

Exercício 7

1. Um sistemata estudou as espécies vogais A, E, I, e U, e as espécies consoantes B, L, N, R e S. Sua análise filogenética resultou nas seguintes relações:
As espécies B, R e U formam um clado sem resolução interna, irmão de todos os demais. E é grupo irmão de S e ambos formam com L um grupo monofilético. Este grupo, juntamente com I e A, é grupo irmão de N.

i. Qual é a topologia que representa a relação entre esses terminais?

ii. Conservando as relações entre os grupos, disponha os terminais de maneira que as letras formem um táxon pertencente à uma categoria hierárquica Ordem (em Botânica) e que compreende a família Bruniaceae.

iii. Proponha dois esquemas de classificação para esta filogenia, um por sequenciação e outro por subordinação, que terminem em nível de espécie.

2. Que tipo de grupo (mono, para ou polifilético) forma o sufixo designativo da categoria do táxon revelado no item **ii** acima?

3. Qual seria a principal vantagem de uma e de outra classificação?

4. Porque grupos monofiléticos são relevantes em classificações biológicas?

5. Muitas das comparações sobre biodiversidade ou sobre sucesso de um grupo são baseadas no número de táxons de uma determinada categoria presentes numa determinada área ou componentes desse grupo (*e.g.*, “A região X possui mais gêneros que a região Y.”, ou ainda, “Os insetos são mais diversos que os crustáceos, pois possuem um maior número de famílias.”). Com base nos resultados das classificações acima, discuta criticamente o uso desses critérios comparativos.

6. Considere a seguinte matriz de dados, com seis táxons e 10 caracteres:

TÁXON\CARATER	1	2	3	4	5	6	7	8	9	10
A	0	0	0	0	0	0	0	0	0	0
B	1	1	0	0	0	0	0	0	0	0
C	1	1	0	0	1	0	0	0	0	0
D	1	0	1	1	0	1	1	0	0	0
E	1	0	1	2	0	0	0	1	1	1
F	1	0	1	3	2	2	2	2	2	2

i. Faça uma análise cladística dessa matriz. Obs: apenas o caráter 4 deve ser considerado ordenado e polarizado (0→1→2→3) e a raiz da topologia deve situada no ramo do táxon ^a

ii. Faça uma análise fenética, utilizando UPGMA, dessa matriz¹.

¹ O protocolo de cálculo está no final deste documento – caso você tenha esquecido como UPGMA é calculado.

iii. Proponha um esquema de classificação por subordinação para ambos os diagramas, i.e., o cladograma e o fenograma obtidos.

iv. Esses esquemas de classificação, aplicados aos resultados dos dois métodos, podem ser idênticos? Justifique.

v. Qual classificação você escolheria? Por que motivo?

UPGMA (Unweighted Pair-Group Method using arithmetic Averages)

Para a construção do fenograma, vários algoritmos são disponíveis para medir a semelhança ou diferença entre as OTUs. Alguns medem a similaridade (coeficientes de similaridade), e os valores são expressos em porcentagem. Outros medem a distância (coeficientes de distância), que pode ser calculada de diversas maneiras. Por exemplo, a distância Manhattan considera valores absolutos; já a distância euclidiana calcula a média entre as OTUs situadas nos vértices de um triângulo. Esses valores podem ser agrupados de diversas formas. Por exemplo: por ligação média sem peso ou UPGMA (*Unweighted Pair-Group Method using arithmetic Averages*); por ligação média com peso ou WPGMA (*Weighted Pair-Group Method using arithmetic Averages*); por ligação simples (*simple linkage*).

No exemplo que segue, será utilizada ligação por UPGMA e as distâncias serão consideradas em seus valores absolutos (distância Manhattan).

O cálculo desta distância é feita aos pares entre todas as OTUs. Por exemplo, considere duas OTUs com as seguintes sequências nucleotídicas:

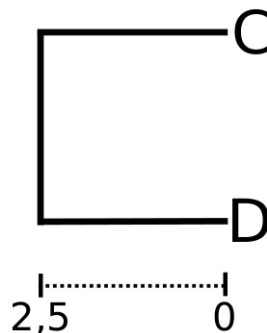
```

A      ACGTACGTACGTACGTACGT
      * * * * * * * * * *
B      CCCTCCCTCCCTCCCTCCCC
  
```

Teríamos que a distância entre A e B é 11 (*s). Com base nesta comparação, construímos então a matriz inicial de distâncias Manhattan e teríamos a matriz abaixo:

	A	B	C	D
A	-			
B	11	-		
C	14	9	-	
D	15	10	5	-

Com base no menor valor encontrado na matriz (**5**) (menor distância = menor diferença), constrói-se o primeiro núcleo, **C/D**. Para a construção do fenograma, divide-se o valor por 2; $5/2 = 2,5$. Esse valor é colocado no eixo de distância, que acompanha o fenograma em elaboração.



A seguir, constrói-se uma nova matriz, considerando-se o núcleo **C/D**.

Para encontrar os valores da matriz, calcula-se a distância do núcleo **C/D** para as outras OTUs. Devem ser considerados os valores existentes na matriz inicial.

$$D_{(C/D), A} = D_{(C, A)} + D_{(D, A)} / 2 = 14 + 15 / 2 = 14,5.$$

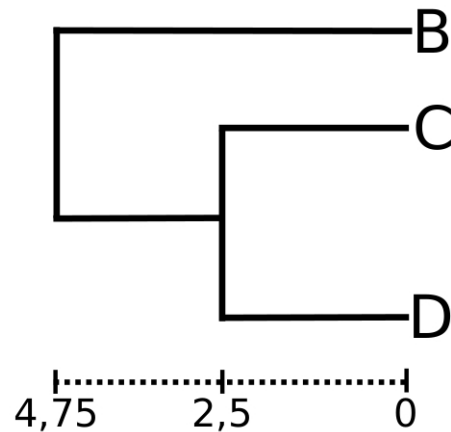
$$D_{(C/D), B} = D_{(C, B)} + D_{(D, B)} / 2 = 9 + 10 / 2 = 9,5.$$

$$D_{(A, B)} = 11 \text{ (observado diretamente na matriz acima).}$$

Nova Matriz de distâncias:

	C/D	A	B
C/D	-		
A	14,5	-	
B	9,5	11	-

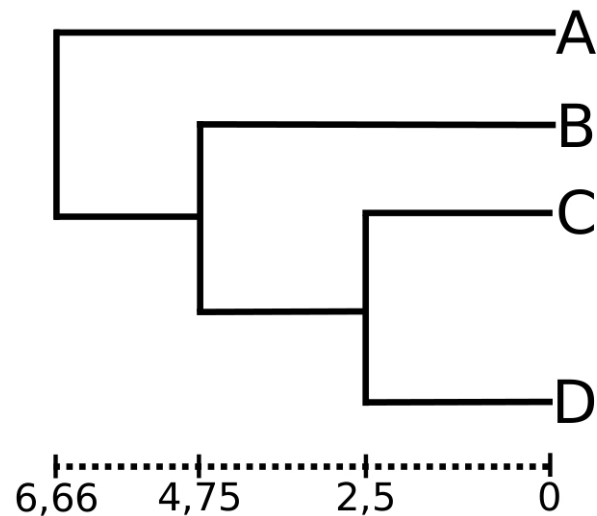
Como o menor valor obtido é **9,5**, B é a nova OTU que vai ser ligada ao núcleo C/D. A distância será igual a $9,5/2 = 4,75$



O núcleo **C/D/B** está formado. Como só resta uma OTU, **A**, basta calcular a distância entre **A** e o núcleo **C/D/B**.

$$D_{(C/D/B), (A)} = D_{(C,A)} + D_{(D,A)} + D_{(B,A)} / 3 = 14 + 15 + 11 / 3 = 40/3 = 13,33 \text{ e } 13,33/2 = 6,66$$

Abaixo está fenograma final, com seu eixo de distâncias.



A distância reflete a similaridade entre duas OTUs. Se possuírem uma distância zero, isso significa que as OTUs são iguais. Quanto maior for a distância que separa duas OTUs, menos similares elas serão.