

TRABAJO PRÁCTICO N°2

ANALISIS DE SECUENCIAS

Actividad 1: Traducción de secuencias

1. Traducir 4 secuencias (de Arabidopsis, especieZ y las dos secuencias de algas que obtuvo en el TP1) en los seis marcos de lectura posibles y elegir el que corresponda a la proteína putativa, prestando atención a los codones stop prematuros. Para esto empleará un programa online, ExPASy - Translate tool (web.expasy.org/translate/). Seleccione la secuencia que desea traducir del archivo gen.fas, abriéndolo en el block de notas, copie y pegue en la pantalla principal de ExPASy - Translate tool, luego oprima el botón TRANSLATE SEQUENCE.

2. Analice los resultados obtenidos e indique para cada secuencia en qué marco se encuentra la proteína?

Actividad 2: Alineamiento con el programa MEGA6.

1. Realice el alineamiento de las secuencias del gen bajo estudio. Para ello, iniciar el programa Mega. File-Open A File/Session, busque el archivo con las secuencias en FASTA: gen.fas y cliqueé la opción Align. Mega posee dos algoritmos para alinear las secuencias: ClustalW2 y Muscle. Realice el alineamiento con Muscle, seleccione todas las secuencias y diríjase a Alignment options –Align By Muscle para elegirlo como el algoritmo a utilizar (no modifique las opciones que aparecen por default) y cliqueé el botón compute.

2. Agregue las dos secuencias de las algas verdes que obtuvo en el práctico 1 y que deben estar guardadas en un archivo en formato FASTA. Para ello, vaya a la opción EDIT- Insert Sequence From File y seleccione el archivo donde tiene las secuencias de las algas y vuelva a realizar el alineamiento.

3. Traducción del alineamiento completo. Dado que las secuencias son genes que codifican para proteínas, vamos a transformar el alineamiento para verlo como proteínas. En este caso, el alineamiento está en frame +1. Vaya la pestaña Translated Protein Sequences, le aparece una ventana que le pregunta si el código genético es estándar, apriete el botón – No, elija el correspondiente a la secuencia que se encuentra analizando (Plant Plastid o Chlorophycean mitocondria).

- a. Observe las secuencias de amino ácidos a lo largo del alineamiento. ¿Observa los siguientes símbolos “?” y “*” en el alineamiento? ¿Por qué?

- b. Ahora volvamos a la secuencia en nucleótidos seleccionando la pestaña DNA Sequences y guardar, Data – Export Alignment – FASTA format, guardar con el nombre de alineamiento1.
4. Agregar al alineamiento la secuencia de la especie Z e intente alinear. Si el alineamiento de alguna de las secuencias no resultó, selecciónela y realice el complemento reverso (Data - reverse complement) de la misma y alinee nuevamente.
5. Incorpore al alineamiento la secuencia de la especie Y que posee un intron y realinee. ¿Cuál es el tamaño del intron? (Para responder esta pregunta realice un BLASTx. Alternativamente, puede observarlo en el alineamiento).

Actividad 3: Tipos de archivos para secuencias y alineamientos.

1. Eliminar la región del intrón del alineamiento. Para ello seleccione desde arriba la región a eliminar en todas las secuencias, vaya a Edit – Delete.
2. Guardar el alineamiento realizado en distintos formatos:
 - a. como nexus, Data-Export alignment- PAUP format, nombrar al nuevo archivo como alignment.nexus
 - b. como fasta, Data-Export alignment- FASTA format.

Observar las diferencias entre los tipos de archivos, abriéndolos en Block de Notas u algún otro programa para textos simples. (NOTA: En el futuro es recomendable utilizar el programa Textedit (u otro similar) que son gratuitos para utilizar en análisis de secuencias. NUNCA utilizar un procesador de texto como Word.)

FASTA: tipo de formato utilizado en NCBI. Todas las entradas en la base de datos están disponibles en formato FASTA. Una gran variedad de programas también es capaz de leer este tipo de archivos. Sirve para secuencias u alineamientos. En el caso de los alineamientos, las secuencias pueden estar alineadas o no.

NEXUS: este formato es muy utilizado por programas como GARLI, MacCLade, ModelTest, MrBayes, PAUP, etc. Los archivos nexus generalmente llevan la extensión .nxs o .nex. En el caso del programa GARLI, el archivo nexus puede poseer la extensión .phy. Los archivos nexus pueden ser “interleaved” o “non-interleaved”.