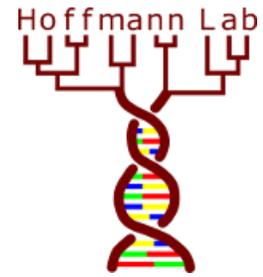




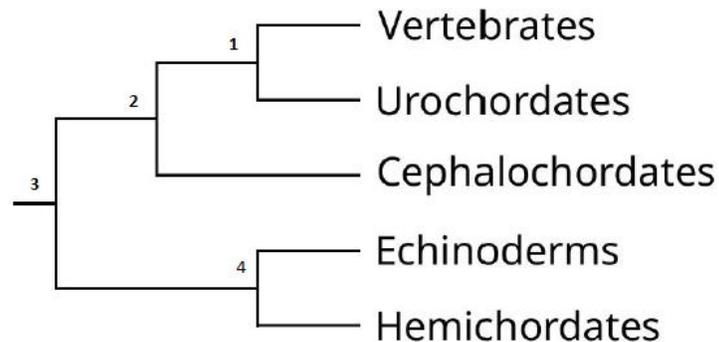
FACULTAD DE  
**CIENCIAS**

UDELAR | [fcien.edu.uy](http://fcien.edu.uy)



### Evolución de familias multigénicas 2020 Ejercicios 27 de abril

Tenemos a la vista la filogenia de deuterostomados como referencia



La finalidad de estos ejercicios es realizar una exploración de algunas herramientas informáticas, tomando como referencia el problema abordado por Hoffmann et al. 2012

#### 1. Explorando genomas en NCBI.

- En el sitio de NCBI, identificar la herramienta Taxonomy para encontrar el genoma (genome, genome project) del gusano bellota (*acorn worm*). Tomar nota de las principales características de la versión disponible del genoma.
- Utilizar la Tabla 1 del artículo para seleccionar un locus putativo (identificado por su *protein record*) como base para una búsqueda con la herramienta Blast. Notas: centrarse en *protein Blast* (notar variantes) y restringir la búsqueda al genoma del gusano bellota o a los hemicordados, usando las opciones de Taxon ID de Blast. Distribuir entre los miembros del grupo varios de los registros de proteínas.
- Examinar y comparar los resultados: ¿Cuántos genes o regiones se recuperan en cada caso? Usando los resúmenes gráficos y los detalles de los resultados, caracterizar los resultados en mayor detalle, incluyendo: grados de similitud entre la secuencia de referencia (*query*) y las otras, tamaños de las secuencias identificadas, anotación de las secuencias obtenidas en el genoma.

- d. Comparar los resultados obtenidos entre sí, así como con los genes reportados en la Tabla 1 del artículo. ¿Qué se recupera y qué se puede estar perdiendo al usar una secuencia de referencia para una búsqueda?

## 2. Explorando ortólogos y parálogos en vertebrados usando Ensembl.

- a. Ingresar a Ensembl y seleccionar el genoma humano. Buscar la mioglobina en dicho genoma. Tomar nota de la caracterización del gen (transcriptos, ortólogos y parálogos, familia de proteínas).
  - b. Explorar en particular la caracterización de genes parálogos. ¿Qué contiene y cómo se correlaciona con los datos de la Fig. 1 del artículo? (Nota: esta parte es laboriosa. Implica ir ubicando los distintos genes, comenzando con la mioglobina misma, en el árbol, tanto para el humano como para otros vertebrados).
  - c. Utilizar la herramienta Gene tree. ¿Qué representa el árbol que aparece en primer lugar? ¿Cómo se interpretan sus distintos componentes?
  - d. Explorar las opciones que ofrece Gene tree para ampliar algunos de los clados que inicialmente aparecen cerrados (triángulos grises), así como para colapsar aquellos grupos que ya exploramos. Notar que este árbol está centrado en nuestra referencia (mioglobina humana). Examinar de manera progresiva las variantes cada vez más lejanas de esta referencia arbitraria (otros primates, otros mamíferos placentados, otros vertebrados...). Observar algunas de las notas al pie del árbol, en particular las referidas a los nodos. ¿A qué se corresponden los nodos de especiación y duplicación (*speciation-duplication nodes*)? ¿Por qué puede haber nodos ambiguos en este sentido?
3. El artículo de referencia se basó en un conocimiento previo de las principales clases de globinas en vertebrados. Identificar dichas clases y sus características principales, tales como la estructura cuaternaria, localización de la expresión, y funciones conocidas.