Éléonore Lemieux

536773222

BIF-2102

Projet de recherche en bio-informatique III

**Outil bio-informatique de métagénomique virale**

Rapport présenté à

Antony Vincent

[antony.vincent@fsaa.ulaval.ca](mailto:antony.vincent@fsaa.ulaval.ca)

19 août 2022  
Durée : 16 semaines

IBIS, Pavillon Eugène-Marchand, Université Laval

Table des matières

[Résumé 3](#_Toc103337204)

[Introduction 4](#_Toc103337205)

[Matériel et méthodes 5](#_Toc103337206)

[Résultats 6](#_Toc103337207)

[Discussion 7](#_Toc103337208)

[Conclusion 8](#_Toc103337209)

[Bibliographie 9](#_Toc103337210)

# Résumé

½ page

# Introduction

*présentant la littérature pertinente permettant de comprendre et de supporter les objectifs et hypothèse de la recherche proposée.*

*Intro sur les phages*

*Lytique/lysogène*

*Intro sur métagénomique*

*Intro sur (co)asssemblage*

*Contigs*

*reads*

## Objectifs

* Outil bio-informatique pour investiguer efficacement les phages dans les métagénomes
  + Polyvalent
  + Utilisable par non bio-informaticiens
  + Comparer ensembles de jeux de données (contrôle et traitement, milieux, etc.)
  + Résultats fiables, utiles et représentés clairement
  + Résultat facilement analysable en R

# Matériel et méthodes

écrite dans le format retrouvé dans les articles scientifiques

scripts Bash

environnement bioconda

virsorter2

CheckV

MEGAHIT

Bowtie2

Fastq

# Résultats

principaux résultats obtenus

comprend habituellement des figures et des tableaux qui peuvent être intégrés au texte ou placés en annexe.

# Discussion

discussion commente et critique les résultats de façon à formuler et soutenir les conclusions découlant du projet.

lien en rapport avec les objectifs du projet décrits à la section Introduction. Lorsque cela s'applique, les perspectives du projet peuvent également être présentées.

# Conclusion

# Bibliographie

1.

Hockenberry, A. J. & Wilke, C. O. BACPHLIP: predicting bacteriophage lifestyle from conserved protein domains. *PeerJ* **9**, e11396 (2021).

2.

Shang, J., Jiang, J. & Sun, Y. Bacteriophage classification for assembled contigs using graph convolutional network. *Bioinformatics* **37**, i25–i33 (2021).

3.

Nayfach, S. *et al.* CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol* **39**, 578–585 (2021).

4.

Ruohan, W., Xianglilan, Z., Jianping, W. & Shuai Cheng, L. I. DeepHost: phage host prediction with convolutional neural network. *Briefings in Bioinformatics* **23**, bbab385 (2022).

5.

Dion, M. B., Oechslin, F. & Moineau, S. Phage diversity, genomics and phylogeny. *Nat Rev Microbiol* **18**, 125–138 (2020).

6.

Yukgehnaish, K. *et al.* PhageLeads: Rapid Assessment of Phage Therapeutic Suitability Using an Ensemble Machine Learning Approach. *Viruses* **14**, 342 (2022).

7.

Zielezinski, A., Deorowicz, S. & Gudyś, A. PHIST: fast and accurate prediction of prokaryotic hosts from metagenomic viral sequences. *Bioinformatics* **38**, 1447–1449 (2022).

8.

Boeckaerts, D. *et al.* Predicting bacteriophage hosts based on sequences of annotated receptor-binding proteins. *Sci Rep* **11**, 1467 (2021).

9.

Buchfink, B., Reuter, K. & Drost, H.-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* **18**, 366–368 (2021).

10.

Puente-Sánchez, F., García-García, N. & Tamames, J. SQMtools: automated processing and visual analysis of ’omics data with R and anvi’o. *BMC Bioinformatics* **21**, 358 (2020).

11.

Tamames, J. & Puente-Sánchez, F. SqueezeMeta, A Highly Portable, Fully Automatic Metagenomic Analysis Pipeline. *Front. Microbiol.* **9**, 3349 (2019).

12.

Kieft, K., Zhou, Z. & Anantharaman, K. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* **8**, 90 (2020).

13.

Guo, J. *et al.* VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* **9**, 37 (2021).