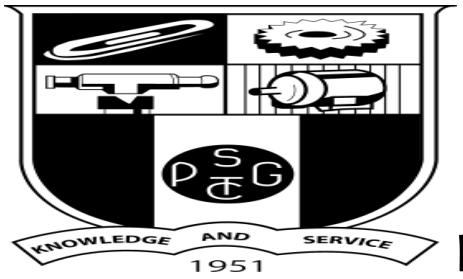


PSG College of Technology

Biology tutorial Report

By

Mohamed Aklamaash .M.R

(1)Msc Ds

Introduction:

Biological databases play a central role in bioinformatics. They offer scientists the opportunity to access a wide variety of biologically relevant data, including the genomic sequences of an increasingly broad range of organisms. Biological Databases are free to use and they has large collection of data with them. There are several databases maintained by the governments across various countries and they provide the data free to use for the people to work on them and make upgradations on them. Scientists across the globe use these data to cure and find aides for the disease.

Types of biological databases:

1. Primary Database
2. Secondary Database
3. Composite Database

Primary Database:

- It can also be called an archival database since it archives the experimental results submitted by the scientists. The primary database is populated with experimentally derived data like genome sequence, macromolecular structure, etc. The data entered here remains uncurated(no modifications are performed over the data).
- It obtains unique data obtained from the laboratory and these data are made accessible to normal users without any change.

- The data are given accession numbers when they are entered into the database. The same data can later be retrieved using the accession number. Accession number identifies each data uniquely and it never changes.

Examples:

- Examples of Primary database- Nucleic Acid Databases are GenBank and DDBJ
- Protein Databases are PDB,SwissProt,PIR,TrEMBL,Metacyc, etc.

Secondary Databases:

- The data stored in these types of databases are the analyzed result of the primary database. Computational algorithms are applied to the primary database and meaningful and informative data is stored inside the secondary database.
- The data here are highly curated(processing the data before it is presented in the database). A secondary database is better and contains more

valuable knowledge compared to the primary database.

Example:

- InterPro (protein families, motifs, and domains)
- UniProt Knowledgebase (sequence and functional information on proteins)

Composite Databases:

- The data entered in these types of databases are first compared and then filtered based on desired criteria.
- The initial data are taken from the primary database, and then they are merged together based on certain conditions.
- It helps in searching sequences rapidly. Composite Databases contain non-redundant data.

Example:

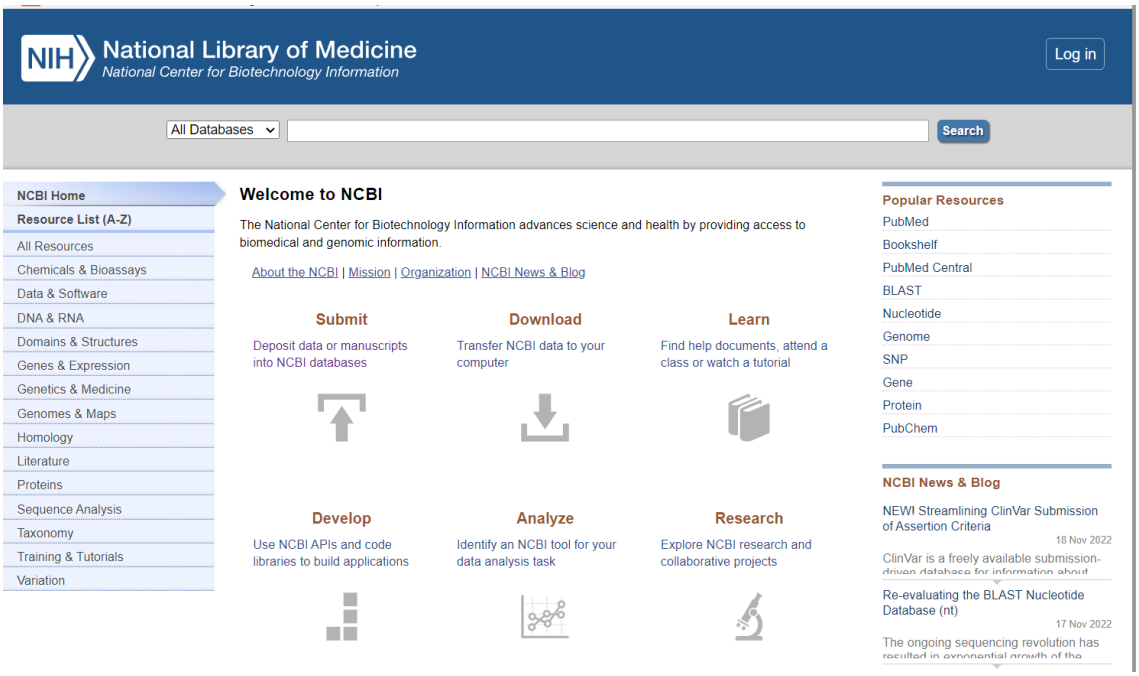
- Composite Databases -OWL,NRD and Swissport +TREMBL

NCBI:

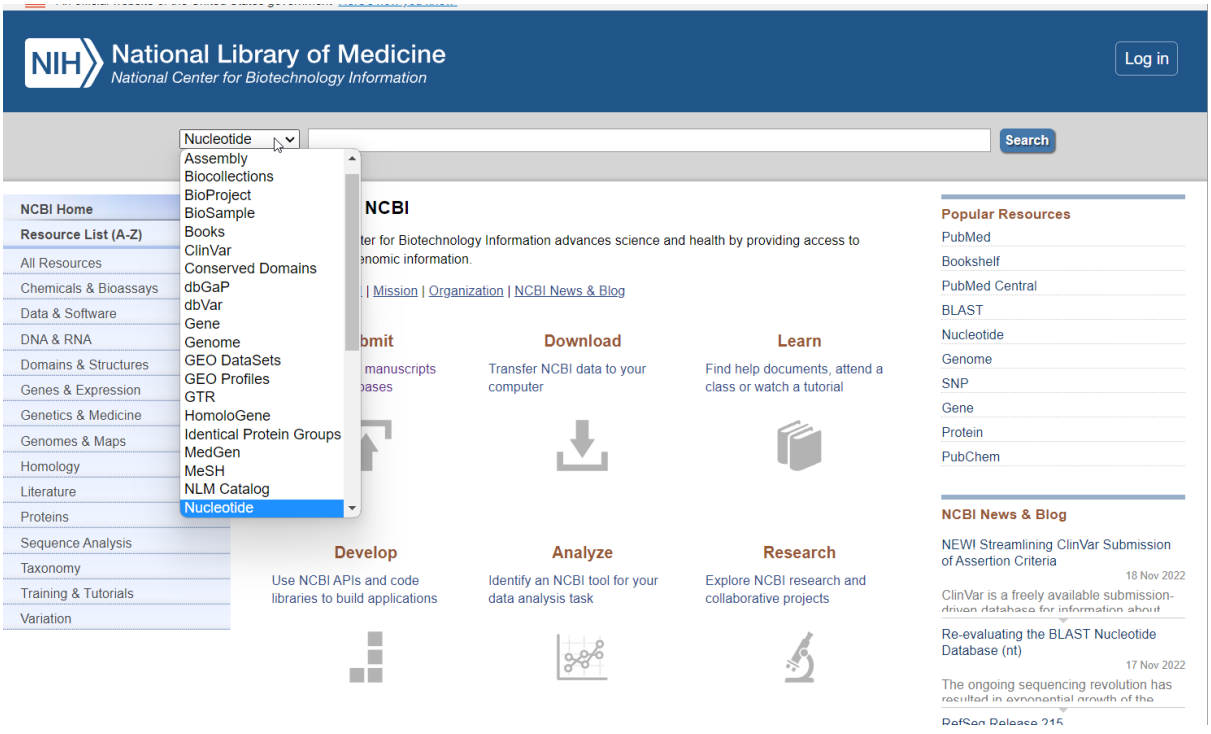
NCBI:- Natinal Centre for Biotechnology Information.

- It is funded by the United States.
- It is a branch of National Institutes of Health.
- The NCBI is located Bethesda , Maryland and was founded in 1988 through legislation sponsored by US Congressman Claude Pepper.

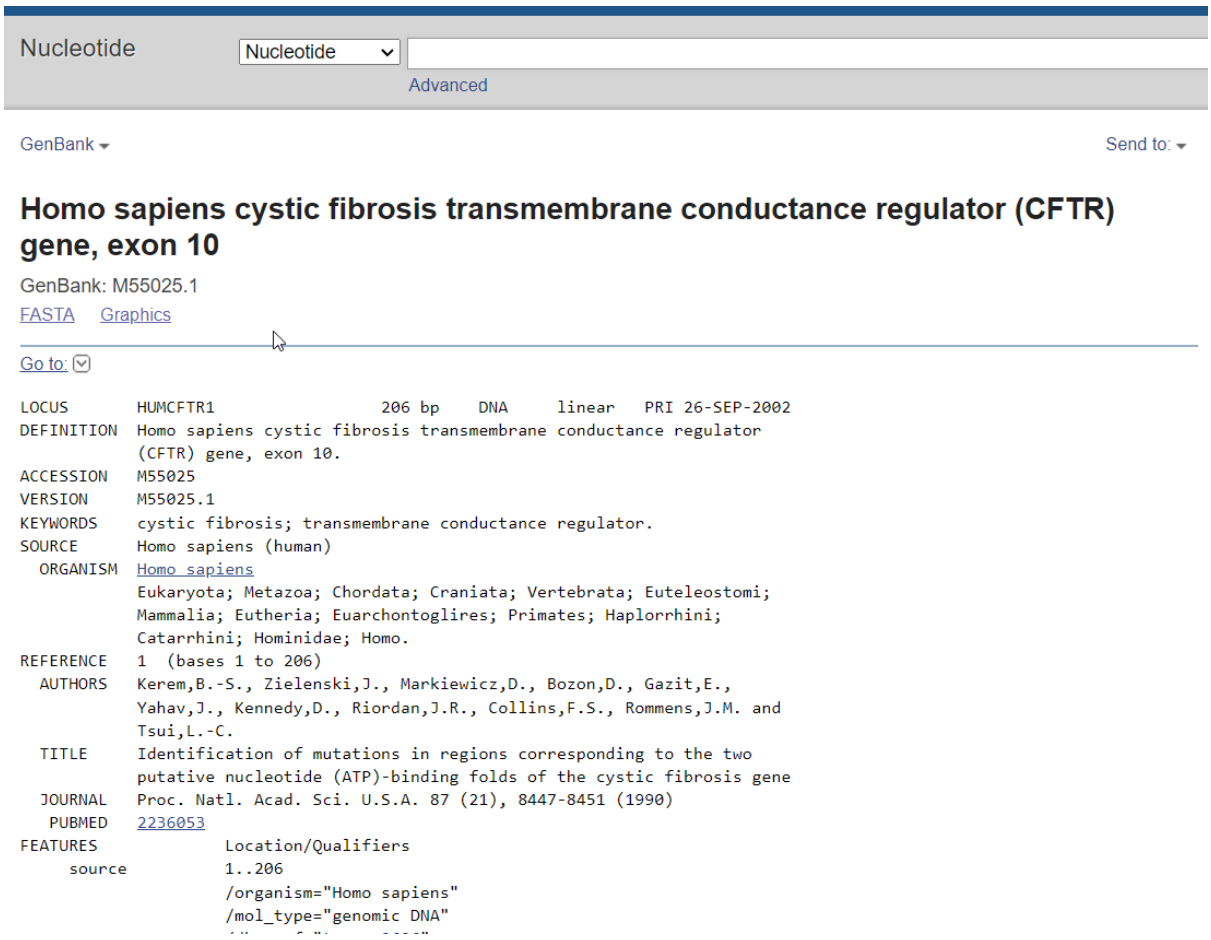
- NCBI ENTRY PAGE:



- Nucleotide :



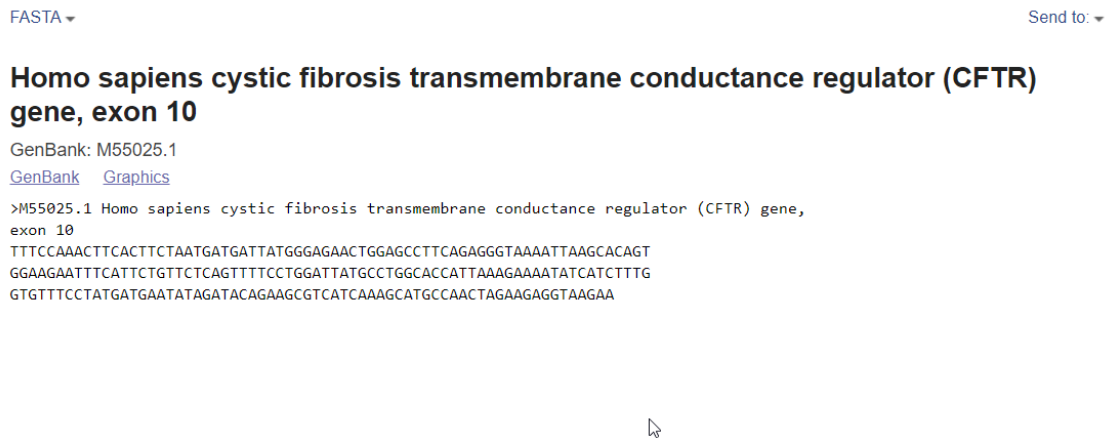
- Search for Fasta:



Steps to enter into ncbi:

1. Open Google.
2. Url : https://www.ncbi.nlm.nih.gov/nuccore/?term=cystic+fibrosis
3. Select nucleotide in the dropdown menu.
4. Enter the disease for which you want the nucleotide sequence
5. Download the FASTA sequence.

- Download the Fasta Sequence:



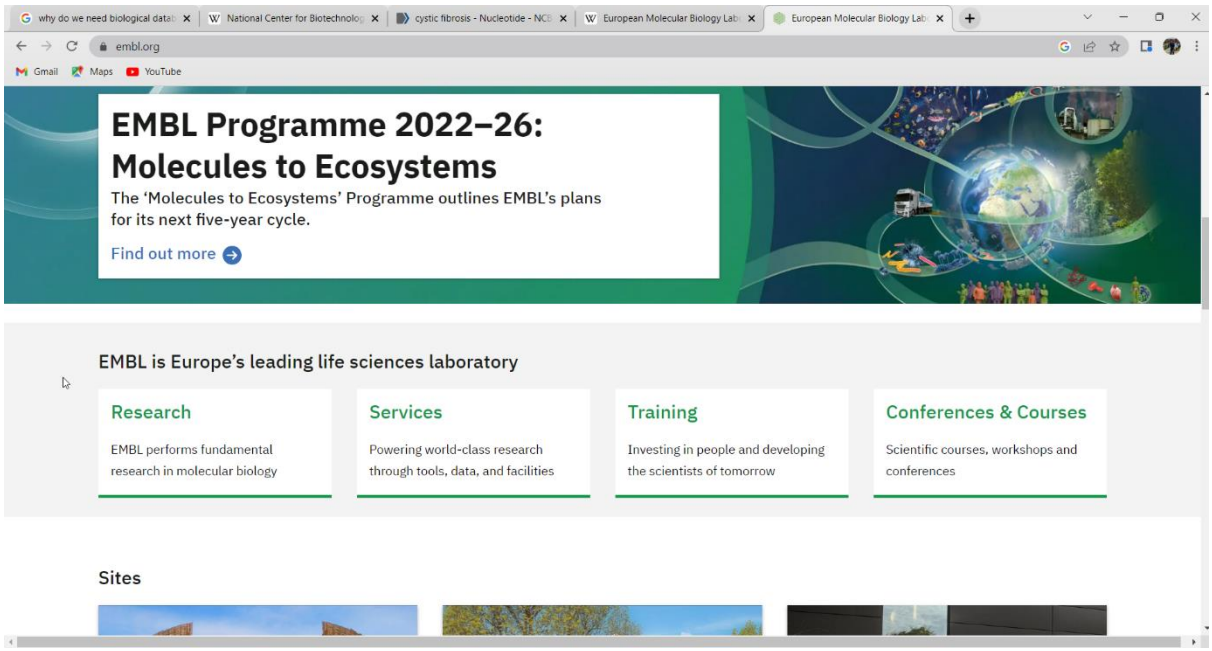
EMBL:

EMBL:- European Molecular Biological Laboratory.

- It is funded by the European community.
- The main lab with respect to this informatica is The European Bioinformatics Institute

- EMBL groups and laboratories perform basic research in molecular biology and molecular medicine as well as train scientists, students, and visitors. The organization aids in the development of services, new instruments and methods, and technology in its member state Israel is the only full member state located outside Europe.

EMBL ENTRY PAGE:



Embl Download Page:

Steps to enter into EMBL:

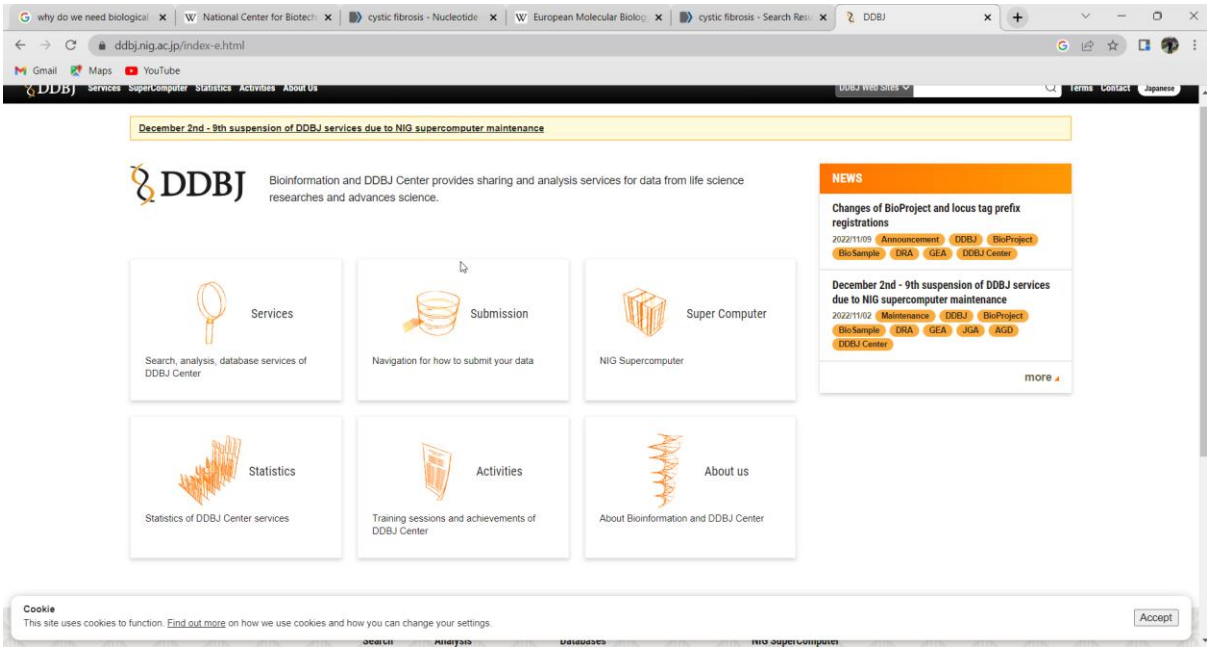
1. ENTER GOOGLE
2. URL: https://pubmed.ncbi.nlm.nih.gov/27292156/
3. Select protein in the dropdown menu.
4. Enter the disease for which you want the nucleotide sequence
5. Download the FASTA sequence.

DDBJ:

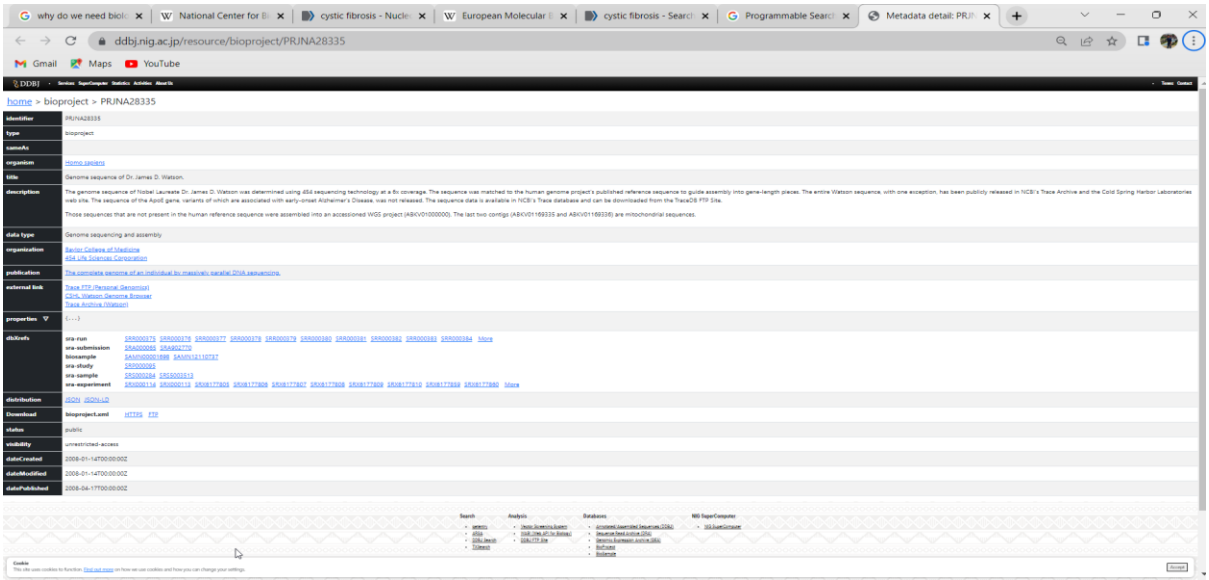
DDBJ:- DNA Databank of Japan

- It’s a biological database which collects DNA sequences.
- It is located at the National Institute of Genetics (NIG) in the Shizuoka Prefecture of Japan.
- It’s funded by the Japan Government
- It accepts sequences all around the globe.
- It exchanges it’s ideas with the Genbank and the EMBL.
- DDBJ began data bank activities in 1987

DDBJ Entry Page:



Details of the searched disease:



Steps to enter into DDBJ:

1. Enter Google.
2. URL: https://ddbj.nig.ac.jp/resource/bioproject/PRJNA28335
3. Enter the Disease for which you need to find the DNA.
4. Download the DNA sequence and save it in a file.

Conclusion:

I would like to conclude by saying that Biological databases play a central role in bioinformatics because, they offer scientists the opportunity to access a wide variety of biologically relevant data, hence

why data bases are useful for the society.