

PRÁCTICA 3: Alineamientos con mafft, con un caso práctico

Elvira Mayordomo

7/Marzo/2018

1 Introducción

En esta tercera sesión se utilizarán los conocimientos adquiridos en un caso real: estudio del DNA mitocondrial. La duración aproximada de esta sesión será de 2 horas.

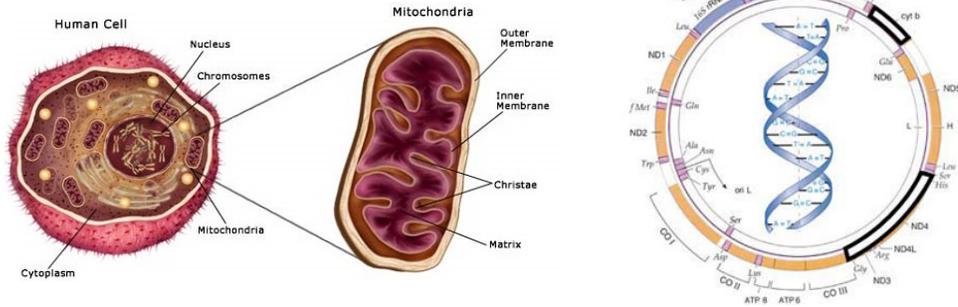
Al principio de la sesión dos estudiantes harán una introducción a las herramientas mafft y WebLogo.

2 Caso de estudio: el DNA mitocondrial

2.1 Introducción

Las mitocondrias son unos orgánulos que existen en la mayoría de las células eucariotas, cuya función es proveerla de energía. Así como en la mayoría de casos la evolución se ve reflejada en mutaciones en el DNA nuclear, las mitocondrias poseen un DNA propio, el mtDNA. La teoría más plausible sostiene que este orgánulo existía fuera de la célula hace millones de años y que, en algún momento, una célula eucariota la absorbió creándose un proceso simbiótico entre ambas.

El mtDNA es una molécula circular de doble hebra compuesta entre 15,000 y 17,000 nucleótidos (en mamíferos), y tiene una serie de propiedades que lo hacen muy útil de cara al estudio de los procesos evolutivos: **i)** su tasa de mutaciones es casi 10 veces superior al DNA nuclear; **ii)** su herencia es exclusivamente por vía materna; **iii)** tiene una asociación directa con muchas



<http://vignette2.wikia.nocookie.net/dragonflyissuesinevolution13/images/4/45/MitochondrialDNA.jpg/revision/latest?cb=20131110214142>

<http://www.mayo.edu/research/-/media/kcms/gbs/research/images/2013/03/07/12/48/lab-ross-project-4.jpg>

enfermedades, la mayoría mortales. Además, entre dos seres humanos cualesquiera, la media de diferencias en su mtDNA es de sólo 40 nucleótidos.

2.2 Descubriendo una nueva tribu indígena

Aunque puede resultar curioso, hoy en día siguen existiendo ciertas tribus que permanecen *ocultas* al mundo moderno. Este hecho suele ser el resultado de la combinación del territorio de su asentamiento (lugares bastante inaccesibles) y de su elevada agresividad como respuesta al más mínimo intento de establecer contacto con ellos. Sin embargo, alguna tribu puede que finalmente acceda a que los extranjeros entren en su territorio, dando comienzo a varios estudios científicos que se realizarán a lo largo de los siguientes meses. Uno de ellos será, cómo no, la secuenciación y análisis del nuevo mtDNA.

Secuenciación

Como ya se ha comentado anteriormente, el proceso de secuenciación incluye varias fases para poder *digitalizar* una secuencia biológica. Las tres primeras partes requieren de equipo especializado del cual no disponemos, pero en el fichero *seq42_reads.fasta* se encuentran los fragmentos ya procesados de una de las nuevas secuencias que hemos extraído de una muestra de sangre de un miembro de la tribu. La primera tarea del alumno será encontrar en GenBank la secuencia de referencia del mtDNA humano



http://farm6.static.flickr.com/5133/5437858371_1a791b9566.jpg

(también conocida como rCRS), descargarla en formato FASTA y alinear todos los fragmentos con la secuencia de referencia utilizando la herramienta MAFFT (<http://mafft.cbrc.jp/alignment/software/>). *Nota: se aconseja al alumno consultar el manual de MAFFT por si es necesario usar algún comando especial.* En concreto se trata de alinear una secuencia de referencia con **fragmentos** de otra parecida.

Una vez hemos obtenido el alineamiento correspondiente, es recomendable realizar una revisión inicial con MSViewer (ver práctica anterior) para confirmar que todo está correcto y el proceso ha finalizado satisfactoriamente.

La última parte de este apartado consistirá en extraer la secuencia digitalizada del miembro de la tribu. Para ello, vamos a utilizar una herramienta web que genera el logo de una secuencia, que, entre otras cosas, refleja la secuencia más común de un alineamiento de forma visual. Para ello nos conectaremos a la web <http://weblogo.berkeley.edu/logo.cgi> y subiremos de nuevo el fichero con nuestro alineamiento. La siguiente tarea del alumno será familiarizarse con esta herramienta y generar un logo útil con la nueva secuencia. *Nota: es posible que la secuencia de referencia genere “ruido” a la hora de extraer la secuencia deseada.*

Agradecimiento

Jorge Álvarez Jarreta colaboró en la primera versión de esta práctica.

2.3 Entrega

Debe entregarse una breve memoria (formato preferido PDF) sobre esta práctica incluyendo los ficheros FASTA necesarios y las visualizaciones y logos de los alineamientos obtenidos (todo ello incluido dentro de un zip). Debe enviarse por correo electrónico a elvira@unizar.es antes de la siguiente sesión (14 de marzo).