

Yi Zhang

+1(217)419-9634 • yiz@ds.dfci.harvard.edu • yi-zhang.github.io

Department of Data Science, Dana-Farber Cancer Institute, Boston, MA 02215 USA

SUMMARY

I am a postdoctoral research fellow focused on biomedical data science. I work with Shirley X. Liu in the Department of Data Science at Dana-Farber Cancer Institute and Harvard University T.H. Chan School of Public Health. My research interest falls at the intersection of biomedical data science, machine learning, cancer genomics and single cell genomics field. I have developed computational methods to study gene regulatory role of disease-associated genetic variants, with current focus on gene regulation in tumor immunity.

ACADEMICS AND EDUCATION

Dana-Farber Cancer Institute, Harvard University, Boston, MA Jul 2019 - Present
Research Fellow, Department of Data Science
Advisor: Dr. X. Shirley Liu

University of Illinois at Urbana-Champaign, Urbana, IL Aug 2014 - May 2019
Ph.D. Bioengineering and Biomedical Engineering

- Dissertation: Computational methods for functional interpretation of cancer-associated genetic variants
- Advisor: Dr. Jun S. Song, Department of Physics, UIUC
- Committee: Dr. Jun S. Song, Dr. Saurabh Sinha, Dr. Dave S. Zhao, and Dr. Pablo Perez-Pinera

University of Science and Technology of China, Hefei, China Aug 2010 - Jun 2014
Bachelor of Natural Science, Biological Science, Talent Program in Life Sciences

PROFESSIONAL EXPERIENCE

Dana-Farber Cancer Institute, Harvard University, Boston, MA Jul 2019 - Present
Research Fellow, Department of Data Science
Advisor: Dr. X. Shirley Liu

- Develop computational methods for integrating single cell genomics to elucidate tumor immunity regulation.
- Cancer Immunologic Data Commons (CIDC) center.

Collaboration with Dr. Myles Brown's lab, DFCI

- Nuclear receptor gene regulation in tumor immunity and immunotherapy efficacy.

Collaboration with Dr. Catherine J. Wu's lab, DFCI

- Transcriptomics analysis for an immunotherapy trial to identify genomic biomarkers in relapsed leukemia.

Collaboration with Dr. David Fisher's lab, Massachusetts General Hospital

- Transcriptomics analysis for discovering immune switch during melanoma timecourse treatment

Collaboration with Dr. Xihong Lin's lab, Harvard University

- Human genetics associated with rare variants using epigenetic data

University of Illinois at Urbana-Champaign, Urbana, IL Jan 2015 - May 2019
Research Assistant, NIH Big Data to Knowledge (BD2K) U54 Center
Advisor: Dr. Jun S. Song
Functional interpretation of breast cancer-associated genetic variants.

- Developed integrative methods for large-scale and multi-omics Next Generation Sequencing (NGS) data analyses for cancer genomics.
- 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.

Machine learning models to predict 3D chromatin structure.

- Built machine learning models using random forest to predict chromatin interaction from epigenetics.

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

- Built an interactive educational web resource of 66 sequencing techniques and analysis pipelines (8.5k users and 17.7k views since Jul 2018).
- Built online data visualization tools for clustering and dimensional reduction.

University of California at Los Angeles, Los Angeles, CA

May 2013 – Aug 2013

Undergraduate Research Assistant, Department of Bioengineering.

- Developed algorithm for burst detection in electromyography signals of stepping spinal cord injured rats.

University of Science and Technology of China, Hefei, CA

May 2013 – May 2014

Undergraduate Research Assistant, Neurophysics Lab.

- Pattern recognition of spatiotemporal neuronal firing in hippocampus on-chip

PUBLICATIONS

Y. Zhang, Y. Liu, X. Shirley Liu. "Neural network architecture search with AMBER". *Nature Machine Intelligence*. (2021) 3(5), 372-373.

L. Penter, **Y. Zhang**, A. Savell, T. Huang, N. Cieri, E.M. Thrash, S. Kim-Schulze, A. Jhaveri, J. Fu, S. Ranasinghe, S. Li, W. Zhang, E.S. Hathaway, M. Nazzaro, H.T. Kim, H. Chen, M. Thurin, S.J. Rodig, M. Severgnini, C. Cibulskis, S. Gabriel, K.J. Livak, C. Cutler, J.H. Antin, S. Nikiforow, J. Koreth, V.T. Ho, P. Armand, J. Ritz, H. Streicher, D. Neuberg, F.S. Hodi, S. Gnjatic, R.J. Soiffer, X.S. Liu, M.S. Davids, P. Bachireddy, C.J. Wu. "Molecular and cellular features of CTLA-4 blockade for relapsed myeloid malignancies after transplantation". **Blood**. (2021) Mar 15: blood.2021010867. doi: 10.1182/blood.2021010867.

S. Gu, W. Zhang, X. Wang, P. Jiang, N. Traugh, Z. Li, C. Meyer, B. Stewig, Y. Xie, X. Bu, M. Manos, A. Front-Tello, E. Gijini, A. Lako, K. Lim, J. Conway, A. Tewari, Z. Zeng, A. Sahu, C. Tokheim, J. Weirather, J. Fu, **Y. Zhang**, B. Kroger, J. Liang, P. Cejas, G. Freeman, S. Rodig, H. Long, B. Gewurz, F. Hodi, M. Brown, and X. S. Liu. "Therapeutically increasing MHC-I expression potentiates immune checkpoint blockade." **Cancer Discovery**, (2021)

M. Manjunath, **Y. Zhang**, S. Zhang, S. Roy, P. Perez-Pinera, J. S. Song. "ABC-GWAS: an interactive web resource for functional annotation of estrogen receptor-positive breast cancer genetic variants." **Frontiers in Genetics, Cancer Genetics**. 11(2020): 730.

M. Manjunath, J. Yan, Y. Youn, K. L. Drucker, T. M. Kollmeyer, A. M. McKinney, V. Zazubovits, **Y. Zhang**, J. F. Costello, P. R. Selvin, R. B. Jenkins, and J. S. Song. "Functional analysis of low-grade glioma genetic variants predicts key target genes and transcription factors." **Neuro-Oncology**. (2020)

S. Gu, X. Wang, X. Hu, P. Jiang, Z. Li, N. Traugh, X. Bu, Qin Tang, C. Wang, Z. Zeng, J. Fu, C. Meyer, **Y. Zhang**, P. Cejas, K. Lim, J. Wang, W. Zhang, C. Tokheim, A. D. Sahu, X. Xing, B. Kroger, Z. Ouyang, H. Long, G. J. Freeman, M. Brown, and X. S. Liu. "Clonal tracing reveals diverse patterns of response to immune checkpoint blockade." **Genome Biology**, 21.1 (2020): 1-28

Wang, X., Tokheim, C., Wang, B., Gu, S.S., Tang, Q., Li, Y., Traugh, N., **Zhang, Y.**, Li, Z., Zhang, B., Fu, J. Xiao, T., Li, W., Meyer, C.A., Chu, J., Jiang, P., Cejas, P., Lim, K., Long, H., Brown, M., Liu, X.S. "In Vivo CRISPR Screens

Identify E3 Ligase Cop1 as a Modulator of Macrophage Infiltration and Cancer Immunotherapy Target.” bioRxiv. Submitted. (2020)

Y. Zhang, M. Manjunath, J. Yan, B. Baur, S. Zhang, S. Roy, and J.S. Song. “The cancer-associated genetic variant rs3903072 modulates immune cells in the tumor microenvironment.” **Frontiers in Genetics, Bioinformatics and Computational Biology**, 10 (2019): 754.

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), NGS web resource of 8.5k users and 17.7k view.

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 Computer Science**, 4:e155, (2018).

CONFERENCES AND POSTER PRESENTATIONS

L. Penter, **Y. Zhang**, A. Savell, S. Ranasinghe, T. Huang, N. Cieri, S. Gohil, W. Zhang, S. Li, R. Zeiser, K. J. Livak, J. Ritz, H. Kim, R. J. Soiffer, S. X. Liu, M. S. Davids, P. Bachireddy, C. J. Wu. “Local and systemic effects of immune checkpoint blockade on relapsed myeloid malignancies following allogeneic hematopoietic stem cell transplantation.” **American Society of Hematology Annual Meeting, Washington, DC** (2020)

Y. Zhang, M. Manjunath, J. Yan, Britany Baur, S. Zhang, S. Roy, and J.S. Song. “A breast cancer-associated genetic variant modulates immune cells in the tumor microenvironment.” **Variant to Function Symposium, Broad Institute of MIT and Harvard, Boston, MA** (2020).

Y. Zhang, M. Manjunath, J. Yan, Britany Baur, S. Zhang, S. Roy, and J.S. Song. “A breast cancer-associated genetic variant modulates immune cells in the tumor microenvironment.” **Program in Quantitative Genomics, Harvard T.H. Chan School of Public Health, Boston, MA** (2019).

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.” **American Association for Cancer Research 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 (CSST) summer research, No.69, *Los Angeles, CA* (2013).

TALKS

“Functional mechanisms of cancer-associated human variants.” *Graduate Student Conference, UIUC*. (2018)

“Prioritizing functional cancer-associated human germline variants.” *UIUC BD2K center talk*. (2017)

PEER REVIEW

Journals: Nature Machine Intelligence, Nature Medicine, Nature Metabolisms, Cell Reports, Biochimica et Biophysica Acta, Frontiers in Immunology, Frontiers in Genetics, PeerJ.

TEACHING EXPERIENCE

Mentorship

- Jiajinlong Kang, *Master student, Harvard University, Computational Biology and Quantitative Genetics*
- Yijia Jiang, *Master student, Harvard University, Master of Biomedical Informatics*
- Yingxiao Shi, *Ph.D. student, Harvard University, Biological and Biomedical Sciences*
- Evan Arnold, *Undergraduate student, Harvard University*

Teaching Assistant, Harvard University 2020

- Developed GWAS series course videos used in STAT115/215 at Harvard University
- Course materials for High Impact Cancer Research Course at Harvard Medical School

Invited Speaker, Bioinformatics Seminar, UIUC 2016, 2018

- Presented bioinformatics lectures and led discussions among graduate students

Teaching Assistant, Biomedical Instrumentation Lab, Bioengineering, UIUC 2014 - 2015

- Led labs and discussions among graduate students.

FUNDING AND AWARDS

DFCI-NIBR Drug Discovery and Translational Research Grant 2021

Best Poster Award, Bioengineering Graduate Student Symposium, UIUC 2017

National Student Innovation Research Grant, USTC 2014

- Pattern recognition of spatiotemporal neuronal firing of cultured hippocampus on-chip

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

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).

Workshops: NHGRI Short Course on NGS: Technology & Statistical Methods, *Birmingham, AL* (2016)
New England Future Faculty Workshop, Northeastern University, *Boston, MA* (2020)

TECHNICAL SKILLS

Programming: Python, R, JavaScript, HTML/CSS, Shell, Julia, MATLAB, C, TensorFlow, Git.

Computation: Bioinformatics, RNA-seq, Genotyping array, ChIP-seq, Hi-C, Statistics, Machine learning.