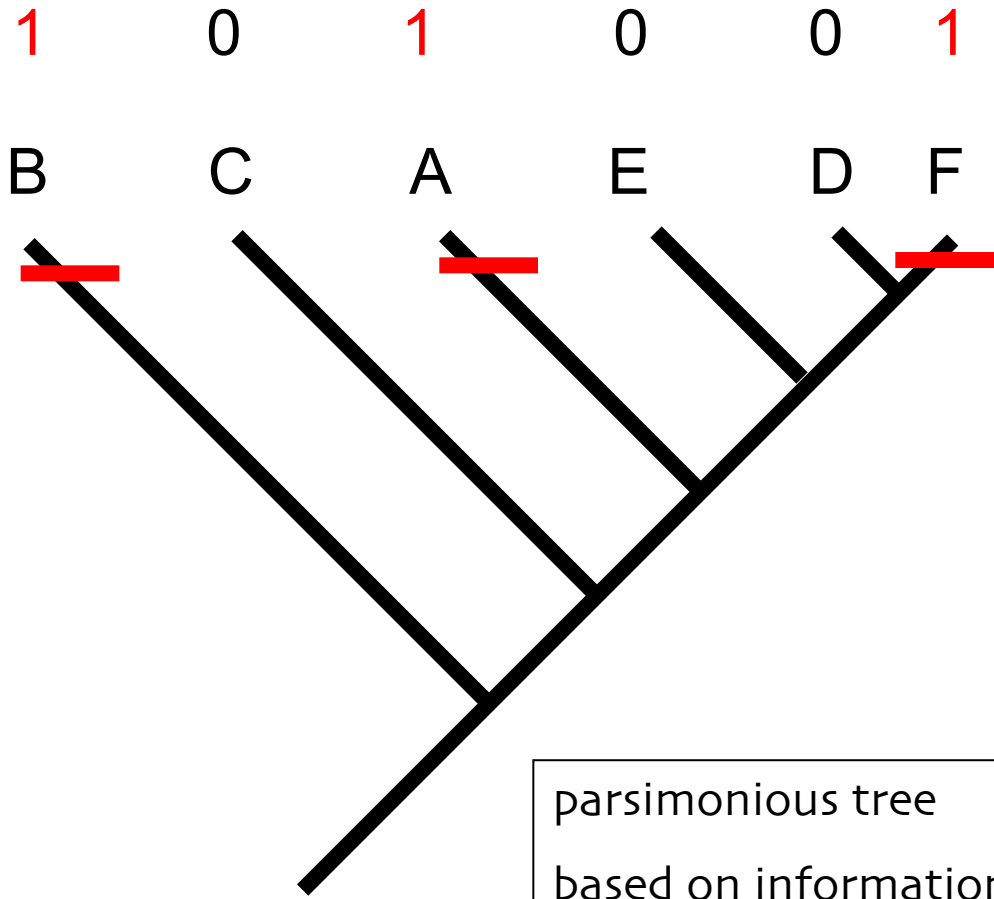
$\sum (e)^2$  minimized



$\sum (e)^2$  minimized

number  
of lvs





characters

trees	0	0	0	0	0	0	0	0	0	0	1	1	1	$\Sigma$
	1	2	3	4	5	6	7	8	9	0	1	2		
1	1	1	2	2	2	2	2	1	2	2	1	2	20	
2	1	1	2	2	2	2	2	1	2	2	2	2	21	
3	1	1	2	2	2	2	2	1	1	2	2	2	20	
4	1	2	2	1	2	2	2	1	2	2	2	2	21	
5	1	2	2	1	2	2	2	1	2	2	2	1	20	
6	1	2	2	1	2	2	2	1	2	2	2	2	21	
7	1	2	2	2	2	2	2	1	2	1	1	2	20	
8	1	2	1	2	2	1	2	1	2	1	2	2	19	
9	1	2	2	2	2	2	2	1	2	1	2	1	20	
10	1	2	2	2	2	2	2	1	1	2	2	1	20	
11	1	2	2	2	2	2	2	1	2	2	1	2	21	
12	1	2	1	2	2	1	2	1	2	2	2	2	20	
13	1	2	2	2	1	2	1	1	1	2	2	2	19	
14	1	2	2	2	1	2	1	1	2	2	2	2	20	
15	1	2	1	2	1	1	1	1	2	2	2	2	18	

COMPARATIVE  
study of structures

analogy

homology

only these should be  
compared with each  
other

phylogenetic analysis

homoplasy

homology

preliminary assumption  
about homology failed

might still define SMALLER  
clades > still valuable

# 1.xi.

1. characters on trees

2. parsimony

**3. apo- & plesiomorphy**

4. outgroup

5. homology & homoplasy

6. character congruence

7. summary

8. home exercise

# Emil Hans WILLI HENNIG

\*20.4.1913 †5.11.1976

Hennig's (1965) 3 primary questions:

1. What is phylogeny?
2. How is it established?
3. How to describe it explicitly?

and his 3 precise answers:

1. Phylogeny is **GENEALOGICAL** relationship where two taxa are more closely related to each other than they are to a third one
2. Relationships are established by **SYNAPOMORPHIES**
3. Relationships can be presented using branching diagrams (=cladograms)

# APOMORPHY

= evolutionary novelty

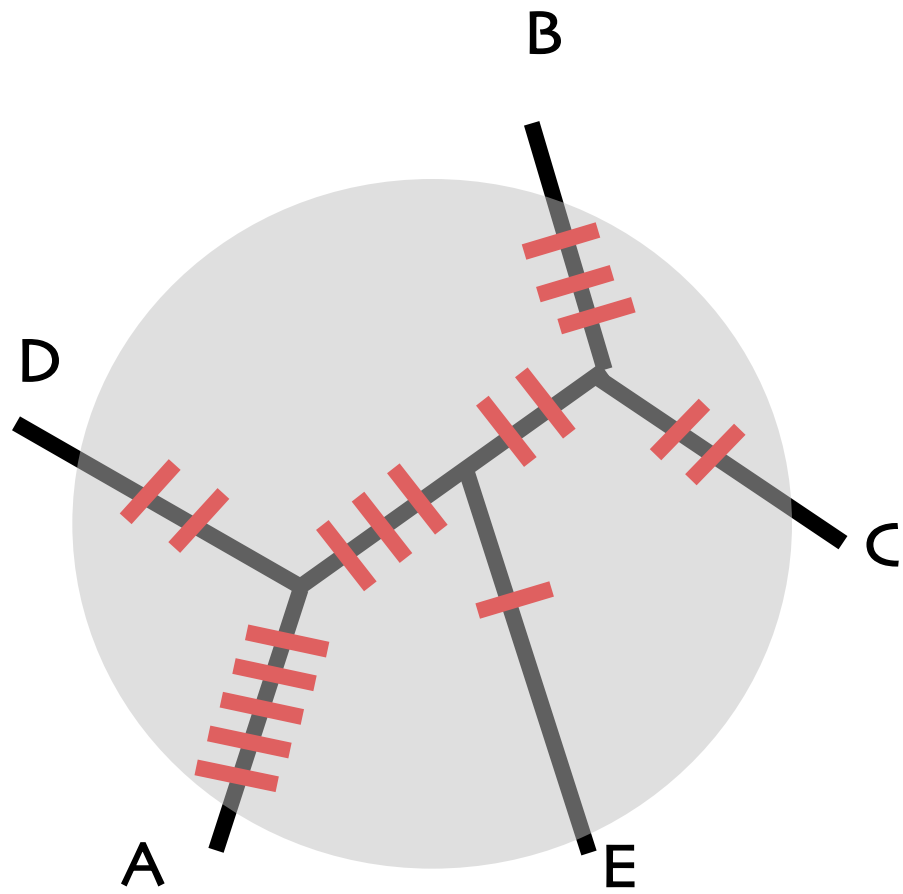
# PLESIOMORPHY

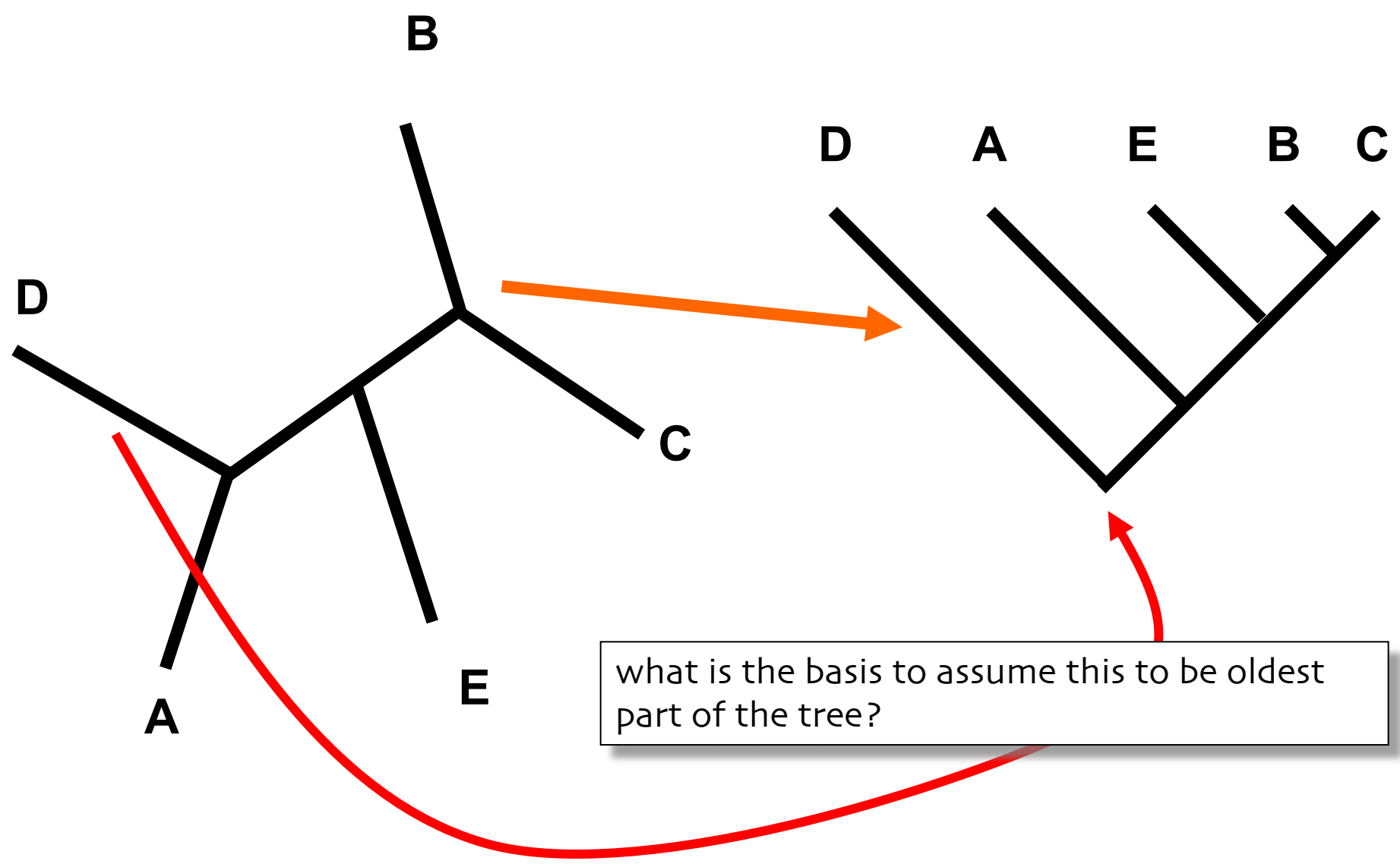
= OLD, inherited character

how to distinguish apo- & plesiomorphies?

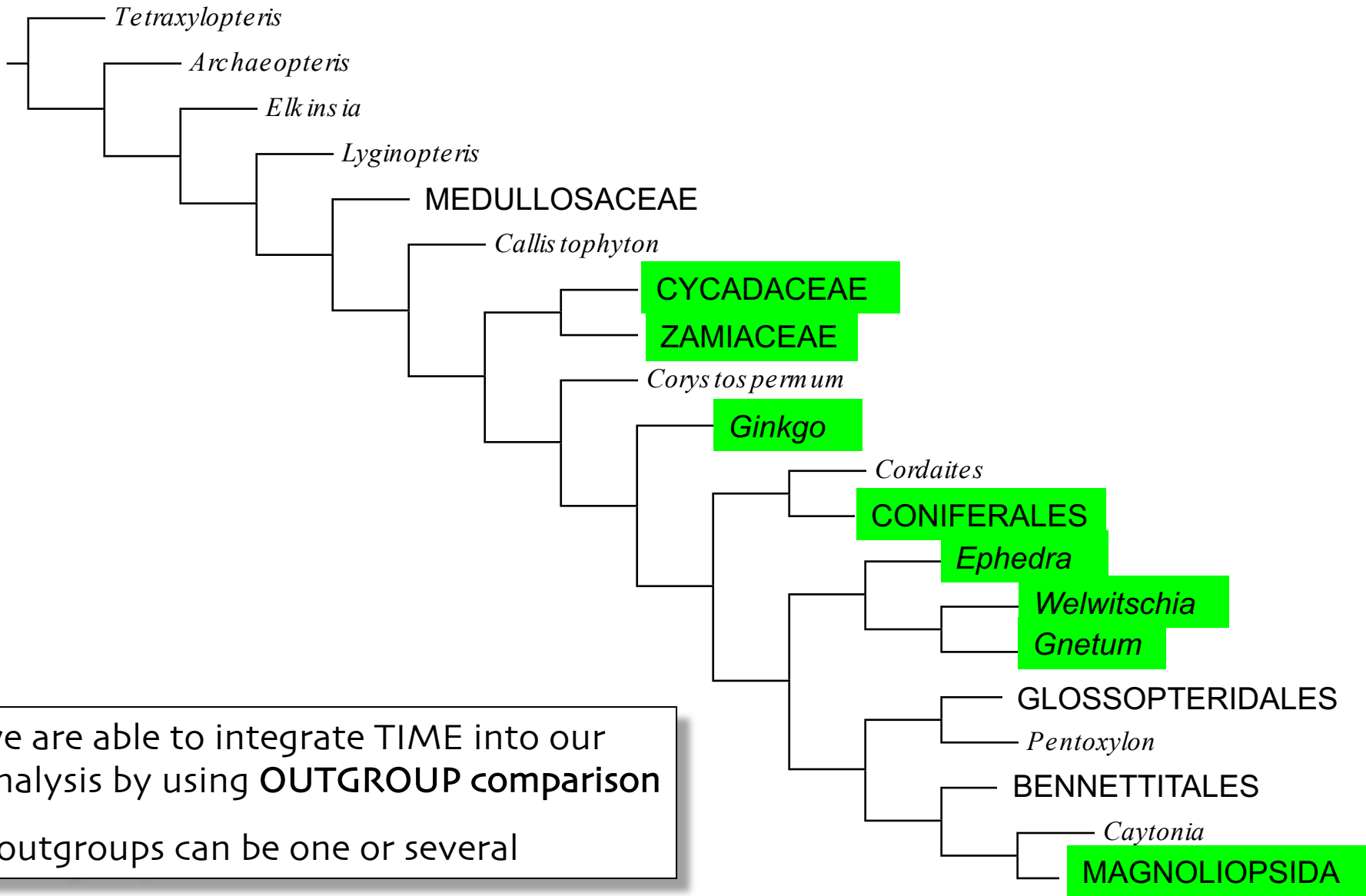
how to introduce time in our hypothesis?



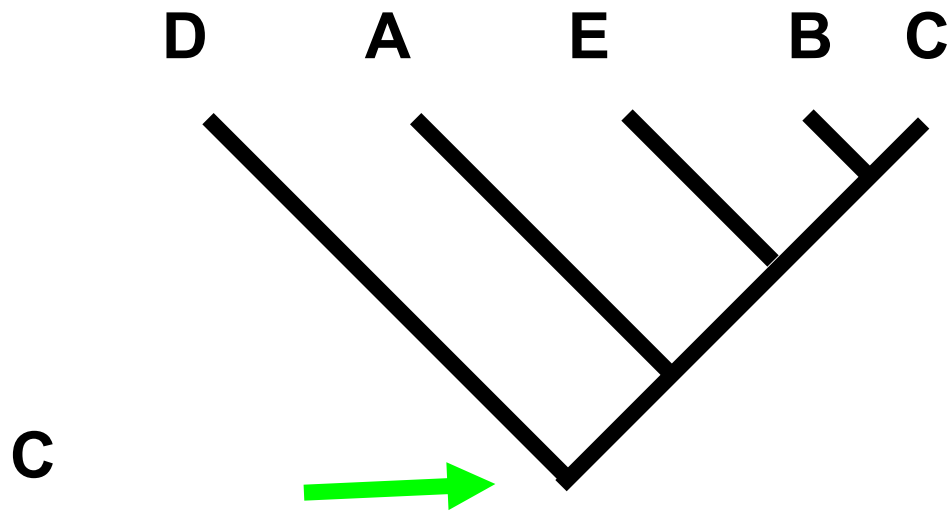
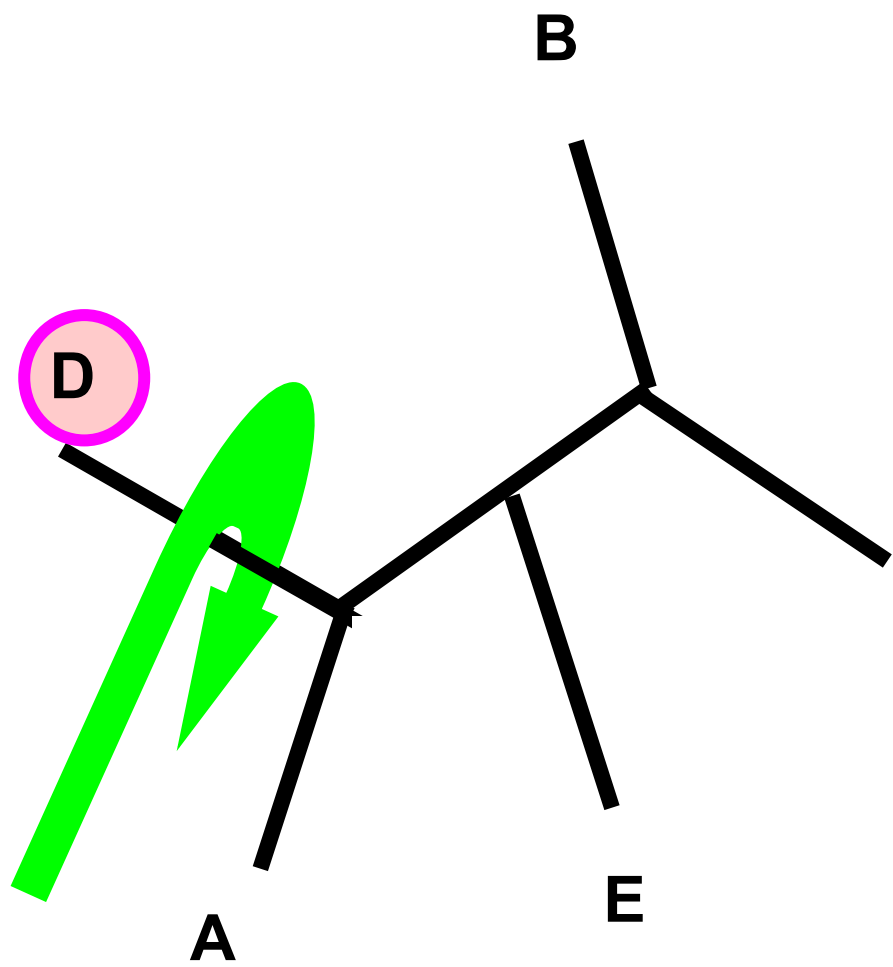




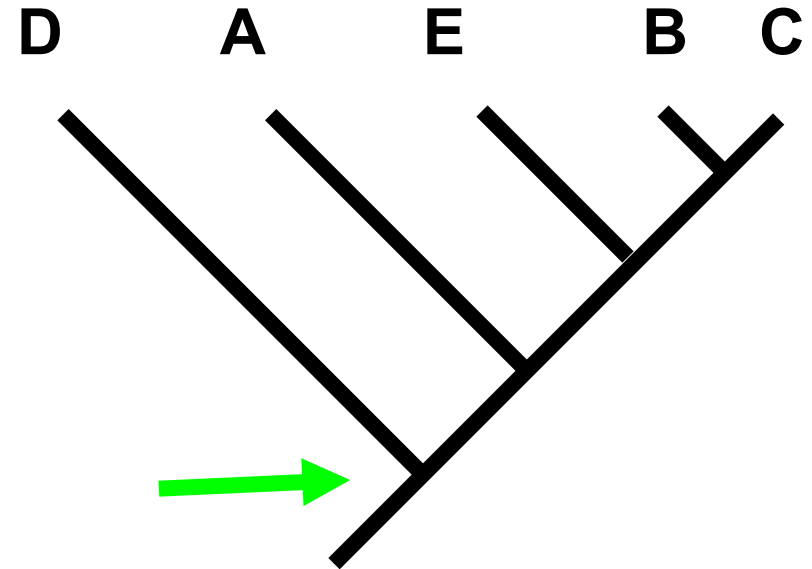
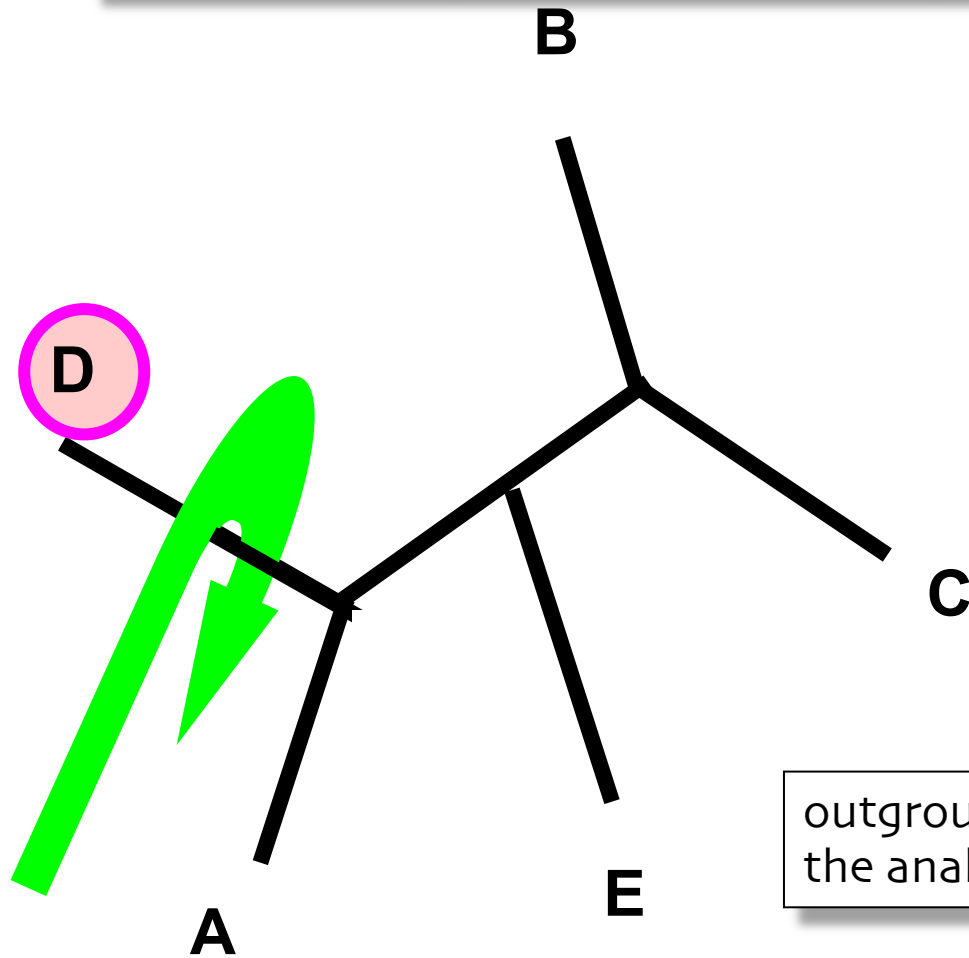
what is the basis to assume this to be oldest part of the tree?



we are able to integrate TIME into our analysis by using **OUTGROUP** comparison  
 - outgroups can be one or several

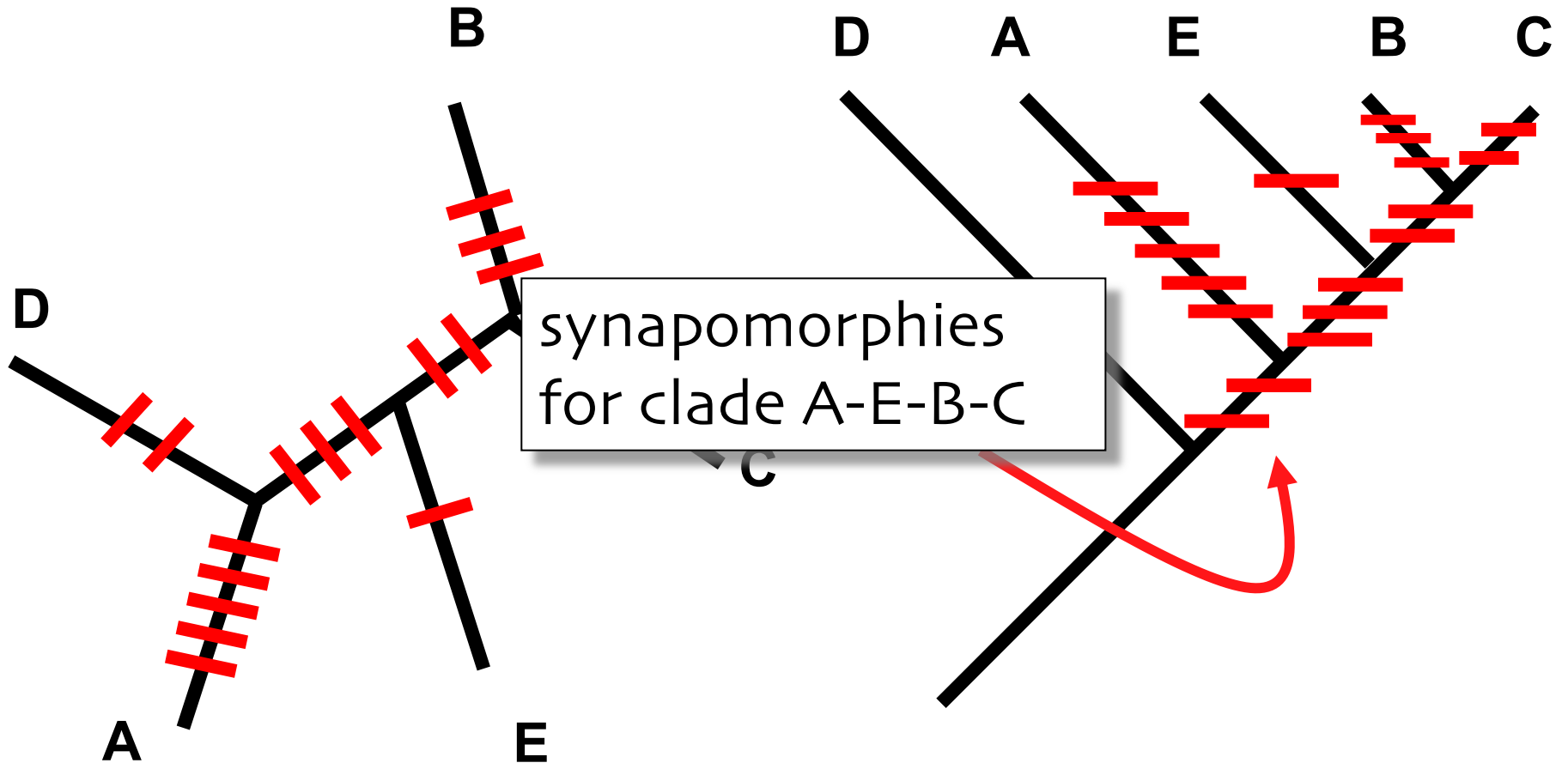


Nixon, K. & Carpenter, J. 1993. On outgroups. Cladistics 9: 413-426.

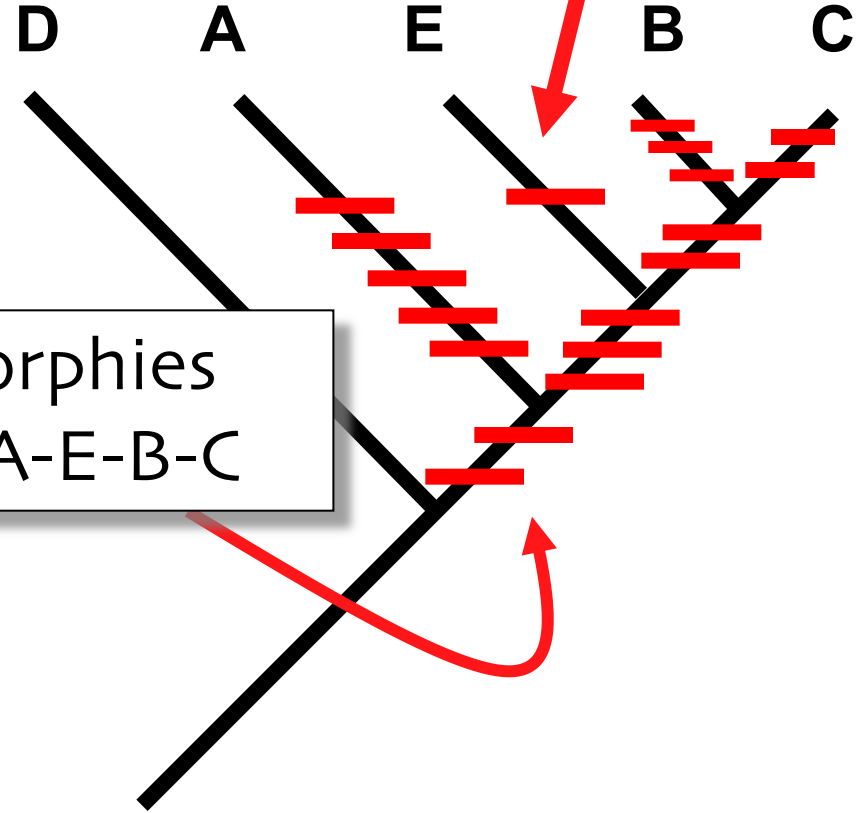


outgroup terminals should be treated in the analyses AS ANY OTHER TERMINAL

they should be included in the analyses FROM THE VERY START



autapomorphy  
for terminal E



synapomorphies  
for clade A-E-B-C

# 1.xi.

1. characters on trees

2. parsimony

3. apo- & plesiomorphy

4. outgroup

**5. homology & homoplasy**

6. character congruence

7. summary

8. home exercise



# HOMOPLASY

earlier assumptions about homology have to be denied (at least at this level)

**HOMOLOGY** = shared feature inherited from common ancestor

for example wings of birds & bats  
flippers of penguins & seals

homologies as limbs **NOT** as wings/flippers

homoplasy is due to convergence or reversals

e.g. during ontogeny development of structure is not completed

# HOMOLOGY & HOMOPLASY



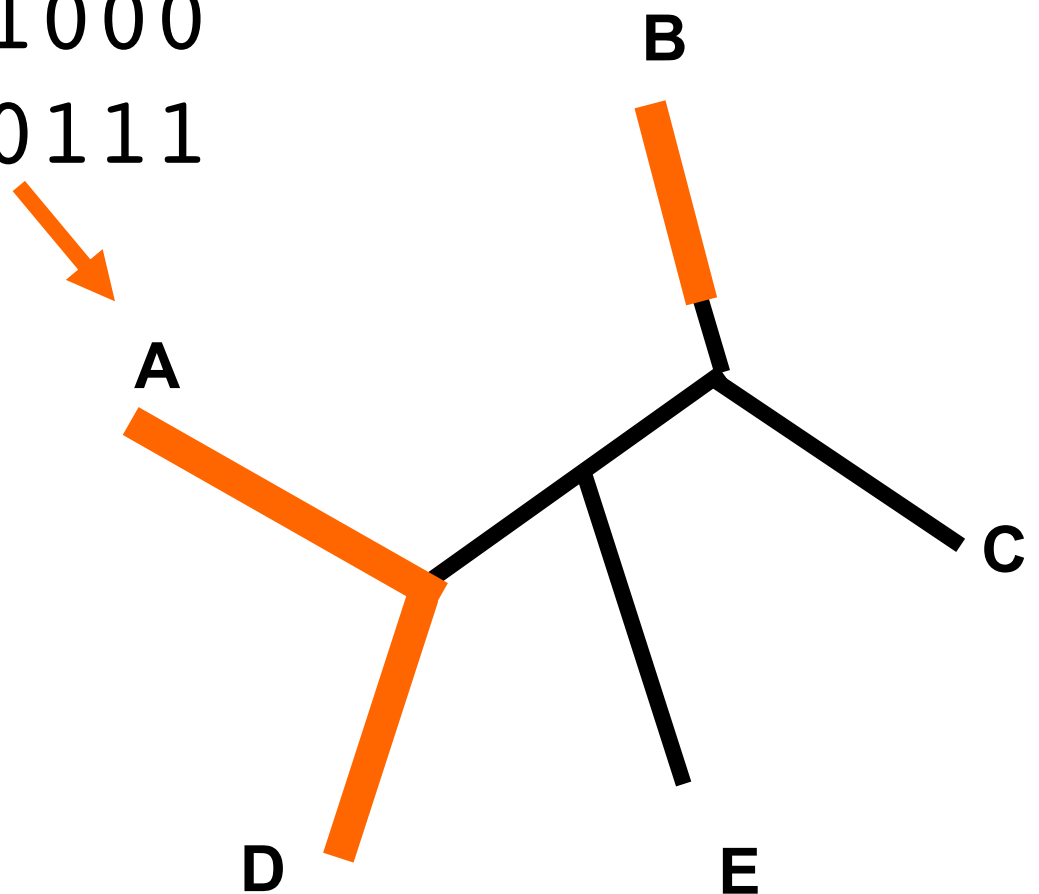
characters are ALWAYS in conflict to some extent →  
**HOMOPLASY** = EXTRA STEPS, i.e. more ch. state changes than the minimum, AFTER phylogenetic analysis has been performed

A	010111011111
B	111000101010
C	101100100001
D	100011001000
E	100001100111

**HOMOPLASY**

(most) parsimonious, i.e.  
shortest tree 18 (16+2)  
evolutionary changes

e.g. character 9  
2 evolutionary changes



# HOMOLOGY & HOMOPLASY



characters are ALWAYS in conflict to some extent →  
**HOMOPLASY** = EXTRA STEPS, i.e. more ch. state changes than the minimum, AFTER phylogenetic analysis has been performed

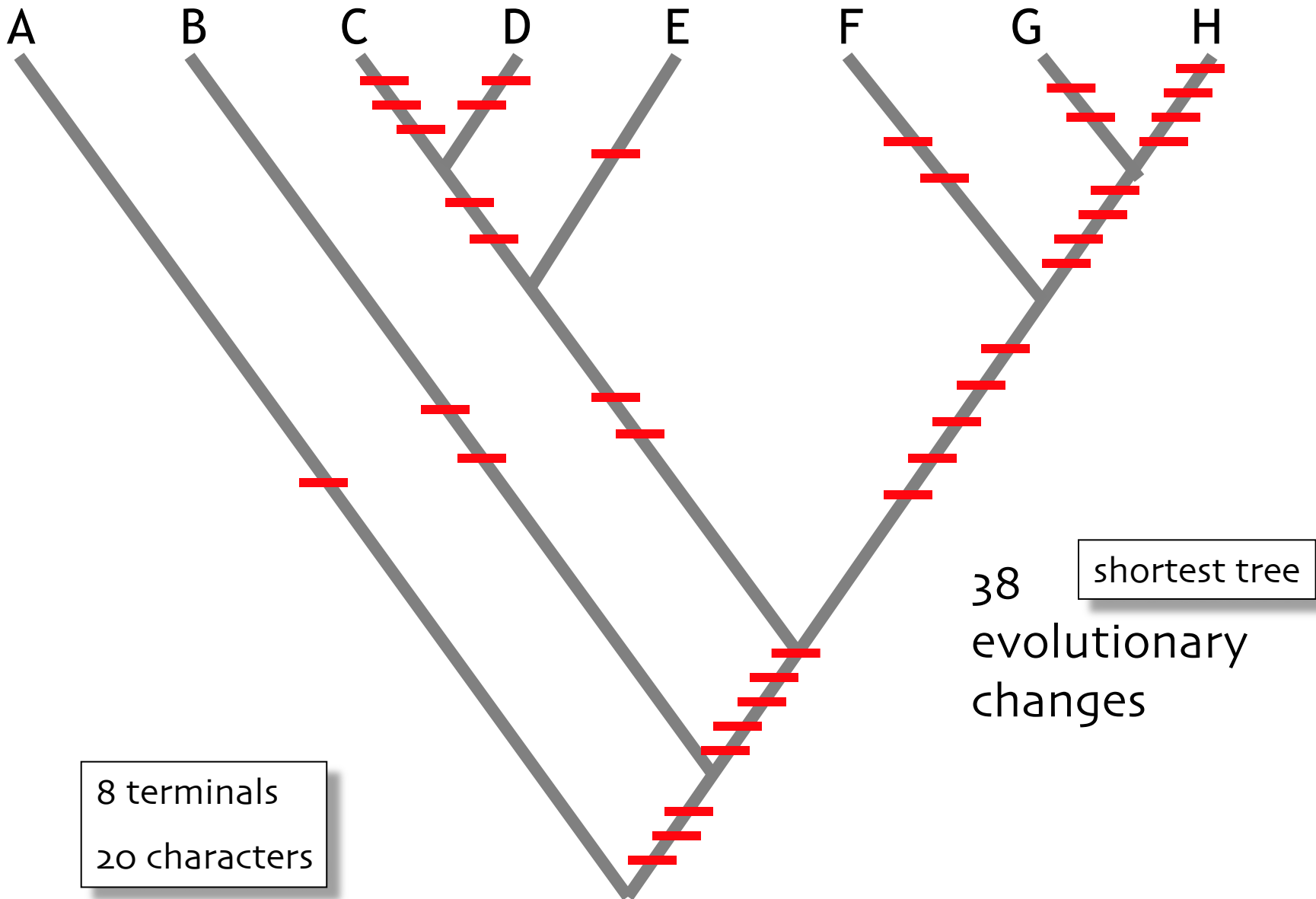
homology informs about history, part of historical signal,  
homoplasy ONLY PARTIALLY, it is NOISE to some extent & at some level

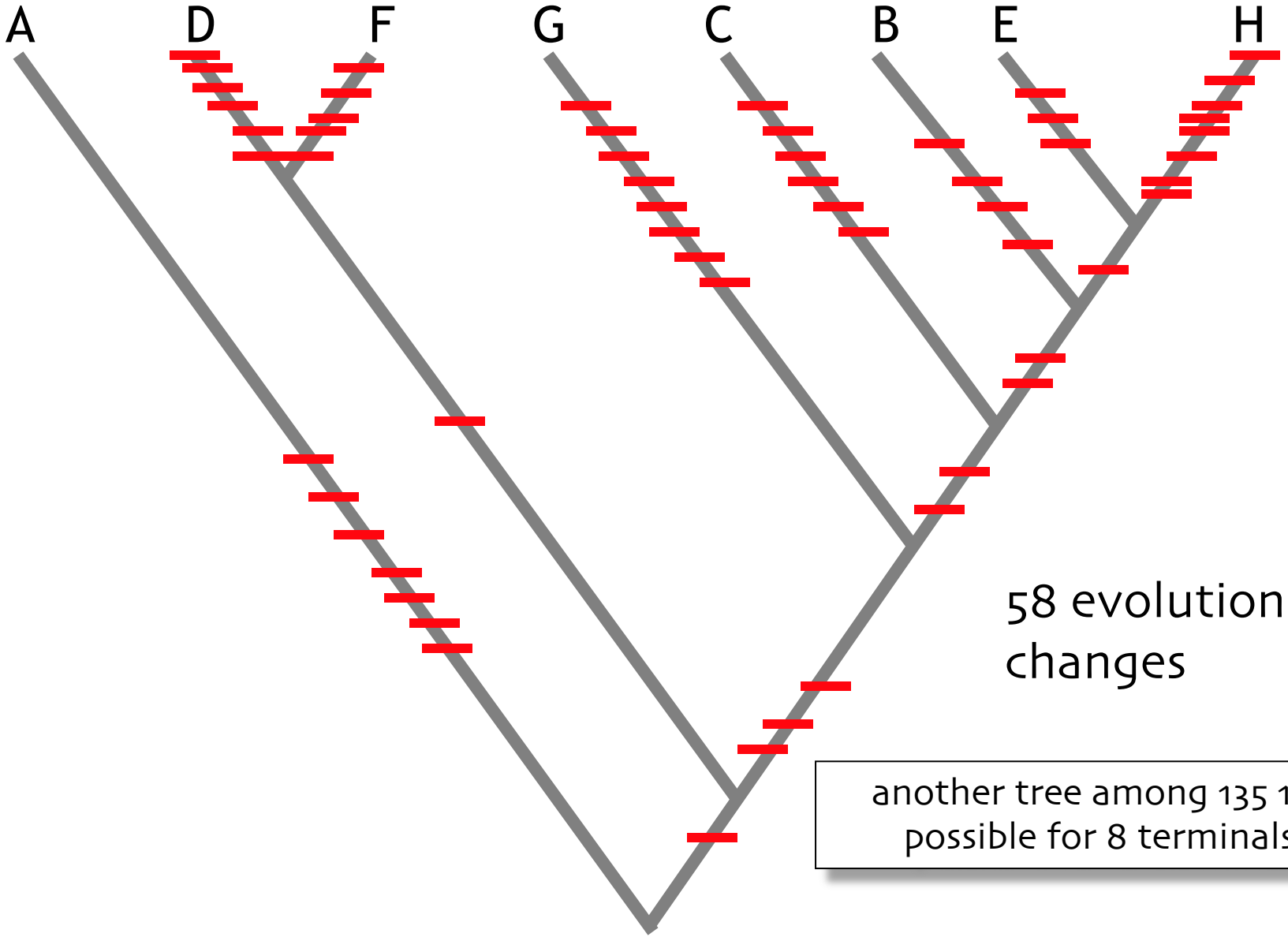
characters

trees	0	0	0	0	0	0	0	0	0	0	1	1	1	$\Sigma$
	1	2	3	4	5	6	7	8	9	0	1	2		
1	1	1	2	2	2	2	2	1	2	2	1	2	20	
2	1	1	2	2	2	2	2	1	2	2	2	2	21	
3	1	1	2	2	2	2	2	1	1	2	2	2	20	
4	1	2	2	1	2	2	2	1	2	2	2	2	21	
5	1	2	2	1	2	2	2	1	2	2	2	1	20	
What is the CONCRETE reason for length differences between trees?														
7	1	2	2	2	2	2	2	1	2	1	1	2	20	
8	1	2	1	2	2	1	2	1	2	1	2	2	19	
9	1	2	2	2	2	2	2	1	2	1	2	1	20	
10	1	2	2	2	2	2	2	1	1	2	2	1	20	
11	1	2	2	2	2	2	2	1	2	2	1	2	21	
12	1	2	1	2	2	1	2	1	2	2	2	2	20	
13	1	2	2	2	1	2	1	1	1	2	2	2	19	
14	1	2	2	2	1	2	1	1	2	2	2	2	20	
15	1	2	1	2	1	1	1	1	2	2	2	2	18	

# 1.xi.

1. characters on trees
2. parsimony
3. apo- & plesiomorphy
4. outgroup
5. homology & homoplasy
- 6. character congruence**
7. summary
8. home exercise

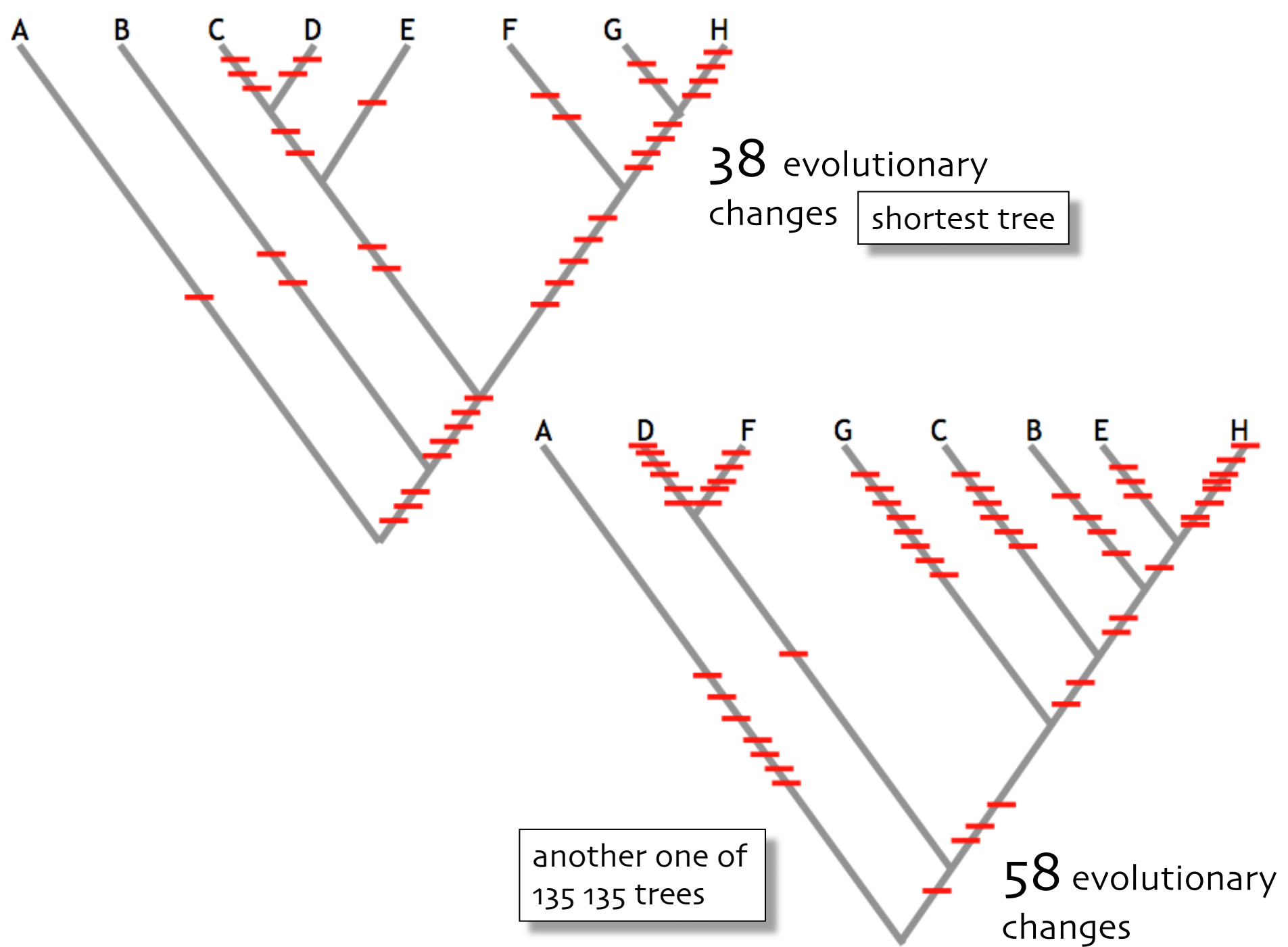




58 evolutionary changes

another tree among 135 135 possible for 8 terminals





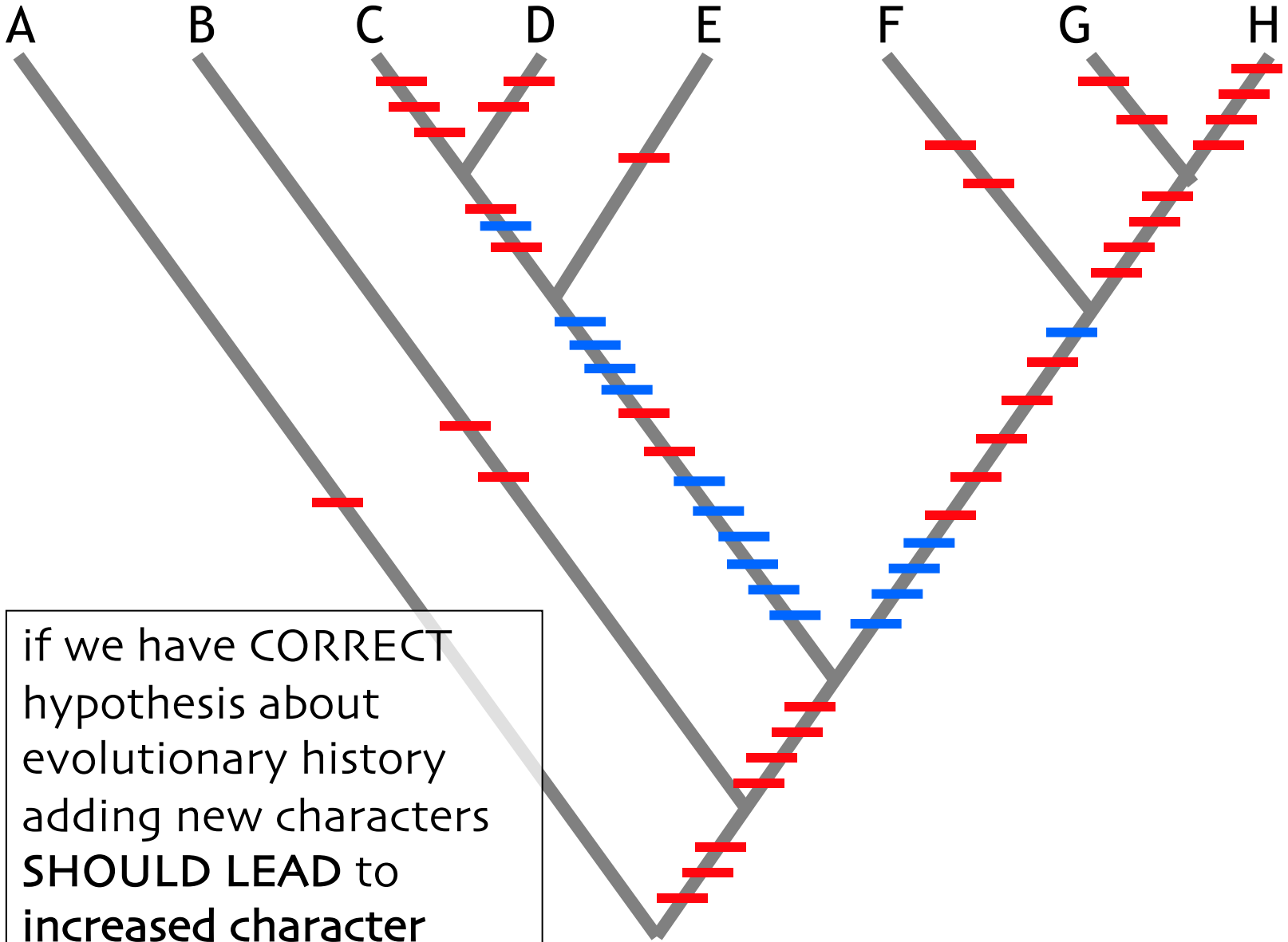
# CHARACTER CONGRUENCE



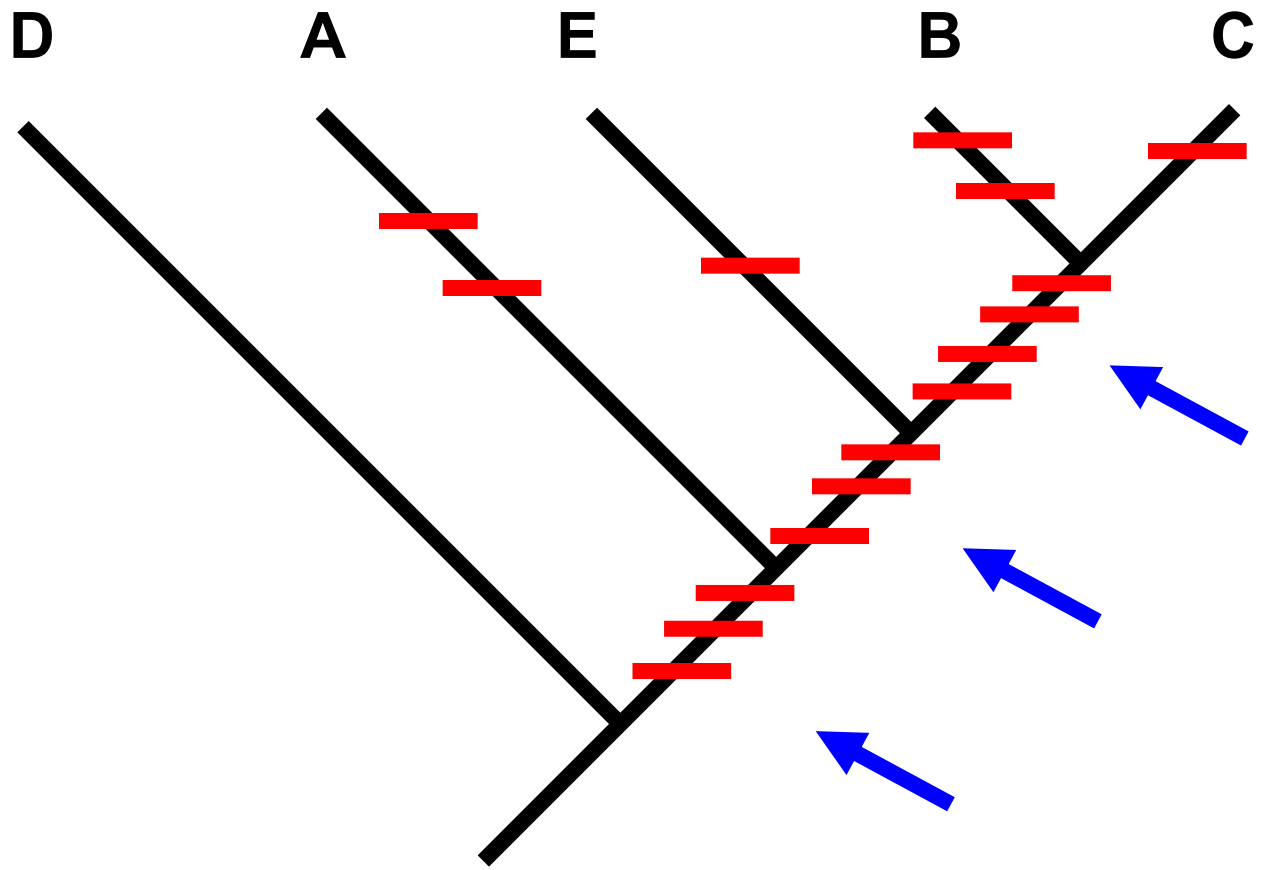
is the concrete reason for the fact that we find short tree(s)

on the shortest tree evolutionary changes are  
CONGRUENT, i.e. CONCENTRATED ON SAME  
BRANCHES of the tree (less homoplasy)

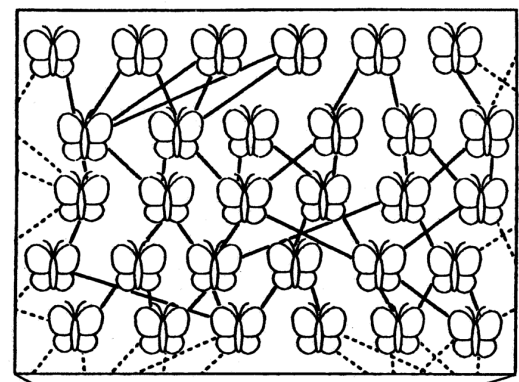
1. Is this accidental?
2. Or have we possibly found a hypothesis that is a good approximation about evolutionary history and this is why so MANY characters are congruent?



if we have CORRECT hypothesis about evolutionary history adding new characters SHOULD LEAD to increased character congruence

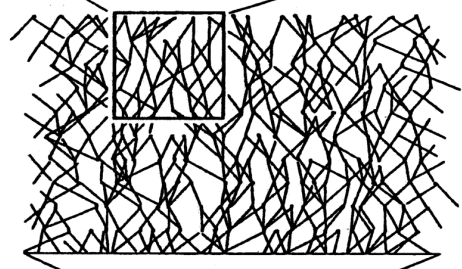


TOKOGENY  
VS.  
PHYLOGENY



1992  
1991  
1990  
1989  
1988

INDIVIDUALS



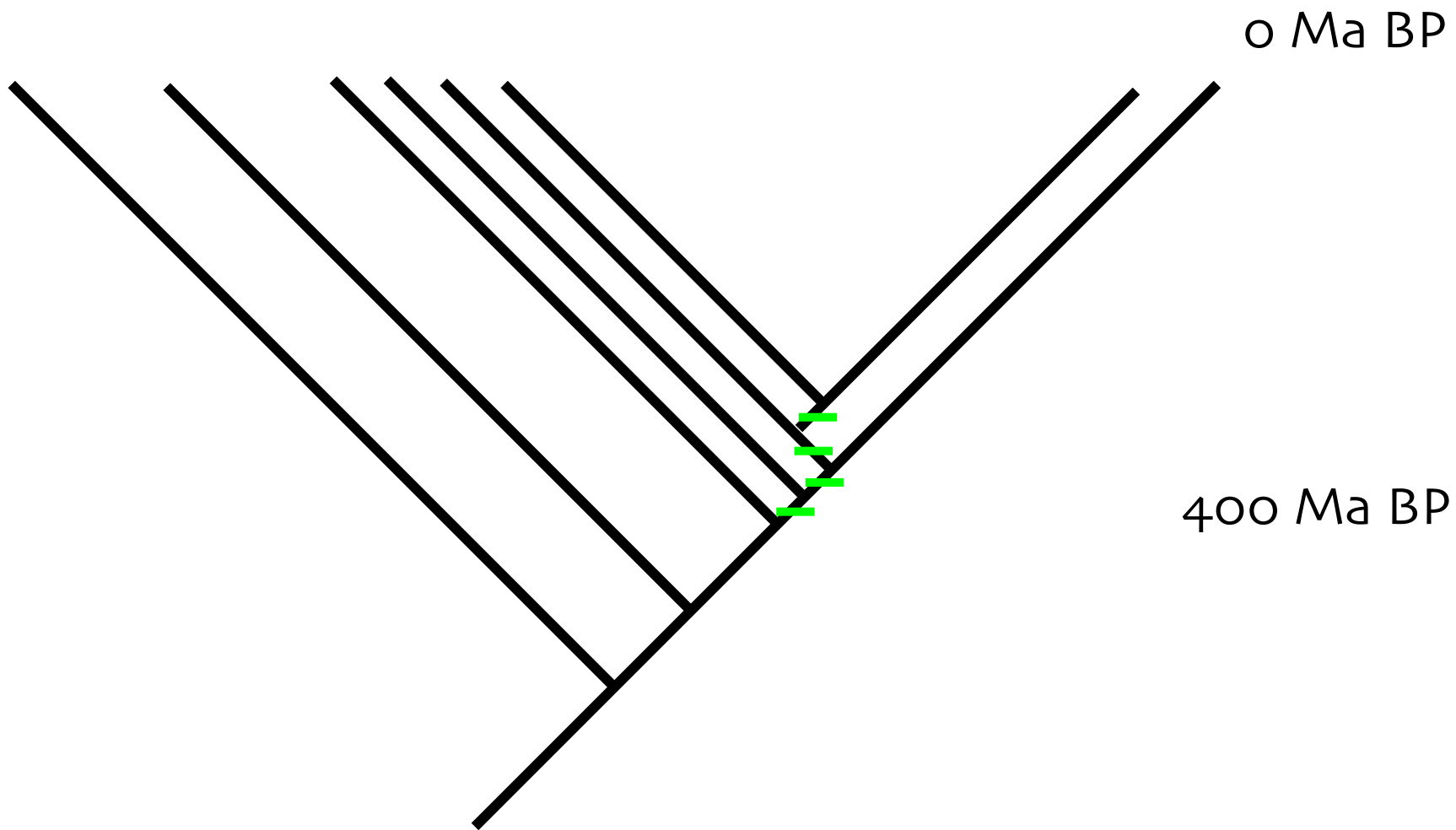
POPULATIONS



SPECIES

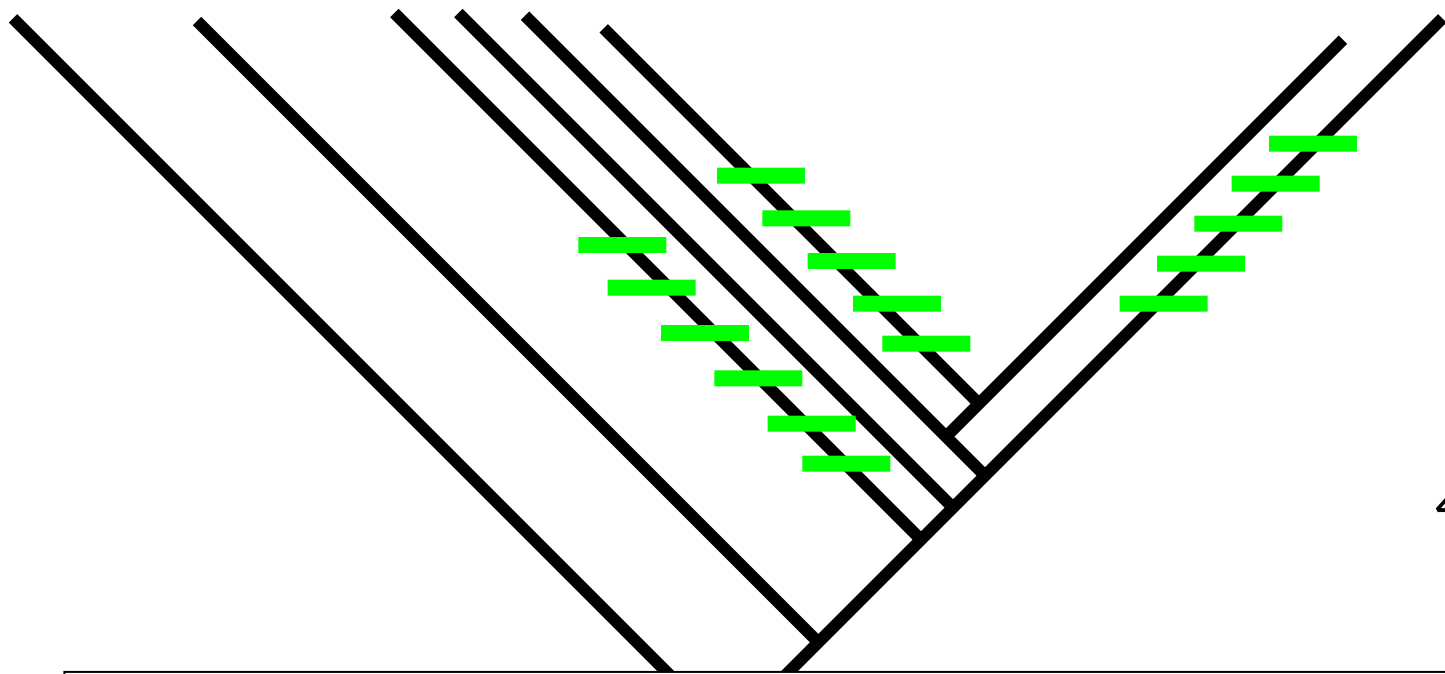
VERY RECENT events of evolutionary history  
NO lineages, NO historical signal

evolutionary process  
evolutionary pattern



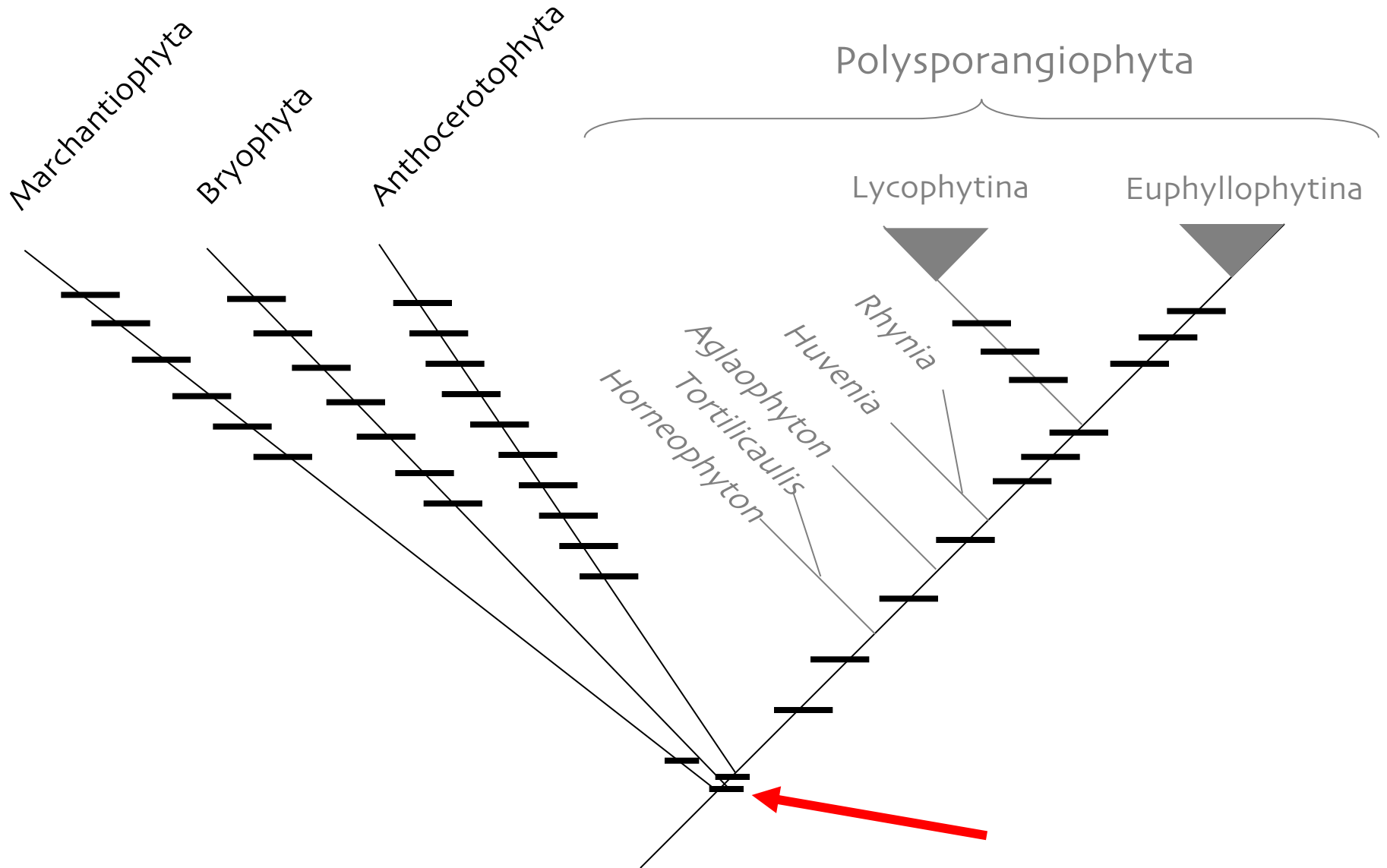
0 Ma BP

400 Ma BP



...or such VERY OLD lineages that have diverged from each other VERY RAPIDLY  
SHARED HISTORY, as compared to the one that followed, is very short

# EMBRYOPHYTES





# SUMMARY

if we use parsimony as our optimality criterion the tree with the smallest number of ch. state changes between the terminals is the best hypothesis of phylogeny

most of the resemblance between terminals is explained by shared evolutionary history

maximizing explanatory power of the hypothesis

apo- & plesiomorphies can be distinguished using outgroup(s)  
concrete reason for length differences between trees is character **CONGRUENCE**

for how many characters does the SAME tree give parsimonious evolutionary explanation

**HOMOLOGY** is the NULL hypothesis for resemblance between terminals

analyses of real matrices will ALWAYS reveal **homoplasy**

# HOME EXERCISE

	characters
	00000000111
	123456789012
terminals	
A	010111011111
B	111000101010
C	101100100001
D	100011001000
E	100001100111