

M2 Mycologie

Outils bioinformatiques



Quelques ordres de grandeur

Podospora 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)

1

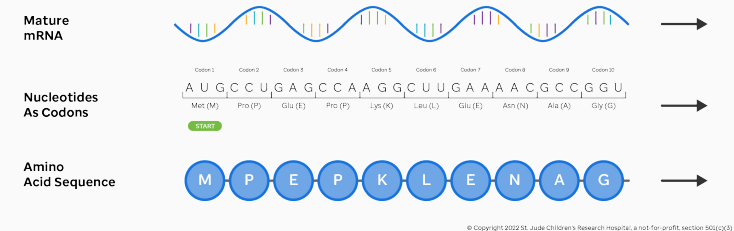
Implications informatiques

- 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)

2

Recherche de motifs

Translation from mRNA to Protein



- blast (shell)
- scripts Python

3

Alignement de séquence

```

RLA0_METVA --MIDAKSEHKIAPWKIEEVNALKEELKSNVIALIDMMEVPAVOLQEIRDK
RLA0_METJA ---METKVKAHVAPWKIEEVKTLKGLIKSKPVVAIVDMMDVPAPOLQEIRDK
RLA0_PYRAB -----MAHVAEWKKKEVEELANLIKSYPIVIALVDVSSMPAYPLSQMRRL
RLA0_PYRHO -----MAHVAEWKKKEVEELAKLIKSYPIVIALVDVSSMPAYPLSQMRRL
RLA0_PYRFU -----MAHVAEWKKKEVEELANLIKSYPIVIALVDVSSMPAYPLSQMRRL
RLA0_PYRKO -----MAHVAEWKKKEVEELANLIKSYPIVIALVDVAGVPAYPLSKMRDK
RLA0_HALMA MSAESERKTETIPEWKQEEYDAIVEMIESYESVGVVNIAGIPRQLQDMRRD
RLA0_HALVO MSAESEVRQTEVIPQWKREEVDLDFIESYESVGVVGVAGIPRQLQSMRRE
RLA0_HALSA MSAEEQRTEEVPEWKQEEVDELVDLLETYSVGVVNVITGIPSKOLODMRRG
RLA0_THEAC -----MKEYSQKKELVNEITRIKASRSVAIVDTAGIRTRQIQDIRGK
RLA0_THEVO -----MRKINPKKEIVSELAODITKSKAVAIVDIKCVRTROMODIRAK
RLA0_PICTO -----MTEPAQWKIDFVKKLENEINSRKVAAIVSIKGLRNNFQKIRNS
    
```

- clustal
- mafft
- muscle¹

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

Mapping et assemblage

Whole Genome Sequencing

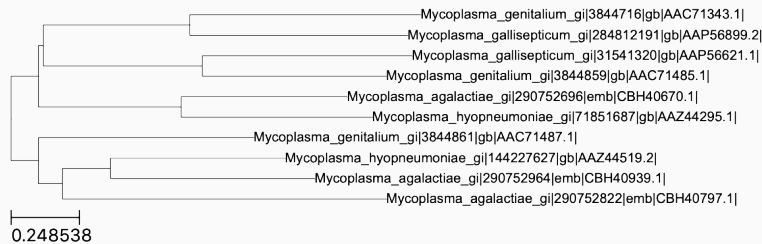
50-60x Coverage



© Copyright 2000 M. Jane Children's Research Hospital, a not-for-profit, section 501(c)(3)

- bowtie2
- unicycler (spades)

Phylogénie moléculaire

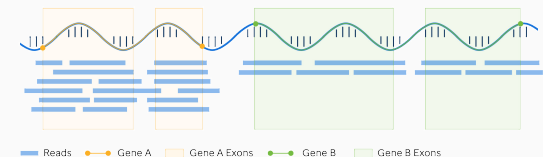


- fasttree
- RaXML
- MEGA
- NGPhylogeny
- seaview (phylip)

RNA-Seq

RNA-Seq

Coverage Dependent on Expression



© Copyright 2000 M. Jane 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/>.