

Haotian Teng

E-mail: havens.teng@gmail.com haotianteng@andrew.cmu.edu

Personal Website: <https://haotianteng.github.io/>

RESEARCH EXPERIENCE

Ph.D. Research Project, combine deep learning and probabilistic graphical model to clustering the spatial transcriptomics data in single cell level. 2019.9 – 2021.2

- Developed a dimensional reduction tool for gene expression profile (Program page: <https://github.com/haotianteng/GECT>).
- Developed a probabilistic Graphical model to clustering the cell from the single-cell spatial gene expression data (Program page: <https://github.com/haotianteng/FICT>).
- Produce a simulation pipeline for validate the spatial transcriptomics clustering tools and benchmark and visualization the clustering result. (Program page: <https://github.com/haotianteng/FICT-SAMPLE>)
- Conduct differential gene expression analysis and GO term annotation.

Ph.D. Research Project, RNA methylation basecalling using Nanopore sequencer. 2019.10 – 2021.12

- Apply VQ-VAE to do semi-supervised deep learning to train a m6A methylation basecaller for the ONT RNA sequencer.

Masters Research Project, Using Machine learning algorithm in Nanopore Basecalling 2017.2 – 2018.7

Advisor: Prof. Lachlan Coin, Institute for Molecular Bioscience, University of Queensland

- Built a deep learning-based basecaller **Chiron** using Tensorflow, for Oxford Nanopore sequencer basecalling (Program page: <https://github.com/haotianteng/Chiron>)
- Developed a preprocessing tool **Nanopre** to identify the polyA region in the Nanopore RNA sequencing platform. (Program page: https://github.com/haotianteng/Nanopre_Tool)
- Prepared training dataset of DNA and RNA Nanopore basecalling reads, using Nanoraw and graphmap to label the data.
- Implemented a pipeline in Google Cloud and Google Compute engine for end-to-end genome analysis.

Internship and Winter Scholar, The development of spontaneous neural activity in the zebrafish 2016.3 – 2017.2

Advisor: Prof. Geoffrey Goodhill, Queensland Brain Institute, University of Queensland

- Built a pipeline for laboratory automation and data analysis in Zebrafish neuron experiment with Aduino, LabVIEW and MATLAB.
- Constructed PHANTOM toolbox for projecting visual stimulation with conformal transformation, used for zebrafish tectum research. Program page in Github: <https://github.com/haotianteng/PHANTOM-toolbox>
- Developed algorithms for functional connectivity reconstruction using regularization method under scale free assumption, correct the false positive correlation due to common input, transition connection and latent common input.

Internship, Feedback in AIY neurons in Thermotaxis behavior of C.elegans 2015.7-2015.12

Advisor: Prof.Aravinthan D.T. Samuel, Center for Brain Science, Harvard University, Boston

- Studied thermotaxis in C.elegans with tracking and multi-neuron fluorescent marked.
- Cross & keep the worm, experiment using a spinning disk confocal microscope and the afterwards data acquisition & processing with combination of ImageJ (Miji) and Matlab
- Proved the derivation dependence between AFD neuron and temperature, designed and conducted the experiment to measure the parameters of the AFD-temperature relationship with temperature signal input under different shape.

Research Assistant, Locomotion and PH sensing mechanism in C.elegans & fast reaction tracking System development 2012.7-2015.6

Advisor: Dr. Louis Tao, Center for Bioinformatics, Peking University, Beijing

- Marked GCaMP6 into the C.elegans ASH, AWC, ASE neurons to testify and determine the neuron responsible for PH sensing.
- Developed a neuro-muscle model of C.elegans motor system and proved the theoretical prediction of gait adaptation in C.elegans.
- Recorded and analyzed long-term locomotion parameter of C.elegans by using a tracking and photographing system.
- Developed a visualization tool with OpenGL to describe and simplify the neuron network in C.elegans, and enabled the tool to search the whole neural pathway through any two given neurons.
- Built a tracking system as one of the contributors, which could achieve high-precision (accuracy below 1 micron) tracking and photographing and simultaneous data collection & processing
- Modified and developed a “snake” model based algorithm for robust and precise C.elegans center line extraction.

Professional EXPERIENCE

Bioinformatics Engineer

Novogene Europe, Beijing, China

2018.9-2019.02

- Optimized the human resequencing and laboratory automation pipeline.
- Designed and developed the long-read sequencing platform.

Senior Research Technician

Institute for Molecular Bioscience, University of Queensland, Australia

2017.6-2018.7

- Worked on Oxford Nanopore Technologies Long-read Nanopore direct RNA sequencing data processing, improve the sequencing accuracy and efficiency, improve the succeed sequencing reads ratio by 15X compared to the original pipeline for long poly-A tail reads.

Intern

2014.3-2014.10

Biodynamics Optical Imaging Center, PKU, Beijing, China

- Designed, fabricated, and tested a micro-fluid chip for fast generating stable linear gradient field.
- Developed a Computational Fluid Dynamics (CFD) module for the microfluid chips fluid field calculation in Fluent, which could draw the flow field from the CAD design sketch.

PUBLICATIONS

- **Teng, H.**, Yuan, Y., & Bar-Joseph, Z. (2021). Clustering Spatial Transcriptomics Data. (in press) Bioinformatics.
- Pitt, M. E., Nguyen, S. H., Duarte, T. P., **Teng, H.**, Blaskovich, M. A., Cooper, M. A., & Coin, L. J. (2020). Evaluating the genome and resistome of extensively drug-resistant *Klebsiella pneumoniae* using native DNA and RNA Nanopore sequencing. *GigaScience*, 9(2), g1aa002.
- **Teng, H.**, Cao, M. D., Hall, M. B., Duarte, T., Wang, S., & Coin, L. J. (2018). Chiron: translating nanopore raw signal directly into nucleotide sequence using deep learning. *GigaScience*, 7(5), giy037.
- Avitan, L., Pujic, Z., Mölter, J., Van De Poll, M., Sun, B., **Teng, H.**, Amor, R., Scott, E.K. and Goodhill, G.J., 2017. Spontaneous activity in the zebrafish tectum reorganizes over development and is influenced by visual experience. *Current Biology*, 27(16), pp.2407-2419.
- **Teng, H.** "A neuron-muscle circuit model of *C.elegans*'s locomotion." *Bachelor of Science Thesis: Peking University*, 2015

EDUCATION

Carnegie Mellon University, Pittsburgh, United States

Ph.D., Computational Biology

- Advisor: Ziv Bar-Joseph, Machine Learning Department and Computational Biology department, School of Computer Science, Carnegie Mellon University.

2020 - present

University of Queensland, Queensland, Australia

M.S., Bioinformatics

2016 – 2019

- Advisor: Prof. Lichlan Coin, Institute of Molecular Bioscience, University of Queensland
- Advisor: Prof. Geoffrey Goodhill, Queensland Brain Institute, University of Queensland

Peking University, Beijing, China

B.S., Physics

2011 – 2015

HONORS AND AWARDS

- The 1st Prize at 27th Chinese Physics Olympiad, Zhejiang Province (rank 1/1232 in theory part) 2011
- The Silver Medal at 27th Chinese Physics Olympiad, Finals 2011
- The 1st Prize at 29th Parts of the National College Students Physics Competition 2012

SKILLS

- Programming: Python, C, C++, Matlab, R, Linux, LaTeX,
- Packages&Platforms: Tensorflow, MXNet, Caffe, CUDA, cuDNN, OpenGL, BWA, SAMtools, Velvet, DIAMOND, BLAST+, Minimap2, H5py, Psychtoolbox, LabVIEW, Arduino.
- Software: PyMOL, Fluent(ANSYS), Origin, AutoCAD, Primer Premier, DNA Man, Microsoft Office,
- Wet-lab experiment skill: Molecular cloning, Microinjection
- Language: Chinese(Mother Language), English(Fluent), Spanish(basic), German (Pizza-orderable)
- TOEFL: Cumulative 103 (R 29, L 29, S 23, W 22); GRE: V 150, Q 169, AW 3.0
- Proficient in Piano playing, accomplished the Piano highest-grade (grade 10) in 2005, learned since 6 years old. Skillful in saxophone.