



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA
UNIDAD IZTAPALAPA

Manual de prácticas de laboratorio

Introducción a la Bioinformática



Ma. Dolores **García Suárez**

Héctor Serrano



UNIVERSIDAD AUTÓNOMA METROPOLITANA

Dr. Salvador Vega y León
Rector General

Mtro. Norberto Manjarrez Álvarez
Secretario General

UNIDAD IZTAPALAPA

Dr. Javier Velázquez Moctezuma
Rector de Unidad

Dr. Miguel Ángel Gómez Fonseca
Secretario de Unidad

Dr. Rubén Román Ramos
Director de la División de Ciencias Biológicas y de la Salud

Dra. Milagros Huerta Coria
Coordinadora de Extensión Universitaria

Lic. Adrián Felipe Valencia Llamas
Jefe de la Sección de Producción Editorial

Primera Impresión 2013

UNIVERSIDAD AUTÓNOMA METROPOLITANA

UNIDAD IZTAPALAPA

Av. San Rafael Atlixco No. 186, Col. Vicentina,
Del. Iztapalapa, C.P 09340, México D.F. Tel.: 5804 4600

Impreso y hecho en México/*Printed in Mexico*

Índice

| | |
|--|----|
| Prólogo | 5 |
| Capítulo 1 Búsqueda de literatura especializada | 7 |
| Capítulo 2 Actualización automatizada de literatura científica | 15 |
| Capítulo 3 Obtención y análisis de secuencias de proteínas utilizando un servidor en línea | 21 |
| Capítulo 4 Simulación de la estructura de genes a partir de la secuencia de aminoácidos. . . | 33 |
| Capítulo 5 Búsqueda de información general de enfermedades genéticas | 43 |
| Capítulo 6 Diseño de oligonucleótidos para amplificación | 49 |
| Referencias generales | 62 |

Prólogo

El material que está a punto de analizar y seguir, fue pensado para permitir que los alumnos del nivel Licenciatura y las personas que se están incorporando a la investigación científica en general, puedan tener una guía simple para poder integrarse a la corriente actual de la actividad de la investigación.

En primer lugar, este manual está enfocado a las necesidades de la Genética Molecular Humana. Sin embargo, dada la generalidad y estructura del texto, puede ser aplicable a otros campos. Las herramientas que se sugieren no son lo más sofisticadas que podrían encontrarse disponibles. De hecho, esta simplicidad permite su amplia aplicación y la facilidad de interpretación que normalmente requiere un curso introductorio. En cada una de las páginas electrónicas que se utilizan se encuentra material de apoyo que permite un uso amigable.

La mayoría de las plataformas utilizadas permite que sean modificadas por lo que si se tienen los conocimientos avanzados, permiten llenar los requerimientos de lo que sería el desarrollo de procesos de búsqueda, comparación, análisis y evaluación de resultados específicos que requiera un laboratorio de investigación en alguna de las disciplinas “ómicas” que se estén desarrollando.

Los autores de este manual esperamos que les sea útil y les permita adentrarse en el uso de herramientas que hagan aún más atractiva la actividad de investigación que estén desarrollando.

Capítulo 1

Búsqueda de literatura especializada

La labor del investigador es una acción continua y que requiere en primera instancia de una documentación adecuada, confiable y que permita el avance del campo del conocimiento en el que se desempeña. En la mayoría de los casos, trata de ser vanguardista por lo que es necesario mantener un acervo no sólo de las evidencias experimentales que apoyan los resultados de uno u otro grupo de trabajo, sino que abarque la importancia de ambos, evalúe en la medida correcta los aciertos y remarque los campos que son necesarios apoyar para dar congruencia a un fenómeno específico.

Podemos dividir en dos grandes etapas la labor de documentación del investigador a partir del desarrollo de una herramienta brindada por la red de redes informáticas. Era común que un investigador pasara varias horas en la biblioteca de la institución a la que pertenecía leyendo o fotocopiando los artículos de las revistas donde comúnmente se publicaban los resultados de su campo. Esta costumbre tenía la limitante de la cantidad de revistas a las que estaba suscrita la institución y era más o menos solventada por los intercambios bibliotecarios, con la consiguiente ampliación del tiempo que requería para leer un artículo que muchas veces cuando llegaba, ya no tenía el impacto de cuando fue requerido.

Un servicio de alerta se ofreció a través de una compañía denominada Institute for Scientific Information (ISI) que tenía sus oficinas generales en Philadelphia, Estados Unidos. El servicio de documentación Current Contents® primeramente por correo ordinario y posteriormente apoyados por computadora permitían una actualización relativamente aceptable. De manera general, con una sola suscripción se podían atender varios laboratorios de una sola Institución y cada uno de los miembros tenía un tiempo y prioridad de revisión especificado para hacer eficiente dentro de lo posible el sistema.

Con el desarrollo de una función alterna a la que se le denominó "Medline", se entró en el uso de las comunicaciones vía computadora. Este servicio, que originalmente tenía un costo de suscripción permitía mantener búsquedas retrógradas de hasta 5 años de la literatura publicada en un tema particular o tener un listado semanal de los artículos publicados sobre temas específicos. Al igual que con Current Contents, el acceso era dependiente de la capacidad económica del laboratorio y la disponibilidad de al menos un aparato de fax o acceso a la red institucional que en México a finales de la década de 1980, era restringido solamente a la UNAM y CONACyT.

La fundación del Centro Nacional de Información Biotecnológica (NCBI) como una división específica de la Biblioteca Nacional de Medicina (NLM) dependiente de los Institutos Nacionales de Salud (NIH) utilizando la plataforma y base de datos de Medline consolidará la herramienta de búsqueda poderosa que conocemos como PubMed y su central de artículos sin costo PubMed Central. Esta base de datos se actualiza cada 24 horas y es alimentada fundamental aunque no exclusivamente por las oficinas editoriales de las revistas.

Objetivo

Mostrar el uso de la base de datos PubMed en el acopio de literatura especializada.

Desarrollo:

Acceda a la página principal del NCBI a través de la dirección <http://ncbi.nlm.nih.gov>

Puede llegar a la página de PubMed ya sea desplegando el menú en la casilla de Búsqueda (SEARCH) o en la liga de Recursos comunes (POPULAR RESOURCES) (Figura 1.1). Si lo hace en la ventana del dialogo, puede incluir inmediatamente la palabra clave que identifique su búsqueda. Si lo hace por la liga de Popular resources, una vez que se despliega la página de PubMed escriba la palabra clave.

En ambos casos, puede utilizar diversos tipos de entradas (AUTOR, KEY WORD, JOURNAL) ya que el identificador incorporado tiene esta facilidad. Al presionar el botón SEARCH se despliega la información requerida. La duración de la búsqueda depende de la complejidad del campo, la actividad que tenga en ese momento y el volumen de información que se tenga.

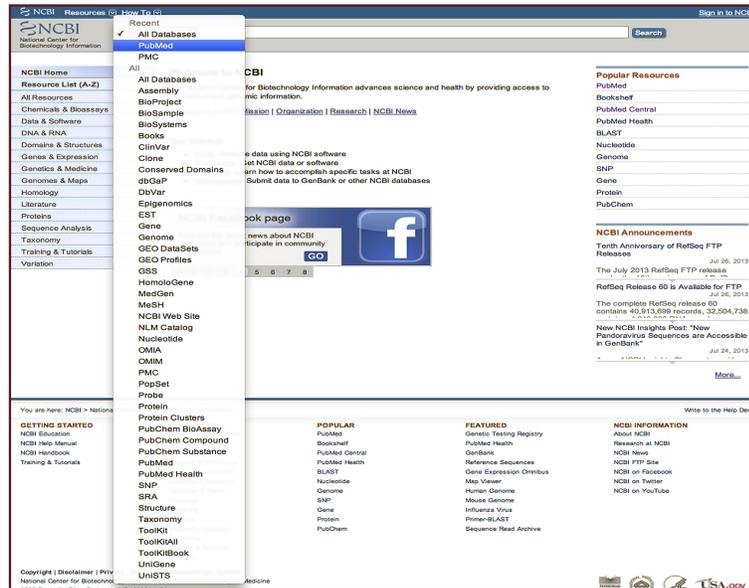


Figura 1.1a. Pantalla de inicio del NCBI.

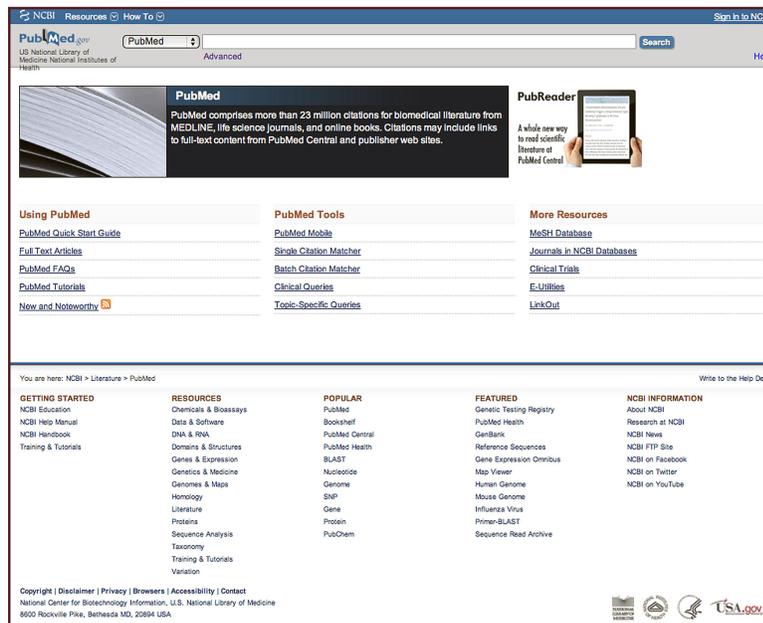


Figura 1.1b. Pantalla PubMed.

La búsqueda en PubMed Central brinda un menor número de artículos referidos pero casi siempre disponibles para ser adquiridos inmediatamente sin costo o requerimientos (Figura 1.2). A esta iniciativa se le conoce como OPEN ACCESS y varias revistas están haciendo esfuerzos por poner a disposición de la comunidad científica su contenido. Algunas otras permiten el Open Access a revistas de sus archivos que fueron publicadas un lapso variable de tiempo atrás; finalmente, algunos autores han pagado el costo que representa a los editores para que la información sea catalogada como Open Access, básicamente mediante apoyos otorgados por el NIH o el gobierno de Estados Unidos.

The screenshot shows the PubMed search results for the keyword "Phytoestrogen". The search was performed on the PubMed website, and the results are sorted by "Recently Added". The search results list 1 to 20 of 9255 results. The first few results are:

- Phytoestrogens mediated anti-inflammatory effect through suppression of IRF-1 and pSTAT1 expressions in lipopolysaccharide-activated microglia.** Jantaratornoi N, Utaisinchatoen P, Sanvavinda P, Thampitthak A, Sanvavinda Y. *Int Immunopharmacol*. 2013 Aug 9. doi:10.1016/j.intimp.2013.07.013. [Epub ahead of print]. PMID: 23923252 [PubMed - as supplied by publisher] [Related citations](#)
- The E-screen test and the MEL N gene-reporter assay used for determination of estrogenic activity in fruits and vegetables in relation to pesticide residues.** Tziانا S, Arianna P, Annalisa L, Sara C, Giorgio G. *Food Chem Toxicol*. 2013 Aug 7. doi:10.1016/j.fct.2013.07.087. [Epub ahead of print]. PMID: 23923258 [PubMed - as supplied by publisher] [Related citations](#)
- Microbial sulfation of 8-crenylnaringenin.** Barmařska A, Tronina T, Huszcza E, Z Natufowech C. 2013 May-Jun;86(6):231-5. PMID: 23923220 [PubMed - in process] [Related citations](#)
- An estrogen receptor dependent mechanism of oroxynin a in the repression of inflammatory response.** Wang H, Guo Y, Zhao X, Li H, Fan G, Mao H, Miao L, Gao X. *PLoS One*. 2013 Jul 29;8(7):e69555. doi:10.1371/journal.pone.0069555. Print 2013. PMID: 23922237 [PubMed - in process] [Free PMC Article](#) [Related citations](#)
- Soy extracts suppressed iodine uptake and stimulated the production of autoimmunogen in rat thyrocytes.** Tran L, Hammuda M, Wood C, Xiao CW. *Exp Biol Med (Maywood)*. 2013 Jun 1;238(6):823-30. doi:10.1177/1535370213489488. PMID: 23918874 [PubMed - in process] [Related citations](#)
- Soy food consumption and breast cancer.** Mourouli N, Panagiotakos DB. *Metulias*. 2013 Aug 2. doi:10.1016/j.metulias.2013.07.006. [Epub ahead of print]. PMID: 23916376 [PubMed - as supplied by publisher] [Related citations](#)
- Association of urinary phytoestrogen concentrations with serum concentrations of prostate-specific antigen in the national health and nutrition examination survey.** Waiser-Domjan E, Richard A, Eichholzer M, Platz EA, Linselsen J, Rohmann S. *Nutr Cancer*. 2013;86(8):819-9. doi:10.1080/01635581.2013.801999. PMID: 23907224 [PubMed - in process] [Related citations](#)

The interface also includes filters for article types, text availability, publication dates, species, and a "Results by year" chart. There are also sections for "PMC Images search for Phytoestrogen" and "Titles with your search terms".

Figura 1.2a. Resultados simples utilizando la palabra clave "Phytoestrogen*" en las bases PubMed.

The screenshot shows the PubMed Central search results for the keyword "Phytoestrogen". The search was performed on the PMC website, and the results are sorted by "Default order". The search results list 1 to 20 of 3346 results. The first few results are:

- Development of an updated phytoestrogen database for use with the SWAN Food Frequency Questionnaire: intakes and food sources in a community-based, multiethnic cohort study** Mei-Hua Huang, Jean Norris, Weijuan Han, Torin Block, Ellen Gold, Sybil Crawford, Gail Greendale *Nutr Cancer*. Author manuscript; available in PMC 2015 June 6. Published in final edited form as: *Nutr Cancer*. 2012; 64(2): 228-244. Published online 2012 January 2. doi: 10.1080/01635581.2012.638434 PMID: PMC3874882 [Article](#) [PubMed](#) [PDF-347K](#)
- Estimated dietary phytoestrogen intake and major food sources among women during the year before pregnancy** Suzan L Carmichael, Amparo G Gonzalez-Feliciano, Chen Ma, Gary M Shaw, Mary E Cogswell *Nutr J*. 2011; 10: 105. Published online 2011 October 8. doi: 10.1186/1475-2891-10-105 PMID: PMC3196893 [Article](#) [PubMed](#) [PDF-282K](#)
- Activation of Southern White Rhinoceros (*Ceratoptherium simum simum*) Estrogen Receptors by Phytoestrogens: Potential Role in the Reproductive Failure of Captive-Born Females?** Christopher Tubbs, Phillip Hartig, Mary Cardon, Nicole Varga, Matthew Milnes *Endocrinology*. 2012 March; 153(3): 1444-1452. Published online 2012 January 17. doi: 10.1210/en.2011-1982 PMID: PMC3281539 [Article](#) [PubMed](#) [Supplemental Material](#)
- Phytoestrogen consumption from foods and supplements and epithelial ovarian cancer risk: a population-based case control study** Elisa V Bandera, Melony King, Urmila Chandran, Lisa E Paddock, Lorna Rodriguez-Rodriguez, Sara H Olson *BMC Womens Health*. 2011; 11: 45. Published online 2011 September 23. doi: 10.1186/1472-6874-11-45 PMID: PMC3196697 [Article](#) [PubMed](#) [PDF-192K](#)
- Vascular Effects of Phytoestrogens and Alternative Menopausal Hormone Therapy in Cardiovascular Disease** Vahide B. Gencel, Mina M. Benjamin, Shaik N. Bahou, Raouf A. Khalil *Mini Rev Med Chem*. Author manuscript; available in PMC 2013 February 1. Published in final edited form as: *Mini Rev Med Chem*. 2012 February 1; 12(2): 149-174. PMID: PMC3288319 [Article](#) [PubMed](#) [PDF-2.2M](#)

The interface includes filters for "Filter your results" (All (3346), NIH Grants (1557)), "PMC Images search for Phytoestrogen", and "Find related data".

Figura 1.2b. PubMed Central.

En el desarrollo del curso, estaremos interesados en tres tópicos fundamentales: la regulación de la esteroidogénesis en mamíferos, el uso de marcadores moleculares de diabetes y enfermedades asociadas al aparato circulatorio y las pruebas de paternidad.

Para cada uno de estos tópicos será necesaria la documentación de los antecedentes acumulados durante los 10 años previos. Para llenar este requisito, abra la liga LIMITS y en la ventana de diálogo donde aparece "PUBLICATION DATE" indique 10 YEARS. Asegúrese de haber introducido la palabra clave en la ventana de diálogo principal (SEARCH). Si es necesario, indique los parámetros adicionales. Nuestro interés es con estudios en animales de sexo indistinto, artículos publicados únicamente en Inglés, Francés, Alemán, Italiano, Portugués y Español. Indique la búsqueda oprimiendo el botón SEARCH del final de la página (Figura 1.3).

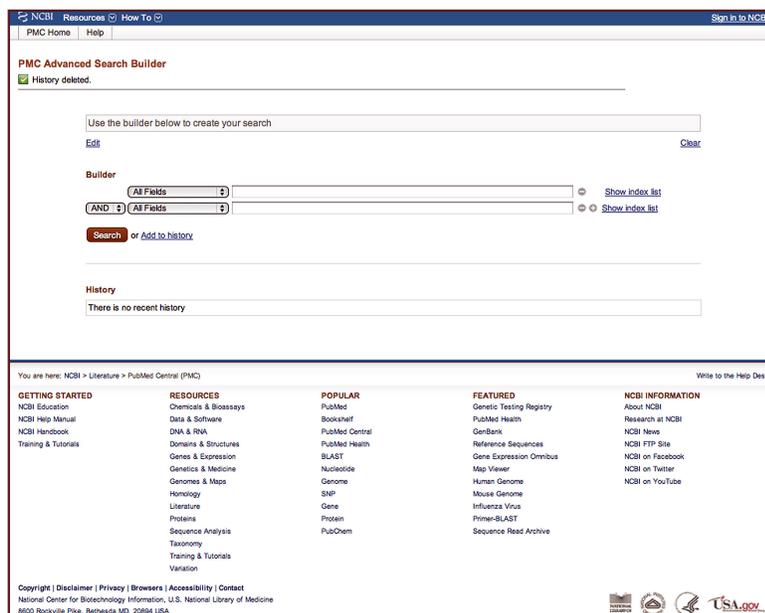


Figura 1.3a. Ventana de diálogo de búsqueda de avanzada de PubMed.

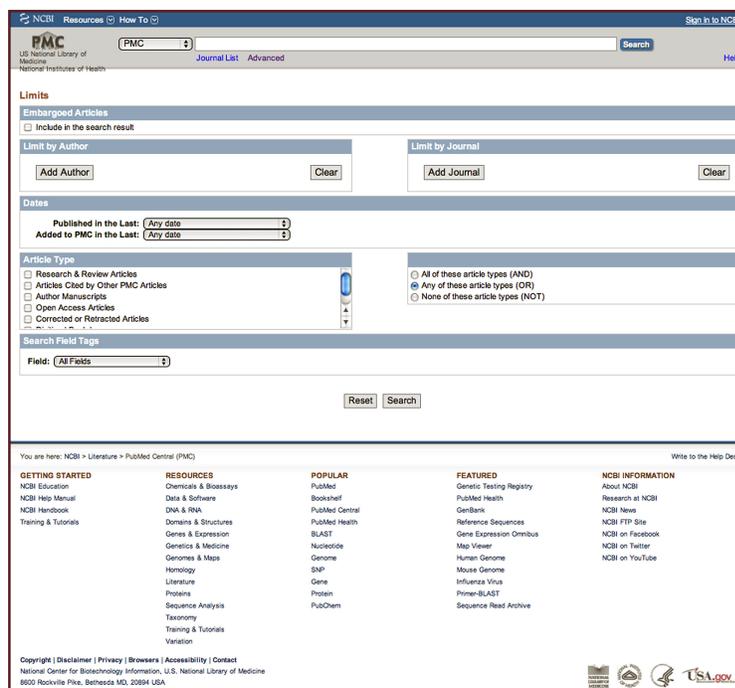


Figura 1.3b. PubMed Central. PMC ofrece la opción de delimitar la búsqueda utilizando los criterios que aparecen en esta ventana. PubMed lo hace directamente en la ventana de resultados (vea el campo izquierdo de la figura de arriba).

Después que el buscador le entregue los resultados, seleccione sólo aquellos Open Access pulsando la liga FREE FULL TEXT que aparece dentro del recuadro REFINE YOUR SEARCH localizado en la porción de herramientas (lado derecho de la pantalla).

Una vez que se tiene acceso al listado de artículos gratuitos, aumente el número de referencias que puede ver por página abriendo el cuadro de diálogo localizado al inicio del encabezado del listado bajo el renglón DISPLAY SETTINGS. Obtendrá un cuadro. Al terminar de hacer la selección, pulse el botón APPLY que está al fondo del cuadro de diálogo

Actividad para reportar

Utilizando las palabras clave STEROID*, MOLECULAR MARKERS, TYPE 2 DIABETES MELLITUS, ADRENERG* RECEPTOR* y PATERNITY TESTING haga una búsqueda de la información que se hubiese publicado sobre el tema en los últimos 10 y 5 años.

Anote la cantidad de artículos originales, de revisión y editoriales que se tenga en cada lapso. Anote la cantidad de artículos disponibles bajo el esquema de OPEN ACCESS.

Anote la cantidad relativa de artículos que se escribieron en cada uno de los idiomas principales que estamos utilizando en el curso.

Anote el número de estudios que se han realizado en animales y humanos así como la proporción de estudios que se hicieron por sexo de los individuos.

En el caso de estudios en humanos, indique el número de artículos en los cuales están contenidos datos de estudios con aplicación clínica.

Obtenga un artículo de investigación relacionado con los temas de estudio del curso y uno de revisión o meta-análisis, si es aplicable.



Figura 1.4a. Ventana de acceso a las bases de datos de la biblioteca UAM-I.

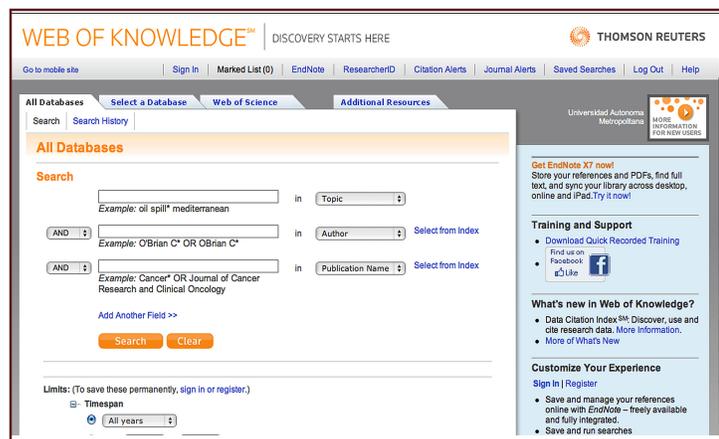


Figura 1.4b. Acceso a ISI Web of knowledge si se está conectado a una computadora de la Unidad o mediante el acceso remoto permitido.

Utilizando la base de datos electrónica ISI WEB of Knowledge (bases de datos disponible en línea dentro del portal de la Biblioteca UAM-I ya sea mientras está utilizando computadoras conectadas a la red de la Universidad o mediante el ingreso de su clave de identificación (DIRECCION IP) para poder tener acceso remoto), realice una búsqueda similar (Figura 1. 4). Utilice al menos otras dos bases de datos electrónicas y realice búsquedas similares. Como punto de partida, utilice la opción de Google Académico (<http://scholar.google.com.mx>) y Microsoft Academic Search (<http://academic.research.microsoft.com>, Figura 1.5), las páginas electrónicas más conocidas y de mayor consulta. En cada caso obtenga un artículo de investigación de cada tema del curso y una revisión o meta-análisis, si es aplicable.



Figura 1.5a. Ventana de acceso a Google Académico.

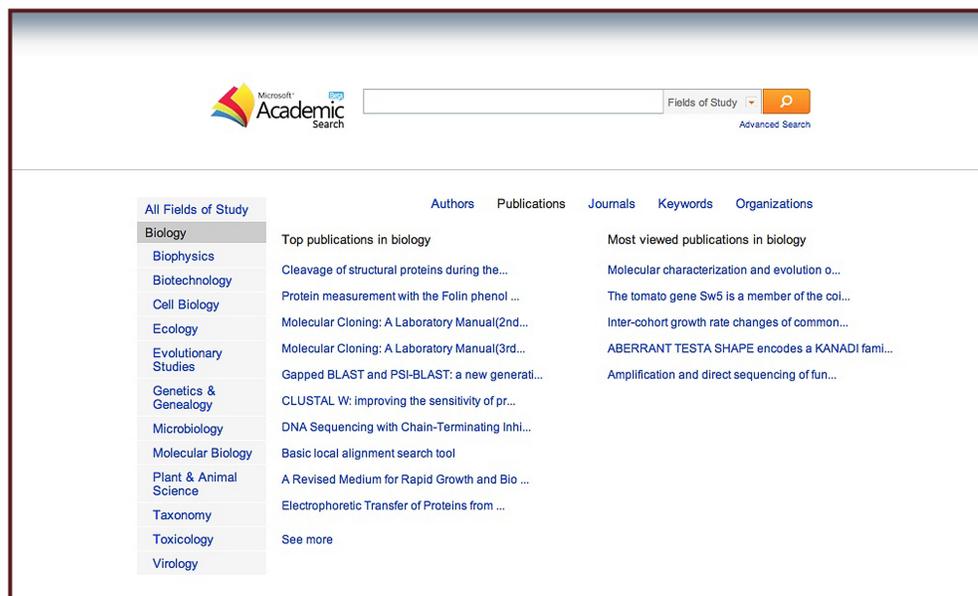


Figura 1.5b. Microsoft Academic Search popular buscadores.

Indique las ventajas y limitantes que presenta el uso de ISI Web of Knowledge y las otras dos bases de datos que consultó cuando se comparan con PubMed y PubMed Central.

Referencias

-  PubMed Help [Internet]. National Center for Biotechnology Information. 2005. Disponible en el URL <http://www.ncbi.nlm.nih.gov/books/NBK3830>
-  Sawyers, E., Barret, T., Benson, D.A., Bolton, E., Bryant, SH y otros 38 co-autores. 2012. Database resources of the National Center for Biotechnology Information. Nucleic. Acids. Res. 40:D13-D25. Disponible en el URL <http://nar.oxfordjournals.org/content/40/D1/D13.full.pdf+html>

Capítulo 2

Actualización automatizada de literatura científica

La generación de conocimiento que implica la investigación requiere de una actualización constante. La búsqueda en las bases de datos puede ser repetitiva y para poder ser eficiente, debería ser al mismo ritmo que es actualizada la base de datos en línea que se consulta. En el caso de PubMed y PubMed Central, esta actividad diaria podría implicar que no se encuentren nuevos artículos y la consulta sea improductiva.

Una posible alternativa es recibir el índice de las revistas en las cuales se publica normalmente o preferencialmente el tipo de estudios en los cuales estamos interesados. Esto tiene como ventaja el estar no sólo actualizado respecto a la literatura impresa sino que incluso se puede tener acceso a los temas que la revista publicará en breve. La desventaja es que perderíamos aquellos artículos en los cuales por no tener el enfoque particular y reportar sólo avances metodológicos, no sean aceptables para la revista en cuestión.

Esta limitante podría tener una salida de búsqueda compulsiva equivalente a ir a la biblioteca al menos una vez a la semana y ver cuales son las novedades o, en la actual era tecnológica, recurrir a la búsqueda rutinaria de las fuentes electrónicas confiables (PubMed, PubMed Central, Scopus, etc) teniendo que iniciar una sesión cada vez que se haga la consulta haciéndolo tedioso y en la mayoría de los casos, frustrante. Una posibilidad es utilizar un servicio de actualización que nos haga la búsqueda rutinaria. Varias compañías han desarrollado diferentes buscadores de literatura con capacidades y costos diferentes. Sin embargo, la comunidad científica ha tratado de mantener una actitud más abierta y sin fines de lucro. Con este objeto, varios servidores ofrecen la posibilidad de ocupar un espacio en el cual se pueda recopilar la literatura específica de un investigador para después enviarle las características mínimas para que el investigador pueda tener acceso a ella, básicamente a través del correo electrónico.

Objetivo

Mostrar un sistema automatizado de actualización bibliográfica mediante un buscador remoto.

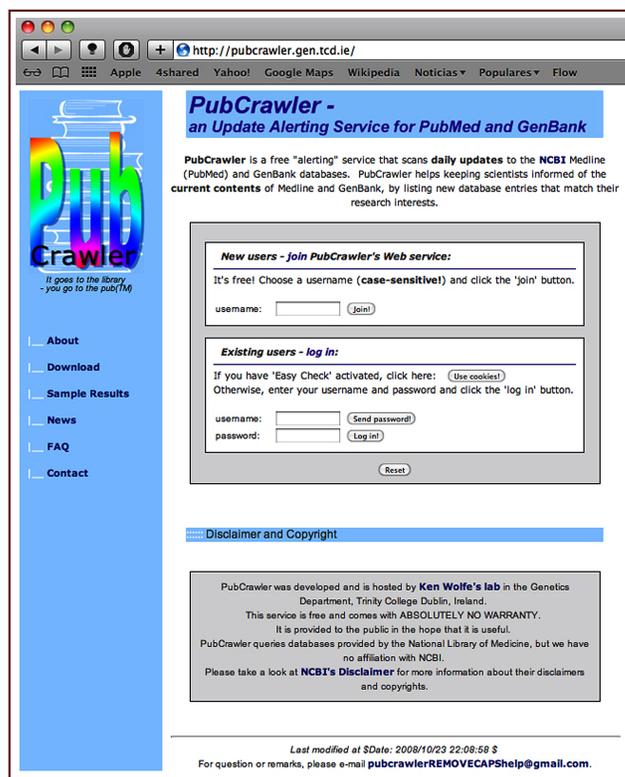
Desarrollo

En el inicio de su buscador, escriba la dirección <http://pubcrawler.gen.tcd.ie>

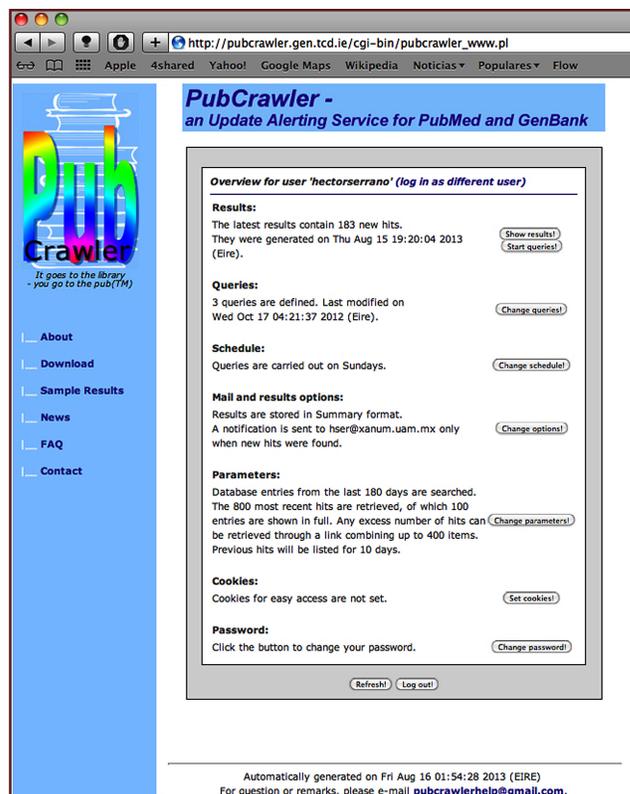
Después de que se visualice la ventana de acceso, registre la apertura de una cuenta ingresando los datos necesarios para ello (Figura 2.1).

Una vez creada la cuenta y dado de alta el buzón electrónico, introduzca los términos de la búsqueda. Asegúrese que en el despliegue del menú se seleccione "ALL FIELDS". Indique que el lapso de actualización como "WEEKLY"

Al terminar de introducir los términos de la misma manera que la sesión anterior, envíe la consulta. En algunos casos, la respuesta es inmediata por lo que puede abrir su dirección de correo para analizar las referencias que sean depositadas en su buzón.



Figuras 2.1a. Ventana de inicio - registro.



Figuras 2.1b. Desarrollo de perfil de búsqueda de PubCrawler.

Cada una de las entradas (QUERY o ALIAS) puede tener varias palabras clave. Si la investigación en la cual está inmiscuido requiere de la separación o agrupación de palabras clave, cada una de ellas puede obtenerse por separado.

Una vez acreditada la cuenta, puede ingresar a ella utilizando las características de usuario y clave de acceso para ampliar, modificar o suprimir las consultas y recopilaciones solicitadas (Figura 2.2)

A lo largo del curso, determine el impacto que tiene el campo de acuerdo al número de citas y la relación con el curso que tenga.

The screenshot shows the PubCrawler web interface. The browser address bar displays the URL: `http://pubcrawler.gen.tcd.ie/cgi-bin/pubcrawler_www.pl`. The page title is "PubCrawler - an Update Alerting Service for PubMed and GenBank". The main heading is "Query configuration for user 'hectorserrano'".

Existing queries:

- 1 **pubmed** 'phytoestrogens' Phytoestrogen* [All Fields] OR phytoestrogen* [All Fields] OR coumestrol [All Fields] OR Estrogen receptor [All Fields] OR Genistein [All Fields] OR Zeaxaralnone
- 2 **pubmed** 'Steroid synthesis' Steroid hormone [All Fields] OR Aromatase [All Fields] OR Androgen* [All Fields] OR Progest* [All Fields]
- 3 **pubmed** 'Placenta' Placenta [All Fields] OR Steroidogenesis [All Fields] OR aromatase [All Fields] OR STAR [All Fields]

Action for selected query: [Dropdown] [Do it]

New PubMed query:

Alias: [Query Alias] (replace with descriptive term)

| search term | search field | connector |
|--------------|--------------|------------------|
| [Text Input] | [Dropdown] | [AND] [Dropdown] |
| [Text Input] | [Dropdown] | [AND] [Dropdown] |
| [Text Input] | [Dropdown] | [AND] [Dropdown] |

[Save] [Reset]

Type: PubMed PubMed Neighbour
 Nucleotide Nucleotide Neighbour

Terms: [4] [Change]

[Finish]

Automatically generated on Fri Aug 16 01:56:29 2013 (EIRE)
 For question or remarks, please e-mail pubcrawlerhelp@gmail.com.

Figura 2.2a. Ventana para especificar las revisiones y cambios.

The screenshot shows the PubCrawler Results page. The browser address bar displays the URL: `http://pubcrawler.gen.tcd.ie/cgi-bin/pubcrawler_www.pl`. The page title is "PubCrawler Results for hectorserrano". The date is "Thursday 15 Aug 2013".

Index of PubCrawler results:

- phytoestrogens: 47 new hits today
- Steroid synthesis: 79 new hits today
- Placenta: 57 new hits today

[retrieval] [query box] [disclaimer and copyright]

Results for 'phytoestrogens' at PubMed

[next] [top]

800 hits after first visit for 'Phytoestrogen* [All Fields] OR phytoestrogen* [All Fields] OR coumestrol [All Fields] OR Estrogen receptor [All Fields] OR Genistein [All Fields] OR Zeaxaralnone'

Warning: retrieved max number of items for this query!

→ Increase value of `getmax` (currently 800) via command-line option or in configuration file for additional (older) results.

Today's new results (47 items in total, sorted by Entrez date):

- 1: **Brinton LA, Westhoff CL, Scoccia B, Lamb EJ, Trabert B, Niwa S, Moghisi KS.** Related Articles
Fertily drugs and endometrial cancer risk: results from an extended follow-up of a large infertility cohort.
Hum Reprod. 2013 Aug 13;. [Epub ahead of print]
PMID: 23943785 [PubMed - as supplied by publisher]
- 2: **Davis PJ, Mousa SA, Cody V, Tang HY, Lin HY.** Related Articles
Small Molecule Hormone or Hormone-Like Ligands of Integrin IαVβ3: Implications for Cancer Cell Behavior.
Horm Cancer. 2013 Aug 14;. [Epub ahead of print]
PMID: 23943159 [PubMed - as supplied by publisher]
- 3: **Astiz M, Acas-Fonseca E, Garcia-Segura LM.** Related Articles
Sex Differences and Effects of Estrogenic Compounds on the Expression of Inflammatory Molecules by Astrocytes Exposed to the Insecticide Dimethoate.
Neurotox Res. 2013 Aug 14;. [Epub ahead of print]
PMID: 23943137 [PubMed - as supplied by publisher]

Figura 2.2b. Resultados acumulados por campo requisitado.

El sitio le permite analizar un máximo de 800 citas que pueden ser modificadas en cantidad, forma (puede almacenar los resúmenes de las diferentes citas recibidas, el resumen en formato gráfico, si está disponible, etc). Cada 10 días, se remueven las 100 citas más antiguas de cada uno de los campos de búsqueda requeridos. Sin embargo y con algunas limitantes, puede tener acceso a ellos utilizando la herramienta "PREVIOUS SEARCHES" que aparece una vez que ha accedido a los listados de resultados.

Actividades para reportar

A partir del portal de la Biblioteca o de PubMed, permita que al menos 10 revistas le envíen su tabla de contenidos. Dentro de ellas asegúrese de incluir Nature (varias de ellas ahora están especializadas), Science y Nucleic Acids Research y para una lectura ligera, Scientific American y La Recherche.

Haga una tabla en la cual muestre el número de artículos sobre el tema que seleccionó y la frecuencia con la cual obtuvo actualizaciones a lo largo del trimestre. Esta actividad será evaluada y discutida en el grupo.

Referencias

-  Hokamp, K., Wolfe, K.H. 2004. PubCrawler: keeping up comfortably with PubMed and GenBank. Nucl. Acids Res. 32:W16-W19. Disponible en la URL http://nar.oxfordjournals.org/content/32/suppl_2/W16.full.pdf+html
-  Ma'ayan, A. 2008. PubMedAlertMe- standalone Windows-based PubMed SDI software application. Comput. Biol. Med. 38:620-622. Disponible en la URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2431148/pdf/nihms51645.pdf>

Búsqueda y análisis de bases de datos

Capítulo 3

Obtención y análisis de secuencias de proteínas utilizando un servidor en línea

En los laboratorios de investigación en Genética Molecular, es necesario conocer los productos génicos con los que se está tratando, sobre todo si están asociados a patologías o características morfológicas que pueden determinarse de manera simple. Como se había planteado desde el inicio del curso, se utilizarán marcadores moleculares asociados a diferentes patologías. En una de ellas, el equilibrio hemodinámico proporcionado por la actividad de los receptores adrenérgicos se encuentra alterada.

Como producto del esfuerzo de la comunidad científica en conocer la secuencia de cada uno de los nucleótidos que conforman nuestro genoma, una gran cantidad de datos se han generado. Durante las últimas dos décadas, la información referente a las secuencias de nucleótidos han dominado el escenario científico. Sin embargo, durante una buena parte del siglo pasado, la función, regulación y constitución de las proteínas mantuvo ocupados a bioquímicos y fisiólogos de diversos campos.

Algunos de los mejores premios en Ciencia se otorgaron a grandes representantes del campo de estudio de las proteínas. Pocas personas han recibido no uno sino dos premios Nobel por el impacto e importancia de sus investigaciones. El primer premio Nobel de Linus Pauling sobre la estructura de una proteína fue seguido por el de Frederick Sanger sobre el desarrollo de una metodología para conocer la secuencia de aminoácidos de una proteína. Mientras Sanger obtuvo su segundo Nobel por su metodología de secuenciación de ADN, Pauling lo obtuvo por su activismo a favor de la paz.

Las publicaciones de las secuencias de aminoácidos crecieron mucho antes de que se tuviese un acervo de secuencias nucleotídicas importante. En ambos casos, la conformación de bases de datos conteniendo información estructural, de las secuencias de aminoácidos, de las funciones y variaciones de las proteínas de diferentes sistemas celulares ha permitido el avance de la investigación con frentes, metodologías y estrategias basadas en la esencia de la investigación básica.

Objetivo

En este ejercicio, haremos el análisis de los diferentes receptores adrenérgicos que se presentan en el humano y se encuentran depositadas en las bases de datos. El objetivo de este análisis es determinar los participantes proteicos de este tipo de proteínas, su alineamiento en cada aminoácido, las características topográficas locales y la posible relación filogenética existente entre ellas.

Desarrollo

En su buscador, ingrese la dirección http://web.expasy.org/docs/swiss-prot_guideline.html Esto desplegará la página inicial de SwissProt and TrEMBL. En el segundo párrafo, active la liga UniProt Website lo que abrirá la página de Uni-Prot en la pestaña SEARCH (Figura 3.1).

Introduzca en la ventana de diálogo la palabra "Human adrenergic receptor" y pulse el botón SEARCH. Esto abrirá una página en la cual se encuentran diferentes proteínas asociadas. Marque los registros de las proteínas correspondientes a los diferentes tipos de receptores adrenérgicos de humano. Al final de la revisión de las páginas consta la búsqueda, pulse el botón "RETRIEVE". En este momento, abrirá la pestaña "RETRIEVE" del servidor en la cual dentro de una ventana aparecen los identificadores de las proteínas seleccionadas (Figura 3.2).

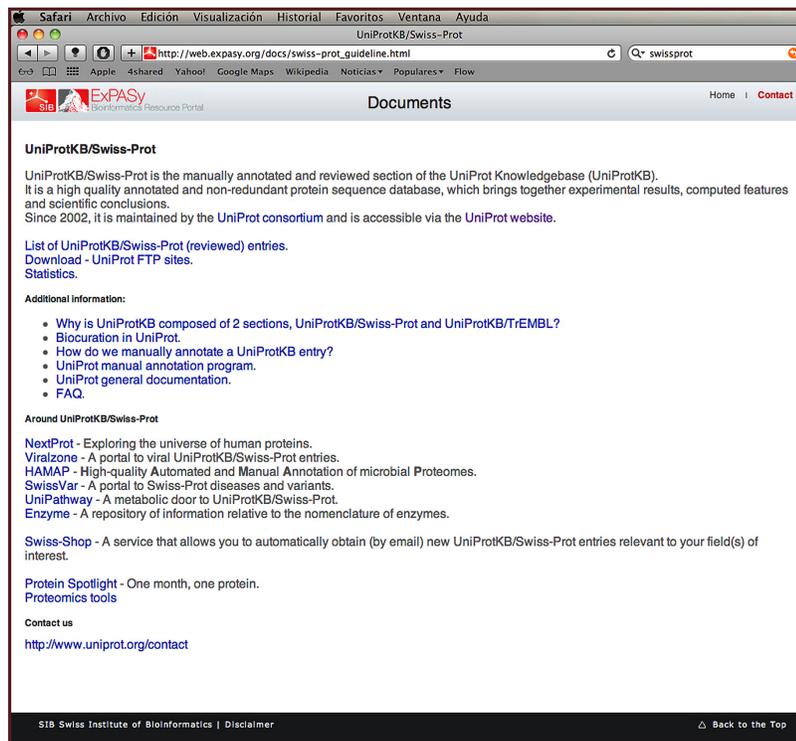


Figura 3.1a. Ventana del portal de inicio de Swissprot.



Figura 3.1b. Arranque del sitio UniProt.

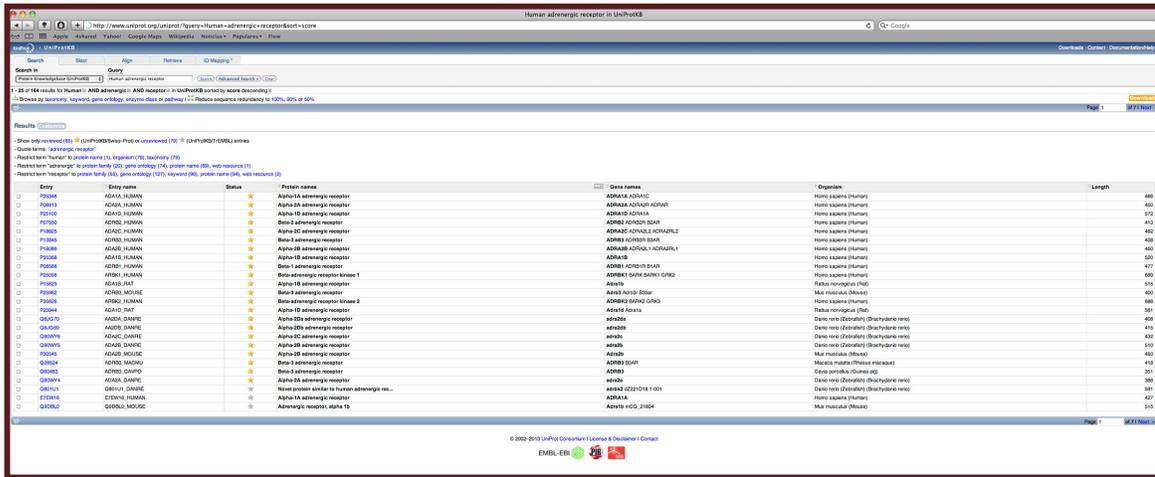


Figura 3.2a. Listado resultante de la búsqueda de receptores adrenérgicos.

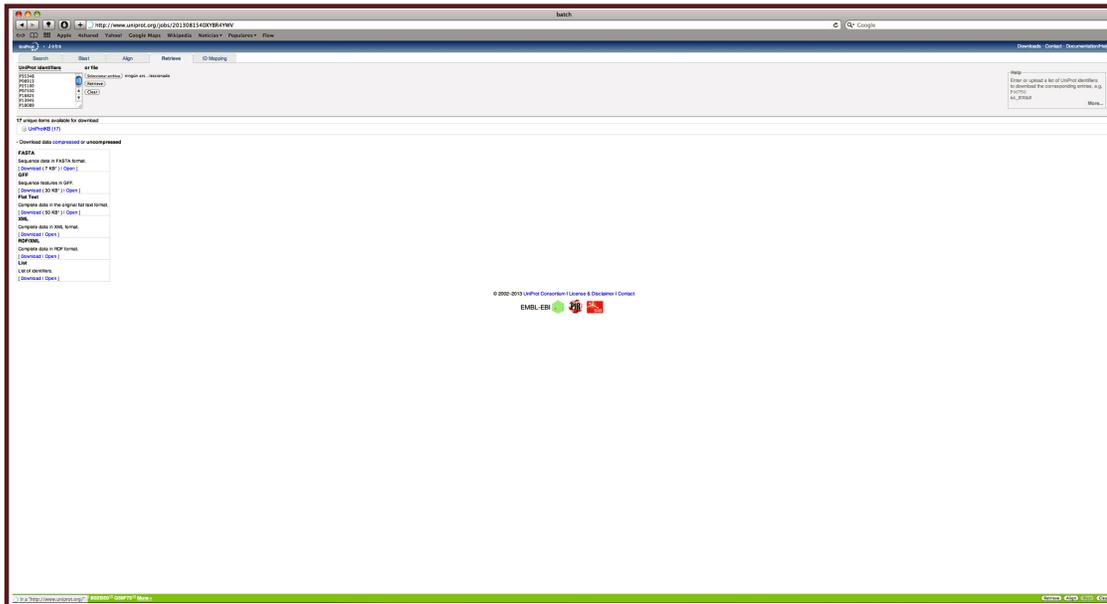


Figura 3.2b. Depuración para humanos de UniProt.

Pulse el botón "ALIGN" localizado en el extremo inferior derecho de la pantalla. Esto abrirá la pestaña correspondiente en el servidor. Si el número de selecciones es pequeño, se abrirá un cuadro que contiene el código de una letra de los aminoácidos que forman las proteínas (Figura 3.3).

Al término de la revisión, la pestaña correspondiente se activa y muestra el alineamiento que tienen las diferentes proteínas seleccionadas. El programa de alineamiento presenta un algoritmo que optimiza la "llamada" del aminoácido y, cuando no existe un alineamiento adecuado, inserta un guión. En este punto es importante hacer notar las herramientas de análisis disponibles en la zona derecha del inicio de la secuencia alineada. En esta herramienta es posible localizar los diferentes tipos de aminoácidos o la presencia de modificaciones y dominios. En ejemplo mostrado en la figura 3.3, se ha pedido la localización de los dominios transmembranales que se ha resaltado en amarillo.

Si activa las ventanas de las características de los aminoácidos (polares, pequeños, hidrofóbicos, etc) se irán marcando en diferentes colores los aminoácidos que cumplan con los requisitos. Todo esto permitirá hacer el análisis de las diferentes regiones que conforman las proteínas y que podrían estar asociadas con la función.

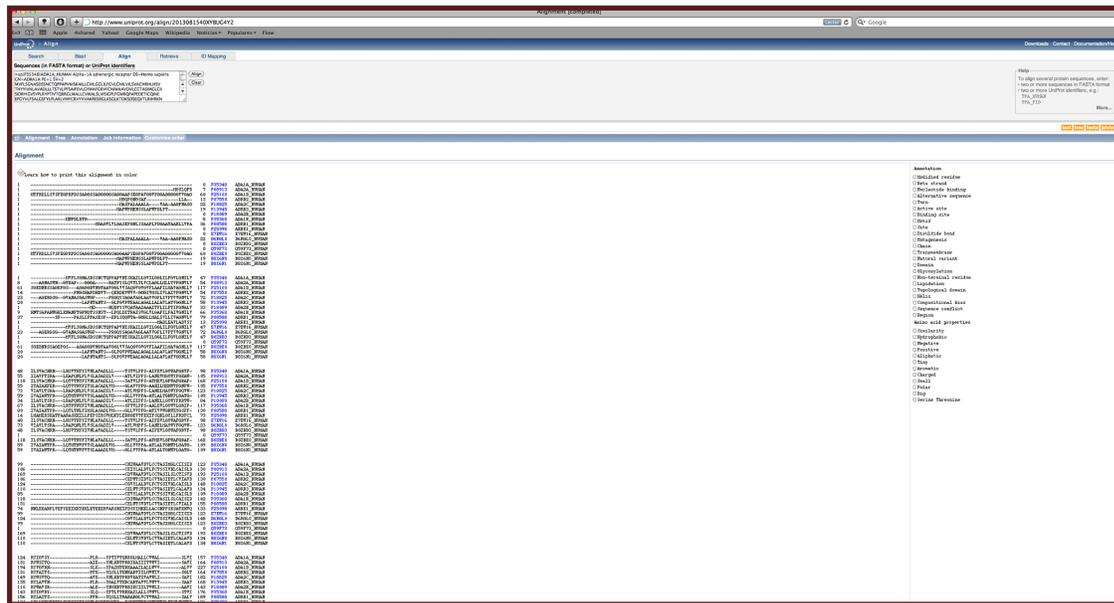


Figura 3.3a. Ventana de alineamiento.

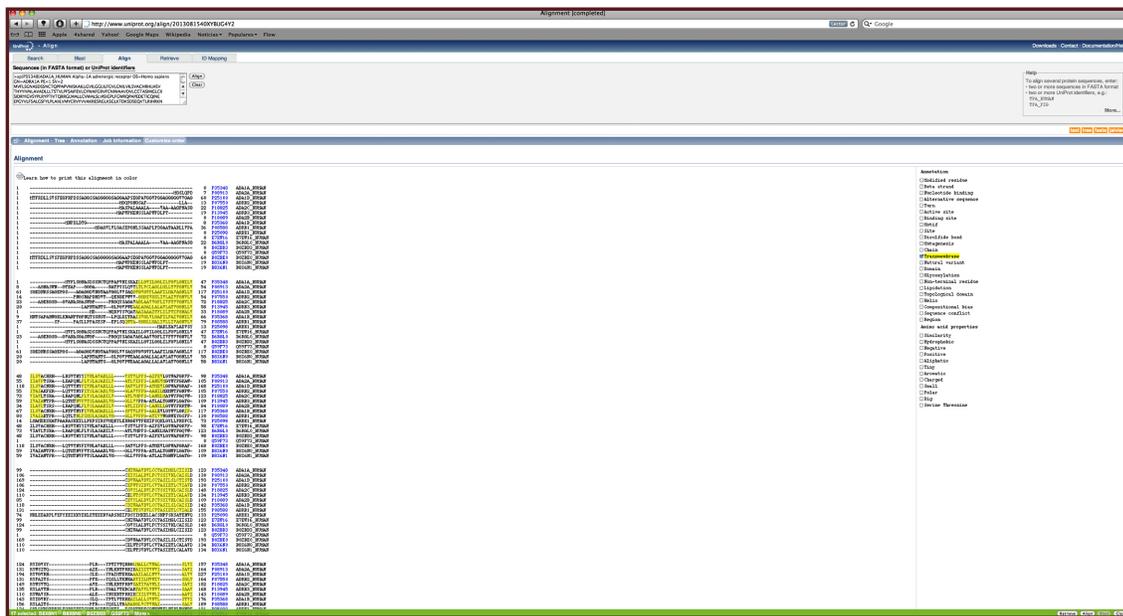


Figura 3.3b. Dominios transmembranales del extremo amino terminal de los receptores adrenérgicos humanos.

Actividades para reportar

Utilizando el alineamiento del receptor adrenérgico humano, identifique los residuos hidrofóbicos, dominios polares, variantes y sitios de glicosilación. Identifique los residuos de aminoácidos polares pequeños, de lipidización y los dominios funcionales que pueda presentar el receptor.

Haga un análisis similar utilizando por separado los diferentes tipos de receptores adrenérgicos resaltando las posiciones donde tenga homología.

Repita el análisis utilizando las secuencias de Calpaína 10 humana y los receptores estrogénicos humanos.

Referencias

-  O'Donovan C, Martin MJ, Gattiker A, Gasteiger E, Bairoch A, Apweiler R. 2002. High-quality protein knowledge resource: Swiss-Prot and TrEMBL. Brief. Bioinf. 3:275-284. Disponible en la URL <http://bib.oxfordjournals.org/content/3/3/275.full.pdf+html>
-  Uniprot Consortium. 2012. Reorganizing the protein space at Universal Protein Resource (UniProt). Nucleic Acid Res. 40:D71-D75. Disponible en la URL <http://nar.oxfordjournals.org/content/40/D1/D71.full.pdf+html>

Actividad 2

Analisis de homología y evolucion de grupos proteicos

Además de comparar e ingresar nuevas secuencias, la labor de los investigadores requiere el conocimiento de las alteraciones que puedan presentarse en una proteína o enzima que afecta una función particular de la célula o tejido que se está estudiando. El estudio detallado de este tipo de alteraciones y que se realizó durante una buena parte del siglo XX, se ha facilitado por el empleo de las metodologías de ADN recombinante y el posterior desarrollo de algoritmos que permiten su alineamiento, como lo vimos en la actividad anterior.

Es posible seguir el camino evolutivo o divergente de una proteína o familia proteica utilizando la información contenida en su propia constitución. El marco teórico en el cual se encuadra la presunción básica es que si cada una de las proteínas realiza una función específica en un organismo determinado, en el origen de ella, el proceso de mutación que afecta al gen que la codifica podría haber afectado una región no importante para la función de la proteína. Si al mismo tiempo el gen codificante se duplicó, el efecto de estas mutaciones y la selección natural en cada uno de los genes sería diferente: mientras uno fuese fuertemente seleccionado para sufrir cambios en lugares diferentes al sitio activo o de importancia para la función proteica, el otro tendría una presión de selección menor. La acumulación de los cambios en cada uno de los genes a lo largo de la historia evolutiva permitiría que si bien se mantienen características similares, existan diferencias que permitan no solo distinguir los genes que codifican a estas dos proteínas, sino de las mismas proteínas que permitirían identificarlas como distintas.

Uno de estos escenarios podría implicar, por ejemplo, que una mutación interna en la región codificante de un exón ensamblara un codón de término. La proteína producto de este gen sería más corta que la proteína homóloga. Si el residuo de la proteína alterada implica, por ejemplo una región de regulación alostérica, esta versión de la proteína no respondería a este tipo de regulación, por lo que la función sería diferente o se llevaría a cabo en un lugar diferente.

Mediante el uso de las herramientas desarrolladas a partir de las bases de datos, es posible hacer una estimación y simulación de este tipo de eventos.

Objetivo

El objetivo de esta actividad es mostrar la forma de analizar las relaciones entre diferentes miembros de las familias proteicas cuyas secuencias se encuentran depositadas en los bancos de secuencias como UniProt.

Desarrollo

En la página inicial de SwissProt (<http://www.uniprot.org>) dirija hacia la página de búsqueda (SEARCH del punto 3 de la actividad anterior) introduzca "Human Adrenergic Receptor" en el cuadro de diálogo; despliegue el cuadro de menú de la base de datos localizada a la izquierda del cuadro de diálogo y seleccione "SEQUENCE CLUSTERS (UniRef)" como base de datos para la búsqueda (Figura 3.4a). Pulse el botón SEARCH para desplegar la hoja de resultados en la pestaña SEARCH.

The screenshot shows the UniProt search interface. The search term is 'Alpha-2B adre'. The results are filtered to show 1 member from the UniProt dataset. The 'Members' section is expanded, showing a list of 236 members. The 'Primates' category is selected in the filter, and the list of members is updated to show only primates, including species like *Otolemur garnettii*, *Simiiformes*, *Catarrhini*, *Cercopithecinae*, *Macaca*, *Macaca mulatta*, *Hominidae*, *Hyllobates lar*, *Pongo abelii*, *Gorilla gorilla gorilla*, *Homo sapiens*, *Platyrrhini*, *Callithrix jacchus*, and *Ateles*.

Figura 3.4c. Receptores adrenérgicos y la selección del filtro para primates.

En esta pestaña, se encontrarán los grupos de proteínas agrupadas en relación a la identidad de los residuos de aminoácidos y longitud de la homología (Figura 3.4 b). Active la liga que aparece como “UniRef50_ P18089” que abrirá la página de resultados parciales. En el listado del filtro, seleccione “PRIMATES” y active el botón “FILTER” para mostrar solamente los resultados de los diferentes primates con que cuenta la base de datos. (Figura 3.4 c)

Cuando aparezca la página de resultados, active el botón “ALIGN” para obtener los resultados con el algoritmo CLUSTALW para análisis de grupos (cluster analysis). Este algoritmo evalúa el grado de identidad, homología y regiones de cambio. Para ver los detalles de las distancias euclidianas entre los diferentes nodos, utilice la herramienta “TREE” que se encuentra al inicio de la zona de alineamiento (Figura 3.5). Coloree los diferentes tipos de aminoácidos y obtenga las relaciones filogenéticas y distancias entre las diferentes proteínas.

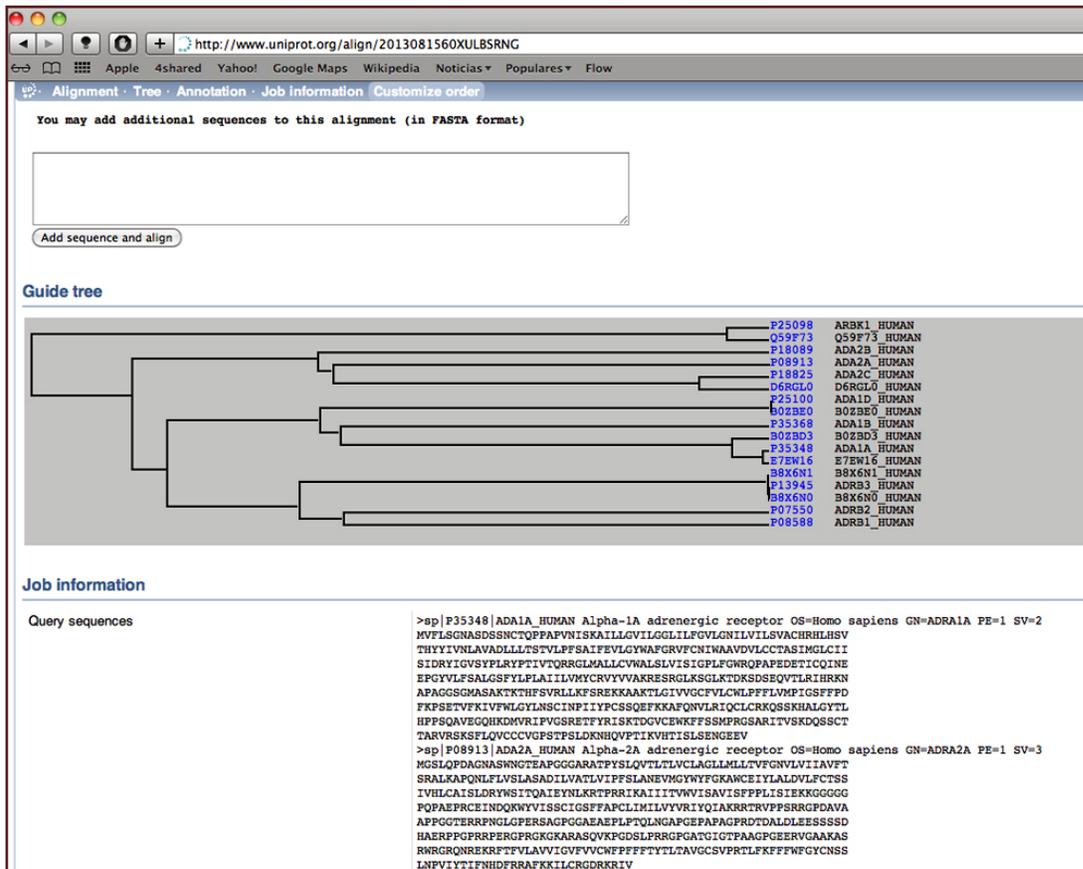


Figura 3.5a. Ventanas de relación filogenética euclidean.

Referencias

-  Pearson WR. 1997. Identifying distantly related protein sequences. *Comput. Appl. Biosci.* 13:325-332. Disponible en la URL <http://bioinformatics.oxfordjournals.org/contents/13/4/325.full.pdf+html>
-  ThompsonJD, Higgins DG, Gibson TJ. 1994. CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weigh matrix choice. *Nucleic Acid Res.* 22:4673-4680. Disponible en la URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC308517/pdf/nar00046-0131.pdf>

Capítulo 4

Simulación de la estructura de genes a partir de la secuencia de aminoácidos

En la mayoría de los casos, la obtención de un gene proviene de la secuencia de aminoácidos de un fragmento de la proteína que nos interesa. Actualmente, el uso de las herramientas bioinformáticas nos permiten no sólo hacer la búsqueda de la proteína que nos interesa, sino también determinar el grado de conservación que presenta en diferentes organismos, como se hizo en el ejercicio de la sesión 3.

Partiendo del hecho que se conoce la proteína que se pretende estudiar, el análisis experimental puede hacerse utilizando una estrategia definida. Si quisiéramos hacer un mapeo fino de la proteína para determinar los parámetros básicos de su función, por ejemplo debida a la presencia de un dominio de aminoácidos polares que le den un dominio hidrofílico con una estructura tridimensional específica para que se exponga un residuo lateral de serina del cual depende la función, una estrategia adecuada sería clonar específicamente o de manera restringida la región en cuestión y expresarla en cantidades abundantes o de manera constante para realizar los experimentos en cantidad y número necesario para obtener resultados confiables.

Usualmente, la clonación de genes o fragmentos de genes requiere el uso de sitios de restricción específicos tanto en el vehículo molecular (plásmido, bacteriófago, cósmido, etc.) por lo que es necesario conocer el sitio potencial de clonación así como aquellas enzimas de restricción que pueden ser utilizadas en el análisis o mapeo fino del gene y así diseñar alelos modificados (mutantes), que carezcan de una región (con deleciones) o que contengan un codón de paro al inicio que evite ser traducido.

Objetivo

Que el alumno pueda desarrollar una estrategia de investigación de un gen a partir del conocimiento de su secuencia de aminoácidos

Desarrollo

Obtenga la secuencia de la proteína aromatasas de la oveja (Cytochrome p450 19A1) utilizando el código de acceso NP_001116472 en la base de datos PROTEIN del NCBI (<http://www.ncbi.nlm.nih.gov/protein> Fig 4.1) o Q9XS28 de UniProt (<http://www.uniprot.org> Fig 4.2).

The screenshot shows the NCBI Protein database homepage. At the top, there is a search bar with the text "Protein" and a "Search" button. Below the search bar, there is a large banner with the text "Protein" and a description: "The Protein database is a collection of sequences from several sources, including translations from annotated coding regions in GenBank, RefSeq and TPA, as well as records from SwissProt, PIR, PRF, and PDB. Protein sequences are the fundamental determinants of biological structure and function." Below the banner, there are three columns of links: "Using Protein" (Quick Start Guide, FAQ, Help, GenBank FTP, RefSeq FTP), "Protein Tools" (BLAST, LinkOut, E-Utilities, BLink, Batch Entrez), and "Other Resources" (GenBank Home, RefSeq Home, CDD, Structure). At the bottom, there is a "GETTING STARTED" section with various links for education, manuals, and training, and a "POPULAR" section with links to various biological databases and tools.

Figura 4.1a. Ventana de acceso a Protein del NCBI.

The screenshot shows the NCBI Protein database entry for cytochrome P450 19A1 [Ovis aries]. The entry is titled "cytochrome P450 19A1 [Ovis aries]" and includes the NCBI Reference Sequence: NP_001116472.1. The entry is displayed in FASTA format. The entry includes the following information:

- LOCUS:** NP_001116472 503 aa linear MAM 18-APR-2013
- DEFINITION:** cytochrome P450 19A1 [Ovis aries].
- ACCESSION:** NP_001116472
- VERSION:** NP_001116472.1 GI:172088165
- DBSOURCE:** REFSEQ; accession NM_001123000.1
- KEYWORDS:** RefSeq.
- SOURCE:** Ovis aries (sheep)
- ORGANISM:** Ovis aries
- REFERENCE:** 1 (residues 1 to 503) Roselli, C.E. and Stormshak, F. The ovine sexually dimorphic nucleus, aromatase, and sexual partner preferences in sheep. *J. Steroid Biochem. Mol. Biol.* 118 (4-5), 252-256 (2010)
- REFERENCE:** 2 (residues 1 to 503) Mura, A., Gadau, S., Lepore, G., Balzano, F., Zedda, M., Mura, E. and Farina, V. Expression and distribution of P450-aromatase in the ovine hypothalamus at different stages of fetal development. *Neuro Endocrinol. Lett.* 31 (5), 690-699 (2010)
- REFERENCE:** 3 (residues 1 to 503) Lepore, G., Gadau, S., Mura, A., Zedda, M. and Farina, V. Aromatase immunoreactivity in fetal ovine neuronal cell cultures exposed to oxidative injury. *Eur J Histochem* 53 (4), E28 (2009)
- REFERENCE:** 4 (residues 1 to 503) Vanselow, J., Selimyan, R. and Furbass, R. DNA methylation of placenta-specific Cyp19 promoters of cattle and sheep

The entry also includes a "References" section with links to "Articles about the CYP19 gene", "Identical proteins for NP_001116472.1", and "Reference sequence information".

Figura 4.1b. Desplegado de la aromatasa de oveja.



Figura 4.2a. Ventana de acceso a UniProt.

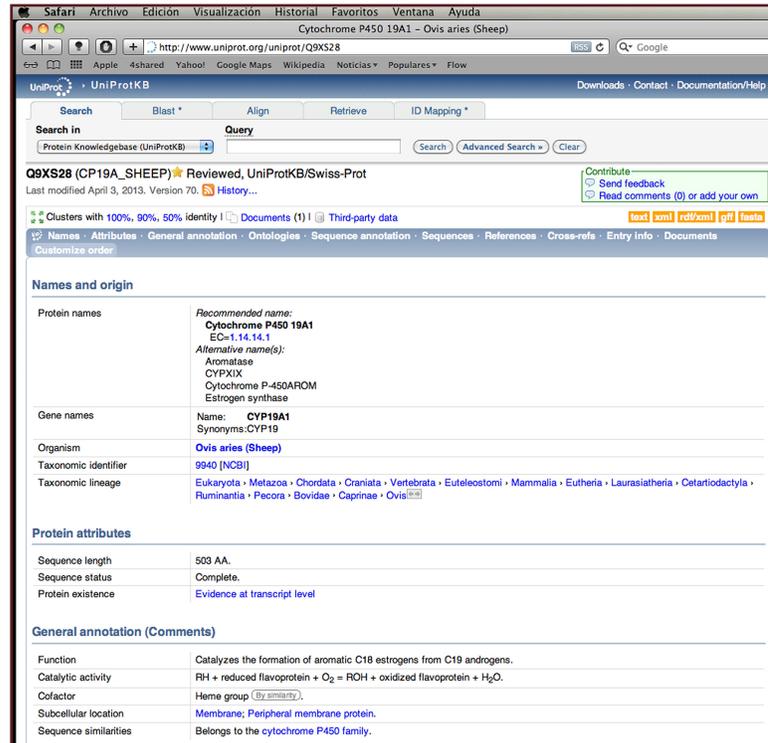


Figura 4.2b. Desplegado de la aromatasa de oveja.

En cualesquiera de los casos que hubiese seleccionado, active la liga para obtener la secuencia en formato FASTA (Figura 4.3) Seleccione la secuencia de aminoácidos de la proteína y cópiela en un archivo de Word o wordpad. El archivo que obtenga de UniProt requiere que elimine los números de los aminoácidos y los espacios que separan los grupos.

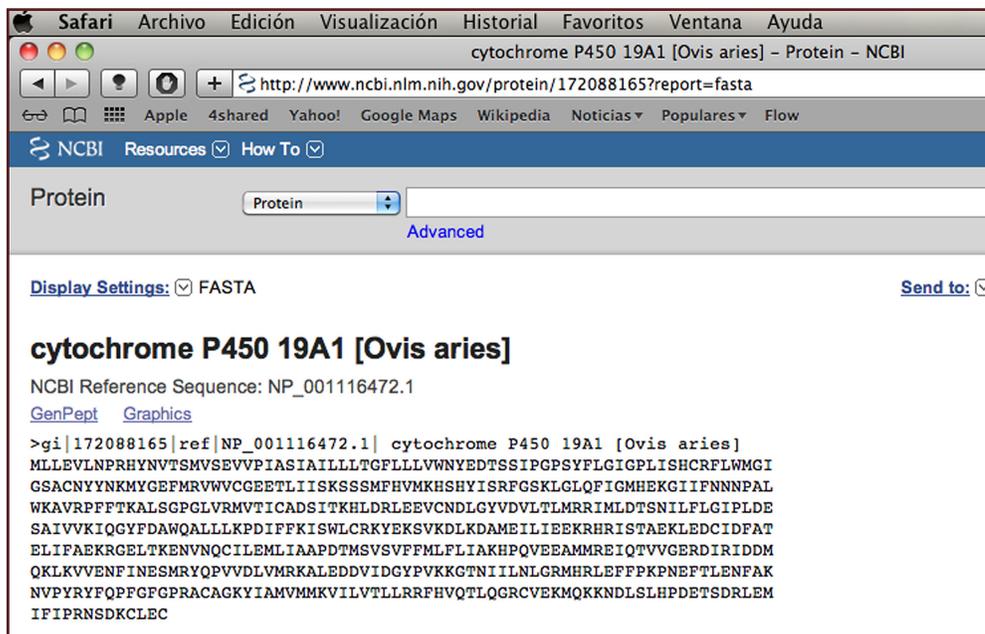


Figura 4.3a. Ventana que muestra la secuencia en lenguaje FASTA de la base PROTEIN NCBI.

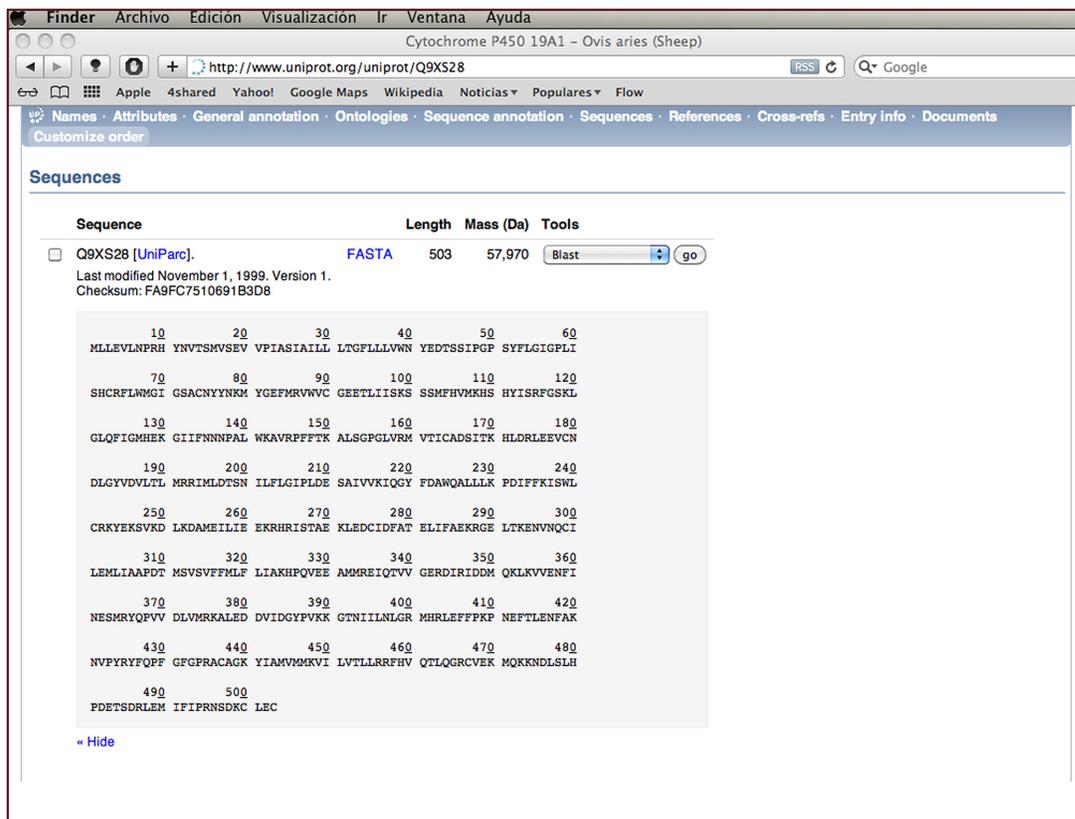


Figura 4.3b. UniProt.

Abra el servidor del sitio GeneDesign en la URL <http://54.235.254.95/gd/> que lo llevará al menú de operaciones del sitio. Seleccione la opción "REVERSE TRANSLATION" que abre la ventana correspondiente (Figura 4.4).

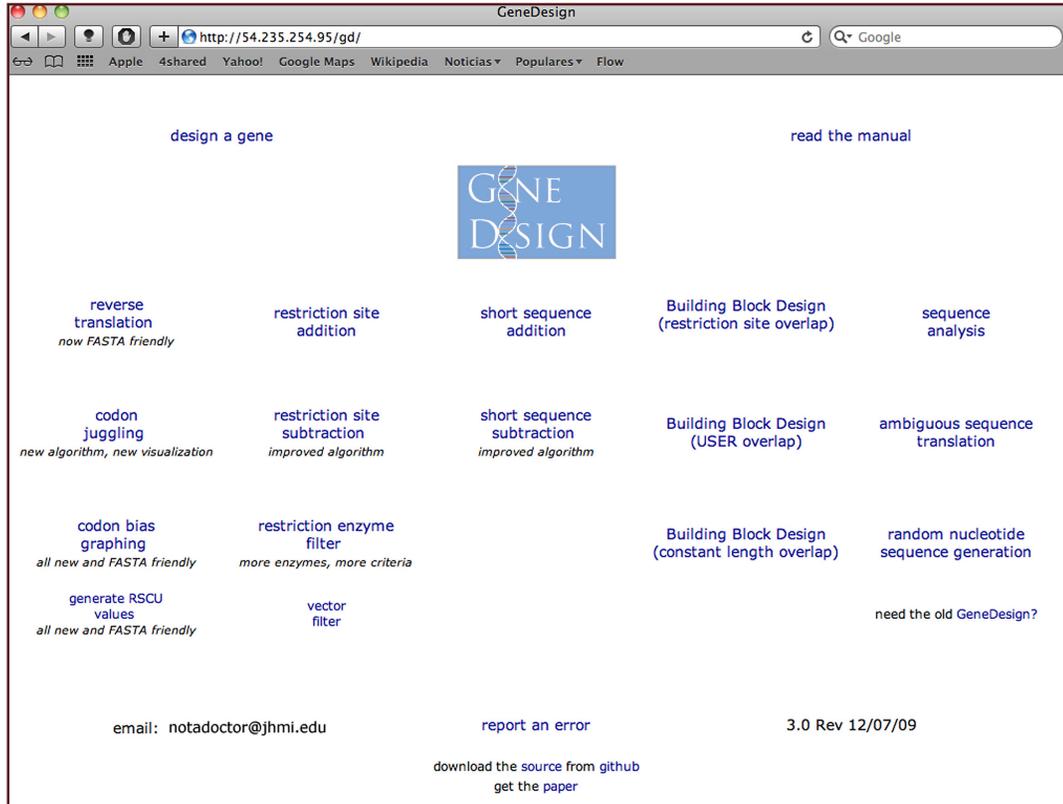


Figura 4.4a. Ventana de ingreso a Gene Design.

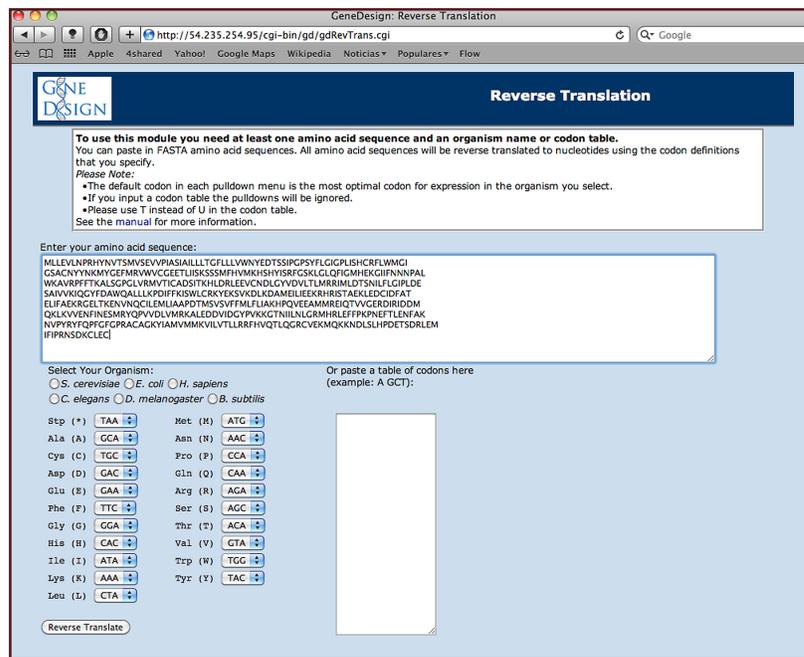
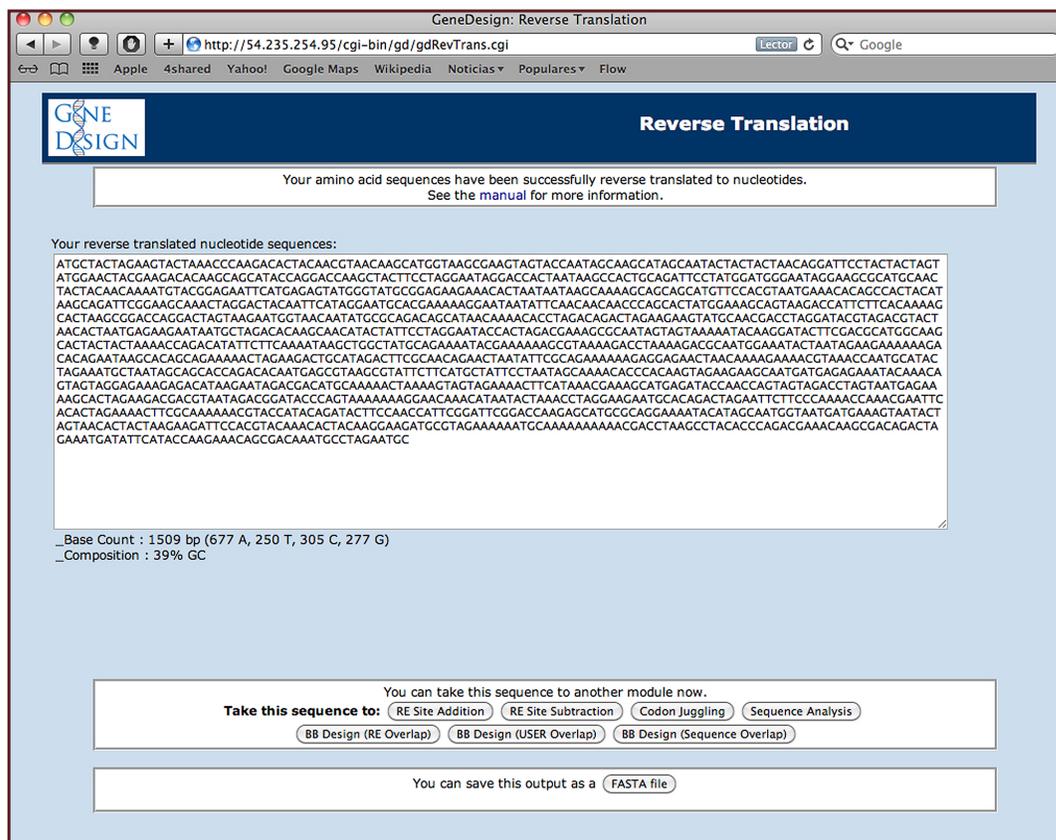


Figura 4.4b. Herramienta REVERSE TRANSLATION con la secuencia FASTA de la aromatasa ovina.

En esta nueva pestaña, se encuentra un encabezado con inscripciones simples y una ventana de diálogo en el cual debe pegar el archivo FASTA que hizo en Word o wordpad. En este sitio, puede seleccionar las características de los codones preferenciales que son utilizados en algunos de los modelos celulares más utilizados sin embargo, tiene la capacidad de utilizar los de la especie con la que se está trabajando. Para una referencia sobre los codones más utilizados en diferentes especies, puede consultar el artículo Sharp. PM, Cowe E, Higgins DG, Shields DC, Wolfe KH. Wright F. Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity. Nucl. Acids Res. 16:8207-8211. 1988. Disponible en la URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC338553/pdf/nar00159-0016.pdf>

Seleccione "H. sapiens" para el análisis. Pulse el botón "REVERSE TRANSLATE" lo que traerá una ventana en la cual se tiene la secuencia de nucleótidos que tendría el mRNA que codificaría para la proteína. Guarde esta secuencia en un archivo con formato FASTA.

Regrese a la página de resultados y pulse el botón "SEQUENCE ANALYSIS" lo que traerá una ventana en la cual está en una ventana simple la secuencia del mRNA que obtuvo, dos mapas verticales en el cual se indican los sitios de restricción que están presentes en el gene que codifica a este mRNA y uno más en el cual se encuentran las posibles pautas de lectura que pueden codificar a la proteína (ORF del inglés Open Reading Frame). En el espacio lateral se encuentra un listado de los sitios de restricción que no se encuentran en las ORF del gen que codificaría para la proteína (Fig 4.5). Guarde toda esta información en un documento que pueda visualizar posteriormente.



The screenshot shows the GeneDesign Reverse Translation web interface. At the top, there is a navigation bar with the GeneDesign logo and the title "Reverse Translation". Below this, a message states: "Your amino acid sequences have been successfully reverse translated to nucleotides. See the manual for more information." The main content area displays the "Your reverse translated nucleotide sequences:" followed by a long, multi-line nucleotide sequence. Below the sequence, the base count is shown as "_Base Count : 1509 bp (677 A, 250 T, 305 C, 277 G)" and the composition as "_Composition : 39% GC". At the bottom, there is a section titled "You can take this sequence to another module now." with several buttons: "RE Site Addition", "RE Site Subtraction", "Codon Juggling", "Sequence Analysis", "BB Design (RE Overlap)", "BB Design (USER Overlap)", and "BB Design (Sequence Overlap)". Below this, another section says "You can save this output as a" followed by a "FASTA file" button.

Figura 4.5a. Secuencia de nucleótidos.

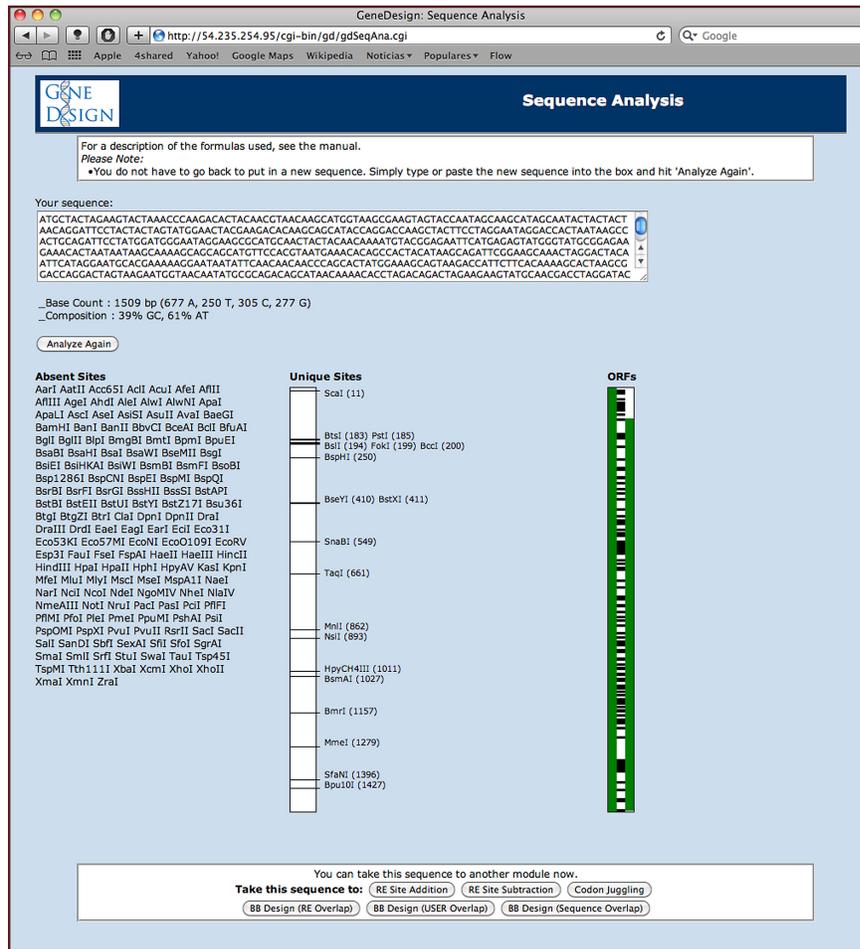


Figura 4.5b. Mapa de restricción obtenidos después de ingresar la secuencia de aminoácidos de la aromatasa de ovinos.

Con la secuencia de nucleótidos en FASTA que guardó, compárela con la que se encuentra depositada en el GenBank. Abra la página de inicio del NCBI y espere a que aparezca la página de inicio. En el campo de las herramientas disponibles, active la liga que lo lleve a la herramienta BLAST (Basic Local Alignment Search Tool). Esta activación lo llevará al menú de inicio de la herramienta. Identifique la porción correspondiente a “nucleotide blast” dentro de la sección “Basic BLAST” y active la liga.

La nueva ventana tiene una ventanilla donde puede pegar la secuencia en FASTA que obtuvo con GeneDesign. Asegúrese de que no se tienen espacios en la secuencia pues de otra forma el programa lo reconoce como una proteína independiente que debe alinear y provocar que se utilice un bucle del algoritmo que reporta un mensaje de error.

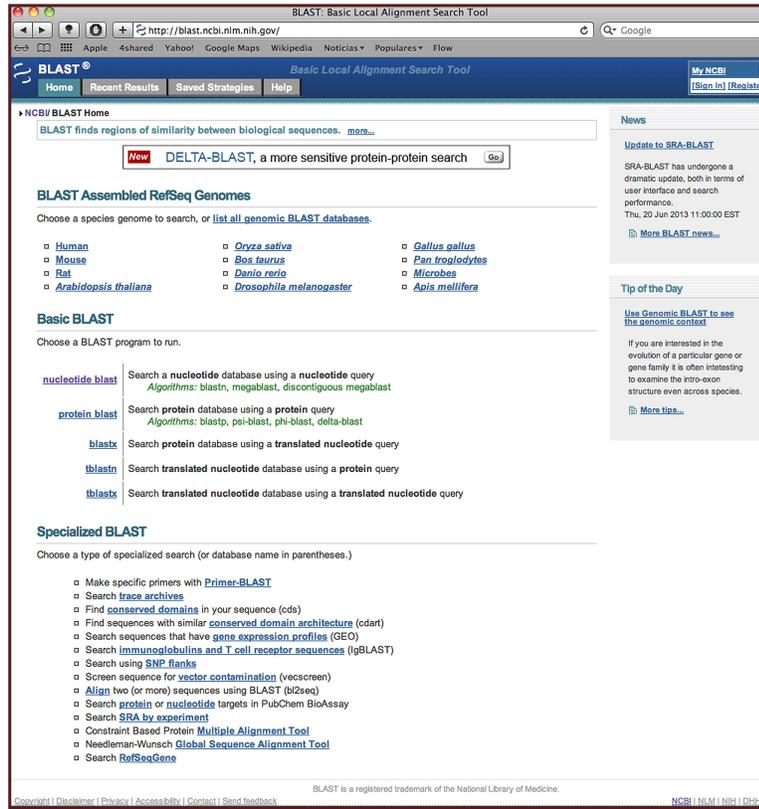


Figura 4.6a. Página de inicio de la herramienta BLAST.

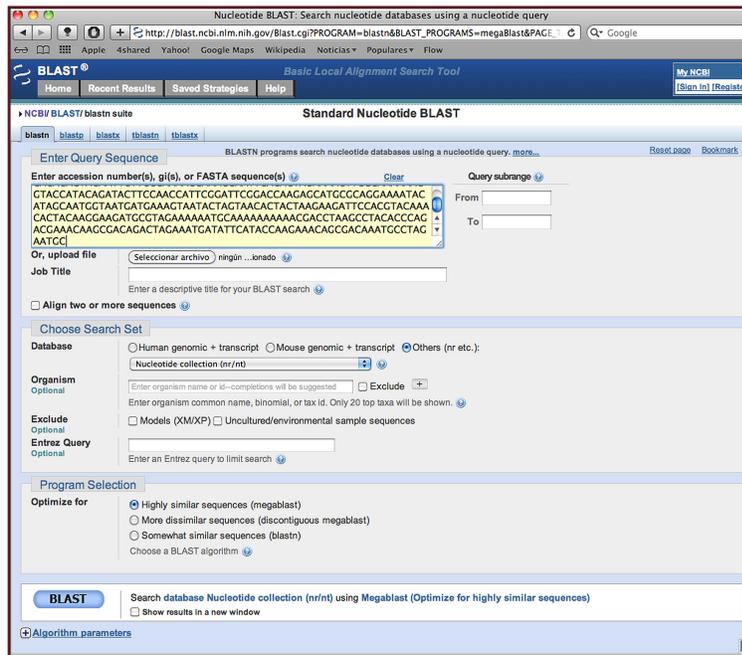


Figura 4.6b. Consulta en donde se incorpora la secuencia FASTA obtenida en GeneDesign.

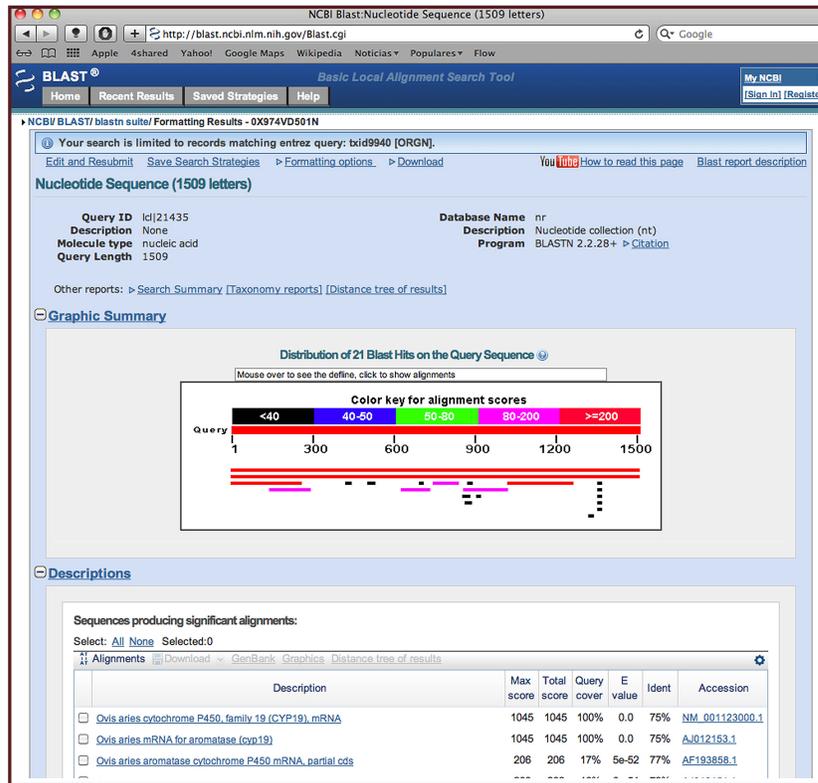


Figura 4.6c. Alineamiento de las secuencias y grado de homología.

Una vez terminada la depuración, indique que se trata de una secuencia para oveja (Sheep) en el cuadro de diálogo correspondiente. Marque el cuadro que indica que quiere los resultados en una nueva ventana y pulse el botón BLAST (Fig 4.5). En la nueva ventana aparecerá un mapa a color de los resultados del alineamiento y los genes o porciones de genes con las que se tiene alineamiento, junto con parámetros importantes para hacer la evaluación. Copie a un archivo los resultados para tener un registro permanente.

Actividades para reportar

Obtenga las secuencias de las proteínas de los receptores adrenérgicos humanos. Clasifíquelos y haga un análisis similar al realizado en las actividades descritas. Identifique las diferencias primordiales que se obtendría si se hiciera una estrategia de clonación en un bacteriófago del tipo λ gt11 en un sitio diferente al reconocido por *EcoRI*.

En un análisis redundante, obtenga la secuencia de nucleótidos del mRNA de la calpaína 10 y replique la actividad realizada. Discuta las diferencias que encontró en el proceso.

Referencias

-  Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T.L., Agarwala, R., Schäffer, A.A. 2008. Database Indexing for Production MegaBLAST Searches, *Bioinformatics* 24:1757-1764. Disponible en la URL <http://bioinformatics.oxfordjournals.org/content/24/16/1757.full.pdf+html>
-  Richardson SM, Wheelan SJ, Yarrington RM, Boeke, JD. 2006. GeneDesign: Rapid, automated design of multikilobase synthetic genes. *Genome Res.* 16:550-556. Disponible en la URL <http://genome.cshlp.org/content/16/4/550.full.pdf+html>
-  Zhang, Z., Schwartz, S., Wagner L., Miller W. 2000. A greedy algorithm for aligning DNA sequences", *J Comput Biol* 7:203-14. Disponible en la URL <http://online.liebertpub.com/doi/pdf/10.1089/10665270050081478>

Aplicación de datos para el diseño de experimentos de genotipificación

Capítulo 5

Búsqueda de información general de enfermedades genéticas

Una de las metas de la investigación en Genética Humana es encontrar la razón fundamental de las enfermedades o la predisposición a desarrollarlas. Los estudios y esfuerzos continuos de Victor McKusick utilizado las metodologías básicas de la citogenética, inició el acopio de las relaciones entre las alteraciones cromosómicas y el desarrollo de enfermedades (o síndromes) y que se asociaban a una herencia mendeliana clásica. McKusick plasmó estas asociaciones en su libro "Mendelian Inheritance of Man" y a quien se refería en forma coloquial como "EL MAPA MORBIDO DEL HUMANO" que concretó primeramente en un texto de 12 volúmenes.

El inicio de la aplicación de las metodologías de manipulación y modificación de los ácidos nucleicos que actualmente confundimos con "Ingeniería Genética" o aún más con "Biología Molecular" permitió que se encontrara el sustrato fino de algunas enfermedades cuya atención médica era imperante. Así pues, la clonación del gene normal que se encuentra alterado en la Fibrosis Quística y su función como transportador (CFTR) permitió visualizar el potencial que tenía esta aplicación metodológica y que superaba pero no reemplazaba a su predecesora.

Con la implementación del proyecto HUGO, el desarrollo de las bases de datos, los adelantos en la informática, los algoritmos de correlaciones entre genes expresados (EST), porciones transcritas (ORF), mapeo de genes por cercanía (CONTIGS) y variaciones de un solo nucleótido (SNP) han permitido que el asentamiento físico de los genes cuya expresión anómala pueda ser asociado a una región y secuencia específica de genes.

La unidad encargada de mantener actualizada estas relaciones se encuentra en la Universidad John Hopkins, última institución a la que perteneció el Dr. McKusick. El proyecto forma parte del acervo en línea que maneja el NCBI. Al estar disponible en línea, la base de datos permite no sólo hacer una revisión de las regiones que han sido asociadas al desarrollo de una enfermedad, sino que también tiene una referencia cruzada con la base de datos en la cual se reportan los diferentes marcadores moleculares asociados, principalmente del tipo SNP y loci microsatelitales.

Además de las anotaciones propias, OMIM (Online Mendelian Inheritance of Man), se puede tener acceso a otras bases de datos, algunas de acceso abierto como las asociadas a las mutaciones del genoma mitocondrial (MITOMAP) mientras que otras requieren de un registro, aunque son libres de pago (Human Gene Mutation Database, HGMD).

Objetivo

Manejar una base de datos especializada en información médica. Utilizar la información relevante para diseñar una estrategia experimental que permita aportar conocimiento de punta al Área de la Salud.

Desarrollo

Adquiera el acceso a la página principal del NCBI (<http://www.ncbi.nlm.nih.gov>). Despliegue el menú de la ventana de selección de la base de datos. Seleccione OMIM. Opcionalmente, puede acceder ingresando directamente la dirección (<http://www.ncbi.nlm.nih.gov/omim>). En la ventana de diálogo, escriba "Non-insulin dependent Diabetes mellitus" y oprima el botón "SEARCH". Esto abrirá la página de OMIM y le brindará en una ventana los primeros 20 registros de la enfermedad. Debajo de cada una de los diferentes registros, se encuentra la localización de los genes asociados, incluyendo los de susceptibilidad (Fig 5.1). Modifique la página para poder ver los diferentes genes que recibe la página. Abra aquellos genes que no sea explícita su participación en la enfermedad.

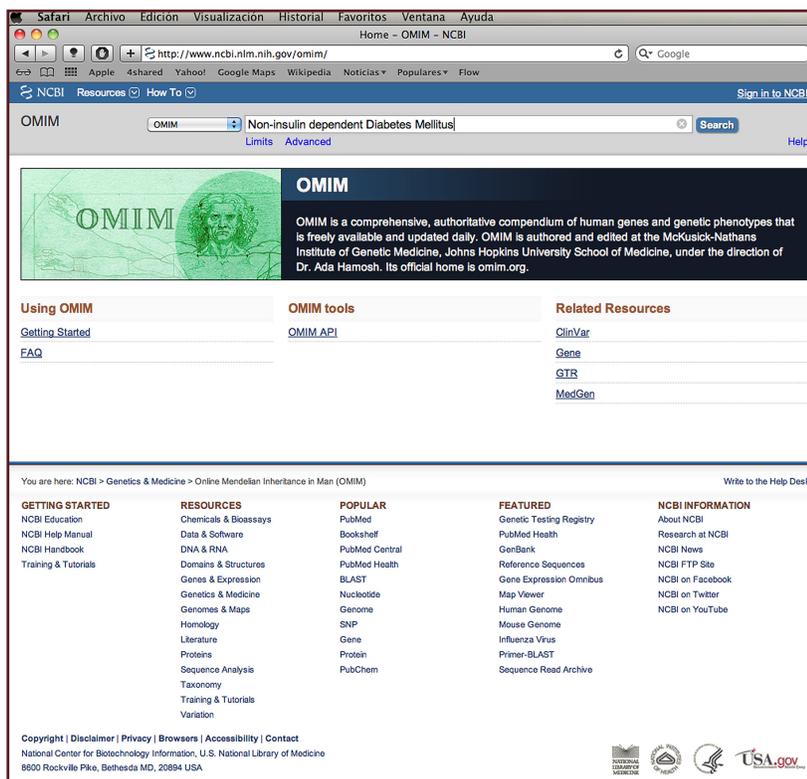


Figura 5.1a. Ventana de acceso a la base de datos OMIM.

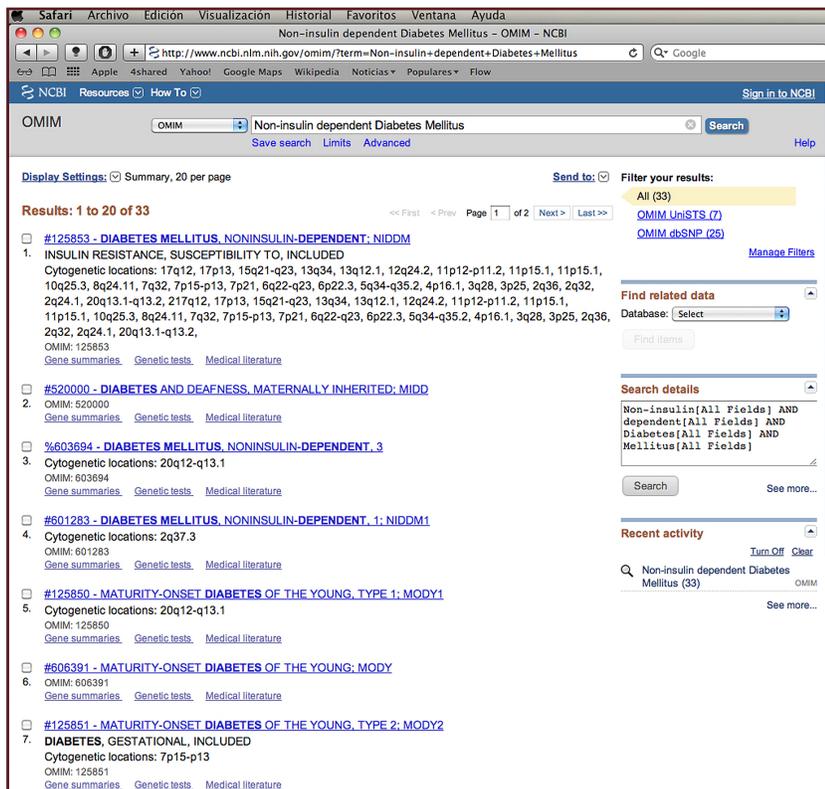


Figura 5.1b. Registros iniciales de Diabetes no dependiente de insulina.

Seleccione los registros asociados a la enfermedad. Elabore una lista de aquellos directamente asociados. En la pestaña de OMIM UniSTS confirme que su listado sea el asociado de acuerdo a la expresión de genes únicos. Haga un listado de la localización de estos genes. Abra la pestaña de OMIM dbSNP para ver la posición de los marcadores (Figura 5.2).

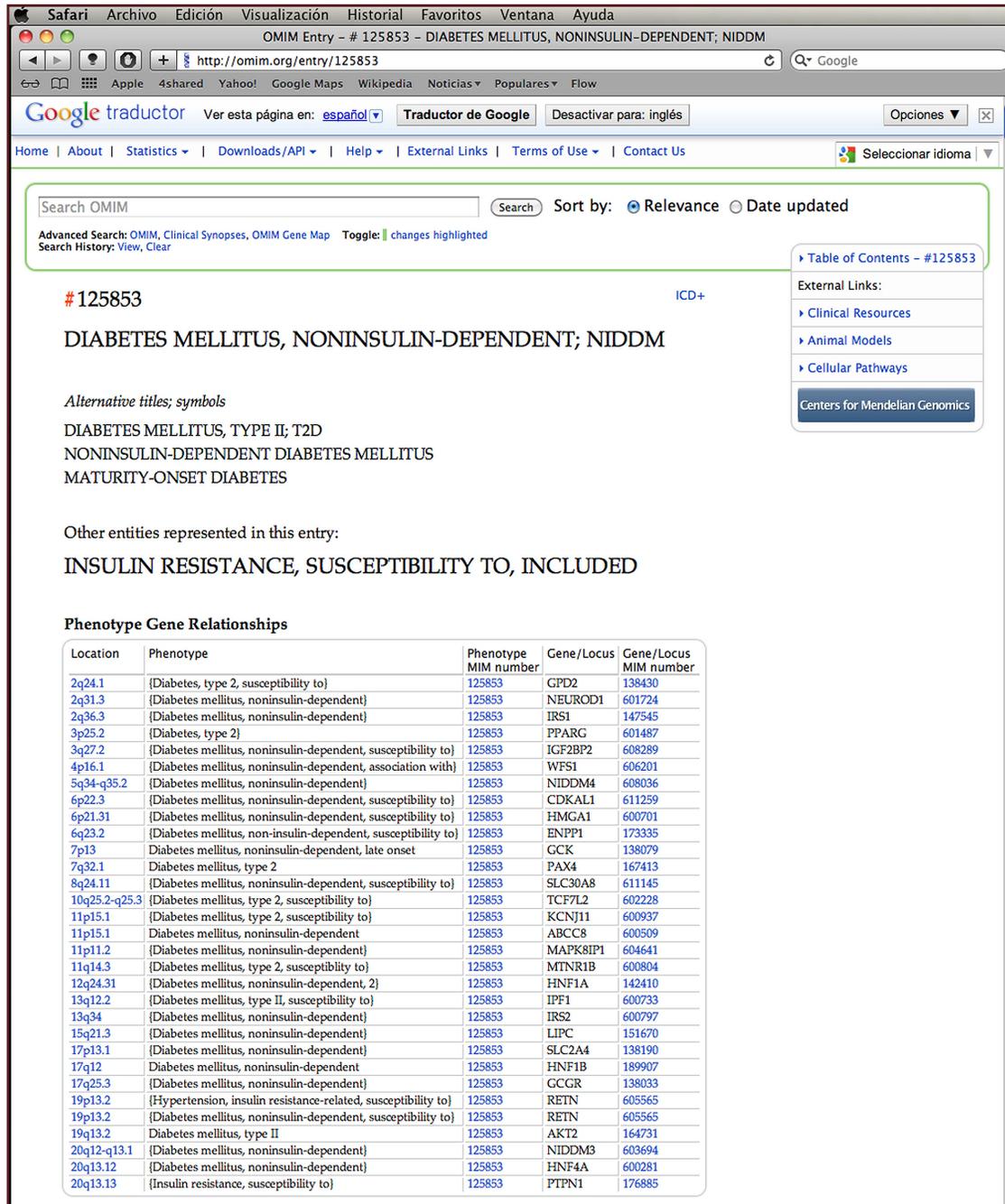
The screenshot shows the OMIM database search results for 'Non-insulin dependent Diabetes Mellitus'. The search results are displayed in a list format, with each entry including the OMIM ID, the disease name, and the cyto-genetic locations. The search filters are set to 'All (33)', 'OMIM UniSTS (7)', and 'OMIM dbSNP (25)'. The search details sidebar shows the search criteria: 'Non-insulin[All Fields] AND dependent[All Fields] AND Diabetes[All Fields] AND Mellitus[All Fields]'. The recent activity sidebar shows 'Non-insulin dependent Diabetes Mellitus (33)'.

Figura 5.2a. Ventana de resultados de las bases de datos de secuencias estandarizadas únicas, UniSTS.

The screenshot shows the OMIM database search results for 'Non-insulin dependent Diabetes Mellitus'. The search results are displayed in a list format, with each entry including the OMIM ID, the disease name, and the cyto-genetic locations. The search filters are set to 'All (33)', 'OMIM UniSTS (7)', and 'OMIM dbSNP (25)'. The search details sidebar shows the search criteria: 'Non-insulin[All Fields] AND dependent[All Fields] AND Diabetes[All Fields] AND Mellitus[All Fields]'. The recent activity sidebar shows 'Non-insulin dependent Diabetes Mellitus (33)'.

Figura 5.2b. Marcadores moleculares de variaciones de una sola base, SNP de Diabetes no dependiente de insulina.

Active la liga de los genes asociados para desplegar el mapa que muestra la localización del gene. Active la liga del número MIM para desplegar las características del gen (estructura, función, evidencias de asociación con la enfermedad, etc). Identifique el papel que tienen en la posible causa de la enfermedad (Figura 5.3). Utilizando la liga MAPVIEWER que aparece, indique la búsqueda de los genes asociados a DIABETES MELLITUS. Identifique por cromosoma la cantidad de genes que se han asociado.



OMIM Entry - # 125853 - DIABETES MELLITUS, NONINSULIN-DEPENDENT; NIDDM

http://omim.org/entry/125853

Google traductor Ver esta página en: **español** Traductor de Google Desactivar para: inglés

Home | About | Statistics | Downloads/API | Help | External Links | Terms of Use | Contact Us

Search OMIM Search Sort by: Relevance Date updated

Advanced Search: OMIM, Clinical Synopses, OMIM Gene Map Toggle: changes highlighted Search History: View, Clear

125853 ICD+

DIABETES MELLITUS, NONINSULIN-DEPENDENT; NIDDM

Alternative titles; symbols

DIABETES MELLITUS, TYPE II; T2D
NONINSULIN-DEPENDENT DIABETES MELLITUS
MATURITY-ONSET DIABETES

Other entities represented in this entry:

INSULIN RESISTANCE, SUSCEPTIBILITY TO, INCLUDED

Phenotype Gene Relationships

| Location | Phenotype | Phenotype MIM number | Gene/Locus | Gene/Locus MIM number |
|---------------|---|----------------------|------------|-----------------------|
| 2q24.1 | [Diabetes, type 2, susceptibility to] | 125853 | GPD2 | 138430 |
| 2q31.3 | [Diabetes mellitus, noninsulin-dependent] | 125853 | NEUROD1 | 601724 |
| 2q36.3 | [Diabetes mellitus, noninsulin-dependent] | 125853 | IRS1 | 147545 |
| 3p25.2 | [Diabetes, type 2] | 125853 | PPARG | 601487 |
| 3q27.2 | [Diabetes mellitus, noninsulin-dependent, susceptibility to] | 125853 | IGF2BP2 | 608289 |
| 4p16.1 | [Diabetes mellitus, noninsulin-dependent, association with] | 125853 | WFS1 | 606201 |
| 5q34-q35.2 | [Diabetes mellitus, noninsulin-dependent] | 125853 | NIDDM4 | 608036 |
| 6p22.3 | [Diabetes mellitus, noninsulin-dependent, susceptibility to] | 125853 | CDKAL1 | 611259 |
| 6p21.31 | [Diabetes mellitus, noninsulin-dependent, susceptibility to] | 125853 | HMGAI | 600701 |
| 6q23.2 | [Diabetes mellitus, non-insulin-dependent, susceptibility to] | 125853 | ENPP1 | 173335 |
| 7p13 | Diabetes mellitus, noninsulin-dependent, late onset | 125853 | GCK | 138079 |
| 7q32.1 | Diabetes mellitus, type 2 | 125853 | PAX4 | 167413 |
| 8q24.11 | [Diabetes mellitus, noninsulin-dependent, susceptibility to] | 125853 | SLC30A8 | 611145 |
| 10q25.2-q25.3 | [Diabetes mellitus, type 2, susceptibility to] | 125853 | TCF7L2 | 602228 |
| 11p15.1 | [Diabetes mellitus, type 2, susceptibility to] | 125853 | KCNJ11 | 600937 |
| 11p15.1 | Diabetes mellitus, noninsulin-dependent | 125853 | ABCC8 | 600509 |
| 11p11.2 | [Diabetes mellitus, noninsulin-dependent] | 125853 | MAPK8IP1 | 604641 |
| 11q14.3 | [Diabetes mellitus, type 2, susceptibility to] | 125853 | MTNR1B | 600804 |
| 12q24.31 | [Diabetes mellitus, noninsulin-dependent, 2] | 125853 | HNF1A | 142410 |
| 13q12.2 | [Diabetes mellitus, type II, susceptibility to] | 125853 | IPF1 | 600733 |
| 13q34 | [Diabetes mellitus, noninsulin-dependent] | 125853 | IRS2 | 600797 |
| 15q21.3 | [Diabetes mellitus, noninsulin-dependent] | 125853 | LIPC | 151670 |
| 17p13.1 | [Diabetes mellitus, noninsulin-dependent] | 125853 | SLC2A4 | 138190 |
| 17q12 | Diabetes mellitus, noninsulin-dependent | 125853 | HNF1B | 189907 |
| 17q25.3 | [Diabetes mellitus, noninsulin-dependent] | 125853 | GCGR | 138033 |
| 19p13.2 | [Hypertension, insulin resistance-related, susceptibility to] | 125853 | RETN | 605565 |
| 19p13.2 | [Diabetes mellitus, noninsulin-dependent, susceptibility to] | 125853 | RETN | 605565 |
| 19q13.2 | Diabetes mellitus, type II | 125853 | AKT2 | 164731 |
| 20q12-q13.1 | [Diabetes mellitus, noninsulin-dependent] | 125853 | NIDDM3 | 603694 |
| 20q13.12 | [Diabetes mellitus, noninsulin-dependent] | 125853 | HNF4A | 600281 |
| 20q13.13 | [Insulin resistance, susceptibility to] | 125853 | PTPN1 | 176885 |

Figura 5.3a. Ventana informativa sobre el registro 125853 de diabetes mellitus no dependiente de insulina. Genes asociados a la enfermedad.

TEXT

A number sign (#) is used with this entry because of evidence that more than one gene is involved in the causation of noninsulin-dependent diabetes mellitus (NIDDM).

See [601283](#) for description of a form of NIDDM linked to 2q, which may be caused by mutation in the gene encoding calpain-10 (CAPN10; [605286](#)). See [601407](#) for description of a chromosome 12q locus, NIDDM2, found in a Finnish population. See [603694](#) for description of a locus on chromosome 20, NIDDM3.

A mutation has been observed in hepatocyte nuclear factor-4-alpha (HNF4A; [600281.0004](#)) in a French family with NIDDM of late onset. Mutations in the NEUROD1 gene ([601724](#)) on chromosome 2q32 were found to cause type II diabetes mellitus in 2 families. Mutation in the GLUT4 glucose transporter was associated with NIDDM in 1 patient ([138190.0001](#)) and in the GLUT2 glucose transporter in another ([138160.0001](#)). Mutation in the MAPK8IP1 gene, which encodes the islet-brain-1 protein, was found in a family with type II diabetes in individuals in 4 successive generations ([604641.0001](#)). Polymorphism in the KCNJ11 gene ([600937.0014](#)) confers susceptibility. In French white families, [Vionnet et al. \(2000\)](#) found evidence for a susceptibility locus for type II diabetes on 3q27-qter. They confirmed the diabetes susceptibility locus on 1q21-q24 reported by [Elbein et al. \(1999\)](#) in whites and by [Hanson et al. \(1998\)](#) in Pima Indians. A mutation in the GPD2 gene ([138430.0001](#)) on chromosome 2q24.1, encoding mitochondrial glycerophosphate dehydrogenase, was found in a patient with type II diabetes mellitus and in his glucose-intolerant half sister. Mutations in the PAX4 gene ([167413](#)) have been identified in patients with type II diabetes. [Triggs-Raine et al. \(2002\)](#) stated that in the Oji-Cree, a gly319-to-ser change in HNF1-alpha ([142410.0008](#)) behaves as a susceptibility allele for type II diabetes. Mutation in the HNF1B gene ([189907.0007](#)) was found in 2 Japanese patients with typical late-onset type II diabetes. Mutations in the IRS1 gene ([147545](#)) have been found in patients with type II diabetes. [Reynisdottir et al. \(2003\)](#) mapped a susceptibility locus for type II diabetes to chromosome 5q34-q35.2 (NIDDM4; [608036](#)). A missense mutation in the AKT2 gene ([164731.0001](#)) caused autosomal dominant type II diabetes in 1 family. A (single-nucleotide polymorphism) SNP in the 3-prime untranslated region of the resistin gene ([605565.0001](#)) was associated with susceptibility to diabetes and to insulin resistance-related hypertension in Chinese subjects. Susceptibility to insulin resistance has been associated with polymorphism in the TCF1 ([142410.0011](#)), PPP1R3A ([600917.0001](#)), PTPN1 ([176885.0001](#)), ENPP1 ([173335.0006](#)), IRS1 ([147545.0002](#)), and EPHX2 ([132811.0001](#)) genes. The K121Q polymorphism of ENPP1 ([173335.0006](#)) is associated with susceptibility to type II diabetes; a haplotype defined by 3 SNPs of this gene, including K121Q, is associated with obesity, glucose intolerance, and type II diabetes. A SNP in the promoter region of the hepatic lipase gene ([151670.0004](#)) predicts conversion from impaired glucose tolerance to type II diabetes. Variants of transcription factor 7-like-2 (TCF7L2; [602228.0001](#)), located on 10q, have also been found to confer risk of type II diabetes. A common sequence variant, [rs10811661](#), on chromosome 9p21 near the CDKN2A ([600160](#)) and CDKN2B ([600431](#)) genes has been associated with risk of type II diabetes. Variation in the PPARG gene ([601487](#)) has been associated with risk of type 2 diabetes. A promoter polymorphism in the IL6 gene ([147620](#)) is associated with susceptibility to NIDDM. Variation in the KCNJ15 gene ([602106](#)) has been associated with T2DM in lean Asians. Variation in the HMGA1 gene ([600701.0001](#)) is associated with an increased risk of type II diabetes. Mutation in the MTNR1B gene ([600804](#)) is associated with

Figura 5.3b. Descripción de síntomas clínicos, localización cromosómica y asociación genética para Diabetes no dependiente de insulina.

Actividades para reportar

Utilizando las herramientas desarrolladas en esta sesión, haga el análisis de las patologías asociadas a los receptores adrenérgicos Alfa.

De la misma manera, analice los genes asociados al síndrome vasovagal.

Referencias

-  Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L. 2007. The human disease network. Proc. Nat. Acad. Sci. 104:8685-8690. Disponible en la URL <http://www.pnas.org/content/104/21/8685.full.pdf+html>
-  Schriml LM, Arze C, Nadendla S, Chang Y-W W, Mazaitis M, Felix V, Feng G, Kibbe WA. 2012. Disease Ontology: a backbone for disease semantic integration. Nucleic Acid Res. 40:D940-D946. Disponible en la URL <http://nar.oxfordjournals.org/content/40/D1/D940.full.pdf+html>

Capítulo 6

Diseño de oligonucleótidos para amplificación

Una de las metodologías que ha permitido el avance de la Biología Molecular y la Genética Molecular es la reacción en cadena de la polimerasa, mejor conocida por sus siglas en inglés, PCR. Esta reacción permite la obtención de cantidades de ADN en cantidades adecuadas para ser utilizadas o analizadas por otras técnicas a partir de pequeñas cantidades de material inicial. Si simplificamos el procedimiento, se trata de una reacción de replicación del material genético bajo condiciones *in vitro*. Como en el caso del proceso *in vivo*, se requiere de un sustrato que puede ser ADN o ARN, reactivos necesarios que servirán de unidades estructurales a partir de las cuales se construirán las nuevas cadenas de ADN, componentes que inician la reacción y aquellos que la mantienen, incluyendo las sales que permiten tener una fuerza iónica adecuada para que el funcionamiento de la enzima pueda realizarse bajo las mejores condiciones. Podemos encuadrar en cada uno de estos rubros a los desoxinucleótidos trifosfatados (dNTP), iones Mg^{2+} , glicerol, amortiguadores de Tris-hidroximetil-amino-metano (Tris), los cebadores de la polimerasa y la propia enzima, comúnmente se utiliza la ADN polimerasa termoestable proveniente de la arqueobacteria *Thermus aquaticus*, mejor conocida como Taq polimerasa.

Bajo condiciones naturales, las polimerasas requieren de un oligonucleótido producido por la enzima DNA primasa por lo que se denomina "primer" que puede ser traducido como cebador o iniciador. Este es un oligonucleótido de sólo 30 a 50 bases que proporcionan el sitio de reconocimiento de la ADN polimerasa: el extremo 3' hidroxilo de un nucleótido apareado a la cadena que se copiará. Al unirse a este extremo, la ADN polimerasa continúa con la elongación de la cadena incorporando el desoxinucleótido trifosfatado que complementa la doble cadena, de acuerdo a las reglas de apareamiento de Watson-Crick: si el molde tiene Adenina, en la cadena complementaria se incorpora Timidina; si el molde presenta Citosina, la complementaria incorporará una Guanina.

En la caso de la reacción *in vitro*, se copia únicamente una región específica por lo que se utiliza un par de iniciadores cuya secuencia es complementaria a las zonas que se van a amplificar. La especificidad, eficiencia y posibles variaciones dependen entonces de un diseño adecuado de estos oligonucleótidos. A manera de características generales, los aspectos que deben cuidarse en el diseño de estos oligonucleótidos son su secuencia, temperatura de alineamiento, que no sean autocomplementarios, que no formen estructuras tridimensionales complejas de más de 3 bases, que no se unan a secuencias similares internas (sitios secundarios de hibridación), y que la unión en el extremo 3' del propio oligonucleótido esté estabilizado por una base capaz de formar 3 puentes de hidrógeno con su base complementaria.

Objetivo

Diseñar un par de cebadores capaces de ser utilizados en una reacción de amplificación, determinar las características que tendría la reacción de amplificación y la posible longitud que tendría el fragmento amplificado (amplicón) utilizando un recurso en línea.

Desarrollo

Obtención de iniciadores mediante el uso de un portal de herramientas bioinformáticas

El inicio del análisis implica tener una secuencia blanco a la cual estará dirigida la metodología. Obtenga la secuencia en FASTA del acceso NM_023085 correspondiente a la calpaina 10 humana depositada en el GenBank (<http://www.ncbi.nlm.nih.gov/gene>). Obtenga la secuencia en FASTA activando la liga que se encuentra en la parte superior derecha del mapa cromosómico (figura 6.1). Seleccione todo el contenido, copie y guarde ya sea en la memoria temporal de su computadora o en un archivo de un procesador de textos.

The screenshot shows the NCBI Gene database homepage. At the top, there is a search bar with 'EST' selected and a 'Search' button. Below the search bar, there are links for 'Limits' and 'Advanced'. The main content area features a 'Gene' header with a description: 'Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and links to genome-, phenotype-, and locus-specific resources worldwide.' Below this, there are three columns of links: 'Using Gene' (Gene Quick Start, FAQ, Download/FTP, RefSeq Mailing List, Gene News, Factsheet), 'Gene Tools' (Submit GeneRIFs, Submit Correction, Statistics, BLAST, Genome Workbench, Snpip), and 'Other Resources' (HomoloGene, OMIM, RefSeq, RefSeqGene, UniGene, Protein Clusters). A 'Representative queries' section follows, listing various search criteria and their corresponding search text examples.

| Find genes by... | Search text |
|---|--|
| free text | human muscular dystrophy |
| chromosome and symbol | ((l[chr] OR 2[chr]) AND adh*[sym]) |
| partial name and multiple species | alive[prop] AND transporter[title] AND ("Drosophila melanogaster"[orgn] OR "Mus musculus"[orgn]) |
| associated sequence accession | M11313[accn] |
| gene name (symbol) | BRCA1[sym] |
| publication (PubMed ID) | 11331580[PMID] |
| Gene Ontology (GO) terms or identifiers | "cell adhesion"[GO] 10030[GO] |
| genes with short variants of medical interest | gene_snp_clin[filter] |
| chromosome and species | Y[CHRI] AND human[ORGN] |
| Enzyme Commission (EC) numbers | 1.9.3.1[EC] |

Figura 6.1a. Ventana de acceso al GenBank.

Gene [Limits](#) [Advanced](#)

Display Settings: Full Report [Send to:](#)

CAPN10 calpain 10 [*Homo sapiens* (human)]

Gene ID: 11132, updated on 4-Aug-2013

Summary

Official Symbol CAPN10 provided by HGNC

Official Full Name calpain 10 provided by HGNC

Primary source [HGNC:1477](#)

See related [Ensembl:ENSG00000142330](#); [HPRD:05595](#); [MIM:605286](#); [Vega:OTTHUMG00000133358](#)

Gene type protein coding

RefSeq status REVIEWED

Organism [Homo sapiens](#)

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

Also known as CANP10; NIDDM1

Summary Calpains represent a ubiquitous, well-conserved family of calcium-dependent cysteine proteases. The calpain proteins are heterodimers consisting of an invariant small subunit and variable large subunits. The large catalytic subunit has four domains: domain I, the N-terminal regulatory domain that is processed upon calpain activation; domain II, the protease domain; domain III, a linker domain of unknown function; and domain IV, the calmodulin-like calcium-binding domain. This gene encodes a large subunit. It is an atypical calpain in that it lacks the calmodulin-like calcium-binding domain and instead has a divergent C-terminal domain. It is similar in organization to calpains 5 and 6. This gene is associated with type 2 or non-insulin-dependent diabetes mellitus (NIDDM), and is located within the NIDDM1 region. Multiple alternative transcript variants have been described for this gene. [provided by RefSeq, Sep 2010]

Genomic context

Location: 2q37.3 [See CAPN10 in Epigenomics, MapViewer](#)

Sequence: Chromosome: 2; NC_000002.11 (241526133..241538526)

Chromosome 2 - NC_000002.11

[241,499,471] [241,612,514]

Genomic regions, transcripts, and products

Genomic Sequence [Go to reference sequence details](#)

[Go to nucleotide](#) [Graphics](#) [FASTA](#) [GenBank](#)

NC_000002.11: 242M..242M (16Kbp)

Figura 6.1b. Región del cromosoma 2 humano donde se localiza el gen de la calpaina 10.

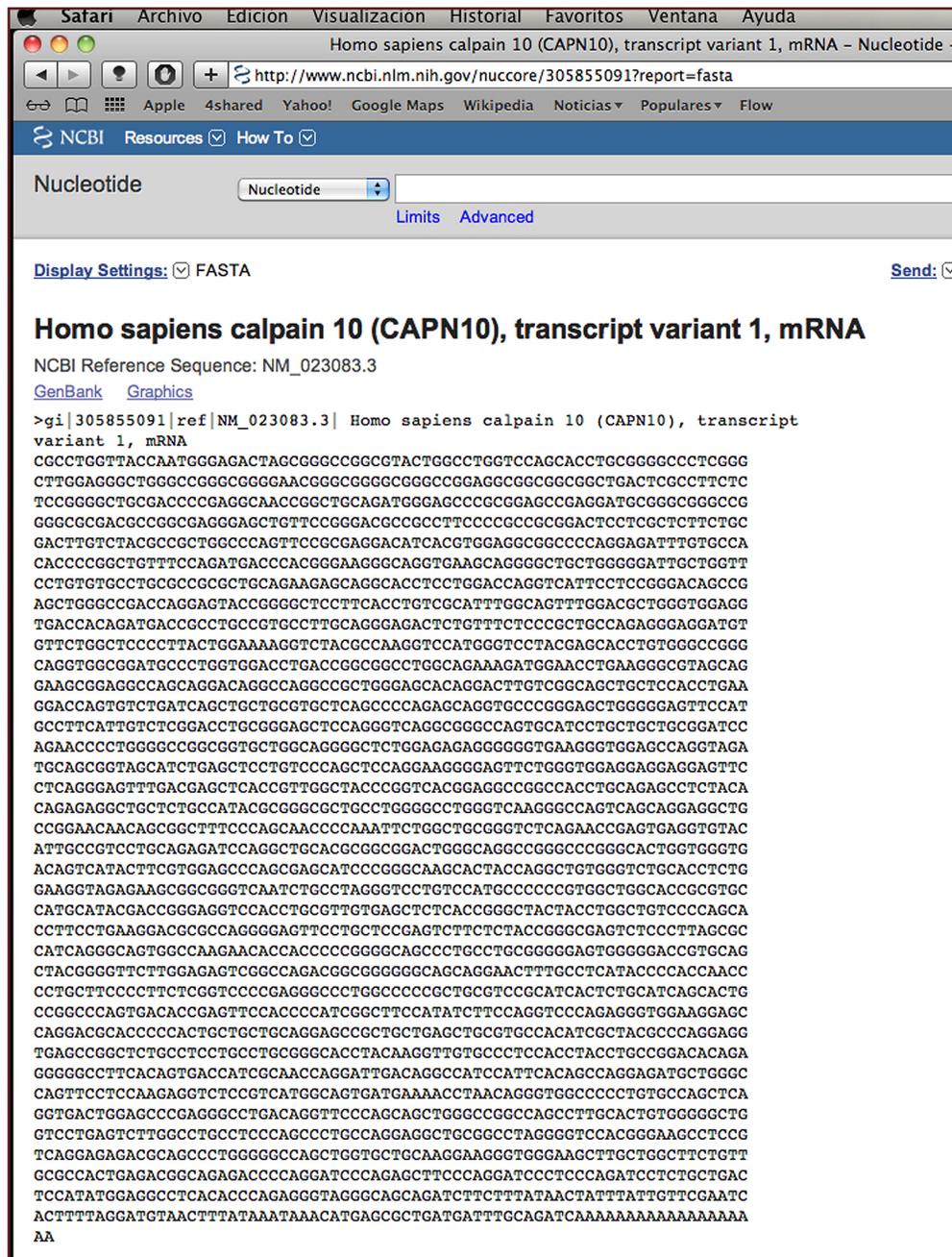


Figura 6.1c Secuencia de nucleótidos del gen capn10.

Ingrese al portal de herramientas de ciencia y tecnología de Israel en la URL <http://www.scienc.co.il/Biomedical/Primer-Tools.asp> lo que le abrirá la página que se muestra en la figura 6.1 Dentro del menú de sitios, seleccione la opción PRIMER3 que lo llevará a la URL <http://frodo.wi.mit.edu/primer3> y le desplegará la ventana que se muestra en la figura 6.2 En la ventana de diálogo, pegue la secuencia FASTA que obtuvo del GenBank, identifique su secuencia como “Calpaina10” y la región que queremos amplificar que en este caso corresponde a la región 4 bases alrededor del nucleótido 1469. Haga esta indicación ingresando “1469,4” en el cuadro de diálogo marcado como “TARGET”. Delimite a un amplicón de 300 a 400 nucleótidos de longitud seleccionando este valor en el cuadro marcado “Product size ranges”

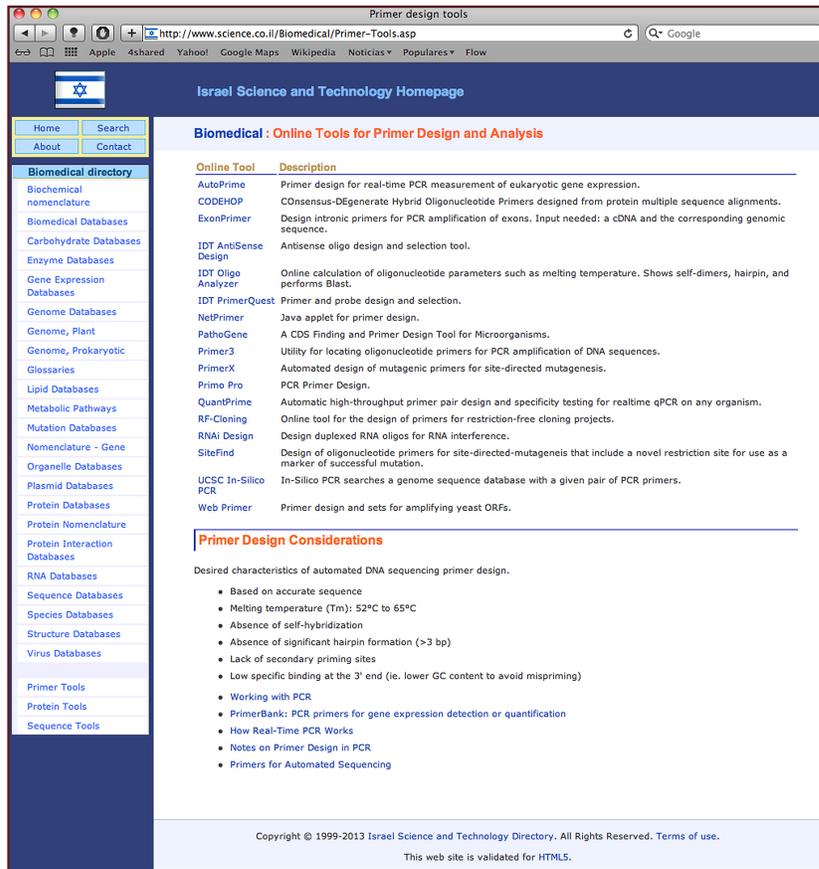


Figura 6.2a. Ventana de acceso al portal de apoyo bioinformático en el área de la Biomedicina de Israel.

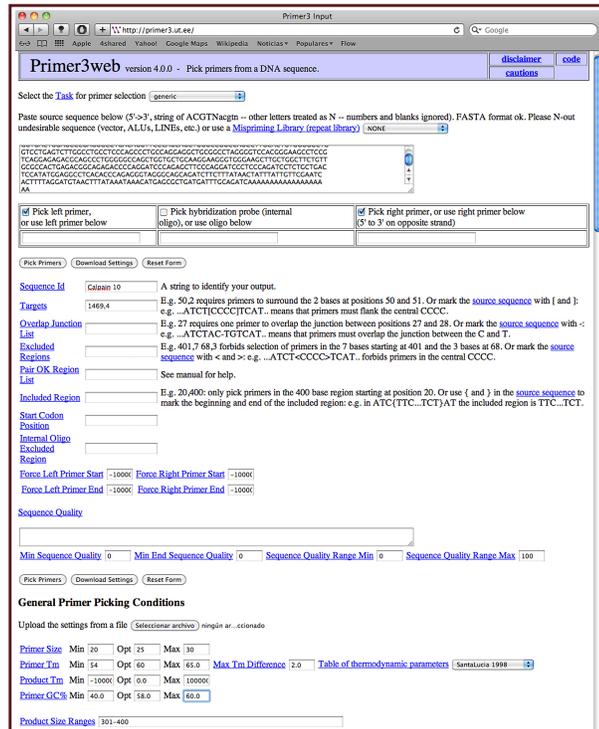


Figura 6.2b. Consulta del algoritmo Primer3 del MIT.

Las características que queremos de los cebadores son tamaño comprendido entre 20 y 30 nucleótidos con un óptimo de 25, la temperatura de linearización entre 54 y 64 °C con una óptima de 60, contenido de CG entre 40 y 60% con óptimo de 58. Coloque estos parámetros en los cuadros correspondientes en la sección "GENERAL PRIMER PICKING CONDITIONS". De los 3 métodos para calcular las características de corrección en el cálculo de la temperatura de linearización (Tm), utilice el modelo de SantaLucia (1998) de corrección de interacciones secuenciales múltiples dependiente de las condiciones de la reacción. Al final de estas indicaciones, active el botón de acción localizado en el final de la sección como "PICK PRIMERS". El programa le indicará que se encuentra trabajando en el diseño de los cebadores. Cuando encuentre algunos, se desplegará la página de resultados (figura 6.3). La tabla de resultados le indica las características de los cebadores con mejor eficiencia. Además de la secuencia, le indica la temperatura de hibridación y en la parte inferior de la tabla, el posible tamaño del amplicon seguido de la secuencia del gene que está tratando de amplificar. Dentro de ella, marcado con una serie de símbolos ">>>". Después de esta sugerencia, el programa regresa pares de iniciadores adicionales. Esto permite hacer una selección dependiendo del interés del investigador.



```

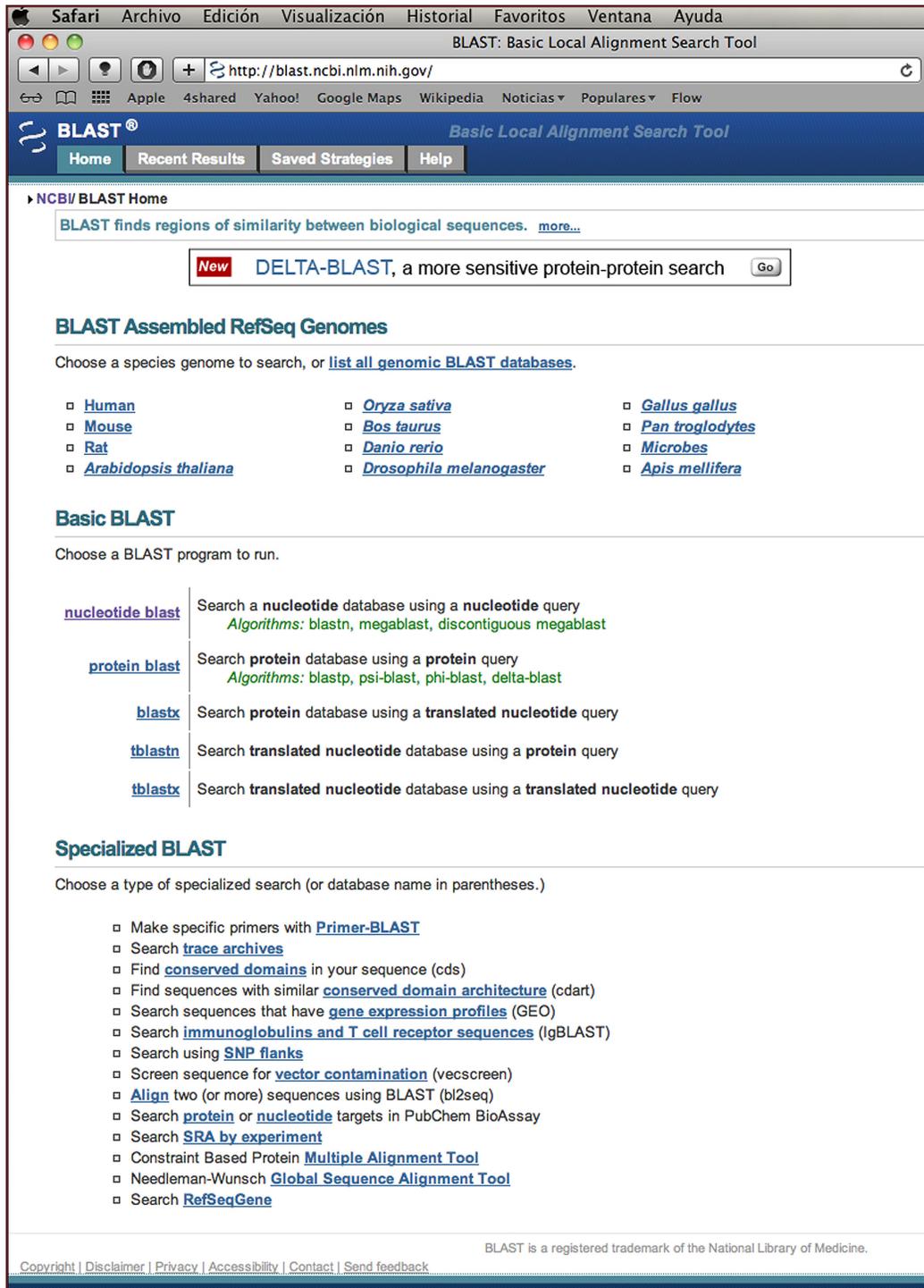
Primer3 Output (primer3_results.cgi release 4.0.0)
http://bioinfo.ebc.ee/cgi-bin/primer3/primer3web_results.cgi
Apple 4shared Yahoo! Google Maps Wikipedia Noticias Populares Flow

Primer3 Output

No mispriming library specified
Using 1-based sequence positions
OLIGO      start  len  tm   gc%  any_th  3'_th  hairpin  seq
LEFT PRIMER 1399  20  55.93 45.00 4.15  0.00  0.00  TGACAGTCATACTTCGTGGA
RIGHT PRIMER 1770  20  58.44 55.00 0.00  0.00  0.00  ACTCTCCAAGAACCCCGTAG
SEQUENCE SIZE: 2662
INCLUDED REGION SIZE: 2662
PRODUCT SIZE: 372, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00
TARGETS (start len
1  CGCCTGGTTACCAATGGGAGACTAGCGGGCCGGCTACTGGCCTGGTCCAGCACCTGCGG
61  GCGCCCTCGGGCTTGGAGGGCTGGCCGGGGGGAACGGGGCGGGCCGGAGGGCGGC
121 GCGCGCTGACTCGCCTTCTCTCCGGGGCTGCGACCCCGAGCAACCGGTGCAGATGGGA
181 GCGCCGGAGCCGAGGATCGGGCGGGCCGGGGCGGACCGCGCAGGGAGCTTTCCG
241 GGACGCCGCCTTCCCGCCGGGACTCTCGCTCTTCTGCGACTTGTCTACGCCGCTGGC
301 CCAAGTCCCGGAGGACATCACGTGGAGCGGCCCCAGGAGATTTGTGCCACACCCGGCT
361 GTTTCAGATGACCCACGGGAAGGGCAGGTGAAGCAGGGGCTGCTGGGGGATTCGTGGTT
421 CCTGTGTGCTGCGCCCGCTGCAGAAAGCAGGACCTCCTGACCCAGGTCAATTCCTCC
481 GGAACAGCCGAGCTGGGCCGACAGGAGTACCGGGCTCCTTACCTGTCCGATTTGGCA
541 GTTTGGACGCTGGGTGGAGGTGACCACAGATGACCCGCTGCCGTGCCTTGCAGGGAGACT
601 CTGTTCTCCCGTCCAGAGGGAGGATGTGTTTGGCTCCCTTACTGGAAAAGGCTTA
661 GCGCAAGGTCCATGGGTCTACGAGCACCTGTGGGCCGGGCAAGTGGCGGATGCCCTGGT
721 GGACCTGACCGCGGCTTGGCAGAAAGATGGAACCTGAAGGGCTAGCAGGAAGCGGAGG
781 CCAGCAGGACAGGCCAGCCGCTGGGAGCACAGGACTTGTTCGCACTGCTCCACCTGAA
841 GGACCAAGTGTCTGATCAGCTGCTGCCTGCTCAGCCCAAGCAGGTGCCCGGAGCTGGG
901 GGAATTCATGCTTCAATGTCTCGGACTGCGGAGCTCCAGGTCAGGCCGGCCAGTG
961 CATCTGCTGCTGCGGATCCAGAACCCCTGGGGCCGGCGGTGCTGGCAGGGGCTCTGGAG
1021 AGAGGGGGTGAAGGGTGGAGCCAGGTAGATGCAGCGGTAGCATCTGAGCTCCTTCCCA
1081 GCTCCAGAAAGGGAGTCTGGGTGGAGGAGGAGGACTCCTCAGGACTTTGACGAGCT
1141 CACCGTTGGCTACCGGCTCACGGAGGCCCGCCACCTGCAGAGCCTTACACAGAGGCT
1201 GCTCTGCCATACCGGGCCGCTGCCCTGGGGCTGGGTCAAGGGCCAGTCAGCAGGAGGCTG
1261 CCGGAACACACCGGCTTCCCAAGAACCCAAATTCCTGGCTGCGGCTCTCAGAACCGAG

```

Figura 6.3a. Resultados iniciales del diseño Prime3 con la secuencia y características de los oligonucleótidos iniciadores así como la posible longitud del amplicón.



Safari Archivo Edición Visualización Historial Favoritos Ventana Ayuda

BLAST: Basic Local Alignment Search Tool

http://blast.ncbi.nlm.nih.gov/

Apple 4shared Yahoo! Google Maps Wikipedia Noticias Populares Flow

BLAST® Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

NCBI/ BLAST Home

BLAST finds regions of similarity between biological sequences. [more...](#)

New DELTA-BLAST, a more sensitive protein-protein search

BLAST Assembled RefSeq Genomes

Choose a species genome to search, or [list all genomic BLAST databases](#).

- [Human](#)
- [Mouse](#)
- [Rat](#)
- [Arabidopsis thaliana](#)
- [Oryza sativa](#)
- [Bos taurus](#)
- [Danio rerio](#)
- [Drosophila melanogaster](#)
- [Gallus gallus](#)
- [Pan troglodytes](#)
- [Microbes](#)
- [Apis mellifera](#)

Basic BLAST

Choose a BLAST program to run.

- [nucleotide blast](#) Search a **nucleotide** database using a **nucleotide** query
Algorithms: blastn, megablast, discontinuous megablast
- [protein blast](#) Search **protein** database using a **protein** query
Algorithms: blastp, psi-blast, phi-blast, delta-blast
- [blastx](#) Search **protein** database using a **translated nucleotide** query
- [tblastn](#) Search **translated nucleotide** database using a **protein** query
- [tblastx](#) Search **translated nucleotide** database using a **translated nucleotide** query

Specialized BLAST

Choose a type of specialized search (or database name in parentheses.)

- Make specific primers with [Primer-BLAST](#)
- Search [trace archives](#)
- Find [conserved domains](#) in your sequence (cds)
- Find sequences with similar [conserved domain architecture](#) (cdart)
- Search sequences that have [gene expression profiles](#) (GEO)
- Search [immunoglobulins and T cell receptor sequences](#) (IgBLAST)
- Search using [SNP flanks](#)
- Screen sequence for [vector contamination](#) (vecscreen)
- [Align](#) two (or more) sequences using BLAST (bl2seq)
- Search [protein](#) or [nucleotide](#) targets in PubChem BioAssay
- Search [SRA by experiment](#)
- Constraint Based Protein [Multiple Alignment Tool](#)
- Needleman-Wunsch [Global Sequence Alignment Tool](#)
- Search [RefSeqGene](#)

BLAST is a registered trademark of the National Library of Medicine.

Copyright | Disclaimer | Privacy | Accessibility | Contact | Send feedback

Figura 6.4a. Ventana de acceso a la herramienta de diseño de iniciadores.

The screenshot displays the NCBI Primer-BLAST web interface. The browser address bar shows the URL: http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome. The page title is "Primer-BLAST: Finding primers specific to your PCR template (using Primer3 and BLAST)".

PCR Template

Enter accession, gi, or FASTA sequence (A refseq record is preferred) [Clear](#)

Range: From To [Clear](#)

Forward primer
Reverse primer

Or, upload FASTA file ningún ...ionado

Primer Parameters

Use my own forward primer (5'->3' on plus strand) [Clear](#)
Use my own reverse primer (5'->3' on minus strand) [Clear](#)

PCR product size: Min Max

of primers to return:

Primer melting temperatures (T_m): Min Opt Max Max T_m difference

Exon/intron selection

A refseq mRNA sequence as PCR template input is required for options in the section

Exon junction span:

Exon junction match: Exon at 5' side Exon at 3' side
Minimal number of bases that must anneal to exons at the 5' or 3' side of the junction

Intron inclusion: Primer pair must be separated by at least one intron on the corresponding genomic DNA

Intron length range: Min Max

Primer Pair Specificity Checking Parameters

Note: Parameter values that differ from the default are highlighted in yellow

Specificity check: Enable search for primer pairs specific to the intended PCR template

Database:

Organism:
Enter an organism name, taxonomy id or select from the suggestion list as you type.
[Add more organisms](#)

Exclusion (optional): Exclude predicted Refseq transcripts (accession with XM, XR prefix) Exclude uncultured/environmental sample sequences

Entrez query (optional):

Primer specificity stringency: Primer must have at least total mismatches to unintended targets, including at least mismatches within the last bps at the 3' end.
Ignore targets that have or more mismatches to the primer.

Misprimed product size deviation:

Splice variant handling: Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input)

[Get Primers](#) Show results in a new window Use new graphic view

Advanced parameters **Note: Parameter values that differ from the default are highlighted in yellow**

Copyright | Disclaimer | Privacy | Accessibility | Contact | Send feedback on new interface NCBI | NLM | NIH | DHHS

Figura 6.4b. Características de la ventana de dialogo inicial.

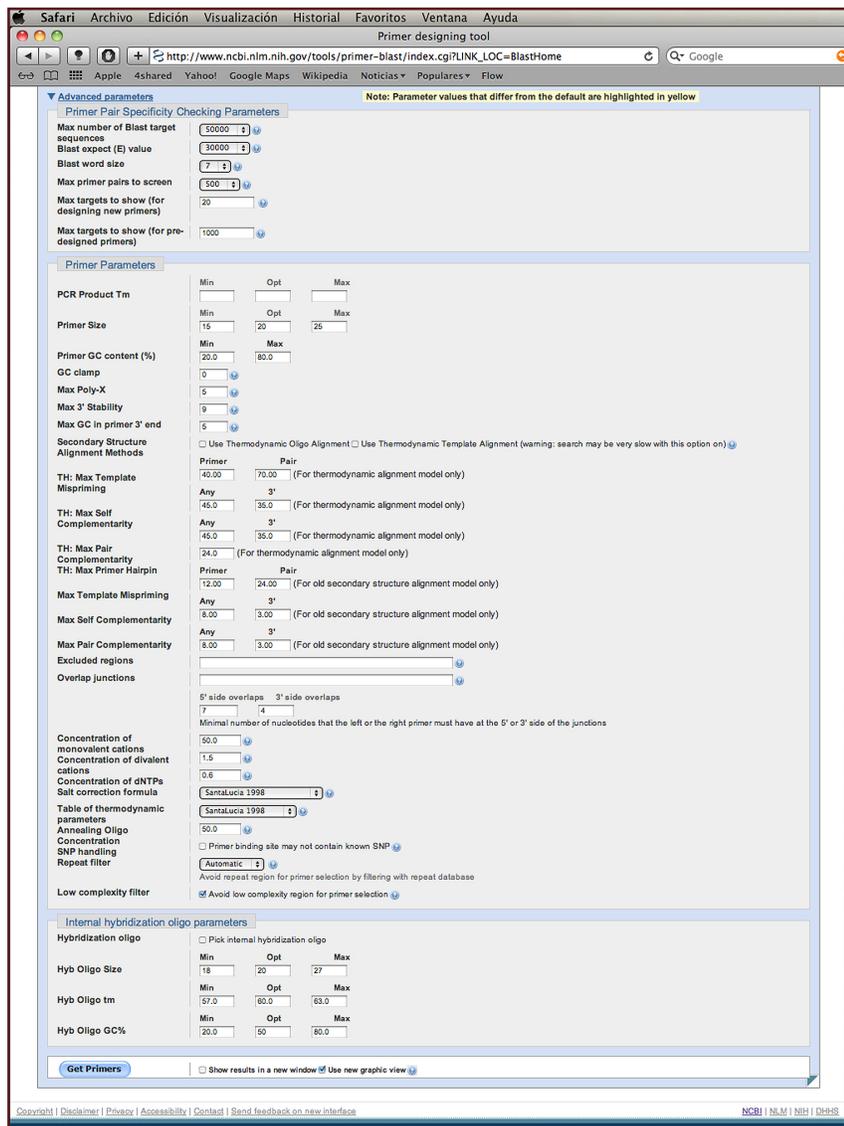


Figura 6.4c. Ventana de característica Avanzadas.

Actividad para reportar

Utilizando el programa Primer3 del MIT, obtenga 5 pares de iniciadores que podría utilizar para hacer la amplificación del receptor estrogénico beta del ratón a partir de ADN. Localice dentro del mRNA producto de la transcripción del gen cada uno de los pares de iniciadores sugeridos. Realice un análisis similar utilizando la herramienta Primer-BLAST.

Construya el perfil térmico que debería programar en el Termociclador para cada par de iniciadores obtenido en cada uno de los programas suponiendo que utiliza Taq polimerasa como enzima de amplificación. De acuerdo a las características anotadas en cada uno de los programas relacionadas a la concentración de iones monovalentes y bivalentes, indique la concentración de $MgCl_2$ y dNTP necesarias para la reacción de amplificación.

Referencias

-  SantaLucía, J. 2007. Physical principles and visual OMP software for optimal PCR design. En: Methods in Molecular Biology vol 40, PCR primer design (A. Yurev, ed). Humana Press (Totowa, NJ): 3-33.
-  Untergrasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G. 2012. Primer3- new capabilities and interfaces. Nucleic Acids res. 40:e115. Disponible en la URL <http://nar.oxfordjournals.org/content/40/15/e15.full.pdf+html>
-  Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache I., Rozen, S., Madden, T.L. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13:134 Disponible en el URL <http://www.biomedcentral.com/content/pdf/1471-2105-13-134.pdf>

Referencias generales

-  Baxevanos AD, Ouellette BFF. 2002. Bioinformatics: A practical guide to the analysis of genes and proteins. Jonh Wiley and Sons (New York).
-  McEntyre A.L. 1998. Coffe break: tutorials for NCBI tools. Nat. Academic Press (Washington).
-  Olson S, Beachy SH, Giammaria CF, Berger AC. 2002. Integrating large-scale genomic information into clinical practice. Nat. Academic Press (Washington).
-  Pool R, Esnayra J. 2000. Bioinformatics: Converting data into knowledge. Nat. Academic Press (Washington).

Manual de prácticas de laboratorio. Introducción a la Bioinformática

Se terminó de imprimir en septiembre de 2013,
con un tiraje de 200 ejemplares, más sobrantes para reposición.



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

UNIDAD IZTAPALAPA

División de Ciencias Biológicas y de la Salud

Av. San Rafael Atlixco No.186, Col. Vicentina
C.P. 09340, Del. Iztapalapa, México D.F.
Tel.: (01) 58044600