

# Siddharth Reed

✉ [sreed@andrew.cmu.edu](mailto:sreed@andrew.cmu.edu) | [DJ Siddharth Vader](https://github.com/DJSiddharthVader) | [in Sid-Reed](https://www.linkedin.com/in/Sid-Reed)

## Education

### Master of Science - Computational Biology

CARNEGIE MELLON UNIVERSITY

- Relevant Coursework: Algorithms, Statistics, Programming, Machine Learning, Genomics

Sep. 2020 - Apr. 2022

Pittsburgh, PA, USA

### Honours Bachelor of Science - Molecular Biology & Genetics Co-op

MCMASTER UNIVERSITY

- Relevant Coursework: Software Development, Mathematical Modelling, Bioinformatics, Microbial & Human Genetics

Sep. 2015 - Apr. 2020

Hamilton, ON, Canada

## Experience

### MCDS, Carnegie Mellon University

11-637 COURSE DEVELOPER ([COURSE PAGE](#))

- Helped to develop a new course project for MSc. students developing data science skills
- Created teaching materials that relate to core data science concepts with a focus on bioinformatics
- Collaborating with professors to provide comprehensive background and support learning objectives for students
- Provided weekly updates on materials and process to senior staff and TAs

May 2021 - Aug. 2021

### Hoffman Lab, University Of Toronto

RESEARCH ASSISTANT CO-OP ([LAB SITE](#))

- Comparing transcriptomic data between publicly available placental, cancer and "normal" tissues
- Extensively cleaning and pre-processing data in a clear, explainable, reproducible manner
- Self-taught and applied linear models to differential expression analysis and correcting for batch effects
- Creating clear intuitive visuals to help communicate and understand patterns in biological data

May 2019 - Aug. 2019

### Golding Lab, McMaster University

UNDERGRADUATE THESIS STUDENT ([GITHUB](#))

- Investigated the relationship between horizontal gene transfer rates and CRISPR-Cas systems using novel techniques
- Built and end-to-end pipeline to download, process, analyze and visualize public genomic data
- Applied methods in network theory and statistics to reach conclusions about biological phenomena
- Presented my work in a clear, engaging way as both a final presentation and manuscript

Sep. 2018 - Apr. 2019

### Adapsyn, McMaster University

DATA SCIENCE CO-OP ([GITHUB](#))

- Performed various analyses with many kinds of biological data to accelerate natural products discovery
- Applied statistics and statistical learning techniques to help direct chemists in natural products discovery
- Communicated effectively with wet lab scientists to address problems from a computational perspective
- Became adept at working under constant deadlines and effectively multitasking

Jan. 2018 - Aug. 2018

## Extracurriculars

### fainhD: filtering and identifying non-host DNA pipeline

PERSONAL PROJECT ([GITHUB](#))

- Helped assemble pipeline to detect viral reads in human RNASeq data and perform functional analysis
- Use Snakemake to create a simple, easy to use pipeline that handles all dependencies itself
- Collaborated with group members using git

Feb. 2021 - March. 2021

### SELEXzyme: Generating DNAzymes with a Genetic Algorithm

PERSONAL PROJECT ([GITHUB](#))

- Implemented a genetic algorithm and trained an SVM model to simulate SELEX sequence optimization *in silico*
- Developed a simple, organized, well-documented codebase using Go and python
- Wrote a detailed README explaining implementation details and choices, usage details and a video demo
- Empirically validated my program results on real world data and writing a summary report

Sep. 2020 - Dec. 2020

## Skills

### Programming

Proficient: Python, Bash, Vim, Git, LaTeX, R, Go, Jupyter

Familiar: Bioconductor, ggplot2, Pandas, NCBI, Cluster Computing (slurm, sungrid)

### Research Languages

Literature Review, Data Analysis, Data Visualization, Experimental Design, Scientific Writing

English, French