Bowen Jin

70 Tillotson Rd, Needham MA 02494 jinbowen@hotmail.com +1 216-702-6928 jin-bowen
#bowennjin ORCID: 0000-0003-1442-6259

Summary

I am a highly skilled computational biologist with extensive experience in multiple programming languages, especially in Python programming and Cromwell workflow. I develop cutting-edge computational tools to study genetic risk and molecular mechanisms in complex diseases. I have successfully led multiple projects, including designing calling pipelines to study somatic mutations, customizing single-cell RNA-seq data to study transcriptional kinetics in Alzheimer's, and modeling genetic variants in 3D protein structures.
I have collaborated closely with bench scientists to test experimental conditions that led to the best sequencing results and worked with computational biologists to develop single-cell somatic mutation caller. My strong organizational skills, along with a proven ability to communicate complex research findings clearly to both technical and non-technical stakeholders, make me well-suited to drive impactful genomics research.

Skills

Technical skills

- Independently design and lead biological research projects
- Comprehensive data analysis, interpretation, and conclusion derivation
- Clear and concise communication of research findings
- First Programming language: Python, R, unix/linux shell, SQL
- Second Programming language: C, Fortran, Perl
- Additional: High performance computing, Cloud computing, Containerization, Cromwell
- NGS data including single-cell RNA-seq, single-cell WGS
- genetic variant calling, somatic mutation calling, variant annotation

Relevant Business and Social skills

- Effective communicator, Lead collaborative teams, Establishes strong relationships
- Innovative, Enthusiastic, Highly motivated
- Highly productive, Consistently generates high quality work, Works Independently
- Excellent presentation skills, Adaptable to different audiences

Research Experience

Brigham and Woman's Hospital, Postdoctoral Research Fellow
Boston, MA
July 2023 – Now

- Lead the data analysis in single neuron somatic mutation study in Alzheimer's disease.
- Contributed to develop single-cell somatic mutation caller

  My work has significantly improved the running time for single-cell somatic mutation calling. This tool successfully calls somatic SNV and indel in 50 neurons within less than two days compared to the previous one-month run time.
- Collaborated with bench scientist to improve a strand-aware somatic mutation calling pipeline.

  I collaborate with bench scientists to improve the sequencing quality by testing out different Protease condition. In addition, I developed an approach to rescue reads with barcode issues due to low levels of spike-in reads. This approach helps rescue multiple experimental data and accommodate the pipeline with variable experimental conditions.

Case Western Reserve University, Graduate Research Assistant — Cleveland, OH
July 2018 – March 2023

- Designed and built the pipeline to study germline genetic risk on protein structure in Alzheimer's disease.

  This approach has significantly increased the power of rare genetic variant studies and identified three AD risk genes by evaluating rare missense variants on their spatial distribution within proteins rather than allele frequency. Applying this method to 5,522 AD cases and 4,919 controls, I screened 5,969 genes with structures from the Protein Data Bank and 18,345 with Alpha Fold2 predicted structures. Our findings highlighted multiple candidate genes, including TREM2, SORL1, and EXOC3L4 that all harbor AD-related genetic variation hotspots.
- Designed and built the pipeline to study transcriptional kinetics with single-cell RNA-seq data in aging and Alzheimer's disease.

  I developed a pipeline to characterize stochastic gene expression by integrating scRNA-seq data with phased genotype profiles. Through this approach, I systematically explored the regulatory impact of 52,000 expression quantitative trait loci (eQTLs) and 150 transcription factor bindings at both the single-cell and single-allelic levels.
- Automated the extraction and processing of UKBB data with SQL and Python.
- Built Perl-based Ensembl plugins to predict premature-stop codon and Nonsense-mediated mRNA decay and apply it with large Alzheimer's disease sequencing consortium data

Education

Ph.D. **Case Western Reserve University**, Systems Biology and Bioinformatics — July 2018 – March 2023

Advisor: William S. Bush, Ph.D.

Dissertation: Transcriptome-wide Study of Transcriptional Kinetics in Human Cells

B.S. **Huazhong University of Science and Technology**, Physics — Sept 2014 – June 2018

Advisor: Shengyou Huang, Ph.D.

Thesis: Simulating protein conformation with coarse-grained models

Publications

Single cell genome and somatic mutation

Jin B, Smirnov D, Kirkham SL, Hennessey EL, Naik SM, Oakley DH, Frosch MP, Hyman BT, Huang AY, Miller MB. Neurons accumulate somatic mutations independently of tau deposition in Alzheimer's disease. (In preparation)

Naik SM, **Jin B**, Kirkham SL, Hennessey EL, Qian X, Zhou Z, Oakley DH, Frosch MP, Hyman BT, Huang AY, Miller MB. Isolation of individual neuronal nuclei by single-cell tau cytopathology in Alzheimer's disease. (In preparation)

Single cell Transcriptional kinetics

Jin B, Feng H, Bush WS. 2022. Allelic Transcription Factor binding shape transcriptional kinetics in human cell lines. bioRxiv

Modeling Rare genetic risk factors in protein structure

Jin B, Capra JA, Benchek P, Wheeler N, Naj AC, Hamilton-Nelson KL, Farrell JJ, Leung YY, Kunkle B, Vadarajan B, et al. 2022. An association test of the spatial distribution of rare missense variants within protein structures identifies Alzheimer's disease–related patterns. Genome Res 32: 778–790.

Li B, **Jin B**, Capra JA, Bush WS. 2022. Integration of Protein Structure and Population-Scale DNA Sequence Data for Disease Gene Discovery and Variant Interpretation. Annual Review of Biomedical Data Science 5: 141–161.

Grunin M, de Jong S, Palmer E, **Jin B**, Rinker D, Moth C, Capra JA, Haines JL, Bush WS, Den Hollander AI. 2022. Spatial modeling of variants in complement genes associated with age-related macular degeneration. Investigative Ophthalmology Visual Science 63: 2627.

Tang Z-Z, Sliwoski GR, Chen G, **Jin B**, Bush WS, Li B, Capra JA. 2020. PSCAN: Spatial scan tests guided by protein structures improve complex disease gene discovery and signal variant detection. Genome Biol 21: 217.

Protein-peptide docking

Zhou P, **Jin B**, Li H, Huang S-Y. 2018. HPEPDOCK: a web server for blind peptide-protein docking based on a hierarchical algorithm. Nucleic Acids Res 46: W443–W450.

Projects

Single neuron somatic mutation in Alzheimer's disease github*

- Contributed to developing a somatic mutation calling pipeline, significantly speeding up mutation calling in single-cell whole genome sequencing data.
- Contributed to developing a strand-aware somatic mutation calling pipeline.
- Tools Used: C++, Python, R, unix/linux shell

Modeling rare genetic risk factors in protein structure github*

- Designed and built the pipeline to study rare germline genetic variants on protein structure for Alzheimer's disease.
- Tools Used: Python, SQL, Docker

Modeling single-cell Transcriptional kinetics github*

- Designed and built the pipeline to study transcriptional kinetics with single-cell RNA-seq data in normal aging and Alzheimer's disease.
- Tools Used: Python

Teaching

- **Fall 2022** A Data-Driven Introduction to Genomics and Human Health, Teaching Assistant/Lecturer, Cleveland OH
- **Summer 2022** Introduction to R Programming, Teaching Assistant/Lecturer, Cleveland OH
- **Summer 2022** Introduction to Scientific Computing, Teaching Assistant/Lecturer, Cleveland OH

Awards

- 2024 Stanley Robbins Memorial Research Fund Award, *Brigham and Women's Hospital*.
- 2022 SYBB Graduate Student Travel Award, *Case Western Reserve University*.
- 2022 Graduate Student Travel Award, *Case Western Reserve University*.

Conference Presentations

ORAL PRESENTATIONS

- **Jin B**, Capra JA, Benchek P, Wheeler N, Naj AC, Hamilton-Nelson KL, Farrell JJ, Leung YY, Kunkle B, Vadarajan B. 2022. Spatial Distribution of Rare Missense Variants Within Protein Structures is Associated with AD Risk.

  Alzheimer's Association International Conference, San Diego CA
- **Jin B**, Feng H, Bush WS. 2022. Allelic-specific transcription factor binding and chromatin accessibility shape promoter

kinetics in human cell lines.

American Society of Human Genetics Annual Meeting, Los Angeles CA

- **Jin B**. 2022. Allelic-specific transcription factor binding and chromatin accessibility shape promoter kinetics in human cell lines.

  Tohoku Symposium, Cleveland OH

POSTER PRESENTATIONS

- **Jin B**, Naik SM, Kirkham SL, Oakley DH, Frosch MP, Hyman BT, Huang AY, Miller MB. 2023. Single-neuron somatic mutation in the context of Alzheimer's disease tau pathology.

  American Society of Human Genetics Annual Meeting, Denver CO

- **Jin B**, Naik SM, Kirkham SL, Qian X, Zhou Z, Oakley DH, Frosch MP, Hyman BT, Huang AY, Miller MB. 2023. Isolation of single neuronal nuclei by phospho-tau accumulation in Alzheimer's disease brain.

  Alzheimer's Association International Conference, Philadelphia PA

- **Jin B**, Bush WS. 2021. Transcriptomic mapping of bursting kinetics in human cells.

  American Society of Human Genetics Annual Meeting, Virtual Meeting

- **Jin B**, Capra JA, Benchek P, Wheeler N, Naj AC, Hamilton-Nelson KL, Farrell JJ, Leung YY, Kunkle B, Vadarajan B. 2020. A protein structure-based association test for studying rare variants. *American Society of Human Genetics Annual Meeting, Houston TX*