

ANÁLISIS FILOGENÉTICO EJERCICIOS

Ejercicio 1

Análisis filogenético con MEGA.

Vamos a realizar el análisis filogenético de las siguientes 13 globinas:

- Mioglobinas de mamíferos: canguro, delfín (marsopa) y foca
- Alfaglobinas de mamíferos: caballo, canguro y perro
- Betaglobinas de mamíferos: perro, conejo y canguro
- Globinas de peces: lamprea de río y de mar
- Globinas de planta e insecto: larva de mosquito y soja

Intentad encontrar estas secuencias por vosotros mismos. Comparadlas con las secuencias que yo encontré*. ¿Se corresponden? ¿Hay replicación de ids o secuencias?

* Mioglobinas de mamíferos
P02194 (canguro), P68278 (delfín) y P68081 (ballena)
Alfaglobinas de mamíferos
P01958 (caballo), P01975 (canguro), P60529 (perro)
Betaglobinas de mamíferos
XP_537902 (perro), NP_001075729 (conejo), P02106 (canguro)
Globinas de peces
690951A (lamprea de río) y P02208 (lamprea de mar)
Globinas de planta e insecto
P02229 (larva de mosquito) y Z29376 (soja)
El fichero con todas estas secuencias está disponible (por compilación) en <http://vts.usal.es/rodrigo/documentos/bioinformatica/filogenia/13globinas.fasta>, pero es recomendable que intentéis construir el vuestro propio para coger soltura en el manejo de las bases de datos y formatos.

Ahora descargamos e instalamos el software MEGA de <http://www.megasoftware.net/>

Podemos realizar el alineamiento múltiple de las globinas mediante una de las herramientas online vistas (clustal, tcoffee, muscle, etc.), o usar las implementaciones de clustalw o muscle disponibles con MEGA:

- Para realizar MSA con MEGA debemos usar:
 - Align → Edit/Build Alignment → Retrieve sequences from a file
 - Y cargar un fichero fasta con las secuencias y realizar el alineamiento
- Para cargar un fichero con el alineamiento múltiple realizado con otra herramienta, debemos usar:
 - “Data → Open a File/Session” abre un editor de texto con el fichero, de modo que en “Utilities → Convert to MEGA format” podemos transformarlo a un fichero que MEGA entienda, y abrirlo luego.
 - NOTA: puede ser que la conversión no sea perfecta, en cuyo caso nos saldrá un mensaje de error y una ventana para editar el fichero y corregir las imperfecciones que hayan surgido.

Relativo a las herramientas:

¿Es fácil de instalar MEGA?

¿Y de cargar los ficheros?

¿Has tenido problemas de formato?

Relativo al alineamiento:

¿Se observan diferencias propias de las especies o de los tipos de globina?

¿Cuáles?

Construid distintos árboles para el alineamiento, probando con distintos modelos de sustitución, distancias y el algoritmos (de distancia –UPGMA y NJ-, de máxima parsimonia y máxima similitud)

¿Hay diferencias entre los árboles? ¿Son las estructuras de los árboles coherentes con lo observado en el MSA?

Evaludad los árboles con bootstrapping: los números en las ramas son los valores de bootstrap. ¿Qué significan estos números? ¿El árbol de consenso difiere del original?

Ejercicio 2 (complejo)

Con MEGA hemos podido probar los métodos de distancia (UPGM y NJ), los de máxima parsimonia y máxima similitud. Nos queda por probar la inferencia bayesiana, en las mismas 13 globinas del ejercicio anterior.

El programa para ello es MrBayes. Descárgalo e instálalo, es muy fácil si tienes Windows (<http://mrbayes.csit.fsu.edu/download.php>).

Una vez instalado, MrBayes es un programa de la vieja escuela. Sólo funciona mediante línea de comandos, y tiene muchas restricciones al formato: sólo podemos introducir alineamientos en formato Nexus (.nex). Seguramente tendrás muchos problemas para dejar el formato bien, de entrada y luego para interpretar el árbol de salida. Los problemas de formato son, de manera reconocida, una de las tareas más engorrosas y que hace perder más tiempo en cualquier análisis o programación de nuevas herramientas bioinformáticas, de ahí la importancia de estándares.

Como decimos, MrBayes sólo permite alineamientos en formato nexus, pero la mayoría de los programas de alineamiento retornan formatos clustal o fasta. Existe la herramienta web ReadSeq, del BCM (búscala con el navegador) que convierte fasta a muchos formatos, entre ellos a nexus. Así:

- Haz el alineamiento múltiple con la herramienta que quieras. MAFFT del EBI, por ejemplo, permite obtener la salida en formato fasta. Si no, igual tienes que convertir el formato clustal a fasta a mano, o buscar alguna herramienta para ello en la red
- Convierte el alineamiento fasta a nexus con ReadSeq (digamos a un fichero aln.nex) y guárdalo directamente en la carpeta de MrBayes
- Abre MrBayes, y teclea `export aln.nex`

Una vez cargado el alineamiento, podemos utilizar distintos modelos:

- Tecleando `lset` podemos cambiar la estructura del modelo (teclea `help lset` para ver las opciones)
- Tecleando `prset` podemos cambiar las distribuciones de probabilidades a priori de los parámetros del modelo (teclea `help prset` para ver las opciones)
- Tecleando `showmodel` podemos ver el modelo actual
- Tecleando `mcmc` podemos cambiar las opciones del método de Monte Carlo que se utiliza para realizar el ajuste. Este método va probando distintas opciones de árbol, hasta minimizar la desviación estándar
- Tecleando `mcmc` invocamos el proceso de Monte Carlo con cadenas de Markov para realizar la reconstrucción filogénica. Cada 1000 iteraciones muestra la desviación estándar, que va bajando hasta el valor puesto de tope (0.01). Si no se llega a ese valor nunca, verás que sigue haciendo iteraciones teniendo siempre valores de desviación similares (dejan de bajar y se estabilizan). Si te ocurre eso modifica el parámetro correspondiente con `mcmc` para incrementar el valor de tope.
- El resultado es, entre otros, un archivo .t que contiene los árboles en todas las iteraciones, en formato Nexus. Sólo hay que tomarlo, quedarnos

con la última iteración, y visualizarlo con alguna herramienta que acepte Nexus como formato.

- La página wiki de ayuda de MrBayes facilita un poco comprender cómo funciona el programa

Finalmente, el fichero en formato nexus puede visualizarse con un visor de árboles filogenéticos, como TreeView, de fácil instalación. Debemos cambiar la extensión de .t a .nex, y conviene eliminar del fichero los árboles intermedios.

Ejercicio 3

Tests de evolución con MEGA: Determina si el ADN mitocondrial de humano y de chimpancé tienen tasas evolutivas iguales. Para ello, utiliza el test de las tasas relativas de Tajima, a través del programa MEGA.

- Descarga e instala el software MEGA
- Obtén las secuencias de ADN mitocondrial de humano, chimpancé, orangután, bonobo, gorila y gibón:
<http://vis.usal.es/rodrigo/documentos/bioinformatica/filogenia/mitDNAprimates.fasta>
- Para determinar qué especie sirve como grupo externo entre chimpancé y humano, realiza un alineamiento múltiple y construye el árbol filogenético asociado, bien con MEGA o con otros programas.
- Realiza el test de Tajima mediante MEGA

¿Tienen humano y chimpancé la misma tasa de evolución? ¿Qué organismo eliges como grupo externo? ¿Por qué? ¿Cambia el resultado si eliges otro organismo como grupo externo?