

Introduction to bioinformatics

Practical exercise



Aalto-yliopisto
Aalto-universitetet
Aalto University

Antonina Kruglova

14.03.2019

Phylogenetic tree exercise

Tools > Multiple Sequence Alignment > Kalign

Results for job kalign-I20190313-105124-0850-64394923-p1m

Alignments

Result Summary

Phylogenetic Tree

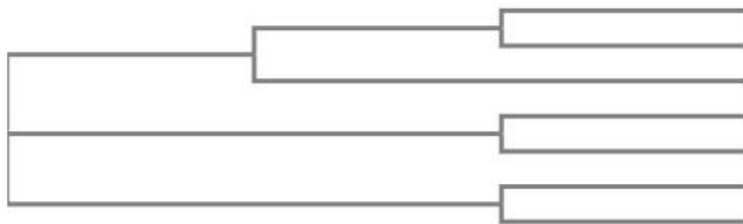
Submission Details

Phylogenetic Tree

This is a Neighbour-joining tree without distance corrections.

Download Phylogenetic Tree Data

Branch length: Cladogram Real



Lactobacillus-bulgar 0.06814
Lactobacillus-acidop 0.07019
Lactobacillus-casei 0.07515
Streptococcus-thermo 0.00546
Streptococcus-saliva 0.00271
Bifidobacterium-adol 0.03236
Bifidobacterium-bifi 0.03221

Why we need computers to answer microbiological questions?

- DNA contains information about all organisms existing
- Excluding human genome over 1000 trillion pairs of bases (AT GC) from over 165 000 species had been sequenced by 2014

Just reading them would take 30 million years!



- ✓ algorithms
- ✓ software tools
- ✓ web servers (storage for big data)

NCBI Resources How To Sign in to NCBI

NCBI National Center for Biotechnology Information

All Databases Search

NCBI Home

- Resource List (A-Z)
- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Training & Tutorials
- Variation

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

Submit
Deposit data or manuscripts into NCBI databases

Download
Transfer NCBI data to your computer

Learn
Find help documents, attend a class or watch a tutorial

Develop
Use NCBI APIs and code libraries to build applications

Analyze
Identify an NCBI tool for your data analysis task

Research
Explore NCBI research and collaborative projects

Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI News & Blog

MANE Select v0.5 is now available
12 Mar 2019
In October last year, we announced the launch of an exciting new collaboration between NCBI and EMBL-EBI called

IgBLAST 1.13.0 now available
11 Mar 2019
IgBLAST is a popular NCBI package for classifying and analyzing immunoglobulin and T cell receptor

Improved search makes it easier to find antimicrobial resistance protein information
11 Mar 2019
It's now easier to find known antimicrobial

Exercise description

- choose one of given unknown sequences from the results “we received analyzing our activated sludge in wastewater treatment plant
- BLAST chosen sequence and report the results of identification (mentioning the % of similarity)
- search literature about the microorganism and prepare its description (with the list of references)
- suggest the role/effect (one or several) this organism may have in wastewater treatment process

In your HW3 you will have to submit the report on the identified organism: classification, important characteristics, contribution to the wastewater treatment process. According to the literature information you can evaluate importance of this microorganism and propose operational parameters which would benefit treatment (to remove or to grow this organism, depending on your decision).