



PHYLOGENETIC TREE ANALYSIS USING MEGA

Process of construction of the phylogenetic tree

- The whole process of construction of the phylogenetic tree is divided into five different steps.
 - i. Choosing an appropriate markers for the phylogenetic analysis.
 - ii. Multiple sequence alignments.
 - iii. Molecular phylogenetic tree building methods.
 - iv. Phylogenetic reconstruction.
 - v. Valuation of the phylogenetic tree.

Construction of phylogenetic tree using MEGA

Phylogenetic information marker: RUBISCO large subunit-which is encoded by the chloroplast genome.

Organisms:

- Chlamydomonas raudensis
- Riccia fluitans
- Lycopodium clavatum
- Cycas szechuanensis
- Oryza sativa Japonica
- Arabidopsis thaliana

Software used:

- » MEGA
- » UniProt

UniProtKB 2020_05 results

 Filter byⁱ
 BLAST Align Download Add to basket Columns

◀ 1 to 25 of 167,967 ▶ Show 25 ▾

 Reviewed (1,008)
Swiss-Prot

 Unreviewed (166,959)
TrEMBL

Popular organisms

[A. thaliana \(62\)](#)
[Rice \(39\)](#)
[B. subtilis \(4\)](#)
[Bovine \(4\)](#)
[Human \(4\)](#)

Other organisms

Search terms

Filter "large" as:

[gene name \(12\)](#)
[gene ontology \(11\)](#)
[organism \(49\)](#)
<https://www.uniprot.org/taxonomy/39947>

Quote terms: "rubisco large" ✕

<input type="checkbox"/>	Entry ▾	Entry name ▾	Protein names ▾	Gene names ▾	Organism ▾	Length ▾	
<input type="checkbox"/>	Q43088	RBCMT_PEA	Ribulose-1,5 bisphosphate carboxyla...	RBCMT	Pisum sativum (Garden pea)	489	
<input type="checkbox"/>	P00875	RBL_SPIOL	Ribulose bisphosphate carboxylase l...	rbcl	Spinacia oleracea (Spinach)	475	
<input type="checkbox"/>	P00876	RBL_TOBAC	Ribulose bisphosphate carboxylase l...	rbcl rbcA	Nicotiana tabacum (Common tobacco)	477	
<input type="checkbox"/>	P00880	RBL_SYNP6	Ribulose bisphosphate carboxylase l...	cbbL rbcA, rbcl, syc0130_c	Synechococcus sp. (strain ATCC 27144 / PCC 6301 / SAUG 1402/1) (Anacystis nidulans)	472	
<input type="checkbox"/>	P00877	RBL_CHLRE	Ribulose bisphosphate carboxylase l...	rbcl	Chlamydomonas reinhardtii (Chlamydomonas smithii)	475	
<input type="checkbox"/>	O03042	RBL_ARATH	Ribulose bisphosphate carboxylase l...	rbcl AtCg00490	Arabidopsis thaliana (Mouse-ear cress)	479	
<input type="checkbox"/>	P0C512	RBL_ORYSJ	Ribulose bisphosphate carboxylase l...	rbcl LOC_Osp1g00420, Nip064	Oryza sativa subsp. japonica (Rice)	477	
<input type="checkbox"/>	P0C2C2	RBL1C_CUPNE	Ribulose bisphosphate carboxylase l...	cbbL1 cbbL, cbxLC, cfxLC	Cupriavidus necator (Alcaligenes eutrophus) (Ralstonia eutropha)	486	
<input type="checkbox"/>	P11383	RBL_WHEAT	Ribulose bisphosphate carboxylase l...	rbcl	Triticum aestivum (Wheat)	477	
<input type="checkbox"/>	P21240	CPNB1_ARATH	Chaperonin 60 subunit beta 1, chlor...	CPN60B1 Cpn60-B(3), LEN1, At1g55490, T5A14.11	Arabidopsis thaliana (Mouse-ear cress)	600	
<input type="checkbox"/>	Q31NB3	RBL_SYNE7	Ribulose bisphosphate carboxylase l...	cbbL rbcl, Synpcc7942_1426	Synechococcus elongatus (strain PCC 7942 / FACHB-805) (Anacystis nidulans R2)	472	

Downloading the RUBISCO large subunit sequence from UniProt

>AEM75128.1 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit, partial (chloroplast)
[Chlamydomonas raudensis]

MVPQTETRAGAGFKAGVKDYRLTYYPDYVVKDLDILAAFRMTPQPGVPAEECGAAVAESSTGTWTTVWTDGL
TSLDRYKGRICYDIEPVPGEDNQYIAYVAYPIDLFEEGSVTNLFTSIVGNVFGFKALRALRLEDLRIPPAYVKTFSGPP
HGIQVERDKINKYGRGLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNPQPFMRWRDRFLFCAEAIYK
AQSETGEVKGHYLNATAATSEEMIKRVVCAKEFGVPIIMHDYLTGGFTANTSLSNYCRDHGLLLHIHRAMHAVIDR
QRNHGIIHFRVLAKALRLSGGDHLHSGTVVGKLEGEREVTLGFDLMDRNYVEKDRSRGIYFTQDWCSMGGVMP
VASGGIHVWHMPALVEIFGDDACLQFGGGTGLGHPWGNAPGAVANRVALEACTQARNEGRDLAREGGDV

>YP_009659421.1 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (chloroplast) [Riccia fluitans]

MSPQTETKAGAGFKAGVKDYRLTYYPDYETKDTDILAAFRMTPQPGVPAEEAGAAVAESSTGTWTTVWTDGL
TNLDRYKGRICYDIEPVAGEQNQYIAYVAYPLDLFEEGSVTNMFTSIVGNVFGFKALRALRLEDLRIPPAYTKTFQG
PPHGIQVERDKLNKYGRPLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFRWRDRFLFVAEA
LFKAQAETGEIKGHYLNATAGTCEEMLKRAACARELGVPIVMHDYLTGGFTANTTLAFYCRDNGLLLHIHRAMHAV
IDRQKNHGMHFRVLAKGLRLSGGDHIIHAGTVVGKLEGDREVTLGFDLLRDDYIEKDRSRGVYFTQDWVSLPGV
LPVASGGIHVWHMPALTEIFGDDSVLQFGGGTGLGHPWGNAPGAVANRVALEACVQARNEGRDLARQGNEIIREA
AKWSPELAAACEVWKEIKFEFDIIDL

>BAJ40178.1 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit, partial (chloroplast) [Lycopodium clavatum]

YTPNYKTKDLDILAAFRMTPQPGVPPEEAGAAVAESSTGTWTTVWTDGLTSLDRYKGRICYDIEPVVGEKNQYIA
YVAYPLDLFEEGSVTNLFTSIVGNVFGFKALRALRLEDLRIPPAYSKTFIGPPHGIQVERDKLNKYGRPLLGCTIKPK
LGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFRWRDRFLFVAEALNKAQVETGEIKGHYLNATAGTYEEML
KRAQCAKELGVPIIMHDYLTGGFTANTS LAHYCRDNGLLLHIHRAMHAVIDRQKNHGIHFRVLAKALRMSGGDHIIH
AGTVVGKLEGERDVTLGFDLLRDDYIEKDRSRGIYFTQDWVSMGPVLPVASGGIHVWHMPALTEIFGDDSVLQF
GGGTGLGHPWGNAPGAVANRVALEA

>YP_009647987.1 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (chloroplast) [Cycas szechuanensis]

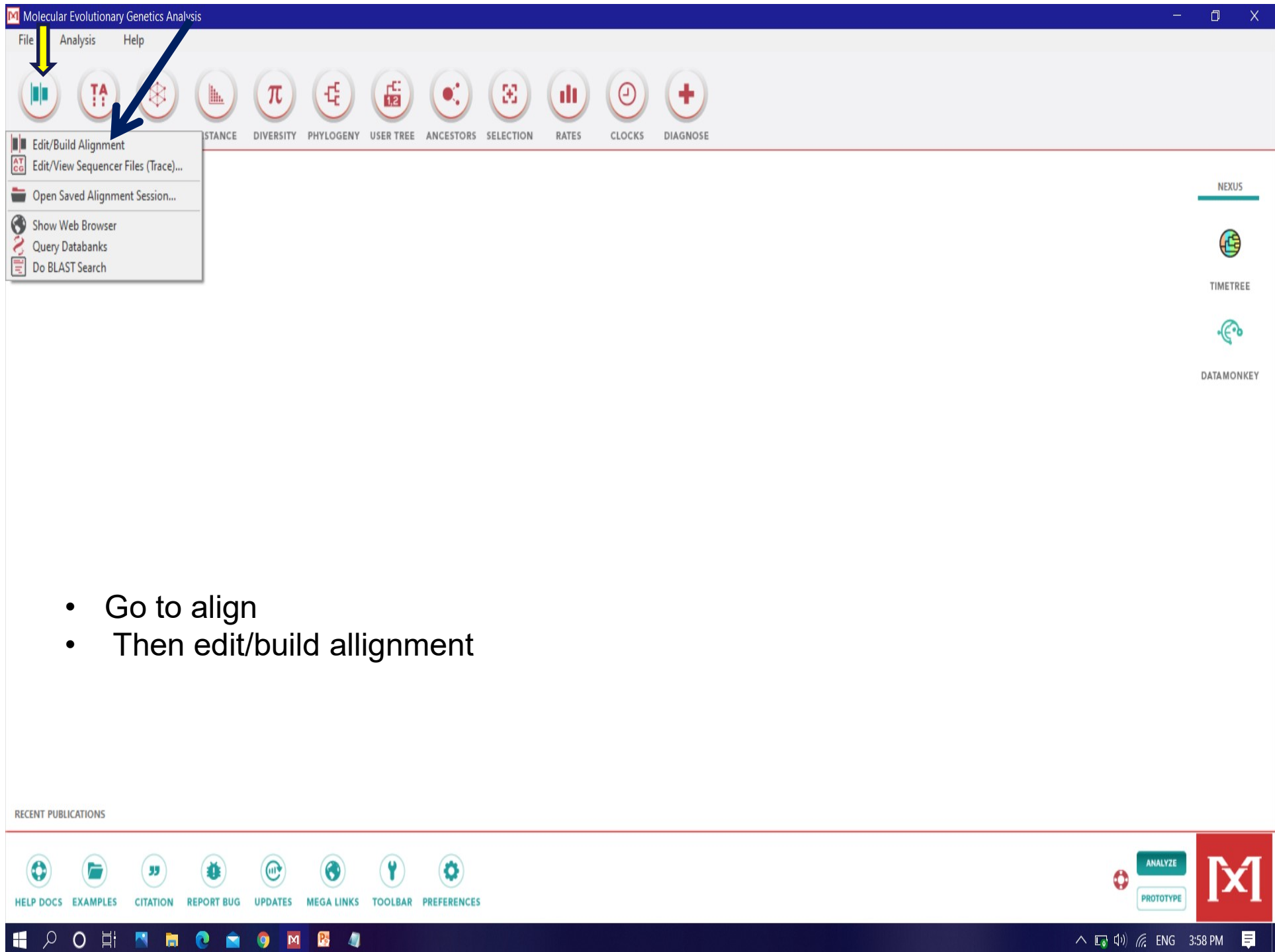
MSPKTETKASVGFKAGVKDYRLTYYTPEYQTKDTDILAAFRVTPQPGVPPEEAGAAVA AESSTGTWTTVWTDGL
TSLDRYKGRICYDIEPVPGEENQFIAYVAYPLDLFEEGSVTNMFTSIVGNVFGFKALRALRLEDLRVPPAYSKTFQG
PPHGIQVERDKLNKYGRPLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMRWRDRFCFAEAI
YKAQAETGEIKGHYLNATAGTCEEMIKRAVFARELGVPIVMHDYLTGGFTANTSLAHYCRDNGLLLHIHRAMHAVI
DRQRNHGMHFRVLAKALRMSGGDHIHAGTVVGKLEGERDVTLGFDLLRDDFIERDRSRGIYFTQDWVSM PGVL
PVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAVANRVALEACVQARNEGRDLAREGNEVIREA
SKWSPELAAACEVWKEIKFEFEAMDVL

>ADD62840.1 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (chloroplast) [Oryza sativa Japonica Group]

MSCREGRMSPQTETKASVGFKAGVKDYKLTYYTPEYETKDTDILAAFRVTPQPGVPPEEAGAAVA AESSTGTWT
TVWTDGLTSLDRYKGRICYHIEPVVGEDNQYIAYVAYPLDLFEEGSVTNMFTSIVGNVFGFKALRALRLEDLRIPPT
YSKTFQGGPPHGIQVERDKLNKYGRPLLGCTIKPKLGLSAKNYGRACYECLRGGLDFTKDDENVNSQPFMRWRDR
FVFCAEAIYKSQAETGEIKGHYLNATAGTCEEMIKRAVFARELGVPIVMHDYLTGGFTANTSLAHYCRDNGLLLHIH
RAMHAVIDRQKNHGMHFRVLAKALRMSGGDHIHAGTVVGKLEGEREMTLGFVDLLRDDFIEKDRARGIFFTQDW
VSM PGVIPVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAAANRVALEACVQARNEGRDLAREG
NEIIRSACKWSPELAAACEIWKAIKFEFEPVDKLDS

>NP_051067.1 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (chloroplast) [Arabidopsis thaliana]

MSPQTETKASVGFKAGVKEYKLTYYTPEYETKDTDILAAFRVTPQPGVPPEEAGAAVA AESSTGTWTTVWTDGLT
SLDRYKGRICYHIEPVPGEETQFIAYVAYPLDLFEEGSVTNMFTSIVGNVFGFKALAALRLEDLRIPPAYTKTFQGGP
HGIQVERDKLNKYGRPLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMRWRDRFLFC AEAIYK
SQAETGEIKGHYLNATAGTCEEMIKRAVFARELGVPIVMHDYLTGGFTANTSLSHYCRDNGLLLHIHRAMHAVIDR
QKNHGMHFRVLAKALRLSGGDHIHAGTVVGKLEGDRESTLGFVDLLRDDYVEKDRSRGIFFTQDWVSLPGVLPV
ASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAVANRVALEACVQARNEGRDLAVEGNEIIREACKW
SPELAAACEVWKEITFNFTIDKLDGQE



- Go to align
- Then edit/build alignment



Alignment Editor

Select an Option

- Create a new alignment
- Open a saved alignment session
- Retrieve a sequence from a file

? Help X Cancel ✓ OK

Done Loading



RECENT PUBLICATIONS



Molecular Evolutionary Genetics Analysis

File Analysis Help

ALIGN DATA MODELS DISTANCE DIVERSITY PHYLOGENY USER TREE ANCESTORS SELECTION RATES CLOCKS DIAGNOSE

NEXUS

MEGA Alignment Explorer

Data Edit Search Alignment Web Sequencer Display Help

DNA Sequences Translated Protein Sequences

Species/Abbrev

1. Sequence 1

Data Type for Alignment

Are you building a DNA or protein sequence alignment?

DNA Protein Cancel

RECENT PUBLICATIONS

HELP DOCS EXAMPLES CITATION REPORT BUG UPDATES MEGA LINKS TOOLBAR PREFERENCES

➤ As I am building protein sequence alignment

MX: Alignment Explorer

Data Edit Search Alignment Web Sequencer Display Help

Protein S
Species/A
1. Seque

- Undo
- Copy
- Cut
- Paste
- Delete
- Delete Gaps
- Insert Blank Sequence
- Insert Sequence From File
- Select Site(s)
- Select Sequence(s)
- Select All
- Allow Base Editing
- Modify All Bases To Upper Case

Site # 0 with w/o gaps

- Go to Edit
- Then insert sequence from file

The screenshot shows the MX: Alignment Explorer software interface. The 'Alignment' menu is open, displaying options for alignment methods and site management. The protein sequences are listed on the left, and the alignment view shows multiple sequences aligned. The 'Auto-Fill Gaps' option is checked in the menu.

Alignment Menu Options:

- Align by ClustalW
- Align by ClustalW (Codons)
- Align by MUSCLE
- Align by MUSCLE (Codons)
- Mark/Unmark Site
- Align Marked Sites
- Unmark All Sites
- Delete Gap-Only Sites
- Auto-Fill Gaps (checked)

Protein Sequences:

Species/Abbrv	Sequence
1. AEM75128.1 ribulosin	M V P Q
2. YP_009659421.1 ribulosin	M S P Q
3. BAJ40178.1 ribulosin	Y T P N
4. YP_009647987.1 ribulosin	M S P K
5. ADD62840.1 ribulosin	M S C R
6. NP_051067.1 ribulosin	M S P Q
7. Sequence 1	

Alignment View:

```
L T Y Y T P D Y V V K D T D I L A A F R M T P Q P G V P A E E C G A A V A A E S S T G T W T T V W T D G L T S L D R Y K G R C Y D I E P V P G E D N Q Y I A Y V A Y P I D L F E E G S V T N L F
L T Y Y T P D Y E T K D T D I L A A F R M T P Q P G V P A E E A G A A V A A E S S T G T W T T V W T D G L T N L D R Y K G R C Y D I E P V A G E Q N Q Y I A Y V A Y P L D L F E E G S V T N M F
P G V P P E E A G A A V A A E S S T G T W T T V W T D G L T S L D R Y K G R C Y D I E P V V G E K N Q Y I A Y V A Y P L D L F E E G S V T N L F T S I V G N V F G F K A L R A L R L E D L R I P
L T Y Y T P E Y Q T K D T D I L A A F R V T P Q P G V P P E E A G A A V A A E S S T G T W T T V W T D G L T S L D R Y K G R C Y D I E P V P G E E N Q F I A Y V A Y P L D L F E E G S V T N M F
A G V K D Y K L T Y Y T P E Y E T K D T D I L A A F R V T P Q P G V P P E E A G A A V A A E S S T G T W T T V W T D G L T S L D R Y K G R C Y H I E P V V G E D N Q Y I A Y V A Y P L D L F E E
L T Y Y T P E Y E T K D T D I L A A F R V T P Q P G V P P E E A G A A V A A E S S T G T W T T V W T D G L T S L D R Y K G R C Y H I E P V P G E E T Q F I A Y V A Y P L D L F E E G S V T N M F
```

Site # 1 with w/o gaps

- Select all the sequence
- Go to alignment
- Then chose align by ClustalW for MSA

MX: Alignment Explorer

Data Edit Search Alignment Web Sequencer Display Help

Create New
 Open
 Open a Recently Used File
 Close

Phylogenetic Analysis

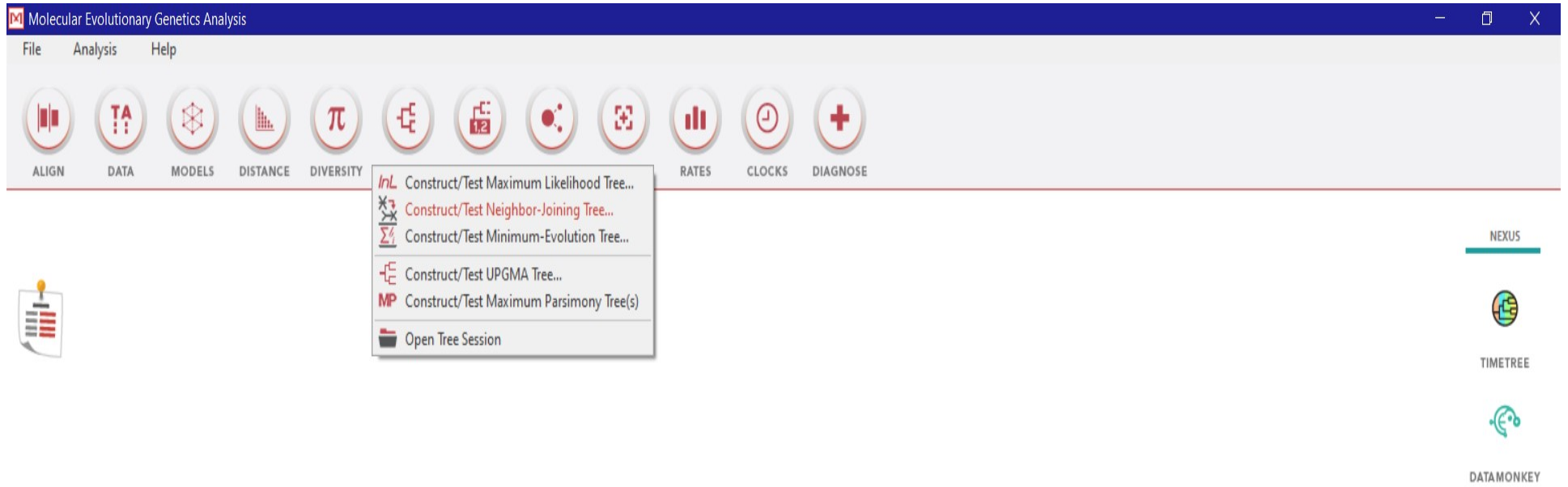
Save Session
 Export Alignment
 DNA Sequences
 Protein Sequences
 Translate/Untranslate Genetic Code
 Reverse Complement
 Reverse
 Complement
 Quit

MEGA Format
 FASTA Format
 NEXUS/PAUP Format

CSMGGVMPVASGGIHVWHMPALVEIFGDDACLQFGGGTLGHPWGNAPGAVANRVALEACTQARNEGRDLAREGGDV
 VSLPGVLPVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAVANRVALEACVQARNEGRDLARQGNEIIREAAKWSPELAAACEVWKEIKFEFDIIDTL
 VSMGPVLPVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAVANRVALEA
 IHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAAANRVALEACVQARNEGRDLAREGNEIIRSACKKWSPELAAACEIWKAIKFEFEPVDKLDSE

Site # 486 with w/o gaps

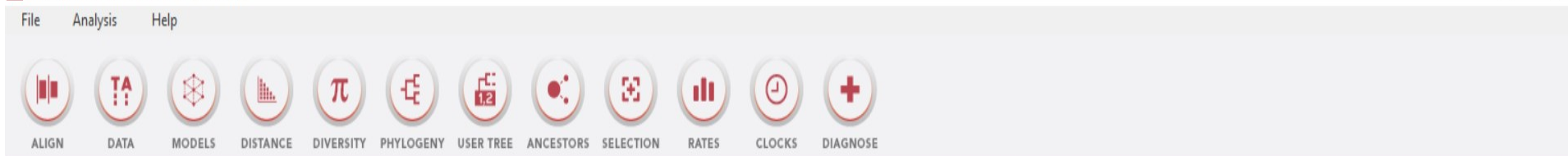
- After getting the MSA
- Go to DATA
- Then export alignment-in MEGA format



- Go to phylogeny
- And chose required option
- Then chose the MEGA file which was exported earlier

RECENT PUBLICATIONS





- Let the analysis preferences as default.
- Or set according to your requirement.

MX: Analysis Preferences

Phylogeny Reconstruction

Option	Setting
ANALYSIS	
Scope →	All Selected Taxa
Statistical Method →	Neighbor-joining
PHYLOGENY TEST	
Test of Phylogeny →	Bootstrap method
No. of Bootstrap Replications →	1000
SUBSTITUTION MODEL	
Substitutions Type →	Amino acid
Model/Method →	Poisson model
RATES AND PATTERNS	
Rates among Sites →	Uniform Rates
Gamma Parameter →	Not Applicable
Pattern among Lineages →	Same (Homogeneous)
DATA SUBSET TO USE	
Gaps/Missing Data Treatment →	Pairwise deletion
Site Coverage Cutoff (%) →	Not Applicable
SYSTEM RESOURCE USAGE	
Number of Threads →	3

? Help × Cancel ✓ OK

NEXUS



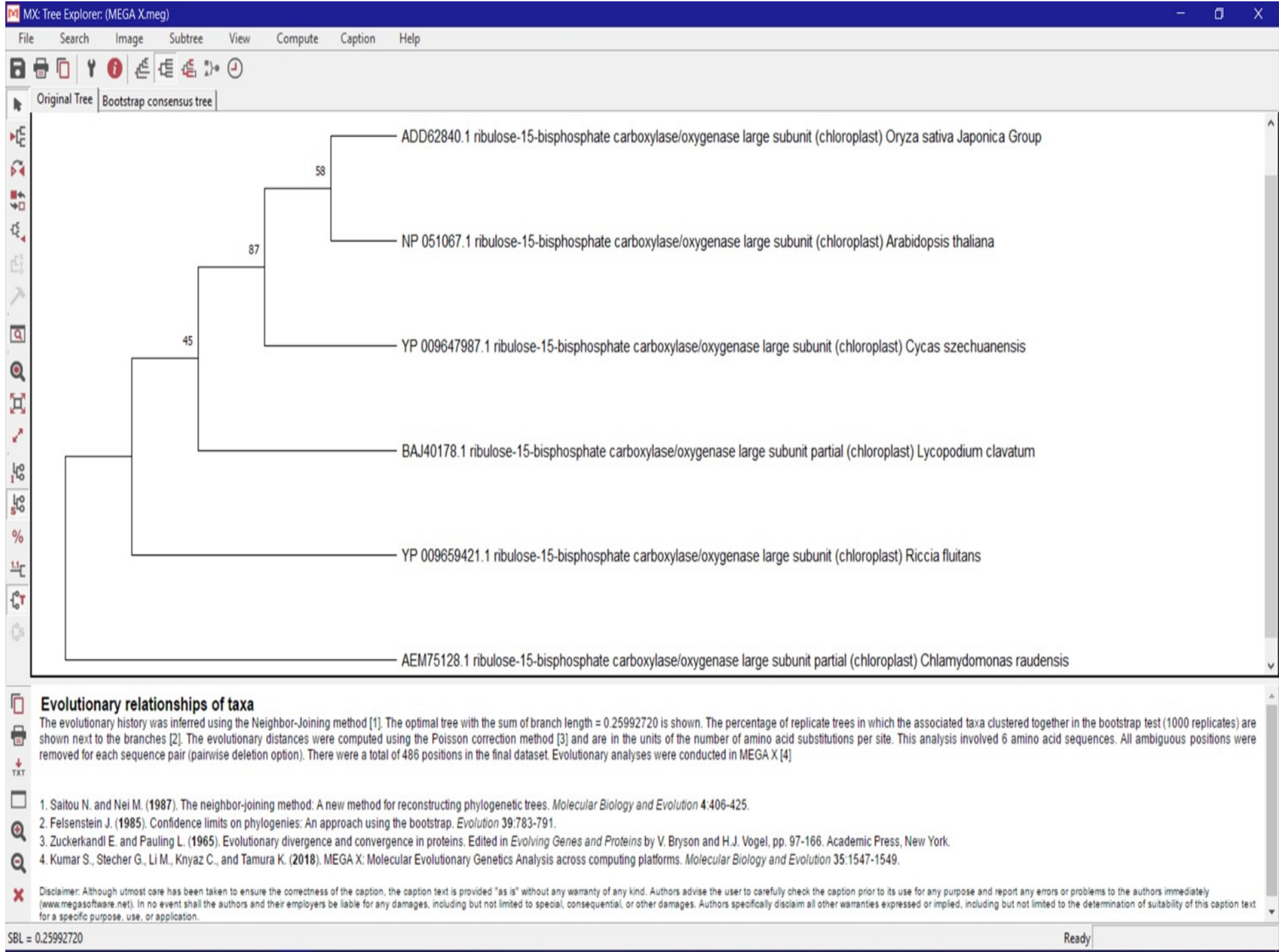
TIMETREE



DATAMONKEY

RECENT PUBLICATIONS





CONCLUSION

By constructing PHYLOGENETIC TREE:-

- We can find out the evolutionary history.
- Can measure phylogenetic diversity using phylogenetic trees.
- Can trace evolutionary history of infectious bacteria and viruses.
- To guide our search for new species.
- Find out how species spread geographically in their evolution.
- To tell us when taxa originated and where.