MEGA

Molecular Evolutionary Genetics Analysis

Ignacio Bitrián Arcas - 717901
Nicolás Lera López - 721808
David Márquez Calavia - 700940

¿Qué es?

- Software de carácter privativo que permite realizar análisis estadísticos de la evolución molecular y la construcción de árboles filogenéticos.

- Desarrollado en 1993 en la Universidad Estatal de Pensilvania.

Características

- Alineamiento y visualización de secuencias
- Métodos de estimación de distancias
- Visualización y métodos de creación de árboles

Y mucho más...

Características

	Likelihood	Distance	Parsimony	Bayesian	Visual Explorer	Caption Expert
Phylogeny	✓	✓	✓		✓	✓
Bootstrap	✓	✓	✓		✓	✓
Distance/Diversity	✓	✓			✓	✓
Model Selection	✓					
Substitution Pattern	✓				✓ XL	✓
Rate Variation	✓				✓ XL	✓
Ancestral Sequence			✓	✓	✓	✓
Clock Test	✓	✓			✓ XL	✓
Time Tree	✓	✓			✓	✓
Selection Test	✓	✓			✓	✓
Disease Mutation	✓	✓			✓ XL	✓

Instalación

- Windows
- MacOs
- Ubuntu / Debian
- RedHat
- Otros sistemas Linux

Interfaz gráfica y línea de comandos

Windows (GUI/CLI)

Descargar el .exe de la web y seguir los pasos indicados por el asistente.

MacOs (GUI/CLI)

Descargar el .dmg de la web, abrirlo y arrastrar el icono de MEGA a la carpeta de aplicaciones en la ventana emergente.

Ubuntu/Debian(GUI/CLI)

Descargar el .deb de la web, navegar con una terminal hasta el directorio que lo contenga y ejecutar:

```
GUI: sudo dpkg -i megax_10.0.5-1_amd64.deb
CLI: sudo dpkg -i megax-cc_10.0.5-1_amd64.deb
```

RedHat(GUI/CLI)

Descargar el .rpm de la web, navegar con una terminal hasta el directorio que lo contenga y ejecutar:

```
GUI: sudo rpm -ivh megax-10.0.5-1.x86_64.rpm
CLI: sudo rpm -ivh megax-cc-10.0.5-1.x86_64.rpm
```

En caso de tener una versión anterior descargada, a lo mejor es necesario ejecutar:

```
GUI: sudo rpm -ivh --replacefiles megax-10.0.5-1.x86_64.rpm
CLI: sudo rpm -ivh --replacefiles megax-cc-10.0.5-1.x86_64.rpm
```

Otros Linux (CLI)

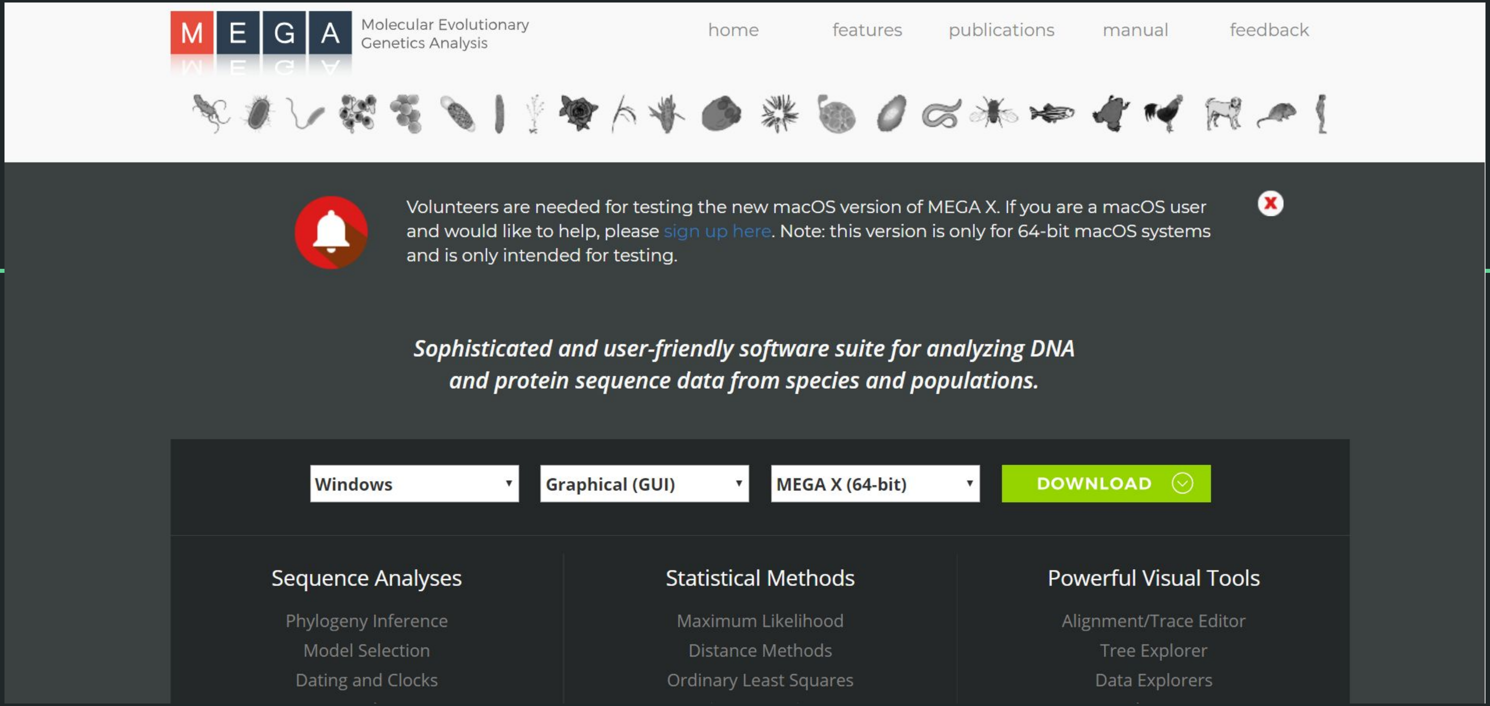
Descargar el .tar.gz de la web y extraerlo mediante el comando:

```
tar -zxvf megacc-7.0.26-1.x86_64.tar.gz
```

Después, mover los binarios y archivos dependientes al directorio que se prefiera.

Windows (GUI)

Descargar el .exe de la web. (https://www.megasoftware.net)



Volunteers are needed for testing the new macOS version of MEGA X. If you are a macOS user and would like to help, please sign up here. Note: this version is only for 64-bit macOS systems and is only intended for testing.

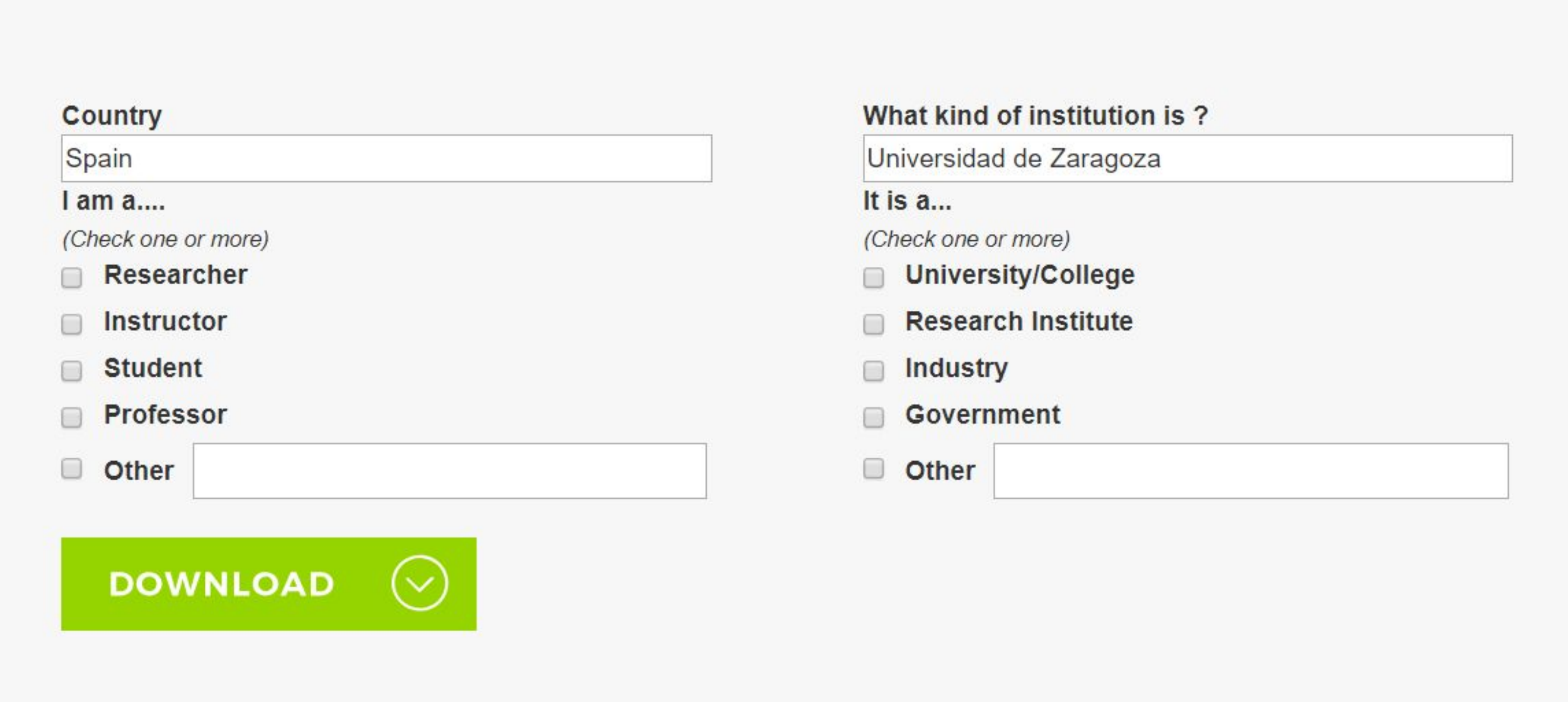
Sophisticated and user-friendly software suite for analyzing DNA and protein sequence data from species and populations.

Windows ▾ | Graphical (GUI) ▾ | MEGA X (64-bit) ▾

Sequence Analyses	Statistical Methods	Powerful Visual Tools
Phylogeny Inference	Maximum Likelihood	Alignment/Trace Editor
Model Selection	Distance Methods	Tree Explorer
Dating and Clocks	Ordinary Least Squares	Data Explorers

Windows (GUI)

Rellenar perfil.



Country

Spain

I am a....

(Check one or more)

☐ Researcher
☐ Instructor
☐ Student
☐ Professor
☐ Other

What kind of institution is ?

Universidad de Zaragoza

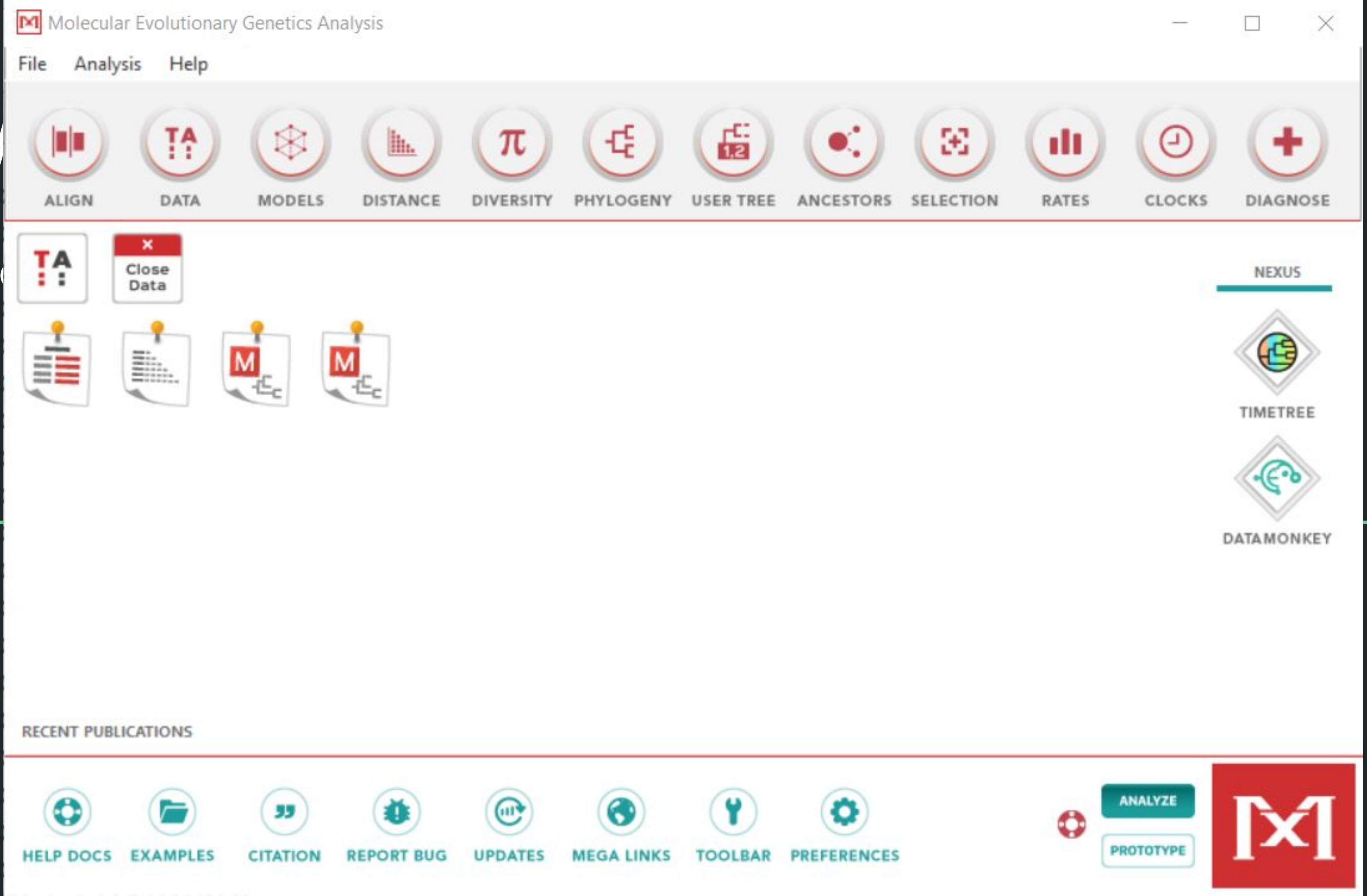
It is a...

(Check one or more)

☐ University/College
☐ Research Institute
☐ Industry
☐ Government
☐ Other

W

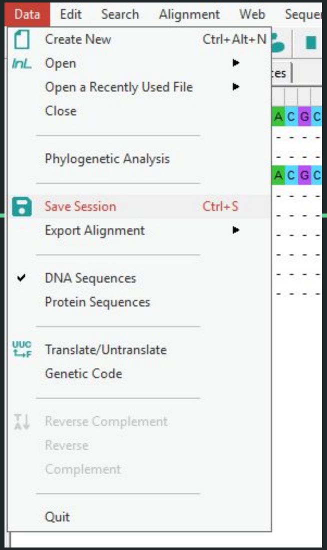
S


Molecular Evolutionary Genetics Analysis
File
Analysis
Help
ALIGN
DATA
MODELS
DISTANCE
DIVERSITY
PHYLOGENY
USER TREE
ANCESTORS
SELECTION
RATES
CLOCKS
DIAGNOSE
Close Data
NEXUS
TIMETREE
DATAMONKEY
RECENT PUBLICATIONS
HELP DOCS
EXAMPLES
CITATION
REPORT BUG
UPDATES
MEGA LINKS
TOOLBAR
PREFERENCES
ANALYZE
PROTOTYPE

Convertir archivo Fasta a Mega

- File -> Convert File to MEGA

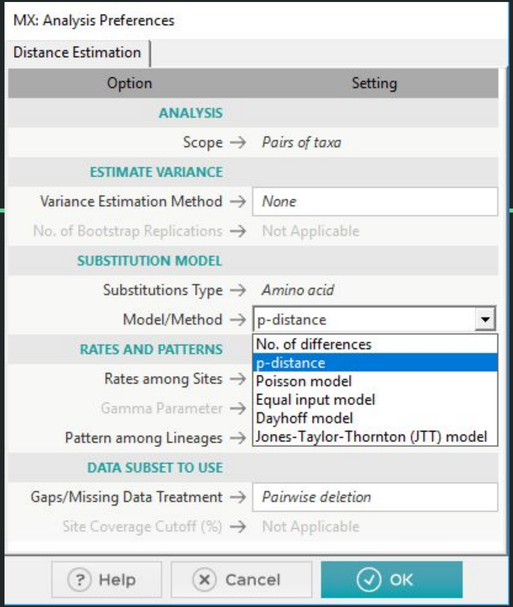
- Abrir archivo fasta con mega y guardar la sesión
  Data -> Save session



Cálculo de distancias

- Distance -> Compute Pairwise Distances

- Opciones:

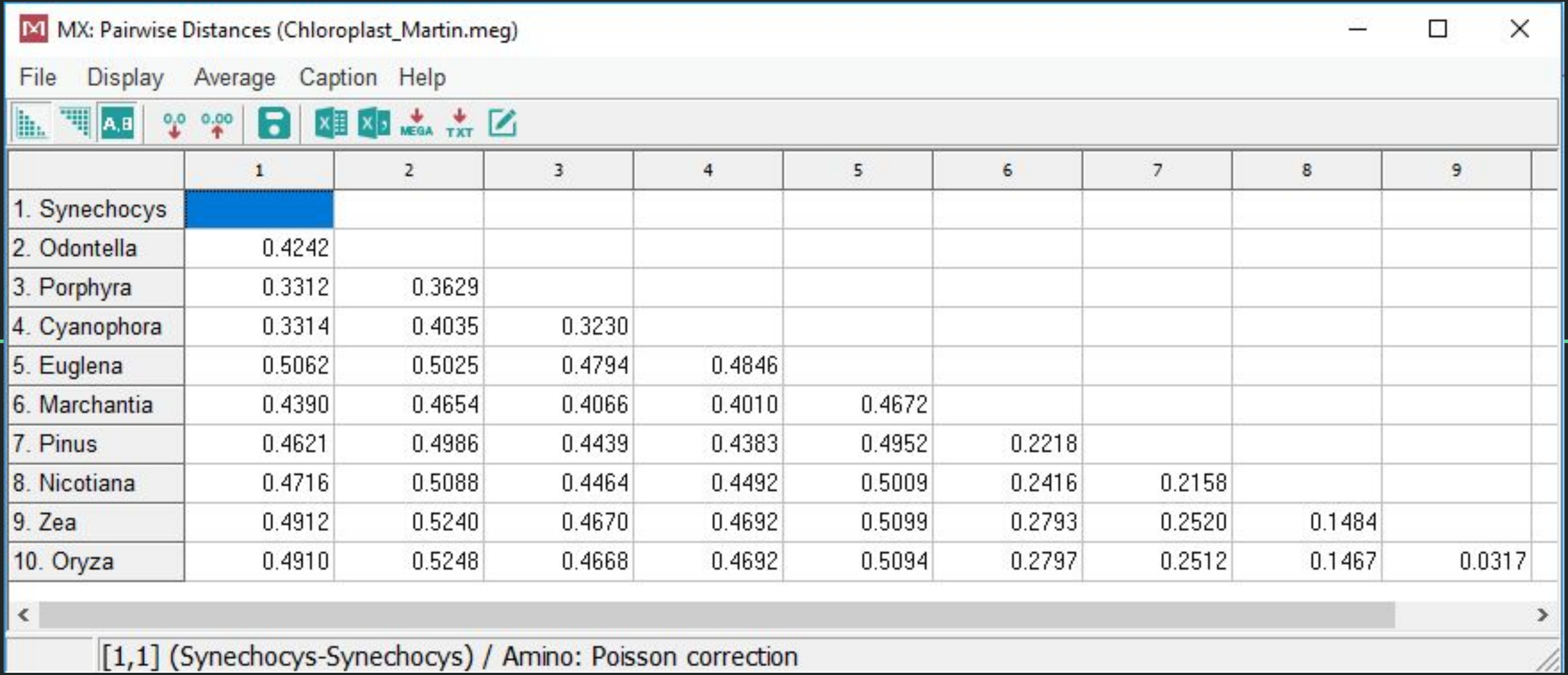


MX: Analysis Preferences

Distance Estimation

Option	Setting
ANALYSIS	
Scope →	*Pairs of taxa*
ESTIMATE VARIANCE	
Variance Estimation Method →	*None*
No. of Bootstrap Replications →	Not Applicable
SUBSTITUTION MODEL	
Substitutions Type →	*Amino acid*
Model/Method →	p-distance
RATES AND PATTERNS	
Rates among Sites →	
Gamma Parameter →	
Pattern among Lineages →	
DATA SUBSET TO USE	
Gaps/Missing Data Treatment →	*Pairwise deletion*
Site Coverage Cutoff (%) →	Not Applicable

Cálculo de distancias



	1	2	3	4	5	6	7	8	9
1. Synechocys									
2. Odontella	0.4242								
3. Porphyra	0.3312	0.3629							
4. Cyanophora	0.3314	0.4035	0.3230						
5. Euglena	0.5062	0.5025	0.4794	0.4846					
6. Marchantia	0.4390	0.4654	0.4066	0.4010	0.4672				
7. Pinus	0.4621	0.4986	0.4439	0.4383	0.4952	0.2218			
8. Nicotiana	0.4716	0.5088	0.4464	0.4492	0.5009	0.2416	0.2158		
9. Zea	0.4912	0.5240	0.4670	0.4692	0.5099	0.2793	0.2520	0.1484	
10. Oryza	0.4910	0.5248	0.4668	0.4692	0.5094	0.2797	0.2512	0.1467	0.0317

[1,1] (Synechocys-Synechocys) / Amino: Poisson correction

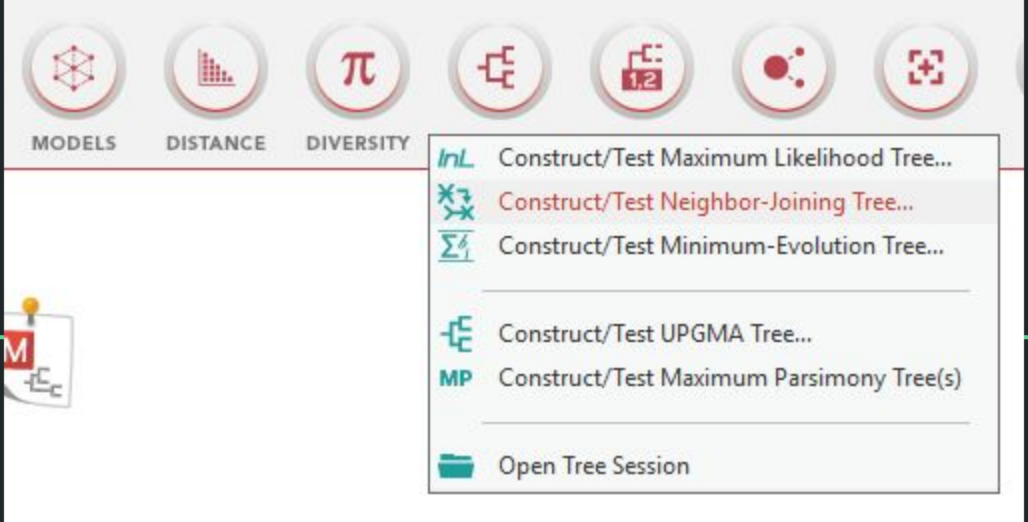
Construcción de árboles

Por distancias:

- NJ
- UPGMA

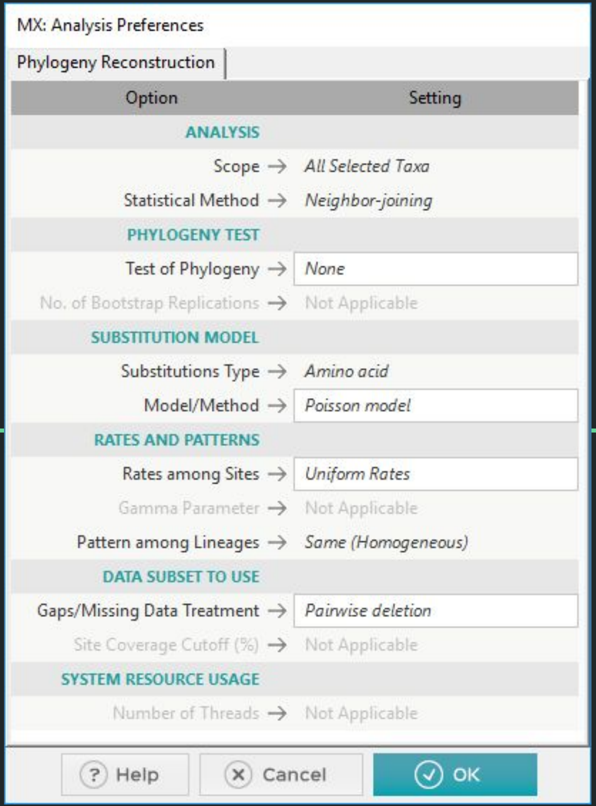
Por máxima parsimonia:

- MP



NJ

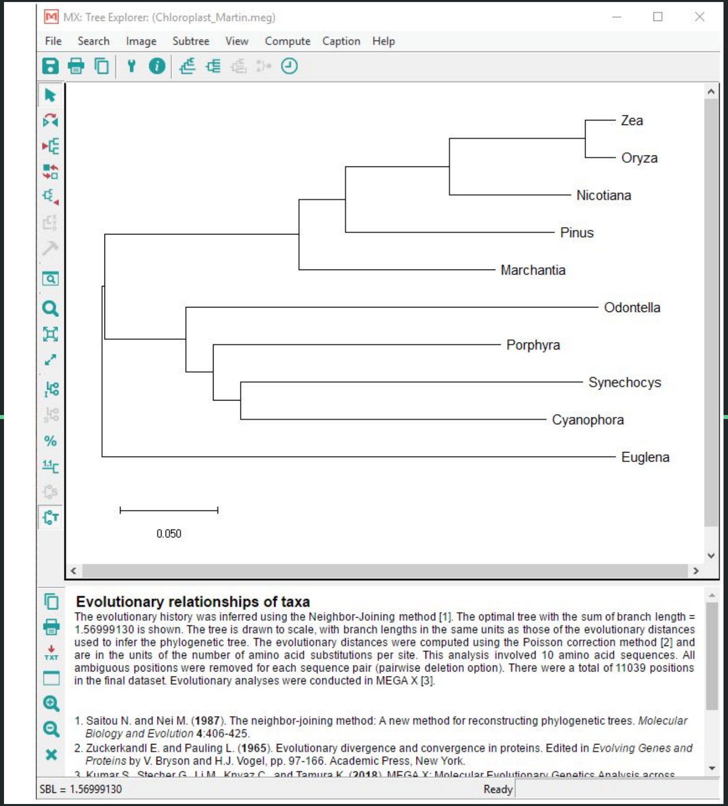
Opciones:



Option	Setting
ANALYSIS	
Scope →	*All Selected Taxa*
Statistical Method →	*Neighbor-joining*
PHYLOGENY TEST	
Test of Phylogeny →	*None*
No. of Bootstrap Replications →	Not Applicable
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 →	Not Applicable

NJ

Árbol generado:



Evolutionary relationships of taxa

The evolutionary history was inferred using the Neighbor-Joining method [1]. The optimal tree with the sum of branch length = 1.56999130 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [2] and are in the units of the number of amino acid substitutions per site. This analysis involved 10 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 11039 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [3].

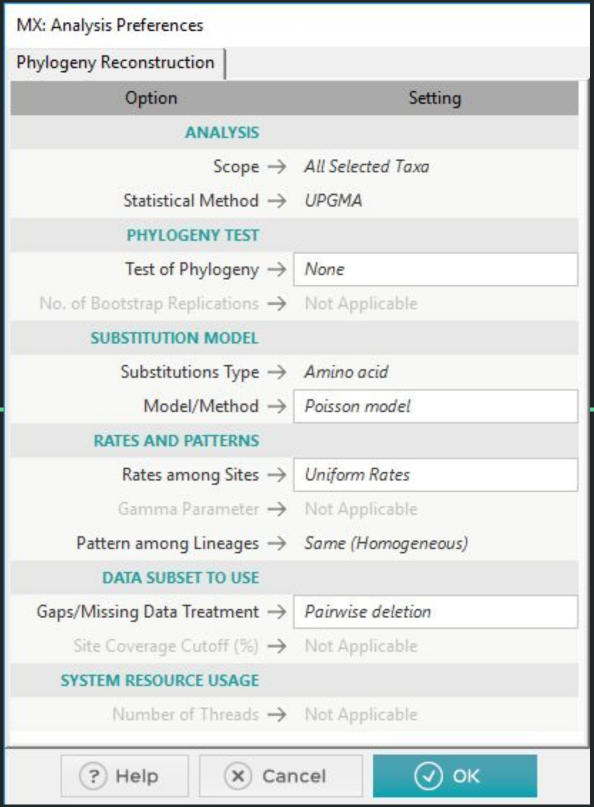
1. Saitou N. and Nei M. (**1987**). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**:406-425.
2. Zuckerkandl E. and Pauling L. (**1965**). Evolutionary divergence and convergence in proteins. Edited in *Evolving Genes and Proteins* by V. Bryson and H.J. Vogel, pp. 97-166. Academic Press, New York.
3. Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (**2018**). MEGA X: Molecular Evolutionary Genetics Analysis across

SBL = 1.56999130

Ready

UPGMA

Opciones:



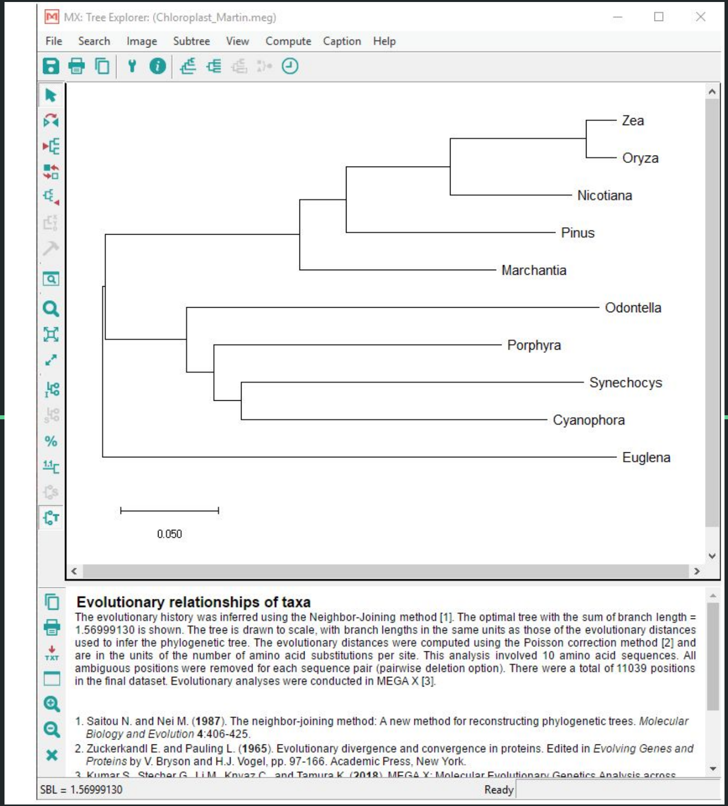
MX: Analysis Preferences

Phylogeny Reconstruction

Option	Setting
ANALYSIS	
Scope →	*All Selected Taxa*
Statistical Method →	*UPGMA*
PHYLOGENY TEST	
Test of Phylogeny →	*None*
No. of Bootstrap Replications →	Not Applicable
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 →	Not Applicable

UPGMA

Árbol generado:



Evolutionary relationships of taxa

The evolutionary history was inferred using the Neighbor-Joining method [1]. The optimal tree with the sum of branch length = 1.56999130 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [2] and are in the units of the number of amino acid substitutions per site. This analysis involved 10 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 11039 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [3].

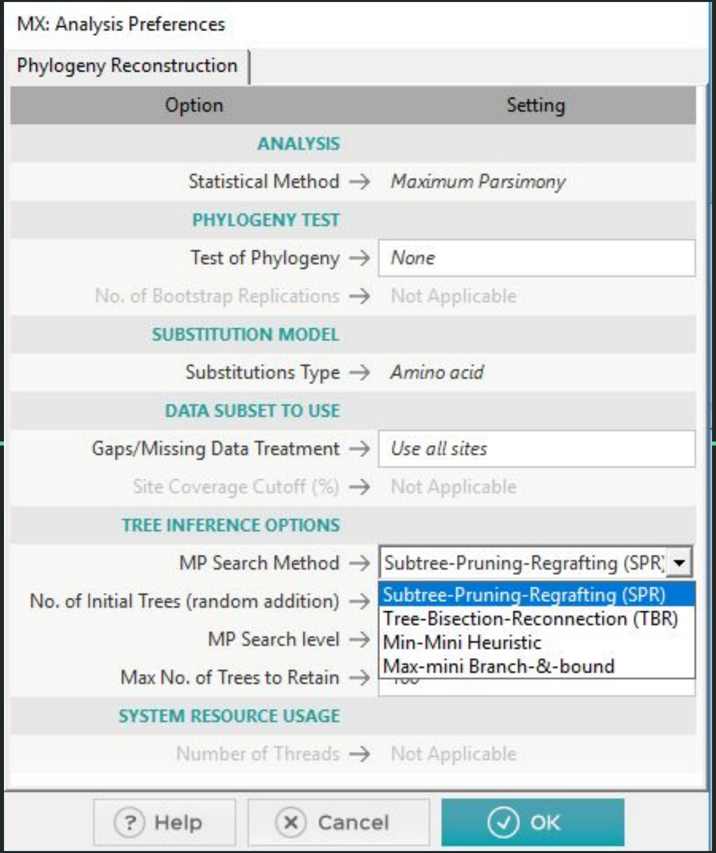
1. Saitou N. and Nei M. (**1987**). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**:406-425.
2. Zuckerkandl E. and Pauling L. (**1965**). Evolutionary divergence and convergence in proteins. Edited in *Evolving Genes and Proteins* by V. Bryson and H.J. Vogel, pp. 97-166. Academic Press, New York.
3. Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (**2018**). MEGA X: Molecular Evolutionary Genetics Analysis across

SBL = 1.56999130

Ready

MP

Opciones:



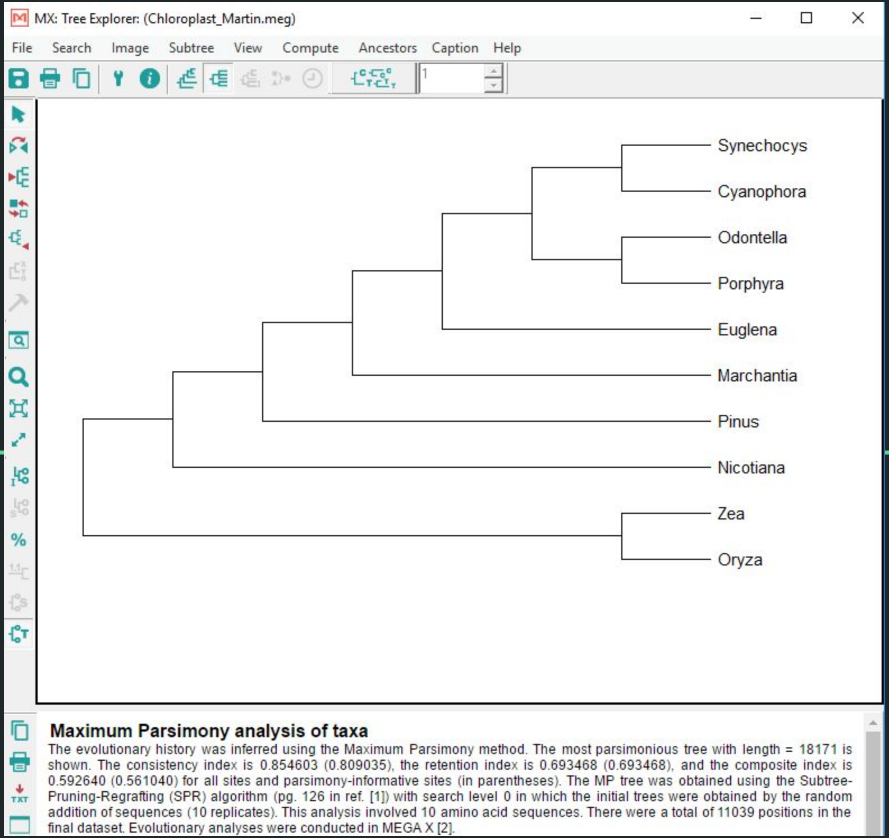
MX: Analysis Preferences

Phylogeny Reconstruction

Option	Setting
ANALYSIS	
Statistical Method →	*Maximum Parsimony*
PHYLOGENY TEST	
Test of Phylogeny →	*None*
No. of Bootstrap Replications →	Not Applicable
SUBSTITUTION MODEL	
Substitutions Type →	*Amino acid*
DATA SUBSET TO USE	
Gaps/Missing Data Treatment →	*Use all sites*
Site Coverage Cutoff (%) →	Not Applicable
TREE INFERENCE OPTIONS	
MP Search Method →	Subtree-Pruning-Regrafting (SPR)
No. of Initial Trees (random addition) →	
MP Search level →	
Max No. of Trees to Retain →	
SYSTEM RESOURCE USAGE	
Number of Threads →	Not Applicable

MP

Árbol generado:

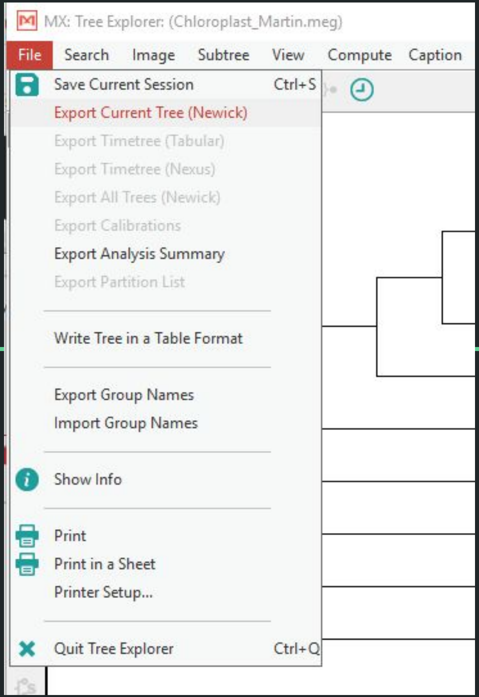


Maximum Parsimony analysis of taxa

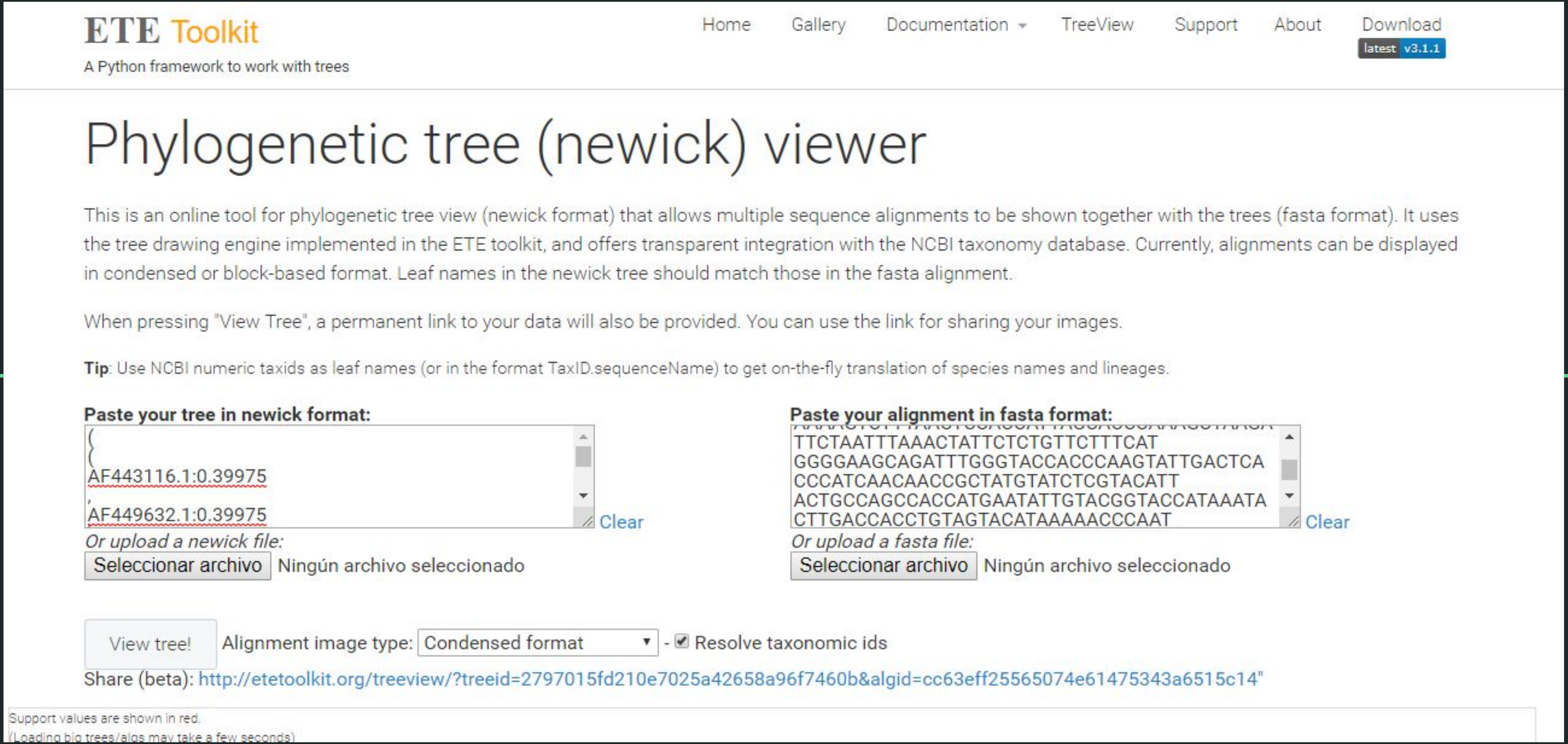
The evolutionary history was inferred using the Maximum Parsimony method. The most parsimonious tree with length = 18171 is shown. The consistency index is 0.854603 (0.809035), the retention index is 0.693468 (0.693468), and the composite index is 0.592640 (0.561040) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm (pg. 126 in ref. [1]) with search level 0 in which the initial trees were obtained by the random addition of sequences (10 replicates). This analysis involved 10 amino acid sequences. There were a total of 11039 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [2].

Exportar árboles

- File -> Export Current Tree (Newick)



ETE Toolkit



Phylogenetic tree (newick) viewer

This is an online tool for phylogenetic tree view (newick format) that allows multiple sequence alignments to be shown together with the trees (fasta format). It uses the tree drawing engine implemented in the ETE toolkit, and offers transparent integration with the NCBI taxonomy database. Currently, alignments can be displayed in condensed or block-based format. Leaf names in the newick tree should match those in the fasta alignment.

When pressing "View Tree", a permanent link to your data will also be provided. You can use the link for sharing your images.

Tip: Use NCBI numeric taxids as leaf names (or in the format TaxID.sequenceName) to get on-the-fly translation of species names and lineages.

Paste your tree in newick format:

```
(
(
AF443116.1:0.39975
,
AF449632.1:0.39975
```

Or upload a newick file:

Seleccionar archivo Ningún archivo seleccionado

Paste your alignment in fasta format:

```
TTCTAATTTAAACTATTCTCTGTTCTTTCAT
GGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCA
CCCATCAACAACCGCTATGTATCTCGTACATT
ACTGCCAGCCACCATGAATATTGTACGGTACCATAAATA
CTTGACCACCTGTAGTACATAAAAACCCAAT
```

Or upload a fasta file:

Seleccionar archivo Ningún archivo seleccionado

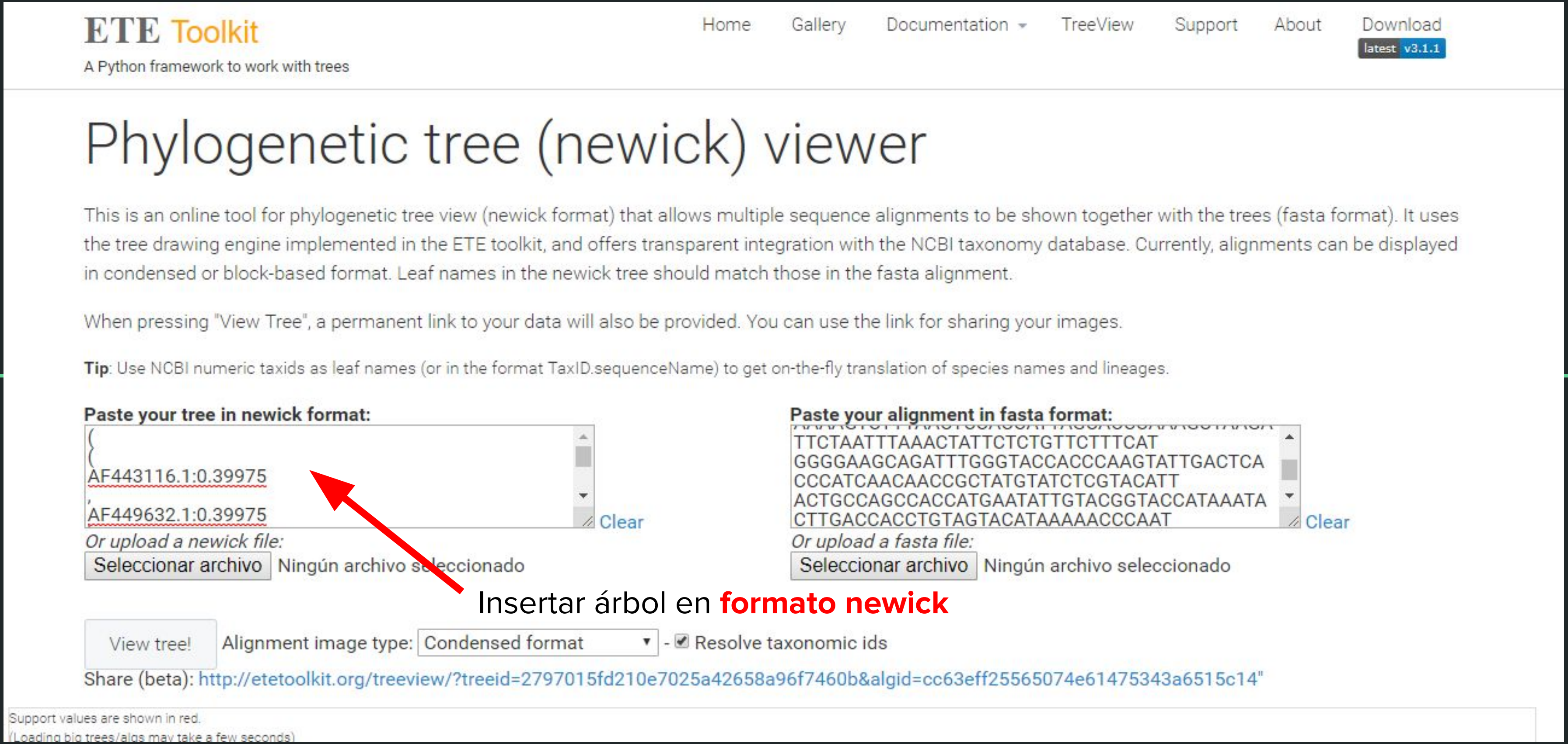
Alignment image type: Condensed format - ☑ Resolve taxonomic ids

Share (beta): http://etetoolkit.org/treeview/?treeid=2797015fd210e7025a42658a96f7460b&algid=cc63eff25565074e61475343a6515c14"

Support values are shown in red.

(Loading big trees/algs may take a few seconds)

ETE Toolkit



Phylogenetic tree (newick) viewer

This is an online tool for phylogenetic tree view (newick format) that allows multiple sequence alignments to be shown together with the trees (fasta format). It uses the tree drawing engine implemented in the ETE toolkit, and offers transparent integration with the NCBI taxonomy database. Currently, alignments can be displayed in condensed or block-based format. Leaf names in the newick tree should match those in the fasta alignment.

When pressing "View Tree", a permanent link to your data will also be provided. You can use the link for sharing your images.

Tip: Use NCBI numeric taxids as leaf names (or in the format TaxID.sequenceName) to get on-the-fly translation of species names and lineages.

Paste your tree in newick format:

```
(
(
AF443116.1:0.39975
,
AF449632.1:0.39975
```

Or upload a newick file:

Ningún archivo seleccionado

Paste your alignment in fasta format:

```
TTCTAATTTAAACTATTCTCTGTTCTTTCAT
GGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCA
CCCATCAACAACCGCTATGTATCTCGTACATT
ACTGCCAGCCACCATGAATATTGTACGGTACCATAAATA
CTTGACCACCTGTAGTACATAAAAACCCAAT
```

Or upload a fasta file:

Ningún archivo seleccionado

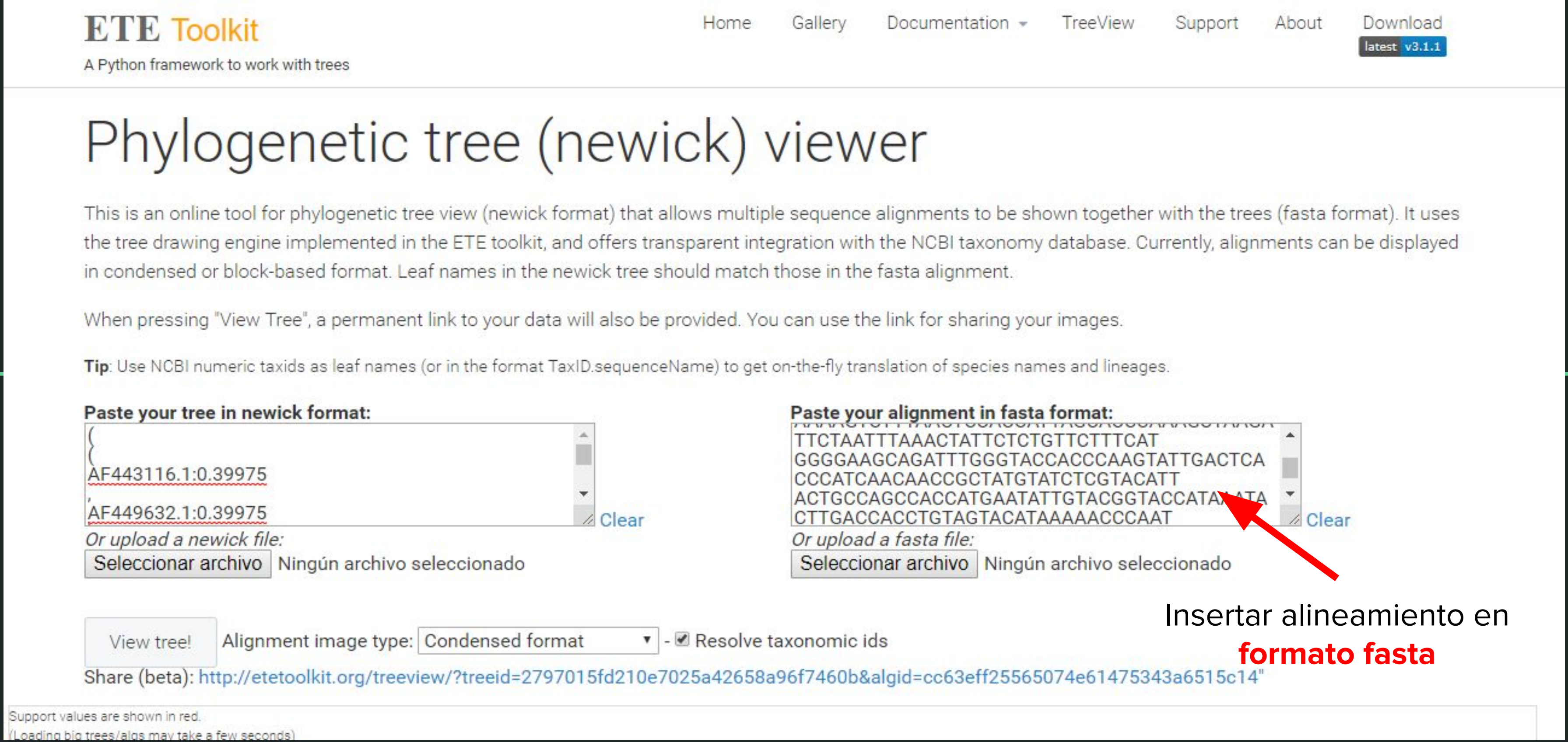
Insertar árbol en **formato newick**

Alignment image type: Condensed format - ☑ Resolve taxonomic ids

Share (beta): http://etetoolkit.org/treeview/?treeid=2797015fd210e7025a42658a96f7460b&algid=cc63eff25565074e61475343a6515c14"

Support values are shown in red.
(Loading big trees/algs may take a few seconds)

ETE Toolkit



Phylogenetic tree (newick) viewer

This is an online tool for phylogenetic tree view (newick format) that allows multiple sequence alignments to be shown together with the trees (fasta format). It uses the tree drawing engine implemented in the ETE toolkit, and offers transparent integration with the NCBI taxonomy database. Currently, alignments can be displayed in condensed or block-based format. Leaf names in the newick tree should match those in the fasta alignment.

When pressing "View Tree", a permanent link to your data will also be provided. You can use the link for sharing your images.

Tip: Use NCBI numeric taxids as leaf names (or in the format TaxID.sequenceName) to get on-the-fly translation of species names and lineages.

Paste your tree in newick format:

```
(
(
AF443116.1:0.39975
,
AF449632.1:0.39975
```

Or upload a newick file:

Seleccionar archivo Ningún archivo seleccionado

Paste your alignment in fasta format:

```
TTCTAATTTAAACTATTCTCTGTTCTTTCAT
GGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCA
CCCATCAACAACCGCTATGTATCTCGTACATT
ACTGCCAGCCACCATGAATATTGTACGGTACCATAAATA
CTTGACCACCTGTAGTACATAAAAACCCAAT
```

Or upload a fasta file:

Seleccionar archivo Ningún archivo seleccionado

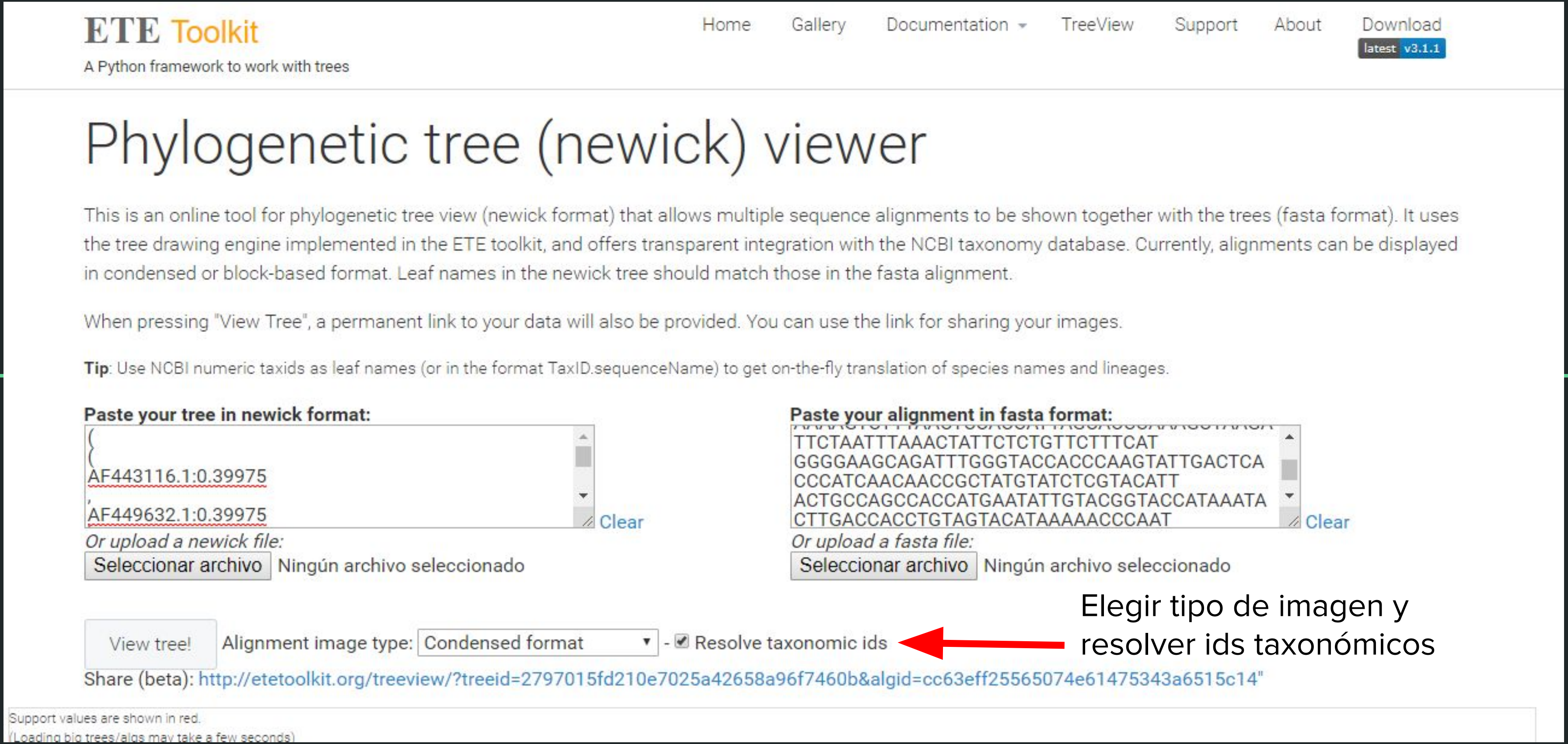
Alignment image type: Condensed format - ☑ Resolve taxonomic ids

Share (beta): http://etetoolkit.org/treeview/?treeid=2797015fd210e7025a42658a96f7460b&algid=cc63eff25565074e61475343a6515c14"

Support values are shown in red.
(Loading big trees/algs may take a few seconds)

Insertar alineamiento en **formato fasta**

ETE Toolkit



ETE Toolkit
A Python framework to work with trees

Phylogenetic tree (newick) viewer

This is an online tool for phylogenetic tree view (newick format) that allows multiple sequence alignments to be shown together with the trees (fasta format). It uses the tree drawing engine implemented in the ETE toolkit, and offers transparent integration with the NCBI taxonomy database. Currently, alignments can be displayed in condensed or block-based format. Leaf names in the newick tree should match those in the fasta alignment.

When pressing "View Tree", a permanent link to your data will also be provided. You can use the link for sharing your images.

Tip: Use NCBI numeric taxids as leaf names (or in the format TaxID.sequenceName) to get on-the-fly translation of species names and lineages.

Paste your tree in newick format:

```
(
(
AF443116.1:0.39975
,
AF449632.1:0.39975
```

Clear

Or upload a newick file:

Seleccionar archivo Ningún archivo seleccionado

Paste your alignment in fasta format:

```
TTCTAATTTAAACTATTCTCTGTTCTTTCAT
GGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCA
CCCATCAACAACCGCTATGTATCTCGTACATT
ACTGCCAGCCACCATGAATATTGTACGGTACCATAAATA
CTTGACCACCTGTAGTACATAAAAACCCAAT
```

Clear

Or upload a fasta file:

Seleccionar archivo Ningún archivo seleccionado

View tree! Alignment image type: Condensed format ▾ - ☑ Resolve taxonomic ids

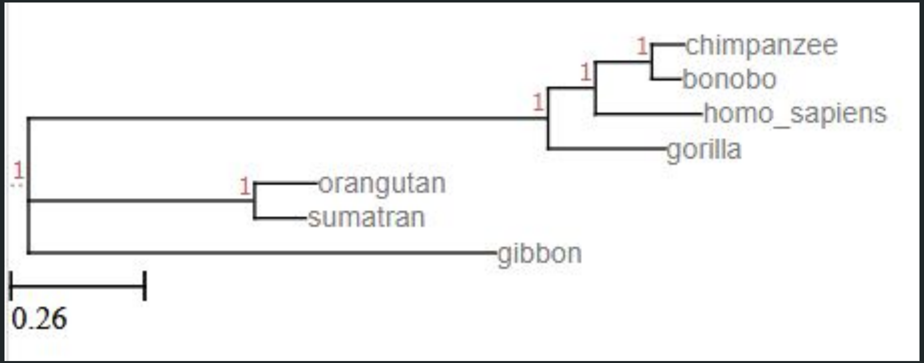
Share (beta): http://etetoolkit.org/treeview/?treeid=2797015fd210e7025a42658a96f7460b&algid=cc63eff25565074e61475343a6515c14"

Support values are shown in red.
(Loading big trees/algs may take a few seconds)

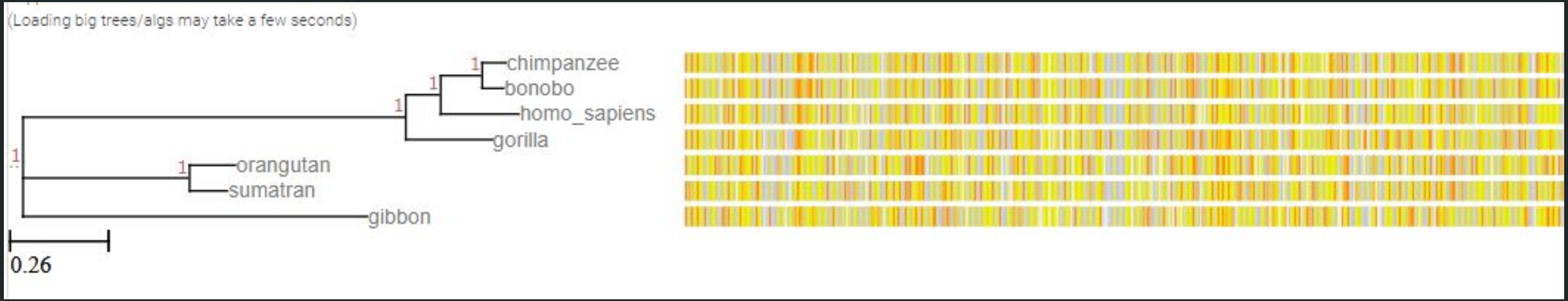
Elegir tipo de imagen y resolver ids taxonómicos

ETE Toolkit

Do not display alignment

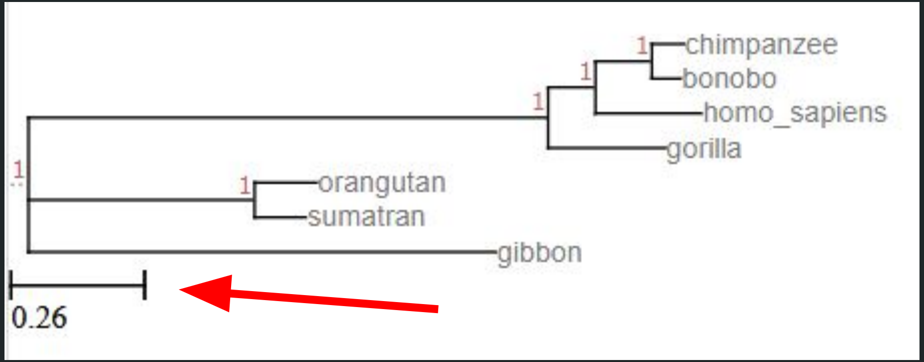


Condensed format

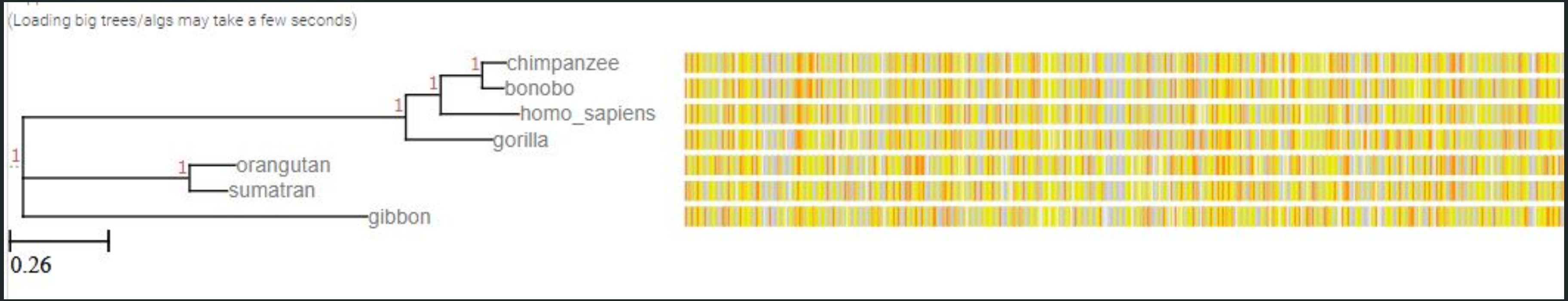

(Loading big trees/algs may take a few seconds)

ETE Toolkit

Do not display alignment



Condensed format



ETE Toolkit

aligned blocks