

TP 6 - INFERENCIAS EVOLUTIVAS

RETO I: Detalla las tácticas y/o metodologías que deberían utilizarse para darles una respuesta a los padres del niño.

A) Dadas las secuencias de Mosca, humano y Moscahumano ¿Qué criterios se les ocurren para comparar las secuencias? ¿Qué resultados obtienen del análisis anterior?

Lo primero que podríamos hacer es comparar la secuencia de Moscahumano con la secuencia de humano una vez alineadas para ver qué características de nuestra especie conserva aún el niño. Luego podríamos comparar la secuencia de Moscahumano con la de Mosca para ver qué características de la nueva especie (mosca) posee. También podríamos comparar las secuencias de Humano y Mosca para ver qué características compartían desde antes de la "fusión", pero no creemos que haya un nivel de coincidencia muy alto.

B) ¿Qué resultado esperaríamos obtener si utilizara el resto de las secuencias en el análisis? ¿Por qué?

Si además incluyéramos el resto de las secuencias podríamos observar qué similitudes poseen las secuencias de Cyt P450 de esas especies con la secuencia de Cyt P450 del niño moscahumano.

RETO II: Como vimos anteriormente existen algunos softwares optimizados para confeccionar alineamientos de secuencias. En particular hemos trabajado con Clustal (Larkin et al. 2007). Confecciona los alineamientos para los del punto Ia y Ib análisis.

Comparación de las secuencias de Cyt P450 de humano, mosca de fruta y moscahumano

```

CLUSTAL O(1.2.4) multiple sequence alignment

NP_001285984.1 cytochrome_c_proximal_isoform_B [Drosophila_melanogaster]
NP_061820.1 cytochrome_c [Homo_sapiens]
bartmosca
MGVPAGDVEKGGKLFVQRCACQCHTVEAGGKHVGNLHGLIGRKTGOAAGFAYTDANKAK 60
---MGDVEKGGKIFIMKCSQCHTVEKGGKHTGPNLHGLFGRKTGOAPGYSYTAANKNK 56
--MGSFDAENGKKIFVQKCAQCHTYEVGGKHTGPNLHGLFGRKTGOAPGYSYTAANKNK 58
      * * * * * . * * * * * * * * * * * * * * * * * * * * * *
      * * * * * . * * * * * * * * * * * * * * * * * * * * * *

NP_001285984.1 cytochrome_c_proximal_isoform_B [Drosophila_melanogaster]
NP_061820.1 cytochrome_c [Homo_sapiens]
bartmosca
GITWNEDTLFEYLENPKKIYIPGTMIFAGLKKPNERGDLIAYLKSATK- 108
GIIWGEDTLMEYLENPKKIYIPGTMIFVGIIKKKEERADLIAYLKKATNE 105
GIIWGEDTLMEYLENPKKIYIPGTMIFVGIIKKKEERADLIAYLKKATNE 107
      * * * * * . * * * * * * * * * * * * * * * * * * * * * *

```

Comparación de las secuencias de Cyt P450 de humano, mosca de fruta, moscahumano y todas las otras especies del documento.

```

CLUSTAL O(1.2.4) multiple sequence alignment

WP_037127040.1 cytochrome_c_family_protein [Rhizobium_sp. CF097]
NP_001285984.1 cytochrome_c_proximal_isoform_B [Drosophila_melanogaster]
NP_061820.1 cytochrome_c [Homo_sapiens]
bartmosca
CBJ31344.1 Cytochrome_c [Ectocarpus_siliculosus]
KAE8767936.1 Cytochrome_c [Hordeum_vulgare]
QEE59979.1 cytochrome_c [Betula_platyphylla]
KAE9612850.1 Cytochrome_c [Lupinus_albus]
NP_001363126.1 cytochrome_c [Vigna_radiata]
MRNQILMIGLSAAVLCPIGALAQEGDAEAGAAIF-KKCATCHIVESD-TNKVGPLKGLF 58
-----MGVPAGDVEKGGKLFVQRCACQCHTVEAGGKHVGNLHGLI 41
-----MGDVEKGGKIFIMKCSQCHTVEKGGKHTGPNLHGLF 37
-----MGSFDAENGKKIFVQKCAQCHTYEVGGKHTGPNLHGLF 39
-----MSTKAGAKIFKTKCSQCHTVEKDAGHKQGNLNLGF 36
-----MAS-FGEAPAGNAAGGEKIFKTKCAQCHTVERDGAHKQGNLNLGF 45
-----MAS-FDEAPPGNPKVGEKIFKTKCAQCHTVEKGAGHKQGNLNLGF 45
-----MAS-FDQAPPGDQKVGKIFKTKCAQCHTVDKGAGHKQGNLNLGF 45
-----MAS-FDEAPPGNKSGEKIFKTKCAQCHTVDKGAGHKQGNLNLGF 45
      . * * * * * * * * * * * * * * * * * * * * * *

WP_037127040.1 cytochrome_c_family_protein [Rhizobium_sp. CF097]
NP_001285984.1 cytochrome_c_proximal_isoform_B [Drosophila_melanogaster]
NP_061820.1 cytochrome_c [Homo_sapiens]
bartmosca
CBJ31344.1 Cytochrome_c [Ectocarpus_siliculosus]
KAE8767936.1 Cytochrome_c [Hordeum_vulgare]
QEE59979.1 cytochrome_c [Betula_platyphylla]
KAE9612850.1 Cytochrome_c [Lupinus_albus]
NP_001363126.1 cytochrome_c [Vigna_radiata]
GRKAGTHPDFSYSSAMKAAGEGLVWDETTLRDYLHNPKAKVKGTMAFVGKVKDDDEITN 118
GRKTGOAAGFAYTDANKA--KGITWNEDTLFEYLENPKKIYIPGTMIFAGLKKPNERGD 98
GRKTGOAPGYSYTAANKN--KGIIWGEDTLMEYLENPKKIYIPGTMIFVGIIKKKEERAD 94
GRKTGOAPGYSYTAANKN--KGIIWGEDTLMEYLENPKKIYIPGTMIFVGIIKKKEERAD 96
GRTAGTVDGYSYSTANKT--SGVLWDESLFDYLLAPKKYIKGTMVFAIGKKPAERKE 93
GRQSGTTAGYSYSTANKN--MAVNWEEKTLDYLLNPKKIYIPGTMVFPGLKKPQDRVD 102
GRQSGTTAGYSYSTANKN--MAVNWEEKTLDYLLNPKKIYIPGTMVFPGLKKPQDRAD 102
GRQSGTTAGYSYSTANKN--MAVIWEEKTLDYLLNPKKIYIPGTMVFPGLKKPQDRAD 102
      * * * * * . * * * * * * * * * * * * * * * * * * * * * *

WP_037127040.1 cytochrome_c_family_protein [Rhizobium_sp. CF097]
NP_001285984.1 cytochrome_c_proximal_isoform_B [Drosophila_melanogaster]
NP_061820.1 cytochrome_c [Homo_sapiens]
bartmosca
CBJ31344.1 Cytochrome_c [Ectocarpus_siliculosus]
KAE8767936.1 Cytochrome_c [Hordeum_vulgare]
QEE59979.1 cytochrome_c [Betula_platyphylla]
KAE9612850.1 Cytochrome_c [Lupinus_albus]
NP_001363126.1 cytochrome_c [Vigna_radiata]
LIAYLKQFP-- 127
LIAYLKSATK- 108
LIAYLKKATNE 105
LIAYLKKATNE 107
LIAYLKEATP 104
LIAYLKQSTA- 112
LISYLKESTAS 113
LIAYLKESTAQ 113
LIAYLKESTA- 112
      * * * * *

```

RETO III: Mediante el uso del servidor de IQtree (Trifinopoulos et al. 2016), confecciona los árboles filogenéticos para los alineamientos obtenidos en el punto II.

A) Como vemos, el servidor nos permite elegir el modelo de sustitución ¿A qué se refiere?

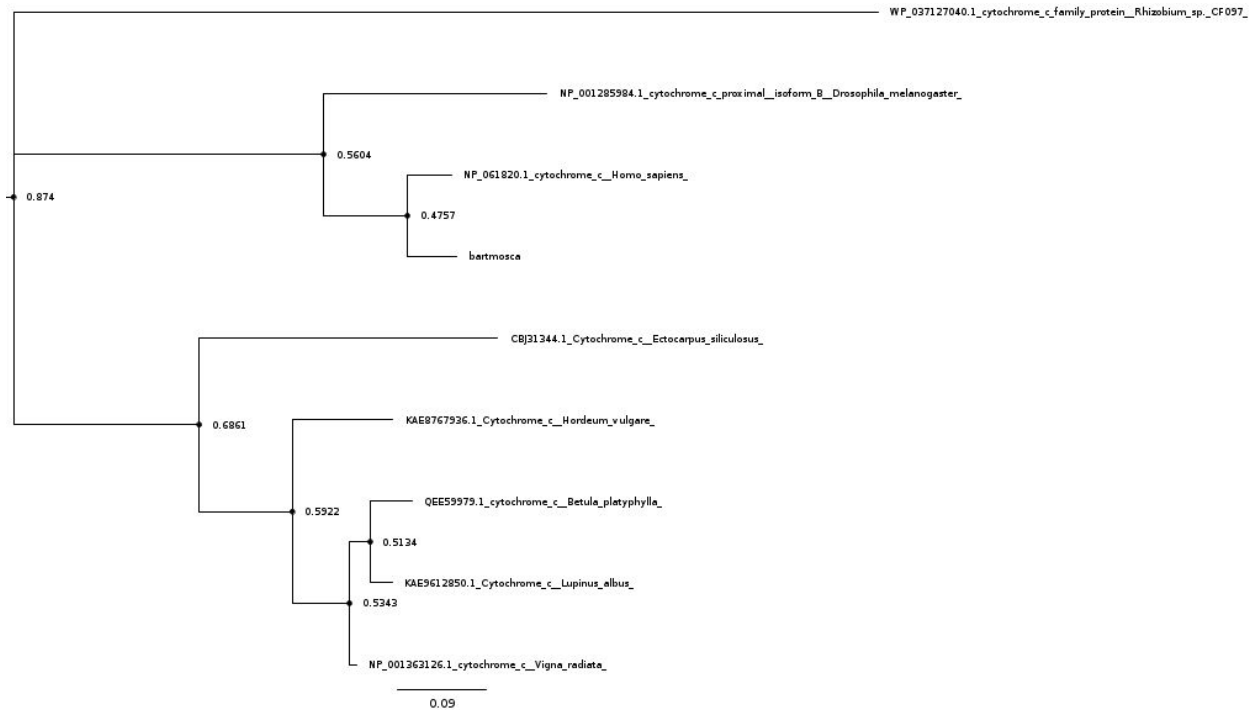
Según Wikipedia, un modelo de sustitución describe el proceso por el cual una secuencia cambia a lo largo del tiempo. Estos modelos se utilizan para calcular la similitud de secuencias a lo largo del árbol filogenético y distancias evolutivas, es decir, la cantidad de sustituciones ocurridas desde que dos secuencias se separaron a partir de un ancestro en común.

B) ¿Qué es el Bootstrap? ¿De qué manera nos habla de la calidad de nuestro árbol? ¿Cómo influye el número de Bootstraps en el resultado?

Es un proceso estadístico que genera muchas muestras simuladas a partir de un único conjunto de datos. Cuando hablamos de Bootstrap nos referimos a un valor que representa cuántas veces fue observada la misma rama a la hora de reconstruir un árbol filogenético a partir de la información dada. Se utiliza como medio para estimar los niveles de confianza de los clados dentro de un árbol filogenético. A mayor valor de Bootstrap, más certeza tenemos sobre la estimación de los clados a través del árbol.

C) Interpreten los resultados obtenidos, mediante la visualización de los árboles con la herramienta FigTree. ¿Es necesario realizar algún paso extra, previo a la interpretación del árbol? ¿Por qué?

Este es el resultado que obtuvimos al subir el archivo con la comparación de secuencias:



Es necesario rootear el árbol previamente. Si observamos la imagen, vemos que desde el nodo de más a la izquierda salen 3 ramas. Esto nos indica que nuestro árbol aún no está rooteado. Para poder rootearlo debemos indicar dónde está la raíz. Es decir, dónde está el nodo principal a partir del cual vamos a empezar a bifurcar. Para ello es necesario seleccionar un organismo que esté bastante alejado en el árbol de la sección que queremos analizar.



Una vez rooteado el árbol nos quedaremos con la sección del nodo marcado en rosa, que es el ancestro en común



Podemos notar que la secuencia analizada de bartmosca tiene mayor similitud con la de humano que con la de mosca. Así que podríamos bien decirle a los padres del niño que aún sigue siendo humano :D