

## Herramientas informáticas útiles

Dra Adriana Woods

adrianawoods@gmail.com

### Diseño de primers

Existen numerosos programas para el diseño de primers, como el Blast, entre otros.

El diseño de los primers se efectuó utilizando el programa Oligo 4.1 (sistema operativo MS DOS), dado que me parece un programa amigable.

### Lectura de los electroferogramas:

Se utilizó el software Chromas Lite (<http://www.technelysium.com.au>) para la lectura de los electroferogramas de secuenciación.

Como testigo, seguir la secuencia publicada por Mancuso DJ, 1989.

Para acceder al algoritmo Blast (basic local alignment search tool) se utilizó la página web de acceso pública del NCBI (<http://www.ncbi.nlm.nih.gov>).

### Análisis de predicción *in silico*

Existen herramientas de predicción *in silico* que están siendo usadas cada vez más por diferentes grupos de trabajo. El objetivo de estos programas de predicción es identificar el daño producido o no por un cambio de nucleótido y cómo este cambio afecta o no la estructura y función de la proteína en estudio, y por ende el fenotipo; utilizando estos programas previo a los estudios de mutagénesis, se podría reducir el número de ensayos funcionales requeridos. Dado que son una herramienta muy útil para la caracterización de nuevas sustituciones nucleotídicas cuando los estudios de expresión son difíciles de realizar, utilizamos diferentes aplicaciones informáticas: PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>), SIFT (Sorting Intolerant from Tolerant) (<http://sift.bii.a-star.edu.sg/>), Provean (<http://provean.jcvi.org/index.php>) para obtener una primera caracterización de las mutaciones halladas en los pacientes.

PolyPhen-2 es una herramienta automática para la predicción de posibles efectos de una sustitución de aa en la estructura y función de una proteína humana. Esta predicción se basa en una serie de características que comprenden la información de la secuencia,

filogenéticos y estructurales, que caracterizan a la sustitución. Para una sustitución de un aa dado en una proteína, PolyPhen-2 utiliza varias secuencias diferentes y características basadas en la estructura del sitio de sustitución y las suministra a un clasificador probabilístico. PolyPhen-2 intenta identificar la proteína de consulta como una entrada en el subconjunto de proteínas humanas en las base de datos UniProtKB/Swiss-Prot. Los parámetros calculados dan como resultados distintos tipos de clasificación en cuanto a la sustitución que se evalúa, estos son "benigno", "posiblemente perjudicial", "probablemente dañino" o "desconocido".

SIFT toma la secuencia que se consulta y utiliza la información para predecir la alineación de múltiples sustituciones toleradas y deletéreas para cada posición de la secuencia que se consulta. SIFT es un procedimiento de múltiples etapas que comienza con la búsqueda de secuencias similares, luego elige secuencias estrechamente relacionadas que pueden compartir función similar a la secuencia que se desea analizar, obtiene la alineación de estas secuencias elegidas, y calcula las probabilidades normalizadas para todas las sustituciones posibles de la alineación. Las posiciones con probabilidades normalizadas de  $<0,05$  se predicen como perjudiciales, mientras que aquellas posiciones  $\geq 0,05$  se predicen como toleradas.

### **Análisis de alineamiento**

Dado que el VWF está altamente conservado en un alto número de especies, se realiza el análisis de alineamiento del VWF humano donde están ubicadas las mutaciones no reportadas en la base de datos y que encontramos en los pacientes, con el de otras especies; si la posición en estudio se corresponde con un residuo altamente conservado a lo largo del árbol filogenético, es altamente improbable que dicha posición tolere sustituciones sin cambios en las características funcionales del VWF. El análisis de alineamiento se realiza usando la aplicación informática UniProt KB (<http://www.uniprot.org>). Para ello, la aplicación compara diferentes especies de mamíferos, entre ellos: Chimpanzee; Crab-eating macaque; Lowland gorilla; Northern white-cheeked gibbon; Sumatran orangutan; Rhesus macaque; White-tufted-ear marmoset; Rattus norvegicus; Mus musculus; Sus scrofa; Thirteen-lined ground squirrel; David's myotis; Giant panda; Naked mole rat.

Para chequear si existen registros previos de variantes que hallamos en pacientes y la frecuencia del alelo menor [(minor allele frequency (MAF)] , accedemos a bases de datos de acceso público: Varsome (<https://varsome.com>) y ExAC browser (<http://exac.broadinstitute.org/variant/>).

Mancuso DJ, Tuley EA, Westfield LA, Worrall NK, Shelton-Inloes BB, Sorace JM, Alevy YG, Sadler JE. Structure of the gene for human von Willebrand factor. J Biol Chem 1989; 264(33):19514-27. PMID: 2584182.