JHU THESIS TEMPLATE TITLE

by

Yunfan Fan

A dissertation submitted to Johns Hopkins University
in conformity with the requirements for the degree of
Doctor of Philosophy

Baltimore, Maryland
August, 2022

Abstract

While next generation sequencing (NGS) has enabled massively parallel DNA sequencing for lower and lower cost, the development of third generation nanopore sequencing offers several key advantages over older sequencing methods. Nanopore sequencers are pocket-sized, making them orders of magnitude cheaper than the next most affordable alternative and the ideal option for wide deployment. They are capable of providing data in real-time, saving valuable hours before data analysis can begin. Additionally, they are able to sequence reads several thousand basepairs long, as opposed to the hundreds of basepairs NGS platforms are capable of, and they embed base modification data without the need for specific treatment beforehand. Given these advantages, in this thesis I examine the application of nanopore sequencing to the study of human pathogens.

First, we use nanopore sequencing to characterize anti-microbial resistance (AMR) in forty clinical isolates. We analyzed real-time data to quickly identify AMR genes, assembled genomes to identify chromosomal mutations, and used short-read sequencing data to correct the errors in the assemblies. With sequencing data, we found that time to effective antibiotic therapy could be shortened by as much as 20 hours compared to standard antimicrobial

susceptibility testing (AST).

Second, we leverage the long reads of nanopore sequencing to assemble the genome of a pathogenic yeast, /textitCandida nivariensis. Previous efforts to assemble this yeast genome relied solely on NGS data, resulting in a highly fragmented genome. Using nanopore data, we achieve a much higher contiguity, capture previously missing portions of the genome. Furthermore, we demonstrate that our more contiguous genome can be used to better study long and repetative genes, such as those involved in pathogenticity to humans.

Third, we use the base modification information embedded in nanopore sequencing data to call methylation in metagenomic assemblies. These calls enable the binning of metagenomic contigs according to methylation signature without the need to collect additional data. We demonstrate the efficacy of this method on a synthetic community sample, a simple two-bacteria system, and a clinical sample with matched proximity ligation binning data.

These applications of nanopore sequencing demonstrate its potential and its utility for all fronts of pathogen genomics research.

Thesis Committee

Primary Readers

Dr. Winston Timp (Primary Advisor)
Associate Professor
Department of Biomedical Engineering
Department of Molecular Biology and Genetics
Johns Hopkins University School of Medicine

Dr. Patricia Simner
Associate Professor
Department of Pathology
Johns Hopkins University School of Medicine

Dr. Steven Salzberg
Bloomberg Distinguished Professor
Department of Computer Science
Johns Hopkins University Whiting School of Engineering
Department of Biomedical Engineering
Johns Hopkins University School of Medicine
Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health

Alternate Readers

First Lastname
Professor
Department of ChangeMe
Johns Hopkins Bloomberg School of Public Health

First Lastname
Assistant Professor
Department of ChangeMe
Johns Hopkins Bloomberg School of Public Health

Acknowledgments

I have tremendous gratitude
to those people,
numerous and uncountable,
who have contributed,
directly or in subtler ways,
to this work.

Some of them are listed here.

To my advisor, Winston: I remember writing to you as a sophomore in college many years ago, asking to do research in your brand new lab, which at the time was but a few months old. Back then, I hardly knew what research was and had no relevant skills or credentials to offer, only my time and my interest to learn. Over these years I've learned so much from you, and will always be grateful to you for building the place where I was able to grow.

To my thesis committee, Trish and Steven: Thank you for your patient guidance, encouragement, advice, and for helping me to keep an eye on the bigger picture.

To the @yfan arc of the #core channel - @isac, @brochael, @shao, @gilfunk,

@narley, @broham, @gmoney, @Brittany, @sherbear, @Sam Sholes, @Paul Hook, @amymeltzer39, and @alice: Thank you for those times when you patiently watched over me as I learned new lab techniques, answered my dumb questions, and generally saved me from my own buffoonery. Thank you even more for commiserating with me as we struggled together through the singular challenges of research, and celebrating the equally singular triumphs.

To the crew that moved me into Boonique, and Charles, and Charlotte, and Sven, and Manolo: Thanks for being there.

To mom and dad, and family further away: It was your labor that first cultivated my growth. Accomplishments in my name are as much yours as they are mine. I flourish for you.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Introduce your thesis (Aardvark, 1900)

References

Aardvark, A. A. (1900). “Article title”. In: *Journal One* 1.1, pp. 1–8.

Chapter 2

Genome assembly of Candida nivariensis

This document was generated on Mon May 2 16:13:51 2022.

As shown in Figure 2.1 and in Table 2.1 we can see that bla bla (Abramson and Barbie, 1900).

A	B
1	0.00
2	0.16
3	-0.56
4	-1.40
5	0.05
6	0.06
7	1.08
8	0.14
9	-2.29
10	0.02

Table 2.1: Table title Table description

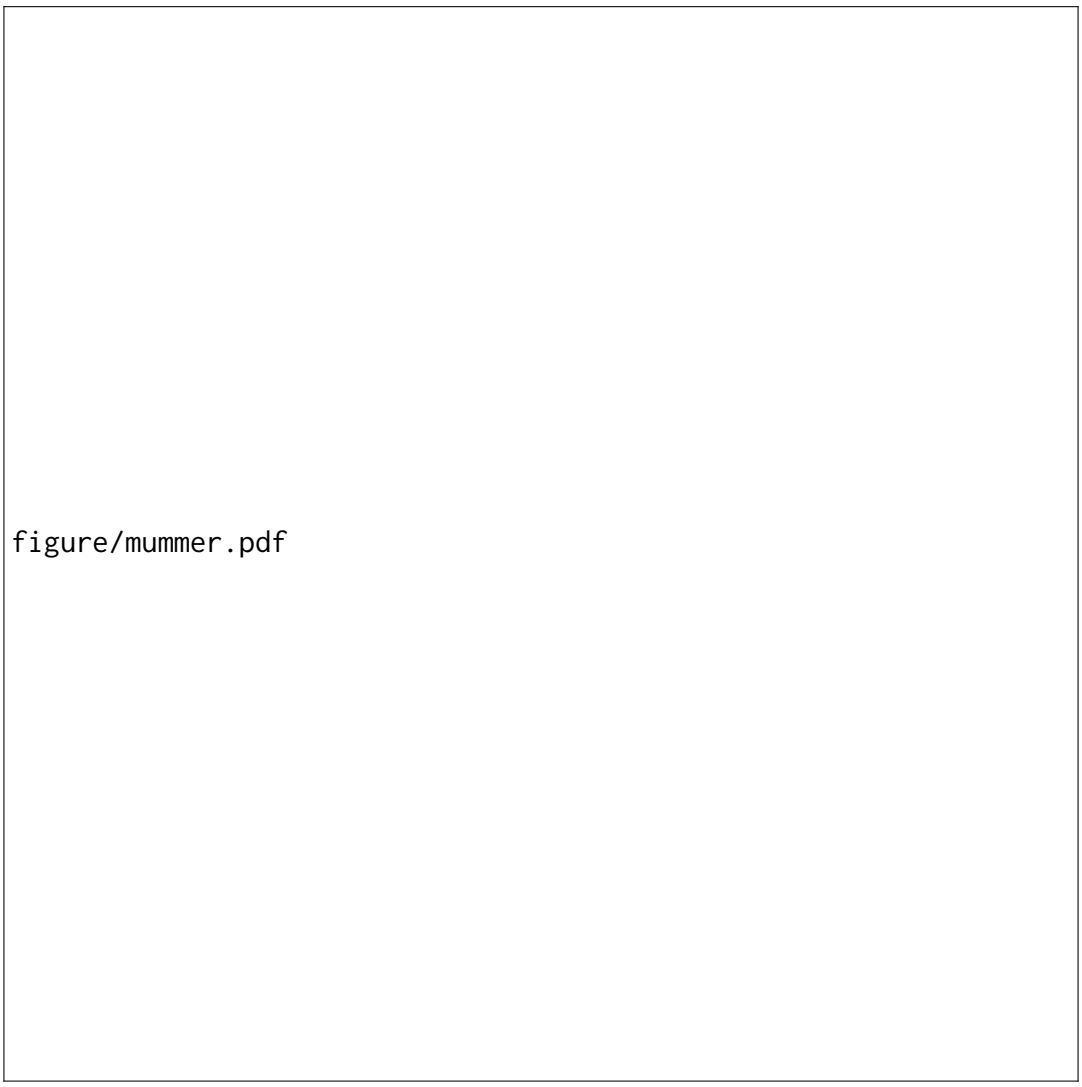

Figure 2.1: Mummer between stuff. (a) blah first point **(b)** blah second point **(c)** blaih thrid panel

References

Abramson, A. A. and B. B. Barbie (1900). "Article title". In: *Journal Two* 1.1, pp. 9–17.

Chapter 3

Discussion and Conclusion

Discuss and conclude your thesis (Abramson, Barbie, and Rider, 1900)

References

Abramson, A. A., B. B. Barbie, and C. C. Rider (1900). "Article title". In: *Journal Three* 1.1, pp. 192–244.

John Doe

Resumé title

Some quote

Education

year–year **Degree**, *Institution*, City, *Grade*.
Description

year–year **Degree**, *Institution*, City, *Grade*.
Description

Master thesis

title *Title*

supervisors Supervisors

description Short thesis abstract

Experience

Vocational

year–year **Job title**, *Employer*, City.
General description no longer than 1–2 lines.
Detailed achievements:

- Achievement 1;
- Achievement 2, with sub-achievements:
  - Sub-achievement (a);
  - Sub-achievement (b), with sub-sub-achievements (don't do this!);
    - Sub-sub-achievement i;
    - Sub-sub-achievement ii;
    - Sub-sub-achievement iii;
  - Sub-achievement (c);
- Achievement 3.

year–year **Job title**, *Employer*, City.
Description line 1
Description line 2

Miscellaneous

street and number – postcode city – country
+1 (234) 567 890 • *+2 (345) 678 901* • *+3 (456) 789 012*
john@doe.org • *www.johndoe.com* • *john.doe* • *jdoe*
jdoe • *additional information*

year–year **Job title**, *Employer*, City.
Description

Languages

Language 1	Skill level	*Comment*
Language 2	Skill level	*Comment*
Language 3	Skill level	*Comment*

Computer skills

category 1	XXX, YYY, ZZZ	category 4	XXX, YYY, ZZZ
category 2	XXX, YYY, ZZZ	category 5	XXX, YYY, ZZZ
category 3	XXX, YYY, ZZZ	category 6	XXX, YYY, ZZZ

Interests

hobby 1	Description
hobby 2	Description
hobby 3	Description

Extra 1

- Item 1
- Item 2
- Item 3. This item is particularly long and therefore normally spans over several lines. Did you notice the indentation when the line wraps?

Extra 2

- Item 1
- Item 2
- Item 3
- Item 4
- Item 5[3]
- Item 6. Like item 3 in the single column list before, this item is particularly long to wrap over several lines.

References

Category 1

- Person 1
- Person 2
- Person 3

Category 2

Amongst others:

- Person 1, and
- Person 2

(more upon request)

All the rest & some more

That person, and **those** also (all available upon request).

Publications

[1] John Doe. Title, year.

[2] John Doe. Title, year.

[3] John Doe and Author 1. *Title*. Publisher, edition edition, year.

[4] John Doe and Author 2. *Title*. Publisher, edition edition, year.

[5] John Doe and Author 3. Title, year.

Company Recruitment team January 01, 1984
Company, Inc.
123 somestreet
some city

Dear Sir or Madam,

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis ullamcorper neque sit amet lectus facilisis sed luctus nisl iaculis. Vivamus at neque arcu, sed tempor quam. Curabitur pharetra tincidunt tincidunt. Morbi volutpat feugiat mauris, quis tempor neque vehicula volutpat. Duis tristique justo vel massa fermentum accumsan. Mauris ante elit, feugiat vestibulum tempor eget, eleifend ac ipsum. Donec scelerisque lobortis ipsum eu vestibulum. Pellentesque vel massa at felis accumsan rhoncus.

Suspendisse commodo, massa eu congue tincidunt, elit mauris pellentesque orci, cursus tempor odio nisl euismod augue. Aliquam adipiscing nibh ut odio sodales et pulvinar tortor laoreet. Mauris a accumsan ligula. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Suspendisse vulputate sem vehicula ipsum varius nec tempus dui dapibus. Phasellus et est urna, ut auctor erat. Sed tincidunt odio id odio aliquam mattis. Donec sapien nulla, feugiat eget adipiscing sit amet, lacinia ut dolor. Phasellus tincidunt, leo a fringilla consectetur, felis diam aliquam urna, vitae aliquet lectus orci nec velit. Vivamus dapibus varius blandit.

Duis sit amet magna ante, at sodales diam. Aenean consectetur porta risus et sagittis. Ut interdum, enim varius pellentesque tincidunt, magna libero sodales tortor, ut fermentum nunc metus a ante. Vivamus odio leo, tincidunt eu luctus ut, sollicitudin sit amet metus. Nunc sed orci lectus. Ut sodales magna sed velit volutpat sit amet pulvinar diam venenatis.

Albert Einstein discovered that $e = mc^2$ in 1905.

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

Yours faithfully,

John Doe

Attached: curriculum vitæ

John Doe
street and number – postcode city – country
+1 (234) 567 890 • *+2 (345) 678 901* • *+3 (456) 789 012*
john@doe.org • *www.johndoe.com* • *john.doe* • *jdoe*
jdoe • *additional information*