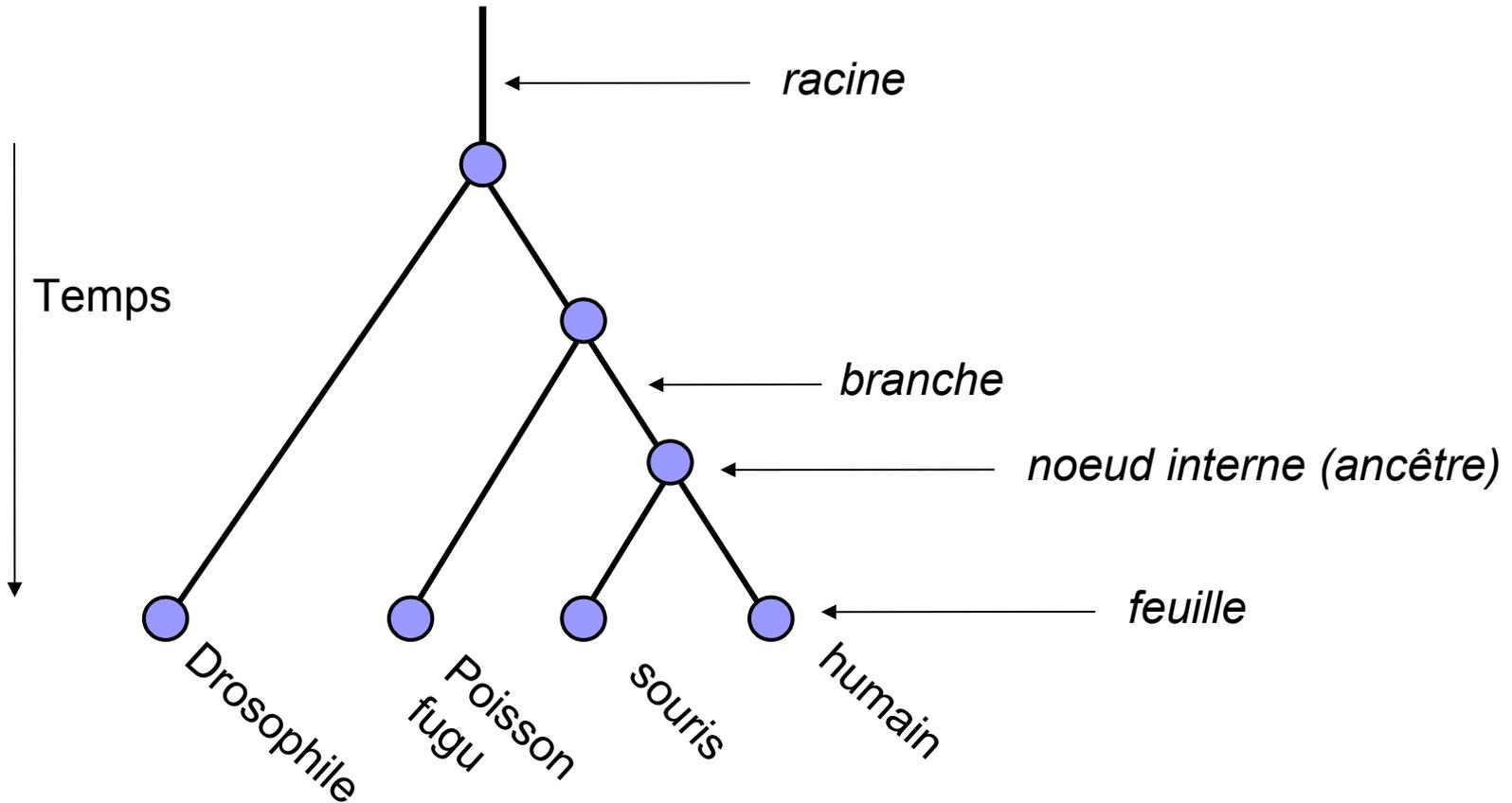
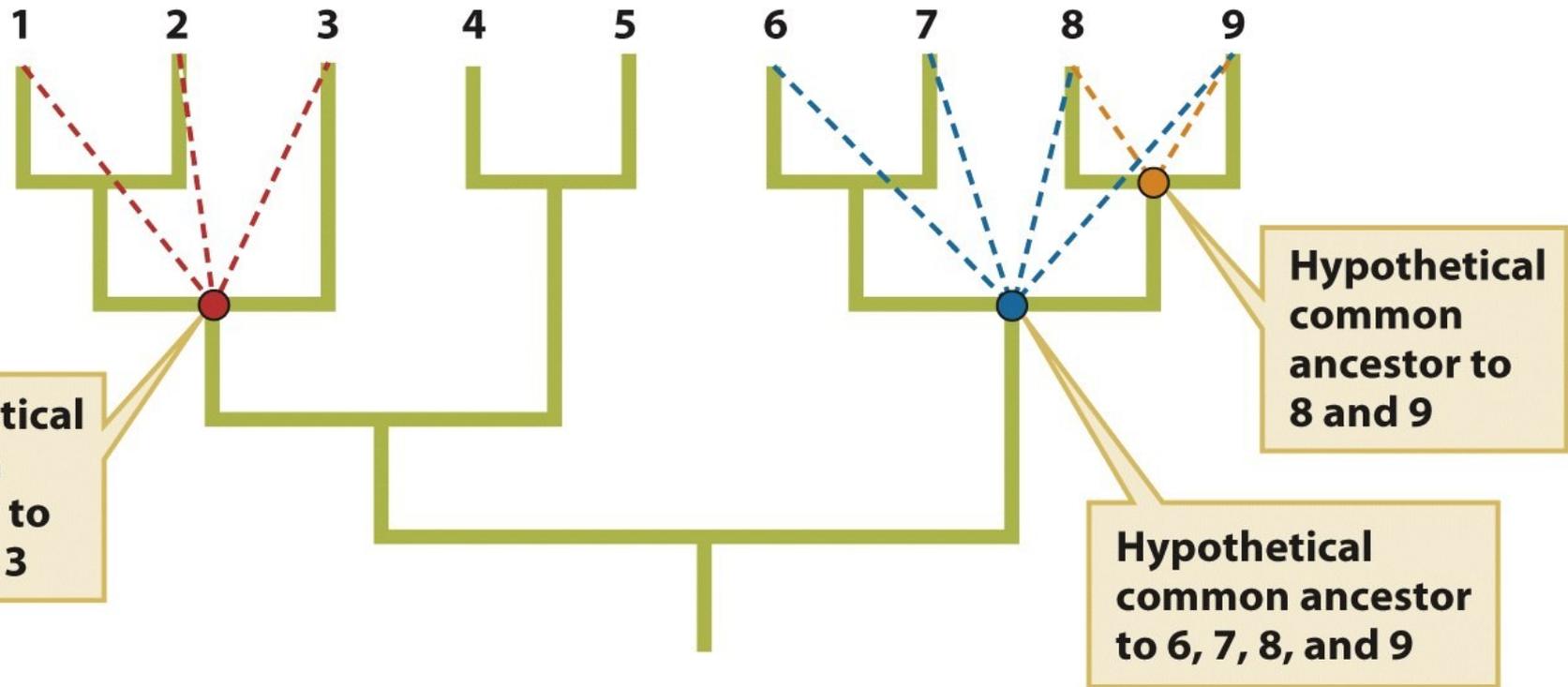
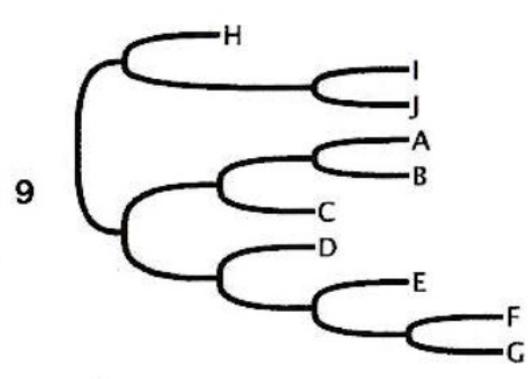
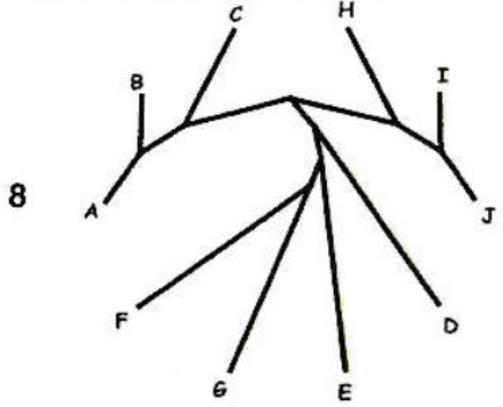
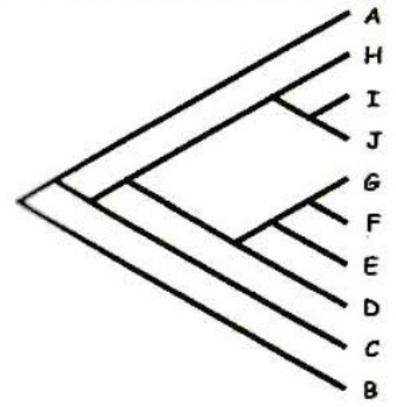
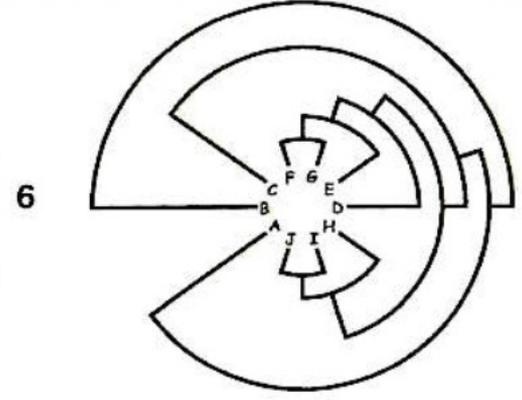
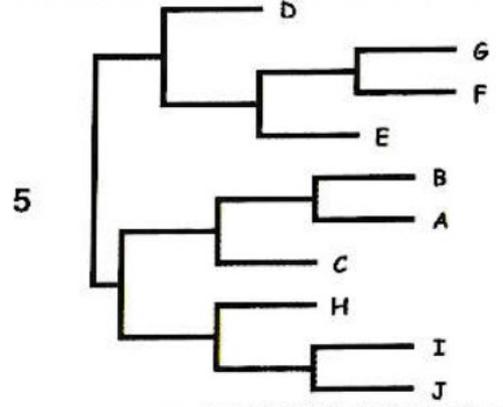
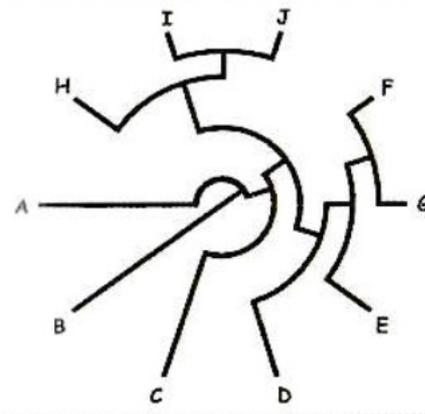
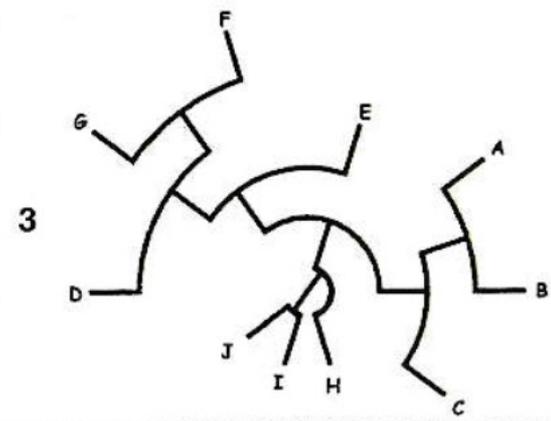
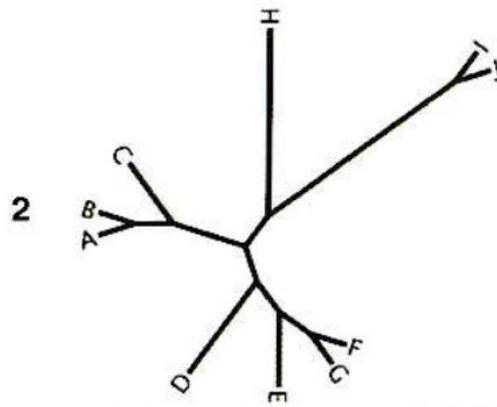
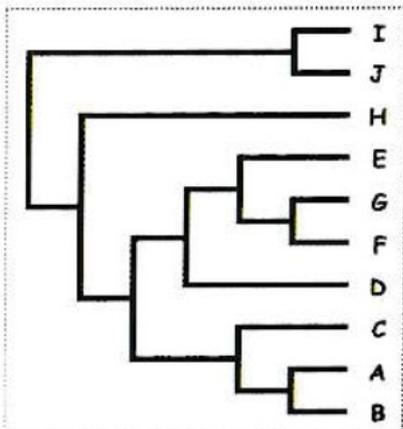


Les arbres phylogénétiques

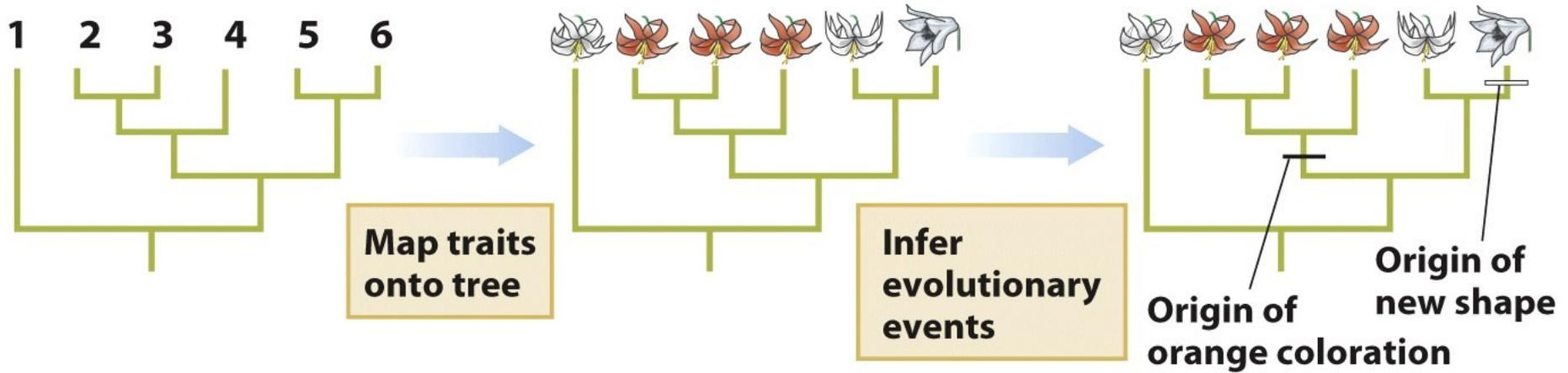


Les noeuds représentent les ancêtres communs

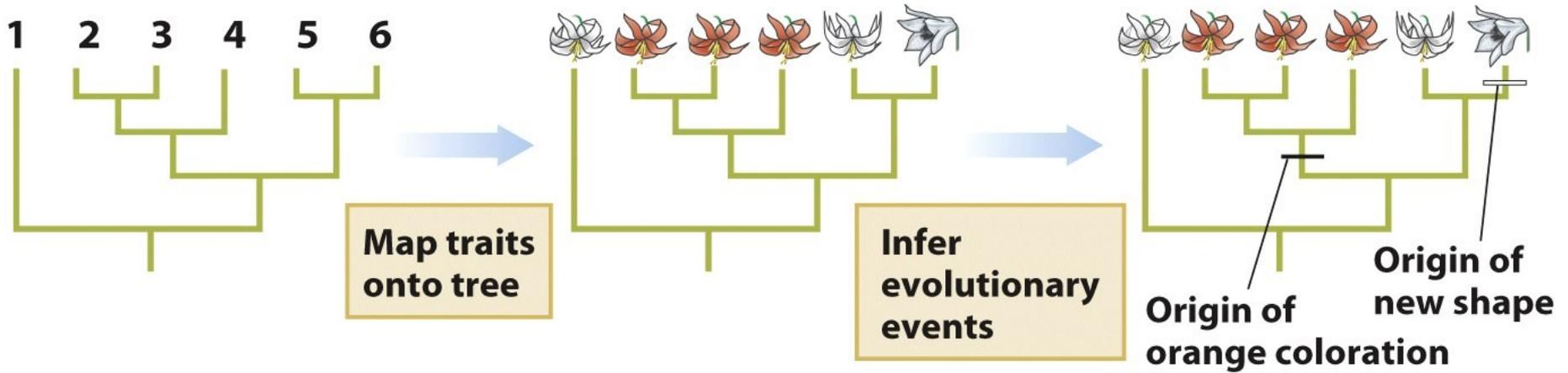




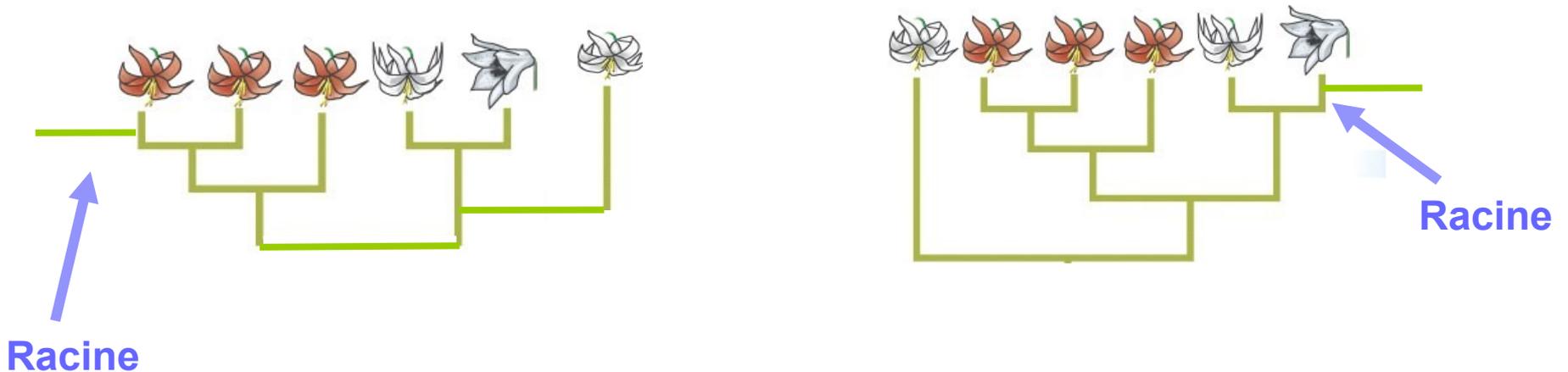
0.1. Parmi ces 9 arbres non enracinés, lequel a une topologie différente des autres ?

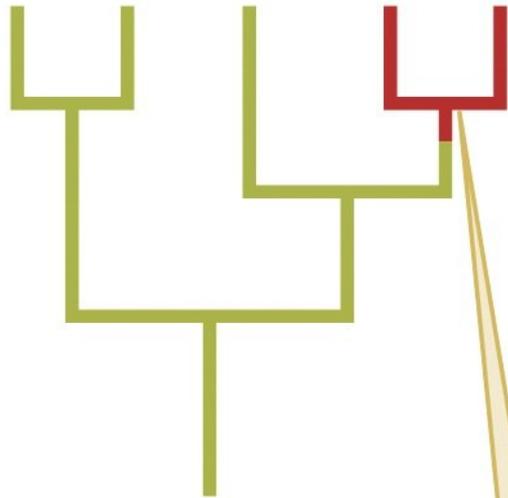


0.2. Si l'arbre n'est pas enraciné, peut-on inférer l'ordre des événements d'apparition de la couleur rouge et de la forme de la fleur ?

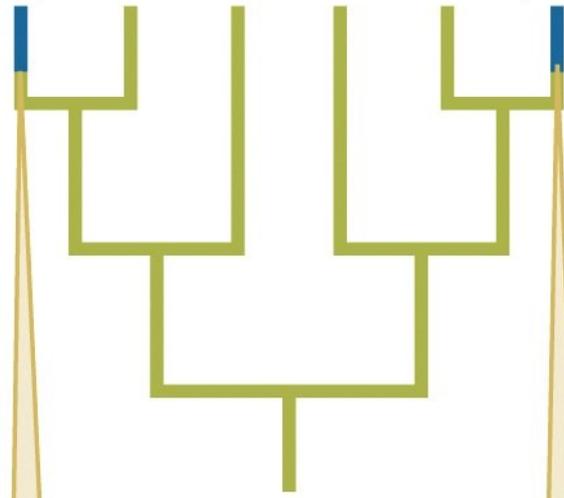
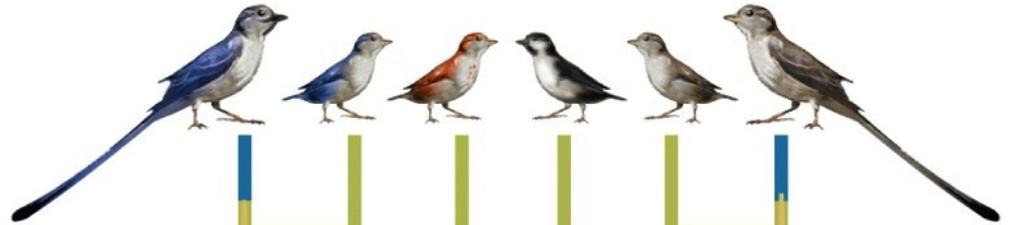


0.2. Si l'arbre n'est pas enraciné, peut-on inférer l'ordre des événements d'apparition de la couleur rouge et de la forme de la fleur ?



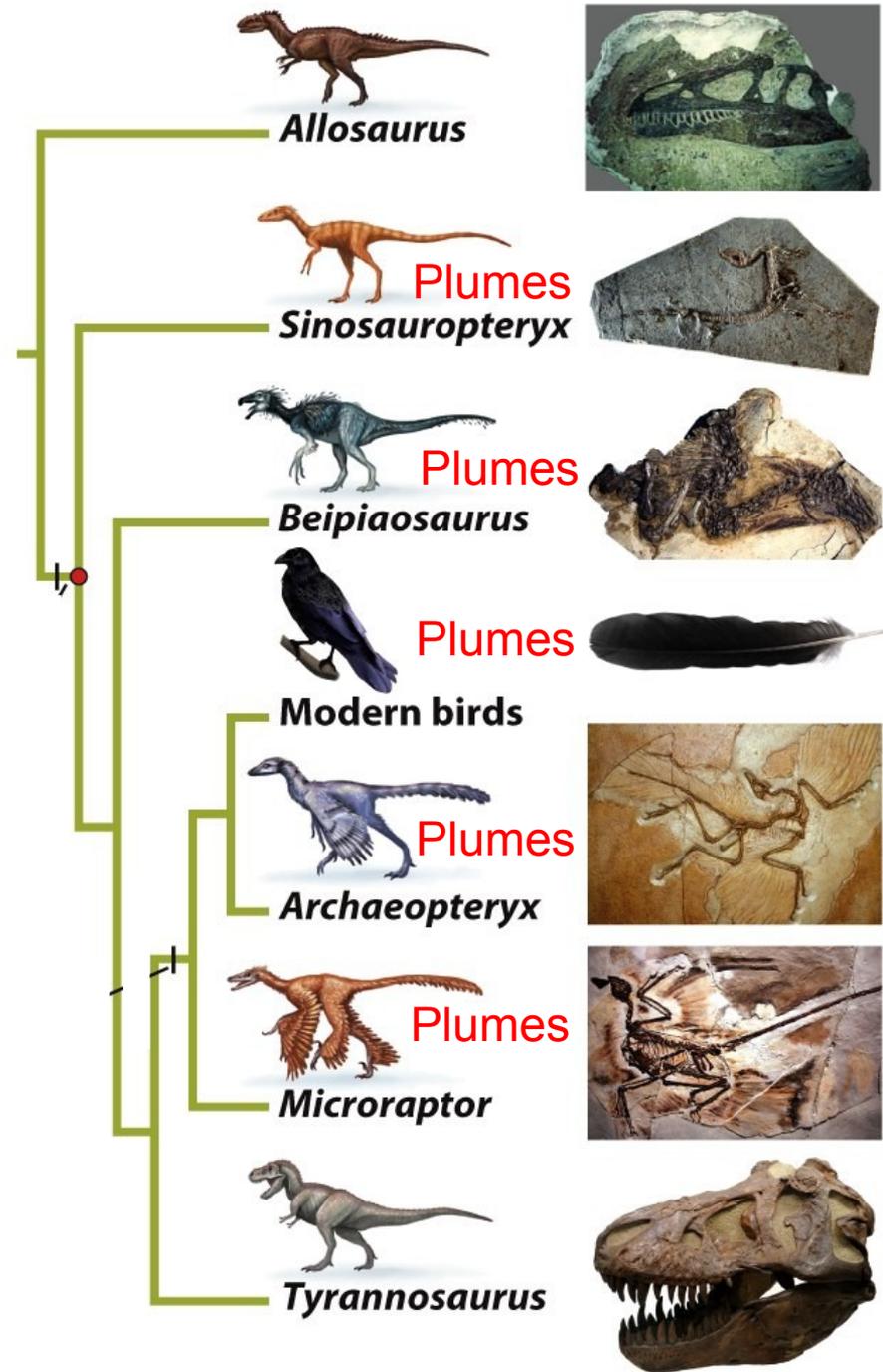


Long legs are a homologous trait

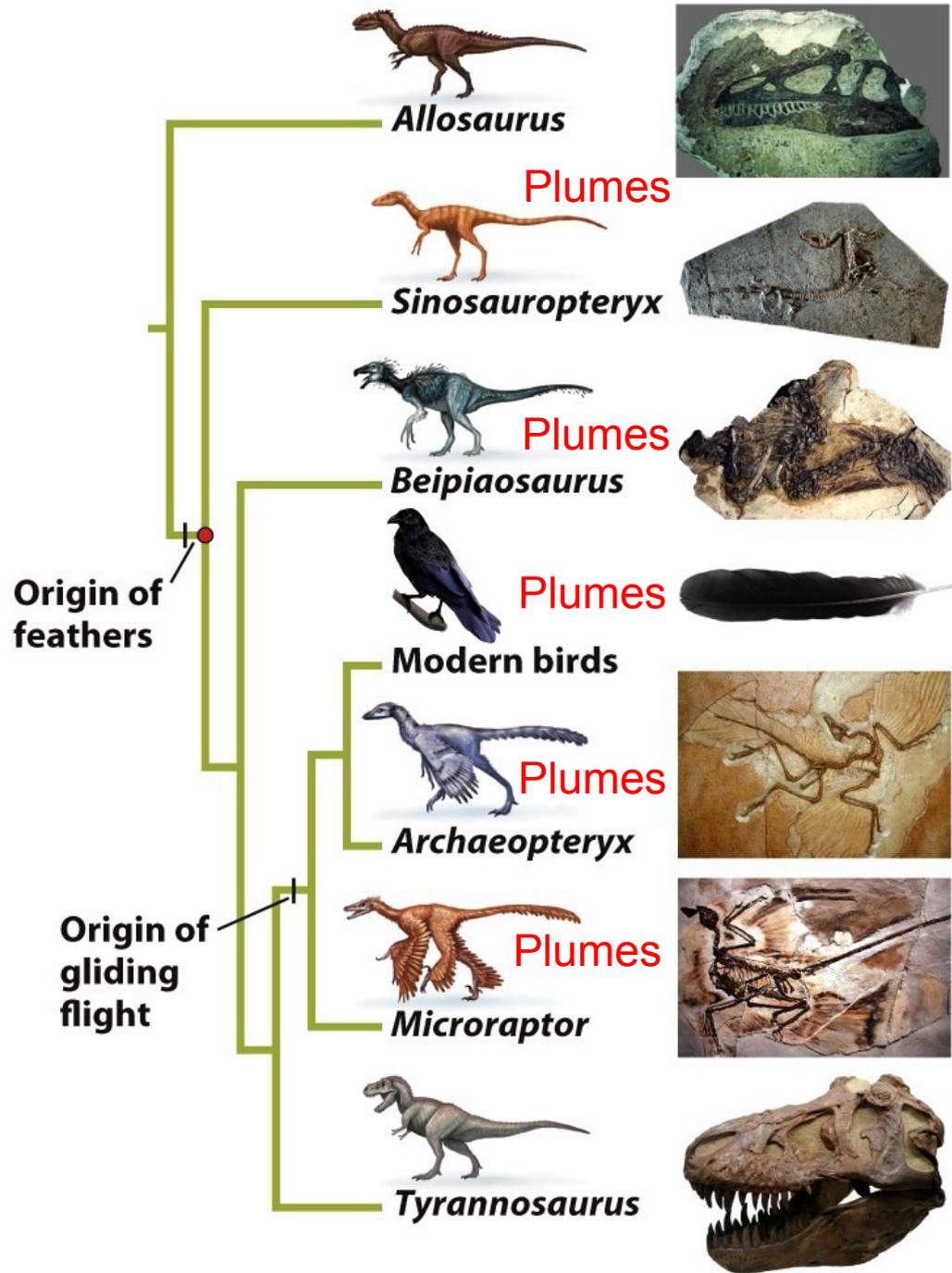


Long tails are an analogous trait

Les plumes sont-elles apparues en même temps que le vol ?



Les plumes sont-elles apparues en même temps que le vol ?



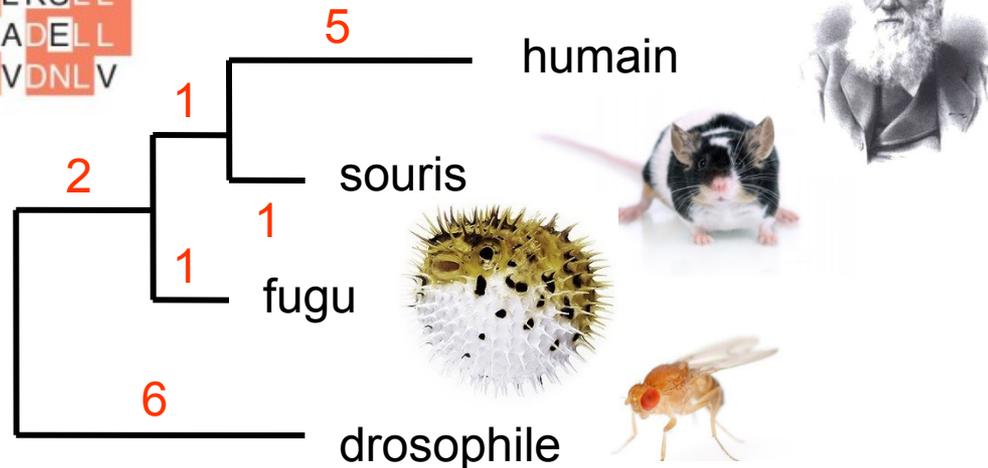
Comment construire
des arbres phylogénétiques?

Méthodes basées sur les distances

Distance évolutive (entre séquences) = différences entre séquences

EGKTCAF I DA	EHALDPI YAR	KLGVD I DNLL
LGGTAAF I DA	EHALDPI YAA	KLGVNVDDL
KGGICAFV DA	EHALDPVYAR	KLGVDLQNL
AGGVAAF I DA	EHALDVSYAR	KLGV RVEELL
NGGVCAF I DA	EHALDVHYAK	RLGVDTQNL
EGGIAAFV DA	E QALDPTYAS	ALGVDIQNLL
AGGQVAFV DA	EHALDPEYAK	KLGVDI DNL I
Q-RTSAF I DA	EHALDPVYAQ	KLGVNIEELL
LGGAAAV I DA	EHALDPVYAK	RLGVNI DDLV
EGGIAAF I DA	EHALDPVYAK	ALGVNVAELW
MGGVAAV I DA	EHALDPNYAA	LI GANI NDLM
AGGIAAF I DA	EHA FDRFYAA	KLGV D VDNLF
AGGLAAI I DA	EHA FDRTYAE	KLGV NVDNLW
MGGVAAF I DA	EHALDPVYAK	NLGV DLKSL
AGGTCAF I DA	EHALDPVYAR	ALGV NADELL
RGGVAVF I DA	EHALDPKYAK	KLGV D VDNLV

humain	x			
souris	6	x		
fugu	7	3	x	
drosophile	14	10	9	x
	humain	souris	fugu	Drosophile



Méthodes basées sur les distances :

UPGMA (Unweight Pair Group Method with Arithmetic mean)

- On identifie d'abord les deux séquences les plus proches
- Ce groupe est ensuite traité comme un tout, on recalcule la matrice de distances (moyenne de la paire aux autres)
- On recherche la séquence la plus proche et ainsi de suite jusqu'à ce qu'il n'y ait plus que deux groupes.

Tracez l'arbre UPGMA des taxa dont la matrice de distances est :

EGKTCAF I DA	EHALDP I YAR	KLGVD I DNLL
LGGTAAF I DA	EHALDP I YAA	KLGVNVDDL
KGG I CAFVDA	EHALDPVYAR	KLGVDLQNL
AGGVAAF I DA	EHALDVSYAR	KLGV RVEELL
NGGVCAF I DA	EHALDVHYAK	RLGVDTQNL
EGG I AAFVDA	E QALDPTYAS	ALGVD IQNL
AGGQVAFVDA	EHALDPEYAK	KLGVD I DNLI
Q-RTSAF I DA	EHALDPVYAQ	KLGVNIEELL
LGGAAAV I DA	EHALDPVYAK	RLGVNIDDLV
EGG I AAF I DA	EHALDPVYAK	ALGVNVAELW
MGGVAAV I DA	EHALDPNYAA	LI GANI NDLM
AGG I AAF I DA	EHA FDRFYAA	KLGVDV DNLF
AGGLAA I I DA	EHA FDRTYAE	KLGVNV DNLW
MGGVAAF I DA	EHALDPVYAK	NLGVDLKSL
AGGTCAF I DA	EHALDPVYAR	ALGVNADELL
RGGVAVF I DA	EHALDPKYAK	KLGVDV DNLV

	A	B	C	D	E
B	2				
C	4	4			
D	6	6	6		
E	6	6	6	4	
F	8	8	8	8	8

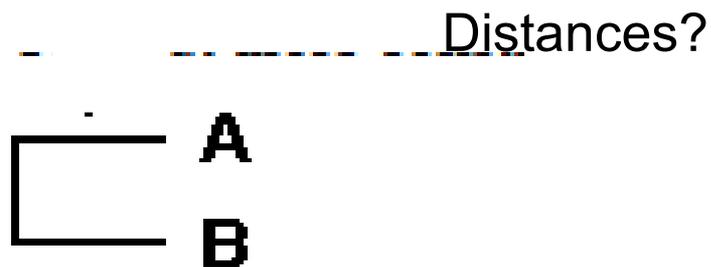
	A	B	C	D	E
B	2				
C	4	4			
D	6	6	6		
E	6	6	6	4	
F	8	8	8	8	8

On groupe tout d'abord les deux taxa avec la distance la plus faible.
Le point de branchement est positionné à la moitié des distances

Premiers à brancher?

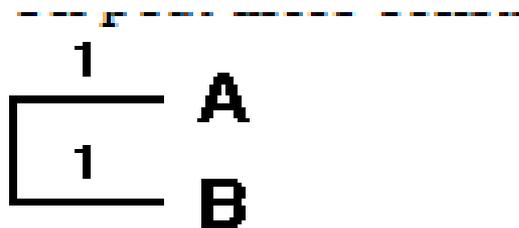
	A	B	C	D	E
B	2				
C	4	4			
D	6	6	6		
E	6	6	6	4	
F	8	8	8	8	8

On groupe tout d'abord les deux taxa avec la distance la plus faible (A et B).
Le point de branchement est positionné à la moitié des distances



	A	B	C	D	E
B	2				
C	4	4			
D	6	6	6		
E	6	6	6	4	
F	8	8	8	8	8

On groupe tout d'abord les deux taxa avec la distance la plus faible (A et B).
Le point de branchement est positionné à la distance $2/2=1$.



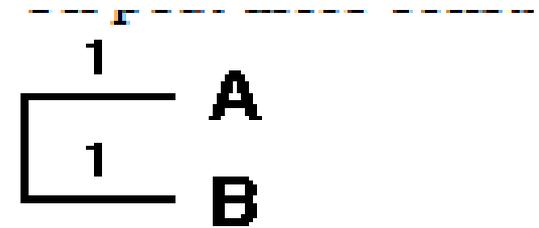
	A	B	C	D	E
B	2				
C	4	4			
D	6	6	6		
E	6	6	6	4	
F	8	8	8	8	8



Dans la suite, le clade (A,B) est considéré comme un tout et on peut calculer une nouvelle matrice de distances comme moyennes des distances à A et B:

	A,B	C	D	E
C				
D		6		
E		6	4	
F		8	8	8

	A	B	C	D	E
B	2				
C	4	4			
D	6	6	6		
E	6	6	6	4	
F	8	8	8	8	8



Dans la suite, le clade (A,B) est considéré comme un tout et on peut calculer une nouvelle matrice de distances :

$$\text{dist}(A,B),C = (\text{dist}AC + \text{dist}BC) / 2 = 4$$

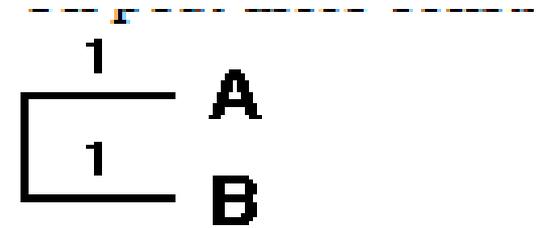
$$\text{dist}(A,B),D = (\text{dist}AD + \text{dist}BD) / 2 = 6$$

$$\text{dist}(A,B),E = (\text{dist}AE + \text{dist}BE) / 2 = 6$$

$$\text{dist}(A,B),F = (\text{dist}AF + \text{dist}BF) / 2 = 8$$

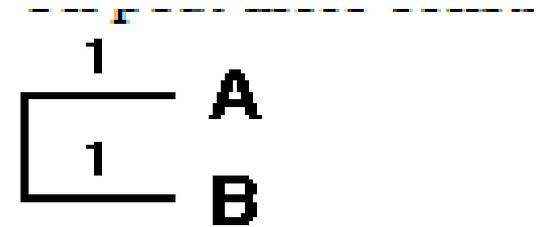
	A,B	C	D	E
C	4			
D	6	6		
E	6	6	4	
F	8	8	8	8

		A,B	C	D	E
Cycle 2	C	4			
	D	6	6		
	E	6	6	4	
	F	8	8	8	8

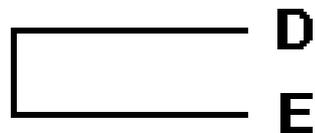


Prochain à brancher?

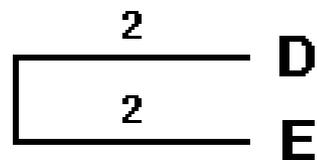
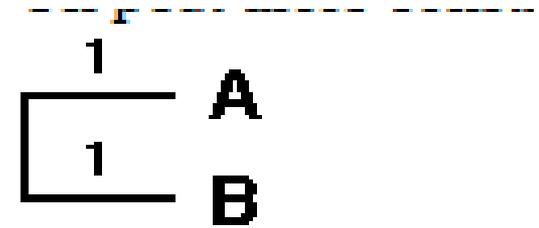
		A,B	C	D	E
Cycle 2	C	4			
	D	6	6		
	E	6	6	4	
	F	8	8	8	8

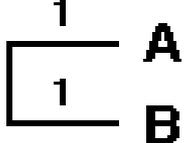
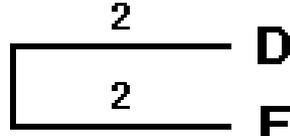


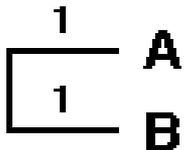
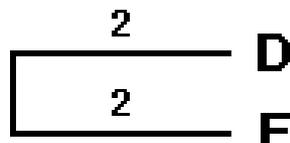
Longueurs des branches?

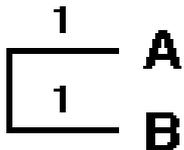
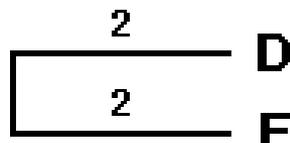
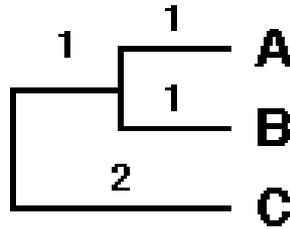


		A,B	C	D	E
Cycle 2	C	4			
	D	6	6		
	E	6	6	4	
	F	8	8	8	8

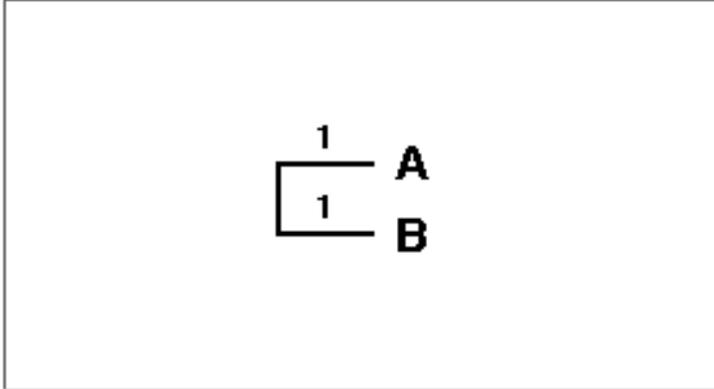


<p>Cycle 1</p>	<table border="1"> <thead> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <th>B</th> <td>2</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <th>C</th> <td>4</td> <td>4</td> <td></td> <td></td> <td></td> </tr> <tr> <th>D</th> <td>6</td> <td>6</td> <td>6</td> <td></td> <td></td> </tr> <tr> <th>E</th> <td>6</td> <td>6</td> <td>6</td> <td>4</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A	B	C	D	E	B	2					C	4	4				D	6	6	6			E	6	6	6	4		F	8	8	8	8	8	
	A	B	C	D	E																																	
B	2																																					
C	4	4																																				
D	6	6	6																																			
E	6	6	6	4																																		
F	8	8	8	8	8																																	
<p>Cycle 2</p>	<table border="1"> <thead> <tr> <th></th> <th>A,B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <th>C</th> <td>4</td> <td></td> <td></td> <td></td> </tr> <tr> <th>D</th> <td>6</td> <td>6</td> <td></td> <td></td> </tr> <tr> <th>E</th> <td>6</td> <td>6</td> <td>4</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A,B	C	D	E	C	4				D	6	6			E	6	6	4		F	8	8	8	8												
	A,B	C	D	E																																		
C	4																																					
D	6	6																																				
E	6	6	4																																			
F	8	8	8	8																																		
<p>Cycle 3</p>	<table border="1"> <thead> <tr> <th></th> <th>A,B</th> <th>C</th> <th>D,E</th> </tr> </thead> <tbody> <tr> <th>C</th> <td>4</td> <td></td> <td></td> </tr> <tr> <th>D,E</th> <td></td> <td></td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td></td> </tr> </tbody> </table>		A,B	C	D,E	C	4			D,E				F	8	8		<p>Nouvelle matrice?</p>																				
	A,B	C	D,E																																			
C	4																																					
D,E																																						
F	8	8																																				

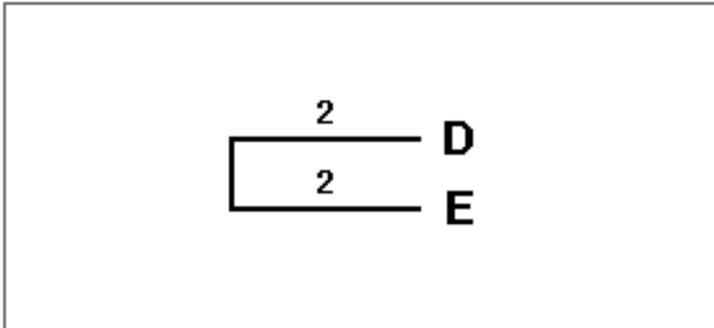
<p>Cycle 1</p>	<table border="1"> <thead> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <th>B</th> <td>2</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <th>C</th> <td>4</td> <td>4</td> <td></td> <td></td> <td></td> </tr> <tr> <th>D</th> <td>6</td> <td>6</td> <td>6</td> <td></td> <td></td> </tr> <tr> <th>E</th> <td>6</td> <td>6</td> <td>6</td> <td>4</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A	B	C	D	E	B	2					C	4	4				D	6	6	6			E	6	6	6	4		F	8	8	8	8	8	 <p>A diagram showing a horizontal line with 'A' at the right end and 'B' at the left end. A vertical line segment of length 1 is drawn above the horizontal line, and another vertical line segment of length 1 is drawn below the horizontal line, forming a rectangular frame around the branch.</p>
	A	B	C	D	E																																	
B	2																																					
C	4	4																																				
D	6	6	6																																			
E	6	6	6	4																																		
F	8	8	8	8	8																																	
<p>Cycle 2</p>	<table border="1"> <thead> <tr> <th></th> <th>A,B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <th>C</th> <td>4</td> <td></td> <td></td> <td></td> </tr> <tr> <th>D</th> <td>6</td> <td>6</td> <td></td> <td></td> </tr> <tr> <th>E</th> <td>6</td> <td>6</td> <td>4</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A,B	C	D	E	C	4				D	6	6			E	6	6	4		F	8	8	8	8	 <p>A diagram showing a horizontal line with 'D' at the right end and 'E' at the left end. A vertical line segment of length 2 is drawn above the horizontal line, and another vertical line segment of length 2 is drawn below the horizontal line, forming a rectangular frame around the branch.</p>											
	A,B	C	D	E																																		
C	4																																					
D	6	6																																				
E	6	6	4																																			
F	8	8	8	8																																		
<p>Cycle 3</p>	<table border="1"> <thead> <tr> <th></th> <th>A,B</th> <th>C</th> <th>D,E</th> </tr> </thead> <tbody> <tr> <th>C</th> <td>4</td> <td></td> <td></td> </tr> <tr> <th>D,E</th> <td>6</td> <td>6</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A,B	C	D,E	C	4			D,E	6	6		F	8	8	8	<p>Prochain à brancher? Longueur des branches?</p>																				
	A,B	C	D,E																																			
C	4																																					
D,E	6	6																																				
F	8	8	8																																			

<p>Cycle 1</p>	<table border="1"> <thead> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <th>B</th> <td>2</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <th>C</th> <td>4</td> <td>4</td> <td></td> <td></td> <td></td> </tr> <tr> <th>D</th> <td>6</td> <td>6</td> <td>6</td> <td></td> <td></td> </tr> <tr> <th>E</th> <td>6</td> <td>6</td> <td>6</td> <td>4</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A	B	C	D	E	B	2					C	4	4				D	6	6	6			E	6	6	6	4		F	8	8	8	8	8	
	A	B	C	D	E																																	
B	2																																					
C	4	4																																				
D	6	6	6																																			
E	6	6	6	4																																		
F	8	8	8	8	8																																	
<p>Cycle 2</p>	<table border="1"> <thead> <tr> <th></th> <th>A,B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <th>C</th> <td>4</td> <td></td> <td></td> <td></td> </tr> <tr> <th>D</th> <td>6</td> <td>6</td> <td></td> <td></td> </tr> <tr> <th>E</th> <td>6</td> <td>6</td> <td>4</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A,B	C	D	E	C	4				D	6	6			E	6	6	4		F	8	8	8	8												
	A,B	C	D	E																																		
C	4																																					
D	6	6																																				
E	6	6	4																																			
F	8	8	8	8																																		
<p>Cycle 3</p>	<table border="1"> <thead> <tr> <th></th> <th>A,B</th> <th>C</th> <th>D,E</th> </tr> </thead> <tbody> <tr> <th>C</th> <td>4</td> <td></td> <td></td> </tr> <tr> <th>D,E</th> <td>6</td> <td>6</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A,B	C	D,E	C	4			D,E	6	6		F	8	8	8																					
	A,B	C	D,E																																			
C	4																																					
D,E	6	6																																				
F	8	8	8																																			

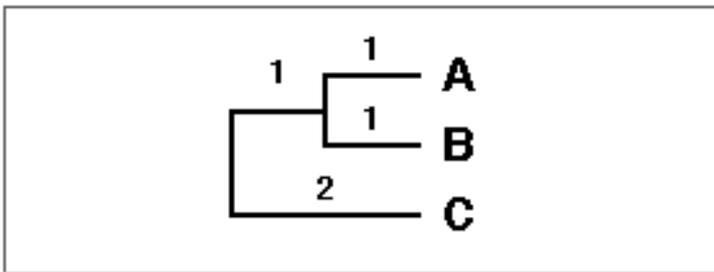
		A	B	C	D	E
Cycle 1	B	2				
	C	4	4			
	D	6	6	6		
	E	6	6	6	4	
	F	8	8	8	8	8



		A,B	C	D	E
Cycle 2	C	4			
	D	6	6		
	E	6	6	4	
	F	8	8	8	8



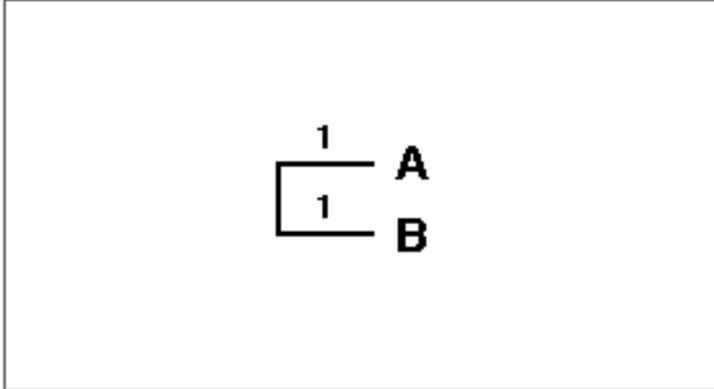
		A,B	C	D,E
Cycle 3	C	4		
	D,E	6	6	
	F	8	8	8



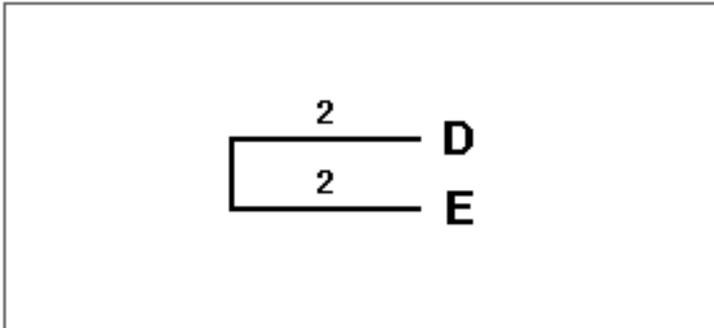
		AB,C	D,E
Cycle 4	D,E		
	F		8

Nouvelle matrice?

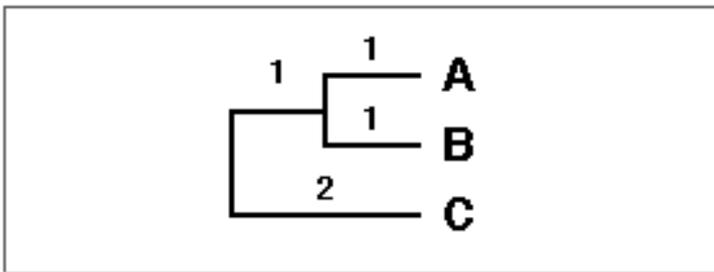
		A	B	C	D	E
Cycle 1	B	2				
	C	4	4			
	D	6	6	6		
	E	6	6	6	4	
	F	8	8	8	8	8



		A,B	C	D	E
Cycle 2	C	4			
	D	6	6		
	E	6	6	4	
	F	8	8	8	8



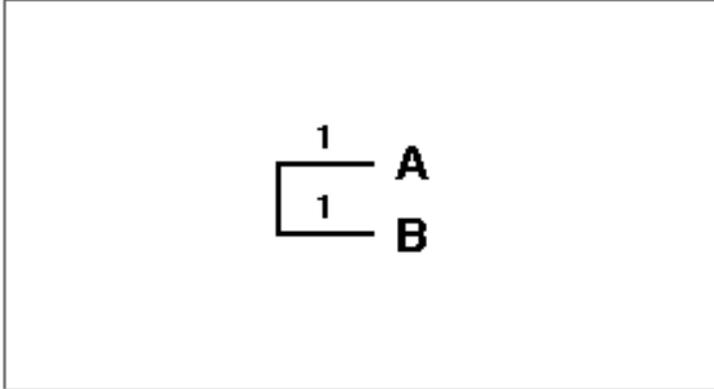
		A,B	C	D,E
Cycle 3	C	4		
	D,E	6	6	
	F	8	8	8



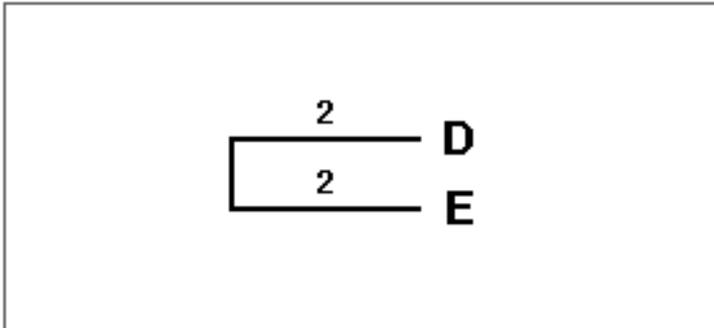
		AB,C	D,E
Cycle 4	D,E	6	
	F	8	8

Prochain à brancher?
Longueur des branches?

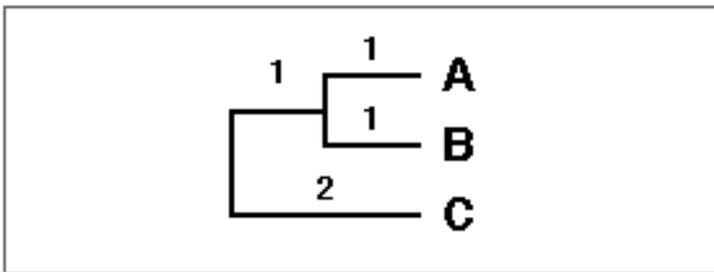
		A	B	C	D	E
Cycle 1	B	2				
	C	4	4			
	D	6	6	6		
	E	6	6	6	4	
	F	8	8	8	8	8



		A,B	C	D	E
Cycle 2	C	4			
	D	6	6		
	E	6	6	4	
	F	8	8	8	8

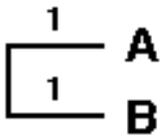
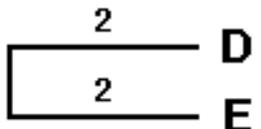
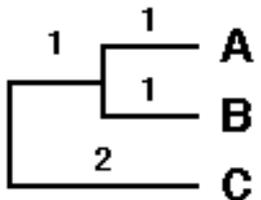
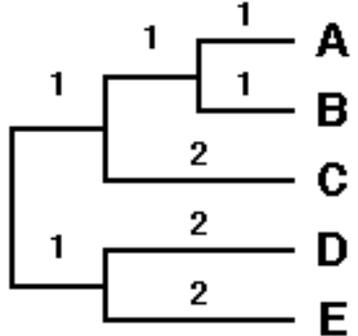


		A,B	C	D,E
Cycle 3	C	4		
	D,E	6	6	
	F	8	8	8



		AB,C	D,E
Cycle 4	D,E	6	
	F	8	8

Prochain à brancher?
Longueur des branches?

Cycle 1	<table border="1"> <thead> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <th>B</th> <td>2</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <th>C</th> <td>4</td> <td>4</td> <td></td> <td></td> <td></td> </tr> <tr> <th>D</th> <td>6</td> <td>6</td> <td>6</td> <td></td> <td></td> </tr> <tr> <th>E</th> <td>6</td> <td>6</td> <td>6</td> <td>4</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A	B	C	D	E	B	2					C	4	4				D	6	6	6			E	6	6	6	4		F	8	8	8	8	8	
	A	B	C	D	E																																	
B	2																																					
C	4	4																																				
D	6	6	6																																			
E	6	6	6	4																																		
F	8	8	8	8	8																																	
Cycle 2	<table border="1"> <thead> <tr> <th></th> <th>A,B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <th>C</th> <td>4</td> <td></td> <td></td> <td></td> </tr> <tr> <th>D</th> <td>6</td> <td>6</td> <td></td> <td></td> </tr> <tr> <th>E</th> <td>6</td> <td>6</td> <td>4</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A,B	C	D	E	C	4				D	6	6			E	6	6	4		F	8	8	8	8												
	A,B	C	D	E																																		
C	4																																					
D	6	6																																				
E	6	6	4																																			
F	8	8	8	8																																		
Cycle 3	<table border="1"> <thead> <tr> <th></th> <th>A,B</th> <th>C</th> <th>D,E</th> </tr> </thead> <tbody> <tr> <th>C</th> <td>4</td> <td></td> <td></td> </tr> <tr> <th>D,E</th> <td>6</td> <td>6</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> <td>8</td> </tr> </tbody> </table>		A,B	C	D,E	C	4			D,E	6	6		F	8	8	8																					
	A,B	C	D,E																																			
C	4																																					
D,E	6	6																																				
F	8	8	8																																			
Cycle 4	<table border="1"> <thead> <tr> <th></th> <th>AB,C</th> <th>D,E</th> </tr> </thead> <tbody> <tr> <th>D,E</th> <td>6</td> <td></td> </tr> <tr> <th>F</th> <td>8</td> <td>8</td> </tr> </tbody> </table>		AB,C	D,E	D,E	6		F	8	8																												
	AB,C	D,E																																				
D,E	6																																					
F	8	8																																				

<p>Cycle 4</p>	<table border="0"> <tr> <td></td> <td></td> <td>AB,C</td> <td>D,E</td> </tr> <tr> <td>D,E</td> <td></td> <td>6</td> <td></td> </tr> <tr> <td>F</td> <td></td> <td>8</td> <td>8</td> </tr> </table>			AB,C	D,E	D,E		6		F		8	8	
		AB,C	D,E											
D,E		6												
F		8	8											

<p>Cycle 5</p>	<table border="0"> <tr> <td></td> <td>ABC,DE</td> </tr> <tr> <td>F</td> <td></td> </tr> </table>		ABC,DE	F		<p>Nouvelle matrice?</p>
	ABC,DE					
F						

<p>Cycle 4</p>	<table border="1"> <tr> <td></td> <td>AB,C</td> <td>D,E</td> </tr> <tr> <td>D,E</td> <td>6</td> <td></td> </tr> <tr> <td>F</td> <td>8</td> <td>8</td> </tr> </table>		AB,C	D,E	D,E	6		F	8	8	
	AB,C	D,E									
D,E	6										
F	8	8									

<p>Cycle 5</p>	<table border="1"> <tr> <td></td> <td>ABC,DE</td> </tr> <tr> <td>F</td> <td>8</td> </tr> </table>		ABC,DE	F	8	<p>Prochain à brancher? Longueur des branches?</p>
	ABC,DE					
F	8					

<p>Cycle 4</p>	<table border="0"> <tr> <td></td> <td>AB,C</td> <td>D,E</td> </tr> <tr> <td>D,E</td> <td>6</td> <td></td> </tr> <tr> <td>F</td> <td>8</td> <td>8</td> </tr> </table>		AB,C	D,E	D,E	6		F	8	8	
	AB,C	D,E									
D,E	6										
F	8	8									

<p>Cycle 5</p>	<table border="0"> <tr> <td></td> <td>ABC,DE</td> </tr> <tr> <td>F</td> <td>8</td> </tr> </table>		ABC,DE	F	8	
	ABC,DE					
F	8					

Méthodes basées sur les distances : UPGMA (Unweight Pair Group Method with Arithmetic mean) :

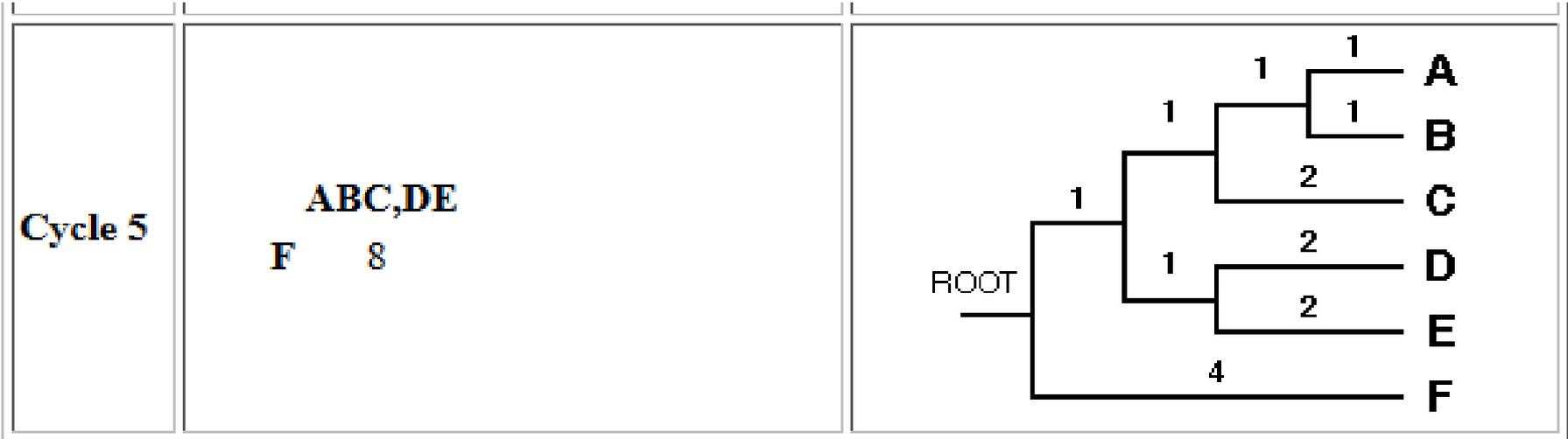
Enracinez l'arbre avec la méthode du "mid-point rooting" : la racine de l'arbre est mise à équidistance des taxa les plus distants.



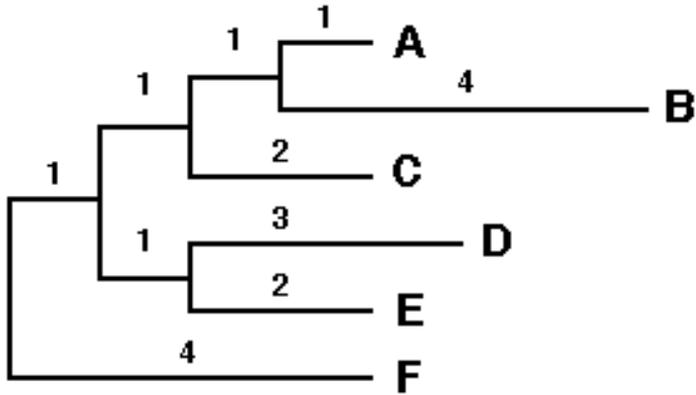
Méthodes basées sur les distances : UPGMA (Unweight Pair Group Method with Arithmetic mean) :

Enracinez l'arbre avec la méthode du "mid-point rooting" : la racine de l'arbre est mise à équidistance des taxa les plus distants.

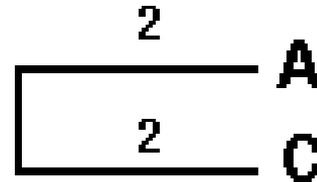
soit $(ABCDE), F / 2 = 4$



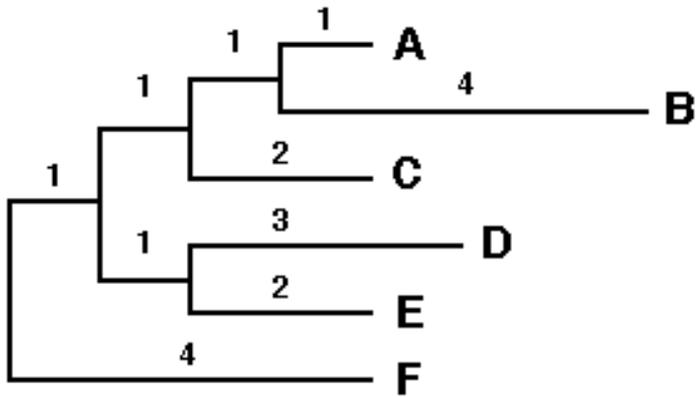
Les inconvénients de la méthode UPGMA



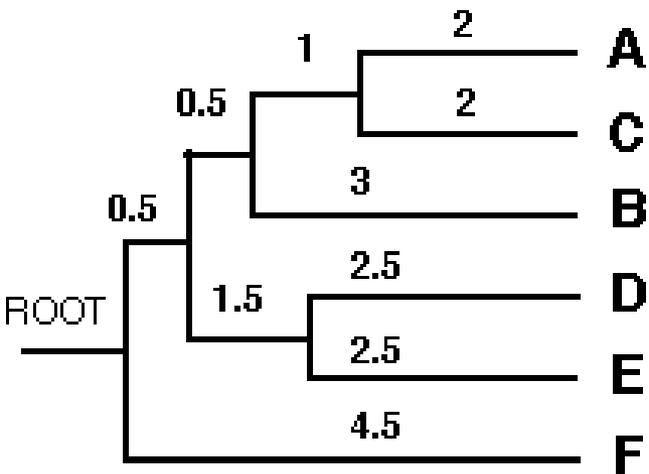
	A	B	C	D	E
B	5				
C	4	7			
D	7	10	7		
E	6	9	6	5	
F	8	11	8	9	8



Les inconvénients de la méthode UPGMA

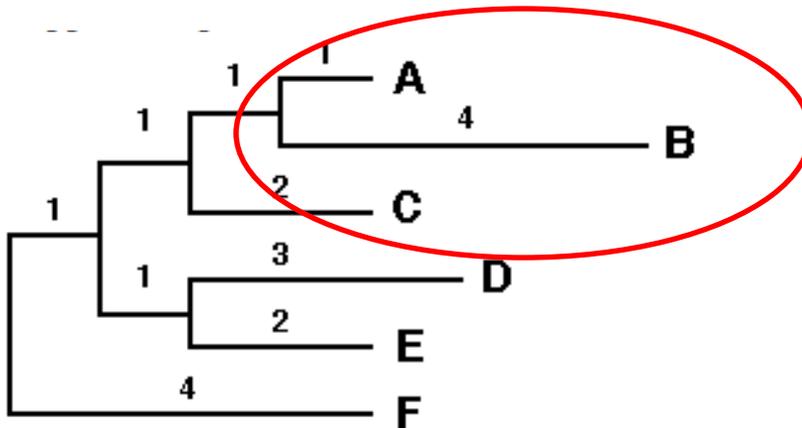


	A	B	C	D	E
B	5				
C	4	7			
D	7	10	7		
E	6	9	6	5	
F	8	11	8	9	8

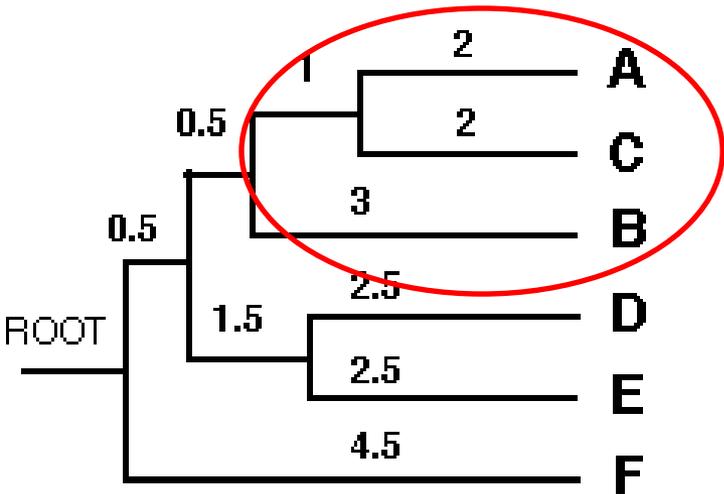


Que constatez-vous ?

Les inconvénients de la méthode UPGMA



	A	B	C	D	E
B	5				
C	4	7			
D	7	10	7		
E	6	9	6	5	
F	8	11	8	9	8



La topologie obtenue est fautive à cause des taux inégaux de substitutions :

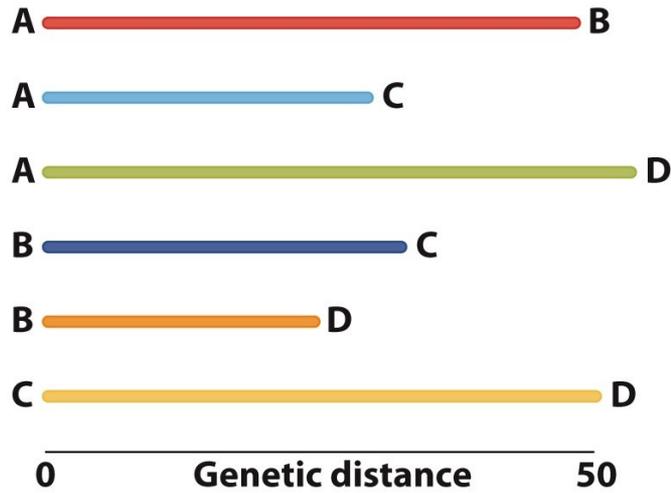
Depuis que A et B ont divergé, B a accumulé beaucoup plus de mutations que A.

C'est l'inconvénient majeur de la méthode UPGMA : sa sensibilité à des taux de mutations différents sur les différentes branches.

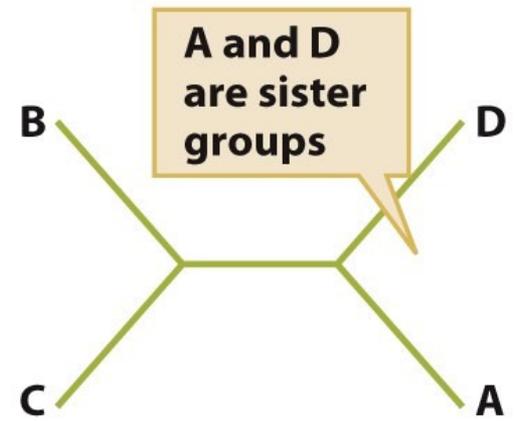
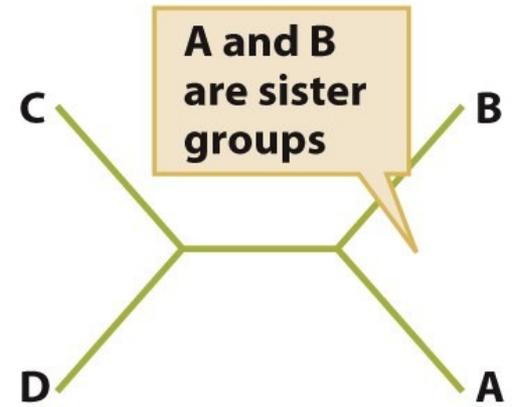
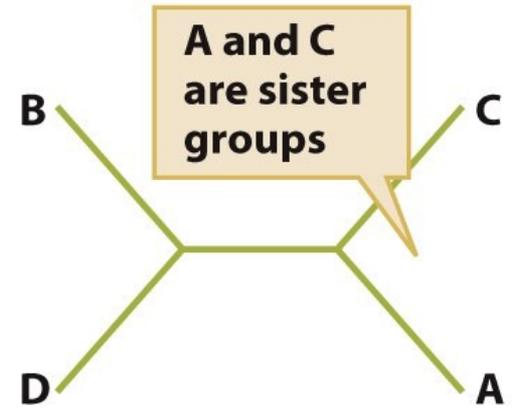
Méthodes basées sur les distances : NJ (Neighbor-Joining)

- Moins sensible aux variations de taux de substitutions, sans pour autant examiner tous les arbres possibles
- Estime les longueurs de branches de l'arbre en minimisant la somme des carrés des différences entre les distances mesurées et les distances dans l'arbre

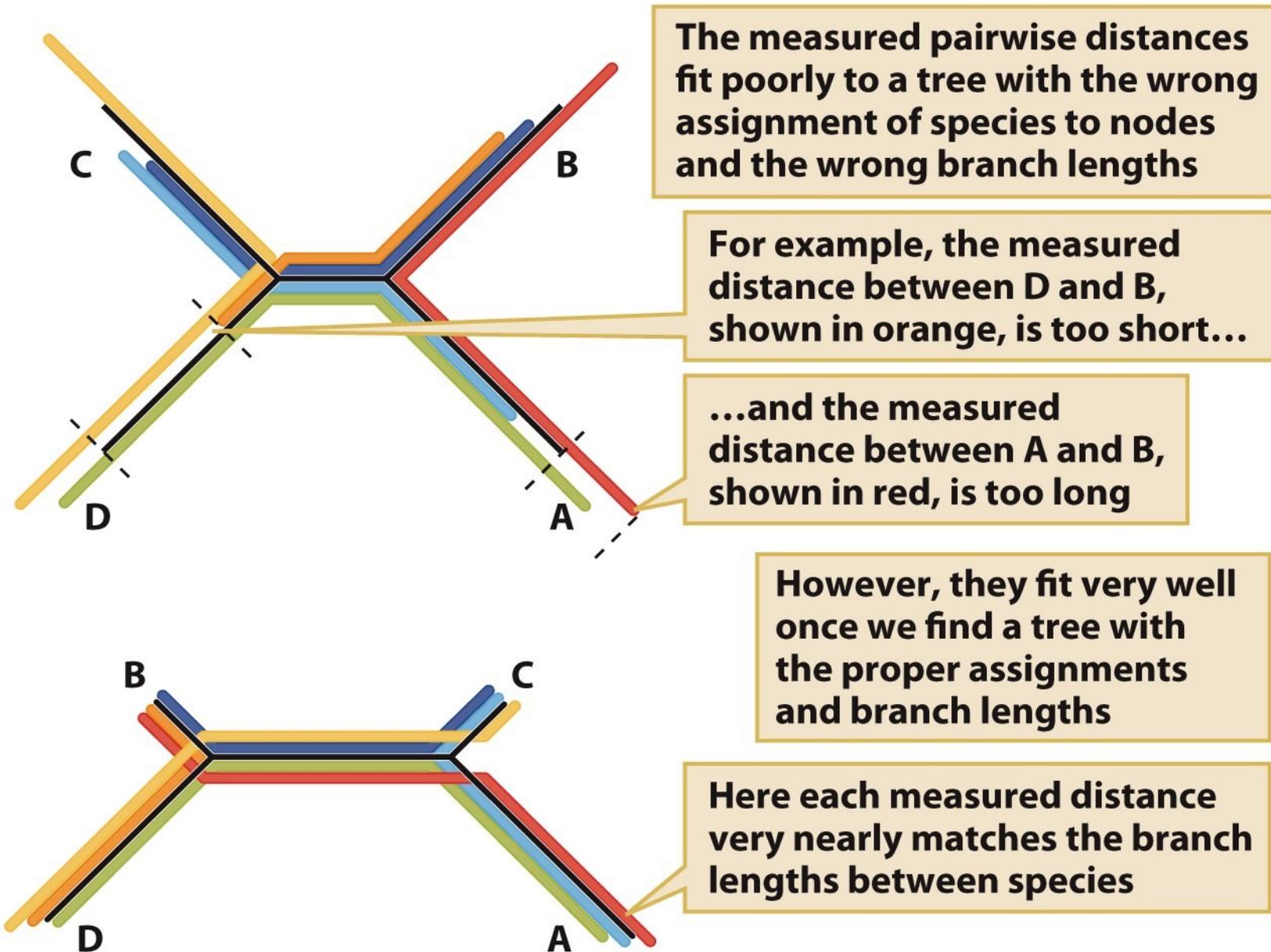
Méthodes basées sur les distances : NJ (Neighbor-Joining)



	A	B	C	D
A	0	48	29	53
B	48	0	32	24
C	29	32	0	50
D	53	24	50	0



Méthodes basées sur les distances : NJ (Neighbor-Joining)



Méthodes basées sur les distances : NJ (Neighbor-Joining)

- On construit au départ un arbre en étoile.
- On recherche d'abord quelle paire de taxa doit être reliée de façon à minimiser la longueur totale de l'arbre.
- Saitou et Nei ont montré qu'il s'agit de la paire qui minimise la distance:

$$M(i,j) = d(i,j) - [r(i) + r(j)] / (N-2)$$

Avec :

$d(i,j)$ = nombre de substitutions (différences) entre i et j

$r(i)$ = somme des divergences du taxon i par rapport aux $N-1$ autres

N = nombre de taxa



Méthodes probabilistes

L'adéquation avec les données est optimisée pour toutes les espèces et pour chaque site de la séquence.

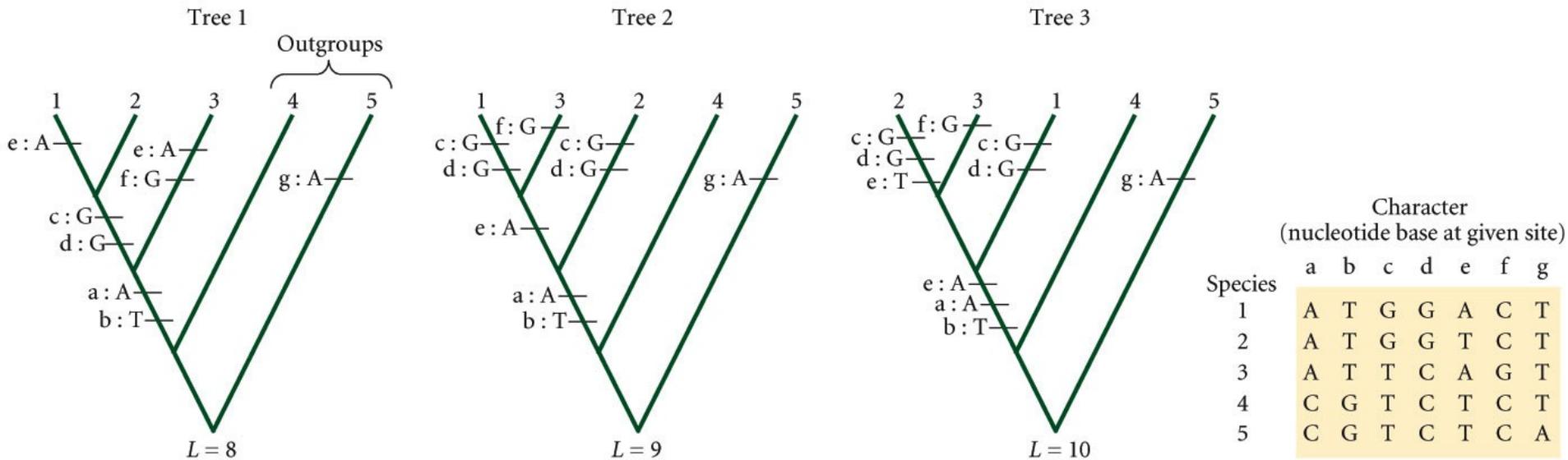
Maximum de parcimonie : nombre minimal de changements.

Maximum de vraisemblance : score (log) de vraisemblance $L = \Pr[\text{Données} | \text{Arbre}]$ optimisé sur la longueur des branches et les paramètres du modèle de substitution.

Bayésien : probabilité postérieure, calculée par l'intégration sur la longueur des branches et les paramètres de substitution.

$P(\text{arbre} | \text{Alignement}) = \frac{P(\text{Alignement} | \text{Arbre}) P(\text{arbre})}{P(\text{alignement})}$

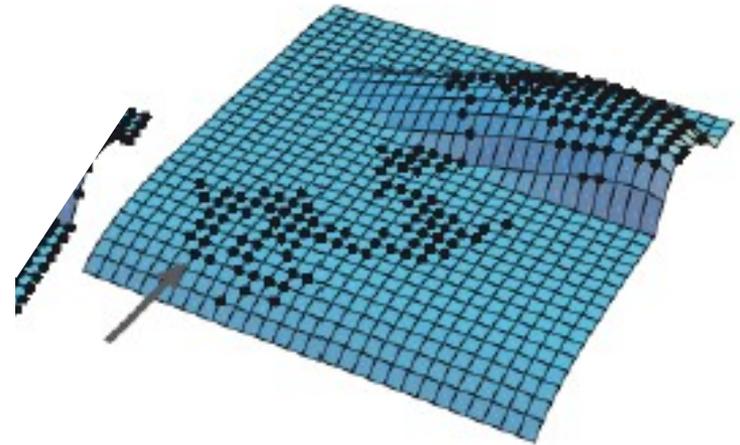
Reconstruction phylogénie par maximum de parcimonie



Minimiser le nombre de changements L (length)

Méthodes phylogénétiques : recherche d'arbres

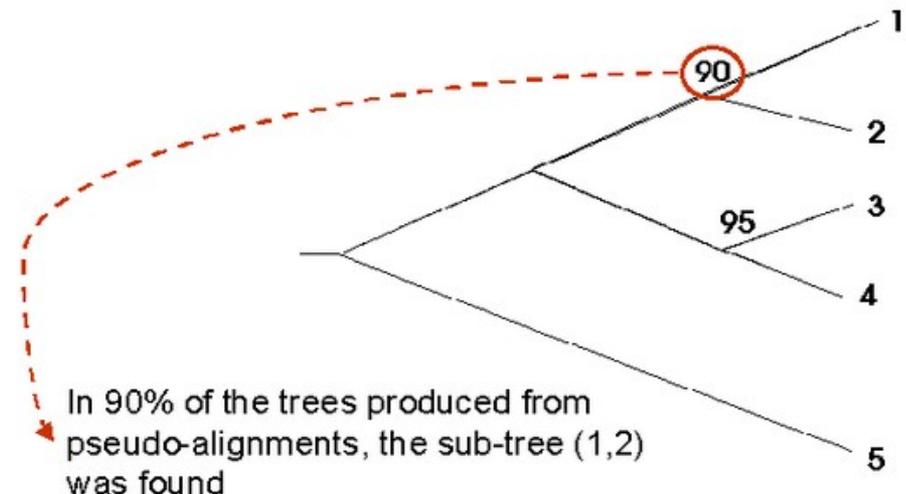
- **Exhaustive** – on obtient un score pour chaque arbre possible. On est sûr de trouver le meilleur arbre. Possible seulement pour des petits jeux de données
- **Heuristique** – algorithmes pour parcourir l'ensemble de l'espace des arbres mais pas chaque arbre possible
- **Stochastique** - recherche aléatoire dans l'espace des arbres. Les plus puissants sont les algorithmes bayésiens de MCMC (chaînes de Markov Monte Carlo)



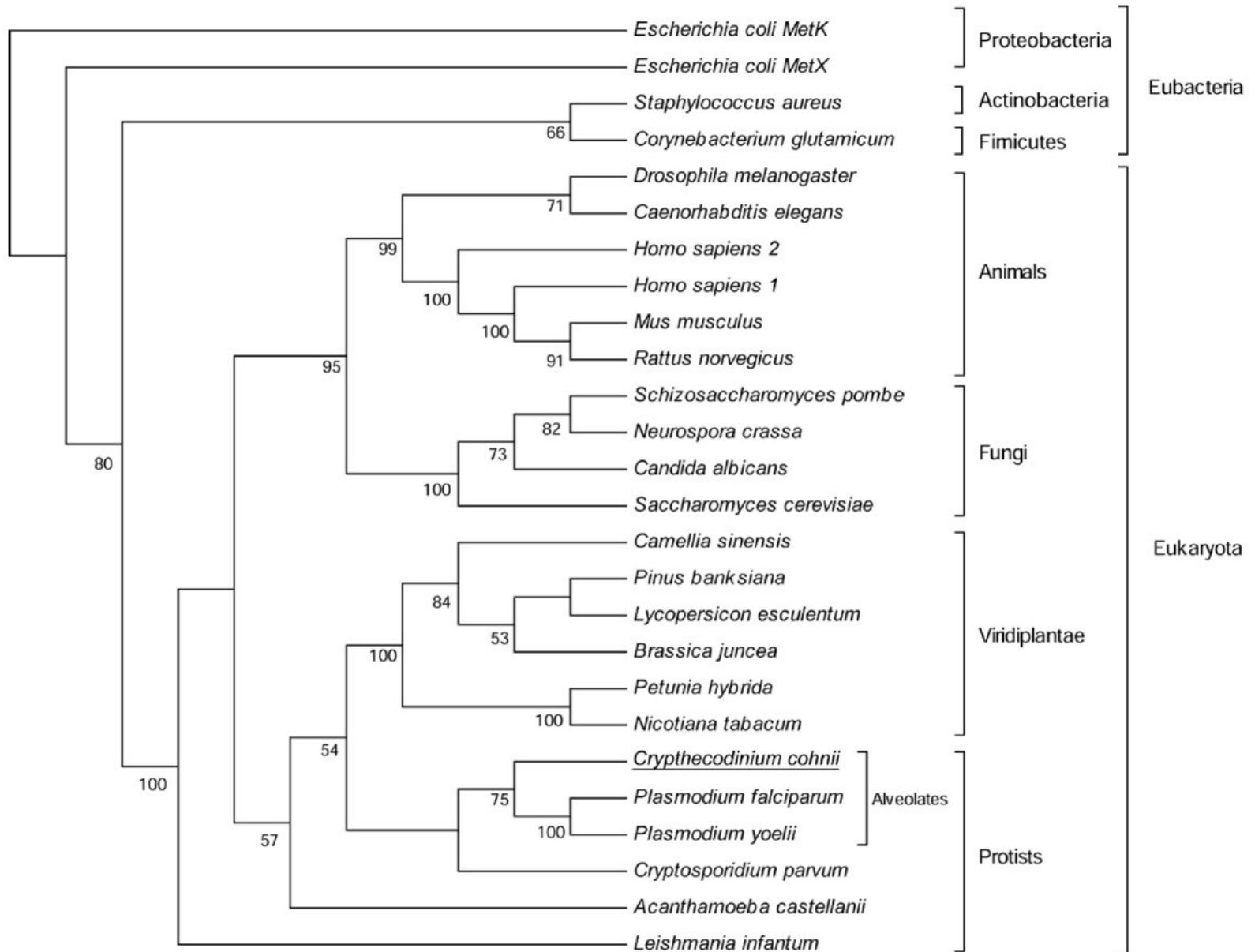
Le bootstrap

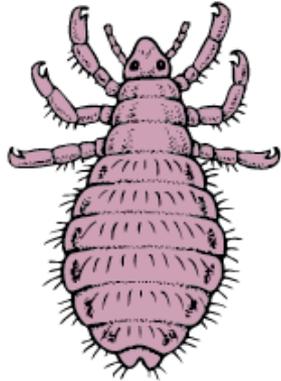
- Évalue la confiance que l'on peut avoir dans l'arbre obtenu
- 3 étapes réalisées chacune au moins 100 fois:
 - Réalisation d'un pseudo-alignement à partir des séquences d'origine en prenant arbitrairement n colonnes (avec remplacements) de l'alignement d'origine
 - Estimation de l'arbre obtenu A'
 - Comparaison des arbres A et A' : pour chaque sous-arbre de A , on regarde s'il est présent dans A' .

On compte ensuite pour chaque sous-arbre le nombre de fois où il est présent dans les A' . Cette fréquence avec laquelle on retrouve un sous-arbre est la valeur de bootstrap. Plus elle est élevée, plus la fiabilité de la branche est importante.

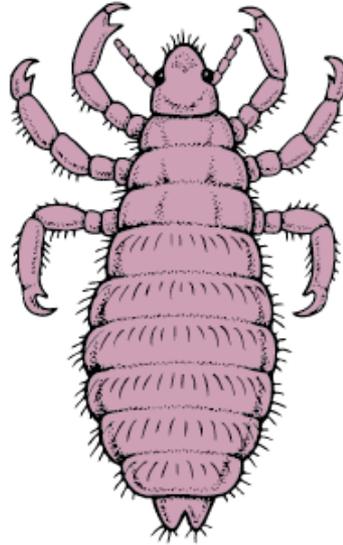


Valeurs de bootstrap

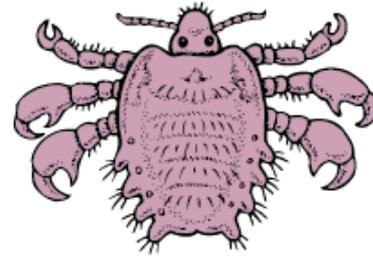




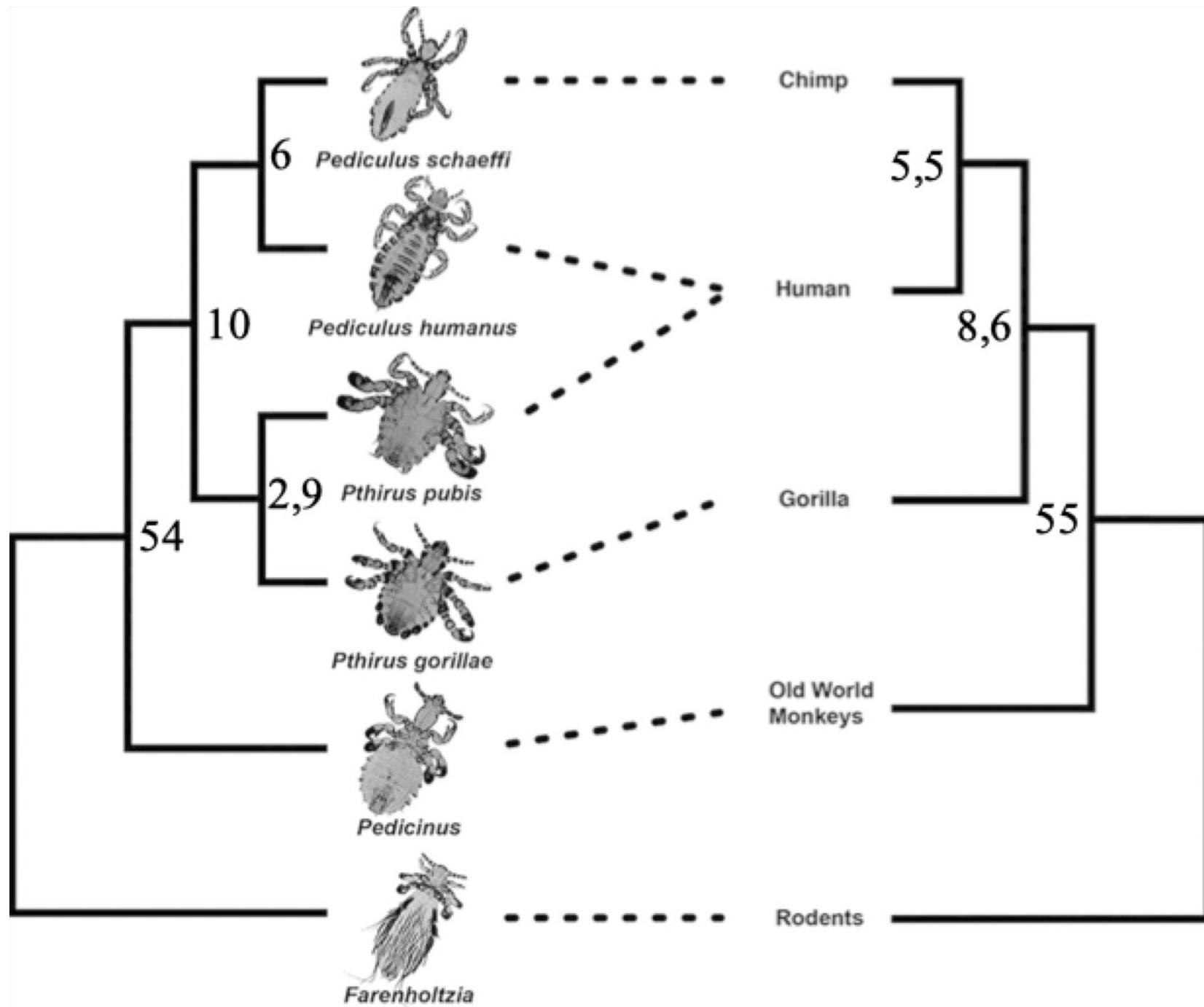
Pou de tête



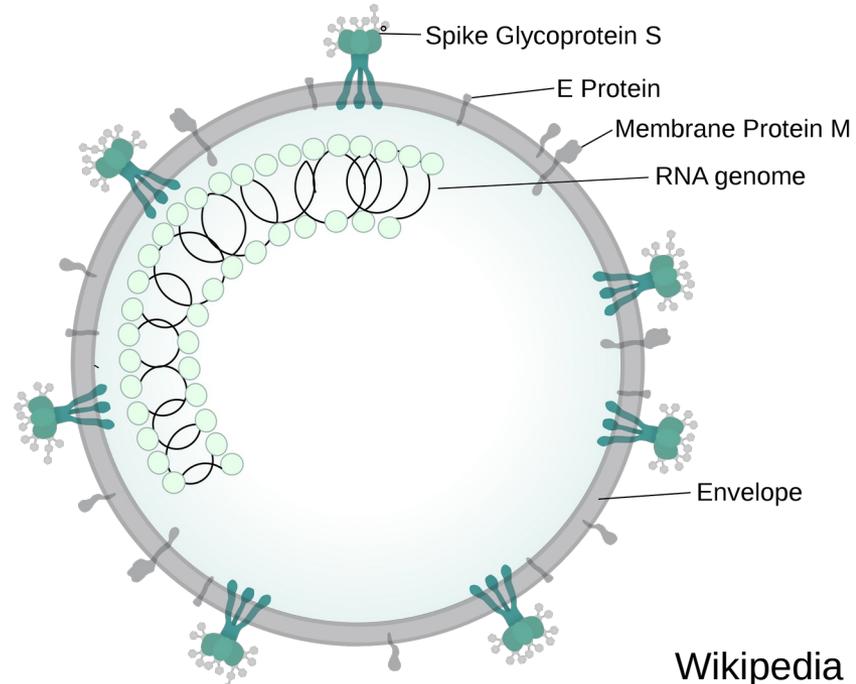
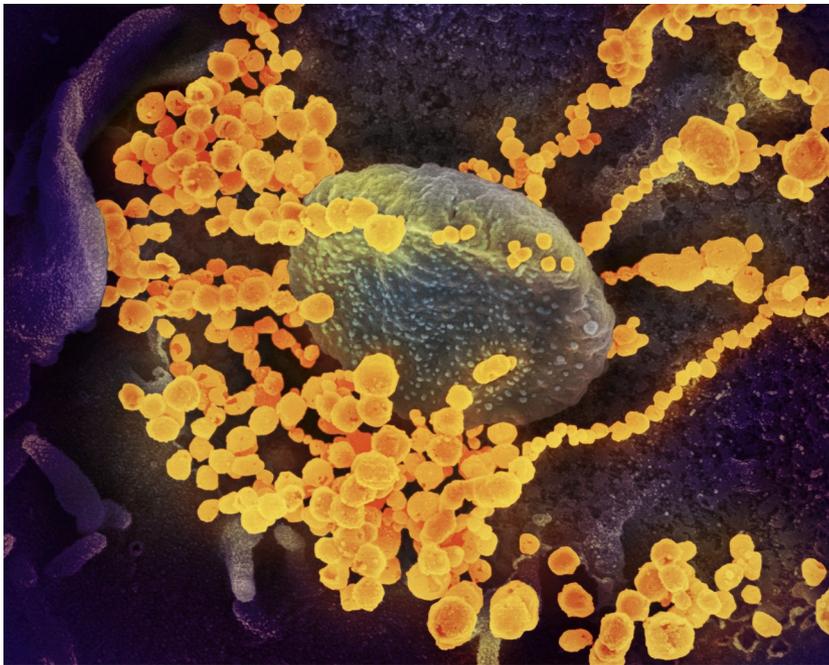
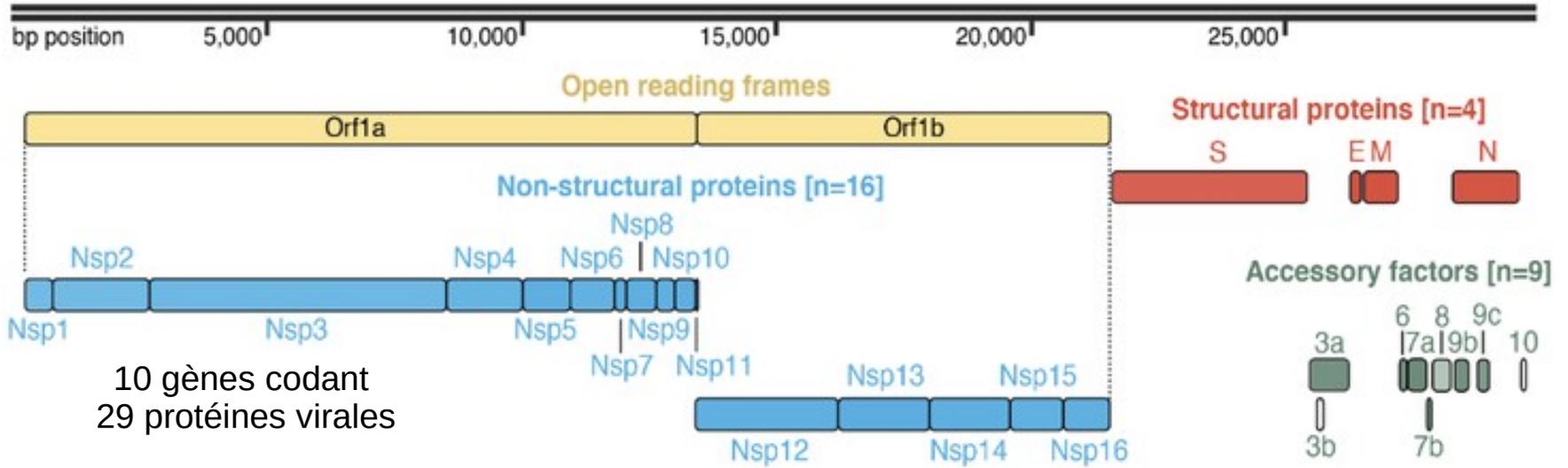
Pou de corps



Pou pubien

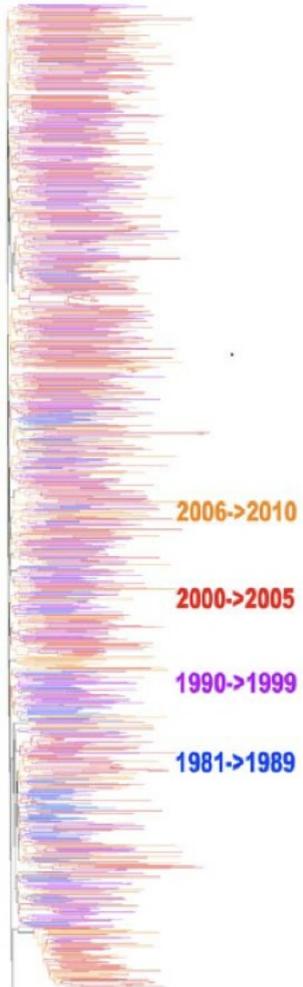


Génome de SARS-CoV-

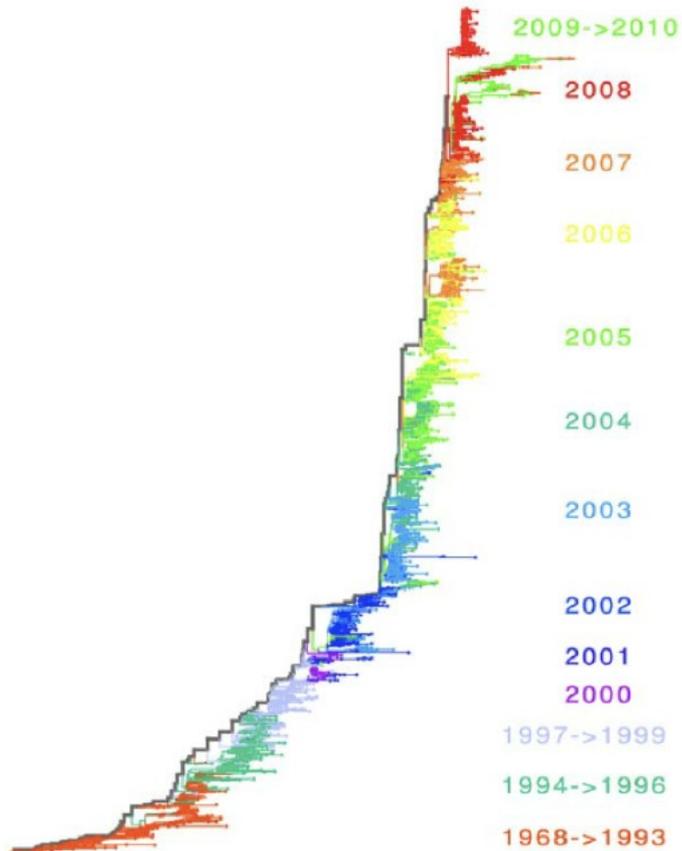


Phylogénies différentes selon les virus

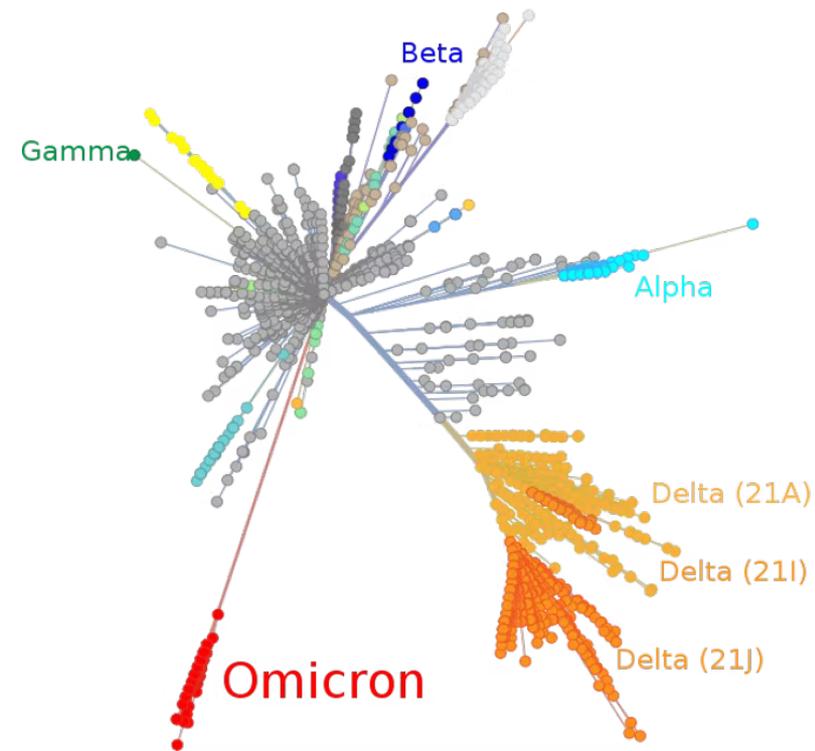
VIH



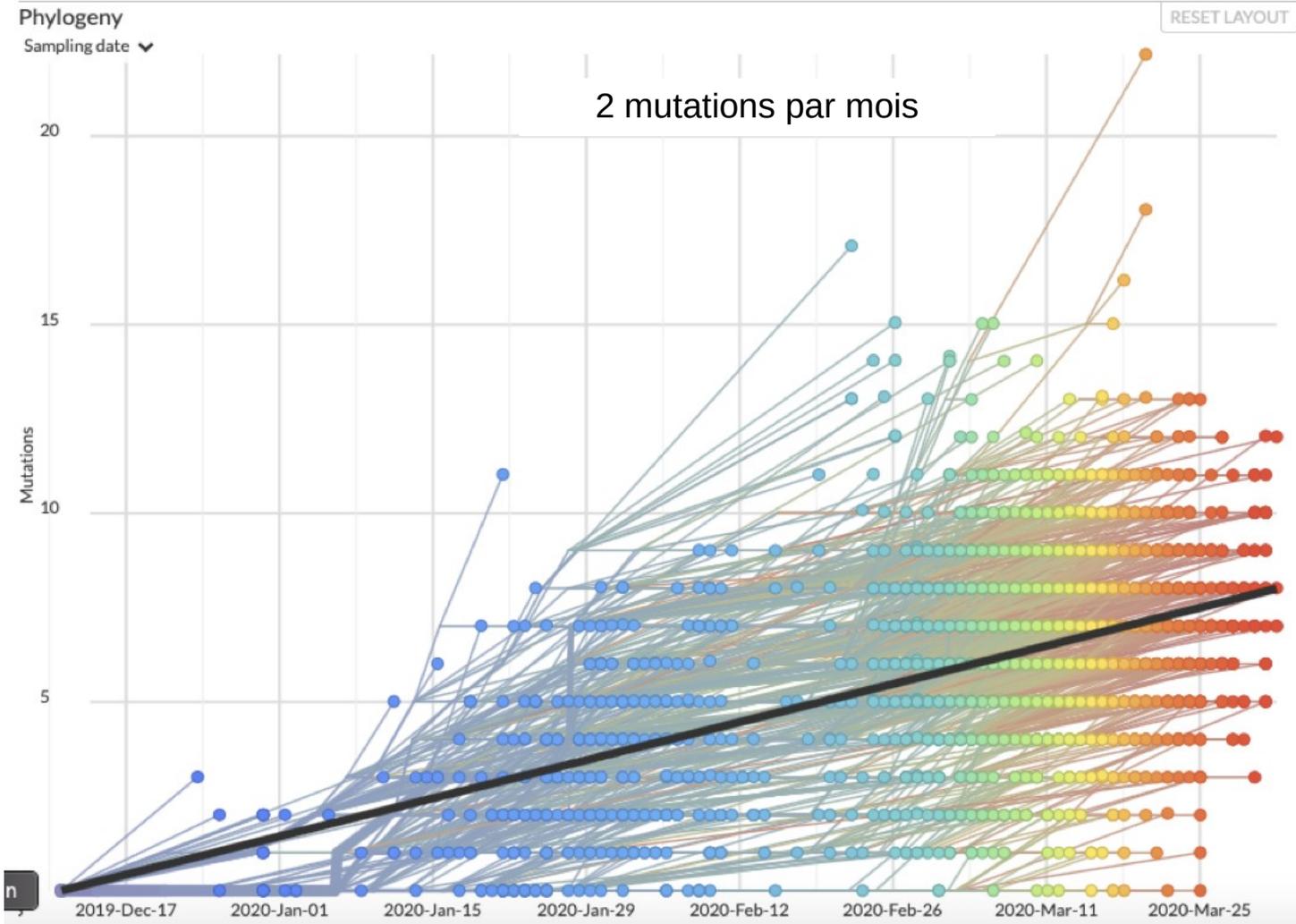
Grippe

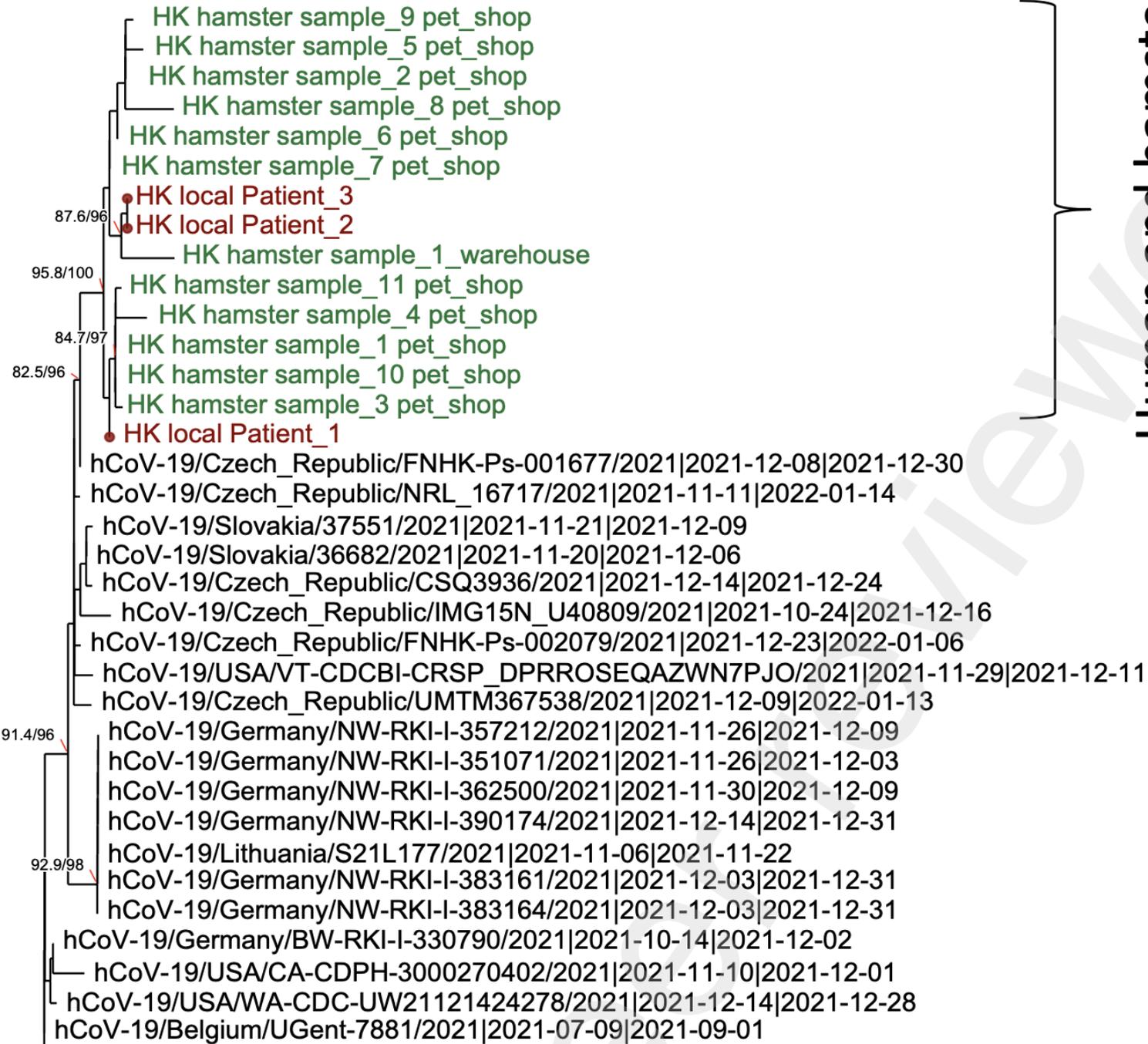


SARS-CoV-2



Estimation du taux de mutation pour SARS-CoV-2

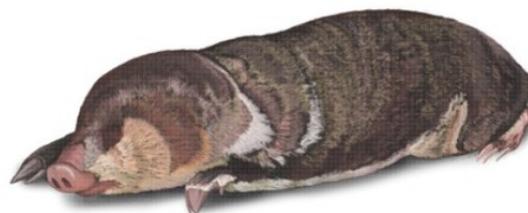
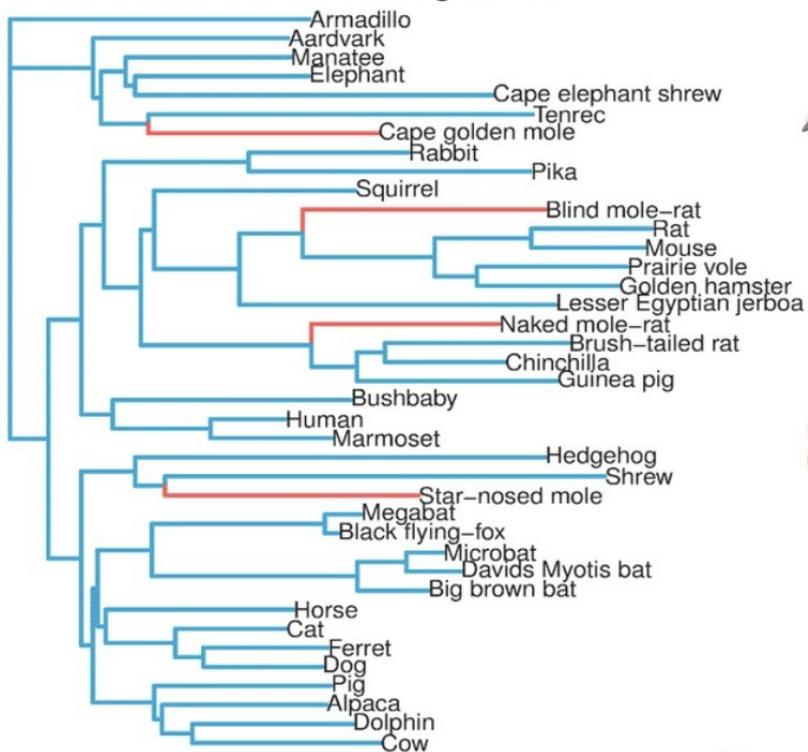




Human and hamster cases in this outbreak

A

Proteome-wide average tree



Cape golden mole



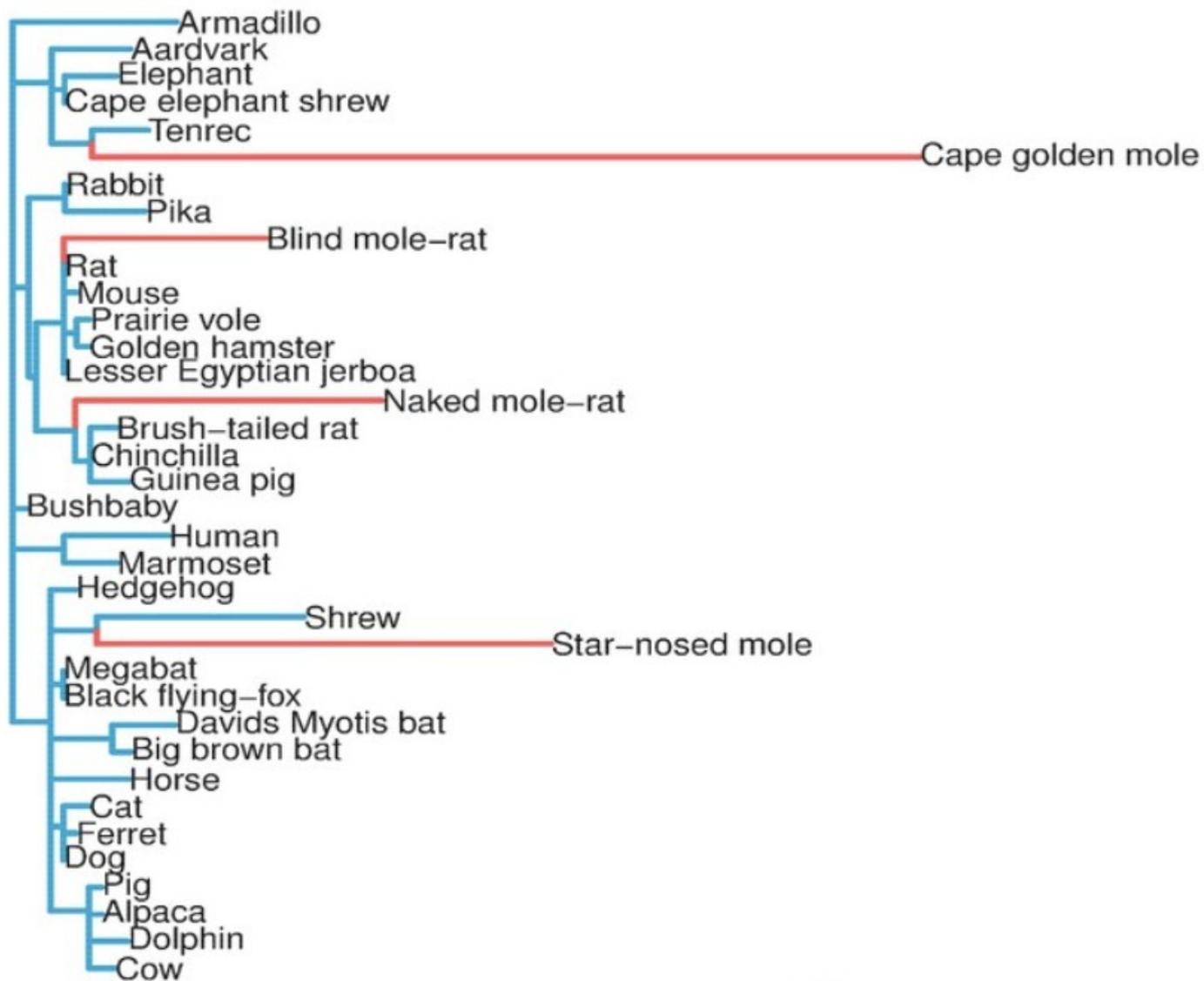
Blind mole-rat



Naked mole-rat



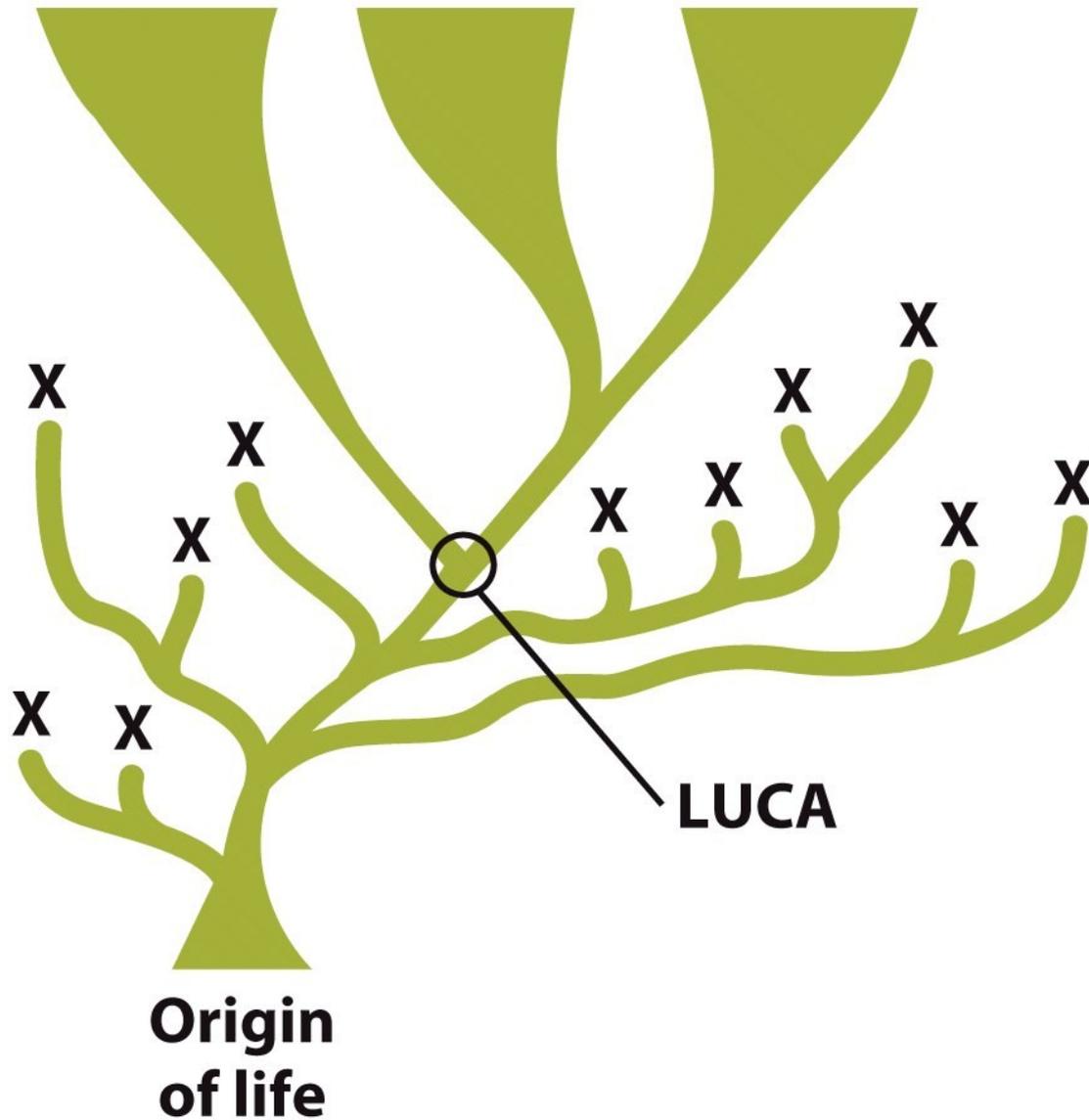
Star-nosed mole

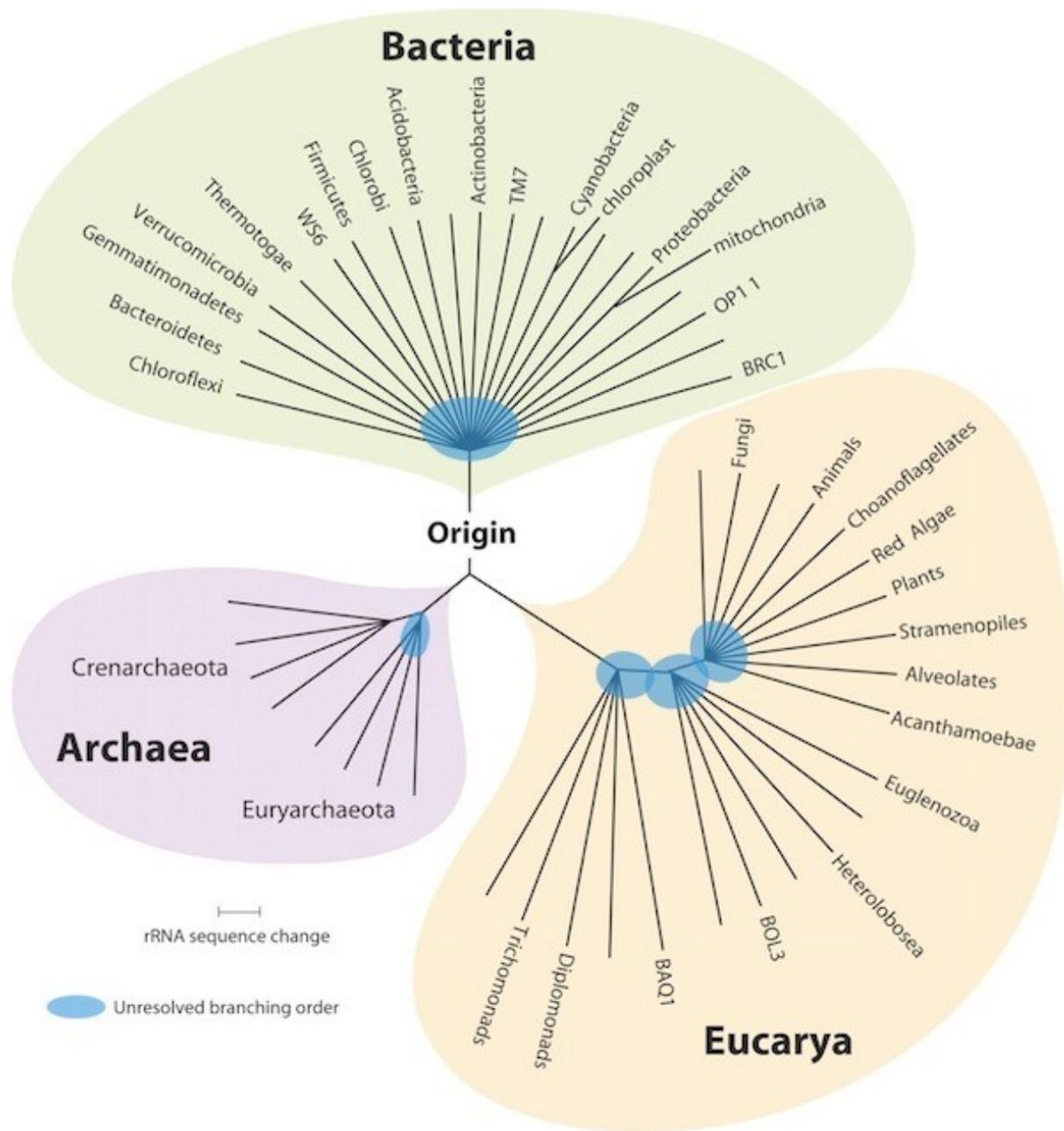


— 0.05

nombre moyen de changement d'acides aminés par site

Bacteria **Archaea** **Eukaryota**

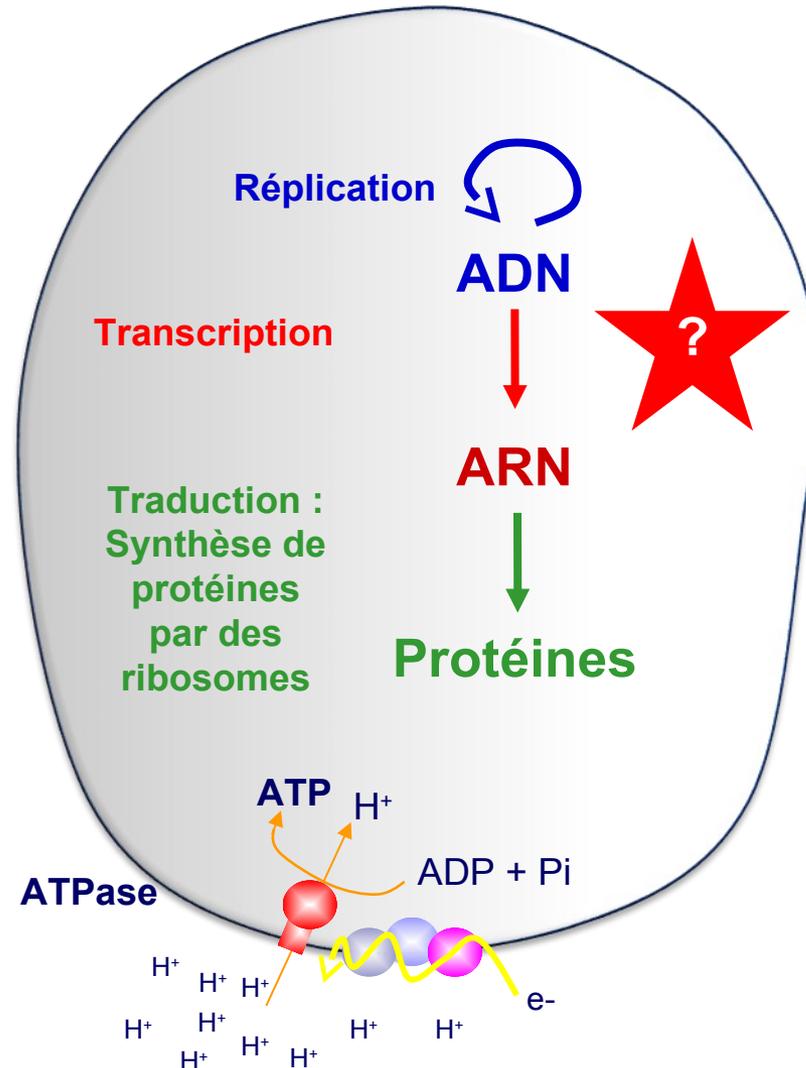




Portrait robot de LUCA

Organisme complexe,
génomme relativement
important
> 600 gènes

Synthèse d'ATP
par une ATPase
de membrane
(chaîne de
transport e⁻)



Cinq grandes extinctions

