**Entrega TP 5 - Bioinformática:**

* **RETO I y II:**

**a)** A través de NCBI se obtuvo la secuencia del ser humano, de la mosca y de la “Bartmosca”:

>NP\_061820.1 cytochrome c [Homo sapiens]

MGDVEKGKKIFIMKCSQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGYSYTAANKNKGIIWGEDTLMEYLE

NPKKYIPGTKMIFVGIKKKEERADLIAYLKKATNE

>AFP62086.1 Cytochrome c [Musca domestica]

MGVPAGDVEKGKKLFVQRCAQCHTVEAGGKHKVGPNLHGLFGRKTGQAAGFAYTDANKSKGITWNEDTLF

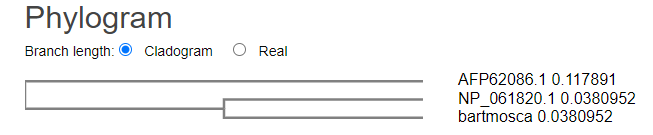
EYLENPKKYIPGTKMIFAGLKKPGDRADLIAYLKSATA

>bartmosca

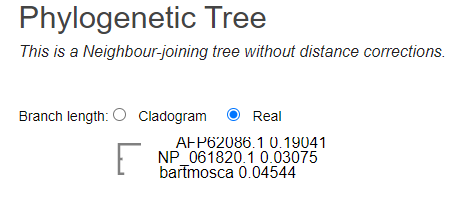
MGSGDAENGKKIFVQKCAQCHTYEVGGKHKTGPNLHGLFGRKTGQAPGYSYTAANKN

KGIIWGEDTLMEYLENPKKYIPGTKMIFVGIKKKEERADLIAYLKKATNE

Luego, se compararon las 3 secuencias a través de Clustal Omega, obteniendo como resultado que la Bartmosca posee mas similitudes para con el Humano que para con la Mosca (los papás del chico se pueden quedar tranquilos…):



El árbol filogenético también muestra más similitud con el Humano (si bien no sirve mucho esta comparación porque solo hay tres muestras):



**b)** Para el caso en el que se usaran todas las secuencias que se utilizaron en el TP anterior (Gallo, Chimpancé, Caballo, Perro, Rana, etc..), la BartMosca se alinearía más hacia la secuencia del Humano y mostraría como ancestro a la Mosca y al Humano.

* **RETO III:**

**a)** Los modelos de sustitución son aquellos que nos permiten calcular las distancias entre las distintas secuencias que se están comparando. Hay modelos simples y complejos, y su precisión está relacionada a su complejidad. Para elegir el modelo de sustitución óptimo para realizar una comparación hay que tener en cuenta muchas variables, como cantidad de datos a evaluar, formato de la información, etc.

Dicho esto, la herramienta IQtree nos permite elegir el modelo de sustitución si sabemos cuál es el más óptimo para el Input que estamos realizando, y una opción “auto” si no lo sabemos; de esta forma delegamos esta responsabilidad a la herramienta.

**b)** El Bootstrap es un método de validación de árboles. Básicamente lo que realiza este método es crear réplicas de distintos alineamientos a partir de un alineamiento original, eliminando cierto número de posiciones al azar en cada réplica. El número final de posiciones se mantiene constante, añadiendo duplicaciones de los sitios que han permanecido. Luego, para cada réplica se aplica el método de reconstrucción filogenética y se genera un árbol. El paso final es evaluar para cada nodo el porcentaje de árboles en los que aparece: los nodos con un alto valor de Bootstrap tienen una probabilidad alta de ser correctos (se estiman en >= 90%), mientras los que tienen un bajo valor de Bootstrap podrían haberse generado por azar (se estiman en <= 50%).

Cuantas más réplicas se generen, más probabilidad tendremos de quedarnos con los nodos correctos. Ésta posibilidad permite IQtree, ya que permite seleccionar la cantidad de alineamientos de Bootstraps.

**c)** Antes de interpretar un árbol debiera realizarse una validación previa (como es el caso de Bootstrap o algún otro método). De esta manera, nos quedaremos con los Clados que realmente comparten ancestros.