

Análisis filogenético de la diversidad biológica

II. Objetivo: reconstruir las relaciones filogenéticas entre la especie humana y sus parientes más cercanos utilizando una secuencia genética mitocondrial.

A partir de una base de datos con la secuencia de nucleótidos de los 500 primeros sitios del gen del citocromo b (una proteína mitocondrial) del ADN mitocondrial de varias especies de Primates inferir la historia evolutiva de las especies consideradas (ver figura 1). Estas secuencias están en el archivo *cytb_primates.meg* que puede descargarse de: proyectodepopularización.ccccc

Suborden	Familia	Género	Nombre común	Distribución
Strepsirrhini	Lemuridae	<i>Lemur</i>	lemur	Madagascar
		<i>Microcebus</i>	lemur	Madagascar
Platyrrhini	Cebidae	<i>Cebus</i>	mono capuchino	Neotrópico
		<i>Saimiri</i>	mono ardilla	
Catarrhini	Cercopithecidae	<i>Papio</i>	babuino	Africa
		<i>Macaca</i>	macaco	Africa
	Hylobatidae	<i>Hylobates</i>	gibón	Asia, Indonesia...
	Hominidae	<i>Pongo</i>	orangután	Borneo
		<i>Gorilla</i>	gorila	Africa
		<i>Pan</i>	chimpancé	Africa
<i>Homo</i>		humano	cosmopolita	

Figura 1.- Una clasificación actual de los primates incluidos en la base de datos. Se incluye además la distribución geográfica de cada taxa.

Para esta tarea se utilizará el programa MEGA que deberá descargarse instalarse. MEGA es un programa que permite alinear secuencias, inferir árboles filogenéticos, estimar distancias entre taxa y tasas de evolución, así como testar hipótesis evolutivas. Puede descargarse gratuitamente su última versión de la siguiente dirección: <http://www.megasoftware.net/>.

Para comenzar la tarea se deberá abrir la base de datos con el programa MEGA (por detalles ver *abrir_basededatos.mp4*).

Al finalizar podrá verse el siguiente una ventana que permite explorar los datos con que se trabajará (ver figura 2)

El programa tiene diversas herramientas. Explorando la base podemos resaltar distintos tipos de caracteres: conservados (clickeando en el icono con la letra "C") o variables (icono con "V"), informativos (clickeando "P") como no informativos (ver explorar_basededatos.mp4).

Figura 2.- Explorador de los datos. Cada fila representa la secuencia génica de un individuo distinto y en la primer columna está indicado a qué especie pertenece ese ejemplar. Cada una de las siguientes columnas indica un sitio en la secuencia génica. El encabezado de la columna indica la base más frecuente, y en cada casillero se indica la base presente en cada individuo: con un punto si es la misma que en el encabezado o indicando la base presente en caso de cambio.

A partir de la base de datos se puede hacer una reconstrucción filogenética utilizando distintos métodos, entre ellos máxima parsimonia (ver arbol_filogenetico.mp4).

Además con este programa podemos calcular las distancias entre pares de especies en distintos tipos de secuencias, absolutas o relativas, incluso utilizando distintos modelos (ver distancias.mp4 para el cálculo de distancias absolutas).

- 1.- ¿Cuántos individuos, especies y caracteres contiene la base de datos?
- 2.- ¿Cuántos caracteres son informativos? ¿Por qué algunos de los sitios variables no son informativos?
- 3.- A partir de una reconstrucción filogenética utilizando el criterio de *Máxima Parsimonia* (usar las opciones que vienen por defecto y definir como grupo externo a Lemuridae): ¿cuál sería el

género más emparentado con el nuestro? ¿Es coherente la filogenia obtenida con la clasificación taxonómica de estos géneros?

4.- Utilizando el criterio de parsimonia, ¿qué origen geográfico habría tenido el género *Homo*?

5.- A partir de una tabla de distancias absolutas pareadas y suponiendo que la tasa de divergencia para el citocromo b es de 2% cada millón de años, datar el momento de divergencia entre el linaje humano y el de los chimpancés.