

Práctico 9

Evolución en familias multigénicas - “Globinas”

El objetivo de esta actividad consiste en familiarizarse con la identificación de familias multigénicas, mediante el estudio de los genes de Globinas en primates a modo de ejemplo.

Una familia génica –o familia multigénica– es un grupo de loci cromosómicos cuya secuencia de nucleótidos es similar y derivan de una secuencia común ancestral. Puede incluir copias de genes ligeramente diferentes y/o pseudogenes, en uno o varios cromosomas.

Los genes codificantes para las globinas representan un ejemplo clásico de familia multigénica. En general, estas proteínas portan un grupo “heme” y se caracterizan por unirse y transportar oxígeno. Presentan dominios homólogos en varios taxa: a) hemoglobinas tetraméricas de vertebrados, las cuales conforman el principal componente proteico de los eritrocitos; b) flavohemoglobinas en microorganismos; c) hemoglobinas homodiméricas bacterianas; d) leghemoglobinas en plantas, asociadas al metabolismo del nitrógeno en plantas con rizomas; e) hemoglobinas de invertebrados; y f) mioglobinas monoméricas usualmente encontradas en el músculo animal entre otras.

En primates existen varias clases de globinas, generalmente designadas con letras griegas (por ejemplo, α , β , γ , δ , ϵ , ξ y Mioglobina MB). Recientemente se ha identificado una nueva clase, la citoglobina (CYGB). En humanos, el gen que la codifica parece expresarse en todos los tejidos y se localiza en el brazo largo del par cromosómico 17. Se sugiere que su tamaño está conservado en varios mamíferos y es de 190 aminoácidos.

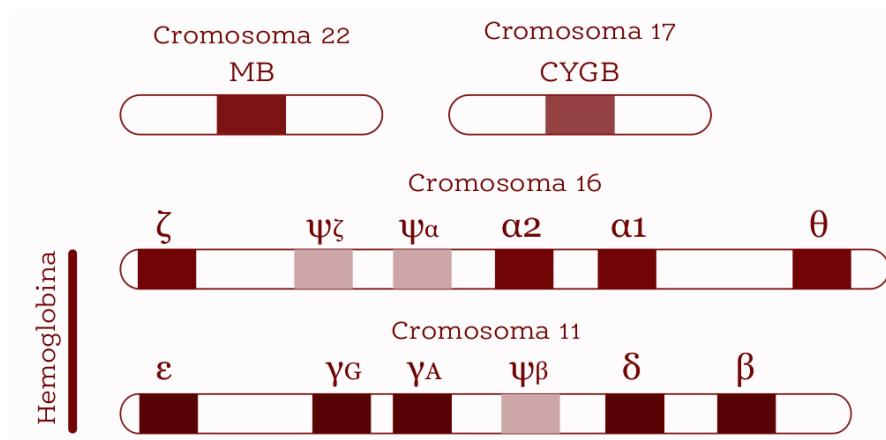


Figura 1 - Localización de los genes codificantes para globinas en humanos.

Conceptos generales

- ¿Qué procesos podrían estar involucrados en la aparición de las familias multigénicas?
- ¿Qué ventajas podrían conferir estas familias génicas?

Alineamiento e identificación de globinas

Análisis del alineamiento de secuencias proteicas de globinas, identificación de regiones conservadas

Cargue el alineamiento en el programa MEGA X: Ir a [Align](#) → [Edit/Build Alignment](#) → [Retrieve a sequence from a file](#) → cargar el archivo “**globinas_aa.fas**” y elegir [Alignment by Muscle](#). Ahora nos detenemos unos segundos a ver la estructura de los alineamientos.

- Identifique los sitios conservados en todas las secuencias y sugiera su posición en la proteína. Justifique.
- ¿Visualiza algún patrón particular en la secuencia de estas proteínas? Interprete.
- Seleccione y copie (ctr + c) la región conservada (al menos 50 aa.) en una de las secuencias y haga un [blastp](#) (blast de sec. aminoacídicas) contra la base de datos en UniProt (<https://www.uniprot.org/blast>). ¿A qué corresponde dicha región?
- En la opción [Structure](#) podemos apreciar la estructura 3D de esta proteína.
- Opcional: Según los colores del modelo, el cual se corresponde con la calidad de la reconstrucción de la proteína –el color depende del pLDDT: “predicted Local Distance Difference Test”, el cual va desde 0 – 100, siendo la calidad de la reconstrucción directamente proporcional al valor–, ¿qué podemos inferir sobre la fiabilidad de la reconstrucción en 3D la proteína?

Cargue los datos en el programa MEGA ([File/Open Data/globinas_nt.meg](#)). Esta base de datos tiene las secuencias codificantes (ADNc) reportadas de genes miembros de la familia de las globinas en los siguientes primates: *Homo sapiens*, *Gorilla gorilla*, *Pongo abelii*, *Macaca mulatta*, *Callithrix jacchus*, *Papio anubis*, *Pan troglodytes*, *Microcebus murinus* y *Otolemur garnettii* (ver Figura 2).

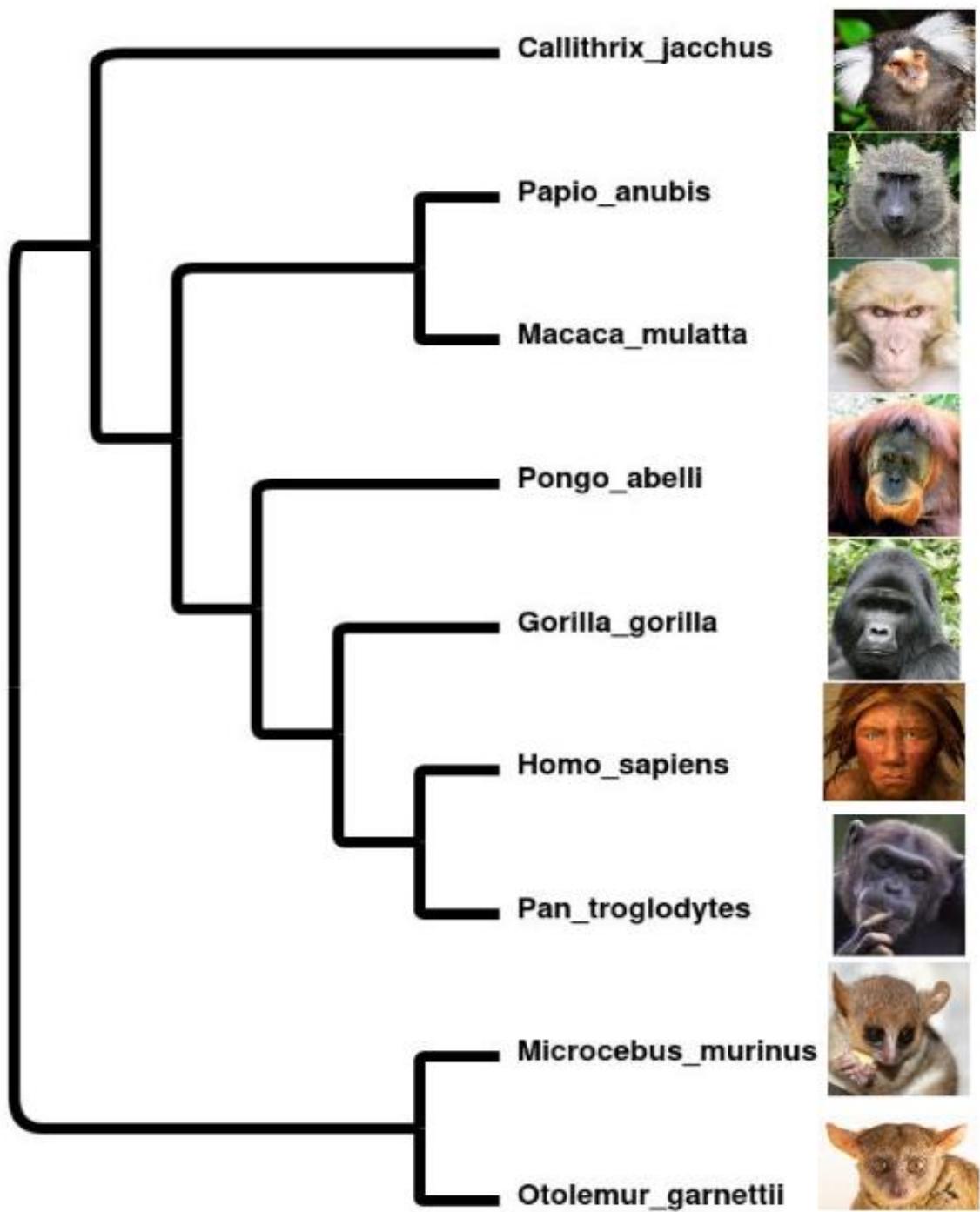


Figura 2 - Cladograma de las especies incluidas en los datos

- En base a los datos disponibles (ver Figuras 1 y 2), ¿cómo piensa usted que ocurrió la evolución de los diferentes tipos de globinas dentro de los primates? Genere un esquema representando las relaciones filogenéticas de estos genes.

Distribución filética de los parálogos

Realice la reconstrucción filogenética en base a los datos que se le brindan (Método: [Neighbor-Joining](#), Model/Method: [p-distance](#), Gaps: [Pairwise deletion](#), Test of Phylogeny: [Bootstrap method](#), No. of Bootstrap Replications: [100](#); dejando las demás opciones por defecto).

- ¿Son los resultados coincidentes con el esquema que usted generó?
- Identifique aquellos nodos que se corresponden con eventos de especiación y/o duplicación.
- Analice e interprete algunos casos particulares:
 - clase delta en humanos
 - ausencia de γ -globina en *M. murinus* y *O. garnettii*
 - ausencia de μ -globina en *M. murinus*
 - relación de las citoglobinas con las otras globinas
- A grandes rasgos ¿a qué se corresponde la distribución filética de los distintos parálogos de globinas?
- ¿En cuál cromosoma buscaría la μ -globina en humanos?
- ¿Qué información adicional sería útil para establecer el origen de un nuevo gen de globinas?

Estimación de las tasas de evolución en los distintos parálogos

Estimación del parámetro ω (dN/dS) entre clases y dentro de las clases

- ¿Cuál espera sea el resultado dentro y entre clases? ¿Qué relación tiene esto con el árbol reconstruido? ¿Cómo piensa que será la estimación de la tasa sinónima?

Como se notó en las actividades anteriores, los genes ya están agrupados en las clases correspondientes en la base de datos. Calcular la distancia media dentro (ortólogos) y entre clases (parálogos). En MEGA, ir a [Distances/Compute within group mean](#) y luego [Compute between group mean](#), incluyendo sustituciones sinónimas (dS) y no-sinónimas (dN), estimadas con el modelo de Kimura 2 parámetros ([Kimura 2-parameters model](#)) y Gaps: [Pairwise deletion](#).

- Abrir el archivo “Cálculo dn-ds familias multigénicas”
- A partir de los valores de distancia media sinónima y no-sinónima calcule la relación entre ellas ($\approx dN/dS$) dentro de cada parálogo (entre ortólogos) y entre parálogos, y tome nota. Todas estas estimaciones ya están realizadas en el archivo mencionado en el punto anterior.
- Interprete ambos resultados, relacione los mismos con procesos evolutivos dados en clase.