



www.esaunggul.ac.id

PENGANTAR BIOINFORMATIKA

IBT 431

By Seprianto S.Pi, M.Si



Pertemuan 4

Pengenalan Data Base

Sasaran Perkuliahan

- Mahasiswa Dapat menjelaskan tentang GenBank (NCBI)
- Mahasiswa Memahami dan mampu menjelaskan berbagai akses online dalam kajian bioinformatika ENSEMBL, UniProt, Genbank, EMBL, SWISS-PROT

Pangkalan Data

- **Pangkalan data primer:** digunakan untuk menyimpan sekuens primer asam nukleat dan protein.
- **Pangkalan data sekunder:** digunakan untuk menyimpan motif sekuens protein,
- **Pangkalan data struktur:** digunakan untuk menyimpan data struktur protein dan asam nukleat.

Pangkalan data primer

- Pangkalan data primer untuk sekuens asam nukleat saat ini adalah
 - [GenBank](#) (Amerika Serikat),
 - [EMBL](#) (the European Molecular Biology Laboratory, Eropa), dan
 - [DDBJ](#) (DNA Data Bank of Japan, [Jepang](#)).
- Ketiga pangkalan data tersebut bekerja sama dan bertukar data secara harian untuk menjaga keluasan cakupan masing-masing pangkalan data.
- Sumber utama data sekuens asam nukleat adalah submisi (pengumpulan) langsung dari
 - **peneliti** individual,
 - proyek sekuensing [genom](#), dan
 - pendaftaran [paten](#).

Pangkalan data Sekunder

- Selain asam nukleat, beberapa contoh pangkalan data penting yang menyimpan sekuens primer protein adalah [PIR](#) (Protein Information Resource, Amerika Serikat), [Swiss-Prot](#) (Eropa), dan [TrEMBL](#) (Eropa).
- Ketiga pangkalan data tersebut telah digabungkan dalam [UniProt](#), yang didanai terutama oleh Amerika Serikat.
- Entri dalam UniProt mengandung informasi tentang sekuens protein, nama organisme sumber protein, pustaka yang berkaitan, dan komentar yang pada umumnya berisi penjelasan mengenai fungsi protein tersebut.

Pangkalan data primer: **Genbank**

- **Genbank**, dioperasikan oleh **NCBI** (*National Center for Biotechnology Information*)
- Penjelasan meliputi identifikasi suatu gen, produk gen, link informasi lain yang terkait dengan sumber database lain.
- NCBI berisi informasi dari sekuens DNA yang sama dengan sekuens DNA dalam **EMBL** (*European Molecular Biology Laboratory*) dan **DDBJ** (*DNA Data Bank of Japan*)

Pengenalan NCBI

- NCBI merupakan server yang memuat data base tentang informasi kesehatan dan bioteknologi (DNA, RNA dan Protein)
- Menyediakan berbagai macam software untuk analisis DNA, protein 3D, pencarian primer, pencarian conserve doomain dan lain sebagainya

National Center for Biotechnology Information

Search for

Resources

- NCBI Home
- All Resources (A-Z)
- Literature
- DNA & RNA
- Proteins
- Sequence Analysis
- Genes & Expression
- Genomes
- Maps & Markers
- Domains & Structures
- Genetics & Medicine
- Taxonomy
- Data & Software

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[More about the NCBI](#) | [Mission](#) | [Organization](#) | [Research](#) | [RSS](#)

Genome

1000 prokaryotic genomes are now completed and available in the Genome database.



1 2 3 4

How To...

Popular Resources

- PubMed
- PubMed Central
- Bookshelf
- BLAST**
- Gene
- Nucleotide
- Protein
- GEO
- Conserved Domains
- Structure
- PubChem

NCBI News

November and 02 Dec 2009

Tools NCBI

- GenBank
- BioSystems
- Database of Expressed Sequence Tags (dbEST)
- Database of Genome Survey Sequences (dbGSS)
- BLAST (Basic Local Alignment Search Tool)

Tools Bioinformatika : BLAST

- Perangkat bioinformatika yang berkaitan erat dengan penggunaan pangkalan data sekuens Biologi ialah BLAST (*Basic Local Alignment Search Tool*).
- Penelusuran BLAST (*BLAST search*) pada pangkalan data sekuens memungkinkan ilmuwan untuk mencari sekuens baik asam nukleat maupun protein yang mirip dengan sekuens tertentu yang dimilikinya.
- Hal ini berguna misalnya untuk menemukan gen sejenis pada beberapa organisme atau untuk memeriksa keabsahan hasil sekuensing atau untuk memeriksa fungsi gen hasil sekuensing.
- Algoritma yang mendasari kerja BLAST adalah penyejajaran sekuens.

Pangkalan data primer: **OMIM**

- **OMIM**, (*Online Mendelian Inheritance in Man—woman*), adalah insiklopedia gen-gen manusia dan penyakit genetik, merupakan penghubung untuk *entry* gen pada GenBank dan literatur ilmiah pada PubMed, berisi informasi berbagai gen manusia komplit dan paling baru.

Pangkalan data primer: **PDB**

- **PDB** (*Protein Data Bank*) berisi semua publisitas yang ada secara eksperimen telah dideterminasi (oleh *x-ray crystallography* dan *NMR*) sebagai model structural proteins dan asam nukleat.
- Tidak berisi model homologi atau tipe model teoritis lainnya.

Pangkalan data primer: PubMed

- Diskripsi pada Wikipedia sebagai “suatu kebebasan mengakses sititasi database MEDLINE dan abstrak artikel riset biomedik.
- Subjek utama adalah riset di bidang kedokteran, dan PubMed juga mempublikasi bidang yang terkait dengan bidang kedokteran, seperti kebidanan dan disipiin kesehatan lainnya.
- Hal ini secara menyeluruh meliputi keilmuan yang berhubungan dengan ilmu seperti biokemia dan biologi sel.
- Situs ini ditawarkan oleh *the United States National Library of Medicine* di *the National Institutes of Health* sebagai bagian dari *the Entrez information retrieval system*.”

Pangkalan data primer : UniProt *Knowledgebase*

- **UniProt *Knowledgebase* (Swiss-Prot and TrEMBL)**, dioperasikan oleh **SIB** (*Swiss Institute of Bioinformatics*) dan **EBI** (*European Bioinformatics Institute*), berisi sebagian besar publikasi yang ada berupa **sekuens protein** (bukan DNA atau RNA).
- Sekuens dalam *Swiss-Prot* dijelaskan secara manual dan menyediakan atau menghubungkan pengguna dengan semua informasi publisitas yang berisi sekuens tersebut.
- Sequences pada TrEMBL dikoleksi dan dijabarkan secara otomatis dari sekuens database, dan akan membuat jalannya menuju Swiss-Prot, tetapi tidak hanya setelah mereka secara manual menjabarkan *Swiss-Prot standards*.

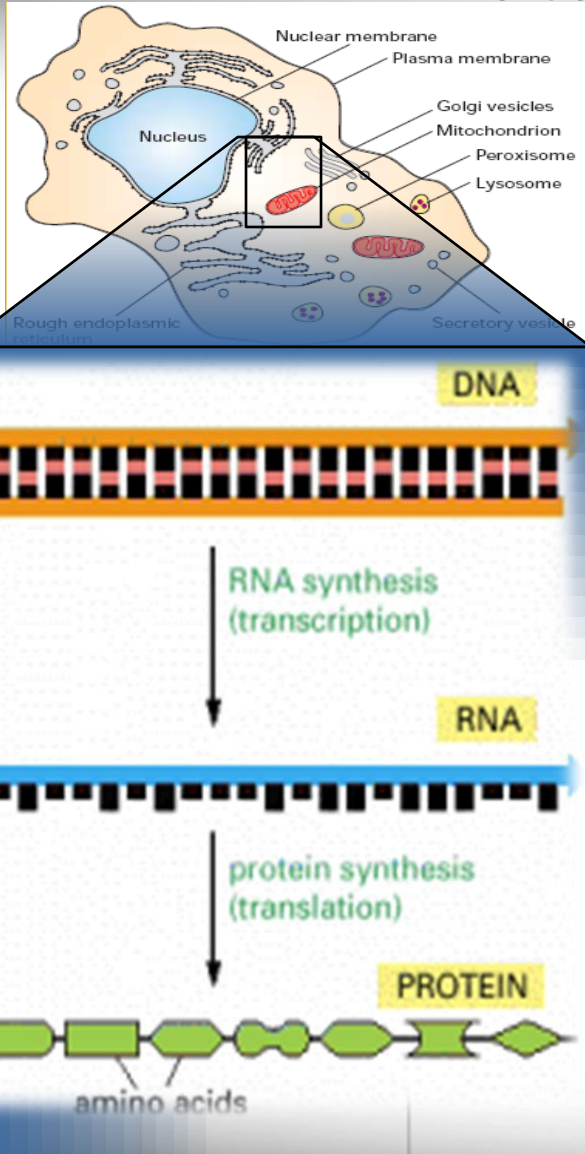
Situs Informasi database

- NCBI: www.ncbi.nlm.nih.gov
- EMBL: www.ebi.ac.uk
- DDBJ: www.ddbj.nig.ac.jp
- SWISS-
PROT: www.expasy.ch/sprot/sprot_details.html
- ENSEMBL: www.ensembl.org
- Univeristy California Santa Cruz:
genome.cse.ucsc.edu
- MGD the Jackson Lab: www.informatics.jax.org

Tools Bioinformatika : PDB

- PDB (*Protein Data Bank*, Bank Data Protein) ialah pangkalan data tunggal yang menyimpan model struktur tiga dimensi protein dan asam nukleat hasil penentuan eksperimental (dengan kristalografi sinar-X, spektroskopi NMR, dan mikroskopi elektron). PDB menyimpan data struktur sebagai koordinat tiga dimensi yang menggambarkan posisi atom-atom dalam protein atau pun asam nukleat.

Data & Website Bioinformatika



Sel

J.Craig Venter

Genome

NCBI

Transcrip-
tome

EMBL

DDBJ

Proteome

Expasy

PDB

PIR



Program
/Tool

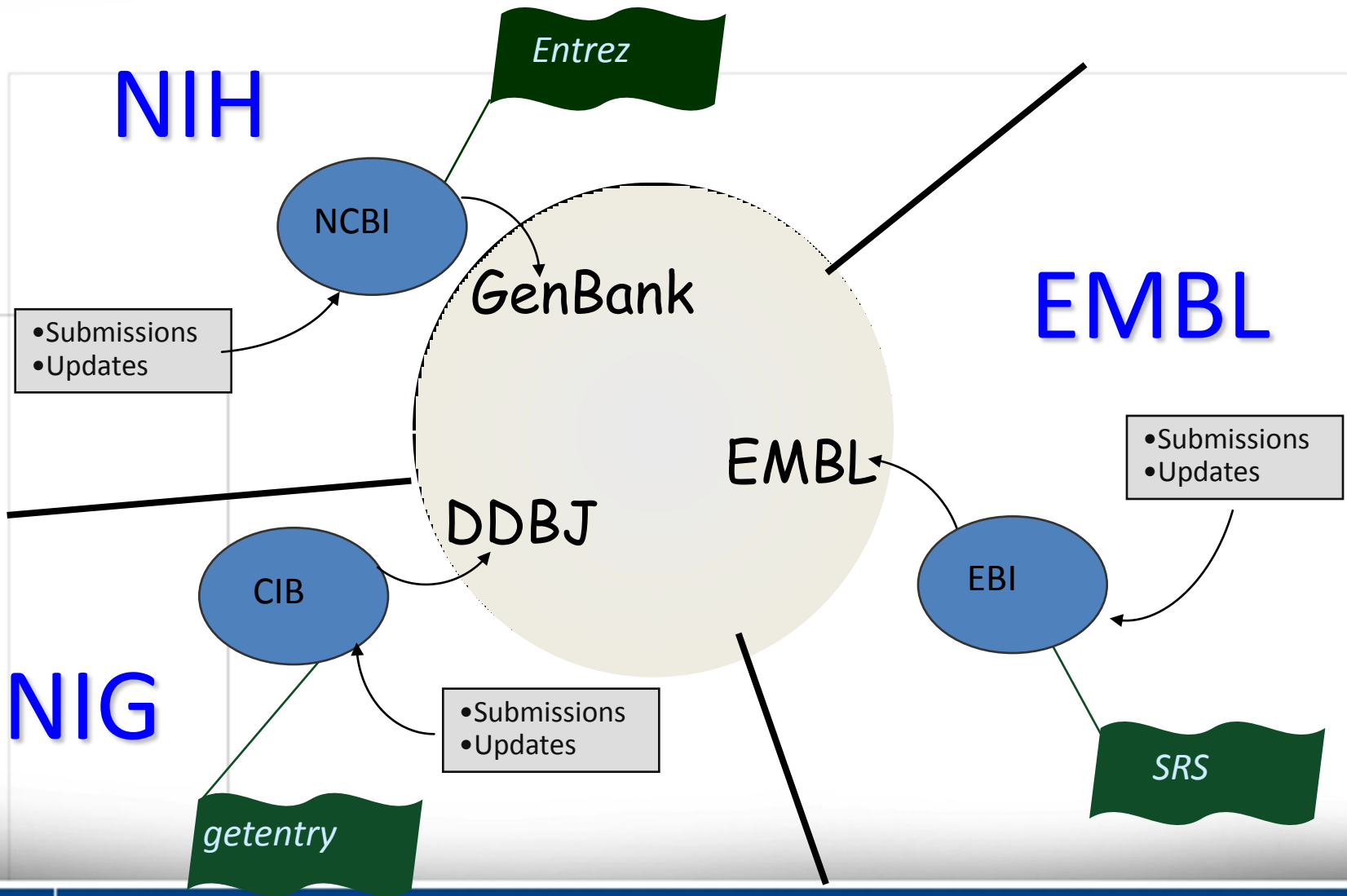
Primer
(Primer3)

Alignment
(Clustal, Muscle)

Filogenetik
(PAUP, MEGA)

DATA NUKLEOTIDA

The International Sequence Database Collaboration



The screenshot shows the NCBI homepage with the following elements:

- Navigation Menu:** PubMed, All Databases, BLAST, OMIM, Books, TaxBrowser, Structure.
- Search Bar:** 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.
- Tools:** Data mining.
- Research at NCBI:** People, projects, and seminars.
- Software engineering:** Tools, R&D, and databases.
- Education:** Teaching 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
- Clusters of orthologous groups
- Coffee Break, Genes & Disease, NCBI Handbook
- Electronic PCR
- Entrez Home
- Entrez Tools
- Gene expression omnibus (GEO)
- Human genome resources
- Influenza Virus Resource
- Map Viewer
- dbMHC
- Mouse genome resources
- My NCBI
- ORF finder
- Rat genome resources
- Reference sequence project
- SAGEmap
- SKY/CGH database

New Protein Clusters
Entrez Protein Clusters database
The new Entrez Protein Clusters database is a collection of Reference Sequence (RefSeq) proteins, from the complete genomes of prokaryotes, plasmids, and organelles, that have been grouped and annotated based on sequence similarity and protein function. Click here to find out more about the [Protein Clusters](#) database.

New dbGaP
NCBI's dbGaP Genome Wide Association Database
NCBI's dbGaP (database of Genotype and Phenotype) provides data from Genome Wide Association (GWA) studies. The resource is intended to help elucidate the link between genes and disease. For each study, users have access to detailed information about the phenotypic variables measured and pre-computed associations between subjects' phenotypes and genotypes. Click here to read the [press release](#). To read more about GWA projects, see NCBI's [GWA resource page](#).

PubMed Central
An archive of biomedical and life sciences journals

- Free fulltext
- Over 1,100,000 articles from over 340 journals
- Linked to PubMed and fully searchable

Use of PubMed Central requires no registration or fee. Access it from any computer with an Internet connection.

1. Situs ini dikembangkan oleh NCBI - USA
2. Perkembangan sangat pesat & pengguna dari situs ini para ahli molecular biologi & bioinformatika
3. Fasilitas/option sangat luas (DNA, protein, Blast, pubmed, tools, struktur protein dll)

Situs website - Bioinformatika - Europe database

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

The screenshot displays the SRS@EBI website interface. At the top, there is a navigation bar with links for 'All Databases', 'Enter Text Here', 'Go', 'Reset', 'Advanced Search', and 'Give us feedback'. Below this is a secondary navigation bar with categories like 'Databases', 'Tools', 'EBI Groups', 'Training', 'Industry', 'About Us', and 'Help'. The main content area is divided into several sections:

- Quick Text Search:** A search box with a dropdown menu set to 'Nucleotides' and a 'Search' button. Below it, it says 'Searches Databanks: EMBL Nucleotides'.
- News and Announcements:** A section titled 'Important announcements:' with two entries:
 - 30.07.07:** PDB has been modified to use the Remediated PDB data. If you use external structure viewer with PDB data from SRS@EBI you may need to update your software (see [Software Resources for Remediated Data](#) for details) to work with this data. The implementation has been changed to use a virtual library to reflect the reorganisation of the PDB data. If you have any feedback/comments please send them to us using the [support form](#).
 - 27.07.07:** Due to resource limitations we have removed a number of databanks from SRS@EBI. For details and links to alternative services please see the [SRS@EBI wiki](#).
- List Search:** A section for pasting a list of sequence ID's. It includes instructions: 'Paste in a list of sequence ID's. The list must be of the format DATABASE:ID. e.g. EMBL:AB046566 or UNIPROT:104K_THEAN. For more details see the [wiki](#).' It also states: 'Ensure each entry is on a single line and that the database(s) exists on this server. Multiple databases can be searched simultaneously. There is a maximum limit of 500 ID's.' Below this is a large empty text box and a 'List file' button with a 'Choose File' dialog and 'no file selected' text. A 'Search' button is at the bottom right.

On the left side, there is a 'Tips' section with three items:

- Want to know more about using SRS?** - go to the [Help Center](#) for online searchable help. - look in our [SRS@EBI FAQ](#) for answers to commonly asked questions
- Linking to SRS?** - Please read our [Linking to SRS](#) guide for important information regarding linking to our SRS server.
- Public SRS servers worldwide**

The bottom of the page features the 'biowisdom SRS' logo.

1. Situs ini adalah database/gen bank versi Eropa khusus dikembangkan oleh bioinformatika di UK.
2. Fasilitas tools yg dimiliki sangat lengkap khususnya untuk menganalisis DNA (Clustal dll) & protein
3. Di indonesia kurang populer dibandingkan NCBI

CIB-DDBJ

Center for information biology

- <http://www.ddbj.nig.ac.jp/>
- Jepang

The screenshot shows the DDBJ website homepage. At the top left is the DDBJ logo with the text "DNA Data Bank of Japan". To the right is a search bar with a "Go" button and radio buttons for database types: DDBJ (selected), UniProt, PDB, DAD, PRF, and Patent. Below the search bar is a navigation menu with items: HOME, Submission, Search/Analysis, FTP/WebAPI, Documents, Statistics, Contact Us, and Japanese. On the left side, there is a sidebar with a "Search" section containing links for getentry, ARSA, SRS, TXSearch, and BLAST. The main content area features a banner for "DDBJ : DNA Data Bank of Japan" with a background image of Mount Fuji. Below the banner is a "Hot Topics" section with three entries dated Mar. 07, 2008, Mar. 7, 2008, and Aug. 1, 2007. At the bottom, there are two yellow buttons: "Sequence Data Submission" and "FTP/Web API".

DDBJ
DNA Data Bank of Japan

Accession [DNA](#) [Protein](#) [Taxonomy](#) [Site Search](#)
Accession numbers
 DDBJ UniProt PDB DAD PRF Patent >>more

HOME Submission Search/Analysis FTP/WebAPI Documents Statistics Contact Us [Japanese](#)

▶ About DDBJ
▶ DDBJ Release Note
▶ Q and A

Sequence Submission
▶ SAKURA
▶ Mass Submission
▶ Data Updates

Search
▶ [getentry](#)
▶ [ARSA](#)
▶ [SRS](#)
▶ [TXSearch](#)
▶ [BLAST](#)

DDBJ : DNA Data Bank of Japan
DDBJ (DNA Data Bank of Japan) is one of three summit databanks that construct DDBJ/EMBL/GenBank International Nucleotide Sequence Database, through close collaboration with EBI in Europe and NCBI in USA.

Hot Topics ▶ More
▶ Mar. 07, 2008 [\[Hot Topic of March, 2008\] DDBJ started to release patent DNA by KIPO](#)
▶ Mar. 7, 2008 [\[Hot Topic of March, 2008\] The directory of anonymous FTP was changed](#)
▶ Aug. 1, 2007 [\[Caution\] Revision of DDBJ flat file format](#)

Maintenance ▶ More
▶ Mar. 10, 2008 [Correction of the invalid sequence version number for BA000010 and BA000044](#)

Sequence Data Submission FTP/Web API

The screenshot shows the ENA website interface. At the top, there is a navigation bar with links for Services, Research, Training, and About us. Below this is the ENA logo and a search bar with a 'Search' button. The search bar contains the text 'Examples: BN000085, histone'. Below the search bar is a navigation menu with links for Home, Search & Browse, Submit & Update, Software, About ENA, and Support. The main content area is titled 'European Nucleotide Archive' and contains a description of the archive and a list of popular links. The 'Text Search' section has a search bar and a 'Search' button. The 'Sequence Search' section has a search bar and a 'Search' button. The 'Popular' section lists links for 'Submit and update', 'Sequence submissions', 'Genome assembly submissions', 'Submitting environmental sequences', 'Citing ENA data', 'Rest URLs for data retrieval', and 'Rest URLs to search ENA'. The 'Latest ENA news' section has a link for '03 Apr 2017: ENA Release 131'. The bottom of the screenshot shows a Windows taskbar with icons for the Start menu, File Explorer, Google Chrome, and Adobe Reader. The system tray shows the date and time as 16:09 on 20-Apr-17.

www.ebi.ac.uk/ena/

EMBL-EBI

Services Research Training About us

ENA
European Nucleotide Archive

Search

Examples: BN000085, histone

Advanced
Sequence

Home Search & Browse Submit & Update Software About ENA Support

European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world's nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation. [More about ENA](#)

Access to ENA data is provided through the browser, through search tools, large scale file download and through the API.

Text Search

Search

Examples: BN000065, histone

Advanced search

Sequence Search

Enter or paste a nucleotide sequence or accession number

Popular

- Submit and update
- Sequence submissions
- Genome assembly submissions
- Submitting environmental sequences
- Citing ENA data
- Rest URLs for data retrieval
- Rest URLs to search ENA

Latest ENA news

03 Apr 2017: [ENA Release 131](#)

Release 131 of ENA's assembled/annotated sequences now available

Interface dari European Nucleotide Archive

Terhubung dengan database lain seperti NCBI dan DDBJ

ENA relatif mudah diakses dengan pengelompokan data yang baik

EMBL-EBI European Nucleotide Archive

Services Research Training About us

Home Search & Browse Submit & Update Software About ENA Support

European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world's nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation. [More about ENA](#)

Access to ENA data is provided through the browser, through search tools, large scale file download and through the API.

Text Search

Search

Examples: [BN000065](#), [histone](#)

Advanced search

Sequence Search

Enter or paste a nucleotide sequence or accession number

Popular

- Submit and update
- Sequence submissions
- Genome assembly submissions
- Submitting environmental sequences
- Citing ENA data
- [Rest URLs for data retrieval](#)
- [Rest URLs to search ENA](#)

Latest ENA news

03 Apr 2017: [ENA Release 131](#)

Release 131 of ENA's assembled/annotated sequences now available

Step 1.

Gunakan kata kunci untuk mencari database transkriptomik dan genomik

Kasus A. Gunakan kata kunci . **Oryza officinalis**

Kasus B. Gunakan kata kunci. **Human Mycobacterium**

www.ebi.ac.uk/ena/data/search?query=oryza+officialis

Assembly (1)
Sequence (Update) (1)
Sequence (Release) (109,167)
Contig set
Genome assembly contig set (1)
Transcriptome assembly contig set (1)
Coding (Update) (37)
Coding (Release) (895)
Non-coding (Release) (105)
Read
Experiment (7)
Run (8)
Study (4)
Study (Sequence) (15)
Taxon (1)
Sample (17)
Submission

Assembly (1 results found)
GCA_000717455.1 O. officinalis chromosome 3 short arm assembly for Oryza officinalis
View all 1 results

Sequence (Update) (1 results found)
AP011469 Oryza officinalis genomic DNA, chromosome 6, BAC clone: OO_Ba0014A07, strain: IRGC 100896.
View all 1 results

Sequence (Release) (109,167 results found)
AC225786 Oryza officinalis, complete sequence.
View all 109,167 results

Genome assembly contig set (1 results found)
JJMQ01000000 Oryza officinalis, WGS project JJMQ01000000 data
View all 1 results

Transcriptome assembly contig set (1 results found)
GBRJ01000000 Oryza officinalis, TSA project GBRJ01000000 data
View all 1 results

Coding (Update) (37 results found)
BAX24998 Oryza officinalis hypothetical protein
View all 37 results

Windows taskbar: File Explorer, Chrome, PowerPoint, PDF Reader, 16:11, 20-Apr-17

Step 2a. Kasus A. Kata kunci ? Oryza officinalis?

Klik data Transcriptome assembly contig set

Penting: Dalam analisis komparasi ekspresi gen, lebih mudah membandingkan data transkrip yang sudah valid seperti RNA sequencing

www.ebi.ac.uk/ena/data/view/GBRJ01000000

TSA Sequence Set: GBRJ01000000

Oryza officinalis, TSA project GBRJ01000000 data

View: [TEXT](#) [XML](#) Download: [XML](#) [TEXT](#)

Lineage	Molecule type	Topology	Data class	Taxonomic Division
Oryza officinalis	transcribed RNA	linear	SET	PLN

Number of sequences: 476 Set Version: 1

Keywords: Transcriptome Shotgun Assembly, TSA.

Lineage: [Eukaryota](#), [Viridiplantae](#), [Streptophyta](#), [Embryophyta](#), [Tracheophyta](#), [Spermatophyta](#), [Magnoliophyta](#), [Liliopsida](#), [Poales](#), [Poaceae](#), [BOP clade](#), [Oryzoideae](#), [Oryzeae](#), [Oryzinae](#), [Oryza](#)

Navigation: [Source Feature\(s\)](#) [Comments](#) [Publications](#) [Submission Details](#)

[1] **De novo assembly and characterization of Oryza officinalis leaf transcriptome by using RNA-seq.**
Bao Y., Xu S., Jing X., Meng L., Qin Z.
School of Life Science, Qufu Normal University, Qufu, Shandong 273165, China.
Biomed Res Int 2015982065 (2015)
[Show abstract](#)
[PubMed](#)
[PDF](#)
DOI: [10.1155/2015/982065](#)
[doi](#)

Publications inherited from Study

[1] ~~De novo assembly and characterization of Oryza officinalis leaf transcriptome by using RNA-seq.~~

Windows taskbar: 16:17 20-Apr-17

Step 3a. Kasus A. Kata kunci [?](#) Oryza officinalis [?](#)

Publikasi dari data tersebut dapat dicek pada laman **Publications**

www.ebi.ac.uk/ena/data/view/GBRJ01000000

rizaputranto - Yahoo | Gmail - Boîte de réception | Webmail - Main | Google Agenda | Facebook | Twitter / Accueil | Welcome! | LinkedIn | Riza Arief Putranto | Riza-Arief Putranto - | Google Maps

TSA Sequence Set: GBRJ01000000

Oryza officinalis, TSA project GBRJ01000000 data

View: [TEXT](#) [XML](#) Download: [XML](#) [TEXT](#)

Lineage	Molecule type	Topology	Data class	Taxonomic Division
Oryza officinalis	transcribed RNA	linear	SET	PLN

Number of sequences: 476 Set Version: 1

Keywords: Transcriptome Shotgun Assembly, TSA.

Lineage: [Eukaryota](#), [Viridiplantae](#), [Streptophyta](#), [Embryophyta](#), [Tracheophyta](#), [Spermatophyta](#), [Magnoliophyta](#), [Liliopsida](#), [Poales](#), [Poaceae](#), [BOP clade](#), [Oryzoideae](#), [Oryzeae](#), [Oryzinae](#), [Oryza](#)

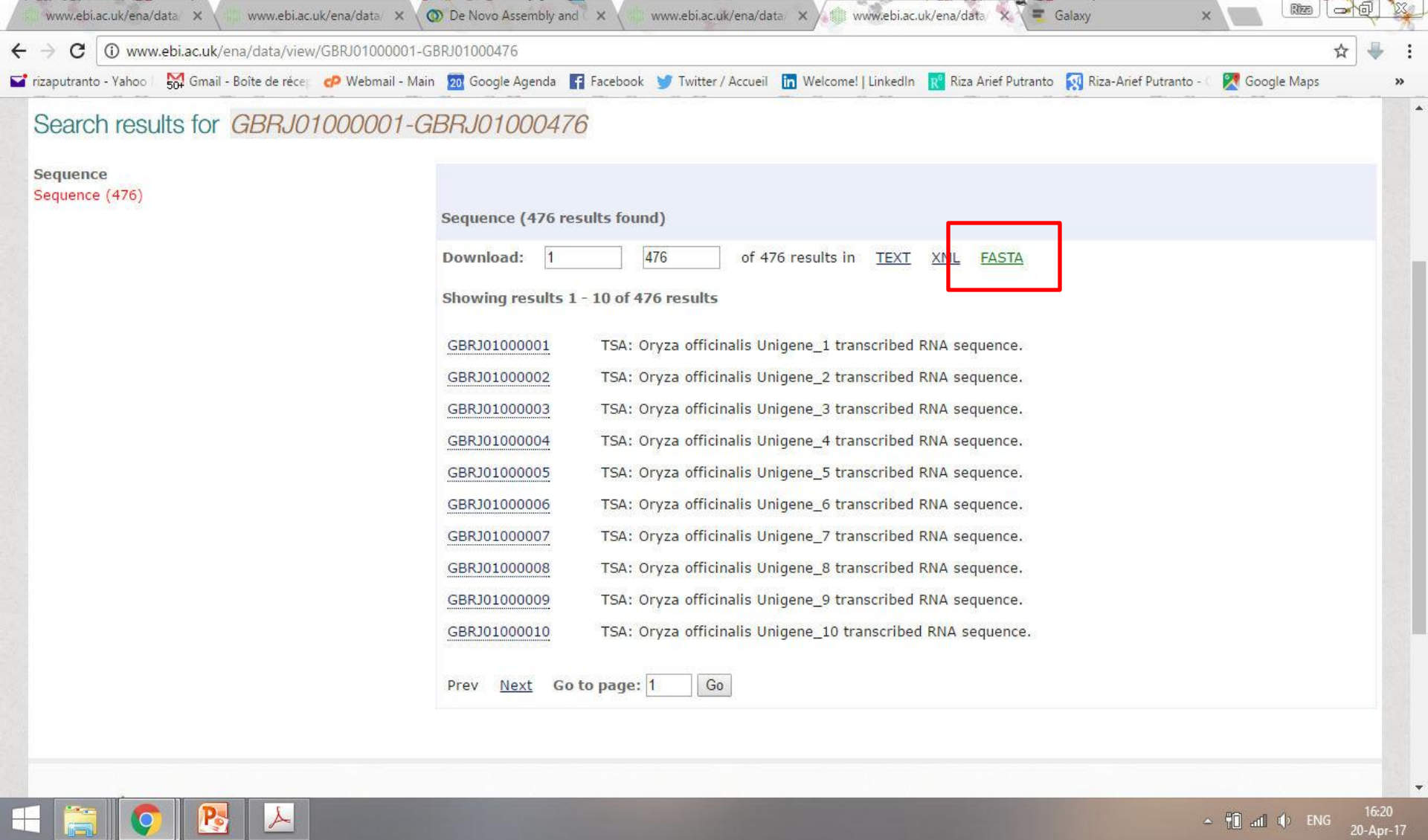
Navigation | Source Feature(s) | Comments | Publications | Submission Details

↓	Transcriptome assembly contig	GBRJ01000001-GBRJ01000476
↑	Study:	PRJNA261634
↓	Sample:	SAMN03074946
🏠	Taxon:	Taxon:4535

Windows taskbar: 16:16 20-Apr-17

Step 4a. Kasus A. Kata kunci [?](#) Oryza officinalis [?](#)

Data transkriptomik dapat diakses di laman **Navigation**
Klik kode sekuen **GBR1000001** untuk mengunduh data



Step 5a. Kasus A. Kata kunci **Oryza officinalis**

Data sekuen transkrip mRNA dapat diunduh dalam bentuk **.fasta** dengan langsung klik **FASTA**

PROTEIN

Situs website - Bioinformatika

protein database. <http://www.expasy.org/sprot/>

ExPASy - UniProt Knowledgebase: Swiss-Prot and TrEMBL

http://www.expasy.org/sprot/

Search for

Swiss-Prot
Protein knowledgebase
TrEMBL
Computer-annotated supplement to Swiss-Prot

UniProt
the universal protein resource

The UniProt Knowledgebase consists of:

- **UniProtKB/Swiss-Prot**; a curated protein sequence database which strives to provide a high level of annotation (such as the description of the function of a protein, its domains structure, post-translational modifications, variants, etc.), a minimal level of redundancy and high level of integration with other databases [[More details](#) / [References](#) / [Linking to Swiss-Prot](#) / [User manual](#) / [Recent changes](#) / [Disclaimer](#)].
- **UniProtKB/TrEMBL**; a computer-annotated supplement of Swiss-Prot that contains all the translations of EMBL nucleotide sequence entries not yet integrated in Swiss-Prot.

These databases are developed by the Swiss-Prot groups [at SIB](#) and [at EBI](#).

UniProt Knowledgebase Release 12.1 consists of:
UniProtKB/Swiss-Prot Release 54.1 of 21-Aug-2007:
277883 entries ([More statistics](#))
UniProtKB/TrEMBL Release 37.1 of 21-Aug-2007:
4754787 entries ([More statistics](#))

> Swiss-Prot headlines
More than 18'500 phosphorylation sites identified by mass spectrometry in UniProtKB/Swiss-Prot (Read [more...](#))

Access to the UniProt Knowledgebase

- **SRS** - Access to UniProtKB/Swiss-Prot, UniProtKB/TrEMBL and other databases using the Sequence Retrieval System
- **Full text search** in the UniProt Knowledgebase
- **Advanced search in the UniProt Knowledgebase** by description, gene name and organism (can be used to create html links to UniProt Knowledgebase queries)
- **Taxonomy browser (NEWT)**
- **BLAST** similarity search
- **by description or identification** (any word in the DE, OS, OG, GN and ID lines)
- **by citation** (RL line; UniProtKB/Swiss-Prot only)
- **Retrieve a list of UniProtKB entries**
- **Randomly retrieve a UniProtKB entry**
- **UniProtKB Sequence/Annotation Version Database** new!
- **Swiss-Prot ID tracker**

Documents and services

Swiss-Prot documents - user manual, release notes, indices and lots of other important documents and lists

1. Situs ini menghususkan pada protein database
2. Fasilitas meliputi : Blast, struktur dll

PDB

Protein data bank



A MEMBER OF THE PDB

An Information Portal to Biological Macromolecular Structures

As of Tuesday Mar 11, 2008 there are 49426 Structures | [PDB Statistics](#) [CONTACT US](#) | [HELP](#) | [PRINT PAGE](#) PDB ID or keyword Author | [Adv](#)[Home](#) [Search](#)

Are you missing data updates? The PDB archive has moved to <ftp://ftp.wwpdb.org>.
For more information click [here](#).

Welcome to the RCSB PDB

The **RCSB** PDB provides a variety of tools and resources for studying the structures of biological macromolecules and their relationships to sequence, function, and disease.

The RCSB is a member of the **wwPDB** whose mission is to ensure that the PDB archive remains an international resource with uniform data.

News

- [Complete News](#)
- [Newsletter](#)
- [Discussion Forum](#)
- [Job Listings](#)

11-March-2008

Using RCSB PDB Tools and Resources

A chapter describing the various options and procedures for querying, downloading and

- <http://www.rcsb.org/pdb/home/home.do>
- 3D Struktur

- Home
- Getting Started
- ▶ Download Files
- ▶ Deposit and Validate
- ▶ Structural Genomics
- ▶ Dictionaries & File Formats
- ▶ Software Tools
- ▶ General Education
- ▶ Site Tutorials
- BioSync
- ▼ General Information
 - About the PDB

Uniprot

- <http://www.uniprot.org/>

What we provide

UniProtKB	Protein knowledgebase, consists of two sections: <ul style="list-style-type: none">★ Swiss-Prot, which is manually annotated and reviewed.★ TrEMBL, which is automatically annotated and is not reviewed.
UniRef	Sequence clusters, used to speed up similarity searches.
UniParc	Sequence archive, used to keep track of sequences and their identifiers.
Supporting data	Literature citations , taxonomy , keywords and more .



[SWISS-PROT](#) · [TrEMBL](#)

- > [Forthcoming changes](#)
- > [News archives](#)

SITE TOUR



Learn how to make best use of the tools and data on this site.

PROTEIN SPOTLIGHT

[Take a walk on the genome](#) August 2008

What? No issue in July? A number of our regular readers may have noticed that for the first time in a short decade no article appeared during the month of July. And time

Situs website - Bioinformatika

- Microbial database. <http://www.tigr.org>



J. Craig Venter™
I N S T I T U T E

Home | Site Map | FAQ | Careers

- About
- What's New
- Faculty
- Databases
- Scientific Programs
- Software
- Research Groups
- Related Links

The New JCVI

The new J. Craig Venter Institute was formed in October 2006 through the merger of several affiliated and legacy organizations--The Institute for Genomic Research (TIGR) and The Center for the Advancement of Genomics (TCAG), The J. Craig Venter Science Foundation, The Joint Technology Center, and the Institute for Biological Energy Alternatives (IBEA). Today all these organizations have become one large multidisciplinary genomic-focused organization. With more than 500 scientists and staff, more than 250, 000 square feet of laboratory space, and locations in Rockville, Maryland and La Jolla, California, the new JCVI is a world leader in research. [\[more\]](#)

Contact Us | Privacy Statement

Data Disclaimer | © 2008 J. Craig Venter Institute

1. Banyak digunakan oleh peneliti biologi molekuler dibidang mikroba.

Situs website - Bioinformatika

Genome database. <http://www.jgi.doe.gov/>

DOE Joint Genome Institute

search jgi
contact us
site map
internal

[GO]

ABOUT US
PROGRAMS
SEQUENCING
JGI SCIENCE
MEETINGS
NEWS
EDUCATION
EMPLOYMENT

JGI brings the expertise of five national laboratories, [Lawrence Berkeley](#), [Lawrence Livermore](#), [Los Alamos](#), [Oak Ridge](#), and [Pacific Northwest](#), and the [Stanford Human Genome Center](#) to bear on the frontiers of genome sequencing and related biology. Our sequencing targets encompass a rapidly expanding range of microbes, animals, and plants. The [Community Sequencing Program \(CSP\)](#) aims to broaden the range still further. JGI is operated by the [University of California](#) for the [U.S. Department of Energy](#).

genomes
[Microbial, Eukaryotic](#)
[Integrated Microbial Genomes \(IMG\)](#)
[IMG w/ Microbiome Samples \(IMG/M\)](#)

latest news
[JGI Upgrades IMG system](#)
[JGI announces 2008 Targets](#)

sequencing
This fiscal year : 26.799 billion base pairs sequenced | [More stats](#)
[Sequencing Plans and Progress](#)
[Sequencing for Researchers](#)

bioenergy

Page last updated 7/24/2007 · [Disclaimer](#) · [Comments/Questions](#)
©1997-2007 The Regents of the University of California · [Credits](#)

1. Situs ini relatif baru dibandingkan NCBI, tetapi fasilitas yang ada sangat luas, termasuk database microbial dan eukayotic.

2. Situs ini dilengkapi dengan map metabolisme pathway dari microbial.

PROGRAM/ TOOL

Situs website - Bioinformatika - Program/Tools - Design primer

http://frodo.wi.mit.edu

Salah satu program desain primer secara on line yang paling populer & dikembangkan oleh MIT.

Primer3 Input 0.4.0 (primer3-web/htdocs/input-040.htm)

http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi

Primer3 (v. 0.4.0) Pick primers from a DNA sequence.

[Primer3plus interface](#) [More primer/oligo tools](#) [disclaimer](#) [Primer3 Home](#)
[Old \(0.3.0\) interface](#) [cautions](#) [FAQ/Wiki](#)

Paste source sequence below (5'->3', string of ACGTNacgtn -- other letters treated as N -- numbers and blanks ignored). FASTA format ok. Please N-out undesirable sequence (vector, ALUs, LINES, etc.) or use a [Mispriming Library \(repeat library\)](#):

<input checked="" type="checkbox"/> Pick left primer, or use left primer below:	<input type="checkbox"/> Pick hybridization probe (internal oligo), or use oligo below:	<input checked="" type="checkbox"/> Pick right primer, or use right primer below (5' to 3' on opposite strand):
<input type="text"/>	<input type="text"/>	<input type="text"/>

[Sequence Id:](#) A string to identify your output.

[Targets:](#) E.g. 50,2 requires primers to surround the 2 bases at positions 50 and 51. Or mark the [source sequence](#) with [and]: e.g. ...ATCT[CCCC]TCAT.. means that primers must flank the central CCCC.

[Excluded Regions:](#) E.g. 401,7 68,3 forbids selection of primers in the 7 bases starting at 401 and the 3 bases at 68. Or mark the [source sequence](#) with < and >: e.g. ...ATCT<CCCC>TCAT.. forbids primers in the central CCCC.

[Product Size Ranges](#)

[Number To Return](#) [Max 3' Stability](#)

[Max Repeat Mispriming](#) [Pair Max Repeat Mispriming](#)

[Max Template Mispriming](#) [Pair Max Template Mispriming](#)

General Primer Picking Conditions

[Primer Size](#) Min: Opt: Max:

[Primer Tm](#) Min: Opt: Max: [Max Tm Difference:](#) [Table of thermodynamic parameters:](#)

[Product Tm](#) Min: Opt: Max:

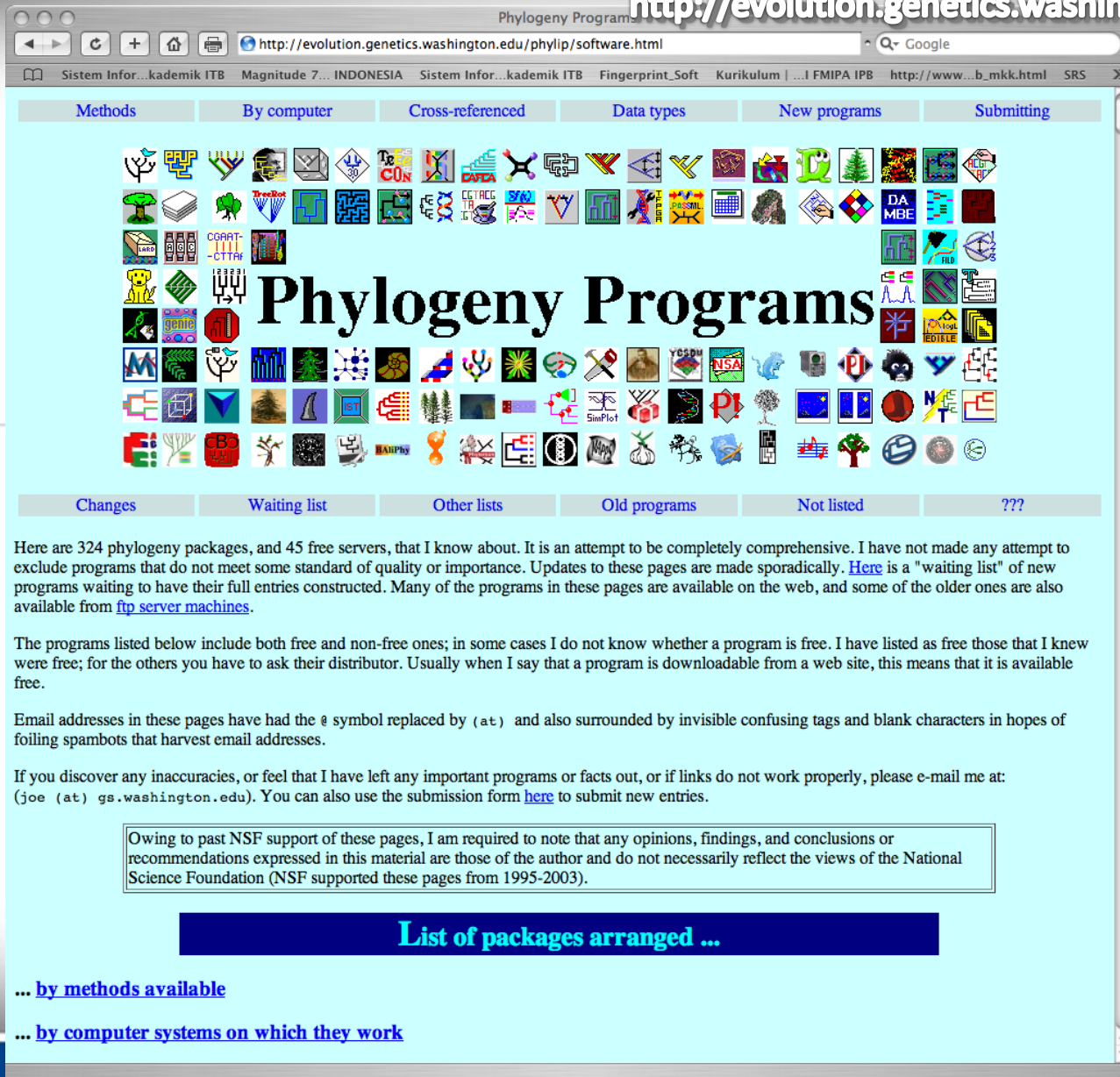
[Primer GC%](#) Min: Opt: Max:

[Max Self Complementarity:](#) [Max 3' Self Complementarity:](#)

[Max #N's:](#) [Max Poly-X:](#)

[Inside Target Penalty:](#) [Outside Target Penalty:](#) [Note: you can set Inside Target Penalty to allow primers inside a target.](#)

Situs website - Bioinformatika - Program/Tools - Phylogeny

<http://evolution.genetics.washington.edu>

Phylogeny Programs

[http://evolution.genetics.washington.edu/phylop/software.html](#)

Methods By computer Cross-referenced Data types New programs Submitting

Phylogeny Programs

Changes Waiting list Other lists Old programs Not listed ???

Here are 324 phylogeny packages, and 45 free servers, that I know about. It is an attempt to be completely comprehensive. I have not made any attempt to exclude programs that do not meet some standard of quality or importance. Updates to these pages are made sporadically. [Here](#) is a "waiting list" of new programs waiting to have their full entries constructed. Many of the programs in these pages are available on the web, and some of the older ones are also available from [ftp server machines](#).

The programs listed below include both free and non-free ones; in some cases I do not know whether a program is free. I have listed as free those that I knew were free; for the others you have to ask their distributor. Usually when I say that a program is downloadable from a web site, this means that it is available free.

Email addresses in these pages have had the @ symbol replaced by (at) and also surrounded by invisible confusing tags and blank characters in hopes of foiling spambots that harvest email addresses.

If you discover any inaccuracies, or feel that I have left any important programs or facts out, or if links do not work properly, please e-mail me at: (joe (at) gs.washington.edu). You can also use the submission form [here](#) to submit new entries.

Owing to past NSF support of these pages, I am required to note that any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation (NSF supported these pages from 1995-2003).

List of packages arranged ...

- ... [by methods available](#)
- ... [by computer systems on which they work](#)

Situs khusus untuk mendownload program-program yg terkait dengan studi moleculer Filogenetik. Banyak program yang dapat diakses secara free.

CLC bio: Home

http://www.clcbio.com/index.php?id=354

Enter your search here Search

Home Info Software Hardware Consulting Download Science Support Contact

Combined Workbench DNA Workbench CLC RNA Workbench Protein Workbench Free Workbench Education Solutions Company

Home

Bioinformatics Solutions

Software Consulting Hardware

Bioinformatics solutions
CLC bio offers a wide range of world class bioinformatics solutions: From highly advanced sequence analysis software to bioinformatics consulting services, and hardware accelerated database searches.

Our Bioinformatics Solutions

Why shop two, three or five different places for DNA, RNA, and Protein sequence analyses, when you can get everything you need under a single roof?

We have introduced one-stop-shopping in sequence analysis, with a wide variety of solutions, such as software, hardware, data analysis, and custom-designed bioinformatics algorithms, to name a few. Oh, and by the way: We're Apple solution providers and value added reseller!

SOLUTIONS
CLC RNA Workbench
The CLC RNA Workbench gives the user easy access to a range of complex algorithms and options. RNA secondary structures can be predicted using state-of-the-art free energy minimization algorithms and parameters, thermodynamic and statistical details are easily extracted.
[Read more...](#)

INFORMATION
Impressive performance per dollar with CLC Bioinformatics Cell 2.0 from CLC bio
With the release of CLC Bioinformatics Cell 2.0, the world's leading bioinformatics solution provider, CLC bio, raises the bar significantly in terms of performance per dollar in high-performance bioinformatics.
[Read more...](#)

Currently...
47234
...workbench users!

Newsletter
Sign up for email news about product releases and relevant notes on bioinformatics...
Enter your email here
[Subscribe](#)

CLC RNA Workbench
Have you tried the new state-of-the-art CLC RNA Workbench yet?
[Download a free 30-day trial!](#)

HAMMER-TIME!
Unleash hidden powers in your computer processor with CLC Bioinformatics Cell - now also featuring HMMER
[Download fact sheet...](#)

Go to "<http://www.clcbio.com/apple>"

1. Salah satu program bioinformatika komersial. Harga dapat mencapai 25 Juta/program atau sewa sekitar 1 jt/3 bulan
2. Program sangat bagus & mudah digunakan serta memiliki koneksi dengan gen bank.
3. Dengan program ini kita dapat mengakses gen bank & menganalisis DNA, protein secara terintegrasi

Molecular Biology Freeware for Windows

http://molbiol-tools.ca/molecular_biology_freeware.htm#Viewing%20three%20dimensional%20(~ Google

Sistem Infor...kademik ITB Magnitude 7... INDONESIA Sistem Infor...kademik ITB Fingerprint_Soft Kurikulum | ...I FMIPA IPB http://www...b_mkk.html SRS >>

Molecular Biology Freeware for Windows

A. General - below


B. Microarray - [next page](#)


C. Java programs - [next page](#)


Good places to start are [Genamics SoftwareSeek](#) and [BioExchange](#) and [eBioinfogen](#). For general software see [Winsite](#). The following sites are arranged in the order that I discovered them. At some point they will be clustered by preference:

A. [DNA, RNA and genomic analysis](#)
B. [Plasmid graphic packages](#)
C. [Primer design](#)
D. [Protein analysis](#)
E. [Viewing three dimensional structures](#)
F. [Alignments](#)
G. [Phylogeny](#)
H. [Miscellaneous](#)
I. [Graphic packages](#)

- **DNA, RNA and genomic analysis:**

 **CLC bio** [CLC Free Workbench](#) - allows basic sequence analysis such as open reading frame determination, restriction site analysis, translation from DNA/ RNA to proteins, alignments, and tree reconstruction in a single window format.

 **geneIOUS** - [Geneious](#) (Alexei Drummond Biomatters Ltd. Auckland, New Zealand) provides an automatically-updating library of genomic and genetic data; for organizing and visualizing data. It provides a fully integrated, visually-advanced toolset for: [sequence alignment](#) and phylogenetics; sequence analysis including BLAST; protein structure viewing, NCBI, EMBL, Pubmed auto-find & much more ...

 **Genome2D** - is for the visualization of transcriptome and other customized data sets (visualizes a bacterial genome with all relevant information such as gene orientation, operon structure, transcriptional terminators or regulator binding sites) on linear chromosome maps constructed from annotated bacterial genome sequences. For a complete list of what this incredible program will do see [here](#). Images and data tables from Genome2D can easily be exported for further use in other presentation programs. (Reference: Baerends, R.J.S., et al. (2004) [Genome Biology](#), 5: R37).

1. Situs freeware program molecular biology & program bioinformatika yg terbaru dapat diperoleh di situs ini

THANK
YOU



607132.wordpress.com

Noviani's Blog

