



**DATE** : 30 June 2022

**PROJECT NAME** : Bioinformatics

**GROUP** : D

**GROUP MEMBERS** : Aachal Tiwari,  
Shraddha Mishra, Chinmayee Shinde,  
Niranjan Kathavate, Shraddha Dubey,  
Akshita Poojary, Sancia Fernandes,  
Sanika Naik.

**30 June,2022**

**Bhavans College, Mumbai :**

Bioinformatics was one among the three projects that were assigned to the interns. Bioinformatics, as related to genetics and genomics, is a scientific subdiscipline that involves using computer technology to collect, store, analyze and disseminate biological data and information, such as DNA and amino acid sequences or annotations about those sequences. Respected Sir Anubrata Das conducted online sessions giving us a vast insight on the topic of bioinformatics and gave us hands on experience increasing our practical skills.

During the tenure of the project we go an overview about systems biology, various databases (FungiDB, NCBI etc), Polymerase Chain Reaction, PCR Primer Design Guidelines, real time PCR machine, sanger sequencing, data flow in software, scoring matrix, number of possible alignments, BLAST, FASTA and multiple sequence alignment.

Inorder to gain practical experience we were assigned with various assignments that involved preparing a group presentation on primary data bases (NCBI/EBI/DDBJ) and databases on fungi ,also each group had to select a genus and each student from the group a particular genera and download one sequence of 18S rDNA each from NCBI –our group selected *penicillium* as the genus and *penicillium samsonianum*, *penicillium hennertii*, *penicillium chrysogenum*, *penicillium citrinum*, *penicillium expansum*, *penicillium viticola* as the genera after which each group members downloaded one sequence from NCBI, we then ran the sequences through BLASTn and downloaded any five 18S rDNA sequences, further in Clustal Omega we pasted the five sequences and downloaded the alignment file and opened it in jalview software. As our final assignment we compiled the five 18S rDNA sequences of each genera downloaded by each intern, got a single alignment sequence after running it through Clustal Omega , downloaded it and ran it through jalview software creating a Multiple Alignment Sequence for our group.







**NIH** National Library of Medicine  
National Center for Biotechnology Information Log in

All Databases  Search

**NCBI Home**

**Resource List (A-Z)**

- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Training & Tutorials
- Variation


**Welcome to NCBI**

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

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)


**Submit**

Deposit data or manuscripts into NCBI databases




**Download**

Transfer NCBI data to your computer




**Learn**

Find help documents, attend a class or watch a tutorial




**Develop**

Use NCBI APIs and code libraries to build applications




**Analyze**

Identify an NCBI tool for your data analysis task



**Research**

Explore NCBI research and collaborative projects



**Popular Resources**

- PubMed
- Bookshelf
- PubMed Central
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

**NCBI News & Blog**

GenBank Release 250.0 is available! 30 Jun 2022

GenBank release 250.0 (6/17/2022) is now available on the NCBI FTP site. [This](#)

Introducing NLM's new NCBI Data sets genome page! 29 Jun 2022

As part of the NIH Comparative

Join Us at the ISMB Codeathon- Tools for Sharable Protein Analysis 28 Jun 2022

NLM's NCBI is gearing up for the Tools for [More...](#)

**COVID-19 Information**

[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

**FOLLOW NCBI**






Connect with NLM





National Library of Medicine  
8 600 Rockville Pike  
Bethesda, MD 20894

Web Policies  
FOIA  
HHS Vulnerability Disclosure

Help  
Accessibility  
Careers

NLM | NIH | HHS | USA.gov

## NCBI DATABASE

**NIH** National Library of Medicine  
National Center for Biotechnology Information Log in

**BLAST** <sup>®</sup> » **blastn suite** Home Recent Results Saved Strategies Help

Standard Nucleotide BLAST

blastn
blastp
blastx
tblastn
tblastx

BLASTN programs search nucleotide databases using a nucleotide query. more...

Reset page
Bookmark

**Enter Query Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) Clear

Query subrange ?

From

To

Or, upload file  No file chosen ?

Job Title

Enter a descriptive title for your BLAST search ?

Align two or more sequences ?

**Choose Search Set**

Database  Standard databases (nr etc.):  rRNA/ITS databases  Genomic + transcript databases  Betacoronavirus

Nucleotide collection (nr/nt) ?

Organism Optional

Enter organism name or id—completions will be suggested   exclude Add organism

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown ?

Models (XM/XP)  Uncultured/environmental sample sequences

Sequences from type material

Entrez Query Optional

Enter an Entrez query to limit search ? YouTube Create custom database

**Program Selection**

Optimize for  Highly similar sequences (megablast)

More dissimilar sequences (discontiguous megablast)

Somewhat similar sequences (blastn)

Choose a BLAST algorithm ?

BLAST
Search database Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences)

Show results in a new window

+ Algorithm parameters

## blastn

Tools > Multiple Sequence Alignment > Clustal Omega

**EMBL-EBI to be HTTPS by default from 1st October**  
 On the 1st October the majority of services hosted on www.ebi.ac.uk will be served over HTTPS by default. Services that are becoming HTTPS by default will automatically redirect users accessing the site on insecure HTTP URLs to secure HTTPS URLs. Users of EMBL-EBI services may wish to update links, bookmarks or API clients to use the HTTPS URLs.

## Multiple Sequence Alignment

Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between **three or more** sequences. For the alignment of two sequences please instead use our [pairwise sequence alignment tools](#).

**Important note:** This tool can align up to 4000 sequences or a maximum file size of 4 MB.

STEP 1 - Enter your input sequences

Enter or paste a set of

PROTEIN

sequences in any supported format:

## CLUSTAL OMEGA

The screenshot displays the Jalview 2.6 interface. The primary window, titled 'MAFFT Multiple Sequence Alignment of Retrieved from Uniprot', shows a multiple sequence alignment of various FER1 domain proteins. The alignment is color-coded by amino acid type. Below the alignment are three horizontal bars: 'Conservation' (yellow), 'Quality' (yellow), and 'Consensus' (black). The consensus sequence is shown as 'KVRLITPFGQIQEFCFDDVFLDAAEEAGTDLFYSCKRAGSLSKAGAVVSGSVDDQSD'. A secondary window on the right shows a dendrogram titled 'Average distance tree using B', illustrating the hierarchical clustering of the sequences. A third window in the foreground shows a 3D ribbon diagram of a protein structure, with specific residues highlighted in different colors (yellow, orange, pink, purple, green).

## JALVIEW SOFTWARE

