BIO 103 Biologia evolutiva

Exercício: Comparações de sequências de macromoléculas.
Prof. Dr. Sergio Russo Matioli, Depto. de Genética e Biologia evolutiva, IB-USP

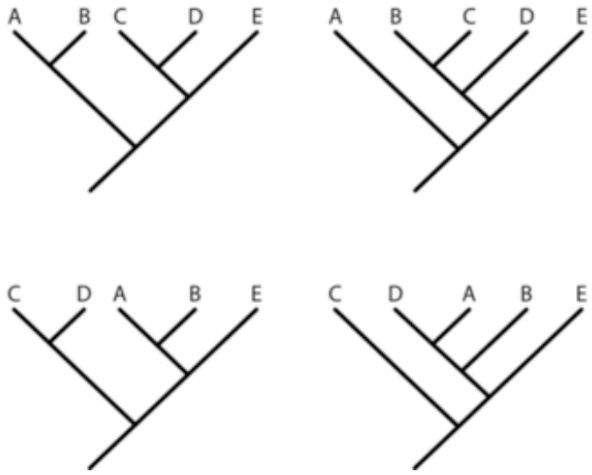
Abaixo, estão representadas as sequências de resíduos de aminoácidos alinhadas da sub-unidade 2 da oxidase do citocromo de algumas espécies de vertebrados

```
                    1        10        20        30        40        50
                    ---------+---------+---------+---------+---------+
Gorilla gorilla     MAHAAQVGLQDATSPIMEELIIFHDHALMIIFLICFLVLYALFLTLTTKL
Pan paniscus        MAHAAQVGLQDATSPIMEELIIFHDHALMIIFLICFLVLYALFLTLTTKL
Latimeria chalumnae MAHPSQLGLQDAASPVMEELLHFHDHALMIVFLISTLVFYIILAMMTTKM
Ambystoma mexicanum MAHPSQLGFQDAASPIMEELLHFHDHALMAVFLISTLVLYIITVMMTTKL
Sibon nebulatus     MPYAAQLTLQEATGPTMEEVIFLHDHVLLLTCLMSLVIMMFAITATTTTL
                    *.:.:*: :*:*:.* ***:: :***.*:   *:. :::       **.:
                    51       60        70        80        90        100
                    ---------+---------+---------+---------+---------+
Gorilla gorilla     TNTN-ISDAQEMETIWTILPAIILVLIALPSLRILYMTDEINDPSFTIKS
Pan paniscus        TNTS-ISDAQEMETVWTILPAIILVLIALPSLRILYMTDEVNDPSFTIKS
Latimeria chalumnae TDKY-ILDAQEIEIVWTLLPAIVLILVALPSLRILYLIDEVENPHLTIKA
Ambystoma mexicanum TNTN-AMDAQEIEMVWTIMPAIILIVIALPSLRILYLMDEINDPHLTVKA
Sibon nebulatus     THNDPTEEVEQLEAAWTAAPIMILILTALPSVRSLYLMEEVFDPYLTIKA
                    *..    :.:::*  **  * ::*:: ****:* **: :*: :* :*:*:
                    101      110       120       130       140       150
                    ---------+---------+---------+---------+---------+
Gorilla gorilla     IGHQWYWTYEYTDYGGLIFNSYMLPPLFLEPGDLRLLDVDNRVVLPVEAP
Pan paniscus        IGHQWYWTYEYTDYGGLIFNSYMLPPLFLEPGDLRLLDVDNRVVLPVEAP
Latimeria chalumnae MGHQWYWSYEYTDYEELSFDSYMTPLQDLNPGQFRLLETDHRMVIPMESL
Ambystoma mexicanum IGHQWYWSYEYTNYDDLVFDSYMLPTQNLNPGEFRLLEVDNRMVVPMESP
Sibon nebulatus     TGHQWYWNYEYSDHTPMSFDSYMTQTKDLQNGSPRLLEVDHRMTMPAGLQ
                     ******.***:::  : *:***     *: *. ***:.*:*:.:*
                    151      160       170       180       190       200
                    ---------+---------+---------+---------+---------+
Gorilla gorilla     VRMMITSQDVLHSWAVPTLGLKTDAIPGRLNQTTFTATRPGVYYGQCSEI
Pan paniscus        VRMMITSQDVLHSWAVPTLGLKTDAIPGRLNQTTFTATRPGVYYGQCSEI
Latimeria chalumnae IRVLISAEDVLHSWAVPALGVKMDAVPGRLNQITFMISRPGLYYGQCSEI
Ambystoma mexicanum IRMLISAEDVLHSWAMPSMGIKTDAIPGRLNQTTFIASRPGVFYGQCSEI
Sibon nebulatus     TRIVVTAEDVLHSWTIPSLGVKVDAVPGRLNQIPLITSYVGVFFGQCSEI
                     *:::::::*******::*::*:* **:****** .:  :  *:::******
                    201      210       220       230
                    ---------+---------+---------+-
Gorilla gorilla     CGANHSFMPIVLELIPLKIFEMGPVFAL---
Pan paniscus        CGANHSFMPIVLELIPLKIFEMGPVFAL---
Latimeria chalumnae CGANHSFMPIVLEAIPLDPFEDWSSSMLEEA
Ambystoma mexicanum CGANHSFMPIVVETTPLGHFQNWSSSMQE--
Sibon nebulatus     CGANHSFMPIAVEAIPLYHFEQWTIKGQ---
                    **********.:*  **  *:  .
```

Os alinhamentos de macromoléculas são representações onde cada letra corresponde a um tipo de monômero (aqui no caso 1 letra para cada um dos 20 aminoácidos) onde está implícita uma hipótese de homologia. Onde há o símbolo (-) em uma sequência significa a ausência de um monômero mas que corresponde a um monômero homólogo em outra sequência, que implica um evento de inserção ou deleção durante a evolução desses organismos. Os asteriscos abaixo do alinhamento indicam que os resíduos são idênticos para as 5 espécies. Os símbolos “ponto”, “dois pontos” e “espaço em branco” indicam graus diferentes de variação na posição.

Questões

A figura abaixo representa algumas relações filogenéticas possíveis entre 5 táxons de tal forma que não haveria ambiguidades (cada vértice interno está conectado a não mais que três arestas)



1) Quantas seriam as relações filogenéticas possíveis entre 5 táxons que satisfaçam as condições acima?

2) Tente estabelecer o parentesco entre essas espécies, baseando-se nas sequências de resíduos de aminoácidos das proteínas fornecidas. Para isso estabeleça uma série de procedimentos que seja geral, ou seja, que possam ser utilizados em qualquer conjunto de dados semelhantes. Isso equivale a propor um algoritmo. O resultado desse procedimento deve ser representado sob a forma de uma filogenia.