

Journées RNG de phylogénie moléculaire et analyse  
comparative

Un tour d'horizon de l'offre logicielle en phylogénie

Manolo Gouy

Biométrie et Biologie Evolutive - UMR CNRS 5558

Compilation exhaustive par Joe Felsenstein des programmes de phylogénie

<http://evolution.genetics.washington.edu/phylip/software.html>



« Here are some 244 of the phylogeny packages, and 28 free servers, that I know about. It is an attempt to be completely comprehensive. »

## Maximum likelihood and Bayesian methods

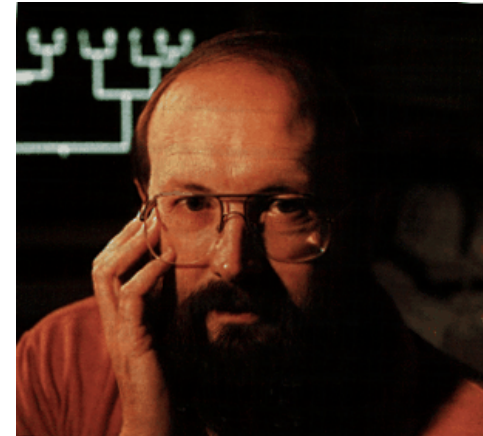
- [PHYLIP](#)
- [PAUP\\*](#)
- [fastDNAmI](#)
- [MOLPHY](#)
- [PAML](#)
- [Spectrum](#)
- [SplitsTree](#)
- [PLATO](#)
- [TREE-PUZZLE](#)
- [Hadtree, Prepare and Trees](#)
- [SeqPup](#)
- [Phylo\\_win](#)
- [PASSML](#)
- [ARB](#)
- [Darwin](#)
- [BAMBE](#)
- [DAMBE](#)
- [Modeltest](#)
- [TreeCons](#)
- [VeryfastDNAmI](#)
- [PAL](#)
- [dnarates](#)
- [TrExMI](#)
- [HY-PHY](#)
- [Vanilla](#)
- [Bionumerics](#)
- [fastDNAmIRev](#)
- [RevDNARates](#)
- [rate-evolution](#)
- [MrBayes](#)
- [Hadtree, Prepare and Trees](#)
- [CONSEL](#)
- [PAUPRat](#)
- [EDIBLE](#)
- [Mesquite](#)
- [PTP](#)
- [Treefinder](#)
- [MetaPIGA](#)
- [RAXML](#)
- [PHASE](#)
- [PHYML](#)
- [BEAST](#)
- [r8s-bootstrap](#)
- [MrBayes tree scanners](#)
- [MTgui](#)
- [MrModeltest](#)
- [BootPHYML](#)
- [p4](#)
- [Porn\\*](#)
- [SIMMAP](#)
- [Spectronet](#)
- [CIPRES](#)
- [Rhino](#)
- [IM](#)
- [Protest](#)
- [ModelGenerator](#)
- [Simplot](#)
- [MDIV](#)
- [MrAIC](#)
- [Modelfit](#)
- [IQPNNI](#)

- Les logiciels de reconstruction phylogénétique
- Les services web de reconstruction phylogénétique
- Logiciels intégrant banque de données et outils phylogénétiques
- Algorithmes et services web pour l'alignement multiple
- Logiciels de dessin d'arbres phylogénétiques

## PHYLIP : Phylogeny Inference Package

<http://evolution.genetics.washington.edu/phylip.html>

Joe Felsenstein,  
University of Washington at Seattle



Logiciel multiplateformes librement accessible, qui rassemble un vaste éventail de méthodes dont le code est public :

- Parcimonie
- Distances
- Maximum de vraisemblance (DNA/RNA et protéines)
- Bootstrap

Utilisation en mode ligne, par succession de commandes, qui nécessite une bonne connaissance de l'outil.

# PAUP: Phylogenetic Analysis Using Parsimony

<http://paup.csit.fsu.edu>

David Swofford,  
Florida State University

## PAUP\* Version 4

...tools for inferring  
and interpreting  
phylogenetic trees

Analyze

- Molecular sequences
- Morphological data
- Other data types

Using

- Maximum likelihood
- Parsimony
- Distance methods



Logiciel disponible sous MacOS avec interface graphique, et sous Windows et Unix en mode commande. Le détail des algorithmes est parfois difficilement accessible.

Coût des licences individuelles : 85 à 150 \$ selon la version

10 utilisateurs = 500 \$

# MEGA: Molecular Evolutionary Genetics Analysis

<http://www.megasoftware.net/>

Kumar, Tamura, Nei

Logiciel librement utilisable pour l'enseignement et la recherche.

Ecrit pour plateforme Windows.

Principaux algorithmes:

- Alignement par ClustalW
- Calcul de distances nucléotidiques et protéiques avec/sans variation du taux entre sites.
- Calcul de distances synonymes et non synonymes (Ka/Ks)
- Reconstruction d'arbre: NJ, ME, Parcimonie
- Tests statistiques: bootstrap, comparaison de taux entre lignées.



# **PHYML - A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.**

<http://atgc.lirmm.fr/phyml>

Stéphane Guindon, Olivier Gascuel  
LIRMM, Montpellier.

Algorithme rapide d'estimation phylogénétique au maximum de vraisemblance.

Catalogue de modèles évolutifs nucléotidiques et protéiques très complet.

Possibilité de tests de bootstrap.

Logiciel multiplateformes sans interface graphique.

Serveur web pour exécuter PHYML librement accessible.



# RaxML : A Fast Program for Maximum Likelihood-based Inference of Large Phylogenetic Trees

<http://www.ics.forth.gr/~stamatak/index.htm>

Alexandros Stamatakis

Institute of Computer Science, Heraklion, Grèce

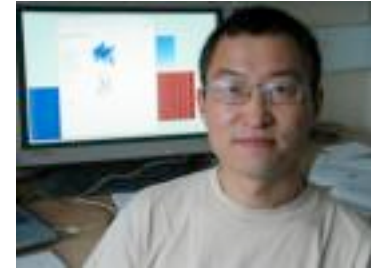
Logiciel librement accessible, pour toutes plateformes, en mode commande.

Une heuristique astucieuse d'exploration de l'espace des topologies conduit à un programme de maximum de vraisemblance très rapide.

# PAML: Phylogenetic Analysis by Maximum Likelihood

<http://abacus.gene.ucl.ac.uk/software/paml.html>

Ziheng Yang, University College London



PAML is a package of programs for phylogenetic analyses of DNA or protein sequences using maximum likelihood. It is maintained and distributed for academic use free of charge.

« PAML is not good for tree making. There are a few options for heuristic tree search, but they do not work well except for small data sets of only a few species.

PAML may be useful if you are interested in the process of sequence evolution. The two main programs, baseml and codeml, implement a number of sophisticated models, which you can use to construct likelihood ratio tests of evolutionary hypotheses. »

HYPHY is a free multiplatform software package to perform maximum likelihood analyses of genetic sequence data and to test various statistical hypotheses.

<http://www.hyphy.org>



Sergei Kosakovsky Pond, Spencer Muse & Simon Frost. UCSD - NC State University

« HYPHY was designed with maximum flexibility in mind and to that end it incorporates a simple high level programming language which enables the user to tailor the analyses precisely to his or her needs. These include relative rate and ratio tests, several methods of ML based phylogeny reconstruction, bootstrapping, model selection, positive selection, molecular clock tests and many more. »

## **TREE-PUZZLE: Maximum likelihood analysis for nucleotide, amino acid, and two-state data.**

<http://www.tree-puzzle.de/>



Heiko Schmidt, Korbinian Strimmer, Martin Vingron, and Arndt von Haeseler. Jülich, Munich, Düsseldorf

« TREE-PUZZLE reconstructs phylogenetic trees from molecular sequence data by maximum likelihood. It implements a fast tree search algorithm, quartet puzzling, that allows analysis of large data sets and automatically assigns estimations of support to each internal branch. TREEPUZZLE also computes pairwise maximum likelihood distances as well as branch lengths for user specified trees. »



# MrBayes: Bayesian Inference of Phylogeny

<http://mrbayes.csit.fsu.edu/index.php>



John Huelsenbeck, Fredrik Ronquist, Bret Larget,  
Paul van der Mark.

Logiciel librement accessible.

- A common command-line interface for Macintosh, Windows, and UNIX operating systems
- Ability to analyze nucleotide, amino acid, restriction site, and morphological data
- Mixing of data types, such as molecular and morphological characters, in a single analysis
- An abundance of evolutionary models, including 4 X 4, doublet, and codon models for nucleotide data and many of the standard rate matrices for amino acid data
- Estimation of positively selected sites in a fully hierarchical Bayes framework

- Les logiciels de reconstruction phylogénétique
- Les services web de reconstruction phylogénétique
- Logiciels intégrant banque de données et outils phylogénétiques
- Algorithmes et services web pour l'alignement multiple
- Logiciels de dessin d'arbres phylogénétiques



- Programmes pour les séquences [ [sequence.doc](#) ]

- ADN

- [dnadist](#) [ [utilisateurs expérimentés](#) ] [ [dnadist.doc](#) ]

- Distances calculées à partir des séquences d'ADN alignées.

- [dnapars](#) [ [utilisateurs expérimentés](#) ] [ [dnapars.doc](#) ]

- Méthode de parcimonie.

- dnaml

- dnaml a été supprimé ; utiliser plutôt : [fastDNAmI](#), équivalent et beaucoup plus rapide.*

- Protéines

- [protdist](#) [ [utilisateurs expérimentés](#) ] [ [protdist.doc](#) ]

- Distances calculées à partir des séquences de protéines alignées.

- [protpars](#) [ [utilisateurs expérimentés](#) ] [ [protpars.doc](#) ]

- Méthode de parcimonie.

- Programmes pour les matrices de distances [ [distance.doc](#) ]

- [neighbor](#) [ [utilisateurs expérimentés](#) ] [ [neighbor.doc](#) ]

- Neighbor-joining et UPGMA.

- [fitch](#) [ [utilisateurs expérimentés](#) ] [ [fitch.doc](#) ]

- Fitch-Margoliash et moindres carrés.

- [kitsch](#) [ [utilisateurs expérimentés](#) ] [ [kitsch.doc](#) ]

- Fitch-Margoliash et moindres carrés avec horloge moléculaire.



## PHYML Online execution

Sequences

Choisir le fichier aucun fichier sélectionné

File   Example file

Data Type

DNA   Amino-Acids

Sequence file

interleaved   sequential

Number of data sets

1  Perform bootstrap

Number of bootstrap data sets

Print bootstrap info

Substitution model

HKY

Transition / transversion ratio  
(DNA models)

4  fixed   estimated

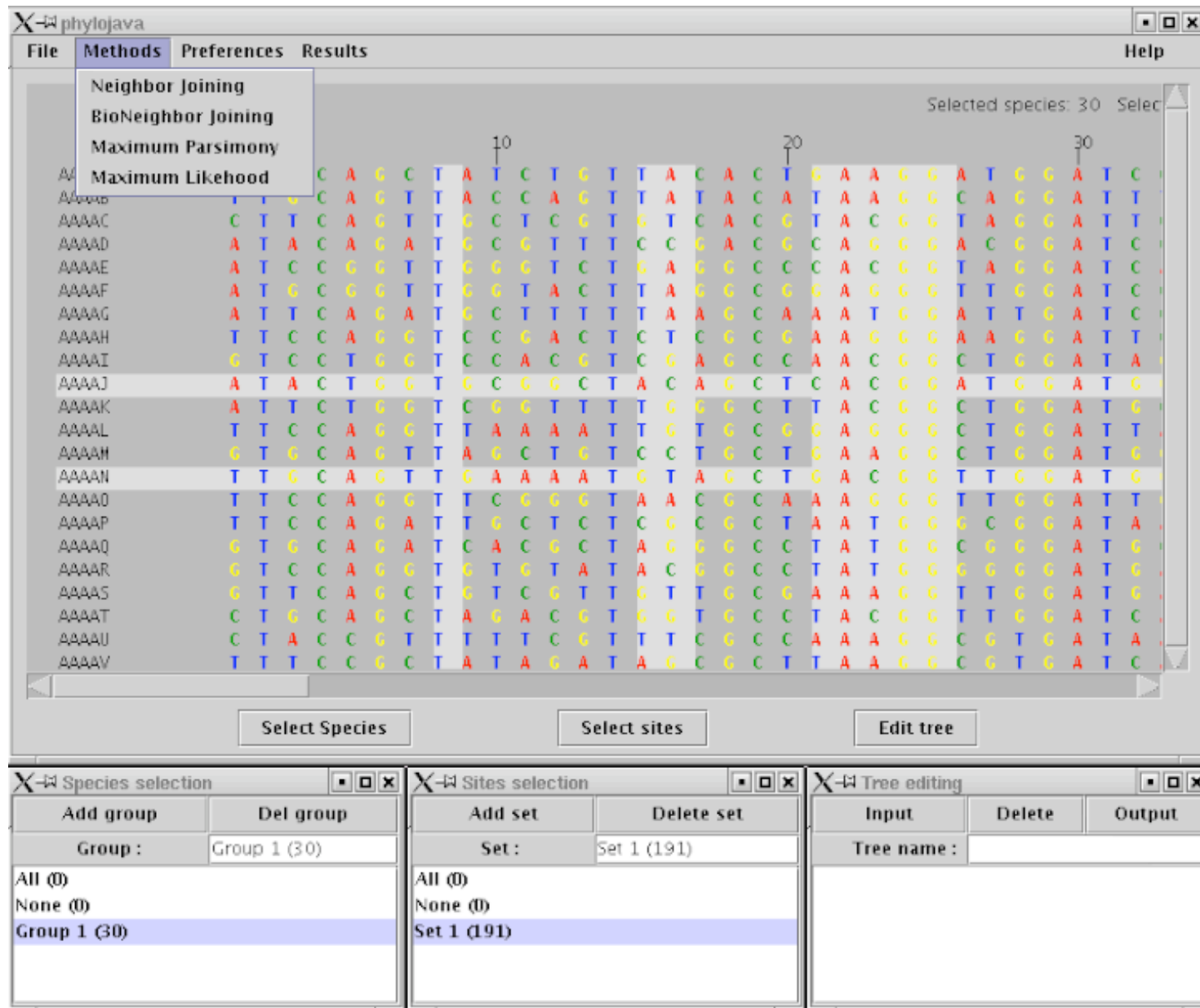
Proportion of invariable sites

0.0  fixed   estimated

Number of substitution  
rate categories

1

PhyloJava (en préparation au PBIL): interface graphique pour lancer des analyses phylogénétiques sur serveur(s) distant(s).



- Les logiciels de reconstruction phylogénétique
- Les services web de reconstruction phylogénétique
- **Logiciels intégrant banque de données et outils phylogénétiques**
- Algorithmes et services web pour l'alignement multiple
- Logiciels de dessin d'arbres phylogénétiques

ARB: A software environment for maintaining databases of molecular sequences, and for analyzing the sequence data, with emphasis on phylogeny reconstruction.

<http://www.arb-home.de/>

Technical University of Munich



Principaux modules d'ARB:

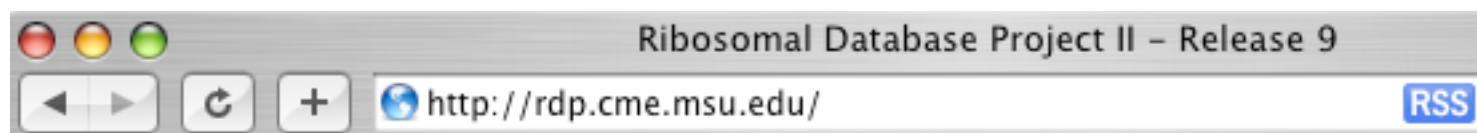
**ARB\_DB:** A central database of aligned homologous sequences taken from public databases or supplied by the user.

**ARB\_NT:** Phylogenetic trees derived from the data or imported from other sources. Different tree topologies, complete trees, and subtrees can be stored and used for walking through the database. Trees can be used to define subsets of data.

**ARB\_ALIGN:** This tool searches for the most similar sequences in the database and inserts the new sequence into an existing alignment.

**ARB\_PHYL:** tools and integrated foreign software (PHYLIP, DE SOETE, fastDNAmI) allow calculation of similarity/distance matrices, conservation profiles, selection masks and phylogenetic tree reconstruction.

# Ribosomal Database Project - II



**Hierarchy Browser** Browse a phylogenetic hierarchy and compile a list of 16S rRNA sequences for download or use.

---

**Classifier** Assign 16S rRNA sequences to our taxonomical hierarchy.

---

**Library Compare** Compare two sequence libraries using the RDP Classifier.

---

**Sequence Match** Upload your sequence and search for its nearest neighbors.

---

**Probe Match** See what your probe targets in our database.

---

BIBI:BioInformatic Bacterial Identification, a tool for bacterial identification

http://pbil.univ-lyon1.fr/bibi/

PBIL PubMed MBE Editor Google GoogleScholar ACNUC BBE AnnuaireUCB Helix CVS Antares WebmailUCB Gmail

**BIBI** Bio Informatic Bacterial Identification version 2  
BBE - UMR CNRS 5558 : Dynamique des populations bactériennes

**Equipe de Dynamique des populations bactériennes**

Unité Mixte de Recherche C.N.R.S. 5558  
[UMR CNRS 5558]

Faculté de Medecine de Lyon Sud BP12  
69921 OULLINS

**BIBI** Bio Informatic Bacterial Identification

« BIBI was developed to simplify sequence exploitation within a bacterial identification framework. This program combines similarity search tools in the sequence databases and phylogeny display programs. It implements a chaining of two well-known tools: BLAST and CLUSTAL W. »

Paste the FASTA nucleic sequence below [example] :

```
> query_sequence_rRNA_16S
CCGCAAGCTNGTGGAGCATGTGGTTTAATTCTGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCCGATG
CTATTTTC
TAGAGATAGGAAGTTTCTTCGGAACATCGGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATG
TTGGGTT
AAGTCCC GCAACGAGCGCAACCCTTATTGTTAGTTGCCATCATTGAGTTGGGCACTCTAGCGAGACTGCCGGT
```

WWWQuery

BLAST

HELP

Contact

### Options

#### General options

# sequences to align

DataBase

Sequence name

30

rrna

QUERY

#### BLAST parameter

Dropoff Value (X)

0

#### Alignment parameters

Alignment program

Mabios

Alignment accuracy

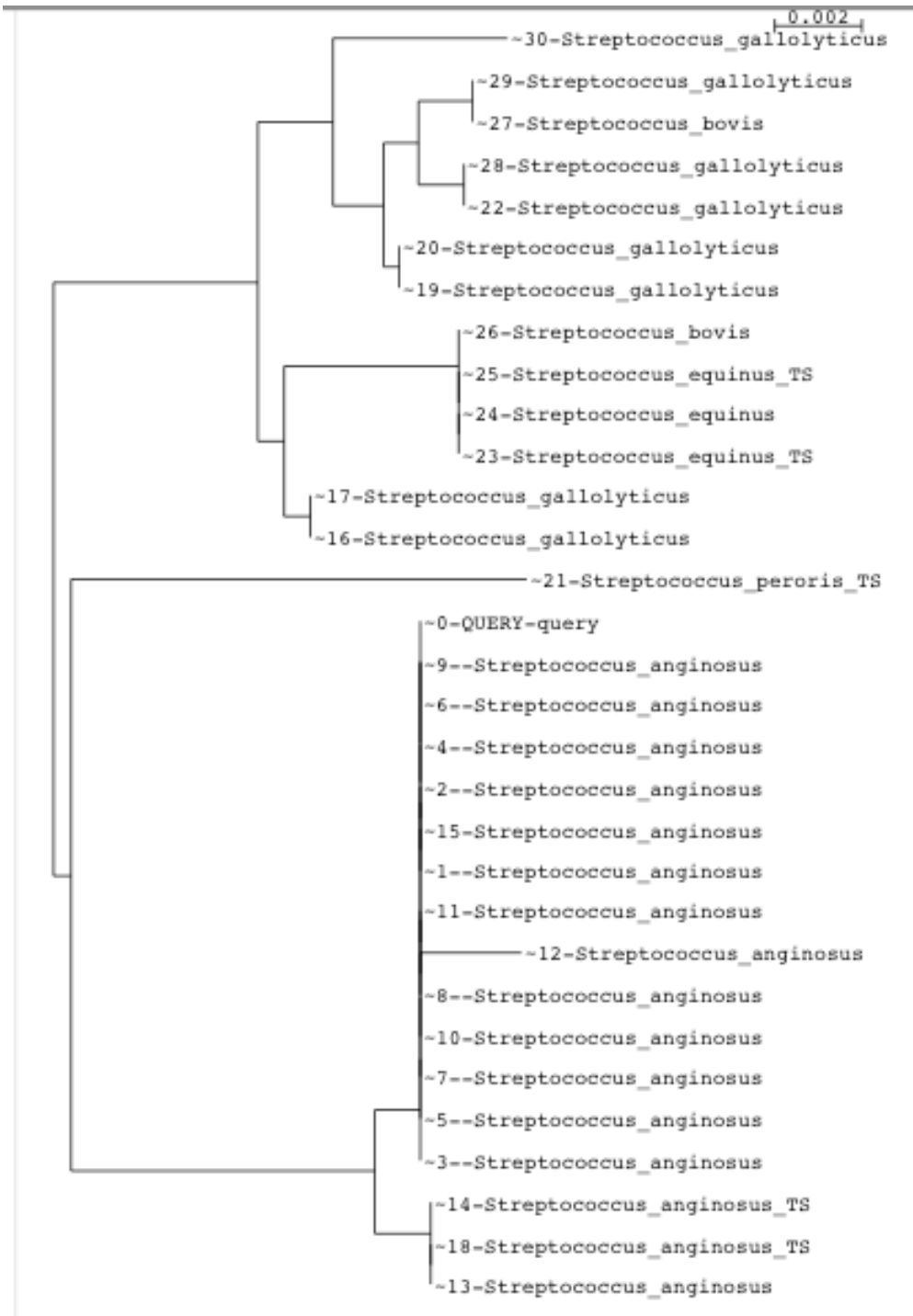
Approximate

#### Phylogenie parameters

Gaps option

Exclude positions with gaps





Positionnement de la séquence requête (QUERY) par rapport à ses 30 plus proches voisines phylogénétiques.

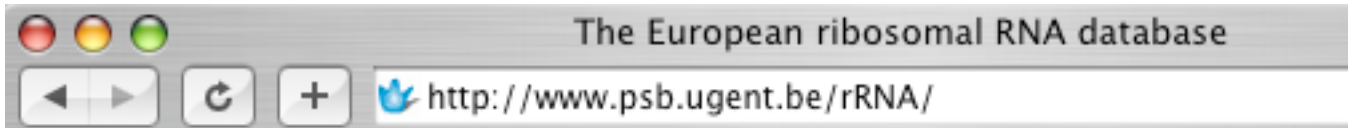




# *The European Ribosomal RNA database*

Department of Plant Systems Biology

University of Gent, BELGIUM



## **Ribosomal RNA quick phylogeny analysis**

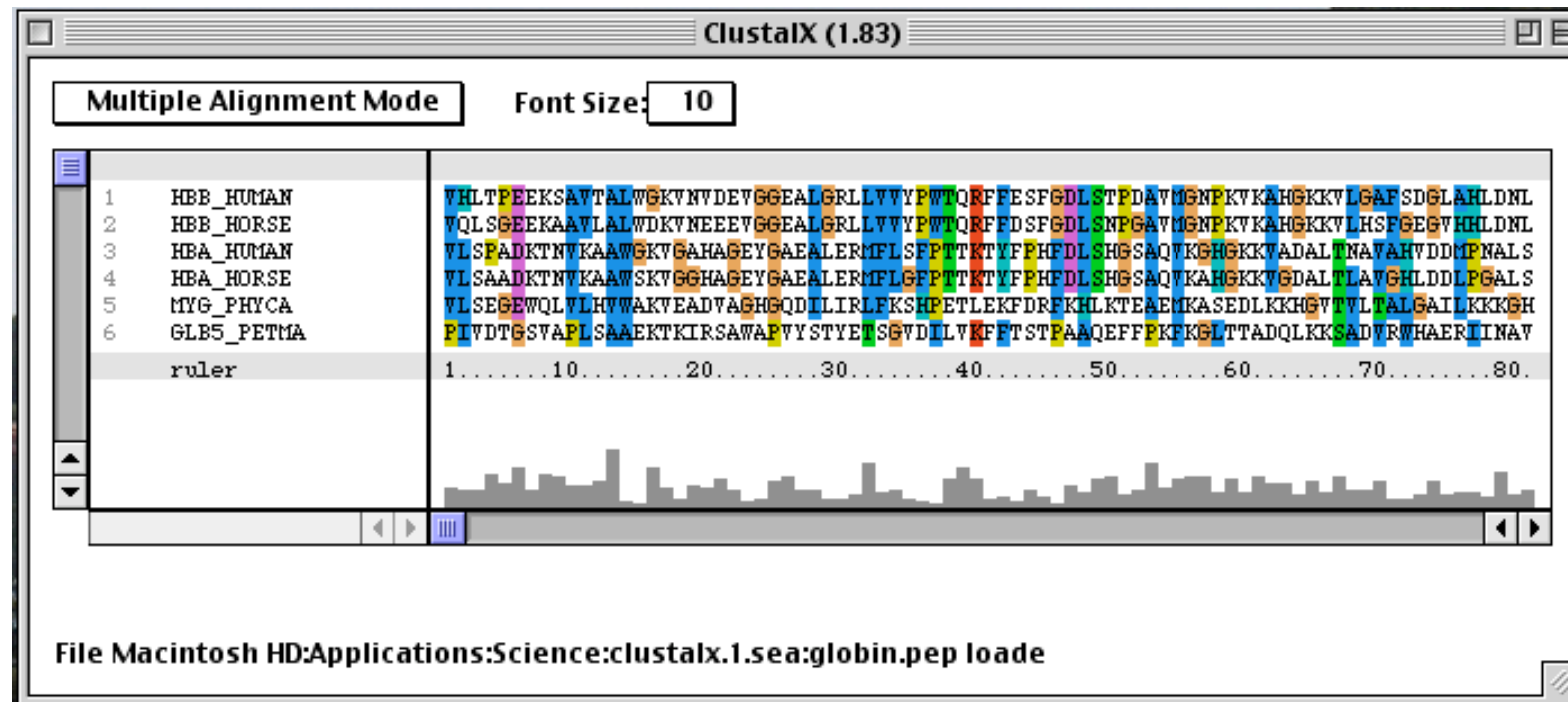
### **Quick phylogeny search**

This service is an implementation of the [bibi program](#) by [Guy Perrière](#) and Gregory Devulder. It uses the programmes BLAST (Altschul et al. 1997) and ClustalW (Thompson et al. 1994) to make a quick approximate identification of a sequence.

- Les logiciels de reconstruction phylogénétique
- Les services web de reconstruction phylogénétique
- Logiciels intégrant banque de données et outils phylogénétiques
- Algorithmes et services web pour l'alignement multiple
- Logiciels de dessin d'arbres phylogénétiques

CLUSTALW: alignement multiple de séquences nucléotidiques ou protéiques  
CLUSTALX: interface graphique de l'algorithme CLUSTALW.

Des Higgins, Julie Thompson, Toby Gibson. Dublin - Strasbourg - Heidelberg.



Logiciel multiplateformes librement accessible.

Une star de la bioinformatique: 16755 citations de l'article introductif de CLUSTALW; 6137 citations pour CLUSTALX.

## ClustalW Submission Form



Clustal W is a general purpose multiple sequence alignment program for DNA or proteins. It produces biologically meaningful multiple sequence alignments of divergent sequences. It calculates the best match for the selected sequences, and lines them up so that the identities, similarities and differences can be seen. Evolutionary relationships can be seen via viewing Cladograms or Phylograms. [New users, please read the FAQ.](#)

 [Download Software](#)

YOUR EMAIL	ALIGNMENT TITLE	RESULTS	ALIGNMENT	CPU MODE
<input type="text"/>	<input type="text" value="Sequenc"/>	<input type="text" value="interactive"/>	<input type="text" value="full"/>	<input type="text" value="single"/>
KTUP (WORD SIZE)	WINDOW LENGTH	SCORE TYPE	TOPDIAG	PAIRGAP
<input type="text" value="def"/>	<input type="text" value="def"/>	<input type="text" value="percent"/>	<input type="text" value="def"/>	<input type="text" value="def"/>
MATRIX	GAP OPEN	END GAPS	GAP EXTENSION	GAP DISTANCES
<input type="text" value="def"/>	<input type="text" value="def"/>	<input type="text" value="def"/>	<input type="text" value="def"/>	<input type="text" value="def"/>

OUTPUT		PHYLOGENETIC TREE		
OUTPUT FORMAT	OUTPUT ORDER	TREE TYPE	CORRECT DIST.	IGNORE GAPS
<input type="text" value="aln w/numbers"/>	<input type="text" value="aligned"/>	<input type="text" value="none"/>	<input type="text" value="off"/>	<input type="text" value="off"/>

Enter or Paste a set of Sequences in any supported format:

Upload a file:  aucun fichier sélectionné

# CLUSTALW

[[Abstract](#)] [[NPS@ help](#)] [[Original server](#)]

Paste a protein sequence databank in Pearson/Fasta format below :

All sequence names must be different !

Output width :

## CLUSTALW Parameters

Output format :

Output oder :

Pairwise alignment type :

Fast pairwise alignment parameters	Slow pairwise alignment parameters
K-tuple (word) size : <input type="button" value="1"/>	Protein weight matrix : <input type="button" value="GONNET"/> Gap openig penalty : <input type="text" value="10.0"/> Gap extension penalty : <input type="text" value="0.1"/>
Number of top diagonals : <input type="text" value="5"/>	
Window size : <input type="text" value="5"/>	
Gap penalty : <input type="text" value="3"/>	
Scoring method : <input type="button" value="Percentage"/>	

## Multiple Alignment Parameters :

Weight matrix :

Gap opening penalty :

Gap extension penalty :

Residue-specific gap penalties OFF :

# Pôle Bioinformatique Lyonnais

[BBE - UMR CNRS 5558](#) / [IBCP - UMR CNRS 5086](#)

[With help from CC-IN2P3](#)



<http://pbil.univ-lyon1.fr/>

Service web pour CLUSTALW  
au PBIL.

# MUSCLE

Protein multiple sequence alignment software



MUSCLE: MUltiple Sequence Comparison by Log-Expectation.  
Public domain multiple alignment software for protein and nucleotide sequences.

<http://www.drive5.com/muscle/index.htm>

R.C. Edgar

« MUSCLE achieves average accuracy statistically indistinguishable from T-Coffee and MAFFT, and is the fastest of the tested methods for large numbers of sequences, aligning 5000 sequences of average length 350 in 7 min on a current desktop computer. »

### MUSCLE WEB SERVER

Multiple sequence comparison by log-expectation. [More information.](#)

Paste sequences in [FASTA format](#):

(Please, no more than 200 sequences on our server.)

OR

[Upload FASTA file:](#)

aucun fichier sélectionné

[Email address](#)  (required)

Confirm email address  (required)

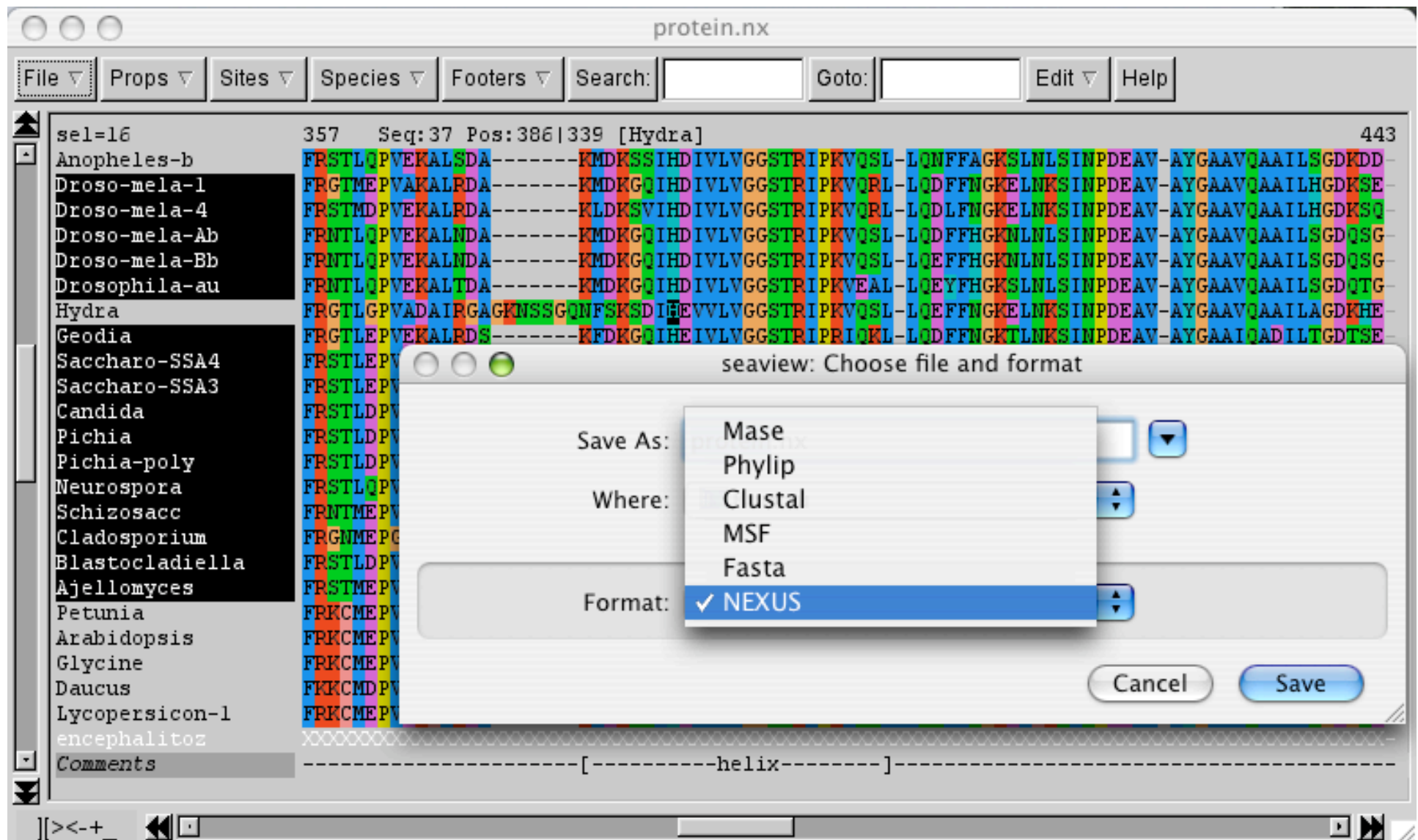
Email subject line



# SEAVIEW: éditeur d'alignement multiple de séquences

<http://pbil.univ-lyon1.fr/software/seaview.html>

Alignement par Clustalw, choix multiples de sites et de séquences, dot-plot, 6 formats de fichiers en lecture et écriture. Logiciel multiplateformes librement accessible.



- Les logiciels de reconstruction phylogénétique
- Les services web de reconstruction phylogénétique
- Logiciels intégrant banque de données et outils phylogénétiques
- Algorithmes et services web pour l'alignement multiple
- Logiciels de dessin d'arbres phylogénétiques





## TreeView

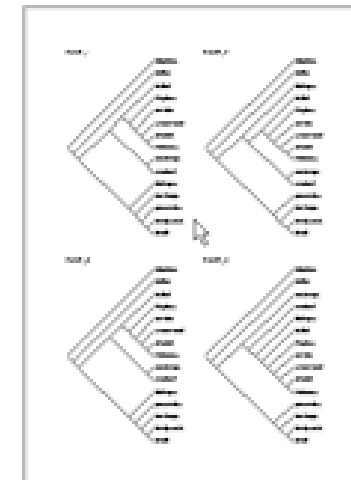
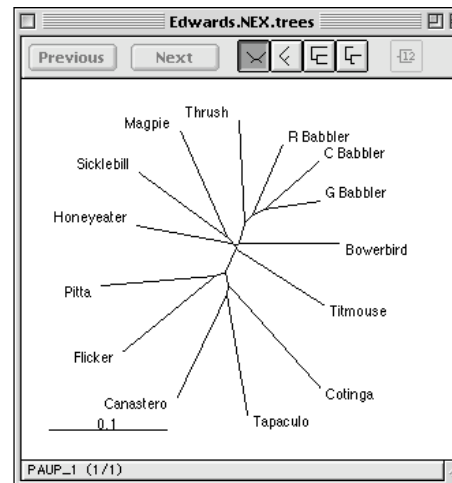
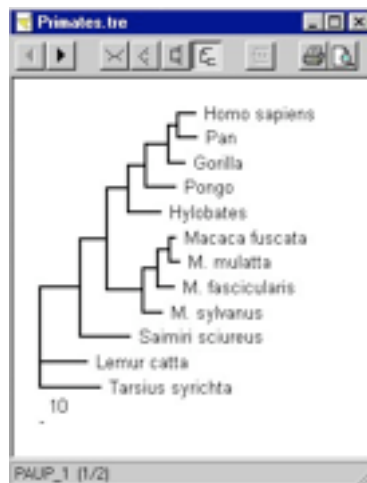
Tree drawing software for Apple Macintosh and Windows

 (and now [Linux and Unix](#))

<http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>

Rod Page,  
Glasgow University

Macintosh and Windows platforms. Reads various tree file formats (NEXUS, PHYLIP, Hennig86, NONA, MEGA, ClustalW/X). Supports PICT on MacOS, Windows metafile on Windows for copying pictures into other applications. Prints multiple trees per page, and one tree over more than one page.



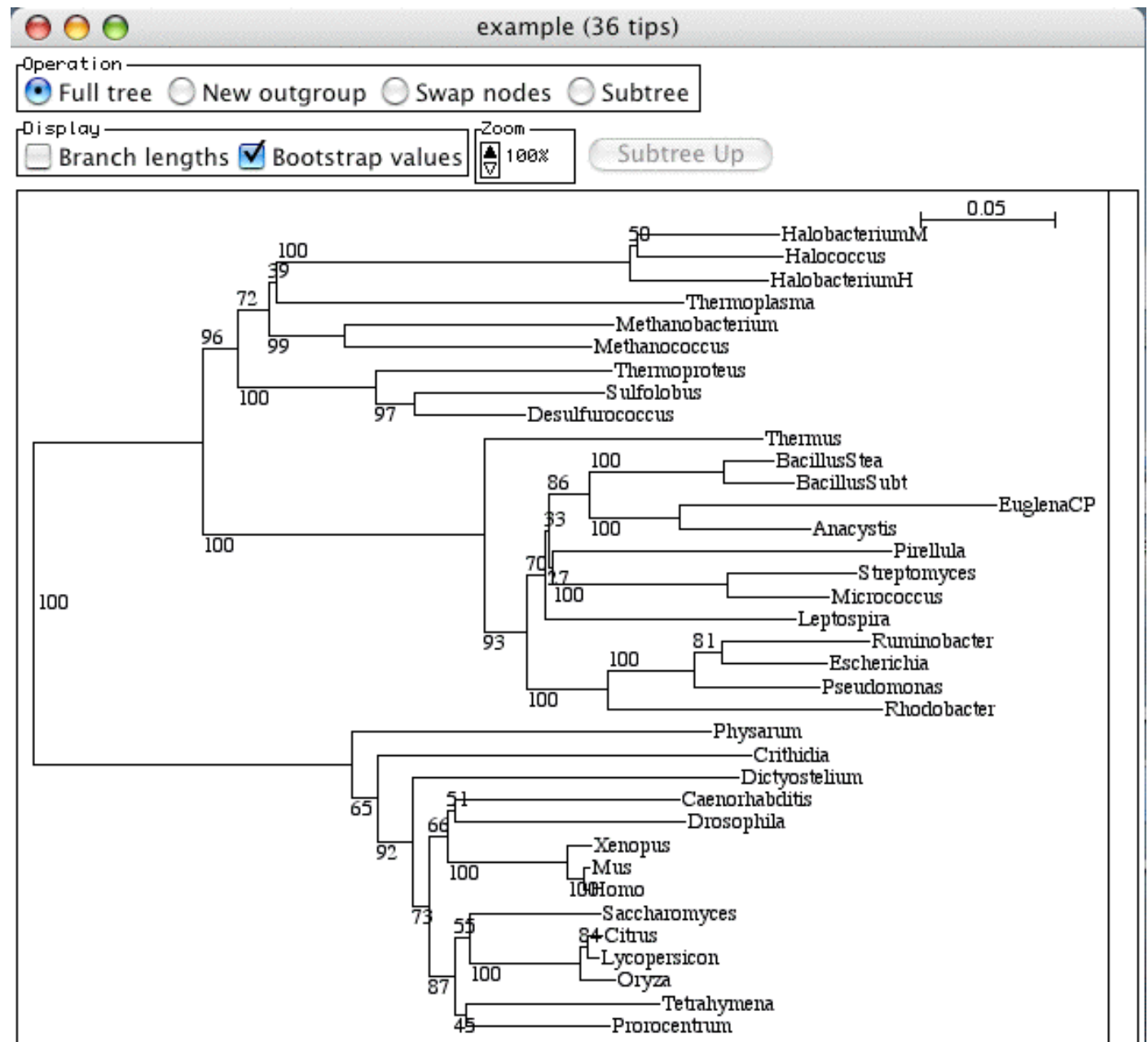
# NJPLOT: dessin d'arbres

<http://pbil.univ-lyon1.fr/software/njplot.html>

Logiciel libre  
multiplateformes.

Fonctions:

- enracinement,
- copier/coller,
- multifurcations,
- zoom,
- recherche cible,
- impression sur plusieurs pages.



**ATV (A Tree Viewer):** a Java tool for the visualization of annotated phylogenetic trees.

<http://www.genetics.wustl.edu/eddy/atv/>  
C.M. Zmasek & S.R. Eddy - Washington University, St Louis

ATV est utilisable comme application et comme Applet.

