

Microbial Genomics 2025

Faculty: Andrew McArthur, Fiona Brinkman, Gary van Domselaar, Jared Simpson, Jimmy Liu, I

November 20-22, 2025

Contents

I	Introduction	5
1	Workshop Info	7
1.1	Pre-work	7
1.2	Class Photo	7
1.3	Schedule	7
2	Meet Your Faculty	9
3	Data and Compute Setup	13
II	Modules	15
4	Module 1 Genome Assembly	17
4.1	Lecture	17
4.2	Lab	17
5	Module 2 Genome Annotation	19
5.1	Lecture	19
5.2	Lab	19
6	Module 3 Antimicrobial Resistant Gene (AMR) Analysis	21
6.1	Lecture	21
6.2	Lab	21

7	Module 4 Mobile Genetic Elements	23
7.1	Lecture	23
7.2	Lab	23
8	Module 5 Pan Genomics	25
8.1	Lecture	25
8.2	Lab	25
9	Module 6 Phylogenetics	27
9.1	Lecture	27
9.2	Lab	27

Part I

Introduction

Chapter 1

Workshop Info

Welcome to the 2025 Microbial Genomics Canadian Bioinformatics Workshop webpage!

1.1 Pre-work

You can find your pre work here:

1. Please read **2023 CARD DB publication**.
2. Please read **2023 CARDShark publication**.
3. Please read **2025 Bait Capture VERSION 2 publication**.
4. Please read **2025 CARD k-mers pre-print**.
5. Please read **2025 CZ-ID publication**.

1.2 Class Photo

Coming soon!

1.3 Schedule

Chapter 2

Meet Your Faculty

2.0.0.1 Andrew McArthur

Professor Michael G. DeGroot Institute for Infectious Disease Research Department of Biochemistry and Biomedical Sciences McMaster University Hamilton, ON, Canada

Dr. McArthur is a Professor and David Braley Chair in Computational Biology at McMaster University. Dr. McArthur has had a career in the United States and Canada, including NIH-funded positions at the Marine Biological Laboratory and Brown University, where he led the genome assembly of the diarrheal pathogen *Giardia intestinalis*, plus 10 years of experience in the private sector. Dr. McArthur's research team focuses on building tools, databases, and algorithms for the genomic surveillance of infectious pathogens. He and his team developed the Comprehensive Antibiotic Resistance Database (card.mcmaster.ca) and the SARS-CoV-2 Illumina GeNome Assembly Line software platform

2.0.0.2 Fiona Brinkman

Distinguished Professor, FRSC, Department of Molecular Biology and Biochemistry; Associate Member, School of Computing Science and Faculty of Health Sciences Simon Fraser University Burnaby, BC, Canada

— brinkman@sfu.ca, <https://www.brinkmanlab.ca/>

Dr. Brinkman is developing bioinformatic resources to better track infectious diseases using genomic data, and improve prediction of new vaccine/drug targets. Her primary aim is to develop more sustainable, integrated approaches

for infectious disease control, however she is also applying her methods to aid allergy, child health, and environmental research.

2.0.0.3 Gary Van Domselaar

Chief, Bioinformatics National Microbiology Laboratory Public
Health Agency of Canada Winnipeg, MB, Canada

— gary.vandomselaar@phac-aspc.gc.ca

Dr. Gary Van Domselaar, PhD (University of Alberta, 2003) is the Chief of the Bioinformatics Section at the National Microbiology Laboratory in Winnipeg Canada and Associate Professor in the Department of Medical Microbiology and Infectious Diseases at the University of Manitoba. Dr. Van Domselaar's lab develops bioinformatics methods and pipelines to understand, track, and control circulating infectious diseases in Canada and globally. His research and development

2.0.0.4 Jared Simpson

Principal Investigator, Ontario Institute for Cancer Research Assistant Prof. Department of Computer Science, University of Toronto
Toronto, ON, Canada

— <https://simpsonlab.github.io/>

Dr. Simpson develops algorithms and software for the analysis of high-throughput sequencing data. He is interested in de novo assembly and the detection of sequence variation in individuals, cancers and populations, with a focus on long read sequencing technologies. Dr. Simpson developed the ABYSS, SGA and nanopolish software packages.

2.0.0.5 Jimmy Liu

Coming soon Coming soon Canada

— <https://simpsonlab.github.io/>

BIO

2.0.0.6 Rob Beiko

Professor, Faculty of Computer Science Dalhousie University Halifax, NS, Canada

— rbeiko@dal.ca

Rob Beiko is a professor in bioinformatics whose research encompasses microbial evolution and gene transfer, comparative genomics, and microbial community analysis. He is the lead of the ARETE project, which is developing a software pipeline for annotation of AMR and mobile genetic elements, phylogenomic analysis, and inference of recombination and transfer in large sets of pathogen genomes. He has also developed and contributed to software tools including STAMP for statistical analysis of metagenome profiles, rSPR for efficient inference of gene transmission, and PICRUSt for the prediction of metagenome composition based on marker-gene surveys.

2.0.0.7 William Hsiao

Associate Professor, Faculty of Health Sciences Simon Fraser University Vancouver, BC, Canada — wwhsiao@sfu.ca, www.cidgoh.ca

William Hsiao is a public health infectious disease researcher with a background in microbial genomics and bioinformatics. He is an associate professor in the Faculty of Health Sciences at Simon Fraser University and an affiliated researcher at BCCDC Public Health Laboratory and at Canada's Michael Smith Genome Sciences Centre. Will leads an interdisciplinary group of researchers interested in solving practical public health and animal health problems through a One Health lens at the Centre for Infectious Disease Genomics and One Health.

2.0.0.8 Charlie Barclay

MSc Graduate Student Researcher, Department of Physics and Astronomy University of British Columbia Vancouver, BC, Canada — cbarcl01@mail.ubc.ca

William Hsiao is a public health infectious disease researcher with a Charlie is an ontology curator at the Centre for Infectious Disease Genomics and One Health (CIDGOH), focusing on data standards for contextual data of genomic epidemiology, including wastewater surveillance. With five years of experience in data management, specializing in biodiversity and genomics data, she also actively contributes to the Public Health Alliance for Genomic Epidemiology (PHA4GE) and the Global Alliance for Genomics and Health (GA4GH).

Chapter 3

Data and Compute Setup

3.0.0.1 Course data downloads

Coming soon!

3.0.0.2 Compute setup

Coming soon!

Part II

Modules

Chapter 4

Module 1 Genome Assembly

4.1 Lecture

4.2 Lab

Chapter 5

Module 2 Genome Annotation

5.1 Lecture

5.2 Lab

Chapter 6

Module 3 Antimicrobial Resistant Gene (AMR) Analysis

6.1 Lecture

6.2 Lab

Chapter 7

Module 4 Mobile Genetic Elements

7.1 Lecture

7.2 Lab

Chapter 8

Module 5 Pan Genomics

8.1 Lecture

8.2 Lab

Chapter 9

Module 6 Phylogenetics

9.1 Lecture

9.2 Lab