Practical 1: Reading and Visualizing DNA Sequences: Sequence Scanner Software

The Importance of Bioinformatics Studies

- Identification
- Taxonomy
- Evolution

Steps of Molecular Identification

- Extraction, PCR, Electrophoresis, Sequencing, and Bioinformatics Analysis

Choosing the 16S rRNA Gene for Molecular Identification

- The stability of the gene's ends allows the synthesis of universal primers.
- A large database is available online for comparison.
- It is a universal gene present in all living organisms.
- It contains stable regions with a low evolutionary rate and unstable regions with a high evolutionary rate.

Alternatives to the 16S rRNA Gene in Molecular Identification

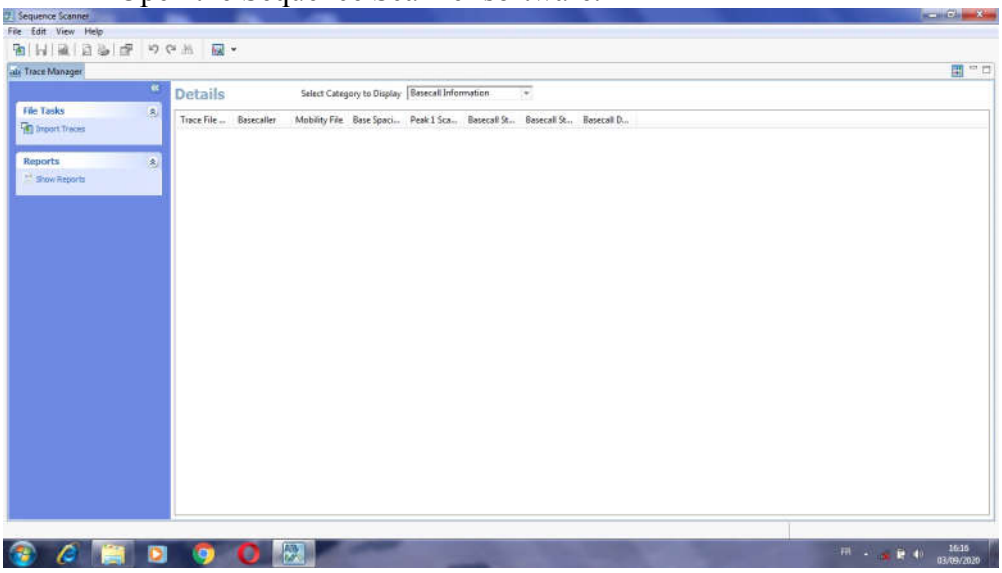
- **tuf** Gene
- Genes encoding: Enzyme, toxin, receptor, hormone.

The Role of Sequence Scanner Software:

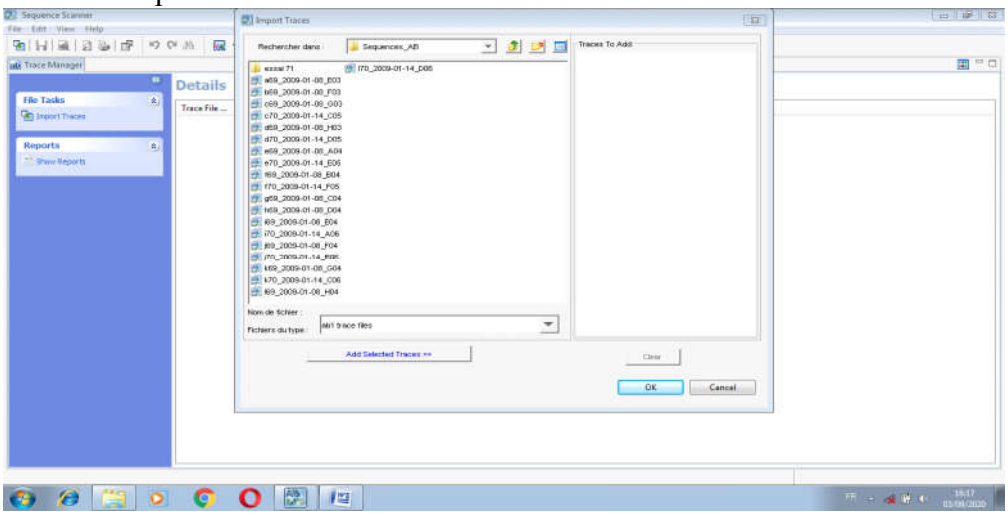
The Sequence Scanner software (Applied Biosystems) allows reading and visualizing AB files from the sequencer.

- Its use involves the following steps:

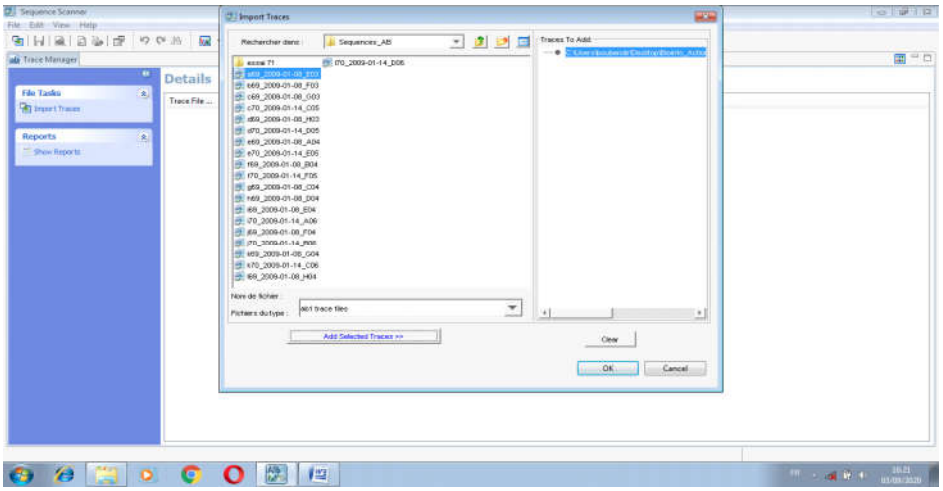
- Open the Sequence Scanner software.



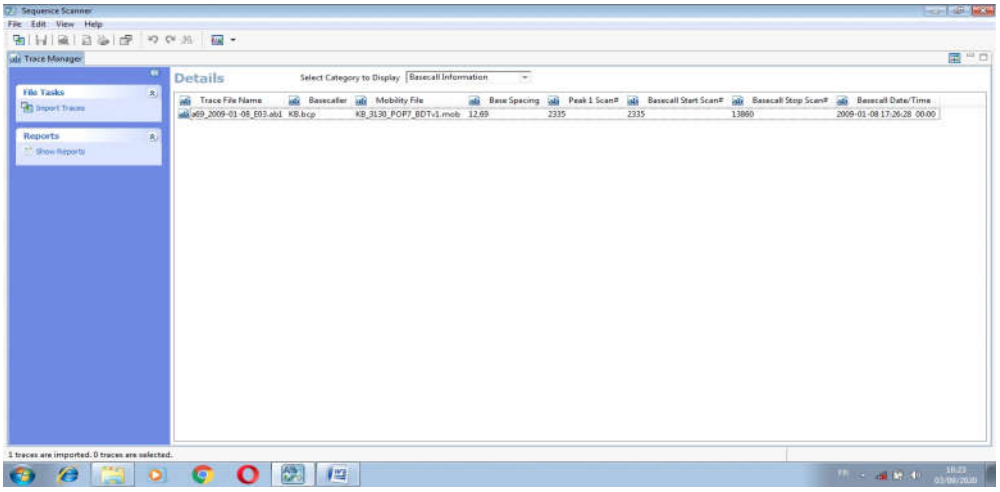
- Click on **Import Traces** (top left) to search for AB files on your computer.



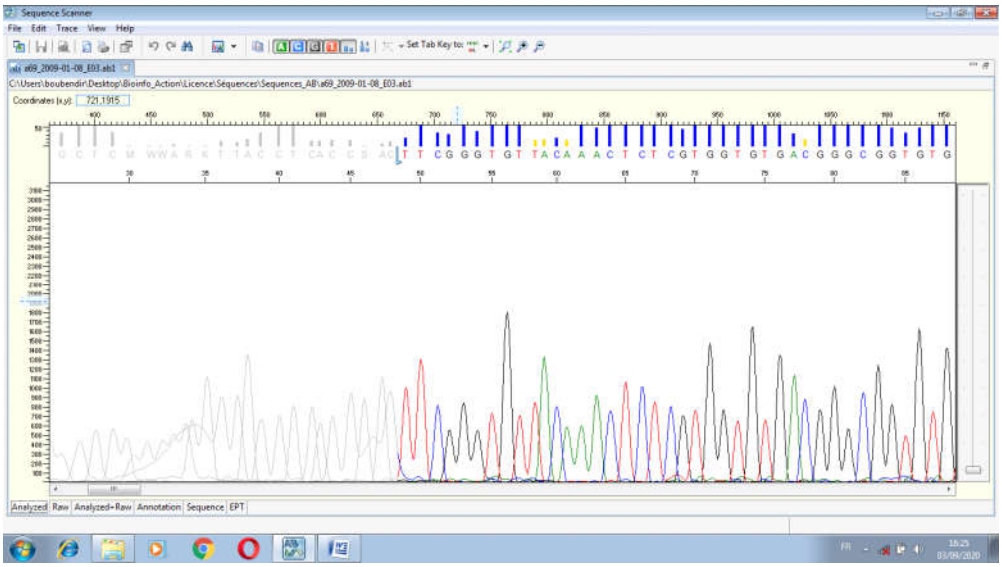
- Select an AB file and click on **Add Selected Traces** (bottom) to import it into the software, then click **OK**.



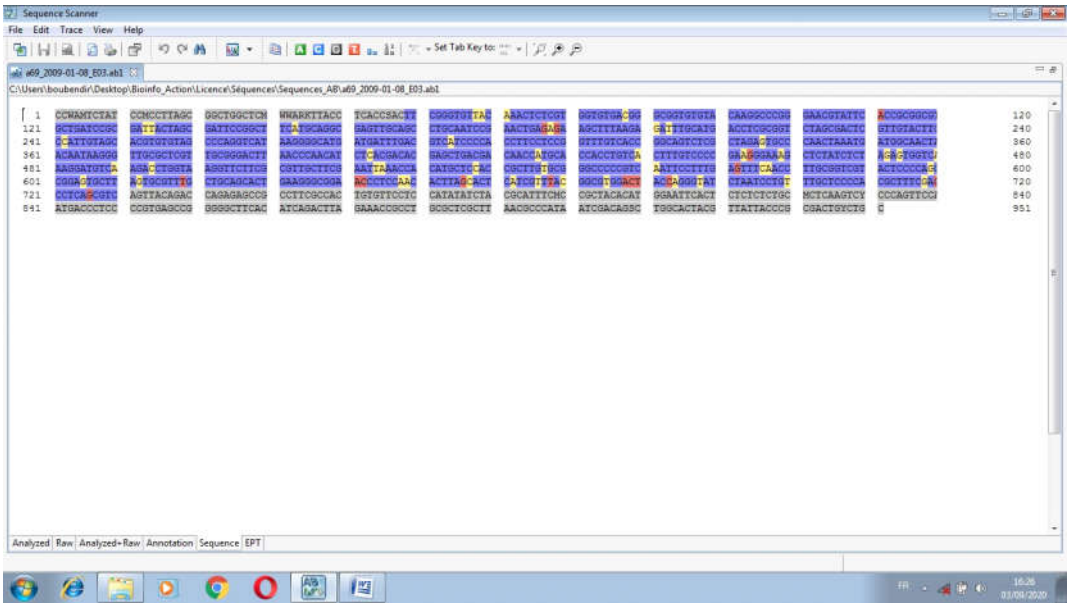
- The AB file is now ready for reading. Double-click on it to open.



- You will see the spectrogram.



- Click on **Sequence** to visualize the details of your DNA.



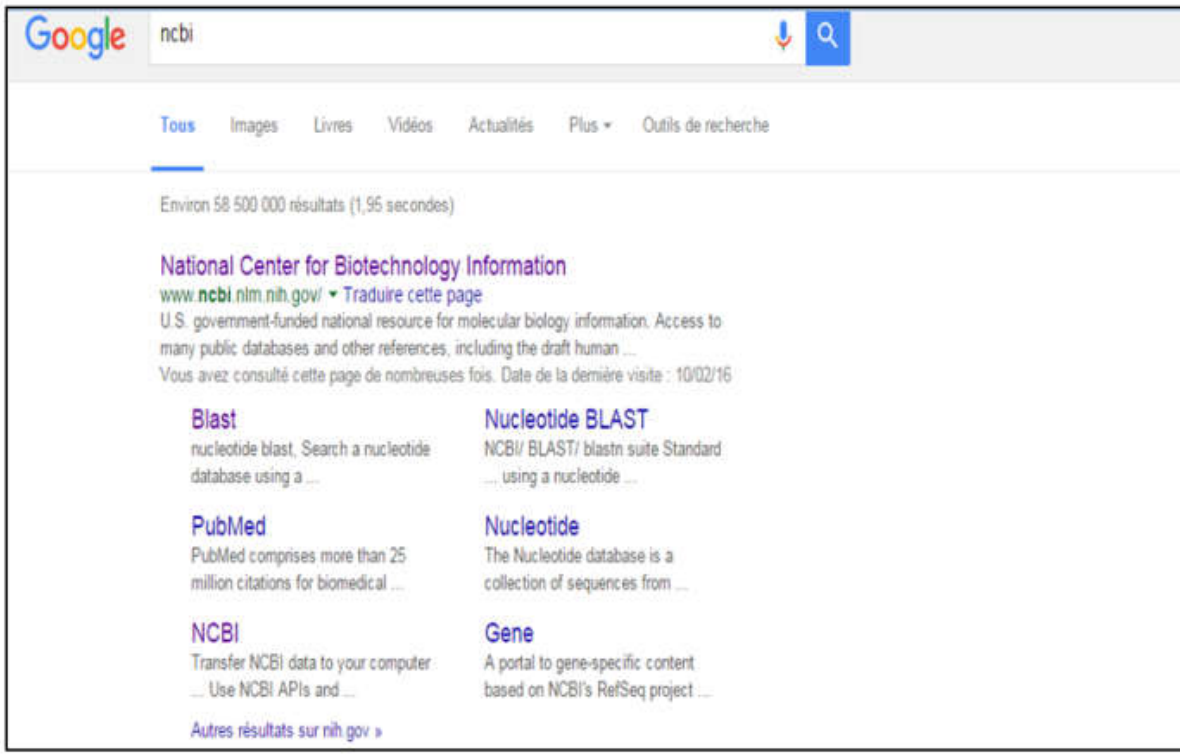
TP 2 : Ouverture du lien NCBI sur internet par l'utilisation du moteur de recherche Google

1. Steps

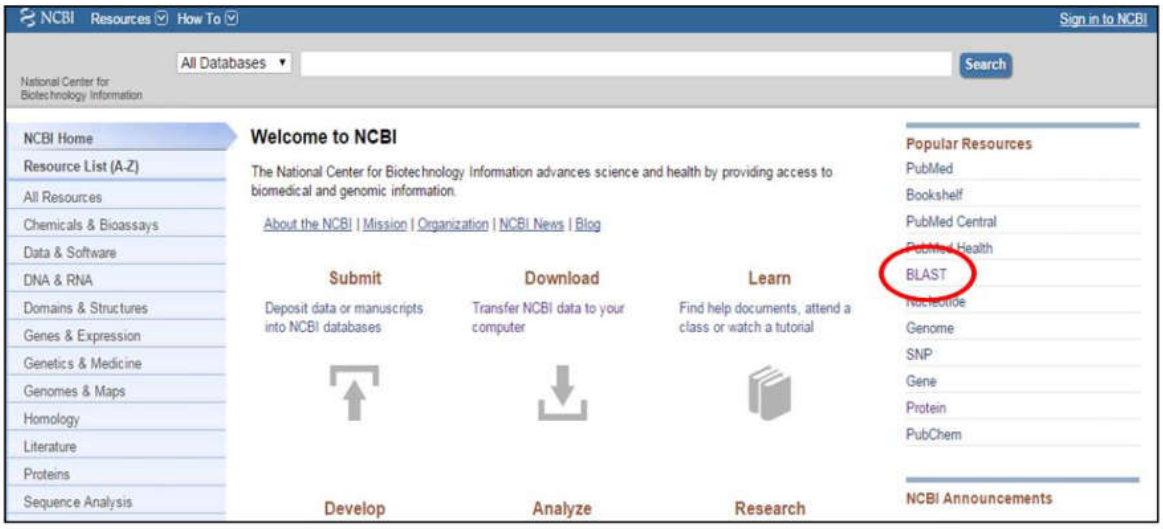
- Open the NCBI link using the Google search engine.
- Select the BLAST program.
- Choose the BLASTn nucleotide tool.
- Insert the DNA sequence or the Accession Number from Gene Bank and activate the BLAST tool.
- Review the list of alignment results.
- Examine the details of the alignment results.
- Collect information about the individual using the Gene Bank Accession Number: author, affiliation, publication, sequence, etc.

2. Step-by-step

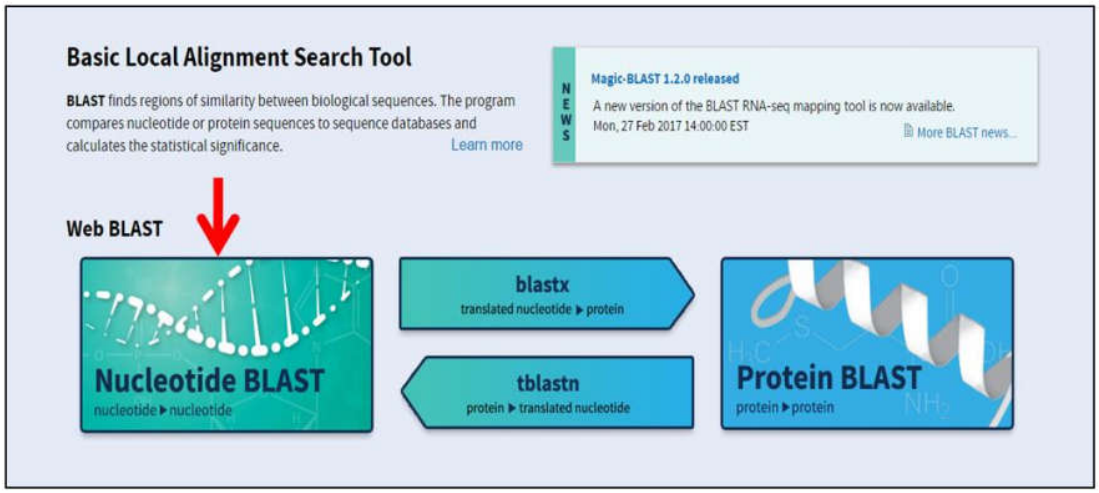
1. Open the NCBI link using Google



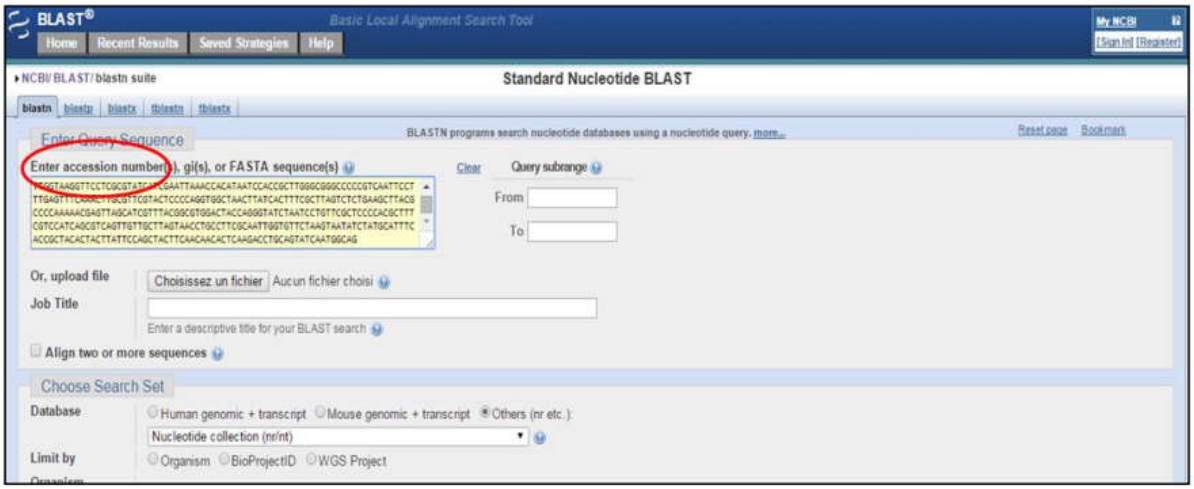
2. **Select the BLAST program.**



3. **Choose the BLASTn nucleotide tool.**

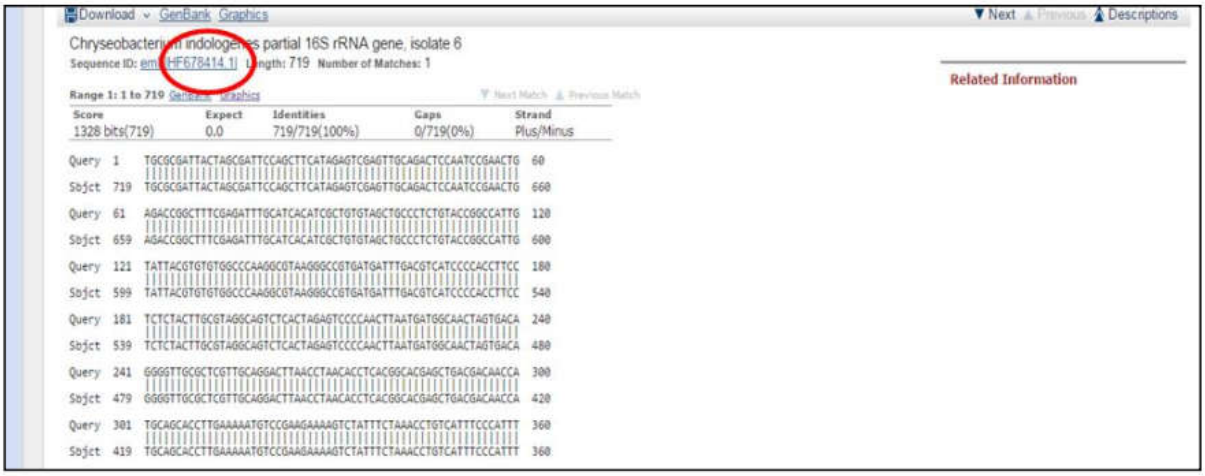


4. **Insert the DNA sequence or the Gene Bank Accession Number and run BLAST.**



5. **Review the list of alignment results.**

6. **Examine the details of the alignment results.**



7. **Gather information about the individual using the Gene Bank Accession Number: author, affiliation, publication, sequence, etc.**

Chryseobacterium indologenes partial 16S rRNA gene, isolate 6

GenBank: HF678414.1

FASTA Graphics

Go to:

```
LOCUS       HF678414               719 bp    DNA     linear   BCT 21-FEB-2013
DEFINITION  Chryseobacterium indologenes partial 16S rRNA gene, isolate 6.
ACCESSION   HF678414
VERSION     HF678414.1  GI:452084714
KEYWORDS    .
SOURCE      Chryseobacterium indologenes
  ORGANISM  Chryseobacterium indologenes
            Bacteria; Bacteroidetes; Flavobacteriia; Flavobacteriales;
            Flavobacteriaceae; Chryseobacterium.
REFERENCE   1
  AUTHORS   Boubendir,A.
  TITLE     Analyse et prevalence du risque infectieux de Listeria
            monocytogenes dans les laits crus recoltes dans deux regions a
            climat different (Zone semi-aride et le Nord-Est algeriens) :
            Modelisation spatiale de la diversite floristique
  JOURNAL   Thesis (2012) Constantine 1 University, Algeria
REFERENCE   2  (bases 1 to 719)
  AUTHORS   Hamidechi,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2013) Constantine University, Constantine, Route
            de Ain El-Bey, 25000, ALGERIA
FEATURES             Location/Qualifiers
     source          1..719
                     /organism="Chryseobacterium indologenes"
                     /mol_type="genomic DNA"
                     /isolate="6"
                     /isolation_source="raw milk"
                     /db_xref="taxon:253"
```

Customize view

Analyze this sequence

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

Related information

Taxonomy

LinkOut to external resources

Ribosomal Database Project II [Ribosomal Database Project II]

SILVA SSU Database [SILVA]

Recent activity

Turn Off Clear

Chryseobacterium indologenes partial 16S rRNA gene, isolate 6 Nucleotide

Nucleotide Sequence (719 letters) BLAST

Practical 3: Introduction to Phylogenetic Analysis:

MEGA6 Software

MEGA6 software is used for phylogenetic analysis and allows you to perform

- Multiple sequence alignment.
- Distance matrix calculation.
- Phylogenetic tree construction.

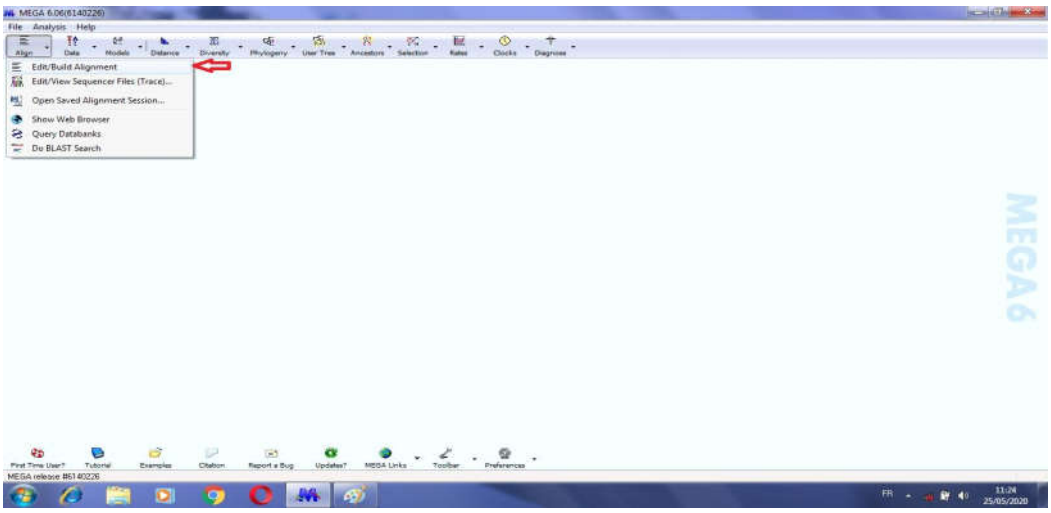
The importance of comparing DNA sequences in multiple alignments

- **Stable regions:** for comparing distant species.
- **Unstable regions:** for comparing closely related species

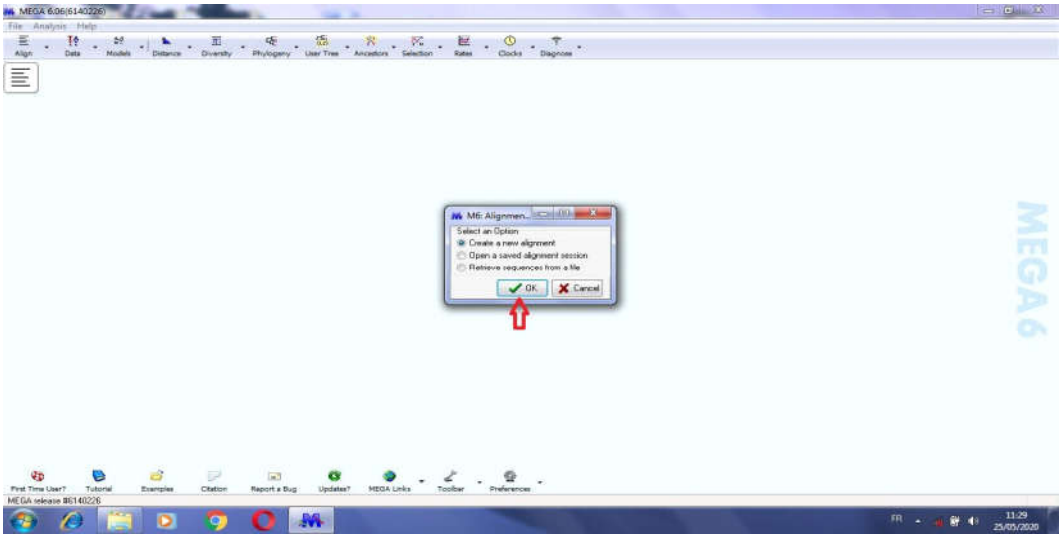
Steps for using the software

1. Multiple sequence alignment: Alignment Explorer

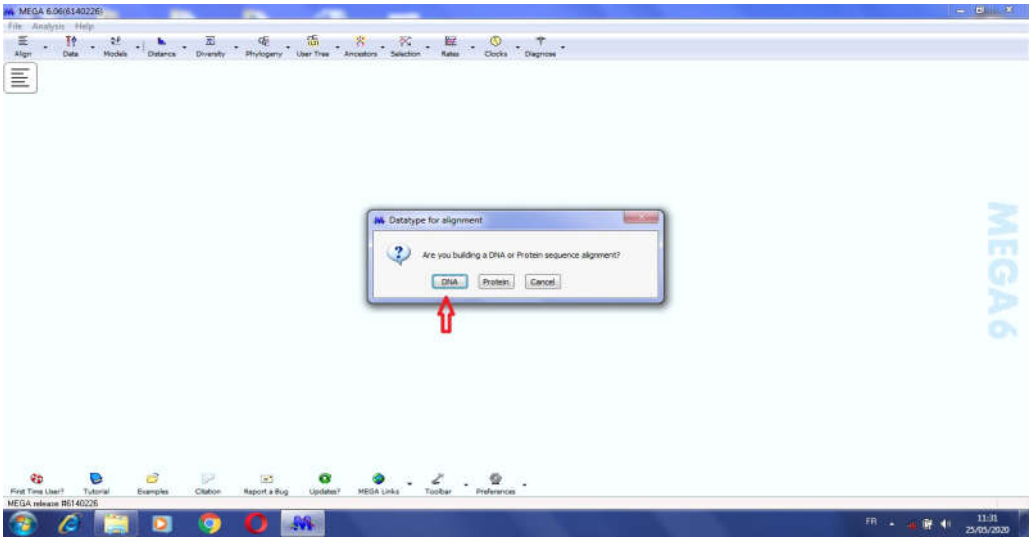
➢ Open the Alignment Explorer program and click on **Edit/Build Alignment**.



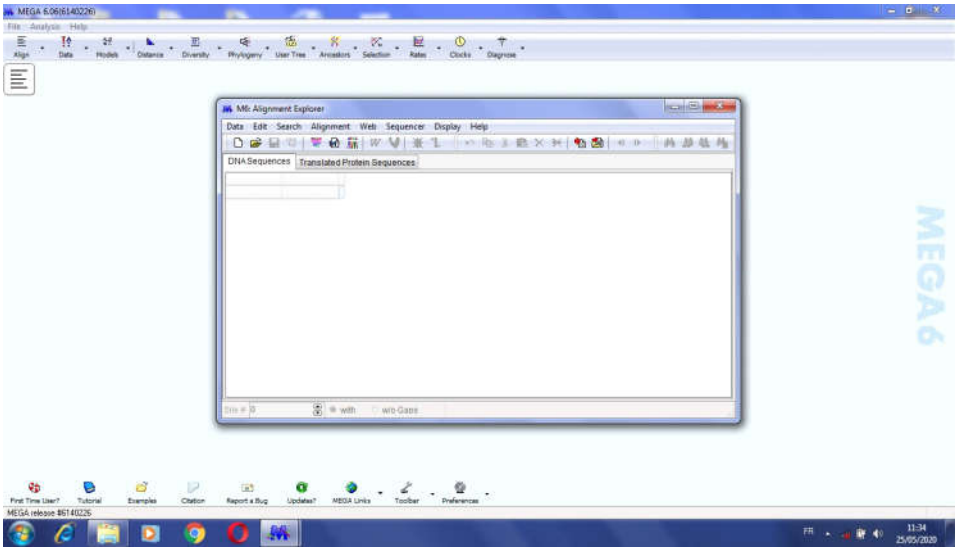
➢ Click **OK** to confirm the creation of a new alignment.



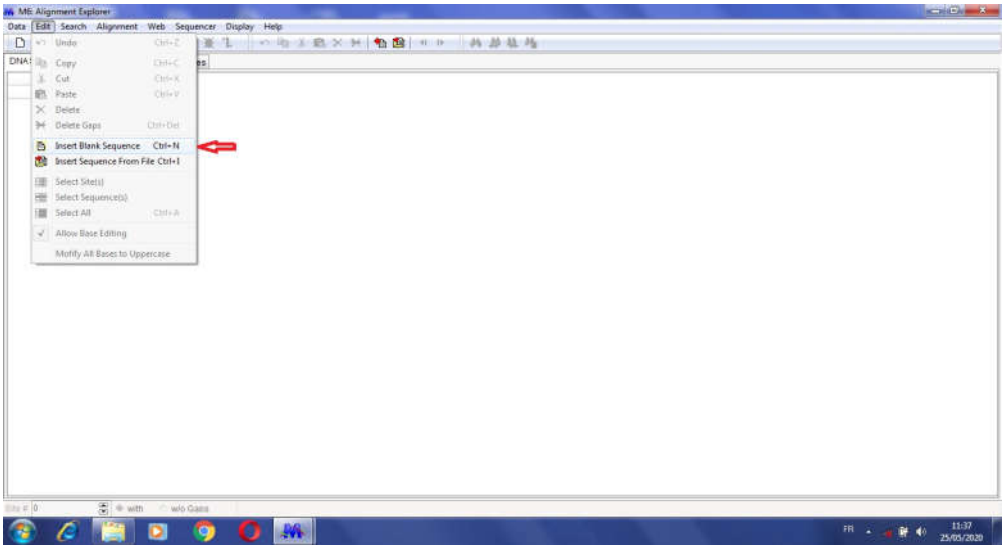
- Confirm your analysis type: DNA or Protein.



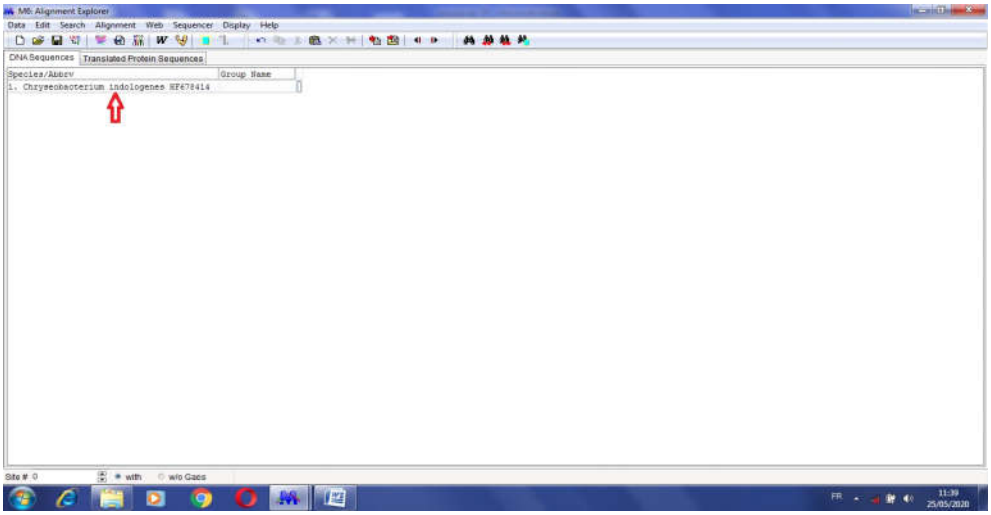
- The Alignment Explorer program opens, enlarge the window.



- Click **Edit_Insert Blank Sequence**.



- Enter the species name and its Gene Bank Accession Number.

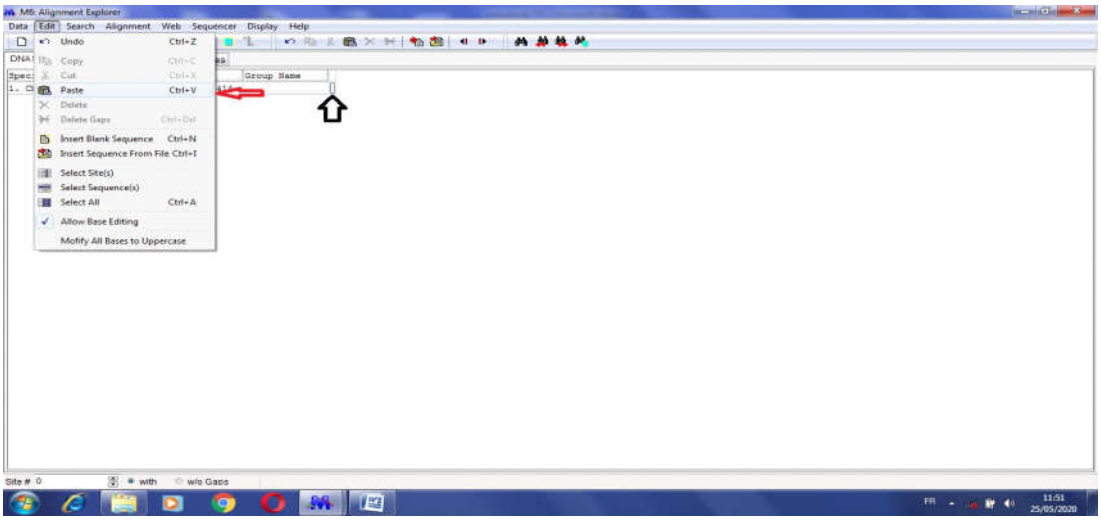


- Select and copy your pre-prepared sequence to analyze.

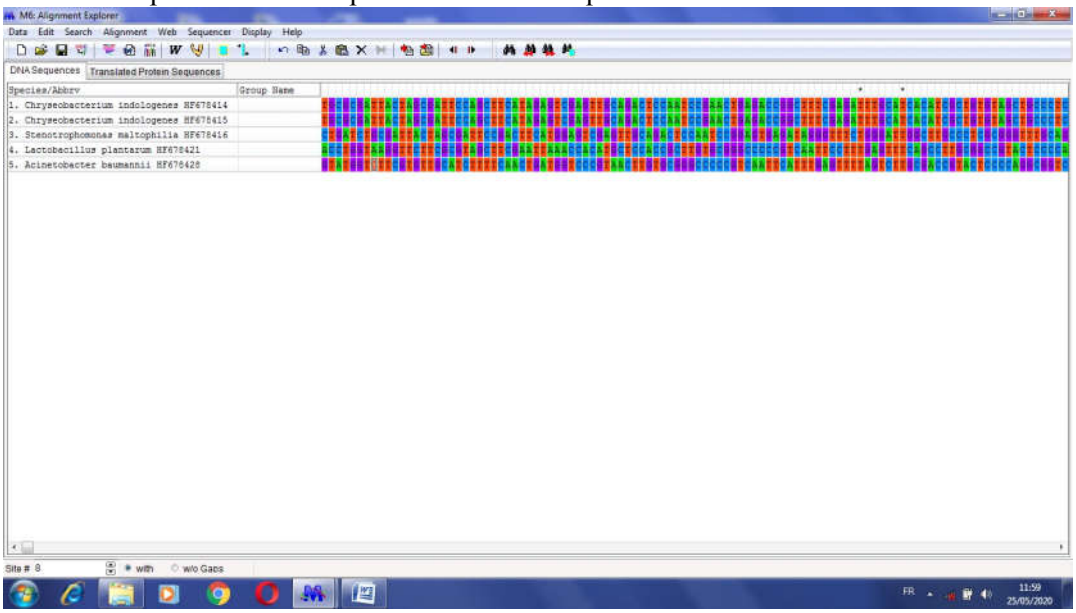
4Hafid(SAMPLE 6) HF678414
Chryseobacterium indologenes 99%
TGCGCGATTACTAGCGATTCCAGCTTCATAGAGTCGAGTTGCAGACTCCAATCCGAACTGAGACCGG
CTTTCGAGATTTGCATCACATCGCTGTGTAGCTGCCCTCTGTACCGGCCATTGTATTACGTGTGTGG
CCCAAGGCGTAAGGGCCGTGATGATTTGACGTCATCCCCACCTTCCTCTCTACTTGCGTAGGCAGTC
TCACTAGAGTCCCCAACTTAATGATGGCAACTAGTGACAGGGGTTGCGCTCGTTGCAGGACTTAACC
TAACACCTCACGGCACGAGCTGACGACAACCATGCAGCACCTTGAAAAATGTCCGAAGAAAAGTCTA
TTTCTAAACCTGTCATTTCCCATTTAAGCCTTGGTAAGGTTCCTCGCGTATCATCGAATTAAACCAC
ATAATCCACCGCTTGGGCGGGCCCCCGTCAATTCCTTTGAGTTTCAAACTTGCGTTCGTACTCCCCA
GGTGGCTAACTTATCACTTTCGCTTAGTCTCTGAAGCTTACGCCCCAAAAACGAGTTAGCATCGTTT
ACGGCGTGGACTACCAGGGTATCTAATCCTGTTCGCTCCCCACGCTTTCGTCCATCAGCGTCAGTTG
TTGCTTAGTAACCTGCCTTCGCAATTGGTGTTCTAAGTAATATCTATGCATTTCACCGCTACACTAC
TTATTCCAGCTACTTCAACAACACTCAAGACCTGCAGTATCAATGGCAG

1Hafid (SAMPLE 3) HF678415
Chryseobacterium indologenes 99%

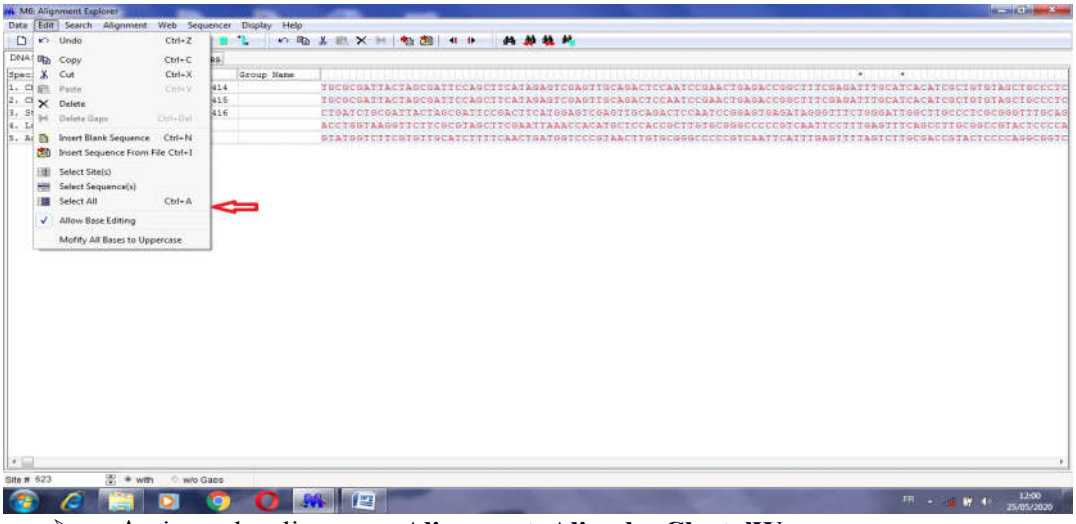
- Click **Edit_Paste** to insert the sequence in the specified location.



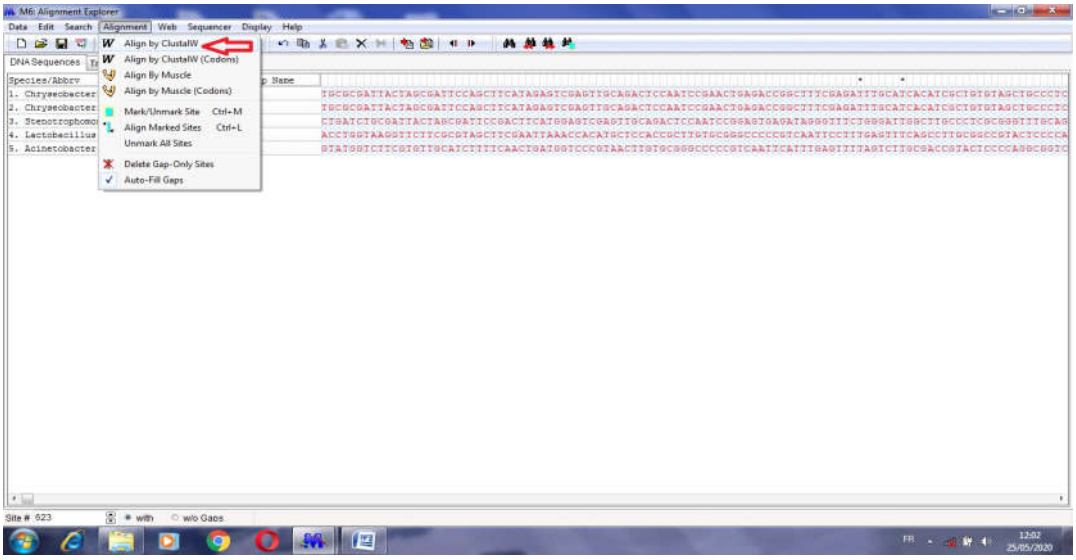
- Repeat the same steps to insert other sequences.



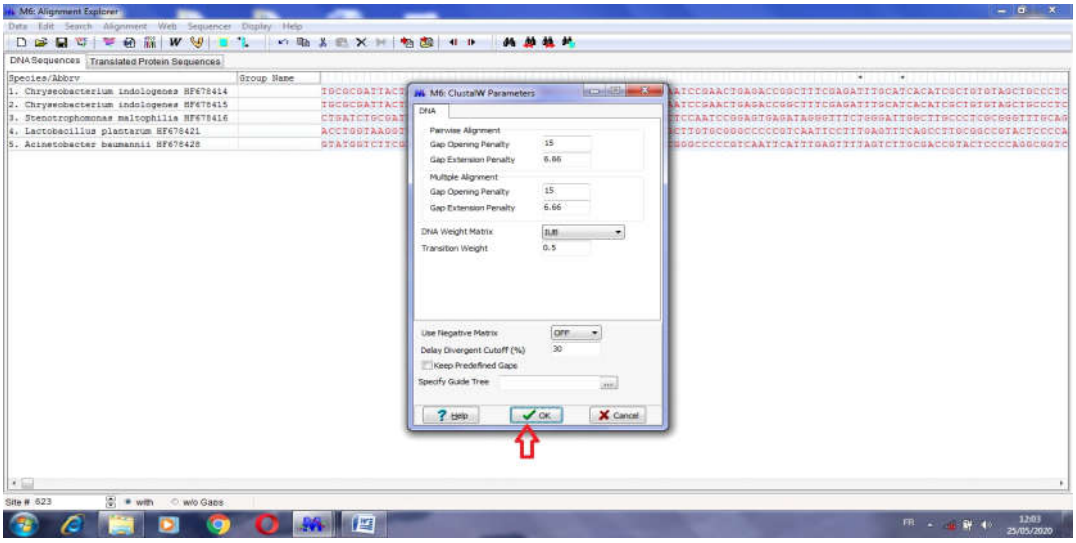
- Select all sequences: **Edit_Select All**.



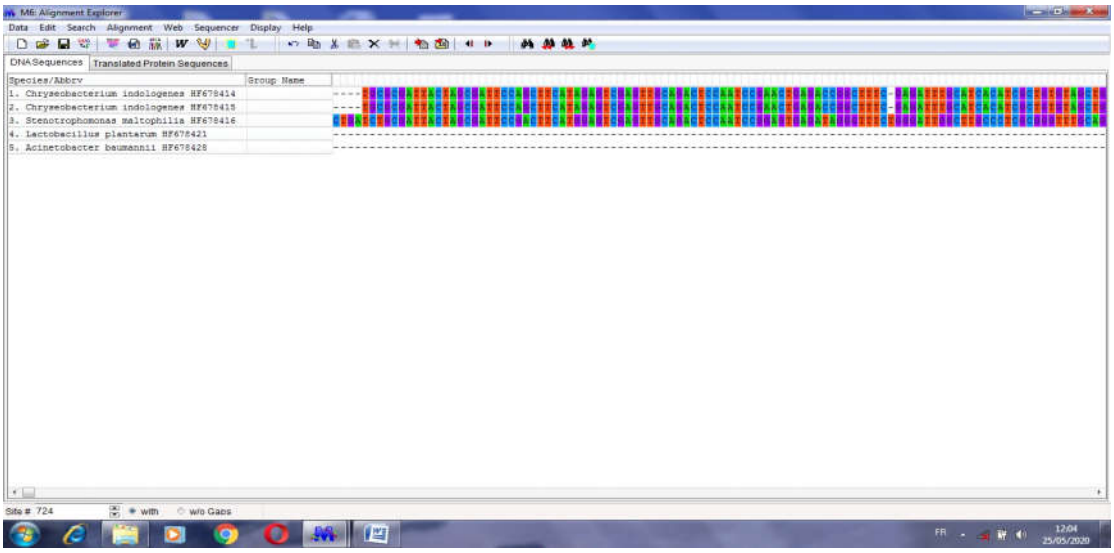
- Activate the alignment: **Alignment_Align by ClustalW**.



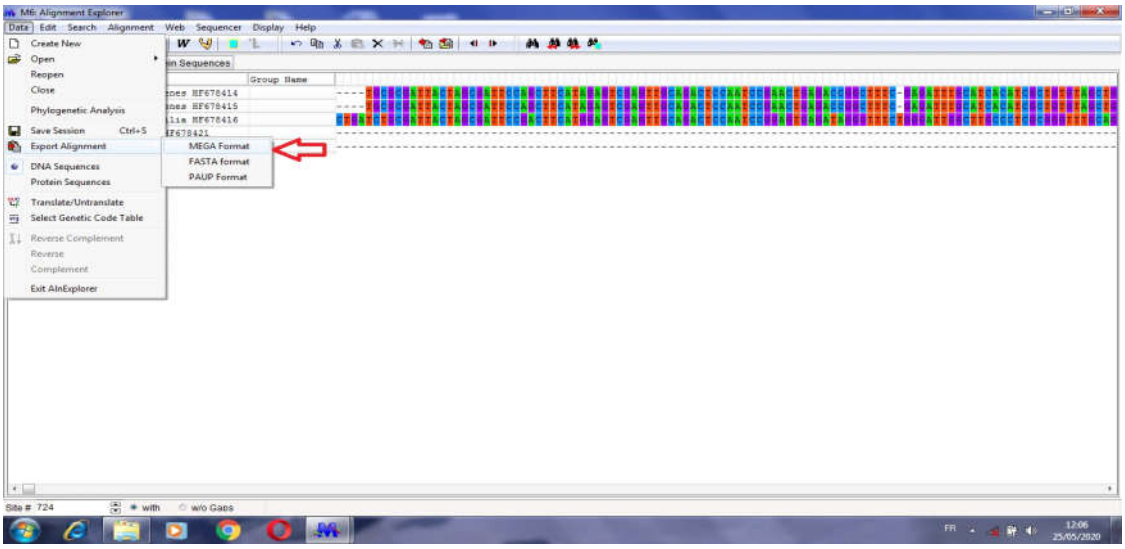
- Confirm the alignment activation: **OK**.



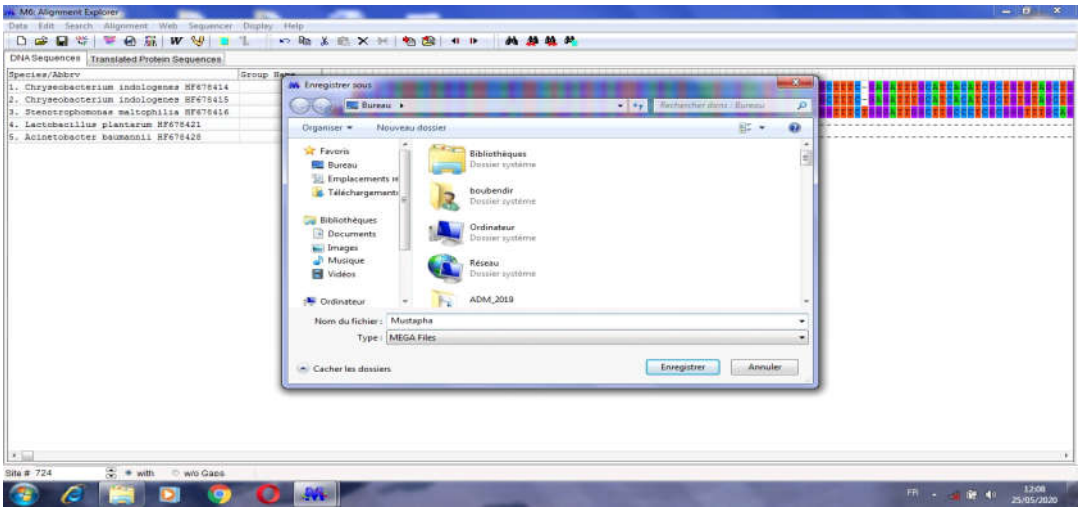
- The appearance of GAPs confirms the alignment is completed.



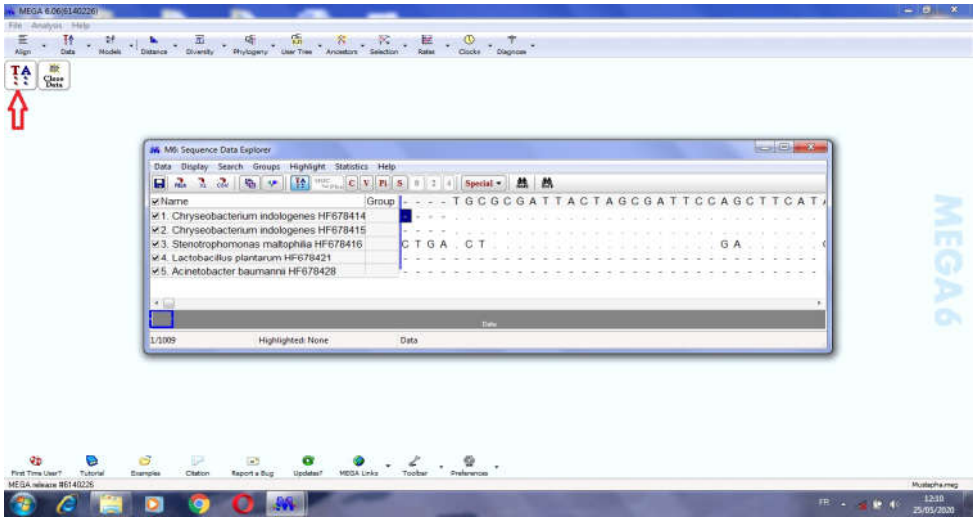
- Save the alignment in MEGA format: **Data_Export Alignment_MEGA Format**.



- Name the MEGA file and save.

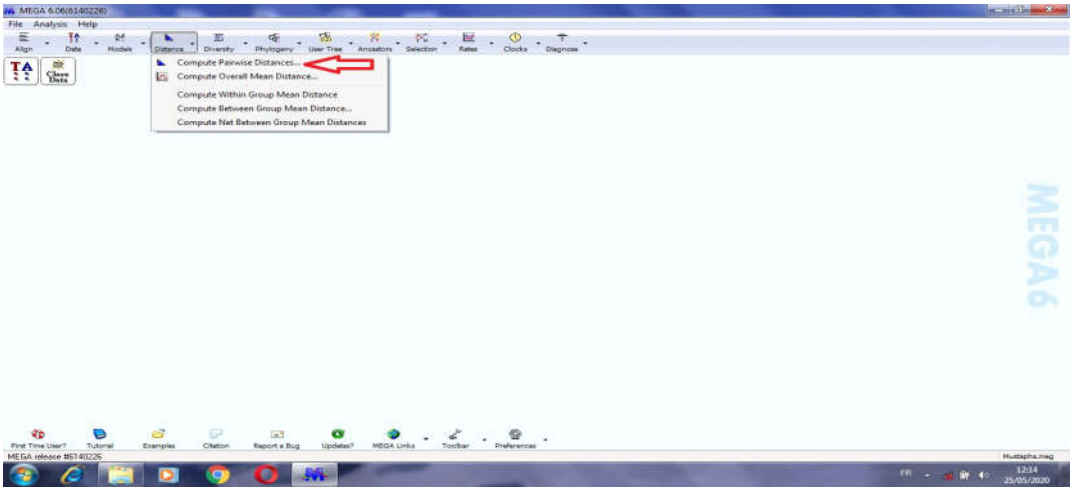


- Finally, open the MEGA file and view your alignment.

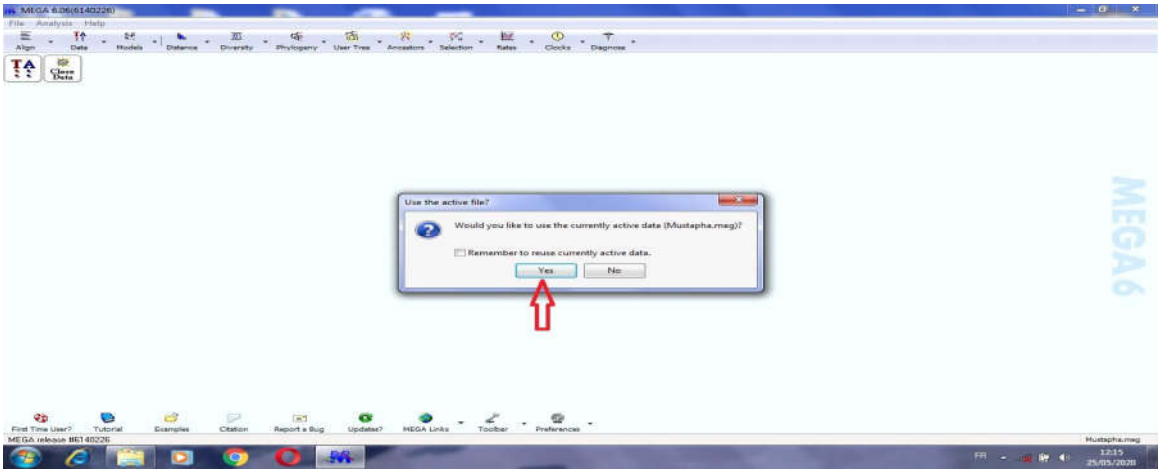


2. Distance matrix: Matrix Distance Explorer

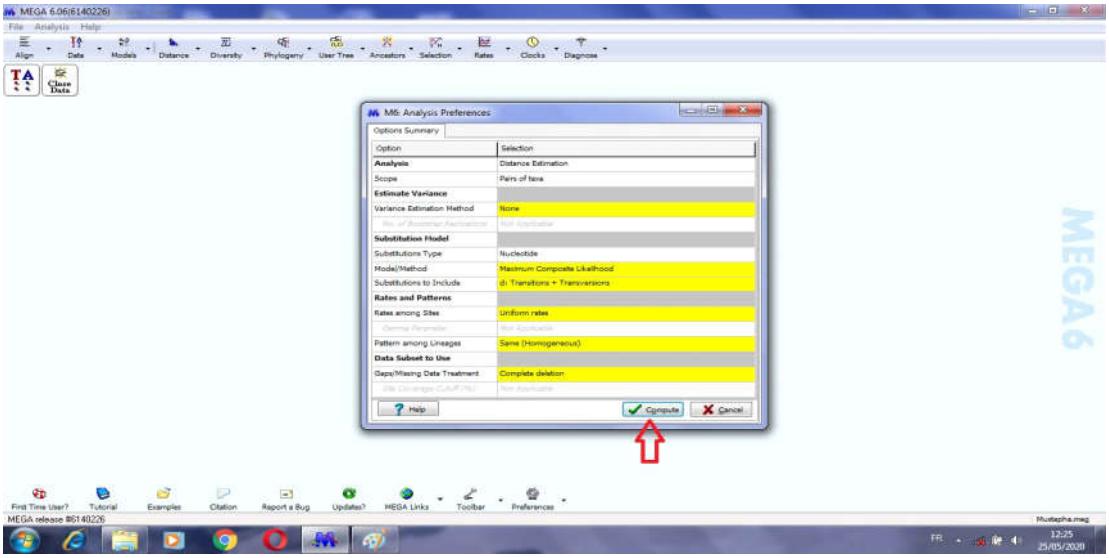
- Open the Matrix Distance Explorer program and select **Compute Pairwise Distance**.



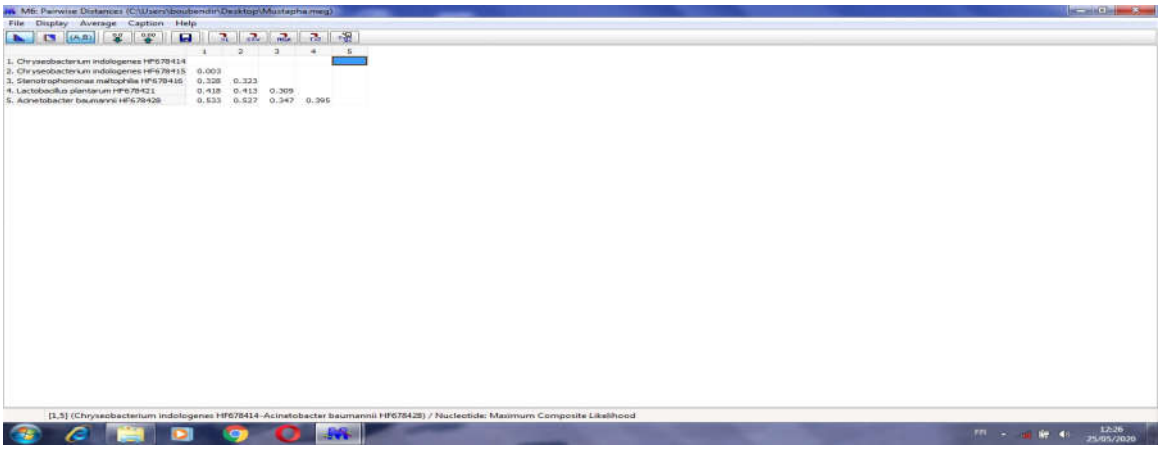
- Confirm the use of active data: **Yes**



- Click **Compute** to generate the distance matrix.

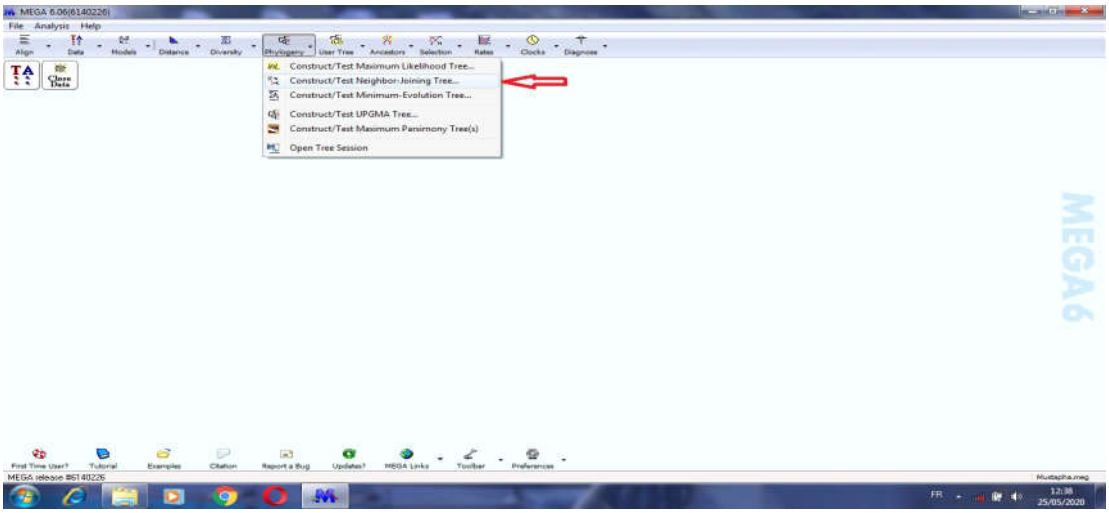


➢ Finally, the distance matrix is displayed.

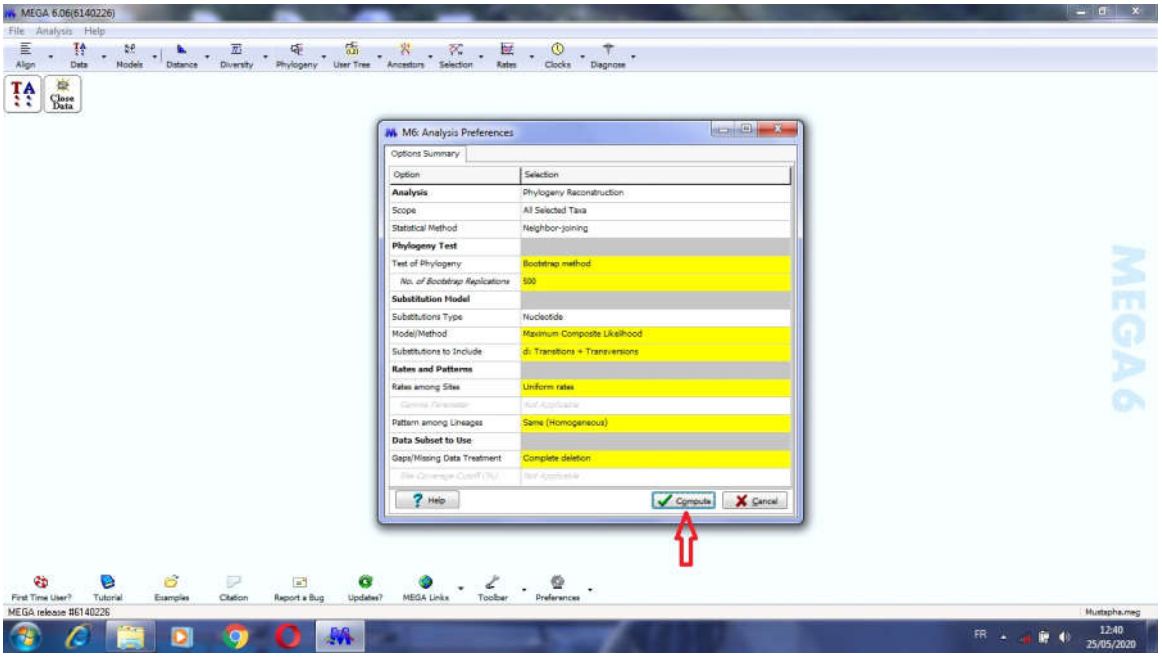


3. Phylogenetic tree topology

➢ Open the Tree Explorer program and select the method to build the tree (NJ or others).



- Start the construction with the bootstrap test.



- Finally, the phylogenetic tree is displayed.