

# Guoshuai Cai, PhD

Department of Environmental Health Sciences  
Arnold School of Public Health  
University of South Carolina  
Phone: 803-777-4120

915 Greene Street  
Discovery Building 517  
Columbia, SC, USA  
E-mail: GCAI@mailbox.sc.edu

## EDUCATION

---

- Aug 2013    **PhD, Biomathematics and Biostatistics**  
Graduate School of Biomedical Sciences, The University of Texas Health Science Center at Houston and MD Anderson Cancer Center, Houston, TX, USA  
Mentor: Shoudan Liang, PhD  
Dissertation: *RNA-Sequencing applications: Gene Expression Quantification and Methylator Phenotype Identification*
- Jun 2009    **Master of Science, Microbiology**  
State Key Laboratory of Virology  
College of Life Sciences, Wuhan University, Wuhan, China
- Jun 2006    **Bachelor of Science, Biotechnology**  
College of Life Sciences, Wuhan University, Wuhan, China

## PROFESSIONAL EXPERIENCE

---

- Aug 2017    **Assistant Professor**  
- Department of Environmental Health Sciences, Arnold School of Public Health,  
Now        University of South Carolina, Columbia, SC, USA
- Nov 2014    **Postdoctoral Fellow**  
- Department of Molecular and Systems Biology, The Geisel School of Medicine at  
Aug 2017    Dartmouth, Hanover, NH, USA  
Mentor: Michael Whitfield, PhD  
Project I: *Methodology development for RNA-seq differential expression analysis with length bias adjustment.*  
Project II: *Identification and selection of transcriptome markers predicting scleroderma disease severity.*  
Project III: *Methodology to infer intercontinental ancestry using genetic data*  
Project IV: *Methodology to infer subset structures using transcriptome expression profiles*  
Project V: *Population stratification identifies gene expression predictors of*

2020-02-04

*survival outcomes in lung adenocarcinoma for both Caucasian and Asian patients.*

Oct 2013 **Postdoctoral Research Associate**

- Department of Clinical Cancer Prevention, MD Anderson Cancer Center,  
Oct 2014 Houston, TX, USA

Mentor: Qiang Shen, MD, PhD

Project I: *Study of mechanisms in cell of new anticancer drug compounds by high-throughput data analysis.*

Project II: *Identification of genes driving breast cancer metastasis.*

Project III: *A computing platform for comprehensive statistical analysis of large-scale and high-dimensional biomedical data.*

Dec 2009 **Graduate Research Assistant**

- Department of Bioinformatics and Biostatistics, The University of Texas Health  
Jun 2013 Science Center at Houston MD Anderson Cancer Center, Houston, TX, USA

Mentor: Shoudan Liang, PhD

Project I: *Modeling the overdispersion in RNA-seq with its dependence on sequencing depth.*

Project II: *Identification of a new bias resource in sequencing technology and a statistical method for correction.*

Project III: *Identification of a new CpG island methylator phenotype in breast cancer by integrating methylation and mRNA expression.*

Dec 2006 **Graduate Research Assistant**

- Department of Neuroscience, University of Minnesota, Minneapolis, MN, USA

Nov 2007 Supervisor: Tongbin Li, PhD

Project: *Development of miRecords, an integrated resource for microRNA-target interactions.*

## RESEARCH INTERESTS

---

Methodology and Application in:

Single cell genomics

Sequencing data analyses

Disease prediction and marker selection

Time series data analyses

Microbiome in disease

Computational toxicity

Tumor microenvironment

Integrative genomics

Biological structure inference

Cancer and scleroderma

Precision medicine

Machine learning

## PUBLICATIONS

---

Published or in press (\*(co-)corresponding author):

1. Huang M, **Cai G**, Baugh LM, Liu Z, Smith A, Watson M, Popovich D, Zhang T, Stawski LS, Trojanowska M, Georgakoudi I, Black LD 3rd, Pioli PA, Whitfield ML, Garlick J. Systemic Sclerosis Dermal Fibroblasts Induce Cutaneous Fibrosis Through LOXL4: New Evidence from 3D Skin-like Tissues. *Arthritis Rheumatol.* (SJR ranking the 3rd in Rheumatology) 2019 Nov 9. doi: 10.1002/art.41163.
2. Franks JM, Martyanov V, **Cai G**, Wang Y, Li Z, Wood TA, Whitfield ML. A Machine Learning Classifier for Assigning Individual Patients with Systemic Sclerosis to Intrinsic Molecular Subsets. *Arthritis Rheumatol.* (SJR ranking the 3rd in Rheumatology) 2019 Mar 28. doi: 10.1002/art.40898.
3. Johnson ME, Franks JM, **Cai G**, Mehta BK, Wood TA, Archambault K, Pioli PA, Simms RW, Orzechowski N, Arron S, Whitfield ML. Microbiome dysbiosis is associated with disease duration and increased inflammatory gene expression in systemic sclerosis skin. *Arthritis Res Ther.* (SJR ranking the 7th in Rheumatology) 2019 Feb 6;21(1):49. doi: 10.1186/s13075-019-1816-z.
4. Xiao F, Luo X, Hao N, Niu YS, Xiao X, **Cai G**, Amos CI, Zhang H. An Accurate and Powerful Method for Copy Number Variation Detection. *Bioinformatics.* (Google Scholar ranking the 1st in Bioinformatics & Computational Biology) 2019 Jan 14. doi:10.1093/bioinformatics/bty1041.
5. Franks JM, **Cai G**, and Whitfield ML. Feature specific quantile normalization enables cross-platform classification of molecular subtypes using gene expression. *Bioinformatics.* (Google Scholar ranking the 1st in Bioinformatics & Computational Biology) 2018 Jun 1;34(11):1868-1874. doi: 10.1093/bioinformatics/bty026.
6. Messemaker TC, Chadli L, **Cai G**, Goelela VS, Boonstra M, Dorjée AL, Andersen SN, Mikkers HM, van t Hof P, Mei H, Distler OMD, Draisma HHM, Li Z, Toes REM, Aarbiou J, Huizinga TW, Whitfield M, DeGroot J, de Vries-Bouwstra JK, Kurreeman Fina. Antisense long non-coding RNAs are deregulated in skin tissue of patients with systemic sclerosis. *Journal of Investigative Dermatology.* (SJR ranking the 4th in Rheumatology) 2018 Apr;138(4):826-835. doi: 10.1016/j.jid.2017.09.053.
7. Zhao Y, Varn F, **Cai G**, Xiao F, Amos CI, and Cheng C. A P53-deficiency gene signature predicts recurrence risk of patients with early stage lung adenocarcinoma. *Cancer Epidemiology, Biomarkers & Prevention.* (SJR ranking the 22rd out of 368 in Oncology) 2018 Jan;27(1):86-95.
8. **Cai G**, Zheng X, Liang S, Xiao F. Local sequence and sequencing depth dependent accuracy of RNA-seq reads. *BMC Bioinformatics.* (Google Scholar ranking the 3rd in Bioinformatics & Computational Biology) 2017 Aug 9;18(1):364.
9. **Cai G**, Xiao F, Cheng C, Li Y, Amos CI, Whitfield ML. Population stratification identifies gene expression predictors of survival outcomes in lung adenocarcinoma for both Caucasian and Asian patients. *Plos One.* 2017 Apr 20;12(4):e0175850.
10. Xiao F, **Cai G (co-first author)**, Zhang H. Segregation Analysis Suggests That a Genetic Reason May Contribute to "the Dress" Colour Perception. *Plos One.* 2016 Oct 21;11(10):e0165095.
11. Li Y, Byun J, **Cai G (co-first author)**, Xiao X, Dennis J, Easton D, Gorlov I, Seldin M, Amos CI. FastPop: a rapid principle component derived method to infer intercontinental ancestry using genetic data. *BMC Bioinformatics.* (Google Scholar ranking the 3rd in Bioinformatics & Computational Biology) 2016 Mar 9;17:122.

12. Bu Y, **Cai G**, Shen Y, Huang C, Cao Y, Liao D, Cao D. Targeting NF- $\kappa$ B RelA/p65 Phosphorylation Overcomes RITA resistance. *Cancer Letters*. 2016 Oct 6;383(2):261-271.
13. Zhou X, Ren Y, Kong L, **Cai G (co-first author)**, Sun S, Song W, Wang Y, Jin R, Qi L, Mei M, Wang X, Kang C, Li M, Zhang L. Targeting EZH2 regulates tumor growth and apoptosis through modulating mitochondria dependent cell-death pathway in HNSCC. *Oncotarget*. 2015 Oct 20;6(32):33720-32.
14. Zhou X, Liu S, **Cai G (co-first author)**, Kong L, Zhang T, Ren Y, Wu Y, Mei M, Zhang L, and Wang X. Long Non Coding RNA MALAT1 Promotes Tumor Growth and Metastasis by inducing Epithelial-Mesenchymal Transition in Oral Squamous Cell Carcinoma. *Scientific report*. (SJR ranking the 6th out of 120 in Multidisciplinary) 2015 Nov 2; 5:15972.
15. Li H, Tong P, Gallegos J, Dimmer E, **Cai G**, Mollidrem J.J, Liang S. PAND: a distribution to identify functional linkage from networks with preferential attachment property. *Plos One*. 2015; 10(7).
16. Lee J, Ji Y, Liang S, **Cai G**, Müller P. Bayesian Hierarchical Model for Differential Gene Expression Using RNA-seq Data. *Statistics in Biosciences*. 2015; 7(1).
17. Tang SH, Wang X, Shen Q, Yang X, Yu C, Cai C, **Cai G**, Xu L, Meng X, and Zou F. Mitochondrial Ca<sup>2+</sup> uniporter is critical for store-operated Ca<sup>2+</sup> entry-dependent breast cancer cell migration. *Biochem Biophys Res Commun*. 2015; 458(1):186-93.
18. Chen H, Yang Z, Ding C, Xiong A, Wild C, Wang L, Ye N, **Cai G**, Flores RM, Ding Y, Shen Q, Zhou J. Eur J Med Chem. Discovery of potent anticancer agent HJC0416, an orally bioavailable small molecule inhibitor of signal transducer and activator of transcription 3 (STAT3). *Eur J Med Chem*. (SJR ranking the 23rd in Drug Discovery) 2014; 82:195-203.
19. Xiao F, Ma J, **Cai G**, Amos CI. Natural and Orthogonal interaction framework for modeling gene-gene interactions applied to cutaneous melanoma. *Human Genetics*. (SJR ranking the 40<sup>th</sup> out of 338 in Genetics) 2014; 133(5):559-74.
20. **Cai G**, Li H, Lu Y, Huang X, Lee J, Müller P, Ji Y, Liang S. Accuracy of RNA-Seq and its dependence on sequencing depth. *BMC Bioinformatics*. (Google Scholar ranking the 3rd in Bioinformatics & Computational Biology) 2012; 13:S5.
21. Lee J, Ji Y, Liang S, **Cai G**, Müller P. On differential gene expression using RNA-Seq data. *Cancer Inform*. (Google Scholar ranking the 19th in Bioinformatics & Computational Biology) 2011; 10:205-15.
22. Xiao F, Zuo Z, **Cai G**, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res*. (SJR ranking the 6<sup>th</sup> out of 338 in Genetics) 2009; 37:D105-10.
23. Kim Y, Xiao H, Du E, **Cai G**, Lu S, Qi Y. Identification and functional analysis of LsMNPV anti-apoptosis genes. *J Biochem Mol Biol*. 2007; 40(4):571-6.

Pre-print (\*(co-)corresponding author):

24. **Cai G\***, Xiao F. SCANNER: A Web Server for Annotation, Visualization and Sharing of Single Cell RNA-seq Data. *bioRxiv*. 2020 Jan. doi: 10.1101/2020.01.25.919712
25. **Cai G\***, Franks JM, Whitfield ML. RNA-seq analyses of molecular abundance (RoMA) for detecting differential gene expression. *bioRxiv*. 2018 Sep. doi: 10.1101/410985

Under review or submitted (\*(co-)corresponding author):

26. **Cai G\***. Tobacco-use disparity in gene expression of ACE2, the receptor of 2019-nCov. *Tobacco Control*. 2020 Feb. Submitted.
27. Ek L, Porter D, Graves D, Cai B, **Cai G**, Scott GI. Enterococci Contamination on Edisto Beach, South Carolina: Frequency, Sources of Contamination and Prospects for Water Quality Improvement. *Frontiers in Public Health*. Revision Submitted
28. Wang L, Luo X, Cheng C, Amos CI, **Cai G\***, Xiao F. A gene expression based immune signature for lung adenocarcinoma prognosis. *Journal of Thoracic Oncology*. Submitted.
29. Zhang L, Cao W, Majewski T, Yuan Y, Young K, Broom B, **Cai G**, Lu K, Sun S, Czerniak B, Weinstein J, Zhang L. Somatic DNA copy number variants in the peripheral blood of patients with solid tumors. Under Revision.
30. **Cai G\***, Xiao F. RNA-seq differential expression detection and its study specific strategy. Under revision.
31. **Cai G**, Qiu P, Verhaak RG, Chen T, Xiao F, Liang S. Expression regulatory methylator phenotype in breast cancer. Under revision.
32. **Cai G\***, Xiao F, Liang S. A new bias in RNA-seq. Under revision.
33. Yang Z, **Cai G (co-first author)**, Chen H, Ding C, Zhang Y, Xiong A, Ester C, Wild C, Zhou J, and Shen Q. Orally active small-molecule inhibitor HJC0152 exhibits potent anticancer activity and suppresses the growth of triple-negative breast cancer xenograft tumors via inhibiting STAT3/JAK signaling. Under revision.

Manuscript (\*(co-)corresponding author):

34. Xu D, Xiao F, **Cai G\***. Pan-cancer analysis of CDKs: expression dysregulation and survival association. Manuscript.
35. Yuan C, Ni S, Xiao F, **Cai G\***. Burden and risk factors of hypertension in east asian: systematic review and meta-analysis. Manuscript.
36. Ni S, Xiao F, **Cai G\***. Assessment of statistical batch effect correction methods for different RNA sequencing data analysis strategies. Manuscript.
37. Zheng X, Lu Y, **Cai G**, Zhang J, Liang S. MTM: An error-correction method for high-throughput sequencing data. Manuscript.
38. **Cai G\***, Franks J, Whitfield M. Subset calibration through concordance of gene expression profiles. Manuscript.
39. **Cai G**, Yang Z, Zhang Y, Jiang S, Price J, Brown P, Shen Q. Role of the AP-1 transcription factor in promoting breast cancer metastasis. Manuscript.
40. **Cai G**, Shen Q. Comparison of STAT3 inhibitors Niclosamide, HJC0152, HJC0123, HJC0416 and Stattic. Manuscript.
41. **Cai G**, Shen Q. Networks involved in the metastatic phenotype of highly metastasis variants of breast cancer cells. Manuscript.

## RESEARCH AWARDS AND GRANTS

---

### Active/ Ongoing Research Support

09/13/2019 – 09/12/2020

NSF XSEDE Startup Allocation (PI: Cai, G)

Title: *A Comprehensive Annotator and Web Viewer for scRNA-seq Data*

2020-02-04

Role: Principle Investigator

07/01/2019 - 09/30/2020

SC ASPIRE-I (PI: Cai, G)

Title: *Characterization of a Novel Methylation Signature in Breast Cancer Progression*

Role: Principle Investigator

07/01/2019 - 06/30/2020

SC INBRE - Bioinformatics Pilot Project/NIH (PI: Sarzynski, M)

Title: *miRNA bioinformatics of peak VO2 response to exercise training in heart failure*

Role: Co-Investigator

08/16/2019 - 08/15/2022

SC Excellent Initiative (PI: Li X)

Title: *Big Data Health Science Center*

Role: Co-Investigator

### **Completed Research Support**

05/01/2018 - 04/30/2019

NIH P20, Centers of Biomedical Research Excellence Pilot Study (PI: Cai, G)

Title: *Transcriptomic Data Analysis to Characterize Inhibitors of Transcription-regulating Kinases*

Role: Principle Investigator

### **Pending**

03/01/2020-02/28/2022

Chan Zuckerberg Initiative (contacting PI: Cai, G; co-PIs: Lead, J and Murphy EA)

Title: *Macrophage Inflammatory Responses: single cell uptake and transcriptome*

Role: Contacting Principle Investigator

07/01/2020-06/30/2025

NIH R01 (PI: Hofseth, L)

Title: *Identifying and fixing molecular anomalies in African American adenomas to reduce colon cancer disparities*

Role: Co-Investigator

07/01/2020 - 06/30/2022

NIH R21 (PI: Xiao, F)

Title: *Copy Number Variation and Lung Cancer: Disease Risk and Mechanisms*

Role: Co-Investigator

07/01/2020 - 06/30/2025

NIH R01 (PI: Broude, E)

Title: *Preventing the emergence of resistance to endocrine therapy by inhibiting transcriptional reprogramming*

Role: Co-Investigator

2020-02-04

## **Unfunded**

04/01/2020 - 03/31/2025

NIH R01 (PI: Steck, S)

Title: *Dietary Inflammatory Potential, the Colonic Environment, and Adenoma Risk*

Objective: The objective of this project is to conduct a case-control study nested among a cohort of African American and European American screening colonoscopy patients to determine the association between inflammatory diet and the colonic environment, including inflammation, bacterial profile, bacterial gene expression, and microbial metabolites, and to examine whether fecal microbiota, their transcripts, and/or metabolites are associated with colorectal adenoma risk.

Role: Co-Investigator

04/01/2020 - 03/31/2022

NIH R21 (PI: Cai, G)

Title: *Complete Bulk Cell Population Deconvolution Harnessing Single Cell RNA-seq Data*

Role: Principle Investigator

07/01/2019 - 06/30/2020

NIH P20, Centers of Biomedical Research Excellence Pilot Study (PI: Cai, G)

Title: *Characterization of Inflammation-Methylation-Expression Axis in Colorectal Tumors*

Role: Principle Investigator

04/01/2020 - 03/31/2025

SPORE, NIH/NCI (PI: Hébert JR)

Title: *Specialized Program of Research Excellence (SPORE) in Racial Disparities in Cancer*

Role: Co-Investigator, Biostatistics, Bioinformatics, and Data Management (BBDM) CORE co-leader

09/01/2019 - 08/31/2022

DOE (PI: Norman, R)

Title: *systems approach to understanding sediment biogeochemical cycling under changing environmental conditions*

Role: Co-PI

01/01/2020 - 12/31/2022

NSF (PI: Lead, J)

Title: *Nanoparticle transformations and biouptake on a single-particle and single-cell basis*

Role: Co-PI

09/01/2019 - 08/31/2023

NIH R01 (PI: Cui, T)

Title: *Vascular Stem Cells in Vein Graft Failure regeneration*

Role: Co-Investigator

12/01/2019 - 11/30/2024

2020-02-04

NIH R01 (PI: Xiao, S)

Title: *Development of a high-throughput in vitro and in silico platform for screening female reproductive toxic chemicals*

Role: Co-Investigator

12/01/2019 - 11/30/2021

NIH R21 (PI: Xiao, F)

Title: *Copy Number Variation and Lung Cancer: Mechanisms and Disease Risk*

Role: Co-Investigator

07/01/2019 - 06/30/2023

NIH R01 (PI: Cui, T)

Title: *Mapping the cellular and molecular atlas of American ginseng-induced vessel regeneration*

Role: Co-Investigator

07/01/2019 - 06/30/2023

SC ASPIRE-II (PI: Xiao, S)

Title: *Side Effect of Epigenetic Anti-Cancer Drugs on Female Reproductive Health and Fertility*

Role: Co-Investigator

07/01/2019 - 06/30/2020

NASA (PI: Gower M)

Title: *The impact of biomaterial implants on simulated microgravity induced adipose tissue gene expression and muscle dysfunction*

Role: Co-Investigator

04/01/2019 - 03/31/2022

Environmental Protection Agency (PI: Xiao S)

Title: *Development of a high-throughput in vitro and in silico platform for female reproductive toxicity evaluation*

Role: Co-PI

08/16/2019 - 08/15/2022

SC Excellent Initiative (PI: Decho A and Tang CB)

Title: *Targeted Antimicrobial Therapeutics Pipeline (TATP) to Avert the Emerging Antibiotic-Resistance Crisis!*

Role: Co-Investigator

08/16/2019 - 08/15/2022

SC Excellent Initiative (PI: Hofseth LJ and Berger FG)

Title: *Early Onset Colorectal Cancer (EOCRC): A PPG to support our ongoing mission at the Center for Cancer Research at the University of South Carolina*

Role: Co-Investigator

2020-02-04

04/01/2019 - 03/31/2024

NIH/National Cancer Institute (PI: Wirth M)

*A Dietary Inflammatory Index-Based Intervention to Improve the Colonic Environment and Systemic Inflammatory Profile in African Americans: A Randomized Controlled Trial*

Role: Co-Investigator

12/01/2018 - 11/30/2023

NIH/National Cancer Institute, R01 (PI: Murphy, E)

Title: *Dietary Inflammatory Potential, the Colonic Environment, and Adenoma Risk*

Role: Co-Investigator

12/01/2018 - 11/30/2020

NIH/National Cancer Institute, R21 (PI: Xiao F)

Title: *Germline Copy Number Variations in Lung Cancer*

Role: Co-Investigator

04/01/2019 - 03/31/2024

NIH/National Cancer Institute, P50 (PI: Woodruff, TK, Duncan, FE)

Title: *Effect of epigenetic modifier on reproductive health*

Role: Project Co-PI

11/01/2018 - 10/30/2019

NIH COBRE pilot program (PI: Moore, D)

Title: *A Pilot Investigation of the Efficacy of Cannabidiol for Reducing Inflammation in Patients with Post-Concussion Syndrome*

Role: Co-Investigator

07/01/2018 - 09/30/2019

SC ASPIRE-I (PI: Cai, G)

Title: *Optimized Method and Pipeline for RNA-seq Differential Expression Analysis*

Role: Principle Investigator

07/01/2018 - 06/30/2023

P01 CA190178-01A1, NIH/NCI (PI: Hébert JR)

Title: *Center for Inflammation, Biobehavioral Risk Factors, and Racial Disparities in Colorectal Cancer.*

Role: Co-Investigator, BIOSTATISTICS, BIOINFORMATICS, AND DATA MANAGEMENT CORE co-leader

09/15/2018 - 09/14/2021

DOE (PI: Chanda A)

Title: *Developing a systems level understanding of fungal bioremediation using integrated omics*

Role: Co-PI

04/30/2018 - 05/31/2019

2020-02-04

SC ASPIRE-III (PI: Shtutman, M)

Title: *Advancing USC transcriptomics research by single cell sequencing*

Role: Co-PI

04/01/2018 -

NICHD (PI: Murphy, EA, Nitin S, Michael W)

Title: *Children's dietary inflammatory index-based intervention and products to promote anti-inflammatory-dietary patterns to reduce inflammation and related conditions in children*

Role: Co-Investigator

04/01/2018 - 03/31/2020

NIH/National Cancer Institute, R21 (PI: Xiao F)

Title: *Germline Copy Number Variations in Lung Cancer*

Role: Co-Investigator

## RESEARCH ADVISING

---

- PhD Dissertation Advisee

Megha Oza, ENHS, ASPH, UofSC, 2019-now. Title: *Development of a Time- and Cost-Efficient Single Cell Multi-omics Measurement Platform.*

Jacob Barber, EXSC, ASPH, UofSC, 2019-now (Secondary Mentor). Title: *Research on Associations between the Circulating Omics and Lipoprotein Traits in Response to Regular Endurance Exercise.*

- PhD Dissertation Committee

Ethan Older, Department of Chemistry and Biochemistry, UofSC, Advisor: Jie Li, 2020-now.  
Yanding Zhao, Department of Molecular and Systems Biology, Dartmouth Geisel School of Medicine, NH, Advisor: Chao Cheng, 2019 Title: *Innovative Frameworks for Elucidating Cancer Mechanisms and Improving Therapy.*

Karlen Correa Velez, ENHS, ASPH, UofSC, Advisor: Sean Norman, 2019-now. Title: *Influence of Environmental Change and Municipal Sewage on Vibrio Vulnificus and Vibrio Parahaemolyticus Virulence and Antibiotic Resistance.*

Maria Zubizarreta, ENHS, ASPH, UofSC, Advisor: Shuo Xiao, 2019-now.

Xizhi Luo, BIOS, ASPH, UofSC, Advisor: Feifei Xiao, 2019-now. Title: *Novel Methods Development of Copy Number Variation Detection from High-throughput Data.*

Kara Clyburn, ENHS, ASPH, UofSC, Advisor: Geoff Scott, Dwayne Porter and Erik Smith, 2019-now. Title: *Nutrient Limitation of Phytoplankton in Lake Wateree, South Carolina: Implications for Future Water Quality Management.*

Ben Torkian, ENHS, ASPH, UofSC, Advisor: Sean Norman, 2017-now. Title: *Analyzing Correlations between Microbial Biological Diversity and Ecosystems with the Application of a Finite State Automaton (FSA) Computational Model.*

Badria Almurshidi, ENHS, ASPH, UofSC, Advisor: Geoff Scott & Jamie Lead, 2018-now. Title: *Selenium, Platinum and Cerium Oxide Nanoparticles Applications in Biomedicine Especially in Metabolic Diseases such as Cancer*.

Daniel Kilpatrick, ENHS, ASPH, UofSC, Advisor: Dwayne Porter, 2018. Title: *Investigating the Relationship of COPD, Lung Cancer, and Polycyclic Aromatic Hydrocarbons from Ambient Air Pollution*.

Rajaa Alyassein, ENHS, ASPH, UofSC, Advisor: Geoff Scott, 2017-2019 Title: *The Effects of Climate Change on the Ecotoxicology of Contemporary Use Pesticides and Pharmaceuticals & Personal Care Products on the Estuarine Grass Shrimp, Palaemonetes Pugio*.

- Master Thesis Committee

Lala Ek, ENHS, ASPH, SC, Advisor: Geoff Scott, 2019-now. Title: *The Impact of Chlorella Biofertilizer on Crop Production and Fungal Pathogen*.

Mirza, Isanovic, ENHS, ASPH, SC, Advisor: Sean Norman, 2019-now. Title: *The presence of antibiotic resistant bacteria in wastewater treatment plants, the aerosolization of ABRs, and the potential risk of exposure for the WWTP workers to these ABRs*.

- Undergraduate Thesis Advisee

Chenxi Yuan, NJMU, NJ, 2019, Title: *Prevalence and Risk Factors of Hypertension in East Asian Population: A Systematic Review and Meta-Analysis*.

Senmiao Ni, NJMU, NJ, 2019, Title: *Assessment of Statistical Batch Effect-correction Methods for RNA-seq Data*.

Dian Xu, NJMU, NJ, 2018, Title: *Pan-cancer Analysis of CDKs: Expression and Cancer Association*.

Lijuan Wang, NJMU, NJ, 2017, Title: *Differential Expression of Immunology Pathway Related Genes in Lung Carcinoma*.

## Mentorship

---

- Current Lab Members

Megha Oza, PhD student, UofSC, 2019-now. Project: *Development of a Time- and Cost-Efficient Single Cell Multi-omics Measurement Platform*.

Wei Xiao, visiting scholar, Tongji Medical College of HUST, 2019-now. Project: *Transcriptomics Analysis of Perioperative Circulating Tumor Cells in Patients with Hepatocellular Carcinoma*.

Xiangyu Hu, PhD student, 2019-now. Project: *A Machine Learning Method for Predicting 3D Spatial Location of Individual Cells from Single Cell Transcriptomics*.

Xizhi Luo, PhD candidate, UofSC, 2018-now. Project: *A New Bayesian Method to Estimate Gene Bursting from Single Cell RNA-seq Data*.

2020-02-04

Fei Qin, PhD student, UofSC, 2018-now. Project: *A New Statistical Simulator for Single Cell RNA-seq Data*.

- Former Lab Members

Huizhong Yang, MS Candidate, 2019.

Chenxi Yuan, visiting undergraduate student, 2019 (now PhD student in Nanjing Medical University).

Senmiao Ni, visiting undergraduate student, 2019 (now PhD student in Nanjing Medical University).

Rongjie Huang, PhD student, 2018.

Yuan Hong, PhD candidate, 2018 (will start her work as a statistician in Bayer soon).

Dian Xu, visiting undergraduate student, 2018 (now PhD student in Nanjing Medical University).

- Incoming Lab Members

Najm Alsadat Madani, PhD student, UofSC, 2020.

Xia Zhu, visiting undergraduate student, Nanjing Medical University, 2020.

Jun Zhou, visiting undergraduate student, Nanjing Medical University, 2020.

Xiang Cui, visiting undergraduate student, Nanjing Medical University, 2020.

## TEACHING

---

- Arnold School of Public Health, University of South Carolina:

Course:

ENHS/ENVR 321 *Environmental Pollution and Health*, Spring 2020

ENHS 793 *Bioinformatics and Computational Biology*, Fall 2019

ENHS/ENVR 321 *Environmental Pollution and Health*, Spring 2019

ENHS 793 *Bioinformatics and Computational Biology*, Fall 2018

ENHS/ENVR 321 *Environmental Pollution and Health*, Spring 2018

Guest Lecture:

*RNA-seq Data Analysis*. A Lecture for BIOS 794, *Biostatistical Modeling in Genetic Data Analysis*, Fall 2019.

- Geisel School of Medicine at Dartmouth:

Workshop:

*High Performance Computing and Sequencing Data Analysis*, Oct 2016

## SOFTWARES AND TOOLS

---

- **SCANNER**: Single Cell RNA-seq Annotated Viewer. <https://thecailab.com/scanner/>
- **miRecords**: an integrated resource for microRNA-target interactions. <http://cl accurascience.com/miRecords/>
- **FastPop**: a rapid principle component derived method to infer intercontinental ancestry using genetic data. <https://sourceforge.net/projects/fastpop/files/>
- **GCAI.bias**: a guided correction approach for sequencing inherent bias. <https://cran.r-project.org/web/packages/GCAI.bias/>
- **RoMA**: RNA-seq Analysis on Molecular Abundance. <https://thecailab.com/RoMA/>
- **BBDG**: RNA-seq DE analysis Beta Binomial modeling on base pairs for RNA-seq differential expression analysis. <https://thecailab.com/BBDG/>
- **ssGSEA-seq**: single sample gene set enrichment analysis for RNA-seq data. (available upon request)
- **gladR**: gene length adjusted differential expression analysis for RNA-seq data. (available upon request)
- **BONE**: a biomedical big data comprehensive analysis system. (available upon request)

## INVITED TALKS

---

1. *Taking the Next Step in Transcriptome Data Analysis: from Bulk to Single Cell*. Environmental Health Sciences Seminar Series, University of South Carolina, Columbia, SC, USA, Sep 2019.
2. *Introduction to Genomics, Bioinformatics and Statistical Genetics*. China Pharmaceutical University, Nanjing, China, July 2019.
3. *Variations in RNA Sequencing: from Bulk to Single Cell*. Zhongnan Hospital of Wuhan University, Wuhan, China, July 2019.
4. *Modeling Transcriptional Bursting by Single Cell RNA-seq Data*. Dartmouth College, NH, USA, Jun 2019.
5. *Transcriptomic Data Analysis to Characterize Inhibitors of Transcription-regulating Kinases*. Annual meeting of the COBRE Center for Targeted Therapeutics (CTT), University of South Carolina, Columbia, SC, USA, Dec 2018.
6. *Sequencing Depth, Local Sequence and RNA-seq Differential Analysis*. Nanjing Medical Univeristy, Nanjing, China, Jun 2018.
7. *Dispersion in RNA-seq Differential Analysis: Characterizing and Modeling*. China Pharmaceutical University, Nanjing, China, Jun 2018.
8. *Modeling Mean-variance Dependency in RNA-seq Differential Analysis*. Fudan University, Shanghai, China, Jun 2018.

9. *Transcriptome Sequencing Reveals Genetic Polymorphisms Associated with SSc Specific Gene Expression Subtypes, Pathways, and Phenotypes*. Dartmouth College, NH, USA, Jun 2018.
10. *Methods of RNA-seq Differential Expression Analysis and the Application in Studying Human Diseases*. Environmental Health Sciences Seminar Series, University of South Carolina, Columbia, SC, USA, Feb 2018.
11. *Informatic Analysis of Transcriptomic Data for Cancer Drug Targeting*. Annual all-day meeting of the COBRE Center for Targeted Therapeutics (CTT), University of South Carolina, Columbia, SC, USA, Dec 2017.
12. *Integrated Genome-wide Analysis of DNA Methylation and RNA Expression*. School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei, China, Jul 2017.
13. *Steps to Precise Disease Subtyping*. Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, SC, USA. Feb 2017.
14. *Differential Expression and Methylation-expression Integration Analyses in RNA-seq*. University of South Carolina Statistics Department Colloquium, Columbia, SC, USA. Mar 2016.
15. *Problems and Methods of RNA-seq Data Differential Expression Analyses*. Genomics Working Group. Dartmouth College, NH, USA. Apr 2015.

## **CONFERENCE PRESENTATIONS AND ABSTRACTS**

---

1. *A Transcriptome Based Nonparametric Method to Deconvolute Immune Cells and Cancer Subtypes*, Invited Talk. The 11th ICSA International Conference, Hangzhou, China, Dec 2019
2. *Computational Methods for Drug Repositioning of Systemic Sclerosis using Gene Fold-Change and Network Analyses*. Abstract. American College of Rheumatology (ACR) Annual Meeting, Atlanta, Georgia, Nov 2019.
3. *Assessment of Statistical Batch Effect Correction Methods for Different RNA Sequencing Data Types*. Poster Presentation. The American Society of Human Genetics (ASHG) 2019 Annual Meeting, Houston, TX, Oct 2019.
4. *An accurate and powerful method for copy number variation detection*. Poster Presentation. The American Society of Human Genetics (ASHG) 2019 Annual Meeting, Houston, TX, Oct 2019.
5. *Precision Weighted RNA-seq Analyses of Molecular Abundance (RoMA) for Detecting Differential Gene Expression*. Poster Presentation. The American Association for Cancer Research (AACR) 2019 Annual Meeting, Atlanta, GA, Mar 2019.
6. *A Machine Learning Classifier for Assigning Individual Patients with Systemic Sclerosis to Intrinsic Molecular Subsets*. Abstract. International Society for Computational Biology (ISCB) the 16<sup>th</sup> annual Rocky Mountain Bioinformatics Conference, Aspen/Snowmass, Colorado, Dec 2018.

7. *A Machine Learning Classifier for Assigning Individual Patients with Systemic Sclerosis to Intrinsic Molecular Subsets*. Abstract. American College of Rheumatology (ACR) Annual Meeting, Chicago, Illinois, Oct 2018.
8. *Multi-Organ RNA-Sequencing of Patients with Systemic Sclerosis (SSc) Finds That Intrinsic Subsets Are Conserved across Organ Systems*. Abstract. American College of Rheumatology (ACR) Annual Meeting, Chicago, Illinois, Oct 2018.
9. *Precision Weighted RNA-seq Analyses of Molecular Abundance (RoMA) for Detecting Differential Gene Expression*. Contributed Talk. 2018 Joint Statistical Meeting (JSM), Vancouver, Canada, Aug 2018.
10. *Integration of DNA Methylation and mRNA Expression Identify a New Regulatory Methylator Phenotype in Breast Cancer*. Invited Talk. The 8th International Forum on Statistics (IFS), Beijing, China, July 2018.
11. *Precision Weighted RNA-seq Analyses of Molecular Abundance (RoMA) for Detecting Differential Gene Expression*. Poster Presentation. American Association for Cancer Research (AACR) Annual Meeting, Chicago, Illinois, Apr 2018.
12. *Transcriptome Sequencing Reveals Genetic Polymorphisms Associated with SSc Gene Expression Subtypes*. Poster Presentation. American College of Rheumatology (ACR) Annual Meeting, San Diego, California, Oct 2017.
13. *Clinically Relevant Serum Proteins in Patients with Early Diffuse Cutaneous Systemic Sclerosis*. Abstract. American College of Rheumatology (ACR) Annual Meeting, San Diego, California, Oct 2017.
14. *Novel Machine Learning Classifier Accurately Predicts Intrinsic Molecular Subsets for Patients with Systemic Sclerosis*. Abstract. American College of Rheumatology (ACR) Annual Meeting, San Diego, California, Oct 2017.
15. *Multi-Organ RNA-Sequencing of Patients with Systemic Sclerosis (SSc) Finds That Intrinsic Subsets Are Conserved across Organ Systems*. Abstract. American College of Rheumatology (ACR) Annual Meeting, San Diego, California, Oct 2017.
16. *Precision Weighted RNA-seq Analyses of Molecular Abundance (RoMA) for Detecting Differential Gene Expression*. Poster Presentation. The International Genetic Epidemiology Society (IGES) Annual Meeting, Cambridge, UK. Sep 2017.
17. *Transcriptome Sequencing Reveals Genetic Polymorphisms Associated with SSc Gene Expression Subtypes*. Poster Presentation. The 15th international workshop on scleroderma research, Pittsburgh, PA, USA. Aug 2017.
18. *Novel Machine Learning Classifier Accurately Predicts Intrinsic Molecular Subsets for Patients with Systemic Sclerosis*. Abstract. The 15th international workshop on scleroderma research, Pittsburgh, PA, USA. Aug 2017.
19. *In Vitro Skin Models Mimic Fibrogenic Signatures of Systemic Sclerosis*. Abstract. The 15th international workshop on scleroderma research, Pittsburgh, PA, USA. Aug 2017.
20. *Intrinsic Subsets Are Conserved Across Multiple Organs in Systemic Sclerosis*. Abstract. The 15th international workshop on scleroderma research, Pittsburgh, PA, USA. Aug 2017.

21. *Reprogramming Glucose Metabolism and Energy Production with a Small Molecule HJC0152 Suppresses Breast Cancer Development and Progression to Metastasis.* Abstract. American Association for Cancer Research (AACR) Annual Meeting, New Orleans, Louisiana, Apr 2016.
22. *Targeting STAT3 with Novel Small Molecule Inhibitors to Sensitize Breast Cancer Cells to Radiation Therapy.* Abstract. The Thirty-Seventh Annual CTRC-AACR San Antonio Breast Cancer Symposium. San Antonio, TX. Dec 2014.
23. *Using Expression Data to Define Patient Specific Predictors for Survival Outcomes in Lung Adenocarcinoma.* Poster Presentation. American Society of Human Genetics Annual Meeting (ASHG), Baltimore, MD, USA. Oct 2015.
24. *Networks Involved in the Metastatic Phenotype of Highly Metastasis Variants of Breast Cancer Cells.* Contributed Talk. 28th Society of Chinese Bioscientists in America (SCBA), Houston, TX, USA. Jun 2014.
25. *Accuracy of RNA-seq and Its Dependence on Sequencing Depth.* Poster presentation. Advances in Genome Biology and Technology (AGBT), Miami, FL, USA. Feb 2012.
26. *Accuracy of RNA-Seq and Its Dependence on Sequencing Depth.* Invited Talk. Biotechnology and Bioinformatics Symposium (BIOT), Houston, TX, USA. Jun 2011.

## **LAB HONOR AND ACTIVITIES**

---

1. *Assessment of Statistical Batch Effect-correction Methods for RNA-seq Data.* Presentation. Shenmiao Ni. Joint Symposium of NMU and UofSC, Columbia, SC, USA. Apr 2009.
2. *Prevalence and Risk Factors of Hypertension in East Asian Population: A Systematic Review and Meta-Analysis.* Presentation. Chenxi Yuan. Joint Symposium of NMU and UofSC, Columbia, SC, USA. Apr 2009.

## **PROFESSIONAL DEVELOPMENT**

---

1. Grant Bootcamp, ASPH, 2019
2. Bioinformatics workshop of Single Cell Sequencing, Oct 2019
3. IBM Watson Cloud Academy Workshop, UofSC, Apr 2019
4. Copyright in the Classroom, University Libraries, Apr 2019
5. Short Course: Statistical Methods for Single-cell RNA-seq Analysis, Joint Statistical Meetings, Vancouver, Canada, Aug 2018
6. Research Cyberinfrastructure Git Version Control Workshop, UofSC, Feb 2018
7. Bioinformatics workshop: CHIP-seq and ATAC-seq for Epigenomics Profiling and Functional Analysis, Nov 2017
8. Grant.gov WORKSPACE training workshop, ASPH, Nov 2017
9. Educational Workshop on Genetic Epidemiology: Fundamentals for Precision Medicine, Cambridge, UK. Sep 2017

## SERVICE AND ACTIVITIES

---

- University Service

ENHS POC, Joint ENHS-EPI Program, ASPH, SC, 2018-now  
Member, ASPH Evaluation Committee, ASPH, SC, 2019-now  
ENHS Assessment Plan Update, ASPH, SC, 2019  
Member, ENHS Admissions Committee, ASPH, SC, 2018-now  
Member, ENHS Awards Committee, ASPH, SC, 2018-now  
Member, ENHS Curriculum Committee, ASPH, SC, 2018-now  
Member, SPARC Review Committee, SC, 2018 Nov  
Judge for poster competition, Annual Meeting NIH Center for Epigenetic Regulation of Inflammation & NIH Center for Dietary Supplements and Inflammation (COBRE), 2018-2019

- Community Service

Program session chair, The 11th ICSA International Conference: Causal inference and related methodology in health sciences, 2019 Dec  
Review Editor, *Frontiers in Environmental Science*, 2019-now  
Reviewer, The University of Tennessee Internal Grant Competition, 2019 Mar  
Attendance, NIH NHGRI Strategic Planning Town Hall meeting, 2019 Oct

- Professional Leadership

Co-Director, the biometry core of South Carolina Center for Conlon Cancer Research (CCCR), 2019-now

- Referee

2019, Peer Reviewer for *Bioinformatics*, *IEEE Transactions on Biomedical Engineering*, *International Journal of Oncology*, *International Journal of Biological Sciences*, *NanoImpact*.

2018, Peer Reviewer for *Circulation*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, and *Frontier in Genetics*.

2017, Peer Reviewer for *Journal of Investigative Dermatology*, *BioData Mining*, *International Journal of Oncology*, *Journal of Thoracic Oncology*.

Before 2017, Peer Reviewer for *European Conference on Computational Biology 2016*, *Molecular and Cellular Biochemistry*, *Journal of Diabetes Investigation*, *Journal of Cancer Therapy*, *PeerJ*.

- Member of Scientific Societies

The International Genetic Epidemiology Society (IGES)  
The American Society of Human Genetics (ASHG)

2020-02-04

The American Statistical Association (ASA)

The American Association for Cancer Research (AACR)

Carolinan Society of Environmental Toxicology and Chemistry (CSETAC)