

# Phylogenetics and Molecular Evolution/Filogenética e Evolução Molecular

Octávio S. Paulo

Computational Biology and Population Genomics Group (CoBiG2)

Faculdade de Ciências, Universidade de Lisboa

[octavio.paulo@fc.ul.pt](mailto:octavio.paulo@fc.ul.pt) gabinete 2.3.46A

TP Bayesian Inference



**Ciências  
ULisboa**

Faculdade  
de Ciências  
da Universidade  
de Lisboa



**Computational  
Biology & Population  
Genomics Group**



# Aula TP MrBayes

---

```
mb.3.2.7-win64 - Short × + ▾  
MrBayes 3.2.7a x86_64  
(Bayesian Analysis of Phylogeny)  
Distributed under the GNU General Public License  
  
Type "help" or "help <command>" for information  
on the commands that are available.  
  
Type "about" for authorship and general  
information about the program.  
  
MrBayes > |
```

# Aula TP MrBayes

---

Prepare nexus files (goto end)

## 2.1 Quick Start Version

There are four steps to a typical Bayesian phylogenetic analysis using MrBayes:

1. Read the Nexus data file
2. Set the evolutionary model
3. Run the analysis
4. Summarize the samples

In more detail, each of these steps is performed as described in the following paragraphs:

# 1- read the nexus file

```
mb.3.2.7-win64 - Short × + ▾  
  
MrBayes 3.2.7a x86_64  
  
(Bayesian Analysis of Phylogeny)  
  
Distributed under the GNU General Public License  
  
Type "help" or "help <command>" for information  
on the commands that are available.  
  
Type "about" for authorship and general  
information about the program.  
  
MrBayes > exec C:\Users\octav\OneDrive\Desktop\FEM_Exercises_2023\Bayes\data\MATLLECB57MBAICcodonpartition.nex
```

# Aula TP MrBayes

---

At the MrBayes > prompt, type execute XXXXX.nex

At the MrBayes > prompt, type lset nst=6 rates=invgamma

At the MrBayes > prompt, type mcmc ngen=20000 samplefreq=100  
printfreq=100 diagnfreq=1000.

If the standard deviation of split frequencies is below 0.01 after 20,000 generations, stop the run by answering no when the program asks Continue the analysis? (yes/no).

>sump

>sumt

# Aula TP MrBayes Iset

---

The lset settings table at the end should look like this:

> help lset

Parameter	Options	Current Setting
Nucmodel	4by4/Doublet/Codon/Protein	4by4
Nst	1/2/6/Mixed	1
Code	Universal/Vertmt/Invermt/Yeast/Mycoplasma/ Ciliate/Echinoderm/Euplotid/Metmt	Universal
Ploidy	Haploid/Diploid/Zlinked	Diploid
Rates	Equal/Gamma/LNorm/Propinv/ Invgamma/Adgamma/Kmixture	Equal
Ngammacat	<number>	4
Nlnormcat	<number>	4
Nmixtcat	<number>	4
Nbetacat	<number>	5
Omegavar	Equal/Ny98/M3	Equal
Covarion	No/Yes	No
Coding	All/Variable/Informative/Nosingletons Noabsencesites/Nopresencesites/ Nosingletonabsence/Nosingletonpresence	All
Parsmodel	No/Yes	No

# Aula TP MrBayes prset

---

> help prset

Model settings for partition 1:

Parameter	Options	Current Setting
Tratioopr	Beta/Fixed	Beta(1.0,1.0)
Revmatpr	Dirichlet/Fixed	Dirichlet(1.0,1.0,1.0,1.0,1.0,1.0)
Aamodelpr	Fixed/Mixed	Fixed(Poisson)
Aarevmatpr	Dirichlet/Fixed	Dirichlet(1.0,1.0,...)
Omegapr	Dirichlet/Fixed	Dirichlet(1.0,1.0)
Ny98omega1pr	Beta/Fixed	Beta(1.0,1.0)
Ny98omega3pr	Uniform/Exponential/Fixed	Exponential(1.0)
M3omegapr	Exponential/Fixed	Exponential
Codoncatfreqs	Dirichlet/Fixed	Dirichlet(1.0,1.0,1.0)
Statefreqpr	Dirichlet/Fixed	Dirichlet(1.0,1.0,1.0,1.0)
Shapepr	Uniform/Exponential/Fixed	Exponential(1.0)
Ratecorrpr	Uniform/Fixed	Uniform(-1.0,1.0)
Pinvarpr	Uniform/Fixed	Uniform(0.0,1.0)
Covswitchpr	Uniform/Exponential/Fixed	Uniform(0.0,100.0)
Symdirihyperpr	Uniform/Exponential/Fixed	Fixed(Infinity)



# Aula TP MrBayes prset

---

Topologypr	Uniform/Constraints/Fixed/ Speciestree	Uniform
Brlenspr	Unconstrained/Clock/Fixed	Unconstrained:GammaDir(1.0,0.100,1.0
Treeagepr	Gamma/Uniform/Fixed/ Truncatednormal/Lognormal/ Offsetlognormal/Offsetgamma/ Offsetexponential	Gamma(1.00,1.00)
Speciationpr	Uniform/Exponential/Fixed	Exponential(10.0)
Extinctionpr	Beta/Fixed	Beta(1.0,1.0)
Fossilizationpr	Beta/Fixed	Beta(1.0,1.0)
SampleStrat	Random/Diversity/Cluster/ FossilTip	Random
Sampleprob	<number>	1.00000000
Popsizepr	Lognormal/Gamma/Uniform/ Normal/Fixed	Gamma(1.0,10.0)
Popvarpr	Equal/Variable	Equal
Nodeagepr	Unconstrained/Calibrated	Unconstrained
Clockratepr	Fixed/Normal/Lognormal/ Exponential/Gamma	Fixed(1.00)

# Aula TP MrBayes prset

---

Clockvarpr	Strict/Cpp/TK02/Igr/Mixed	Strict
Cppratepr	Fixed/Exponential	Exponential(0.10)
Cppmultdevpr	Fixed	Fixed(0.40)
TK02varpr	Fixed/Exponential/Uniform	Exponential(1.00)
Igrvarpr	Fixed/Exponential/Uniform	Exponential(10.00)
Ratepr	Fixed/Variable=Dirichlet	Fixed
Generatepr	Fixed/Variable=Dirichlet	Fixed

# Aula TP MrBayes prset

---

We need to focus on `Revmatpr` (for the six substitution rates of the GTR rate matrix), `Statefreqpr` (for the stationary nucleotide frequencies of the GTR rate matrix), `Shapepr` (for the shape parameter of the gamma distribution of rate variation), `Pinvarpr` (for the proportion of invariable sites), `Topologypr` (for the topology), and `Br1enspr` (for the branch lengths).

# Aula TP MrBayes prset

---

At the MrBayes > prompt, type showmodel

Model settings:

Data not partitioned --

Datatype = DNA

Nucmodel = 4by4

Nst = 6

Substitution rates, expressed as proportions  
of the rate sum, have a Dirichlet prior  
(1.00,1.00,1.00,1.00,1.00,1.00)

Covarion = No

# States = 4

State frequencies have a Dirichlet prior  
(1.00,1.00,1.00,1.00)

# Aula TP MrBayes prset

---

Rates = Invgamma  
The distribution is approximated using 4 categories.  
Shape parameter is exponentially distributed with parameter (1.00).  
Proportion of invariable sites is uniformly distributed on the interval (0.00,1.00).

Active parameters:

Parameters	
Revmat	1
Statefreq	2
Shape	3
Pinvar	4
Ratemultiplier	5
Topology	6
Brlens	7

# Aula TP MrBayes prset

---

1 -- Parameter = Revmat  
Type = Rates of reversible rate matrix  
Prior = Dirichlet(1.00,1.00,1.00,1.00,1.00,1.00)

2 -- Parameter = Pi  
Type = Stationary state frequencies  
Prior = Dirichlet

3 -- Parameter = Alpha  
Type = Shape of scaled gamma distribution of site rates  
Prior = Exponential(1.00)

4 -- Parameter = Pinvar  
Type = Proportion of invariable sites  
Prior = Uniform(0.00,1.00)

5 -- Parameter = Ratemultiplier  
Type = Partition-specific rate multiplier  
Prior = Fixed(1.0)

# Aula TP MrBayes prset

---

```
6 -- Parameter = Tau
   Type        = Topology
   Prior       = All topologies equally probable a priori

   Subparam.   = V

7 -- Parameter = V
   Type        = Branch lengths
   Prior       = Unconstrained:GammaDir(1.0,0.1000,1.0,1.0)
```

# Aula TP MrBayes Iset

---

Iset applyto=() nst=6 # SYM  
prset applyto=() statefreqpr=fixed(equal)

Iset applyto=() nst=6 rates=propinv # SYM + I  
prset applyto=() statefreqpr=fixed(equal)

Iset applyto=() nst=6 rates=gamma # SYM + gamma  
prset applyto=() statefreqpr=fixed(equal)

Iset applyto=() nst=6 rates=invgamma # SYM + I + gamma  
prset applyto=() statefreqpr=fixed(equal)



# Aula TP MrBayes Iset

---

lset applyto=() nst=2 # K2P  
prset applyto=() statefreqpr=fixed(equal)

lset applyto=() nst=2 rates=propinv # K2P + I  
prset applyto=() statefreqpr=fixed(equal)

lset applyto=() nst=2 rates=gamma # K2P + gamma  
prset applyto=() statefreqpr=fixed(equal)

lset applyto=() nst=2 rates=invgamma # K2P + I + gamma  
prset applyto=() statefreqpr=fixed(equal)

# Aula TP MrBayes Iset

---

```
lset applyto=() nst=2 # HKY
lset applyto=() nst=2 rates=propinv # HKY + I
lset applyto=() nst=2 rates=gamma # HKY + gamma
lset applyto=() nst=2 rates=invgamma # HKY + I + gamma
```

# Aula TP MrBayes Iset

---

```
lset applyto=() nst=1 # F81
lset applyto=() nst=1 rates=propinv # F81 + I
lset applyto=() nst=1 rates=gamma # F81 + gamma
lset applyto=() nst=1 rates=invgamma # F81 + I + gamma
```

# Aula TP MrBayes Iset

---

Iset applyto=() nst=1 # JC  
prset applyto=() statefreqpr=fixed(equal)

Iset applyto=() nst=1 rates=propinv # JC + I  
prset applyto=() statefreqpr=fixed(equal)

Iset applyto=() nst=1 rates=gamma # JC + gamma  
prset applyto=() statefreqpr=fixed(equal)

Iset applyto=() nst=1 rates=incgamma # JC + I + gamma  
prset applyto=() statefreqpr=fixed(equal)

# Aula TP MrBayes mcmc

---

At the MrBayes > prompt, type help mcmc

Parameter	Options	Current Setting
Ngen	<number>	1000000
Nruns	<number>	2
Nchains	<number>	4
Temp	<number>	0.100000
Reweight	<number>, <number>	0.00 v 0.00 ^
Swapfreq	<number>	1
Nswaps	<number>	1
Samplefreq	<number>	500
Printfreq	<number>	1000
Printall	Yes/No	Yes
Printmax	<number>	8
Mcmcdiag	Yes/No	Yes
Diagnfreq	<number>	5000
Diagnstat	Avgstddev/Maxstddev	Avgstddev
Minpartfreq	<number>	0.10
Allchains	Yes/No	No
Allcomps	Yes/No	No
Relburnin	Yes/No	Yes
Burnin	<number>	0
Burninfrac	<number>	0.25

# Aula TP MrBayes mcmc

---

At the MrBayes > prompt, type help mcmc

Stoprule	Yes/No	No
Stopval	<number>	0.05
Savetrees	Yes/No	No
Checkpoint	Yes/No	Yes
Checkfreq	<number>	2000
Filename	<name>	primates.nex.<p/t>
Startparams	Current/Reset	Current
Starttree	Current/Random/ Parsimony	Current
Nperts	<number>	0
Data	Yes/No	Yes
Ordertaxa	Yes/No	No
Append	Yes/No	No
Autotune	Yes/No	Yes
Tunefreq	<number>	100

# Aula FigTree and Tracer

---

- 📄 AULAS1.ckp
- 📄 AULAS1.ckp~
- 📄 AULAS1.con
- 📄 AULAS1.mcmc
- 📄 AULAS1.parts
- 📄 AULAS1.run1.p
- 📄 AULAS1.run1
- 📄 AULAS1.run2.p
- 📄 AULAS1.run2
- 📄 AULAS1.trprobs
- 📄 AULAS1.tstat
- 📄 AULAS1.vstat



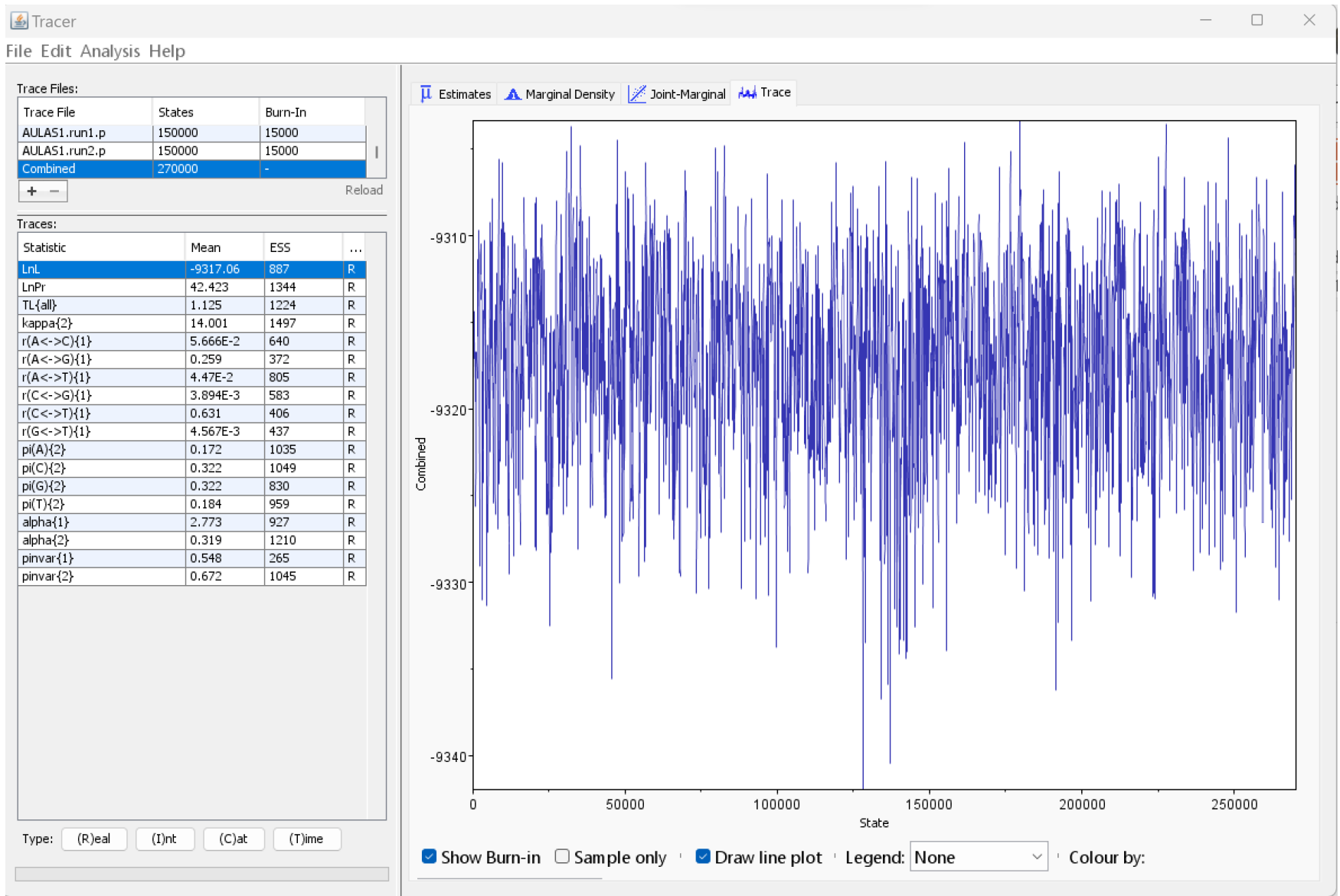
FigTree



Tracer



# Aula Tracer





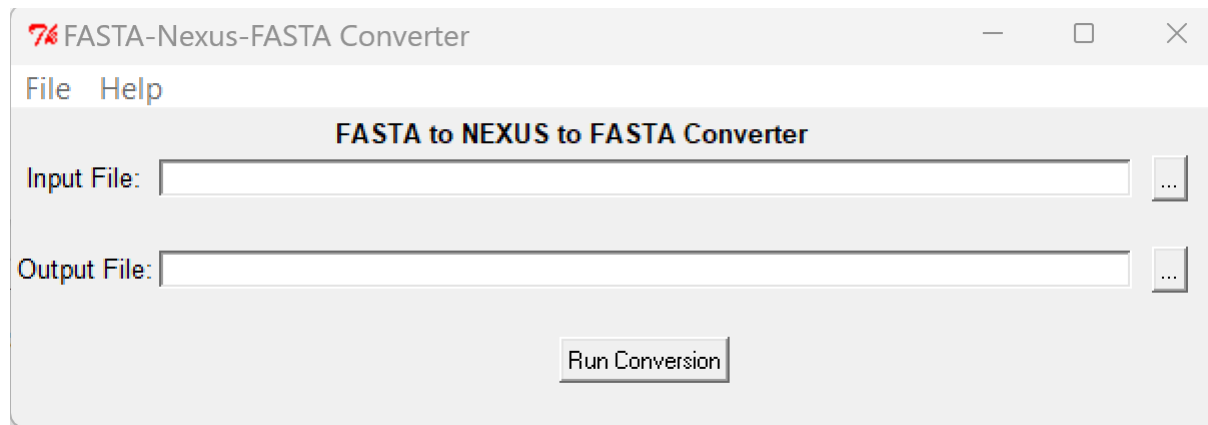
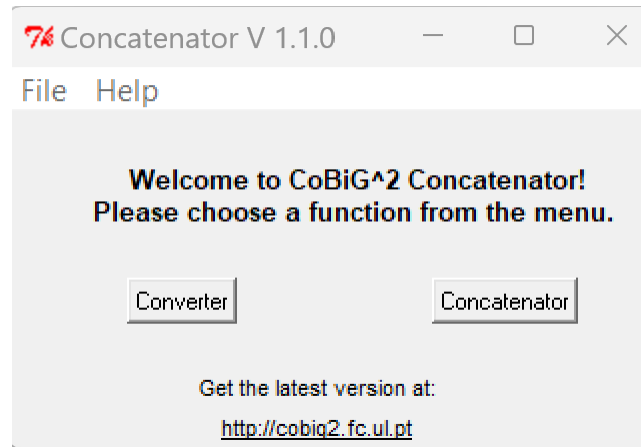
# Aula TP MrBayes

---

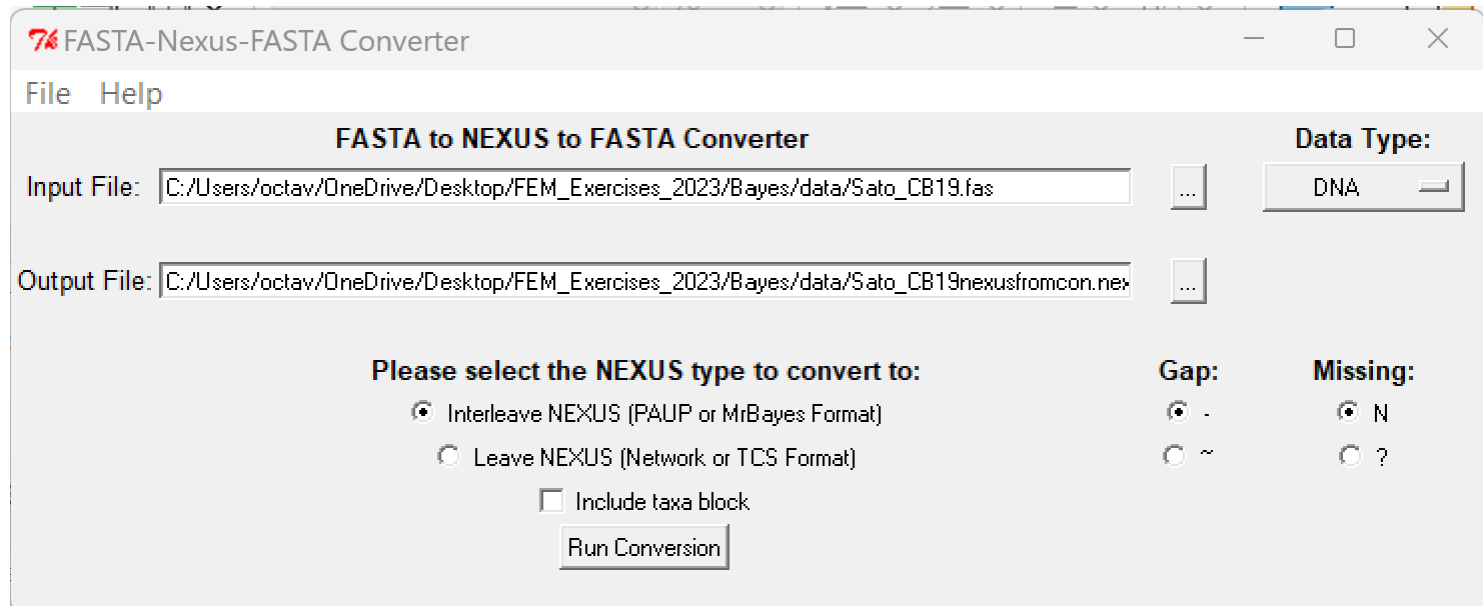
Prepare nexus files:

# Concatenator

---



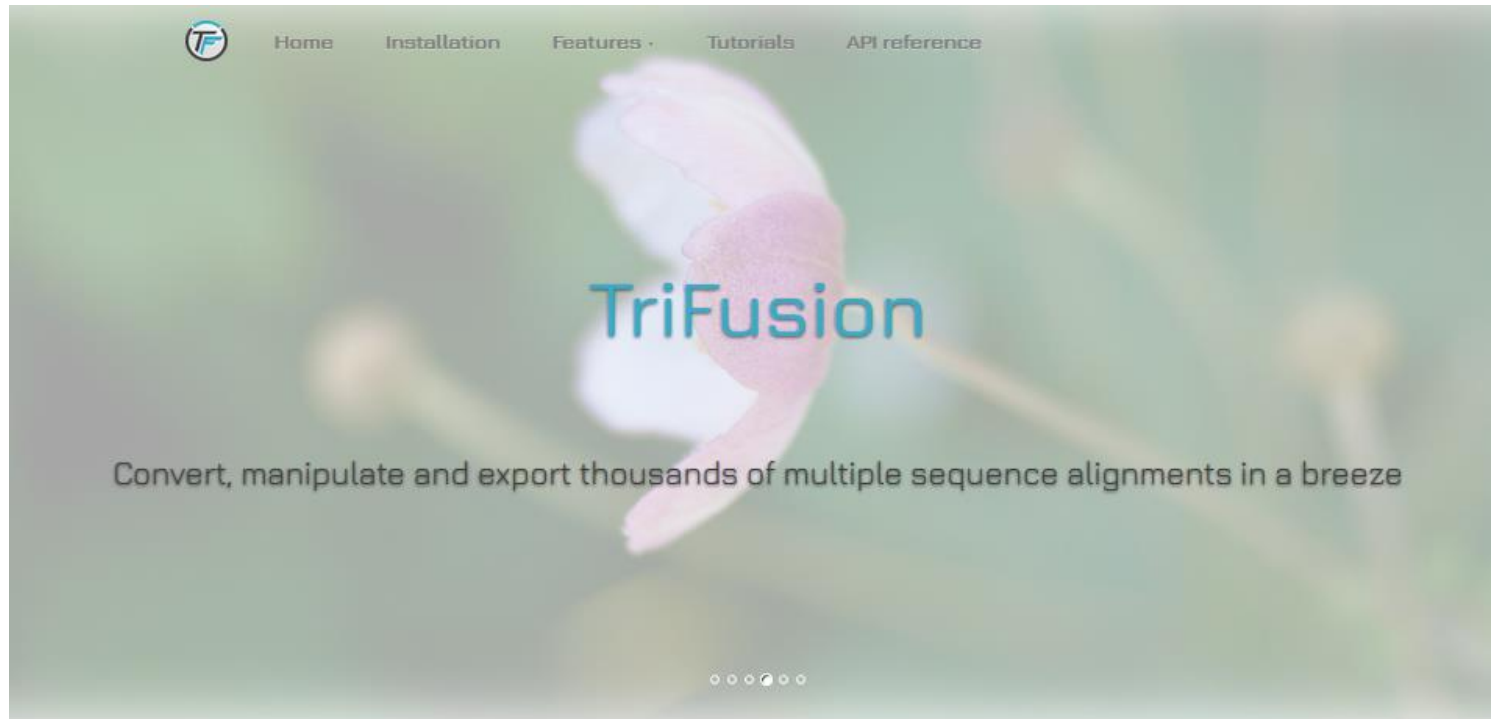
# Concatenator/Converter Fasta-Nexus



# Concatenator/Concatenate (to 5 genes)

The screenshot shows the 'Matrix Concatenator' application window. The title bar reads '74 Matrix Concatenator'. The menu bar contains 'File' and 'Help'. The main area has a title 'Matrix Concatenator' and instructions: 'Inputs must be on NEXUS format. Up to a maximum of five matrices can be joined.' On the right, there is a 'Data Type:' dropdown menu currently set to 'DNA'. Below this, there are two rows for input matrices. 'Matrix 1:' has a text field with the path 'C:/Users/octav/OneDrive/Desktop/FEM\_Exercises\_2023/Bayes/data/Sato\_CB19fromCon.nex' and a browse button '...'. 'Matrix 2:' has a text field with the path 'C:/Users/octav/OneDrive/Desktop/FEM\_Exercises\_2023/Bayes/data/Sato\_IRBP19fromCon.nex' and a browse button '...'. Below these is a button labeled 'Add one more matrix'. The 'Output Matrix:' field contains the path 'C:/Users/octav/OneDrive/Desktop/FEM\_Exercises\_2023/Bayes/data/Sato\_CB\_IRBP19\_nexusfrom' and a browse button '...'. At the bottom, there are three sections: 'Concatenation type:' with radio buttons for 'PAUP ready concatenated matrix' and 'MrBayes ready concatenated matrix' (the latter is selected); 'Gap character:' with radio buttons for '-' (selected) and '~'; and '"Missing" character:' with radio buttons for 'N' (selected) and '?'. There are also two checkboxes: 'Include taxa block' and 'Include "exclude" command for ILD partition test'. A large 'Concatenate!' button is centered at the bottom.

# Trifusion: <https://odiogosilva.github.io/TriFusion/>



## Get Started on TriFusion



### Download

Installers are available for Windows, MacOS and several Unix distro's. PPAs are available Debian and Ubuntu and AUR packages for ArchLinux.

[Get TriFusion](#)



### User Guide

Get the latest documentation on the full features of TriFusion, how to operate GUI and command line versions of TriFusion, and more!

[Download guide](#)



### View on GitHub

TriFusion is and will always be free software (under GPLv3 license). Feel free to view and modify the source code, hosted on GitHub.

[View Source](#)



# Welcome to TriFusion

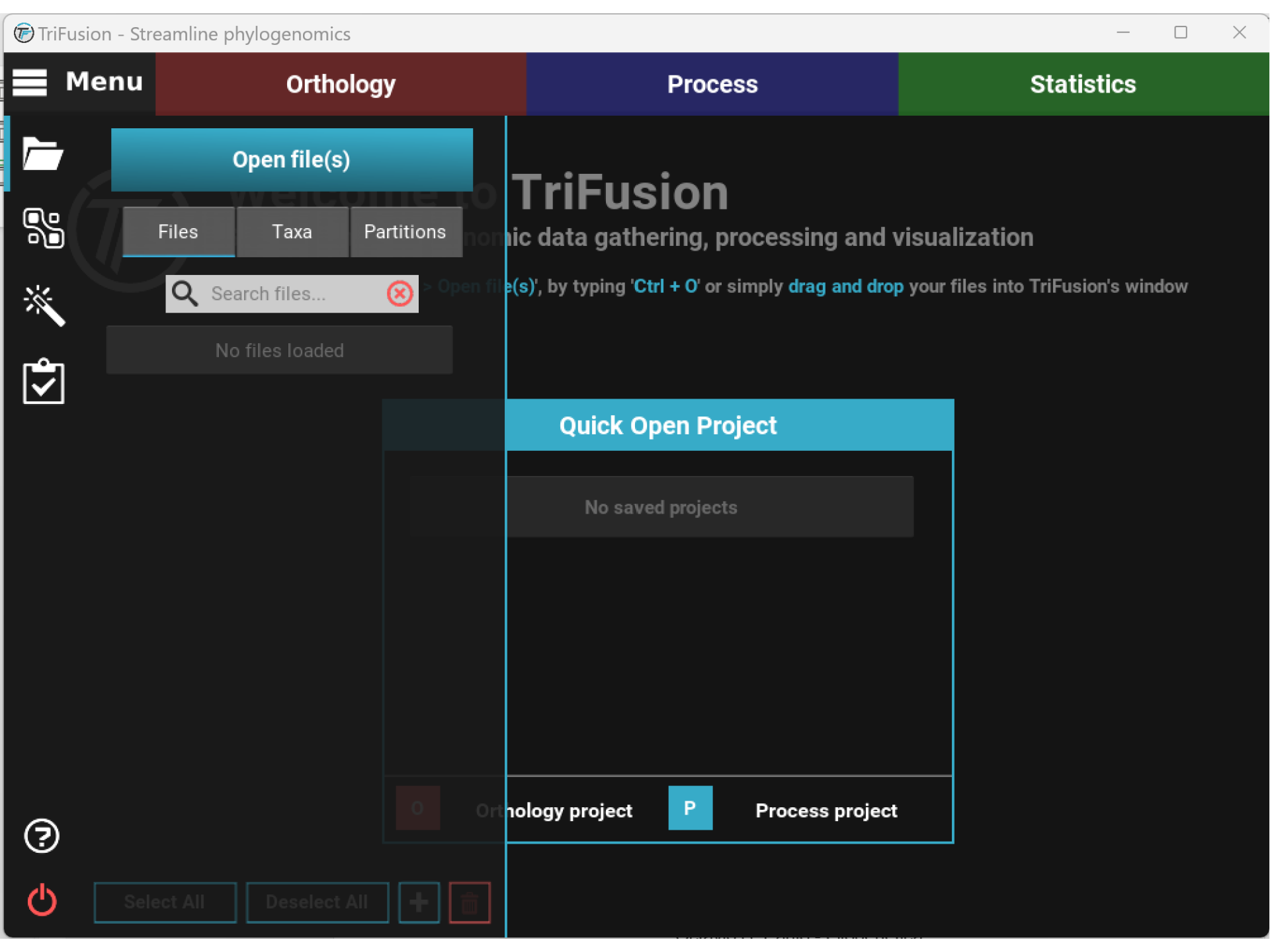
Streamlining phylogenomic data gathering, processing and visualization

Start by clicking 'Menu > Open file(s)', by typing 'Ctrl + O' or simply **drag and drop** your files into TriFusion's window

**Quick Open Project**

No saved projects

**O** Orthology project    **P** Process project



Open file(s)

Files Taxa Partitions

Search files... [x]

No files loaded

# TriFusion

Phylogenomic data gathering, processing and visualization

> Open file(s), by typing 'Ctrl + O' or simply drag and drop your files into TriFusion's window

Quick Open Project

No saved projects



Orthology project



Process project



Select All

Deselect All



Menu

Orthology

Process

Statistics

My Computer

C:

G:

octav

Bookmarks



Input data type

Alignment/Sequence set

Proteome



C:\\Users\\octav\\Dropbox\\AArea\_In\_Out\\1\_temp



Clear Selection

Name

..\\

Camila\_ERGA\_Nature\_2023.pdf

Forestry\_2022\_Nature Ecology and Evolution.pdf

Jupiter.pdf

Mike.docx

Mike\_orituary\_2023.pdf

Nei\_Orituary\_2023.pdf

Sato\_CB19.fas

Sato\_IRBP19.FAS

Video\_teclado1.mp4

video\_teclado2.mp4

Find:

Aa

All

Load selection

Load &amp; go back

Cancel



Choose your destiny

**Conversion**

**Concatenation**

**Execute**

**View Queue**

**Clear Options**

 **Menu****Orthology****Process****Statistics**

Choose your destiny

**Conversion****Concatenation**

General

**Data set**

Select the active taxa data set to perform operations

**Active files****Active taxa****Output Format**

Choose the format of the output file(s). Multiple formats may be chosen

**Fasta****Output directory**

Save output file(s) to the selected directory.

**Select...****Show additional options****Execute****View Queue****Clear Options**

Choose your destiny

**Conversion****Concatenation**

Revert a concatenated file?



OFF

General

**Data set**

Select the active taxa data set to perform operations



Active files

Active taxa

**Output Format**

Choose the format of the output file(s). Multiple formats may be chosen

Fasta

**Output file**

Save output file to the selected file.

Select...

Show additional options

**Execute****View Queue****Clear Options**