

M2 Mycologie

Outils bioinformatiques

Podospira anserina

- génome 36 Mb (Fasta)
- données de séquençage 2 x 500-800 Mb (Fastq)
- 10k gènes annotés
- assemblage nouvelles souches : entre 4 et 8h (12 coeurs 3.5 GHz)
- phylogénie ITS seuls : 4h (phymI)
- phylogénie codes barres : 12 à 15h (phymI)

- serveur de calcul avec beaucoup de RAM (assemblage) et GPU (phylogénie)
- écriture de scripts shell et Python (ou R) pour les prétraitements et le développement de "workflow"
- serveur de stockage : 400 génomes ADN (+ 96 protéines) = 24 Go
- scripts de recherche/blast automatique (NCBI, JGI)

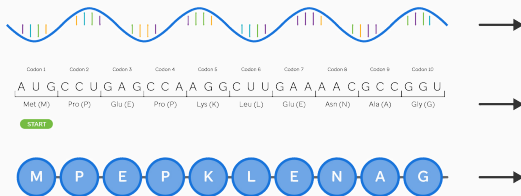
Recherche de motifs

Translation from mRNA to Protein

Mature
mRNA

Nucleotides
As Codons

Amino
Acid Sequence



© Copyright 2022 St. Jude Children's Research Hospital, a not-for-profit, section 501(c)(3)

- blast (shell)
- scripts Python

Alignement de séquence

```
RLA0_METVA  --MIDAKSEHKIAPWKIEEVNALKELLKSANVIALIDMMEVPAVLOEIRDK
RLA0_METJA  ---METKVKAHVADPKIEEVKTLKGLIKSKPVVAIVDMMDVPAPOLOEIRDK
RLA0_PYRAB  -----MAHVAEWKKKEVEELANLIKSYPPVIALVDVSSMPAYPLSQMRRL
RLA0_PYRHO  -----MAHVAEWKKKEVEELAKLIKSYPPVIALVDVSSMPAYPLSQMRRL
RLA0_PYRFU  -----MAHVAEWKKKEVEELANLIKSYPPVIALVDVSSMPAYPLSQMRRL
RLA0_PYRKO  -----MAHVAEWKKKEVEELANLIKSYPPVIALVDVAGVPAAYPLSKMRDK
RLA0_HALMA  MSAESERKTETIPEWKQEEVDAIVEMIESYESVGVVNIAGIPSRLODMRRD
RLA0_HALVO  MSESEVRQTEVIPQWKREEVDELVDFIESYESVGVVGVAGIPSRLOSMRRE
RLA0_HALSA  MSAAEQRTTEEVPEWKQEEVDELVDLLETYSVGVVNVGTGIPSKLODMRRG
RLA0_THEAC  -----MKEVSQKKKELVNEITORIKASRSVAIVDTAGIRTRQIDIRGK
RLA0_THEVO  -----MRKINDKKKEIVSELAODITKSKAVAIVDIKGVRTROMODIRAK
RLA0_PICTO  -----MTEPAQWKIDFVKNLENEINSRKVAAIVSIKGLRNNEFOKIRNS
```

- clustal
- mafft
- muscle¹

¹<https://bioinformaticsreview.com/20151018/multiple-sequence-alignment/>

Mapping et assemblage

Whole Genome Sequencing

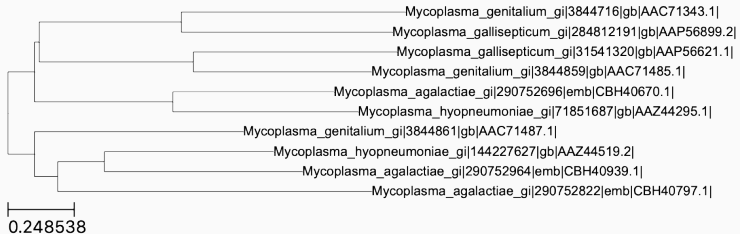
30-60x Coverage



© Copyright 2022 St. Jude Children's Research Hospital, a not-for-profit, section 501(c)(3)

- bowtie2
- unicycler (spades)

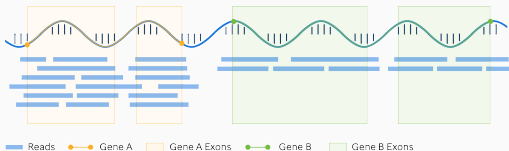
Phylogénie moléculaire



- fasttree
- RaXML
- MEGA
- NGPhylogeny
- seaview (phylip)

RNA-Seq

Coverage Dependent on Expression



© Copyright (2022) St. Jude Children's Research Hospital, a not-for-profit, section 501(c)(3)

- kallisto (R)

- tout ce qui précède sauf Artemis et Seaview
- préférer le serveur Europe
- créer un compte pour le transfert de fichiers et les notifications

Source principale des illustrations : <https://learngenomics.dev/>.