

RESEARCH INTERESTS

I am interested in machine learning for computational biology, in particular:

- Generative models of protein sequence and structure
- Applications and theory of data-driven design algorithms in synthetic biology
- Evolutionary dynamics of proteins and gene regulatory networks

EXPERIENCE

University of Toronto

Research Intern

Aug 2019 – Present

Supervisors: Prof. Philip Kim, Dr. Pedro Alberto Valiente Flores

- Designing novel protein folds with generative models of protein structures (Transformer, GAN).
- Designing oncoprotein inhibitors using Rosetta and molecular dynamics simulations.
- Implemented code to search the PDB for protein surfaces that mimic DNA.

iGEM Toronto

Drylab Lead

Apr 2020 – Present

Drylab Member

Apr 2019 – Apr 2020

Supervisor: Prof. Radhakrishnan Mahadevan

- Designed a plastic-degrading PET hydrolase using generative and discriminative models.
- Compiled dataset which maps protein surface features (MaSIF) to interacting peptides.
- Mapping the protein thermostability landscape with Meltome and Unirep.
- Quantifying benefits of learned protein sequence embeddings on protein function prediction.
- Analyzing PET catalysis dynamics with mixed quantum/classical dynamics simulations.

Canadian Synthetic Biology Education Research Group

Machine Learning Instructor

Sep 2019 – Present

Supervisor: Patrick Diep

- Created three hour-long interactive workshops on protein design, covering both classical (Docking, Rosetta, MD) and recent methods (Sequence-to-function models, generative models, representation learning, active learning).

Toronto General Hospital

Research Student

Feb 2019 – Dec 2019

Supervisor: Dr. Moumita Barua

- Implemented RNA-Seq analysis pipeline using bioinformatics tools (htseq, bamtools, subread).
- Analyzed differential expression data to find genes associated with kidney sclerosis.

EDUCATION

University of Toronto

B. Sc. in Bioinformatics and Computational Biology, Computer Science *Sep 2017 – Jun 2021*
cGPA: 3.92/4.00

SKILLS

Programming Python, PyTorch, NumPy, Bash

Tools PyMOL, VMD, GROMACS, RosettaScripts, Unix, L^AT_EX

TALKS

1. Accelerating Plastic Recycling with PETase, *iGEM Giant Jamboree*, Boston, MA, Oct. 2019.
2. Recurrent Neural Networks for Protein Design, *Ontario iGEM Conference*, Guelph, ON, Jul. 2019.

AWARDS

- NSERC Undergraduate Research Award *2020*
- COVID-19 Student Engagement Award *2020*
- iGEM Gold Medal, Best Manufacturing Project Nomination, Boston MA *2019*
- National Biology Competition, Top 1% *2017*
- University of Toronto Entrance Scholarship *2017*

INTERESTS

Rachmaninoff Piano Concerto No. 2, cycling, public transport, making things with flour