

Chaohui Li

chl221@ucsd.edu | (858) 518-9812 | <https://eric0627.github.io>

RESEARCH INTERESTS

Algorithms, bioinformatics, computational biology, cancer genomics, and machine learning.

EDUCATION

University of California San Diego

PhD student in Computer Science and Engineering

GPA: 4.0/4.0

Advisor: Prof. Vineet Bafna

San Diego, USA

2023 – Present

City University of Hong Kong

Bachelor of Science in Computer Science

GPA: 3.87/4.3 (Top 10% & First Class Honours)

Hong Kong, China

2017 – 2021

PUBLICATIONS

(*joint first authorship)

1. **C. Li***, L. Chen*, G. Pan, W. Zhang, & S. C. Li. “Deciphering complex breakage-fusion-bridge genome rearrangements with Ambigram”. *Nature Communications*, 14, 5528, Sep 2023. DOI: <https://doi.org/10.1038/s41467-023-41259-w>
2. L. Zhang, L. Chen, S. C. Li, M. Wang, **C. Li**, T. Song, Y. Ni, Y. Yang, Z. Liu, M. Yao, B. Shen, W. Li. “Heterogeneity in lung cancers by single-cell DNA sequencing”. *Clinical and Translational Medicine*, vol. 13, issue 9, Aug 2023. DOI: <https://doi.org/10.1002/ctm2.1388>
3. W. Zhang, J. Ju, Y. Zhou, T. Xiong, M. Wang, **C. Li**, S. Lu, Z. Lu, L. Lin, X. Liu, S. C. Li. “IMperm: a fast and comprehensive IMMune Paired-End Reads Merger for sequencing data”. *Briefings in Bioinformatics*, vol. 24, issue 2, Mar 2023. DOI: <https://doi.org/10.1093/bib/bbad080>
4. L. Chen*, Y. Qing*, R. Li*, **C. Li***, H. Li, X. Feng, & S. C. Li. “Somatic variant analysis suite: copy number variation clonal visualization online platform for large-scale single-cell genomics”. *Briefings in Bioinformatics*, vol. 23, issue 1, Jan 2022. DOI: <https://doi.org/10.1093/bib/bbab452>
5. X. Feng, L. Chen, Y. Qing, R. Li, **C. Li**, & S. C. Li. “SCYN: single cell CNV profiling method using dynamic programming”. *BMC Genomics*, vol. 22, Nov 2021. DOI: <https://doi.org/10.1186/s12864-021-07941-3>

WORK EXPERIENCE

University of California San Diego

Graduate Student Researcher, Department of CSE

San Diego, USA

September, 2023 – Present

- Developed algorithms for 3D structure reconstruction of extrachromosomal DNA (ecDNA) from Hi-C data.
- Simulated 3D structures by graphics algorithms, and analyzed data by probability distributions and statistical tests.
- Implemented a toolkit for 3D structure visualization.

City University of Hong Kong

Research Assistant, Department of CS

Hong Kong, China

August, 2021 – June, 2023

- Designed optimization algorithms to reconstruct complex breakage-fusion-bridge (BFB) rearrangements.
- Simulated short and long read sequencing data, and implemented dynamic programming to speed up the algorithms.
- Published a paper in *Nature Communications* as the first author.

CityU Shenzhen Research Institute

Software Engineer, Department of CS

Shenzhen, China

October 2019 – July 2020

- Built an online visualization platform for analysis and visualization of single-cell genomics data.
- Published a paper in *Briefings in Bioinformatics* as a joint first author.

City University of Hong Kong

Teaching Assistant, Department of CS

Hong Kong, China

January – May 2019 & 2020

- Taught undergraduates data structures and algorithms based on C++ programming.
- Designed programming problems for course assignments and tests.

PROJECTS

A Movie Recommender System

September, 2024 – December, 2024

- Built a naive Bayes model with hidden states (types of movie-goers) and partially observed movie ratings.
- Derived the formulas for learning and inference from the model given incomplete data.

- Implemented the expectation-maximization algorithm for movie recommendation.

Transformer in Natural Language Processing

March, 2024 – June, 2024

- Built a transformer encoder in PyTorch, and trained it for speech classification.
- Implemented a word-level, GPT-like transformer decoder, and pretrained it on an autoregressive language modeling task.
- Fine-tuned Transformer-based models, and reviewed their performance on Machine Reading Comprehension datasets.

SurfStore: A Cloud-based File Storage Service

March, 2024 – June, 2024

- Developed a networked file storage application in Go.
- Implemented a fault-tolerant system based on the RAFT protocol.
- Improved the system for multiple clients based on consistent hashing.

HONORS & AWARDS

Silver Award, the 12 th “Challenge Cup” National Business Plan Competition for College Students	December 2020
CityU Scholarship (30,000 HK\$), City University of Hong Kong	March 2020
Dean’s List, College of Engineering, City University of Hong Kong	2018 – 2021
CityU Mainland Student Scholarship (Full Tuition - 120,000 HK\$ per year), City University of Hong Kong	2017 – 2021

SKILLS

Bioinformatics: cancer genomics, structural variations, DNA sequencing, and genome assembly.

Programming Languages: Python, C++, Java, Go, R, and Haskell

Machine Learning: PyTorch, scikit-learn, pandas, and NumPy

Development Tools: Linux, Git, Docker, Jupyter Notebook, Matplotlib, and Seaborn

SOFTWARE

- Ambigram – An algorithm to reconstruct complex BFB rearrangements.
- scSVAS – An online visualization platform for interactive single-cell genomics data analysis.
- Coin Hunter – A multiplayer maze game implemented in Haskell.
- Puzzlengine – A game generator for creating various sliding puzzles with optimal solutions.