

Yi Zhang

+1(217)-419-9634 • yzhan201@illinois.edu • yi-zhang.github.io
2132 Carl R. Woese Institute for Genome Biology, Urbana, IL 61801

EDUCATION

University of Illinois at Urbana-Champaign, Urbana, IL Aug 2014 - May 2019 (expected)
Ph.D. candidate, Bioengineering and Biomedical Engineering. GPA: 3.84/4.0.

- Dissertation: Functional interpretation of cancer-associated genetic variants.
- Advisor: Dr. Jun S. Song.

University of Science and Technology of China, Hefei, China Aug 2010 - Jun 2014
Bachelor of Science, Biosciences.

- Talent Program in Life Sciences, GPA: 3.7/4.3, rank: 3/88.

RESEARCH EXPERIENCE

University of Illinois at Urbana-Champaign, Urbana, IL Jan 2015 - Present
Research Assistant, Carle R. Woese Institute of Genome Biology.

Functional interpretation of breast cancer-associated genetic variants.

- Developed integrative methods for large-scale and multi-omics NGS data analyses of cancer patients.
- Designed haplotype imputation-based statistics on RNA-seq data to reveal variants' allele-specific effects.
- Discovered non-coding variants as modulators of transcription factor binding activity in enhancers.

Functional characterization of a cancer-associated variant in modulating tumor lymphocyte infiltration.

- Proposed a hypothesis that a cancer-protective variant could regulate a gene facilitating immune surveillance.

Interactive web resource for sequencing techniques and machine learning analysis of biomedical data.

- Built an interactive knowledge base of ~70 sequencing techniques (3000+ views in 3 months).
- Built online data visualization tools for clustering and dimensional reduction.

Machine learning models for predicting chromatin interactions and distal gene regulation.

- Predicted enhancer-promoter interaction using Random Forest models and epigenomic data.

University of California at Los Angeles, Los Angeles, CA May 2013 – Aug 2013
Undergraduate Research Assistant, Department of Bioengineering.

- Implemented signal detection algorithms and graphical interfaces for neuromuscular data analysis.

PUBLICATIONS

Y. Zhang*, M. Manjunath*, S. Zhang, D. Chasman, S. Roy, and J.S. Song. "Integrative genomic analysis predicts causative cis-regulatory mechanisms of the breast cancer-associated genetic variant rs4415084." **Cancer Research**, 78(7), 1579-1591, (2018). *co-first authors.

Y. Zhang, M. Manjunath, Y. Kim, J. Heintz, and J.S. Song. "SequencEnG: an Interactive Knowledge Base of Sequencing Techniques." **Bioinformatics**, bty794, (2018).

M. Manjunath, **Y. Zhang**, Y. Kim, S.H. Yeo, O. Sobh, N. Russell, C. Followell, C. Bushell, U. Ravaioli, and J.S. Song. "ClusterEnG: an interactive educational web resource for clustering and visualizing high-dimensional data." **PeerJ Comput. Sci.**, 4:e155, (2018).

Y. Zhang, M. Manjunath, J. Yan, B. Baur, S. Zhang, S. Roy, and J.S. Song. "Can cancer GWAS variants modulate immune cells in the tumor microenvironment?" bioRxiv/2018/493171 (2018), doi: <https://doi.org/10.1101/493171>. Submitted to AJHG.

POSTER PRESENTATIONS

Y. Zhang, M. Manjunath, S. Zhang, D. Chasman, S. Roy, and J.S. Song. "Integrative genomic analysis discovers the causative regulatory mechanisms of a breast cancer-associated genetic variant." **AACR Annual Meeting**, *Chicago, IL* (2018).

Y. Zhang, M. Manjunath, and J.S. Song. "Identifying causal mechanisms of germline risk variants in breast cancer." Bioengineering Graduate Student Symposium, UIUC (2017), Best Poster Award.

Y. Zhang, P. Gad, V. R. Edgerton. "Burst detection in EMG of stepping spinal cord injured rats." Cross-disciplinary Scholars in Science and Technology summer research, No.69, *Los Angeles, CA* (2013).

TECHNICAL SKILLS

Programming: Python, Linux/Bash, R, HTML/CSS, JavaScript, MATLAB, C, TensorFlow, Git.

Bioinformatics: Genotyping array, RNA-seq, DNase-seq, ChIP-seq, ChIA-PET, Hi-C, Statistics, Machine learning.

COURSES AND TRAINING

Related Courses

Mathematical Statistics
Stochastic Processes

Machine Learning
Statistical Learning

Statistical Data Analysis in Physics
Computational Cancer Biology

Course Projects

- *Statistical Learning*: Prediction of Movie votes and campus shooting with gradient boosting models.
- *Machine Learning*: Classification of NGS technique articles using language processing and SVM models.
- *Stochastic Processes*: Segmentation of copy number using HMM based on genotype data in cancer cells.
- *Statistical Data Analysis in Physics*: A cross-entropy method for Cheeger constant estimation.

Study Groups: Deep Learning (2018), Statistical Learning (2017).

Workshop: NHGRI Short Course on NGS: Technology & Statistical Methods, *Birmingham, AL* (2016).

TEACHING EXPERIENCE

Invited Speaker, Bioinformatics Seminar, UIUC	2016, 2018
<ul style="list-style-type: none">• Presented bioinformatics lectures and led discussions among graduate students	
Teaching Assistant, Biomedical Instrumentation Lab, Bioengineering, UIUC	2014 - 2015
<ul style="list-style-type: none">• Led labs and discussions among graduate students.	

HONORS AND AWARDS

Best Poster Award, Bioengineering Graduate Student Symposium, UIUC	2017
Outstanding Student Scholarship, School of Life Sciences, USTC	2013 - 2014
Outstanding Undergraduate Student Honor, USTC	2013
UCLA Cross-disciplinary Scholars in Sci. and Tech. (CSST) Research Scholarship, USTC	2013
Aegon-Industrial Foundation Scholarship, USTC	2012
Panasonic Elite Scholarship, USTC	2011

UNIVERSITY SERVICE

Volunteer, Genome Day, Carl R. Woese Institute for Genomic Biology, UIUC	2017
Vice-chairman of Student Union, School of Life Sciences, USTC	2012 - 2013
Organizer, School Photography Exhibition	2012