

5. Phylogenetic trees

- The tree of life
- The tree, an abstract object
- Building an array of distances
- **The UPGMA algorithm**
- Differences are not always what they look like
- The diversity of bioinformatics algorithms
- The application domains in microbiology

UPGMA

- Unweighted Pair Group Method with Arithmetic Mean
- Quite a complicated name for a quite simple method
- Start with an array of distances

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

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

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



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



$$\begin{aligned} \text{dist}(F,C),A &= (\text{dist } F,A + \text{dist } C,A) / 2 = 4 \\ \text{dist}(F,C),B &= (\text{dist } F,B + \text{dist } C,B) / 2 = 6 \\ \text{dist}(F,C),D &= (\text{dist } F,D + \text{dist } C,D) / 2 = 6 \\ \text{dist}(F,C),E &= (\text{dist } F,E + \text{dist } C,E) / 2 = 8 \end{aligned}$$

	(F C)	A	B	D
A	4			
B	6	6		
D	6	6	4	
E	8	8	8	8



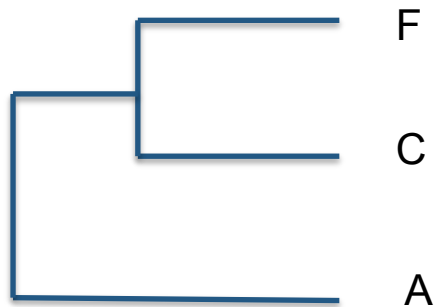
	(F C)	A	B	D
A	4			
B	6	6		
D	6	6	4	
E	8	8	8	8



	(F C)	A	B	D
A	4			
B	6	6		
D	6	6	4	
E	8	8	8	8



	(F C)	A	B	D
A	4			
B	6	6		
D	6	6	4	
E	8	8	8	8



((F C) A)

(D B)

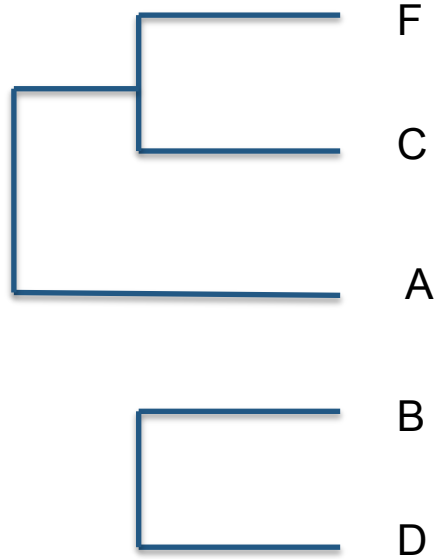
(D B)

6

E

8

8



((F C) A)

(D B)

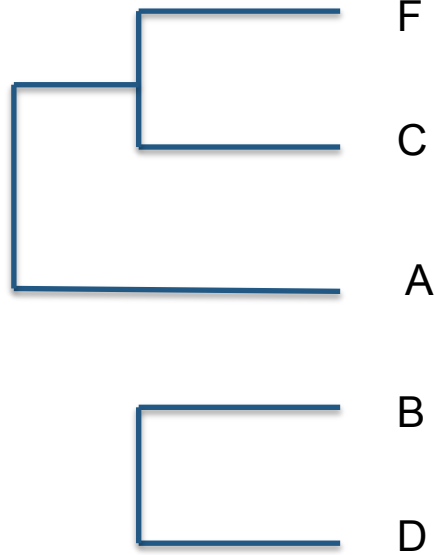
(D B)

6

E

8

8



((F C) A)

(D B)

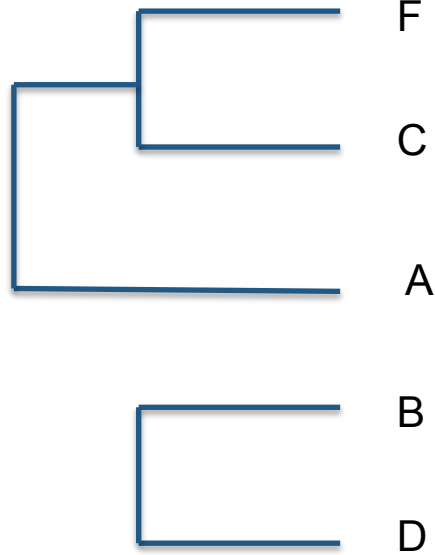
(D B)

6

E

8

8



((F C) A)

(D B)

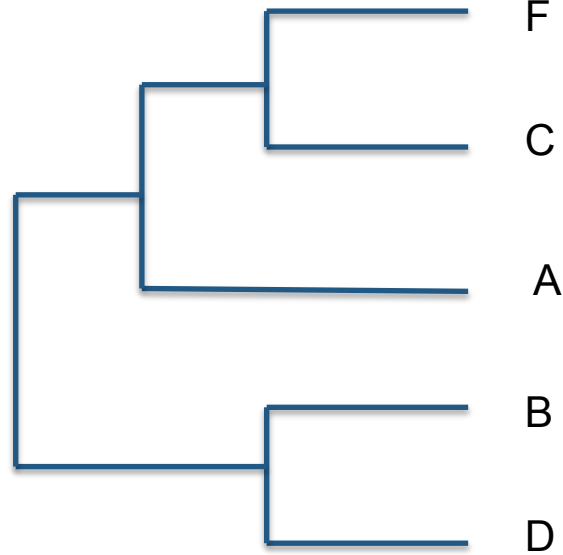
(D B)

6

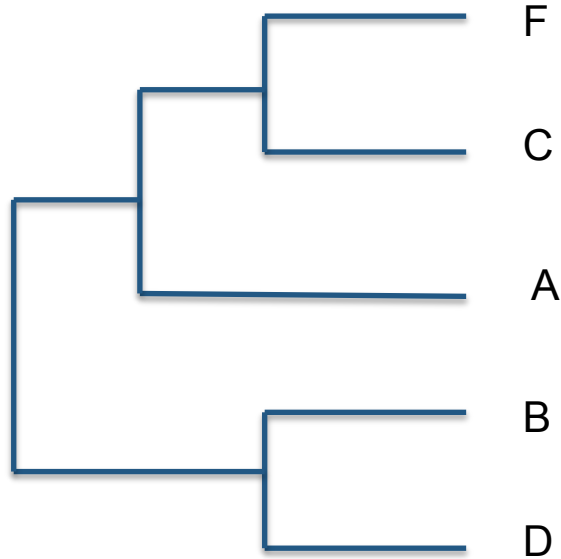
E

8

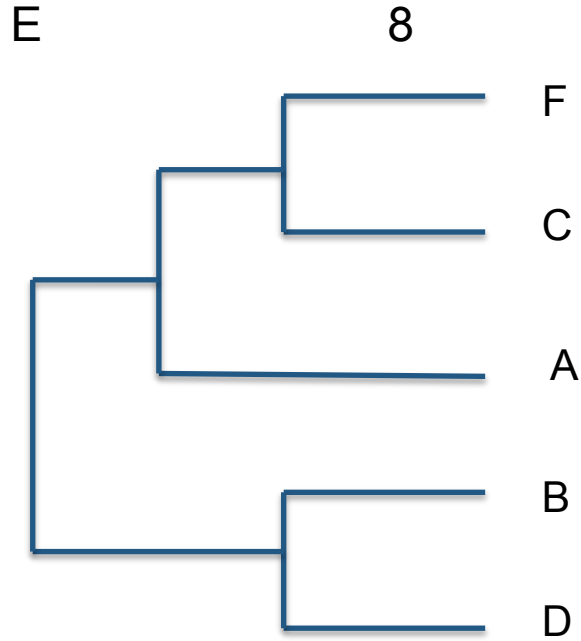
8

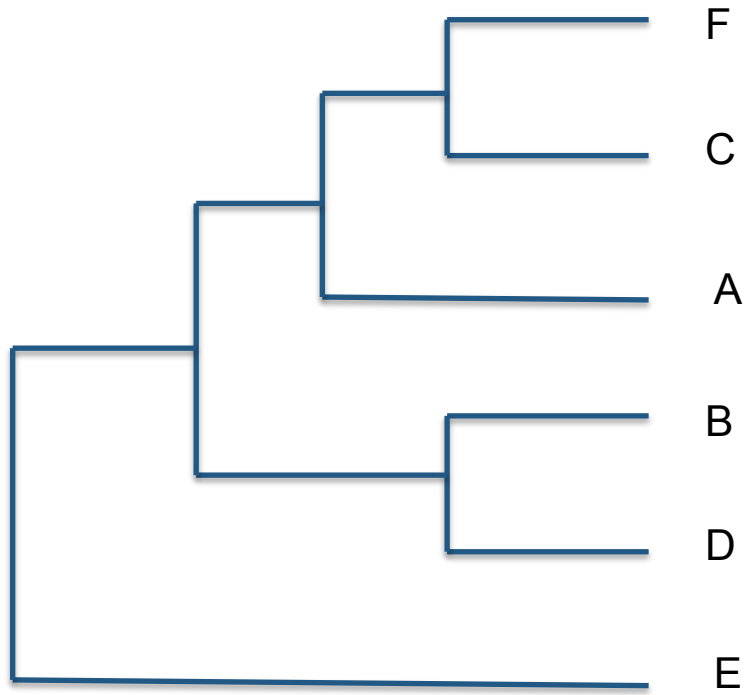


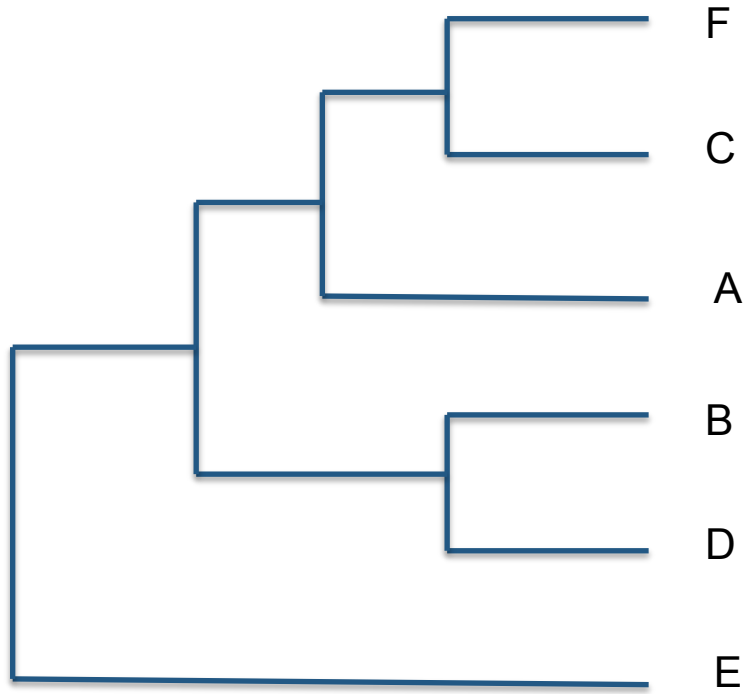
$$\begin{aligned} & \text{dist } (((F C) A)(D B)) E \\ &= (\text{dist } ((F C) A) E + \text{dist } (D B) E) / 2 \\ &= 8 \end{aligned}$$



((F C) A) (D B)







$((F\ C)\ A\ (B\ D)\ E)$