Introducción a la Bioinformática

**Centro de Bioinformática
Instituto de Biotecnología
Universidad Nacional de Colombia**

**Andrés M. Pinzón
cPhD - Universidad de los Andes**

**7'000.000 de
habitantes**

**Capital Mundial del
Libro 2007**

**2600m sobre el nivel
del mar**

“Actualmente en Biología el número de datos es mucho mayor a la capacidad de análisis del investigador”.

- 1958 Primera secuencia de una proteína (insulina. Sanger, Premio Nobel)¹**
- 1975 Primera secuencia de ADN (fago PHI-X174 Sanger, Nobel 1980)**

1986 desarrollo de PCR

... inicio de la era genómica!!

1

<http://www.genetics.org/cgi/content/full/162/2/527>

BD para las ciencias de la vida

- **DNA**
- **Proteínas**
- **Genómicas**
- **Dominios/familias protéicas**
- **Mutación/polimorfismo**
- **Proteómica (2D gel, MS)**
- **3D estructura**
- **Metabolismo**
- **Bibliografía**
- **SNPs, ESTs, Microsatélites, etc..**
-

Algunas estadísticas

- Más de 1000 bases de datos
- Generalmente accesibles vía WEB
 - Biohunt: <http://www.expasy.org/BioHunt/>
 - Amos' links: www.expasy.ch/alinks.html
- Tamaño variable: 100Kb a 100Gb
 - DNA (EMBL): > 68 Gb (comprimida!)
 - 80,591,891 entradas.
 - 146,595,277,574 nucleótidos.

Crecimiento de la base datos EMBL en gigabases.

Crecimiento de la base datos EMBL en millones de entradas.

Sin embargo...

Tenemos las secuencias pero...

¿Como podemos saber qué partes de ese DNA controlan los diversos procesos químicos de la vida?

Conocemos la función y estructura de algunas proteínas pero...

Cómo determinamos nuevas funciones?
Como predecir la conformación espacial de una proteína, basados únicamente en su secuencia?

Entendemos el código genético pero...

Como encontrar nuevas palabras significativas que podamos añadir al diccionario del DNA?

BIOINFORMÁTICA

Biología Computacional
Aplicación de técnicas analíticas y cuantitativas para el modelamiento de sistemas biológicos.

"La bioinformática comprende los métodos matemáticos, estadísticos y computacionales que pretenden solucionar problemas biológicos usando secuencias de ADN y aminoácidos e información relacionada".

Fredj Tekaiia - Instituto Pasteur

"La bioinformática es el estudio de la información biológica desde su almacenamiento en el genoma hasta la obtención de los productos génicos en la célula" esto involucra la creación y desarrollo de tecnologías informáticas y computacionales para la resolución de problemas en biología molecular".

Stanford Center for Profesional Development,

2002

Nuestra definición...

Uso de técnicas computacionales, matemáticas y estadísticas para el análisis, interpretación y **generación de datos biológicos.**

CARACTERÍSTICAS:

- INTERDISCIPLINA Y COLABORACIÓN ENTRE GRUPOS.
- INTEROPERATIVIDAD E INTERDEPENDENCIA DE LOS DATOS.
- FORMACIÓN DE REDES.

-
- **Quiénes?**
 - **Para qué?**
 - **Qué?**
 - **Cómo?**

¿Qué personas pueden hacer Bioinformática?

Necesario:

- Conocimiento y entendimiento del Dogma Central de la Biología molecular.
- Conocimiento en Biología Molecular (bioquímica, biología molecular, biofísica molecular).

Muuuuuuuuuy recomendado:

- Conocimiento en el manejo de sistemas de cómputo.

Recomendado:

- Manejo básico de línea de comandos en ambientes UNIX (GNU/Linux).

Muy deseable:

- Experiencia con algún lenguaje de programación.

¿Qué se busca con el uso de la Bioinformática?

“Profundizar en nuestro entendimiento acerca de los organismos vivos y sus relaciones, partiendo desde el genoma que les codifica”.

Los análisis potenciales en el campo de la Biología molecular son tan diversos como la vida misma.

- Genómica comparativa.
- Análisis de DNA (ORFs, Contenidos GC, etc).
- Recuperación de secuencias.
- Ensamblaje de secuencias.
- Predicción de estructuras protéicas.
- Visualización de estructuras protéicas.
- Microarreglos.
- PCR.
- Filogenia.
- Educación.

La bioinformática provee algoritmos, bases de datos, interfaces y herramientas estadísticas para resolver nuestras preguntas!

Bioinformática y “ómicas”

- **Genómica:** Estudia el genoma de los organismos. Uso sistemático de la información genómica y su asociación con otros datos.
- **Transcriptómica:** transcriptoma es el conjunto de todos los mRNAs (o transcritos), dado un conjunto de condiciones externas. El genoma es fijo (ok, existe polimorfismo genético!) el transcriptoma varía dependiendo del contexto.
- **Proteómica:** es el estudio a gran escala de las proteínas, particularmente su estructura y función.
- **Metabolómica:** estudia el conjunto completo de metabolitos (intermediarios metabólicos, hormonas, metabolitos secundarios etc.) encontrados en una muestra biológica, por ejemplo un organismo.

Genómica comparativa

Alineamiento de genes de diferentes especies buscando secuencias en común.

- Relaciones evolutivas.
- Clasificación de familias de genes.
- Tiempos de divergencia.

Genómica comparativa

Species	Similarity	Source
Human	99.9%	quoted by U.S.A. President Clinton, Jan 2000, State of the Union address; also Human Genome Project
	100%	identical twins
Chimpanzee	98.4%	sources: Americans for Medical Progress ; Jon Entine in the San Francisco Examiner
	98.7%	Richard Mural of Celera Genomics , quoted on MSNBC
Bonobo		equal to chimpanzee
Gorilla	98.38%	based on study of intergenic nonrepetitive DNA in <i>Am J Hum Genet.</i> (2001) Feb; 682 :444-56
Mouse	98%	source: Americans for Medical Progress
	85%	comparing all protein coding sequences, NHGRI
Dog	95%	Jon Entine in the San Francisco Examiner
<i>C. elegans</i>	74%	Jon Entine in the San Francisco Examiner
Banana	50%	source: Americans for Medical Progress
Daffodil	35%	Steven Rose in The Guardian 22 January 2004

Genómica comparativa

**Alineamiento de genes de iguales especies
buscando regiones en común.**

- Identificación de polimorfismos.
- Medicamentos personalizados (farmacogenética/farmacogenómica).
- Tratamiento de enfermedades a nivel genético.

Mismas herramientas ... preguntas diferentes!

Use Ensembl to...

- Run a BLAST search
- Search Ensembl
- Data mining [BioMart]
- Export data
- Download data

Docs and downloads

- Information
- What's New
- About Ensembl
- Ensembl data
- Software

Other links

- Home
- Sitemap
- Vega
- Pre Ensembl
- View previous release of page in Archive!
- Stable Archive! link for this page
- Archive! sites
- Trace server

What's New in Ensembl 39

- ▶ [Honeybee dropped from Ensembl](#) (*Apis mellifera*)
- ▶ [New Opossum assembly and genebuild](#) (*Monodelphis domestica*)
- ▶ [New Mouse assembly and genebuild](#) (*Mus musculus*)
- ▶ [New Ciona savignyi assembly and genebuild](#) (*Ciona savignyi*)
- ▶ [New Zebrafish assembly and genebuild](#) (*Danio rerio*)

[More news...](#)

About Ensembl

Ensembl is a joint project between [EMBL - EBI](#) and the [Sanger Institute](#) to develop a software system which produces and maintains automatic annotation on selected eukaryotic genomes. Ensembl is primarily funded by the [Wellcome Trust](#).

This site provides ~~free access~~ to all the data and software from the Ensembl project. Click on a species name to browse the data.

Access to all the data produced by the project, and to the software used to analyse and present it, is provided free and without constraints. Some data and software may be subject to [third-party constraints](#).

For all enquiries, please [contact the Ensembl HelpDesk](#) (helpdesk@ensembl.org).

Other sites using the Ensembl system

- ▶ [EBI Genome Reviews](#) database - mainly archaea and bacteria.
- ▶ [VEGA](#) - Vertebrate Genome Annotation

[More...](#)

Mammalian genomes

- [Homo sapiens](#)
NCBI 36 | Vega
- [Pan troglodytes](#)
PanTro 1.0
- [Macaca mulatta](#)
MMUL 0.1 | *pre!*
- [Mus musculus](#)
UPDATED! NCBI m36 | Vega
- [Rattus norvegicus](#)
RGSC 3.4
- [Oryctolagus cuniculus](#)
RABBIT
- [Canis familiaris](#)
CanFam 1.0 | Vega | *pre!*
- [Bos taurus](#)
Btau 2.0
- [Dasyurus novemcinctus](#)
ARMA
- [Loxodonta africana](#)
BROAD E1
- [Echinops telfairi](#)
TENREC
- [Monodelphis domestica](#)
UPDATED! MonDom 4

Other species

- [Gallus gallus](#)
WASHUC 1
- [Xenopus tropicalis](#)
JGI 4.1
- [Danio rerio](#)
UPDATED! Zv 6 | Vega
- [Takifugu rubripes](#)
FUGU 4.0
- [Tetraodon nigroviridis](#)
TETRAODON 7
- [Gasterosteus aculeatus](#)
BROAD S1
- [Ciona intestinalis](#)
JGI2
- [Ciona savignyi](#)
NEW! CSAV 2.0
- [Drosophila melanogaster](#)
BDGP 4
- [Anopheles gambiae](#)
AgamP3
- [Aedes aegypti](#)
AEDES 1
- [Caenorhabditis elegans](#)
WS 150
- [Saccharomyces cerevisiae](#)
SGD 1

Genómica funcional

¿Qué genes se expresan, dónde, cuando, función?

Microarreglos

Alineamiento de secuencias

Sec. 1 MALVEDNNAVAVSFLRNSFSEEQILVKYGV**VDG**

Sec. 2 HKSF**EDNNAVAVS**SEFALLEQILVK**DG**

Recuperación de secuencias

NCBI
National Center for Biotechnology Information
National Library of Medicine National Institutes of Health

PubMed All Databases BLAST OMIM Books TaxBrowser Structure

Search All Databases for Go

SITE MAP
Alphabetical List
Resource Guide

About NCBI
An introduction to NCBI

GenBank
Sequence submission support and software

Literature databases
PubMed, OMIM, Books, and PubMed Central

Molecular databases
Sequences, structures, and taxonomy

Genomic biology
The human genome, whole genomes, and related resources

What does NCBI do?
Established in 1988 as a national resource for molecular biology information, NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease. [More...](#)

Hot Spots
Assembly Archive

Whole Genome Association
The NCBI Whole Genome Association (WGA) resource provides researchers with access to genotype associated phenotype information that will help elucidate the link between genes and disease. For more information, click here to see the WGA resource page and click here to read the press release.

CCDS Database
The new Consensus CDS (CCDS) project is a collaborative effort to identify a core set of human protein coding regions that are consistently annotated and of high quality. [More...](#)

PubMed Central
An archive of life sciences journal articles
● Free fulltext
● Over 500,000 articles from over 200 journals
● Linked to PubMed and fully searchable
Use of PubMed Central requires no registration or fee.

LION SRS

Quick Searches Predefined Queries Query Builder Analysis Tools

Quick Searches

Start a permanent link

Quick Text Search

Find: in: Nucleic Acid Sequence

ID Search

Search the following databank: EMBL

for the following identifiers:

or specify a file containing identifiers: Examiner...

Searching: EMBL

Sequence Similarity/Homology Search

Search for sequences of type: Nucleotide Sequences

matching:

(Paste in your sequence here)

Searching: EMBL (Release)

EMBL-EBI
European Bioinformatics Institute

EBI Home About EBI Groups Services Toolbox Databases Downloads Submissions

EMBL-NUCLEOTIDE SEQUENCE DATABASE

- Index
- Access
- Documentation
- News
- Submission
- Publications
- People
- Contact

EMBL Nucleotide Sequence Database

The EMBL Nucleotide Sequence Database (also known as EMBL-Bank) constitutes Europe's primary nucleotide sequence resource. Main sources for DNA and RNA sequences are [direct submissions](#) from individual researchers, genome sequencing projects and patent applications.

The database is produced in an international [collaboration](#) with GenBank (USA) and the DNA Database of Japan (DDBJ). Each of the three groups collects a portion of the total sequence data reported worldwide, and all new and updated database entries are exchanged between the groups on a daily basis. The [current database release](#) (Release 85, Dec 2005), with according [Release notes](#) and [user manual](#) are available from the EBI servers. A sample database entry is shown [here](#).

A publication in [Nucl. Acids Res., 2006, Vol. 34: D10-D13](#) provides further information and details.

The EMBL nucleotide sequence database group is headed by:
Rolf Apweiler.

Forthcoming EMBL Database Changes:

- Starting from the EMBL release 87 (June 2006) the naming of the release files will change...more

Link	Explanation
Access	Database queries, Completed genomes webserver, FTP archives (EMBL release, alignments etc), EMBL sequence version archive (SVA).

INSDC

International Nucleotide Sequence Database Collaboration

EMBL Fetch

Fetch an EMBL record by accession number

TPA

THIRD PARTY ANNOTATION

Users can now submit re-annotations/re-assemblies of sequences already present in EMBL and owned by other groups.

NCBI

Search

Search

for

SITE MAP

[Alphabetical List](#)
[Resource Guide](#)

About NCBI

[An introduction to NCBI](#)

GenBank

[Sequence submission support and software](#)

Literature databases

[PubMed, OMIM, Books, and PubMed Central](#)

Molecular databases

[Sequences, structures, and taxonomy](#)

Genomic biology

[The human genome, whole genomes, and related resources](#)

What does NCBI do?

Established in 1988 as a national resource for molecular biology information, NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease. [More...](#)

Whole Genome Association

The NCBI Whole Genome Association (WGA) resource provides researchers with access to genotype and associated phenotype information that will help elucidate the link between genes and disease. For more information, click here to see the the [WGA resource page](#) and click here to read the [press release](#).

CCDS Database

The new Consensus CDS (CCDS) project is a collaborative effort to identify a core set of human protein coding regions that are consistently annotated and of high quality. [More...](#)

PubMed Central

An archive of life sciences journals

- **Free fulltext**
- **Over 500,000 articles from over 200 journals**
- **Linked to PubMed and fully searchable**

Use of PubMed Central requires no registration or fee.

Hot Spots

- ▶ [Assembly Archive](#)
- ▶ [Clusters of orthologous groups](#)
- ▶ [Coffee Break, Genes & Disease, NCBI Handbook](#)
- ▶ [Electronic PCR](#)
- ▶ [Entrez Home](#)
- ▶ [Entrez Tools](#)
- ▶ [Gene expression omnibus \(GEO\)](#)
- ▶ [Human genome resources](#)
- ▶ [Malaria genetics & genomics](#)
- ▶ [Map Viewer](#)
- ▶ [dbMHC](#)

- Index
- Access
- Documentation
- News
- Submission
- Publications
- People
- Contact

EMBL Nucleotide Sequence Database

The EMBL Nucleotide Sequence Database (also known as EMBL-Bank) constitutes Europe's primary nucleotide sequence resource. Main sources for DNA and RNA sequences are [direct submissions](#) from individual researchers, genome sequencing projects and patent applications.

The database is produced in an international [collaboration](#) with GenBank (USA) and the DNA Database of Japan (DDBJ). Each of the three groups collects a portion of the total sequence data reported worldwide, and all new and updated database entries are exchanged between the groups on a daily basis. The [current database release](#) (Release 85, Dec 2005), with according [Release notes](#) and [user manual](#) are available from the EBI servers. A sample database entry is shown [here](#).

A publication in [Nucl. Acids Res., 2006, Vol. 34: D10-D15](#) provides further information and details.

The EMBL nucleotide sequence database group is headed by:
[Rolf Apweiler](#).

Forthcoming EMBL Database Changes:

- Starting from the EMBL release 87 (June 2006) the naming of the release files will change...more

Link	Explanation
Access	Database queries , Completed genomes webserver , FTP archives (EMBL release, alignments etc), EMBL sequence version archive (SVA).

INSDC

International
Nucleotide Sequence
Database
Collaboration.

EMBL Fetch

Fetch an EMBL record
by accession number

TPA

Users can now submit
re-annotations/
re-assemblies of
sequences already
present in EMBL and
owned by other groups.

NCBI

<http://www.ebi.ac.uk/embl/>

Quick Searches

[Start a permanent project](#)

C
u
s
t
o
m
i
z
e

Quick Text Search ?

Find:

in:

Search

ID Search ?

Search the following databank:

for the following identifiers:

or specify a file containing identifiers:

Searching: EMBL

Search

Sequence Similarity/Homology Search ?

Search for sequences of type:

matching :

(Paste in your sequence here)

Searching: EMBL (Release)

Search

<http://srs.ibun.unal.edu.co:8080/srs81/>

Visualización de estructuras protéicas

¿Visualización o predicción?

GARLI
C

¿Predicción o determinación?

MLDTNMKTQLKAYLEKLT KPVELIATLDDSAKSAEIKELL
TKPVELIATLDDSAEIKELLSAIAMLKSAEIKELLA EIKELL
LKAYLEKLT KPMLDTNMKTQLKKS AEIKELLKSAEIKELL

Predicción
2D

- Modelamiento homólogo (modelamiento comparativo).
- Métodos **Ab initio**

Cristalografía
de
Rayos X

Análisis Filogenéticos

- Evolución molecular de familias de proteínas
- Creación de árboles taxonómicos
- Reconstrucción evolutiva de rutas metabólicas.

- QuickTree: Reconstrucción de árboles filogenéticos.
- Phylip (Phylogenie Inference Package).
- Mavric (Python): Manipulación y visualización de árboles.
- PAUP.

PAUP*
Sinauer Associates, Inc. Publishers
Sunderland, Massachusetts

Recent News

- Beta 10 is the most recent release. Click [here](#) to get an updater.
- There are several known problems with beta 10. We plan to release an update to fix these problems soon. For a list of known problems and possible work-arounds click [here](#).
- The paupinfo email list has moved. To see if you are subscribed to the list and to modify your personal settings go to the [paupinfo email list](#) web site.

Ensamblaje de Secuencias

EST clustering

EMBNet 2002

EST clustering pipeline

Nuestra Bioinformática

The screenshot shows the EMBnet Colombia website. At the top, there is a blue header with the EMBnet Colombia logo and a molecular model. Below the header is a navigation bar with links: Inicio, Quienes somos, Documentos, Cursos, Contacto, and Mapa del Sitio. The date 8 de marzo de 2006 is displayed on the right. The main content area features the Universidad Nacional de Colombia logo and a grid of blue dots. The title "EMBnet Colombia" is centered. The main text describes EMBnet Colombia as the Colombian node of the Red Europea de Biología Molecular (Embnet), a global network of bioinformatics centers. It is managed by the Centro de Bioinformática del Instituto de Biotecnología (CBIB) at the Universidad Nacional de Colombia. A small graphic of two overlapping squares is centered below the text. The footer of the website reads "Centro de Bioinformática - Instituto de Biotecnología 2005 Universidad Nacional de Colombia".

EMBnet Colombia

Inicio :: Centro de Bioinformática del Instituto de Biotecnología :: Mapa del Sitio

Inicio | **Quienes somos** | Documentos | Cursos | Contacto | 8 de marzo de 2006

UNIVERSIDAD NACIONAL DE COLOMBIA

Herramientas

| [EMBOSS](#) | [BLAST](#) | [BLEE](#) | [SMS SEWER](#) | [BACTNAME](#) | [SINCO](#) | [DOTLET](#)

Puede consultar nuestra página de ayuda para mayor [información acerca de estas herramientas](#).

Curso Microarreglos

Del 28 de Noviembre al 2 de Diciembre de 2005 tendrá lugar el curso Internacional: "Introducción al **Análisis de Microarreglos**", organizado por el CBIB en colaboración con varios grupos e institutos a nivel mundial. [Mayor información](#)

Curso Bioinformática

EMBnet Colombia

EMBnet Colombia es el nodo Colombiano de la **Red Europea de Biología Molecular (Embnet)**, una red mundial de centros en bioinformática con nodos en los 5 continentes. EMBnet Colombia se encuentra a cargo del **Centro de Bioinformática del Instituto de Biotecnología (CBIB)** de la Universidad Nacional de Colombia.

Centro de Bioinformática - Instituto de Biotecnología 2005
Universidad Nacional de Colombia

<http://bioinf.ibun.unal.edu.co>

Herramientas

[EMBOSS](#) | [BLAST](#) | [BLEE](#) | [SMS SEWER](#) | [BACTNAME](#) | [SINCO](#) | [DOTLET](#)
Completa suite con más de 100 herramientas para análisis moleculares.
Puede consultar nuestra página de ayuda para mayor [información acerca de estas herramientas](#).

Curso Microarreglos

Del 28 de Noviembre al 2 de Diciembre de 2005 tendrá lugar el curso Internacional: "Introducción al **Análisis de Microarreglos**", organizado por el CBIB en colaboración con varios grupos e institutos a nivel mundial.
[Mayor información](#)

Curso Bioinformatica

EMBnet Colombia

EMBnet Colombia es el nodo Colombiano de la **Red Europea de Biología Molecular (Embnet)**, una red mundial de centros en bioinformática con nodos en los 5 continentes. EMBnet Colombia se encuentra a cargo del **Centro de Bioinformática del Instituto de Biotecnología (CBIB)** de la Universidad Nacional de Colombia.

Centro de Bioinformática - Instituto de Biotecnología 2005
Universidad Nacional de Colombia

Biodiversidad y Bioinformática

SiB Sistema de Información sobre Biodiversidad de Colombia

Indicadores de conocimiento | Catálogo de organismos | Biodiversidad en cifras

Productos de información

¿Qué son los productos de información?

Los productos de información son recursos de información diseñados para una audiencia específica y con un propósito definido, y son resultado de la recopilación y presentación de información analizada e interpretada.

Conozca los productos de información que están disponibles actualmente:

- Biodiversidad en cifras
- Indicadores de conocimiento sobre biodiversidad
- Catálogo electrónico de los organismos presentes en Colombia
- Biota colombiana
- CITES Colombia

ENKI DB

Búsqueda rápida

Eudendrocyclus

Inicio | Blast | Búsqueda | Estadísticas | Registro

Inicio

ACERCA DE ENKI

ENKI mantiene un registro actualizado de toda la información taxonómica (especial) y molecular (DNA, proteínas) de especies colombianas presente en las más importantes bases de datos a nivel mundial.

ENKI es un proyecto financiado por Colciencias y liderado por el Centro de Bioinformática del Instituto de Biología, de la Universidad Nacional de Colombia, con la colaboración del Instituto de Ciencias Naturales y el Instituto Von Humboldt.

BLAST ENKI

BLAST ENKI es nuestra implementación del tradicional BLAST pero que hemos dedicado exclusivamente al registro y análisis de especies propias de la biodiversidad colombiana, de esta manera con BLAST ENKI es posible realizar comparaciones blast colombianas contra nuestra propia base de datos de especies colombianas.

SPICA

Generalidades | Documentación | Productos | Aplicaciones | Registro | Protocolo Web | Contactenos

Módulos de consulta remota

Nombre de la especie

Reino

Género

Reino

Familia

Reino

Usos

COMO CONSULTAR:

Los módulos de búsqueda son independientes, puede consultar por nombre de Especie, Género o Familia dando el criterio en el cuadro correspondiente y pulsando el botón "Iniciar Búsqueda". La búsqueda por usos se hará con respecto a especies que tengan asociado dicho uso.

Los nombres deben iniciar con mayúscula. Se puede escribir la palabra (o palabras) completa o parte de ella. La consulta por nombre de especie hace referencia al nombre vulgar o al

EMBL-EBI European Bioinformatics Institute

EBI Home | About EBI | Groups | Services | Toolbox | Databases | Downloads | Submissions

EMBL NUCLEOTIDE SEQUENCE DATABASE

EMBL Nucleotide Sequence Database

The EMBL Nucleotide Sequence Database (also known as EMBL-Bank) constitutes Europe's primary nucleotide sequence resource. Main sources for DNA and RNA sequences are direct submissions from individual researchers, genome sequencing projects and patent applications.

The database is produced in an international collaboration with GenBank (USA) and the DNA Database of Japan (DDJ). Each of the three groups collects a portion of the total sequence data reported worldwide, and all new updated database entries are exchanged between the groups on a daily basis. The current database release (Release 85, Dec 2005), with accompanying [Sequence User Manual](#) are available from the EBI servers. A sample database entry is shown [here](#).

A publication in [Nucleic Acids Res.](#), 2006, Vol. 34, D10-D13 provides further information and details.

The EMBL nucleotide sequence database group is headed by: **Rolf Apweiler**.

Forthcoming EMBL Database Changes:

- Starting from the EMBL release 87 (June 2006) the naming of the release files will change... more

Link | **Explanation**

[Access](#) | [Database queries](#) | [Consolidated genomes webserver](#) | [FTP archives](#) (EMBL release, alignments etc.) | [EMBL sequence version archive](#) (SVA)

UniProt the universal protein resource

Home | About UniProt | Getting Started | Searches/Tools | Databases | Support/Documentation

Welcome to UniProt

UniProt (Universal Protein Resource) is the world's most comprehensive catalog of information on proteins. It is a central repository of protein sequence and function created by joining the information contained in Swiss-Prot, TrEMBL, and PIR.

UniProt is comprised of three components, each optimized for different uses. The **UniProt Knowledgebase (UniProtKB)** is the central access point for extensive curated protein information, including function, classification, and cross-reference. The **UniProt Reference Clusters (UniRef)** databases combine closely related sequences into a single record to speed searches. The **UniProt Archive (UniParc)** is a comprehensive repository, reflecting the history of all protein sequences.

The sequences and information in UniProt are accessible via [text search](#), [BLAST similarity search](#), and [FTP](#).

[European Bioinformatics Institute](#) | [Swiss Institute of Bioinformatics](#) | [Georgetown University](#)

[About UniProt](#) | [Getting Started](#) | [Searches/Tools](#) | [Databases](#) | [Support/Documentation](#)

HOME | HELP | SITE MAP | © 2006 UniProt Consortium | License & Disclaimer | TERMS OF USE

Biodiversidad y Bioinformática

Busqueda rapida:

ej. Eleutherodactylus

[Inicio](#) | [Blast](#) | [Busqueda](#) | [Estadísticas](#) | [Registro](#)

Ayuda

- ▶ [Guia del Usuario](#)
- ▶ [Preguntas frecuentes](#)

Sugerencias

Contacto

::: Inicio

ACERCA DE ENKI

Enki mantiene un registro actualizado de toda la información **taxonomica (especie)** y **molecular (DNA, proteínas)** de especies colombianas presente en las mas importantes bases

de datos a nivel mundial.

Enki es un proyecto financiado por [Colciencias](#) y liderado por el [Centro de Bioinformatica del Instituto de Biotecnologia](#), de la [Universidad Nacional de Colombia](#), con la colaboracion del Instituto de Ciencias Naturales y el Instituto Von Humboldt.

BLAST ENKI

Blast Enki es nuestra implemetación del tradicional **ncbi BLAST** pero que hemos dedicado exclusivamente al registro y análisis de especies propias de la biodiversidad colombiana, de esta manera con **Blast**

Enki es posible realizar [comparaciones blast](#) cotidianas contra nuestra propia base de datos de especies colombianas.

33308 especies registradas.
413983 registros en Uniprot.
Corresponden a: 11779 (35.4%).

11212 registros en EMBL.

::: Consulta General

Consulte por orden alfabetico:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

o introduzca su termino de busqueda en la casilla que aparece a continuación:

Los siguientes son resultados que contienen su termino de busqueda:

Aa	Aaronsohnia
Abaeis	Abaeis nicippe
Abaeis nicippe	Abaeis nicippe ab. flava
Abaeis nicippe f. pallens	Abaeis nicippe tr. f. callae
Abaeis nicippe tr. f. dammersi	Abarema
Abatia	Abdominea
Abdulmajidia	Abelia
Abeliophyllum	Abelmoschus
Abies	Abobra
Abolboda	Aboriella
Abroma	Abromeiba
Abromeitiella	Abronia
Abrophyllum	Abrotanella
Abrus	Absolmsia
Aburria	Aburria
Aburria	Aburria aburri
Aburria aburri	Aburria aburri
Aburria pipile	Aburria pipile
Aburria pipile	Abuta
Abutilon	Abutilothamnus
Acacia	Acaena
Acalypha	Acalypha alopecuroides

Ayuda

- ▶ [Guia del Usuario](#)
- ▶ [Preguntas frecuentes](#)

Sugerencias

Contacto

::: ariopsis

Informacion Taxonomica:

Informacion Molecular:

Ayuda

- ▶ [Guia del Usuario](#)
- ▶ [Preguntas frecuentes](#)

Sugerencias

Contacto

Entradas en Uniprot (4)

Q70HS1 ▶ [blast] Cytochrome b (Fragment).	Q70HS2 ▶ [blast] Cytochrome b (Fragment).	Q85IQ8 ▶ [blast] NADH dehydrogenase subunit 1 (Fragment).	Q37137 ▶ [blast] Ribulose 1,5-bisphosphate carboxylase large subunit (EC 4.1....
--	--	--	---

No existen entradas en EMBL para ariopsis!

Algunos de nuestros proyectos...

Microsatellites in Phytophthora ESTs: Survey, transferability and association with pathogenesis related genes

Búsqueda e identificación de nuevos candidatos a vacuna contra la Malaria producida por *Plasmodium vivax*.

Búsqueda de dominios específicos para la clase Hexápoda (Phylum Artrópoda).

Modelo para la identificación de genes para Betalactamasas de espectro extendido.

Sequence Retrieve System Federation.

Bioinformática y Genética de la Conservación

...Preguntas?