**Guía para laboratorio 6: Análisis de secuencias génicas en genbank**

Para esta práctica debes haber realizado el prelaboratorio y visualizado los tutoriales. Entonces te resultará más fácil trabajar con tu compañero de laboratorio.

Debes descargar también la plantilla word que contiene la secuencia de trabajo, ya que sobre ella deberán resaltar con sus propios colores, letra cursiva, letra en negrilla, o letra subrayada los elementos solicitados en el presente cuestionario, conforme a sus coordenadas identificadas en la secuencia de trabajo en genbank. Esto debe hacerse en forma muy precisa…No olvides entonces al realizar el ejercicio, incluir una leyenda indicando a qué corresponde cada color o tipo de resaltado empleado.

PROCEDIMIENTO:

* Ingresen a la página principal del NCBI y desde ahí a la base de datos **Nucleotide.**
* Asegúrense que a la izquierda del cuadro de búsqueda figure la base de datos “nucleotide”para que en el cuadro de búsqueda por términos, escriban el siguiente número de acceso: **AF222792**.
* Serán direccionados al registro de la secuencia que analizaremos en esta práctica. Tomense el tiempo para recorrerla e identificar las diferentes secciones que vimos en los tutoriales del prelaboratorio (encabezado, información general asociada, detalles de elementos identificados con sus coordenadas, y al final secuencia completa organizada por bloques de 10 nucleótidos).
* Abran el archivo con la plantilla word, la cual será también un documento de trabajo.

Ya tienen todo lo necesario para hacer la práctica, recuerda que pueden volver a mirar los videos tutoriales en el prelaboratorio, si necesitan recordar algo puntualmente.

**Cuestionario:**

1. ¿Cuál es el nombre del organismo al que pertenece esta secuencia registrada?

2. ¿Qué tipo de organismo es (procariota/eucariota?

3.Cuéntanos más a cerca de su Taxonomía:

4. ¿A qué tipo de molécula correspnde esta accesión? Y cuál es su tamaño?

5. ¿Cuándo fue sometida y cuál fue la fecha de su última actualización?

6. ¿Cuántos artículos científicos publicados y disponibles en Pubmed tiene asociados este registro?

7. ¿Según el/los título(s) o abtract(s) de los artículos asociados, resalte un resultado relevante asociado con esta secuencia y la función global asociada.

8. ¿Cuántas secuencias codificantes (CDS) están contenidas en la totalidad de la secuencia?

9. ¿Cuáles son los nombres asociados a cada secuencia codificante o proteína? (generalmente es un nombre de 3 a 5 letras, por ejemplo “DsxB”) ¿qué tipo de proteína codifican? (se recomienda enumerar los nombres y asociar a cada nombre la función probable de la proteína, por ej. DsxB: proteína flajelar putativa ).

10. La secuencia que figura en cualquier base de datos corresponde a una de las dos cadenas de ADN y por norma su polaridad es de 5’ a 3’ siempre. Dependiendo de la orientación del gen codificado, esta cadena o secuencia, visible al final del registro, corresponderá a la **cadena codificante** si la dirección de transcripción / traducción del gen va hacia la (derecha/ izquierda)\_\_\_\_\_\_\_\_ o a la **cadena molde** si la dirección del gen va hacia la (derecha/ izquierda)\_\_\_\_\_\_\_\_\_\_.

11. Para cada secuencia codificante identificada en la pregunta 9, identifica ahora su respectiva orientación o sentido de lectura (derecha o izquierda). Recuerde que puede haber genes que, estando codificados en la cadena complementaria a la que se visualiza, tienen entonces una dirección o sentido de transcripción y traducción opuesta. Escribe junto al nombre de ORF o CDS su orientación o sentido de lectura (Derecha o izquierda)

A continuación trabajaremos sobre la plantilla word:

Elije colores distintos para señalar, sobre la secuencia contenida en la plantilla Word, cada uno de los elementos solicitados a continuación. Para esto necesitarás ubicar las coordenadas respectivas y por supuesto ir armando una leyenda que te permita ubicarte y entender lo que vas resaltando: ten en cuenta que varias secuencias se pueden traslapar, entonces trata de encontrar una forma de mostrar esos sobrelapamientos de manera que se puedan entender.

1. Resalta en primer lugar, cada una de las secuencias codificantes (CDS) empleando resaltadores de colores diferentes.
2. Después de resaltar todas las secuencias codificantes y en función de lo que contestaste en la pregunta 11 sobre la dirección de cada gen, verifica en los extremos de todos los ORFs que van hacia la derecha el *codón de inicio* (***atg***, o en algunos casos muy raros puede ser ***gtg***) y el codón de parada (***taa, tag* o *tga***) marcándolos en negrilla.
3. Detente ahora sobre cada una de las secuencias codificantes con dirección opuesta (las que van hacia la izquierda, por lo que estarían codificadas por la cadena complementaria) y resalta igualmente la primera y última tripleta de cada ORF en negrilla:

Te darás cuenta que estas tripletas no coinciden con la secuencia *atg* ni con ningún codon de parada…pero es lógico! Porque estás visualizando la cadena que no es codificante para los genes con esta orientación! La cadena que acabas de resaltar, para el caso de estos genes con orientación opuesta, sería la molde…entonces, la cadena codificante corresponde a una secuencia complementaria a la que estás viendo, y además la lectura de la misma se haría de derecha a izquierda al ser antiparalela.

Revisa esto con más detenimiento, al situarte sobre una de las secuencias codificantes u ORF resaltadas en el punto anterior: la tripleta del extremo izquierdo corresponde al codón de parada y la del extremo derecho al codón de inicio, puesto que el gen se lee en sentido contrario. Mentalmente realiza la complementaria de cada una de esas tripletas, y leelas de derecha a izquierda: si lo hiciste bien, deberían ahora sí coincidir con codones stop y de inicio ; )

Si tienen claro esto pasen a los siguientes puntos, de lo contrario, no dudes en discutirlo nuevamente con tu compañero o con el resto del grupo de laboratorio.

1. Empleando la herramienta “*Highlight sequence features*”, pero esta vez seleccionando la categoría “regulatory” identifica y resalta cada una de las secuencias Shine Dalgarno (**Ribosome Binding Site o RBS**).

12. Detente al terminar de resaltar todas las RBS: ¿cuántas se logran identificar en total? ¿Cuántas debería haber?

Fíjate ahora en la posicion de cada RBS con respecto a los ORFs. Lo que vimos en clase es que cada ORF bacteriano debe tener su respectiva RBS unos 10 nt (aprox.) corriente arriba (antes del codón de inicio). Entonces, con base en esto, ajusta el color de cada RBS identificada, empleando el mismo color que el ORF al que debe asociarse (ojo con los genes con dirección opuesta, ahí todo se invierte…)

* + 1. Empleando la misma herramienta anterior, identifica y resalta las cajas -10 y -35 (ej. se leeran como “minus 10 signal”) de los promotores. Ten muy presente cuántas hay y en qué cadena están ubicadas…, y con base en esto usa el mismo color para las dos cajas de un mismo promotor.

Te darás cuenta que falta una caja -35 que los autores no reportan, esto puede pasar.

* + 1. Identifica y resalta también la secuencia terminadora de la transcripción reportada. ¿Faltaría alguna según la organización de los genes?
    2. Emplea ahora la categoría de “Misc-feature” y señala resaltándolos, los respectivos +1 o puntos de inicio de la transcripción: asociálos igualmente con el mismo color que las cajas -10 resaltadas anteriormente.

Detente un momento y observa en tu archivo word los +1 y cajas -10 y -35 resaltadas en esa zona. Observa qué genes están a la izquierda de esa region y que dirección tienen. Fíjate en los genes que están a la derecha de esa región promotora y en la dirección que tienen? Intenta entender cuál es el promotor (+1, caja -10 y -35) del gen que va en dirección de la derecha y del gen que va en dirección opuesta. Para esto recuerda la estructura básica de un gen y de una unidad transcripcional vista en clase y la ubicación de las cajas del promotor y +1, con respecto al ORF.

13. Fíjate ahora donde quedó ubicado el terminador de transcripción reportado: ¿Qué ORFs concretas se encuentran entre la región promotora y el terminador resaltado? \_\_\_\_\_\_\_\_\_\_

Intente mirar la organización del resto de genes que se transcriben en dirección opuesta y verifique el promotor que permite su transcripción, donde debería estar el terminador (pregunta 12 b).

14. Ambas organizaciones contienen varios genes codificantes, ¡esto significa que cada gen codificante hace parte de una unidad transcripcional diferente? \_\_\_\_\_\_¿A qué organización génica propia de este tipo de organismos corresponden?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

15. Teniendo en cuenta las coordenadas del +1 y del terminador de transcripción, cuál sería el tamaño del mRNA que porta el gen ORFX?\_\_\_\_\_\_\_ Y el tamaño aproximado del mRNA que porta el gen merR? \_\_\_\_\_\_\_

16. Ubica la secuencia codificante de la proteína ORFX y escribe los 4 primeros codones del ADN:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Con ayuda de un código genético, realiza la traducción de los 4 primeros aminoácidos respectivos:\_\_\_\_\_\_\_\_\_\_\_\_\_\_

17. Ubica la secuencia codificante de la proteína MerR y escribe correctamente los 4 primeros codones de su respectiva cadena codificante (ten en cuenta orientación!):

5’ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_3’. Realiza la traducción de estos 4 primeros aminoácidos.

18. Observa todo lo que resaltaste en el archivo word y no olvides intentar proponer una posible ubicación de algunos elementos faltantes que los autores no reportaron en este registro.

Conjuntamente con tu compañero,realicen sobre la línea horizontal que encontrarán más abajo y que representa toda la secuencia de ADN trabajada, un diagrama que resumiría la organización de todos los elementos importantes para asegurar la transcripción y traducción de los genes.

Deben emplear los siguientes códigos para representarlos todos (hay más de uno contenido en la secuencia):

(Si les queda más cómodo, realicen el diagrama a mano en una hoja usando colores, pero respetando los mismos códigos de representación:

Secuencias codificantes para todas las proteínas identificadas: colocar el nombre de la proteína en el centro de la flecha y presenta la orientación correcta de su lectura (base flecha= codón inicio, punta = codón parada)

Gira la fecha según cada caso, pero coloca todas las “flechas” que apliquen con sus nombres sobre la misma línea que representa la totalidad de la secuencia de trabajo

\*

T

P

RBS o secuencias de Shine Dalgarno: ubicarlas todas correctamente con respecto a cada secuencia codificante según como las resaltaste

Terminador(es) de transcripción: bien ubicados con respecto a los elementos anteriores

Promotor(es) y +1 colocando la dirección correcta de transcripción (gira la fecha si lo necesitas y según el caso)