

## “Sequence Alignment and Phylogenetics”



Name- P.D.Hansika Kawindi Paranagama

Email- [hansikawindi429@gmail.com](mailto:hansikawindi429@gmail.com)

LinkedIn- Hansika Paranagama

( [https://www.linkedin.com/in/hansika-paranagama-076678332/overlay/about-this-profile/?lipi=urn%3Ali%3Apage%3Ad\\_flagship3\\_profile\\_view\\_base%3BsmFGUb4HSICKjZeiof4Ncw%3D%3D](https://www.linkedin.com/in/hansika-paranagama-076678332/overlay/about-this-profile/?lipi=urn%3Ali%3Apage%3Ad_flagship3_profile_view_base%3BsmFGUb4HSICKjZeiof4Ncw%3D%3D) )

## TASK 01

### 01. Retrieved Human Insulin gene sequence from NCBI Gene Database.

The screenshot shows the NCBI Gene search interface. In the search bar, the query "Human Ins" is entered. Below the search bar, a dropdown menu lists several human insulin-related genes and proteins. A large image of a fluorescence micrograph showing red and green signals on chromosomes is displayed on the left. On the right, there is a sidebar with links for "Nature, Reference Sequences" and "locus-specific resources". The bottom of the screen shows a Windows taskbar with various icons and the date/time (12:09 PM 6/9/2025).

Gene dropdown: Human Ins

- Homo sapiens INS
- Homo sapiens insulin
- Homo sapiens insulin A chain
- Homo sapiens insulin B chain
- Homo sapiens INSR
- Homo sapiens Insulin receptor subunit alpha
- Homo sapiens Insulin receptor subunit beta
- Homo sapiens insulin receptor
- Homo sapiens insulin-like growth factor I
- Homo sapiens insulin like growth factor 1
- Homo sapiens Insulin-like growth factor 1 receptor alpha chain
- Homo sapiens Insulin-like growth factor 1 receptor beta chain
- Homo sapiens insulin-like growth factor 1 receptor
- Homo sapiens insulin like growth factor 1 receptor
- Homo sapiens insulin receptor substrate 1
- Homo sapiens Insulin-like growth factor II Ala-25 Del
- Homo sapiens insulin-like growth factor II

The screenshot shows the detailed view of the Homo sapiens insulin gene page. The search bar at the top contains "Homo sapiens insulin". The main content area displays the gene information, including the gene name "INS – insulin" (circled in red), its organism ("Homo sapiens (human)"), and other details like "Also known as: IDDM, IDDM1, IDDM2, ILPR, IRDN, MODY10, PNDM4". Below this, there are tabs for "RefSeq products", "Orthologs", and "Genome Data Viewer". A "New - Visualize gene across multiple species" link is visible. The bottom of the page includes a "RefSeq Sequences" section and a "Search details" box with the query: ("Homo sapiens"[Organism] OR Homo sapiens[All Fields]) AND insulin[All Fields] AND alive[prop]. The page also features a sidebar with taxon filters and related data options.

Gene dropdown: Gene

Search bar: Homo sapiens insulin

Results by taxon:

- Top Organisms [Tree]
  - Homo sapiens (2221)
  - Arthroderma uncinatum (4)
  - Aspergillus thermomutatus (2)
  - Aspergillus niger (2)
  - Cladophialophora carrionii CBS 160.54 (2)
  - All other taxa (32)
- More...

Find related data:

Database: Select

Search details:

(("Homo sapiens"[Organism] OR Homo sapiens[All Fields]) AND insulin[All Fields] AND alive[prop])

INS insulin [Homo sapiens (human)]

Gene ID: 3630, updated on 24-May-2025

[Download Datasets](#)

**Summary**

**Official Symbol** INS provided by HGNC

**Official Full Name** insulin provided by HGNC

**Primary source** HGNC:HGNC:6081

**See related** Ensembl:ENSG00000254647 MIM:176730; AllianceGenome:HGNC:6081

**Gene type** protein coding

**RefSeq status** REVIEWED

**Organism** Homo sapiens

**Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo

**Also known as** IDDM; ILPR; IRDN; IDDM1; IDDM2; PNMD4; MODY10

**Summary**

This gene encodes insulin, a peptide hormone that plays a vital role in the regulation of carbohydrate and lipid metabolism. After removal of the precursor signal peptide, proinsulin is post-translationally cleaved into three peptides: the B chain and A chain peptides, which are covalently linked via two disulfide bonds to form insulin, and C-peptide. Binding of insulin to the insulin receptor (INSR) stimulates glucose uptake. A multitude of mutant alleles with phenotypic effects have been identified, including insulin-dependent diabetes mellitus, permanent neonatal diabetes mellitus, maturity-onset diabetes of the young type 10 and hyperproinsulinemia. There is a read-through gene, INS-IGF2, which overlaps with this gene at the 5' region and with the IGF2 gene at the 3' region. [provided by RefSeq, May 2020]

Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Expression
- Bibliography
- Phenotypes
- Variation
- HIV-1 interactions
- Pathways from PubChem
- Interactions
- General gene information
- Markers, Readthrough INS-IGF2, Homology, Gene Ontology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences

- Now scrolled down to get FASTA file.

INS insulin [Homo sapiens (human)]

ncbi.nlm.nih.gov/gene/3630

Go to reference sequence details

Genomic Sequence: NC\_000011.10 Chromosome 11 Reference GRCh38.p14 Primary Assembly

Go to nucleotide: Graphics **FASTA** GenBank

Genes, MANE Project (release v1.4)

NCBI RefSeq Annotation GCF\_000001405.40-RS\_2024\_08

Biological regions, aggregate, NCBI RefSeq Annotation GCF...

[https://www.ncbi.nlm.nih.gov/nucleotide/NC\\_000011.10?report=fasta&from=2159779&to=2161209&strand=true](https://www.ncbi.nlm.nih.gov/nucleotide/NC_000011.10?report=fasta&from=2159779&to=2161209&strand=true)

Table of contents

- Diseases, RNA Targets, Tested
- BioProjects
- BioSystems
- Books
- CCDS
- ClinVar
- Conserved Domains
- dbVar
- Full text in PMC
- Full text in PMC\_nucleotide
- Functional Class
- GAP
- Gene neighbors
- Genes with a similar H3K4me3 profile
- Genome
- GEO Profiles
- GTR
- MedGen
- Nucleotide
- OMIM
- Protein

Homo sapiens chromosome 11, GRCh38.p14 Primary Assembly

NCBI Reference Sequence: NC\_000011.10

[GenBank](#) [Graphics](#)

>NC\_000011.10?report=fasta&from=2159779&to=2161209&strand=true

NUCotide NUCotide Advanced Search Help

Send to: ▾

Change region shown

Whole sequence  
 Selected region  
from: 2159779 to: 2161209

Customize view

Display options

Show reverse complement  
 Show gap features

Analyze this sequence

Run BLAST  
Pick Primers

Related information

Assembly  
BioProject

31°C Partly sunny 12:10 PM 6/9/2025

- Now we can download this by “clicking send” to option. Tick the complete record, File and choose FASTA format → create file

ncbi.nlm.nih.gov/nuccore/NC\_000011.10?report=fasta&from=2159779&to=2161209&strand=true

NUCLEOTIDE INNUCLEOTIDE Advanced Search Help

Fasta ▾

## Homo sapiens chromosome 11, GRCh38.p14 Primary Assembly

NCBI Reference Sequence: NC\_000011.10

GenBank Graphics

>NC\_000011.10:c2161209-2159779 Homo sapiens chromosome 11, GRCh38.p14 Primary Assembly

AGCCCTCAGGCAAGCTGCATCAGAAAGGGCATCGAGCTTCCAAAGGGCTTGGCTCAGGT  
GGGGTCAGGATTCCAGGGTGGCTGGACCCCCAGGCCCAAGCTGCAGCAGGGAGGACGTGGCTGGCTCG  
TGAAGCATGTGGGGTGAGCCAGGGGGCCAAGGCAGGGCACCTGGCTCAGGGCTCAGGGCTGC  
CTGCTGGGGCAGACTCTACTCTCTCTCTGGCCATGGCTGTGATCAGCTGGCTCAGGGCTGCTG  
GCCCTCTGGGACCTGACCCAGGCCAGCAGCTTGTGAAACAAACACTGTGGCTCACACCTGGTAGAG  
CTCTCTACTAGTGTGCGGGGAAACGAGGCTCTTCTACACACCCAAAGCCCCGGGAGGCAAGGGACCT  
GCAGGGTAGGCCAACTGGCCATTGCTGCCCTGGCCGGGGCCAGGCCACCCCTGCTCTGGCTCCAC  
CAGCAGTGGCAGAGGGGGCAGGGCTGGCCACCCAGCAGGGGTCAAGTTCACTTTAAAAAGAG  
TTCTCTTGTCAAGCTCTAAAGTACCAAGCTCTCTGTGGCCAGTCAAGATCTCAGGCTGAGGGCTG  
TTGGCTTGGCCAGGGGAGATCATCAGGGTGGGACGCTCTCCCTCAACTGGCCCTAACACAAA  
TGGCCCGCAGGCACTTCTCCACCTTGTGATGGCAGATTCAAGTGTGGTAAGTAAAGCT  
GGGGTACCTGGGGTACAGGGTGGCCCAAGCTGCCTGCTCCGGCAGACACCCCAATCAGCCGGGAGA  
GGGGCTGGCTGGCTGGCTGAGTGGGGCAGACCCCTGGCCAGGGCTCAGGGCAGCTCCATAGTCAGGG  
ATGGGGAAAGATGCTGGGGCAGGGCTGGGGAGAAGTACTGGATCACCTGTTAGGGCTCCACCTGTGAC  
GCTGGCCCTGGGGGGGGGGAGGGAGGTGGGACATGTGGGCTGTGGGGCTGTAGGGTCCACCCAGGTGG  
GTGACCTCCCTAACCTGGCTGGGAGATGGGGCTGGGAGATGGGGCTGGGAGTGGGACCTAGGGCTGGGG  
AGGGGGGGCACTGTGCTCTCTGACTGTGCTCTGGCTCCCTGGCTGGCCAGCTGGGGAAACCTGC  
TCTGGGGGGCAGCTGGCAAGTGGGCACTGGGAGCTGGGGCCCTGGTGCAGGCAAGCAGCTGGCAG  
CCTGG  
CTACCAAGCTGGGAAGACTGCAACTAGCAGCACTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG  
GAGAGAGATGGAAATAAGCCCTTGAACCAGC

Send to: ▾

Complete Record

Coding Sequences

Gene Features

Choose Destination

File

Clipboard

Collections

Analysis Tool

Download 1 item.

Format

FASTA

Show GI

Create File

Analyze this sequence

Run BLAST

Pick Primers

Related information

Assembly

BioProject

## 02. Identified sequence similarities using BLAST.

- Website of NCBI BLAST → nucleotide BLAST

The screenshot shows the NCBI BLAST homepage. At the top, there's a banner about the transition to ClusteredNR. Below it, the "Basic Local Alignment Search Tool" section is visible. It features three main options: "Nucleotide BLAST" (circled in red), "blastx" (translated nucleotide → protein), and "tblastn" (protein → translated nucleotide). To the right, there's a news box about improvements and a link to more news. The bottom of the screen shows a Windows taskbar with various icons and the date/time (12:14 PM, 6/9/2025).

- Downloaded human insulin gene sequence copied and pasted.

The screenshot shows the "Nucleotide BLAST: Search nucleotide" page. The "blastn" tab is selected. In the "Enter Query Sequence" field, the human insulin gene sequence is pasted:  
CTACCAAGCTGGAGAAGTACTGCAACTAGACGCAGCCCGCAGGCCAGCCC  
ACACCCGGCCCTCTGCACC  
GAGAGAGATGGAATAAAGCCCTTGAAACCAGC

Below the sequence, there are fields for "Job Title" and "Align two or more sequences". Under "Choose Search Set", "Database" is set to "Standard databases (nr etc.)", and "Organism" is set to "Core nucleotide database (core\_nt)". A yellow box highlights the "Feedback" button on the right side of the page.

- Choose database – nr/nt - Nucleotide collection.

Core nucleotide database (core\_nt)

RefSeq Select RNA sequences (refseq\_select)  
 Reference RNA sequences (refseq\_rna)  
 RefSeq Reference genomes (refseq\_reference\_genomes)  
 RefSeq Genome Database (refseq\_genomes)

**Nucleotide collection (nr/nt)**

CTACAGCTGGAGAAC  
 ACACCCGCCCTCC  
 GAGAGATGGAATAA

Or, upload file  
 Job Title  
 Align two or more sequences  
**Choose Search Options**

**Database**

Genomic + transcript databases   
 Betacoronavirus   
 Experimental databases

**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 [?](#)

**Exclude** **Optional**

Models (XM/XP)  Uncultured/environmental sample sequences

**Limit to** **Optional**

Sequences from type material

[YouTube](#) [Create custom database](#)

Enter an Entrez query to limit search [?](#)

**Program Selection**

**Optimize for**

Highly similar sequences (megablast)  
 More dissimilar sequences (discontiguous megablast)  
 Somewhat similar sequences (blastn)

Feedback

31°C Partly sunny 12:17 PM 6/9/2025

blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\_TYPE=BlastSearch&LINK\_LOC=blasthome

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 [?](#)

**Exclude** **Optional**

Models (XM/XP)  Uncultured/environmental sample sequences

**Limit to** **Optional**

Sequences from type material

[YouTube](#) [Create custom database](#)

Enter an Entrez query to limit search [?](#)

**Program Selection**

**Optimize for**

Highly similar sequences (megablast)  
 More dissimilar sequences (discontiguous megablast)  
 Somewhat similar sequences (blastn)

Choose a BLAST algorithm [?](#)

**BLAST**  Search database nt using Megablast (Optimize for highly similar sequences)  
 Show results in a new window

Note: Parameter values that differ from the default are highlighted in yellow and marked with ♦ sign

**+ Algorithm parameters**

Feedback

31°C Partly sunny 12:17 PM 6/9/2025

- This is the result that I get.

blast.ncbi.nlm.nih.gov/Blast.cgi NCBI Blast:Nucleotide Sequence

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

BLAST® » blastn suite » results for RID-4D0WEHRE013 Home Recent Results Saved Strategies Help

**Important update** Effective August 2025, the **ClusteredNR** database will become the default Protein BLAST database. [Learn more about ClusteredNR](#)

Save Search Search Summary ? How to read this report? BLAST Help Videos Back to Traditional Results Page

Job Title	Nucleotide Sequence
RID	4D0WEHRE013 Search expires on 06-10 14:47 pm <a href="#">Download All</a>
Program	BLASTN <a href="#">Citation</a>
Database	nt <a href="#">See details</a>
Query ID	IcllQuery_2945319
Description	None
Molecule type	dna
Query Length	1431
reports	<a href="#">Distance tree of results</a> <a href="#">MSA viewer</a> ?

Filter Results

Organism only top 20 will appear  exclude  
Type common name, binomial, taxid or group name [+ Add organism](#)

Percent Identity E value Query Coverage

Feedback

12:19 PM 6/9/2025

blast.ncbi.nlm.nih.gov/Blast.cgi NCBI Blast:Nucleotide Sequence

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments Download Select columns Show 100 ?

select all 100 sequences selected GenBank Graphics Distance tree of results MSA Viewer

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens INS-IGF2 readthrough (INS-IGF2), RefSeqGene on chromosome 11</a>	<a href="#">Homo sapiens</a>	2643	2643	100%	0.0	100.00%	39098	<a href="#">NG_050578.1</a>
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens insulin isoform U2 (INS) mRNA, complete cds, alternatively spliced</a>	<a href="#">Homo sapiens</a>	2643	2643	100%	0.0	100.00%	1707	<a href="#">MT335688.1</a>
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens insulin (INS), RefSeqGene on chromosome 11</a>	<a href="#">Homo sapiens</a>	2643	2643	100%	0.0	100.00%	8416	<a href="#">NG_007114.1</a>
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens chromosome 11, clone RP11-889I17, complete sequence</a>	<a href="#">Homo sapiens</a>	2643	2643	100%	0.0	100.00%	170027	<a href="#">AC132217.15</a>
<input checked="" type="checkbox"/>	<a href="#">Human alpha-type insulin gene and 5' flanking polymorphic region</a>	<a href="#">Homo sapiens</a>	2621	2621	100%	0.0	99.72%	3943	<a href="#">M10039.1</a>
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens insulin (INS) gene, complete cds</a>	<a href="#">Homo sapiens</a>	2619	2619	100%	0.0	99.72%	4969	<a href="#">AH002844.2</a>
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens tyrosine hydroxylase (TH) gene, 3' end: insulin (INS) gene, complete cds; insulin-like g...</a>	<a href="#">Homo sapiens</a>	2619	2619	100%	0.0	99.72%	12565	<a href="#">L15440.1</a>
<input checked="" type="checkbox"/>	<a href="#">Human gene for proinsulin, from chromosome 11. Includes a highly polymorphic region upstream f...</a>	<a href="#">Homo sapiens</a>	2619	2619	100%	0.0	99.72%	4992	<a href="#">V00565.1</a>
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens haplotype ICA tyrosine hydroxylase (TH) gene, partial sequence; insulin (INS) gene, co...</a>	<a href="#">Homo sapiens</a>	2615	2615	100%	0.0	99.65%	7496	<a href="#">AH012037.2</a>
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens chromosome 11, clone RP4-539G11, complete sequence</a>	<a href="#">Homo sapiens</a>	2615	2615	100%	0.0	99.65%	171366	<a href="#">AC130303.8</a>
<input checked="" type="checkbox"/>	<a href="#">Homo sapiens ins gene, partial</a>	<a href="#">Homo sapiens</a>	2545	2545	97%	0.0	99.64%	1393	<a href="#">AJ009655.1</a>
<input checked="" type="checkbox"/>	<a href="#">Pan troglodytes tyrosine hydroxylase (TH) gene, partial cds; and insulin precursor (INS) gene, comple...</a>	<a href="#">Pan troglodytes</a>	2342	2342	95%	0.0	97.73%	7355	<a href="#">AH011814.2</a>

Feedback

12:20 PM 6/9/2025

## TASK 02

01. Retrieved Human Insulin gene and Chimpanzee insulin gene sequences from NCBI Gene Database.
02. Go to EMBOSS Needle website

The screenshot shows the homepage of the EMBOSS Needle website. The title 'EMBOSS Needle' and subtitle 'Pairwise Sequence Alignment (PSA)' are displayed in a white box on a green background. Below this, there is a banner with the text 'Welcome to the Job Dispatcher website! If you need assistance or have feedback, please [contact us](#)'. A blue button labeled 'I agree, dismiss this banner' is visible. The main content area is partially visible, showing a message about cookies and privacy. The browser's address bar shows 'ebi.ac.uk/jdispatcher/psa/emboss\_needle?stype=dna&matrix=EDNAFULL'. The taskbar at the bottom includes icons for Rain showers Tomorrow, Search, and various applications like eWings Education, Kymriah®, and mechanism of action...

03. That downloaded sequences copy and passed into EMBOSS Needle boxes.

- First box – Human Insulin gene sequence
- Second box – Chimpanzee Insulin gene sequence

04. Choose sequence type – DNA

The screenshot shows the sequence input page of the EMBOSS Needle website. It features two main input fields: 'Input sequence' and 'Sequence type'. Under 'Sequence type', the 'DNA' option is selected. Below these, there are two text areas for pasting sequences. The top text area is labeled 'Paste your first sequence here - or use the example sequence' and contains the Human Insulin gene sequence: AGCCCTCCAGGACAGGCTGCATCAGAAGAGGCCATCAAGCAGGTCTTCCAAGGGCCTTCGCGTC AGGT GGGCTCAGGATTCCAGGGTGGCTGGACCCCAGGCCAGCTCTGCAGCAGGGAGGACGTGGCTGG GCTCG TGAAGCATGTGGGGTGAGCCCAGGGCCCCAAGGCAGGGCACCTGGCCTTCAGCCTGCCTCAGCC CTGC CTGTCTCCAGATCACTGTCCTTGCCATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGC To... The bottom text area is labeled 'Paste your second sequence here - or use the example sequence' and contains the Chimpanzee Insulin gene sequence: TCCTGGGGACAGGGGTCTGGGACAGCGGTGCAAAGAGGCCCGCCCTGCAGCCTCCAGCTGTCTGT GTCT AAATCTGCAAAACTGCCCGACGCTGACGCCCTTTGCTCTCTGCAACACATTGCCCCCAGCTGTGACCCAC To... The browser's address bar shows 'ebi.ac.uk/jdispatcher/psa/emboss\_needle?stype=dna&matrix=EDNAFULL'. The taskbar at the bottom includes icons for Rain showers Tomorrow, Search, and various applications like eWings Education, Kymriah®, and mechanism of action...

## 05. Insert title as “Human – Chimpanzee Insulin gene alignment” & submit.

The screenshot shows the EMBoss Needle web application on a Windows desktop. The browser address bar displays the URL [ebi.ac.uk/jdispatcher/psa/emboss\\_needle?stype=dna&matrix=EDNAFULL](https://ebi.ac.uk/jdispatcher/psa/emboss_needle?stype=dna&matrix=EDNAFULL). The main form has a "Choose File" button with "No file chosen" and a dropdown menu set to "pair". Below the form are buttons for "Use the example", "Clear sequence", and "More example inputs". On the left, there's a "Parameters" section with an "OUTPUT FORMAT" dropdown set to "pair" and a "More options" link. On the right, a "Submit" button is visible. The "Submit" section contains a "Title" input field with the text "Human Chimpanzee insulin gene alignment". At the bottom, a message reads: "If you use this service, please consider citing the following publication: [The EMBL-EBI Job](#)". The taskbar at the bottom shows various pinned icons and the system clock indicating 8:28 PM on 6/9/2025.

## 06. View result.

The screenshot shows the same EMBoss Needle application after submission. A green modal dialog box appears with the text "YOUR JOB IS FINISHED" and a "Close" button. Below the dialog, a message states: "Please note that results can only be retrieved for jobs submitted within the last seven days." It also includes the "Job ID: [emboss\\_needle-I20250609-155828-0224-2901608-p1m](#)". At the bottom of the dialog are three buttons: "View Results", "Your Jobs", and "Close". The taskbar at the bottom shows pinned icons and the system clock indicating 8:28 PM on 6/9/2025.

This is the result that I get.

The screenshot shows the EMBoss Needle results page for a pairwise sequence alignment (PSA) job. The top navigation bar includes links for Job Dispatcher, Help & Privacy, Your Jobs, Input form, Feedback, and a contact link. Below the header, a welcome message encourages users to contact support if needed. The main content area displays the results for Job ID emboss\_needle-I20250609-155828-0224-2901608-p1m. It features tabs for Tool Output (selected), Result Files, and Submission Details. The Tool Output tab shows the command-line parameters used for the alignment, including the program (needle), date (Mon 9 Jun 2025 15:58:34), and various options like -auto, -stdout, and sequence files. The command is approximately:

```
#####
# Program: needle
# Rundate: Mon 9 Jun 2025 15:58:34
# Commandline: needle
# -auto
# -stdout
# -asequence emboss_needle-I20250609-155828-0224-2901608-p1m.asequence
# -bsequence emboss_needle-I20250609-155828-0224-2901608-p1m.bsequence
# -datafile EDNAFULL
# -gapopen 10.0
# -gapextend 0.5
# -endopen 10.0
# -endextend 0.5
# -aformat3 pair
# -snucleotide1
# -snucleotide2
# Align_format: pair
# Report_file: stdout
#####
=====
#
# Aligned_sequences: 2
# 1: FMRQSS 001
```

Pan troglodytes isolate AG18354 c × EMBOS Needle < EMBL-EBI × +

ebi.ac.uk/jdispatcher/psa/emboss\_needle/summary?jobId=emboss\_needle-l20250609-155828-0224-2901608-p1m

Gmail YouTube Maps Echem.Ik | ຜົກຄະ... eWings Education Kymriah®, a revolution... mechanism of action... Study Details | Study... All Bookmarks

```
#  
# Aligned_sequences: 2  
# 1: EMBOS_001  
# 2: EMBOS_001  
# Matrix: EDNAFULL  
# Gap_penalty: 10.0  
# Extend_penalty: 0.5  
#  
# Length: 1822  
# Identity: 1354/1822 (74.3%)  
# Similarity: 1354/1822 (74.3%)  
# Gaps: 440/1822 (24.1%)  
# Score: 6612.0  
#  
#=====
```

EMBOSS_001	1	- - - - -	0
EMBOSS_001	1	TCCTGGGGACAGGGTCTGGGGACAGCGGTGCAAAGAGCCCCGCCCTGCA	50
EMBOSS_001	1	- - - - -	0
EMBOSS_001	51	GCCTCCAGCTGTCCTGGTCTAATGTGAAAGTGGCCAGGTGACGGCTT	100

IND - OKC Game score Search

8:30 PM 6/9/2025

Pan troglodytes isolate AG18354 c × EMBOS Needle < EMBL-EBI × +

ebi.ac.uk/jdispatcher/psa/emboss\_needle/summary?jobId=emboss\_needle-l20250609-155828-0224-2901608-p1m

Gmail YouTube Maps Echem.Ik | ຜົກຄະ... eWings Education Kymriah®, a revolution... mechanism of action... Study Details | Study... All Bookmarks

EMBOSS_001	1	- - - - -	0
EMBOSS_001	151	ACCGGGCCCCCTGGTTAACACTCTAACCGCTGGCCCTAACAGAGAGG	200
EMBOSS_001	1	- - - - -	0
EMBOSS_001	201	TGCTGACGACCAAGGAGATCTCCCACAGACCCAGCACAGGGAAATGGT	250
EMBOSS_001	1	- - - - -	0
EMBOSS_001	251	CCGGAAATTGCAGCCTCAGCCCCAGCCATCTGCCGACCCCCCACCCA	300
EMBOSS_001	1	- - - - -	0
EMBOSS_001	301	GGCCCTAACGGCCAGGCCAGGGTTGACAGGTAGGGAGATGGCTC	350
EMBOSS_001	1	- - - - - AGCCCTCCAGGACA	14
EMBOSS_001	351	TGAGACTATAAGCCAGGGGCCAGCAGCCCTCAGCCCTCCAGGACA	400
EMBOSS_001	15	GGCTGCATCAGAAGAGGCCATCAAGCAGGTCTGTTCAAGGGCTTGCG	64
EMBOSS_001	401	GGCTGCATCAGAAGAGGCCATCAAGCAGGTCTGTTCAAGGGCTTGCG	450

## TASK 03

01. Go to Clustal Omega web interface.

The screenshot shows the Clustal Omega website on a browser. The title bar has three tabs: "Mus musculus strain C57BL/6J chr", "Clustal Omega < EMBL-EBI", and "Pan troglodytes isolate AG18354". The address bar shows the URL "ebi.ac.uk/jdispatcher/msa/clustalo?stype=dna". Below the address bar is a navigation bar with links to "Job Dispatcher", "Help & Privacy", "Your Jobs", "Input form" (which is highlighted in blue), and "Feedback". A yellow banner at the top says "Welcome to the Job Dispatcher website! If you need assistance or have feedback, please [contact us](#)." The main content area features a large "Clustal Omega" logo and the text "Multiple Sequence Alignment (MSA)". Below this is a paragraph about the tool's capabilities. The status bar at the bottom shows "Rain showers Tomorrow" and the date "6/9/2025".

02. Paste Human insulin gene sequence and Chimpanzee insulin gene sequence one after the other into input box.

03. Choose type- DNA

The screenshot shows the "Input form" page of the Clustal Omega website. At the top, there are tabs for "Parameters" and "INPUT FORMAT". The main area has two sections: "Input sequence" and "Sequence Type". Under "Sequence Type", the "DNA" option is selected. Below this is a text area titled "Paste your sequence here - or use the example sequence" containing the following DNA sequence:

```
CCTCTCTGTGTCCTCTGCCCTGCCGCTGTCGGAACCTGCTCTGCCTGCCACGCCCTGGCAGTGG  
GCA  
GGTGAGCTGGCGGGGCCCTGGTCAGGCAGCCTGCAGCCCTGGCCCTGGAGGGTTCCCTGC  
AGAAC  
CGTGGATCGTGGAAACAATGCTGTACAGCATCTGCTCCCTTACCAAGCTGGAGAAACTTGCAACTA  
GA  
TGGATAAAGCCCTGAACCAGC
```

Below the sequence area are buttons for "Choose File" (with "No file chosen"), "Use the example", "Clear sequence", and "More example inputs". The status bar at the bottom shows "Rain showers Tomorrow" and the date "6/9/2025".

#### 04. Insert title as “Consensus sequence from Human, Chimpanzee Insulin genes.” And submit.

Screenshot of a web browser showing the Clustal Omega sequence analysis tool interface. The URL is ebi.ac.uk/jdispatcher/msa/clustalo?stype=dna. The page displays parameters for a multiple sequence alignment (MSA) job. Under "Parameters", the "OUTPUT FORMAT" dropdown is set to "ClustalW with character counts". Under "Submit", the "Title" field contains "Consensus sequence from Human Chimpanzee Insulin genes". A large blue "Submit" button is present. At the bottom, a note encourages citing "The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024". The system tray at the bottom shows a weather icon for Rain showers Tomorrow, a search bar, and various application icons.

This is the result I get.

Screenshot of a web browser showing the results of a Clustal Omega MSA job. The URL is ebi.ac.uk/jdispatcher/msa/clustalo/summary?jobId=clustalo-I20250609-163302-0384-74961016-p1m. The main header reads "Clustal Omega" and "Multiple Sequence Alignment (MSA)". Below the header, there are links for "Job Dispatcher", "Help & Privacy", "Your Jobs", and "Input form", along with a "Feedback" button. A yellow banner at the top says "Welcome to the Job Dispatcher website! If you need assistance or have feedback, please [contact us](#)". The main content area shows the results for Job ID clustalo-I20250609-163302-0384-74961016-p1m, with options to "Copy" or "Resubmission". Below this, a navigation bar includes "Tool Output" (which is underlined), "Alignments", "Guide Tree", "Phylogenetic Tree", "Results Viewers", "Result Files", and "Submission Details". The system tray at the bottom shows a weather icon for 29°C Partly cloudy, a search bar, and various application icons.

Alignment with colours

Hide

CLUSTAL 0(1.2.4) multiple sequence alignment

	test1	test2	Length
1	-----	TCCTGGGGACAGGGTCTGGGACAGCGGTCAAAGAGCCCCGCCCTGCAGCCTCAGCT	60
2	-----	GTCTGGTCTAATGTGAAAGTGGCCAGGTGACGGCTTGCTCTCTGGAGACATTG	120
3	-----	CCCCAGCTGTAGCAAGGACAGGCTGGCACCGGCCCTGTTAAGACTCTAATGACC	180
4	-----	CGCTGGCCTAAGGAAGAGGTGCTGACGACCAAGGAGATCTTCCCACAGACCCAGCACCA	240
5	-----	GGGAAATGGTCCGGAAATTGCAGCCTCAGCCCCAGCCATCTGCCGACCCCCAACCCA	300

28°C Partly cloudy 9:23 PM 6/9/2025

Alignment with colours

Hide

CLUSTAL 0(1.2.4) multiple sequence alignment

	test1	test2	Length
1	-----	AGCCCTCAGGACAGGCTGCATCAGAAAGAGGCCA	34
2	-----	AAGCCAGCGGGGGGCCAGCAGCCCTCAGGCCCTCAGGACAGGCTGCATCAGAAAGAGCCA	420
3	TCAAGCAGGTCTGTTCC	AAGGGCCTTGCCTCAGGTGGCTCAGGATTCCAGGGTGGCTG	94
4	TCAAGCAGGTCTGTTCC	AAGGGCCTTGCCTCAGGTGGCTCAGGGTTCCAGGGTGGCTG	480
5	GACCCCAGGCCCCAGCTCTGCAGCAGGGAGGACGTGGCTGGCTGTGAA	AGCATGTGGGG	154
6	GACCCCAGGCCCCAGCTCTGCAGCAGGGAGGACGTGGCTGGCTCTGAA	AGCATGTGGGG	540
7	GTGAGCCCAAGGGCCCCAAGGCAGGGCACCTGGCTTCAGCCTGCC	TCAGGCTGCCGT	214
8	GTGAGCCCAAGGGCCCCAAGGCAGGGCACCTGGCTTCAGCCTGCC	TCAGGCTGCCGT	600
9	CTCCCAGATCACTGTCTTCTGCC	ATGGCCCTGTGGATGCGCCTCTGCCCTGCTGGCG	274
10	CTCCCAGATCACTGTCTTCTGCC	ATGGCCCTGTGGATGCGCCTCTGCCCTGCTGGCG	660
11	CTGCTGGCCCTCTGGGGACCTGACCCAGCCGACGCTTGTGAA	ACCAACACCTGTGGCG	334
12	CTGCTGGCCCTCTGGGGACCTGACCCAGCCGACGCTTGTGAA	ACCAACACCTGTGGCG	720

28°C Partly cloudy 9:23 PM 6/9/2025

## TASK 04

01. Scrolled down in ClustalW result that get at TASK 3 and downloaded the Multiple Sequence Alignment (MSA) in FASTA format.

The alignment in FASTA format converted by Seqret

Percent Identity Matrix

Submission Details

Tool outputs as a single compressed zip file

02. Go to Skylign website and uploaded the that downloaded FASTA file of Clustalw result , pressed generate logo.

Skylign is a tool for creating logos representing both sequence alignments and profile hidden Markov models. Submit to the form on the right in order to produce (i) interactive logos for inclusion in webpages, or (ii) static logos for use in documents.

See an example

Upload an HMM or Multiple sequence alignment ?  
Choose File clustalo-I20...1016-p1m.fa

Alignment Processing

Use Observed Counts ?  
 Use Weighted Counts ?  
 Create HMM - keep all columns ?  
 Create HMM - remove mostly-empty columns ?

Fragment Handling

Alignment sequences are full length ?  
 Some sequences are fragments ?

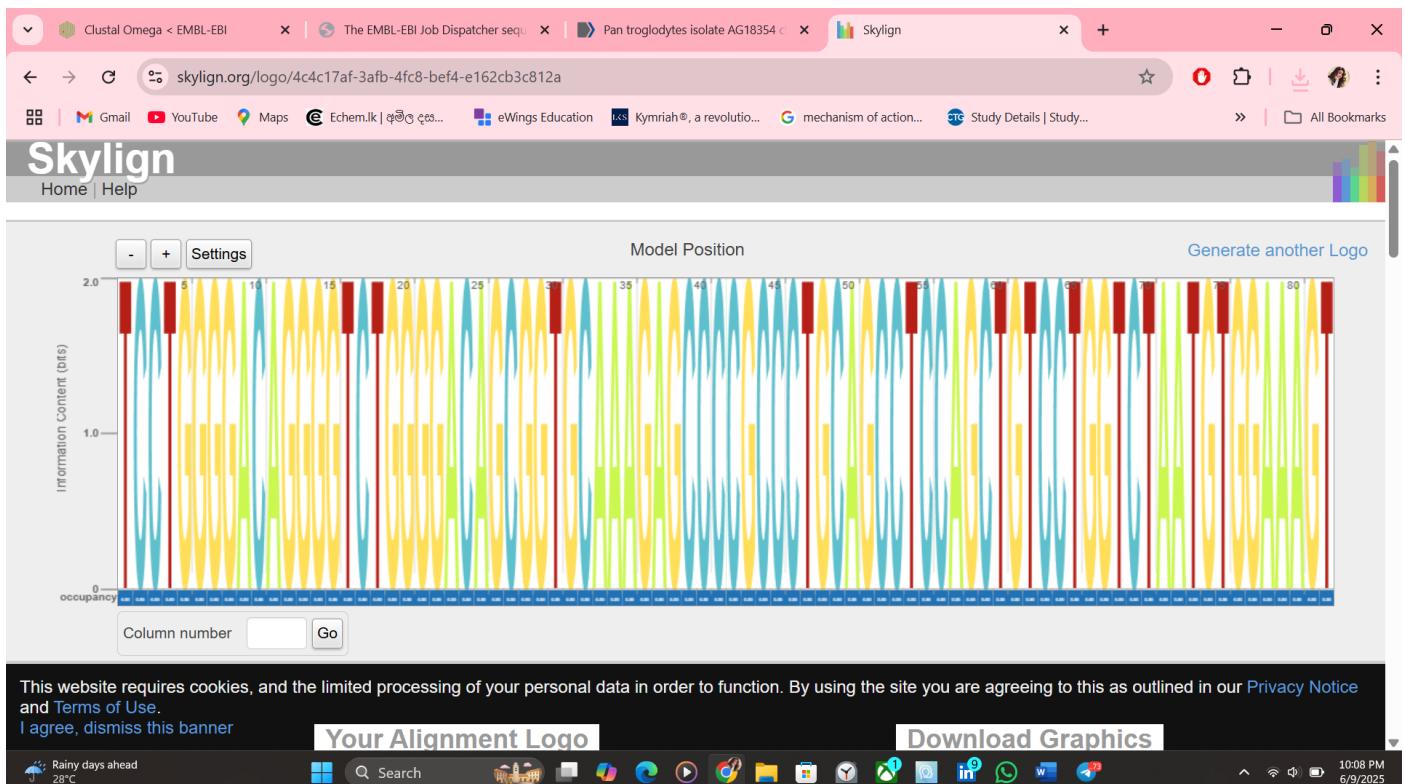
Letter Height

Information Content - All ?  
 Information Content - Above Background ?  
 Score ?

Generate Logo Reset

This website requires cookies, and the limited processing of your personal data in order to function. By using the site you are agreeing to this as outlined in our Privacy Notice and Terms of Use.  
I agree, dismiss this banner

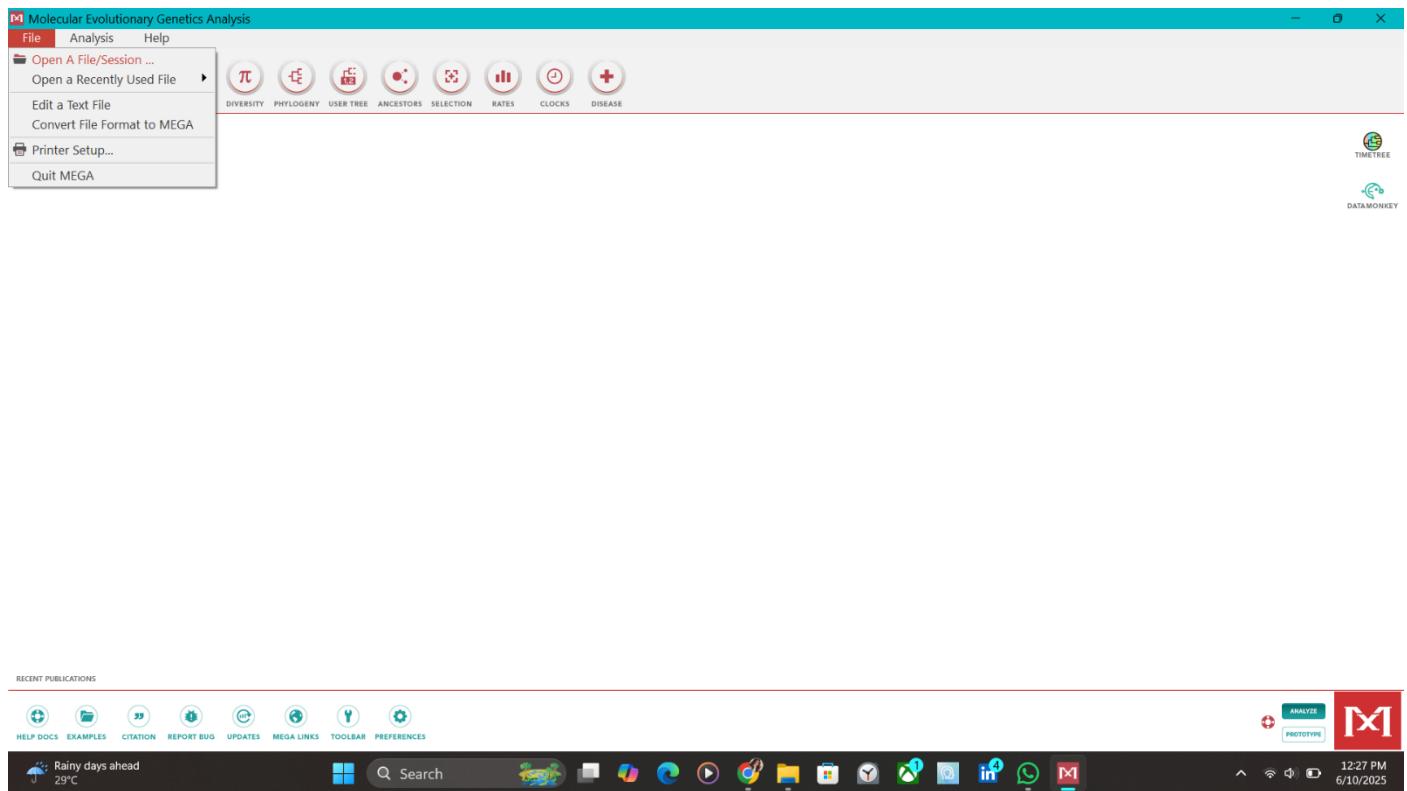
This is the result.



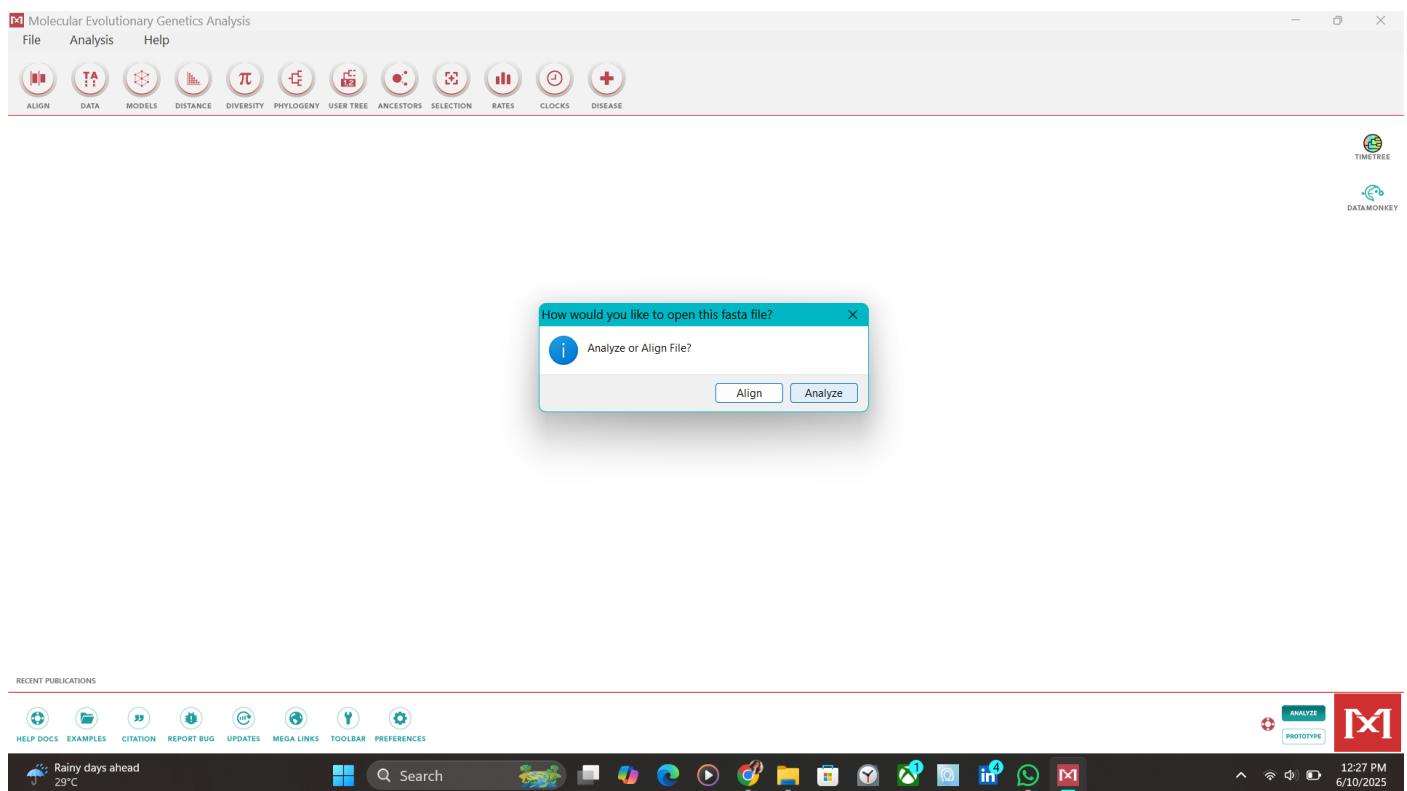
## TASK 05

01. Opened MegaX app.

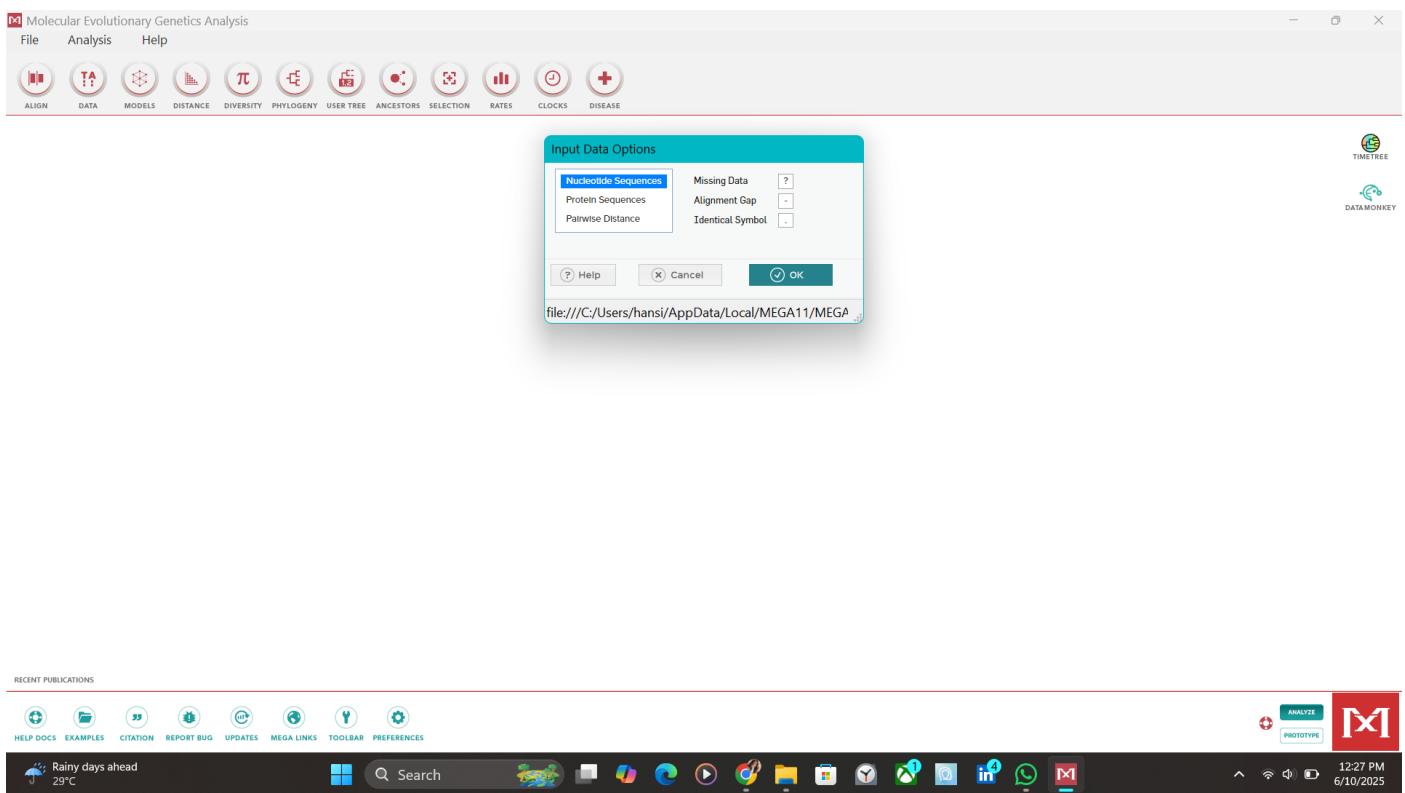
File → Open A file/session → upload file that exported from Bioedit



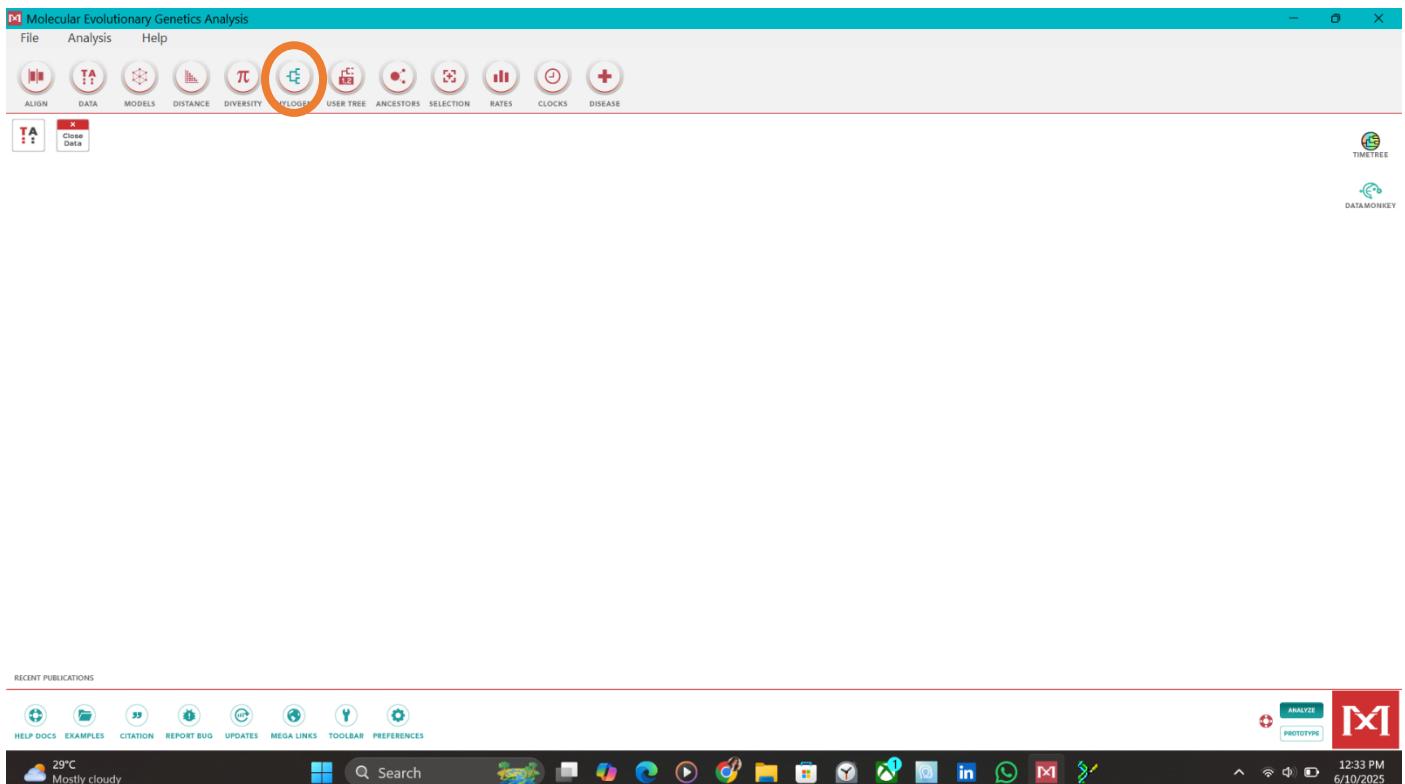
02. Select Analyze.



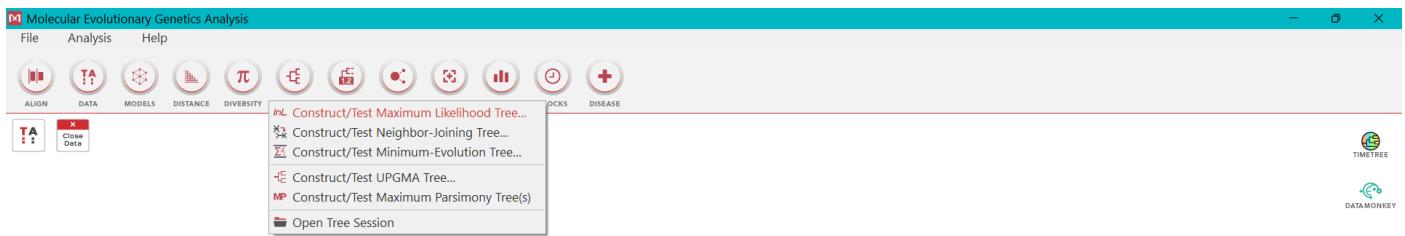
### 03. Sequence type – Nucleotide & OK



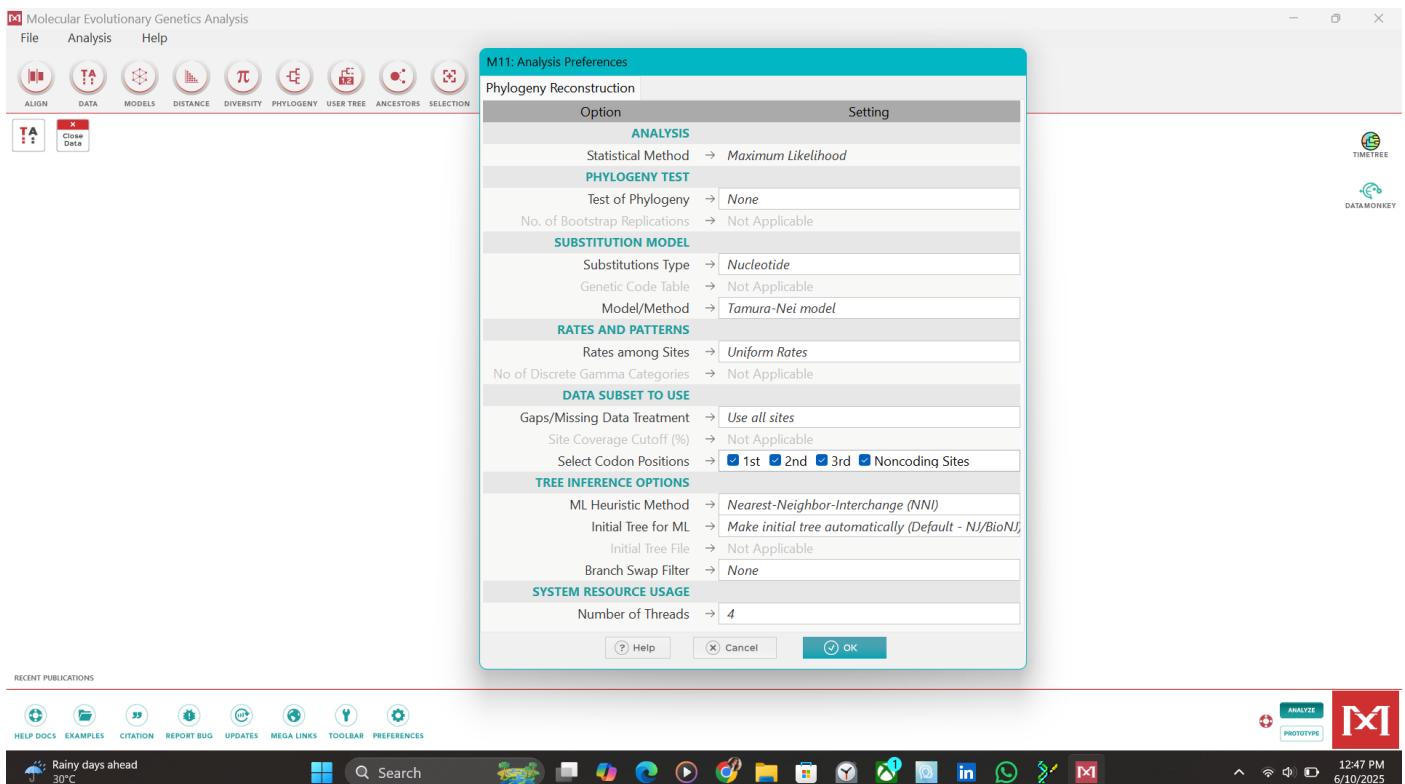
### 04. Clicked on phylogeny in main menu.



## 05. Selected construct/ Test Maximum Likelihood tree



## 06. Set analysis preferences & OK.



This is the final result, that I got phylogenetic tree.

