

TP 6: Inferencias evolutivas

Alumnos: Cuoco Carlos, Markon Mariano, Verdecanna Mariano

Aclaración: Los archivos que utilizamos se pueden encontrar en la carpeta Archivos TP 6

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?

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

a) Lo primero sería alinear las secuencias del citocromo de un humano, de una mosca y de moscahumano, Una vez alineadas se debería hacer el árbol filogenético, y de acuerdo a las diferencias y similitudes encontradas, ahí se vería claramente que tan distante está del humano y de la mosca.

b) Se obtendría un resultado más realista o acertado, ya que como los métodos el análisis filogenético con los que se forman los árboles, tienen en cuenta las posiciones del alineamiento, mientras más secuencias se tengan alineadas, habrá mayor información de cada posición.

En el árbol filogenético, ya no se encontrarían en lugares opuestos el humano y la mosca, habría algún punto en común, situando al moscahumano como un ancestro común de ambos

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.

Alineamiento de reto 1 inciso a:

[illegible]

Alineamiento de reto 1 inciso b:

Alignments	Result Summary	Guide Tree	Phylogenetic Tree	Results Viewers	Submission Details
Download Alignment File	Hide Colors				
CLUSTAL O(1.2.4) multiple sequence alignment					
WP_037127040.1 cytochrome_c family protein [Rhizobium sp. CF097]	MRNQILMIGLSAAVLCPIGALAQEGDAEAGAAIF-KKCATCHIVESD-TNKVGP	SLKGLF	58		
NP_001285984.1 cytochrome_c proximal, isoform_B [Drosophila melanogaster]	-----MGVPAGDVEKGKKLFVQRCACQCHTVEAGGKHVGP	NLHGLI	41		
NP_061820.1 cytochrome_c [Homo sapiens]	-----MGDVEKGKKIFIMKCSQCHTVEKGGKHKTGP	NLHGLF	37		
bartmosca	-----MGSQDAENGKKIFVQKCAQCHTYEVGGKHKTGP	NLHGLF	39		
CBJ31344.1 Cytochrome_c [Ectocarpus siliculosus]	-----MSTKAGAKIFKTKCSQCHTVEKDAGHKQGP	NLNGLF	36		
KAE8767936.1 Cytochrome_c [Hordeum vulgare]	-----MAS-FGEAPAGNAAGGEKIFKTKCAQCHTVERDGAHKQGP	NLSGLF	45		
QEE59979.1 cytochrome_c [Betula platyphylla]	-----MAS-FDEAPPGNPKVGEKIFKTKCAQCHTVEKGAGHKQGP	NLNGLF	45		
KAE9612850.1 Cytochrome_c [Lupinus albus]	-----MAS-FDQAPPDGAKEKIFKTKCAQCHTVDKGAGHKQGP	NLNGLF	45		
NP_001363126.1 cytochrome_c [Vigna radiata]	-----MAS-FDEAPPGNKSGEKIFKTKCAQCHTVDKGAGHKQGP	NLNGLF	45		
		* : * : * : * : * : * : * : * : *			
WP_037127040.1 cytochrome_c family protein [Rhizobium sp. CF097]	GRKAGTHPDFSYSSAMKAAGEGLVWDETTLRDYLHNPKAKVKGTKMAFVGVKKDDEITN		118		
NP_001285984.1 cytochrome_c proximal, isoform_B [Drosophila melanogaster]	-----KGITWNEDTLFEYLENPKKIYIPGTMIFAGLKKPNERGD		98		
NP_061820.1 cytochrome_c [Homo sapiens]	GRKTGQAPGYSYTAANKN-----KGIWGEDTLMEYLENPKKIYIPGTMIFVGIKKKEERAD		94		
bartmosca	GRKTGQAPGYSYTAANKN-----KGIWGEDTLMEYLENPKKIYIPGTMIFVGIKKKEERAD		96		
CBJ31344.1 Cytochrome_c [Ectocarpus siliculosus]	GRTAGTVDGYSYSTANKT-----SGVLWDESSLFDYLLAPKKYIKGTMVFAIKKPAERKE		93		
KAE8767936.1 Cytochrome_c [Hordeum vulgare]	GRQSGTTAGYSYSTANKN-----MAVWVEENTLYDYLLNPKKIYIPGTMVFPGLKKPQERVD		102		
QEE59979.1 cytochrome_c [Betula platyphylla]	GRQSGTTAGYSYSTANKN-----MAVNWEEKTLDYLLNPKKIYIPGTMVFPGLKKPQDRAD		102		
KAE9612850.1 Cytochrome_c [Lupinus albus]	GRQSGTTAGYSYSTANKN-----MAVNWEEKTLDYLLNPKKIYIPGTMVFPGLKKPQDRAD		102		
NP_001363126.1 cytochrome_c [Vigna radiata]	GRQSGTTAGYSYSTANKN-----MAVINWEEKTLDYLLNPKKIYIPGTMVFPGLKKPQDRAD		102		
	** : * : * : * : * : * : * : * : *				
WP_037127040.1 cytochrome_c family protein [Rhizobium sp. CF097]	LIAYLKQFP--	127			
NP_001285984.1 cytochrome_c proximal, isoform_B [Drosophila melanogaster]	LIAYLKSATK-	108			
NP_061820.1 cytochrome_c [Homo sapiens]	LIAYLKKATNE	105			
bartmosca	LIAYLKKATNE	107			
CBJ31344.1 Cytochrome_c [Ectocarpus siliculosus]	LIAYLKEATAP	104			
KAE8767936.1 Cytochrome_c [Hordeum vulgare]	LIAYLKQSTA-	112			
QEE59979.1 cytochrome_c [Betula platyphylla]	LISYLKESTAS	113			
KAE9612850.1 Cytochrome_c [Lupinus albus]	LIAYLKESTAQ	113			
NP_001363126.1 cytochrome_c [Vigna radiata]	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?

En la herramienta nos permite elegir distintos tipos de substitution models o matrices de sustitución categorizados en ADN, proteína, codón, binario y morfológicos. Dependiendo del tipo de muestra que tenemos y en base qué análisis se quiere hacer, se selecciona alguno. Tener en cuenta que cada tipo a su vez tiene distintos modelos para aplicar. También nos da la opción de fijarse automáticamente cuál sería la mejor opción, para esto utiliza ModelFinder

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?

El bootstrap es un algoritmo que se basa en en crear réplicas de los alineamientos en base al original, haciendo modificaciones en cada réplica. El número final de posiciones se mantiene y generamos un árbol en cada réplica. El paso final será evaluar para cada nodo el porcentaje de árboles en los que aparece. Los nodos con un alto valor

de bootstrap tienen, si se cumplen las asunciones del método utilizado, una probabilidad alta de ser correctos. Cuanto mayor sea el número de bootstraps más certezas o confiabilidad tendremos en nuestro resultado.

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?

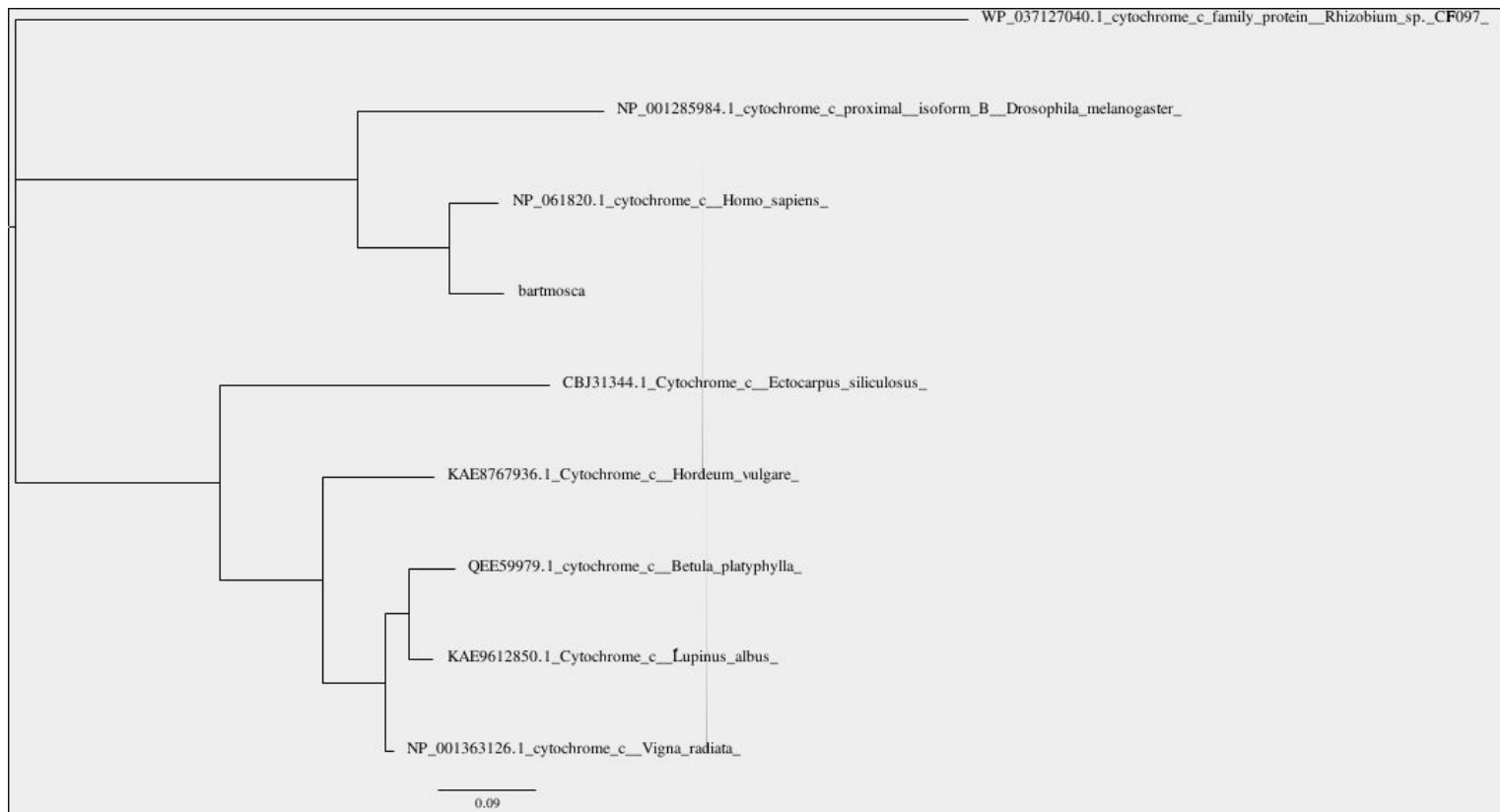
Para poder visualizar en el árbol, tuvimos que procesar el archivo alineado en clustal , lo evaluamos en una instancia local de IQTREE

iqtree -s tp6reto3resultadoalineamiento -m TEST -bb 1000 -alrt 1000

Al ejecutar nos generó varios archivos, para cargar en el figtree utilizamos el archivo **tp6reto3resultadoalineamiento.treefile**.

Luego al abrir el programa buscamos la secuencia más alejada/rara y utilizamos el comando *reroot* para generar un árbol ordenado para analizar. Además también habilitamos el label de los nodos para saber el porcentaje de confianza de los mismos y poder hacer un análisis más certero.

Arbol Sin Rutear



Arbol Ruteado

