


Université de Rennes 1
UFR Sciences de la Vie
et de l'Environnement

Long reads assembly using linear optimization mathematical approach

---

BioInformatic Master 1
- Internship report -

Victor Epain

Inria research laboratory

Parc Alcyone - 1 rue Yvonne et Pierre Meynier

35069 RENNES FRANCE

GenScale Team

Under the responsabilty of :

Rumen Andonov, Inria

Dominique Lavenier, cnrs

2019 april 23 — 2019 june 28

UMR CNRS 6290
Institut de Génétique et Développement de Rennes (IGDR)

Annabelle MONNIER
Equipe Oncogénomique translationnelle

UNIVERSITÉ DE RENNES 1

ENGAGEMENT DE NON PLAGIAT

Je, soussigné (e) ……
Etudiant (e) en ……

Déclare être pleinement informé (e) que le plagiat de documents ou d'une partie de documents publiés sous toute forme de support (y compris l'internet), constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée.

En conséquence, je m'engage à citer toutes les sources que j'ai utilisées pour la rédaction de ce document.

Signature

Laboratoire de Génomique Médicale
BMT-HC - CHU Pontchaillou

2 rue Henri le Guilloux
35033 Rennes Cedex
FRANCE

Annabelle MONNIER
annabelle.monnier@univ-rennes1.fr
TÉL. 33 (0)2 99 28 92 54

Thanks

Contents

Chapter 1

Introduction

1.1 Long read assembly

1.1.1 The advantage of long read assembly against reptitions in genomes

1.1.2 The long read sequencing technology errors issue

1.1.3 The hybrid method to correct short read assembly

1.2 The linear optimization : a mathematical exact approach

Chapter 2

Material and methods

2.1 Material

2.1.1 Eshereshia coli genome

2.1.2 Nanopore long read generator

2.1.3 AMPL optimization mathematical implementation language

2.1.4 Use of GUROBI solver to solve the implemented problem in AMPL

2.2 Methods

2.2.1

Chapter 3

Results

Chapter 4

Discussion

Chapter 5

Conclusion

Bibliography

Appendix A

Host Presentation

Appendix B

Personal Review