



Bioinformática. **José R. Valverde**

**José R. Valverde**  
**EMBnet/CNB, CSIC**

[www.es.embnet.org](http://www.es.embnet.org)  
[sci.cnb.csic.es](http://sci.cnb.csic.es)



**Bases de datos:  
Ejemplos prácticos**

2005-12-30

**Bioinfo 2008**  
Curso Doctorado UAM

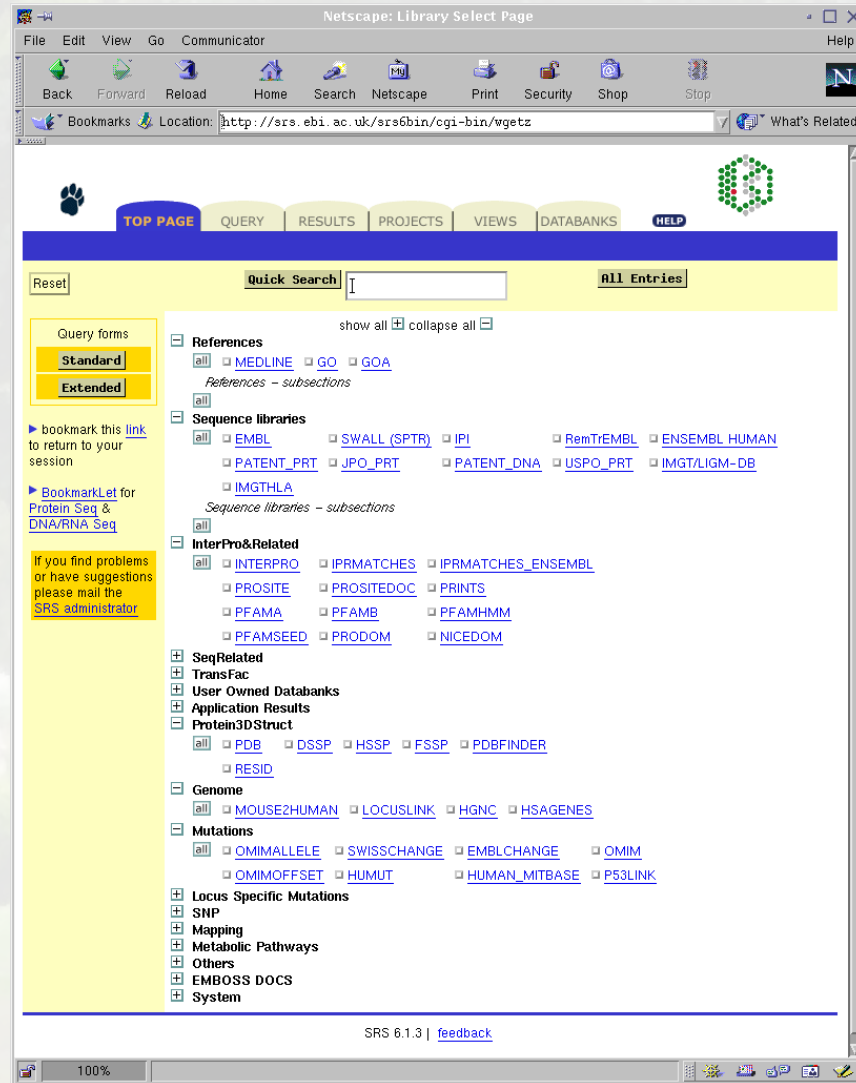
(C) José R. Valverde, 2000-2008 CC-BY.



**srs.ebi.ac.uk**  
**ncbi.nlm.nih.gov**  
**www.biomart.org**  
**mrs.cmbi.ru.nl/cgi-bin/mrs.cgi**



# Temática



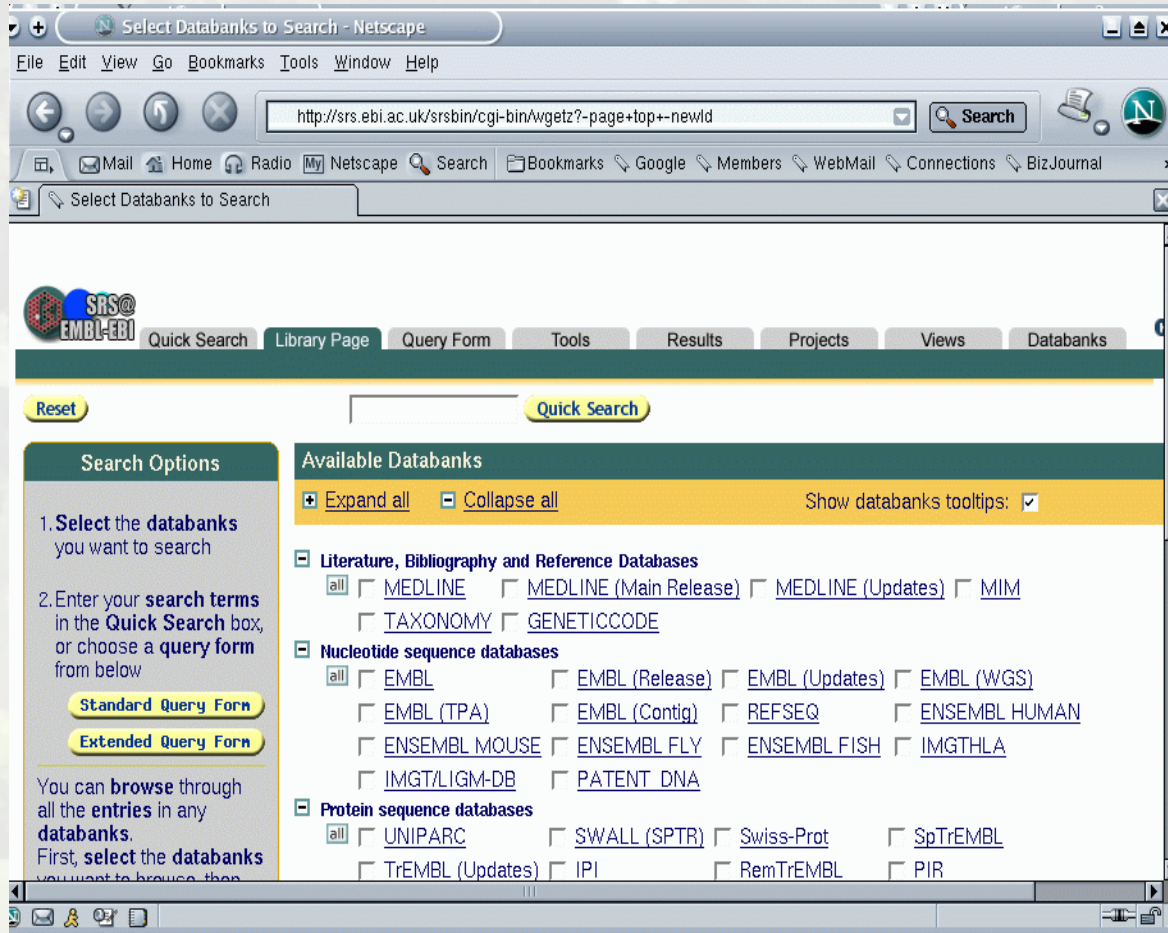
Existen cientos de bases de datos de interés biológico

Constantemente aparecen nuevas bases de datos

La mayoría están especializadas en áreas concretas

Casi todas contienen referencias cruzadas a alguna otra base de datos con información relacionada

# Asunto de una base de datos



- Hacer click sobre el nombre de la base de datos
- ¿De qué trata OMIM?

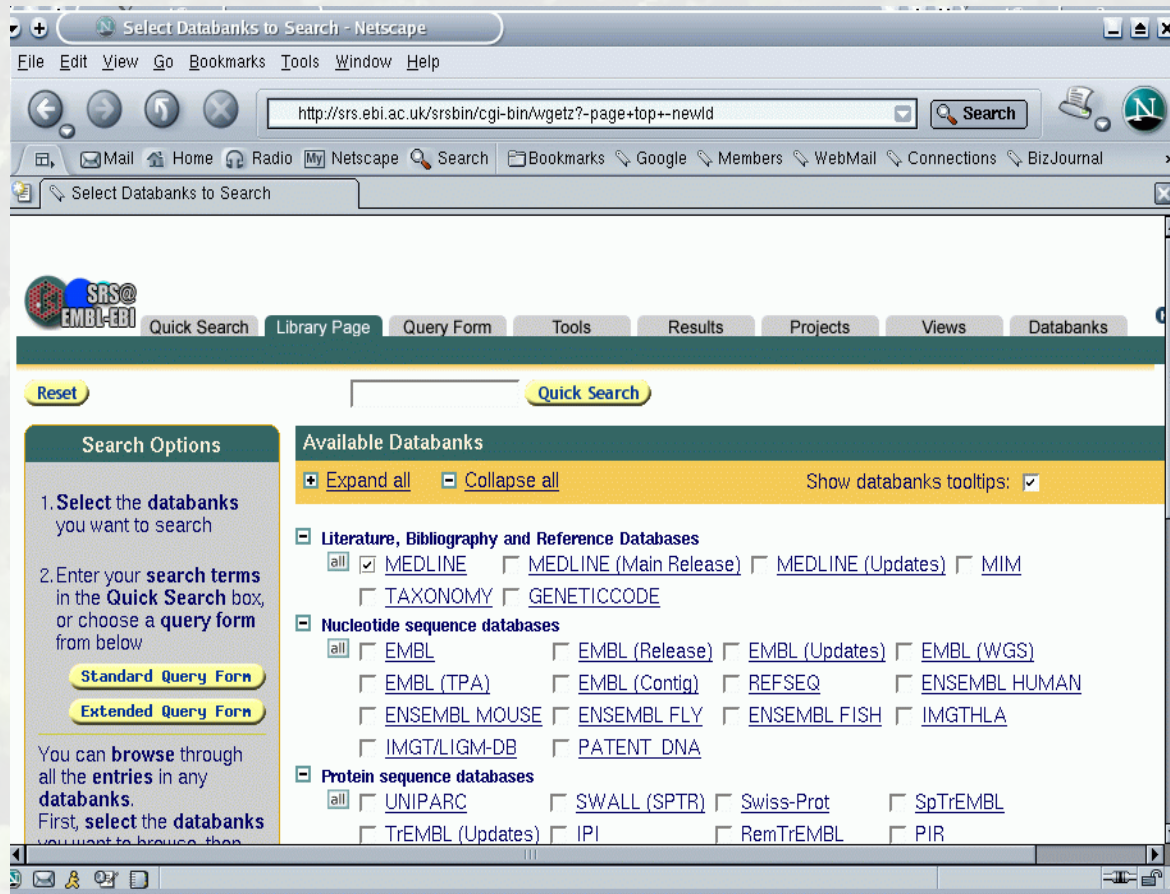
Expandir todo y usar la opción de búsqueda del navegador...



# Búsquedas en bases de datos

- Cada término de búsqueda produce como resultado un conjunto de registros que lo contienen
- La combinación de búsquedas se refleja en una combinación de los conjuntos resultantes
- Basadas en el álgebra de Boole o de conjuntos
  - Unión (O, OR, |,  $\vee$ )
  - Intersección (Y, AND, &,  $\wedge$ )
  - Exclusión (NO, NOT, !,  $\neg$ )

# Selección de una base de datos



- Hacer click sobre la caja de selección adyacente
- O elegir del menú desplegable



# Búsquedas rápidas

- ¿Qué entrada de **EMBL** hace referencia a **Fidel Castro** y uno de sus cigarros (**cigar**)?
- Recuperar la entrada **FUCK\_ECOLI** de **SwissProt**
  - Ojo al carácter “\_” (subrayado o guión bajo)
  - El "código de acceso" de SwissProt está dividido en dos partes:
    - Código de la proteína (3-4 letras)
    - Código del organismo (5 letras)

# Resultado de la búsqueda

- **Una lista de entradas como enlaces**
- Los enlaces llevan a una entrada (registro)
  - Compuesta por muchos campos
- Opciones adicionales
  - Guardar
  - Vistas predefinidas
    - Name only
    - FastSeqs
    - ProteinChart ...
  - Ordenar
  - Analizar



# Búsqueda por campos

The screenshot shows a Netscape browser window titled "Standard Query Form - Netscape" with the URL "http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz". The browser's address bar and menu bar are visible. The page content includes a navigation bar with links: "Quick Search", "Library Page", "Query Form" (active), "Tools", "Results", "Projects", "Views", and "Databanks". Below the navigation bar, there is a "Reset" button and a search input field containing "search Swiss-Prot".

The main content area is divided into two columns. The left column contains "Search Options" and "Result Display Options". The "Search Options" section includes a "Combine search terms" dropdown set to "& (AND)", a "Use wildcards" checkbox checked, and a "Get results of type:" dropdown set to "Entry". The "Result Display Options" section includes a "View results using:" dropdown set to "SeqSimpleView".

The right column contains a table titled "Fields you can search" and "Your search terms". The table has two columns: "Fields you can search" and "Your search terms". The first row shows "Species" selected in the dropdown, with "homo" entered in the search term field. The second, third, and fourth rows show "AllText" selected in the dropdown, with empty search term fields. Below the table, there is a "Create a view" section with a yellow background, containing the text "Select the fields you want displayed in your view and choose the format". Below this, there is a "Choose 1 or more fields:" section and a "Display As:" section with radio buttons for "Table" (selected) and "List".

A "Search" button is located at the bottom right of the table.

- Elegir formulario **“Standard”**
- Elegir campo de búsqueda
- Introducir términos
- Elegir combinación (**AND, OR, NOT**)
- Hacer click en **“Search”**

# Búsquedas sencillas

- ¿Qué es una “cremallera de leucinas”?
  - Expandir la sección “Protein Function”
  - Seleccionar base de datos **PROSITE**
  - Buscar “**leucine zipper**”
  - Visualizar entrada como “**PrositesSimple**” (en “Display Options”).
- ¿Cuántas secuencias de humano (“**homo**”) hay en SwissProt? ¿y en EMBL?
  - Indicar campo “Organismo”
- ¿Hay alguna secuencia de Neanderthal (**homo sapiens neanderthalensis**) en EMBL?



# Búsquedas avanzadas

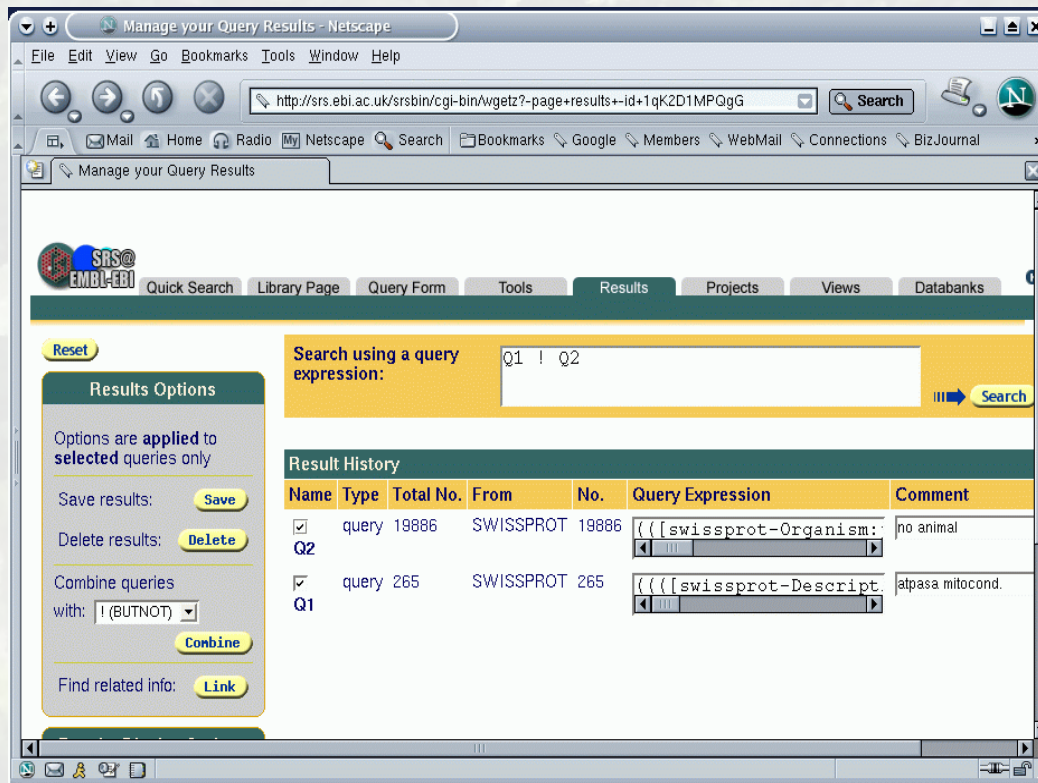
- ¿Cuántas secuencias de homínidos (**hominidae**) no humanos (**homo**) hay en Swissprot?
- Buscar la fuculokinasa de Coli en Swissprot
  - Con una búsqueda rápida
  - Indicando los campos (Descripción, Organismo)
  - ¿Qué está ocurriendo?
  - ¿Cómo podemos arreglarlo?
    - Volvera obtener FUCK\_ECOLI
    - Buscar \*fuculokinase
    - Buscar EC 2.7.1.51

# Búsquedas bibliográficas (**MEDLINE**)

- ¿Qué hizo bien (**right**) Bush con un salto de paracaídas (**parachute jump**)?
- ¿Puede uno suicidarse (**suicide**) con un **enema** de aspirinas (**aspirin**)?
- ¿Qué tienen que ver la Trinidad (**Trinity**) y los castigos corporales (**corporal punishment**)?
- ¿Qué se publicó en Lancet, 1998; vol. 352: 2010-2011?



# Combinar búsquedas



- Elegir “Results”
- Seleccionar búsquedas
- Combinarlas
  - Usando el menú o
  - Escribiendo la expresión

# Navegación por resultados

- **Buscar “ritualistic cannibalism” en SwissProt**
  - ¿Cuál es la página W.W.W. oficial de la enfermedad?
  - Acceder a la secuencia de DNA (**EMBL**)
  - Obtener la descripción de la enfermedad en **MIM**
  - Obtener la “firma” de la proteína en **ProSite**
    - Obtener la descripción de la “firma” (DO)
  - Obtener la estructura 3D de **PDB**
    - Guardar la estructura completa en un fichero



**Esto sonará a Chino**

¡De momento!

¡Probad de nuevo más adelante en el  
curso!

# Análisis de resultados

- Buscar en SwissProt la entrada **las1\_human**
- Seleccionarla y lanzar **BlastP** contra **PDB**
- Pasar a la página de resultados
- Esperar hasta que termine y visualizar siguiendo el enlace
- Seleccionar las cuatro primeras entradas y realizar un alineamiento múltiple sobre las entradas seleccionadas usando **ClustalW**
- ¿Cuáles podemos usar como moldes para construir un modelo de homología?
- <http://www.es.emblnet.org/Services/MolBio/Moldy>