

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

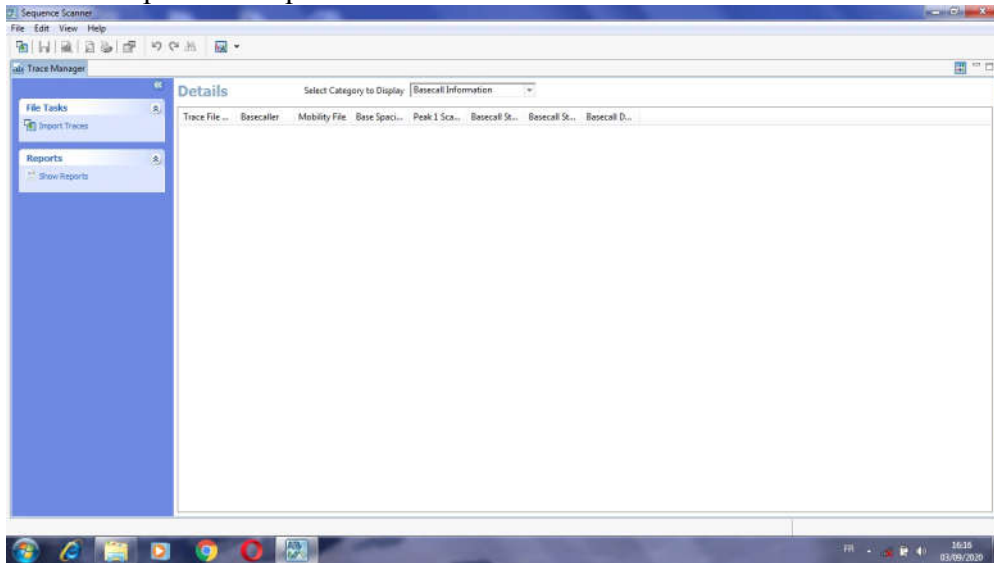
- **tuf** Gene
- Genes encoding: Enzyme, toxin, receptor, hormone.

The Role of Sequence Scanner Software:

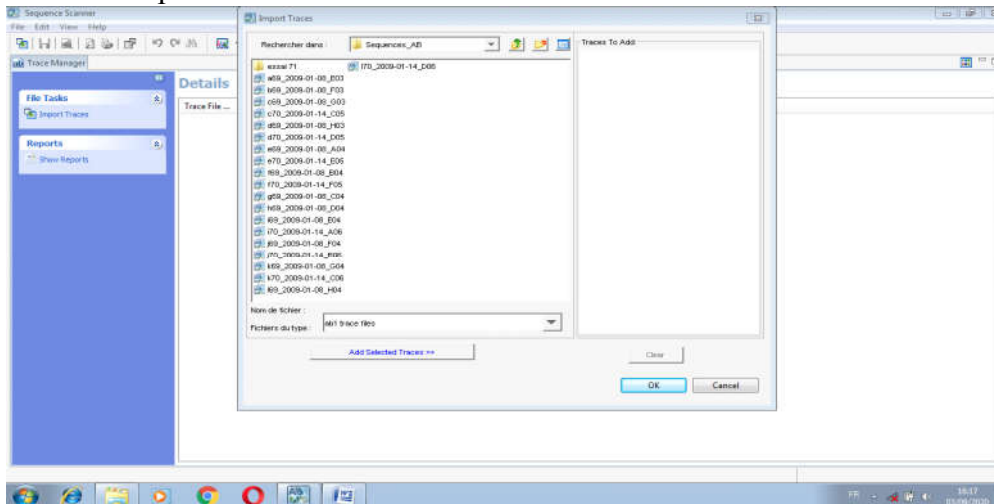
The Sequence Scanner software (Applied Biosystems) allows reading and visualizing AB files from the sequencer.

- Its use involves the following steps:

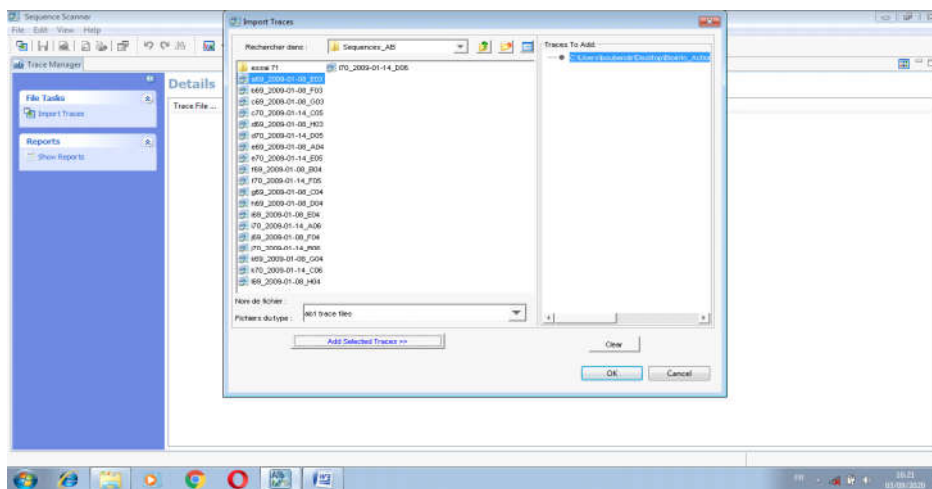
Open the Sequence Scanner software.



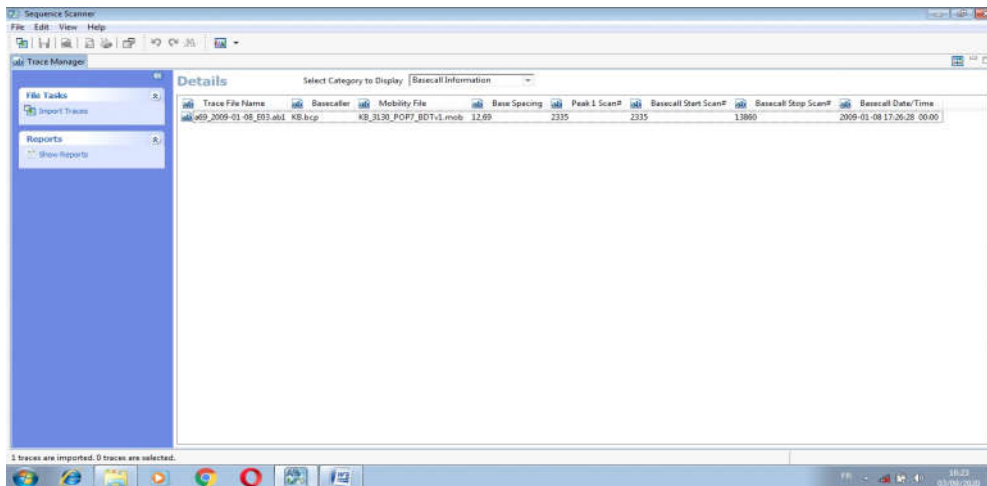
Click on **Import Traces** (top left) to search for AB files on your computer.



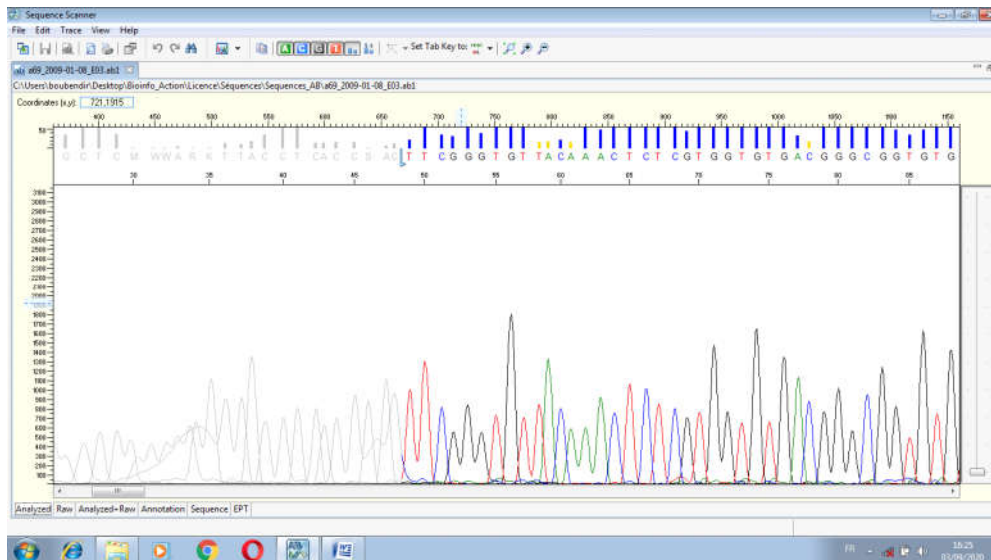
Select an AB file and click on **Add Selected Traces** (bottom) to import it into the software, then click **OK**.



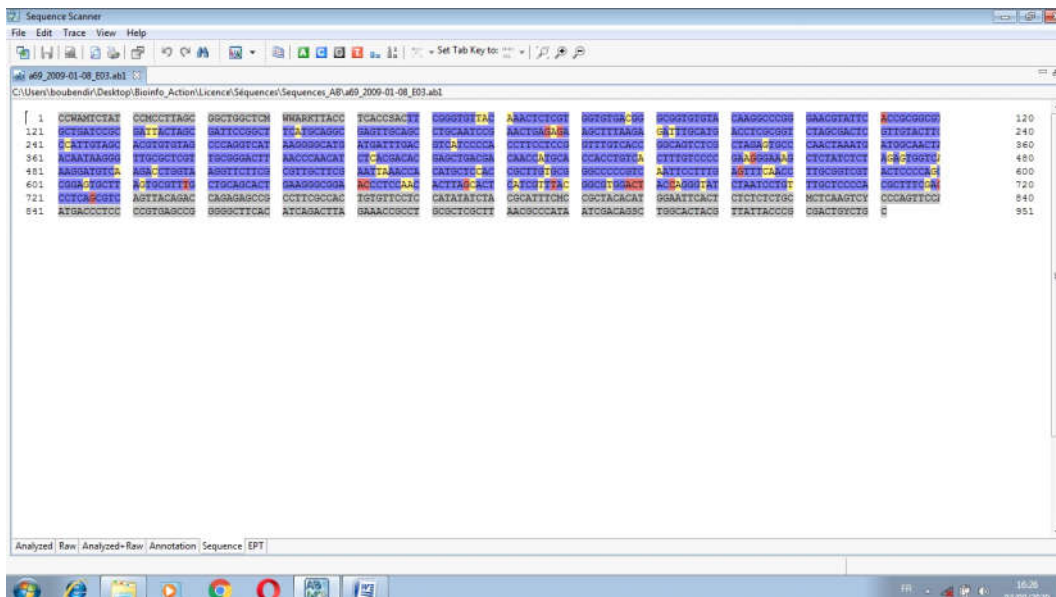
✚ The AB file is now ready for reading. Double-click on it to open.



✚ You will see the spectrogram.



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



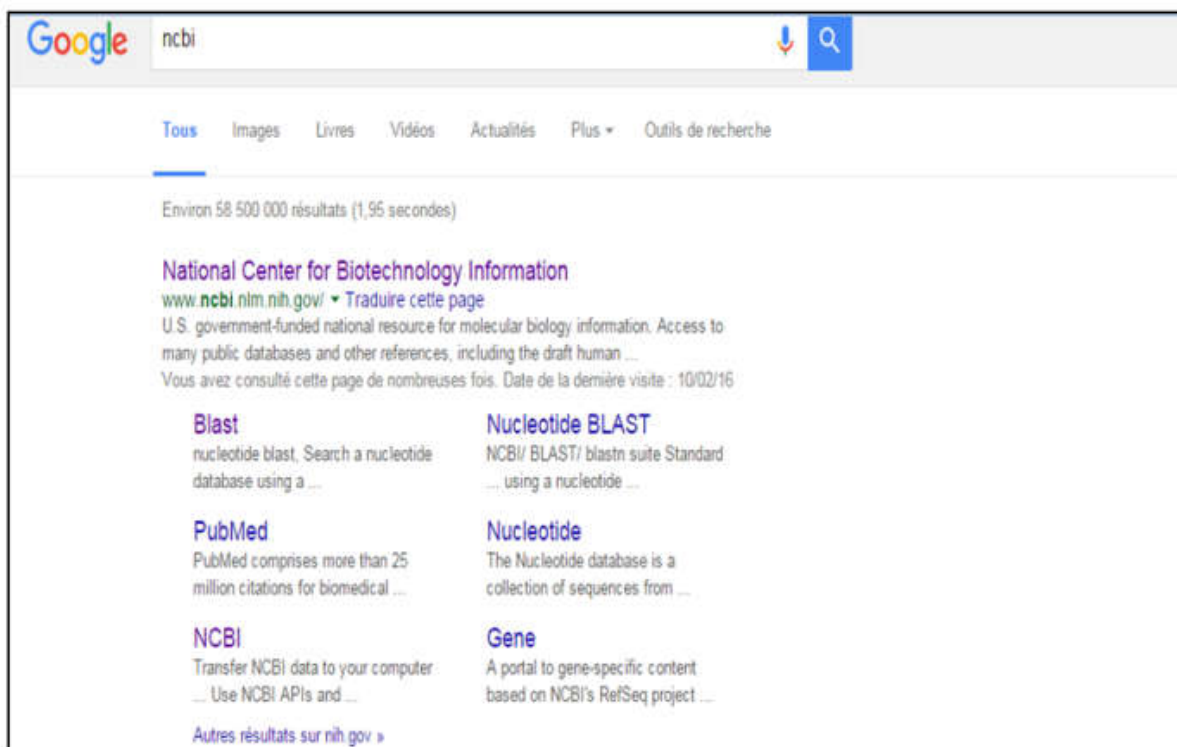
TP 2 : Ouverture du lien NCBI sur internet par l'utilisation du moteur de recherche Google

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.

The screenshot shows the NCBI homepage. On the left is a navigation menu with categories like 'Resource List (A-Z)', 'Chemicals & Bioassays', 'Data & Software', 'DNA & RNA', 'Domains & Structures', 'Genes & Expression', 'Genetics & Medicine', 'Genomes & Maps', 'Homology', 'Literature', 'Proteins', and 'Sequence Analysis'. The main content area features a 'Welcome to NCBI' message and three primary actions: 'Submit' (Develop), 'Download' (Analyze), and 'Learn' (Research). On the right, under 'Popular Resources', the 'BLAST' link is circled in red. Other resources listed include PubMed, Bookshelf, PubMed Central, and PubChem.

3. Choose the BLASTn nucleotide tool.

The screenshot displays the 'Basic Local Alignment Search Tool' interface. A red arrow points to the 'Nucleotide BLAST' option, which is labeled 'nucleotide > nucleotide'. Other options include 'blastx' (translated nucleotide > protein), 'tblastn' (protein > translated nucleotide), and 'Protein BLAST' (protein > protein). A 'NEWS' box on the right announces 'Magic-BLAST 1.2.0 released'.

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

The screenshot shows the 'Standard Nucleotide BLAST' search interface. The 'Enter Query Sequence' field is circled in red and contains a DNA sequence: 'TACGTAAGTTCTCTGCGSTATGACCAATTAAACACATATCCACCGCTTGGGCGGCCCCGTCATTCCTTTGAGTTTGGGCTTGGTACTCCCGAGTGGCTAATCACTTTCCTTATCTCTGAGCTTACGCCCCAAAAGAGTTHGCATGCTTACGGCTGGACTACGAGGATCAATCTCTGCTCCCGACGCTTTGGTCCATCAGGTCAGTGTTCCTAATACCTGCTTCCCAATTGGTGTCTAAGTAATCTCATGATTTCACCGCTACTACTTATTCACGCTACTTCAACCAACTCAAGACTGCAATATCAATGGGAG'. The 'Database' section is set to 'Nucleotide collection (nr/nt)'.

5. Review the list of alignment results.

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

Alignments [Download](#) [GenBank](#) [Graphics](#) [Distance tree of results](#)

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Chryseobacterium indologenes partial 16S rRNA gene, isolate 6	1328	1328	100%	0.0	100%	HF678414.1
<input type="checkbox"/> Chryseobacterium indologenes partial 16S rRNA gene, isolate 12	1323	1323	100%	0.0	99%	HF678419.1
<input type="checkbox"/> Chryseobacterium indologenes partial 16S rRNA gene, isolate 3	1323	1323	100%	0.0	99%	HF678415.1
<input type="checkbox"/> Bacterium 14S134 16S ribosomal RNA gene, partial sequence	1317	1317	100%	0.0	99%	KC734365.1
<input type="checkbox"/> Bacterium 14S132 16S ribosomal RNA gene, partial sequence	1317	1317	100%	0.0	99%	KC734363.1
<input type="checkbox"/> Chryseobacterium enrichment culture clone RA-M137 16S ribosomal RNA gene, partial sequence	1317	1317	100%	0.0	99%	JQ083171.1

6. Examine the details of the alignment results.

Download [GenBank](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

Chryseobacterium indologenes partial 16S rRNA gene, isolate 6
 Sequence ID: [eml:HF678414.1](#) Length: 719 Number of Matches: 1

Range 1: 1 to 719 [Graphics](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
1328 bits(719)	0.0	719/719(100%)	0/719(0%)	Plus/Minus

Query 1 TGC GCGATTACTAGCGATTCCAGCTTCATAGAGTCSA6TTCGAGACTCCAATCCGAACTG 60
 Subject 1 TGC GCGATTACTAGCGATTCCAGCTTCATAGAGTCSA6TTCGAGACTCCAATCCGAACTG 60

Query 61 AGACCGGCTTTCSAGATTTGCATCACATCGCTGTGTAGCTGCCCTCTGTACCGCCATTG 120
 Subject 61 AGACCGGCTTTCSAGATTTGCATCACATCGCTGTGTAGCTGCCCTCTGTACCGCCATTG 120

Query 121 TATTACGTGTGTGGCCCAAGCCGTAAGGCCCCGTGATGATTTGACGTGATCCCACTTCC 180
 Subject 121 TATTACGTGTGTGGCCCAAGCCGTAAGGCCCCGTGATGATTTGACGTGATCCCACTTCC 180

Query 181 TCTCTACTTGCSTAGCAGTCTCACTAGAGTCCCCAACTTAATGATGGCACTAGTGACA 240
 Subject 181 TCTCTACTTGCSTAGCAGTCTCACTAGAGTCCCCAACTTAATGATGGCACTAGTGACA 240

Query 241 GGGGTTGCGCTCGTTGCAAGACTTAACTAACACTCACGGCAGCTGACGACAAACA 300
 Subject 241 GGGGTTGCGCTCGTTGCAAGACTTAACTAACACTCACGGCAGCTGACGACAAACA 300

Query 301 TGCAGCACCTTGAAAAATGCCGAGAAAAGTCTATTTCTAAACCTGTCAATTCCTTCCATTT 360
 Subject 301 TGCAGCACCTTGAAAAATGCCGAGAAAAGTCTATTTCTAAACCTGTCAATTCCTTCCATTT 360

Related Information

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 recoltés 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.

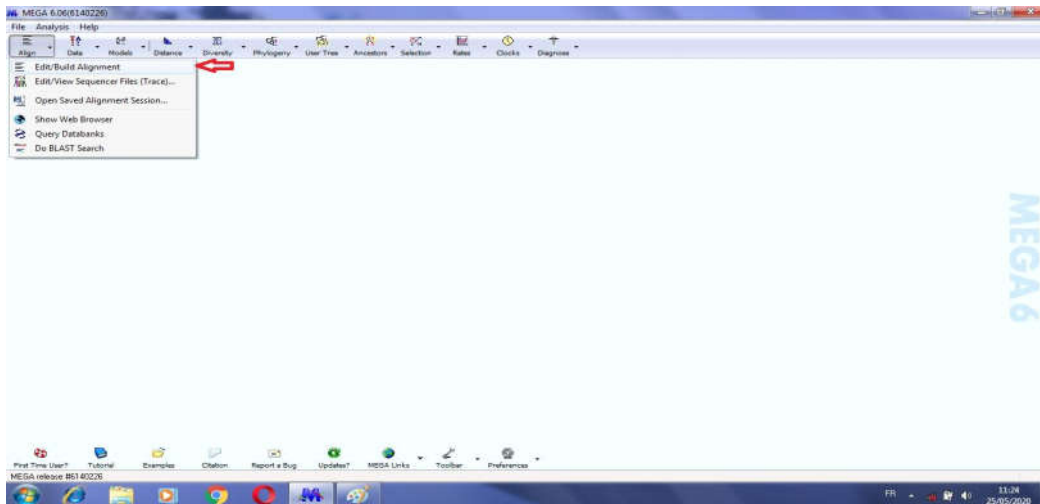
The importance of comparing DNA sequences in multiple alignments

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

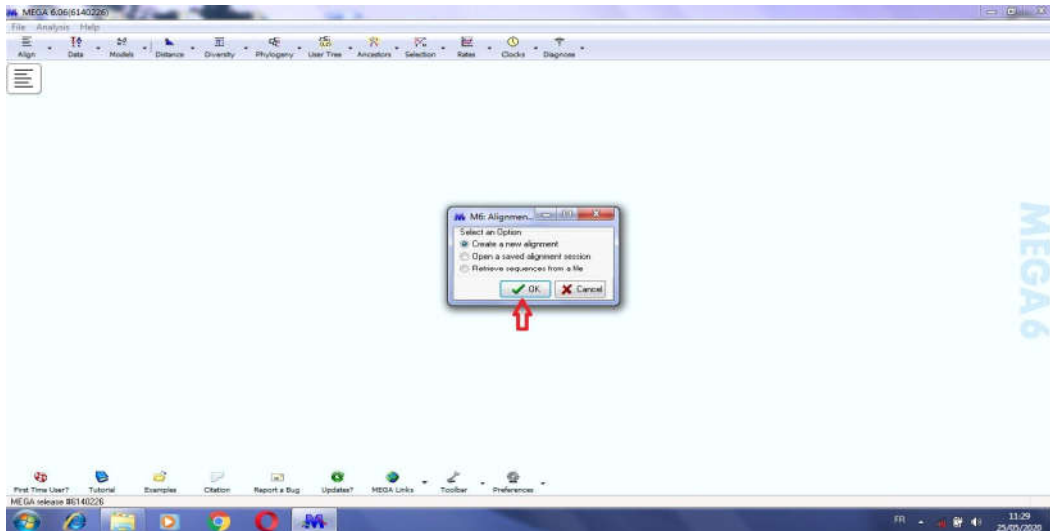
Steps for using the software

1. Multiple sequence alignment: Alignment Explorer

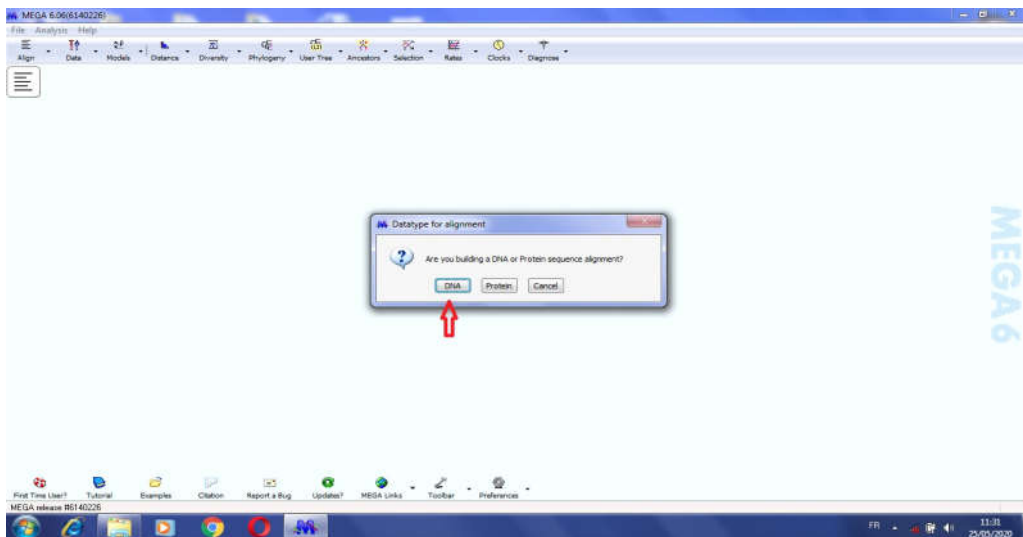
- Open the Alignment Explorer program and click on **Edit/Build Alignment**.



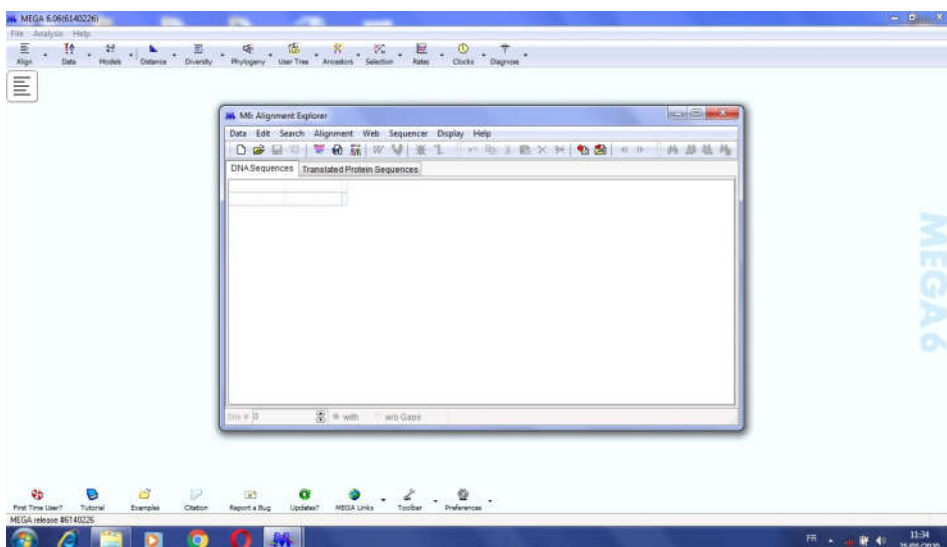
- Click **OK** to confirm the creation of a new alignment.



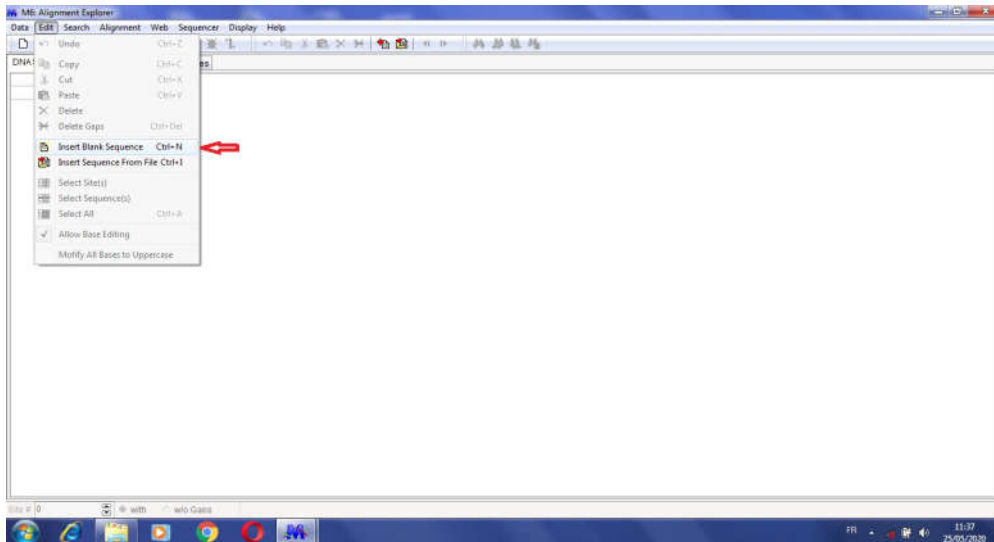
- Confirm your analysis type: DNA or Protein.



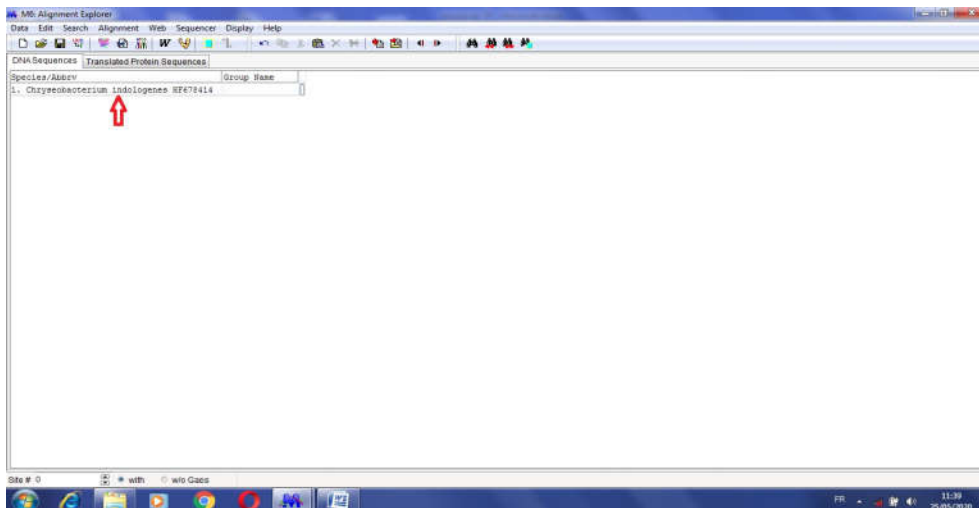
- The Alignment Explorer program opens, enlarge the window.



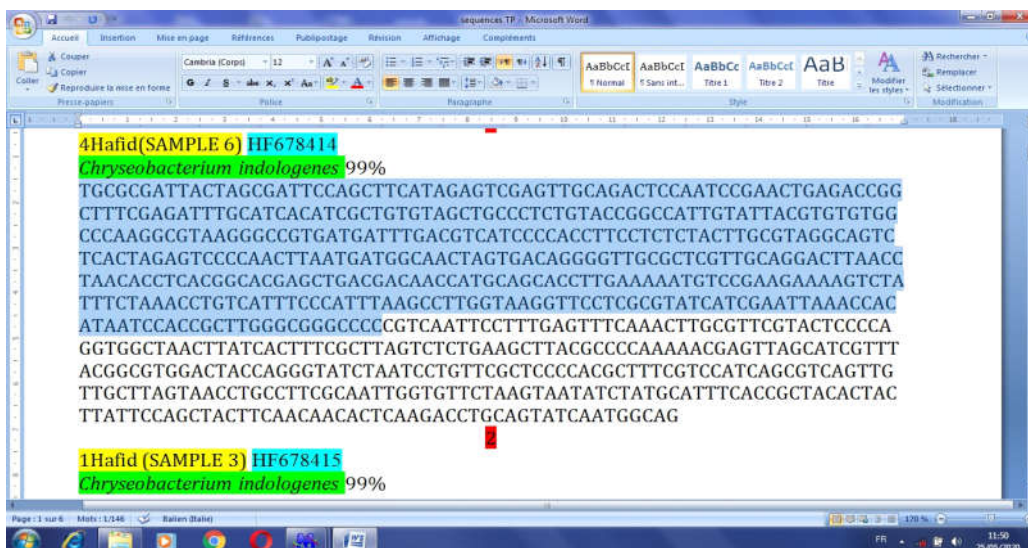
- Click **Edit_Insert Blank Sequence**.



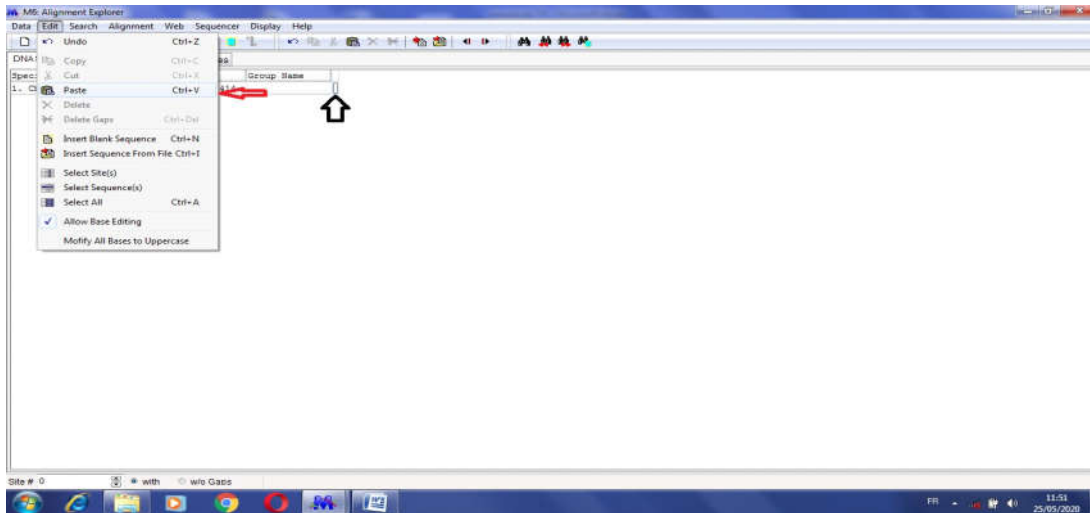
- Enter the species name and its Gene Bank Accession Number.



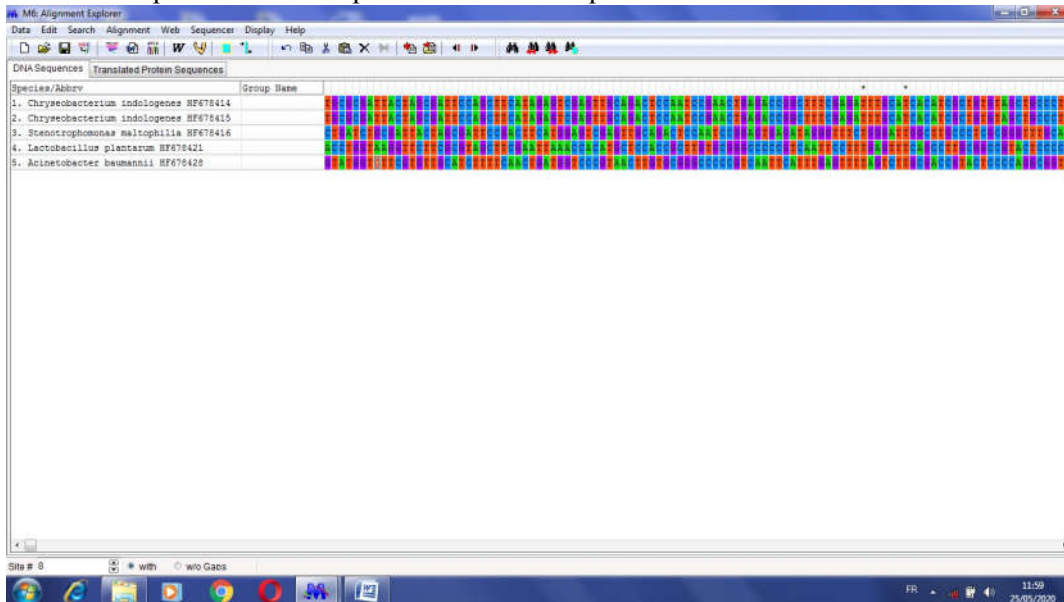
- Select and copy your pre-prepared sequence to analyze.



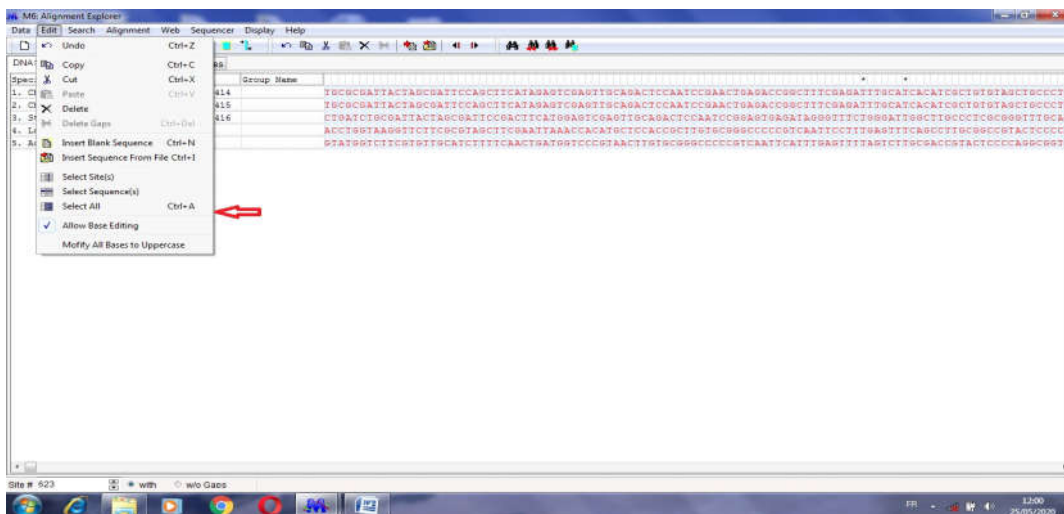
➤ Click **Edit_Paste** to insert the sequence in the specified location.



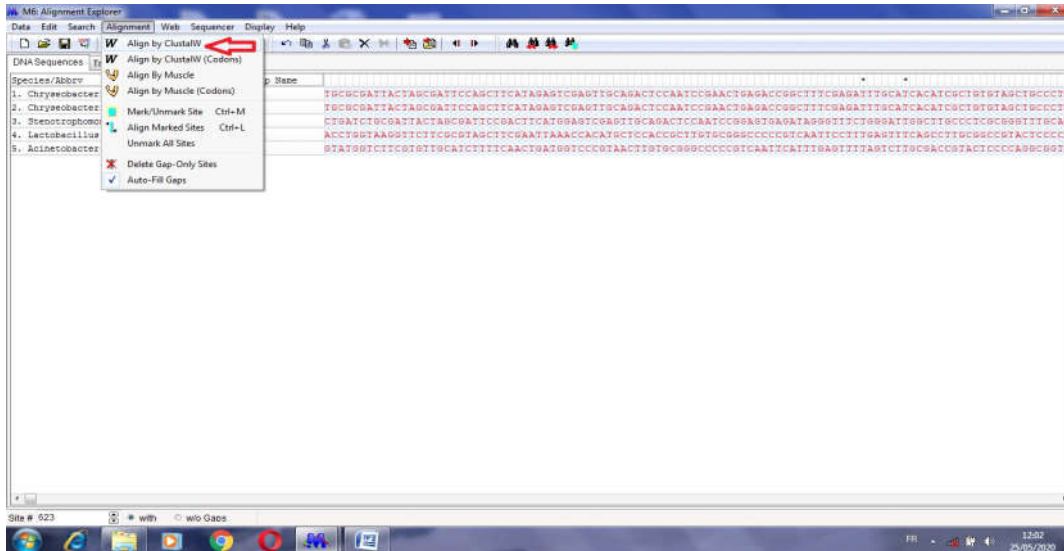
➤ Repeat the same steps to insert other sequences.



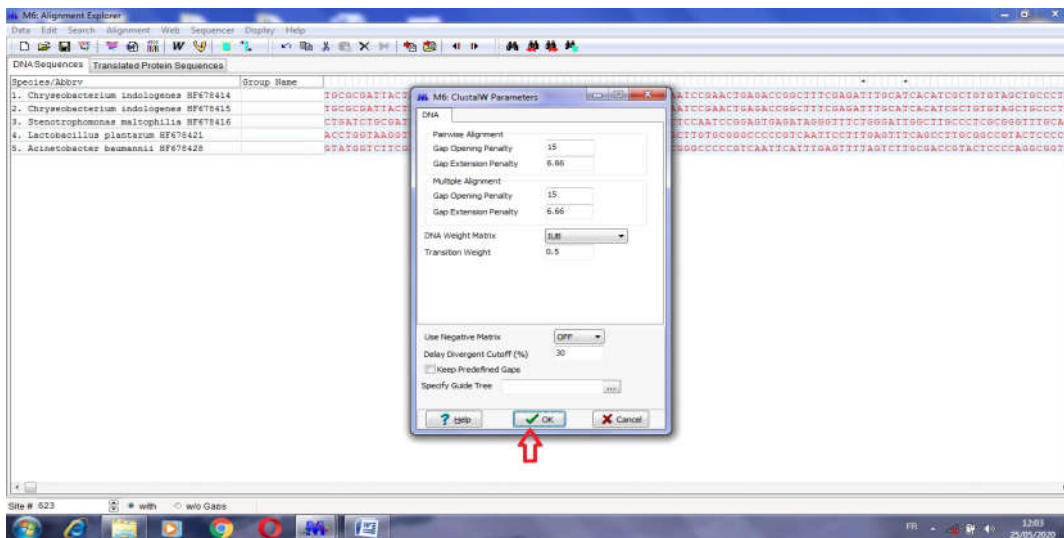
➤ Select all sequences: **Edit_Select All**.



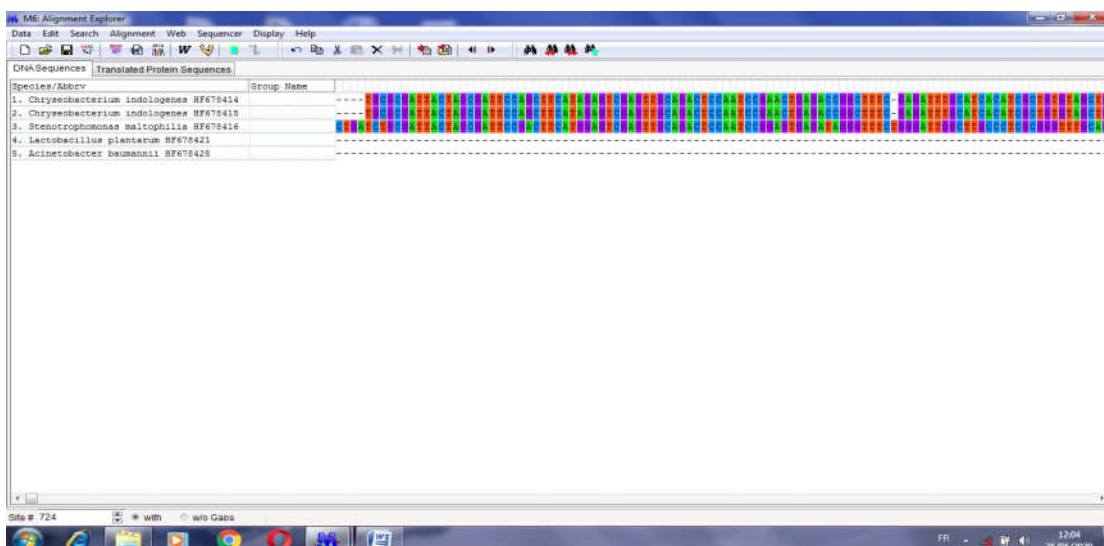
➤ Activate the alignment: **Alignment_Align by ClustalW**.



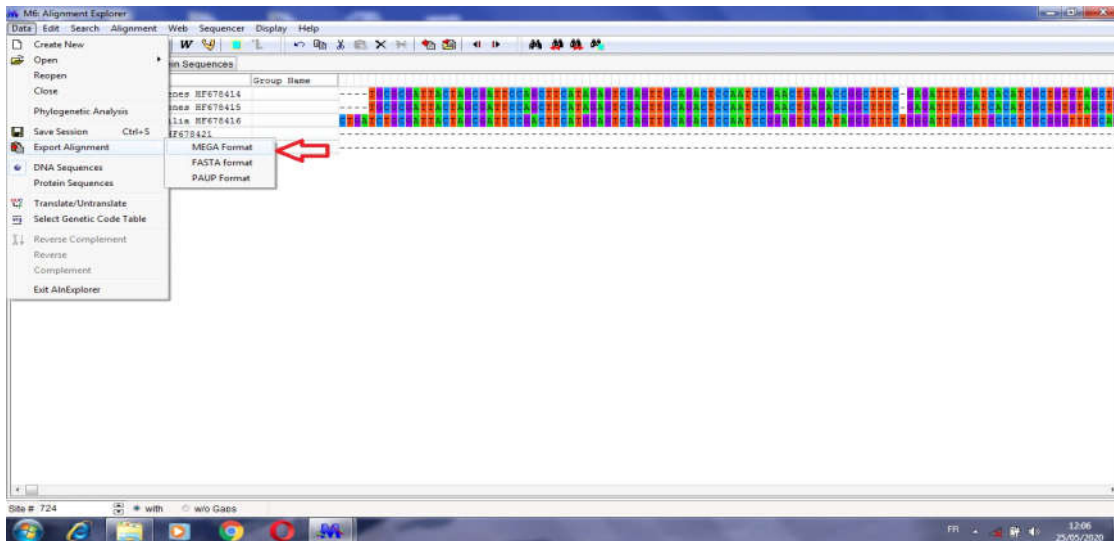
➤ Confirm the alignment activation: **OK**.



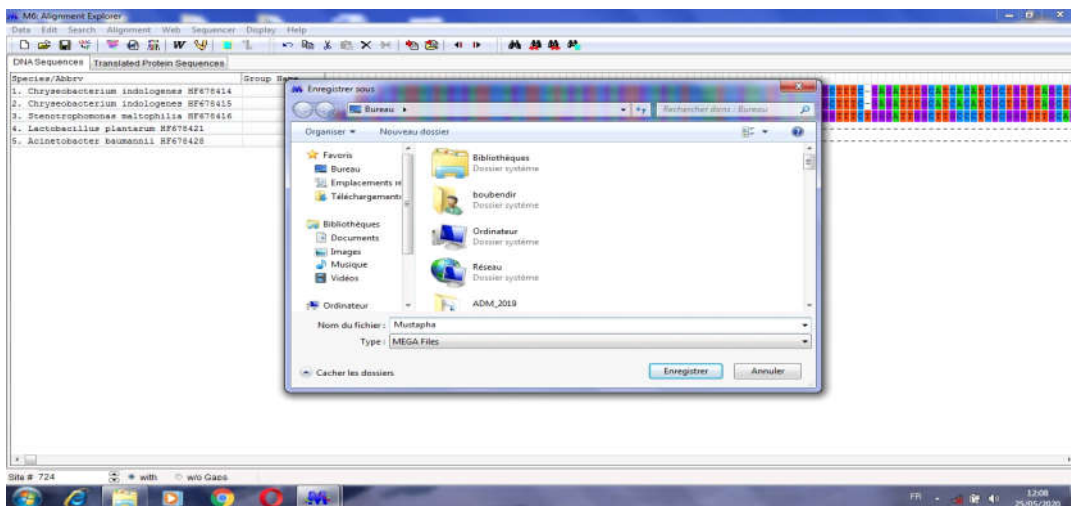
➤ The appearance of GAPS confirms the alignment is completed.



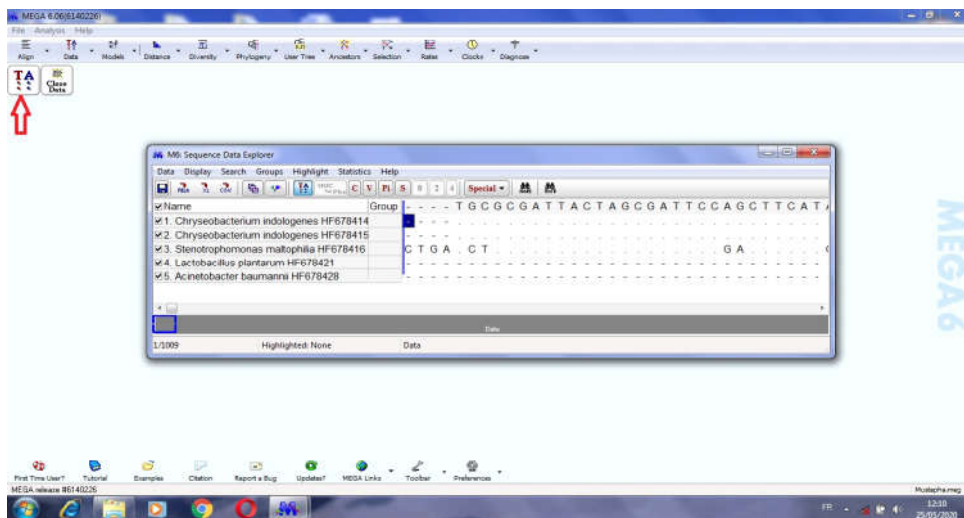
- Save the alignment in MEGA format: **Data_Export Alignment_MEGA Format.**



- Name the MEGA file and save.

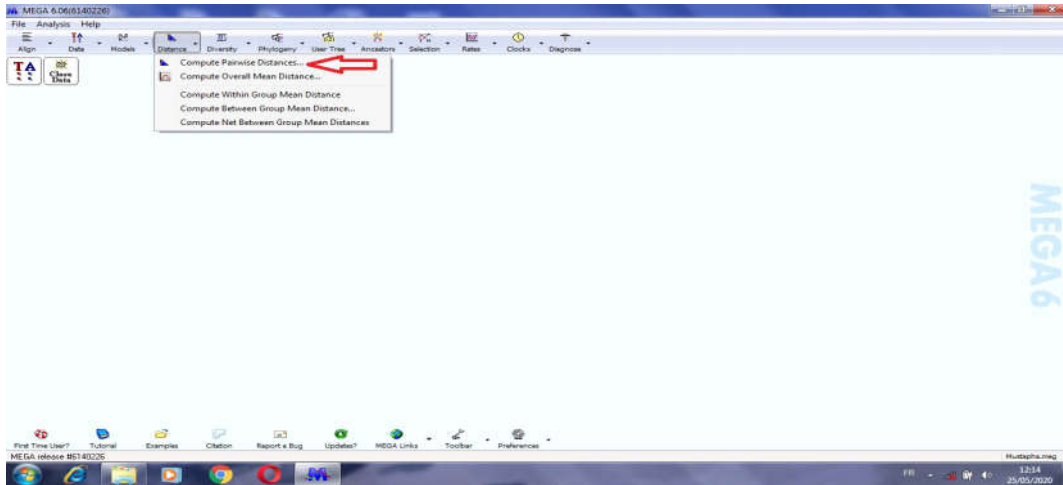


- Finally, open the MEGA file and view your alignment.

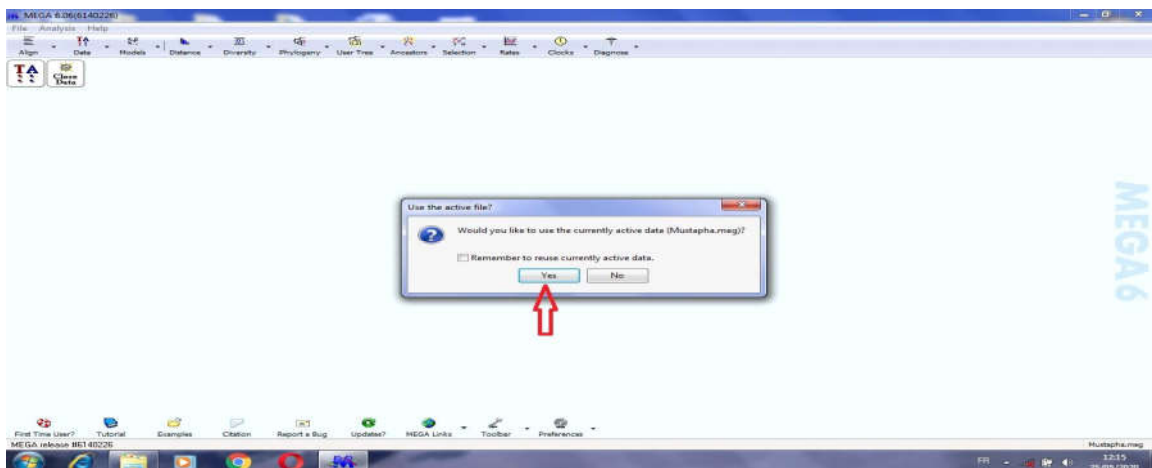


2. Distance matrix: Matrix Distance Explorer

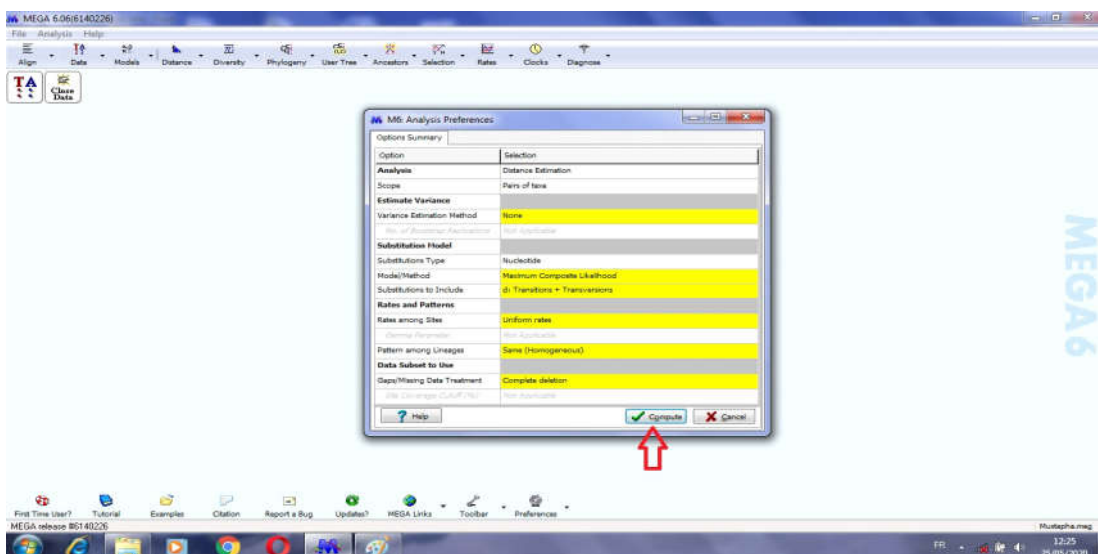
- Open the Matrix Distance Explorer program and select **Compute Pairwise Distance**.



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



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





Finally, the distance matrix is displayed.

Mega Pairwise Distances (C:\Users\bouabendir\Desktop\Mustapha.meg)

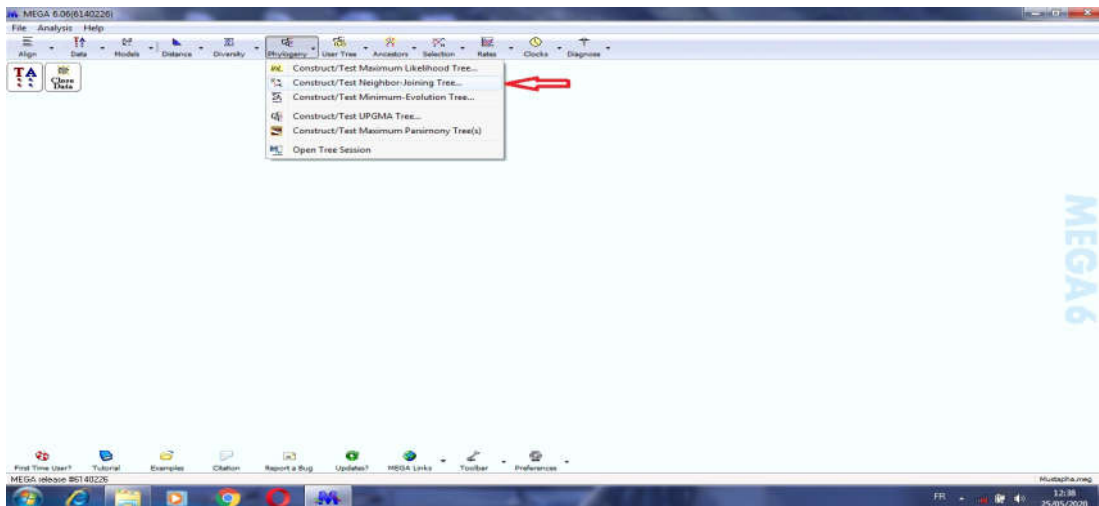
	1	2	3	4	5
1. <i>Chryseobacterium indologenes</i> HF678414					
2. <i>Chryseobacterium indologenes</i> HF678415	0.003				
3. <i>Stenotrophomonas maltophilia</i> HF678416	0.328	0.323			
4. <i>Lactobacillus plantarum</i> HF678421	0.418	0.413	0.309		
5. <i>Acinetobacter baumannii</i> HF678428	0.633	0.627	0.347	0.395	

[1,5] (*Chryseobacterium indologenes* HF678414--*Acinetobacter baumannii* HF678428) / Nucleotide: Maximum Composite Likelihood

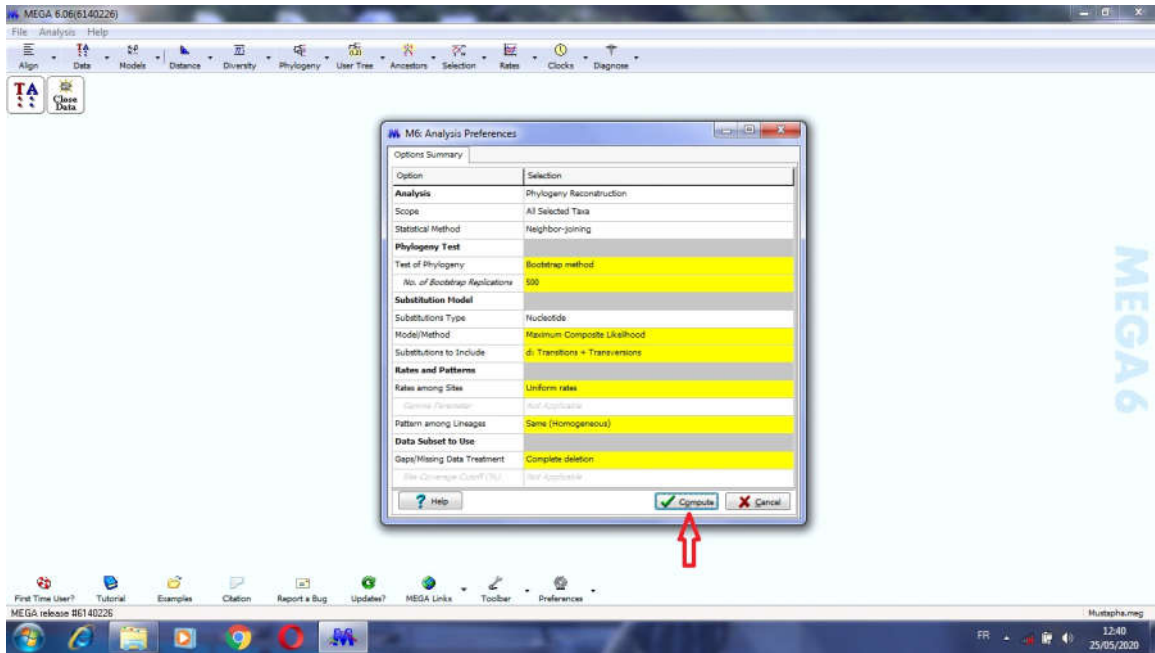
3. Phylogenetic tree topology



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.

