¿Cuántos genes parálogos hay en el genoma? cumplirán la misma función? con cual trabajaría para hacer filogenética? ¿cómo extraer las secuencias de estos genes?

Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamenteLuego de filtrar por porcentaje de identidad mayor al 90% se observa que en realidad aquellos resultantes de la filtración pueden corresponder a hits espurios, dado que el E value es mucho mayor a 0,1 y esto lo que indica es que lo más probable es que dicho alineamiento haya ocurrido por casualidad y no corresponda a una homología a un gen parálogo; adicionalmente si bien presentan porcentajes de identidad bastante altos, se observa que la longitud del alineamiento es relativamente corta, no superando los 23 pares de bases por lo que esto refuerza la noción de que ocurre por chance. Por el contrario, en el set de datos no filtrados puede observarse que los tres primeros resultados si bien no presentan el mayor porcentaje de identidad si presentan la mayor longitud de secuencia y E-values menores a uno, indicando que la probabilidad de que el alineamiento se de por chance es muy baja. Además estos tres primeros resultados corresponden a LATX010021202 por lo que pareciera que son hits diferentes del mismo gen y que es ese gen el que es parálogo.