August 01, 2021

NANCY RUONAN ZHANG

Ge Li and Ning Zhao Professor, Department of Statistics
The Wharton School, University of Pennsylvania
400 Jon M. Huntsman Hall, 3730 Walnut St, Philadelphia, PA 19104,
Philadelphia, PA 19104 | 215-898-8007 | nzh@wharton.upenn.edu
Website: https://statistics.wharton.upenn.edu/profile/nzh/

EMPLOYMENT HISTORY

Post-doctoral Fellow (Mentored by Terence Speed and Mary Wildermuth)	**10.2005 - 07.2006**
Departments of Statistics and Plant Biology, University of California Berkeley	
Assistant Professor	**09.2006 - 06.2011**
Department of Statistics, Stanford University	
Associate Professor	**07.2011 – 06.2018**
Department of Statistics, The Wharton School, University of Pennsylvania	
Professor	**07.2018 - Current**
Department of Statistics, The Wharton School, University of Pennsylvania	
Ge Li and Ning Zhao Professor	**07.2019 - Current**
Department of Statistics, The Wharton School, University of Pennsylvania	
Vice Dean	**07.2019 - Current**
Wharton Doctoral Programs, The Wharton School, University of Pennsylvania	

EDUCATION

Stanford University

Bachelor's in Mathematics	**06.2001**
Master's in Computer Science	**06.2001**
Doctor of Philosophy in Statistics	**09.2005**

Dissertation Title:
Change-point models and sequence alignments: Statistical problems of genomics
Dissertation Advisor: David O. Siegmund

CITIZENSHIP

United States

HONORS

National Defense Science and Engineering Graduate Fellowship	2002
New World Silver Medal for Best Doctoral Thesis in the Mathematical Sciences	2007
Stanford University Terman Fellowship	2006
Sloan Fellowship	2011
American Statistical Association Medallion Lectureship	2021

PUBLICATIONS

PUBLISHED OR FORTHCOMING IN REFEREED JOURNALS

1. **Zhang NR**, Siegmund DO (2007) A modified Bayes information criterion with applications to the analysis of comparative genomic hybridization data, *Biometrics* 63, 22.

* *corresponding or co-corresponding author*
ǂ *alphabetical order*
¤ *co-first authors*

2. Chan HP, **Zhang NR** ‡ (2007) Scan statistics with weighted observations, *Journal of the American Statistical Association*, 102, 595.
3. The ENCODE Project Consortium (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project, *Nature* 447, 799.
4. **Zhang NR**, Wildermuth MC, Speed TP (2008) Transcription factor binding site prediction with multivariate gene expression data, *Annals of Applied Statistics* 2, 332.
5. Lai TL, Xing H, **Zhang NR**‡ (2008) Stochastic segmentation models for array-based comparative genomic hybridization data analysis, *Biostatistics* 9, 290.
6. **Zhang NR**, Senbabaoglu Y, Li J (2010) Joint estimation of DNA copy number from multiple platforms, *Bioinformatics* 26, 153.
7. Siegmund DO, Yakir B, **Zhang NR**‡ (2010) Tail approximations for maxima of random fields by likelihood ratio transformations, *Sequential Analysis* 29, 245.
8. **Zhang NR**, Siegmund DO, Ji H, Li J (2010) Detecting simultaneous changepoints in multiple sequences, *Biometrika* 97, 631.
9. Li F, **Zhang NR**‡ (2010) Bayesian variable selection in structured high-dimensional covariate spaces with applications in genomics, *Journal of the American Statistical Association* 105, 1202.
10. Bickel PJ, Boley N, Brown JB, Huang H, **Zhang NR** ‡ (2010) Subsampling methods for genomic inference, *Annals of Applied Statistics* 4, 1660.
11. Chan HP ¤, **Zhang NR**¤, Chen LHY (2010) Importance sampling of word patterns in DNA and protein sequences, *Journal of Computational Biology* 17, 1697.
12. Chen H, Xing H, **Zhang NR*** (2011) Estimation of parent specific DNA copy number in tumors using high-density genotyping arrays, *PLoS Computational Biology* 7, e1001060.
13. Siegmund DO, Yakir B, **Zhang NR**‡ (2011) Detecting simultaneous variant intervals in aligned sequences, *Annals of Applied Statistics* 5, 645.
14. Efron B and **Zhang NR**‡ (2011) False discovery rates and copy number variation, *Biometrika* 98, 251.
15. Natsoulis G, Bell JM, Xu H, Buenrostro JD, Ordonez H, Grimes S, Newburger D, Jensen M, Zahn JM, **Zhang N**, Ji HP (2011) A flexible approach for highly multiplexed candidate gene targeted resequencing, *PLoS One* 6, e21088.
16. Siegmund DO, **Zhang NR**, Yakir B (2011) False discovery rate for scanning statistics, *Biometrika* 98, 979.
17. Muralidharan O, Natsoulis G, Bell J, Newburger D, Xu H, Keta I, Ji H, **Zhang NR*** (2012) A cross-sample statistical model for SNP detection in short-read sequencing data, *Nucleic Acids Research* 40, e5.
18. Flaherty P, Natsoulis G, Muralidharan O, Winters M, Buenrostro J, Bell J, Brown S, Holodniy M, **Zhang N**, Ji HP (2012) Ultrasensitive detection of rare mutations using next-generation targeted resequencing, *Nucleic Acids Research* 40, e2.
19. Shen J, **Zhang NR*** (2012) Change-point model on nonhomogeneous Poisson processes with application in copy number profiling by next-generation DNA sequencing, *Annals of Applied Statistics* 6, 476.
20. Muralidharan O, Natsoulis G, Bell J, Ji H, **Zhang NR*** (2012) Detecting mutations in mixed sample sequencing data using empirical Bayes, *Annals of Applied Statistics* 6, 1047.
21. **Zhang NR**, Siegmund DO (2012) Model selection for high dimensional, multi-sequence change-point problems, *Statistica Sinica* 22, 1507.
22. Sun Y, **Zhang NR** and Owen A (2012) Multiple hypothesis testing, adjusted for latent variables, with an application to the agemap gene expression data, *Annals of Applied Statistics* 6, 1664.

23. Chen H, **Zhang NR**‡ (2013) Graph-based tests for two-sample comparisons of categorical data, *Statistica Sinica* 23, 1479.
24. Natsoulis G, **Zhang NR**, Welch K, Bell J, Ji HP (2013) Identification of insertion deletion mutations from deep targeted resequencing, *Journal of Data Mining in Genomics and Proteomics* 4, 132.
25. Nadauld LD, Garcia S, Natsoulis G, Bell JM, Miotke L, Hopmans ES, Xu H, Pai RK, Palm C, Regan JF, Chen H, Flaherty P, Ootani A, **Zhang NR**, Ford JM, Kuo CJ, Ji HP (2014) Metastatic tumor evolution and organoid modeling implicate TGFBR2 as a cancer driver in diffuse gastric cancer, *Genome Biology* 15,428.
26. Chen H, Bell JM, Zavala NA, Ji HP, **Zhang NR*** (2015) Allele-specific copy number profiling by next-generation DNA sequencing, *Nucleic Acids Research* 43, e23.
27. Jiang Y, Oldridge DA, Diskin SJ, **Zhang NR*** (2015) CODEX: a normalization and copy number variation detection method for whole exome sequencing, *Nucleic Acids Research* 43, e39.
28. Chen H, **Zhang NR**‡ (2015) Graph-based change-point detection, *The Annals of Statistics* 43, 139.
29. Cushing A, Kamali A, Winters M, Hopmans ES, Bell JM, Grimes SM, Li CX, **Zhang NR**, Moss RB, Holodniy M, Ji H (2015) Emergence of hemagglutinin mutations during the course of influenza infection, *Scientific Reports* 5, 16178.
30. Peixoto LL, Wimmer ME, Poplawski SG, Tudor JC, Kenworthy CA, Liu S, Mizuno K, Garcia BA, **Zhang NR**, Giese K, Abel T (2015) Memory acquisition and retrieval impact different epigenetic processes that regulate gene expression, *BMC Genomics* 16, S5.
31. Yue M, Han X, De Masi L, Zhu C, Ma X, Zhang J, Wu R, Schmieder R, Kaushik RS, Fraser GP, Zhao S, McDermott PF, Weill FX, Mainil JG, Arze C, Fricke WF, Edwards RA, Brisson D, **Zhang NR**, Rankin SC, Schifferli DM (2015) Allelic variation contributes to bacterial host specificity, *Nature Communications* 6, 8754.
32. Wang X, Chen M, Yu X, Pornputtapong N, Chen H, **Zhang NR**, Powers RS, Krauthammer M (2016) Global copy number profiling of cancer genomes, *Bioinformatics*, 32, 926.
33. **Zhang NR**, Yakir B, Xia LC, Siegmund DO (2016) Scan statistics on Poisson random fields with applications in genomics, *Annals of Applied Statistics* 10, 726.
34. Xia LC, Sakshuwong S, Hopmans ES, Bell JM, Grimes SM, Siegmund DO, Ji HP, **Zhang NR*** (2016) A genome-wide approach for detecting novel insertion-deletion variants of mid-range size, *Nucleic Acids Research* 44, e126.
35. Jiang Y, Qiu Y, Minn AJ, **Zhang NR*** (2016) Assessing intratumor heterogeneity and tracking longitudinal and spatial clonal evolutionary history by next-generation sequencing, *Proceedings of the National Academy of Sciences* 113, E5528.
36. Wang X, Chen H, **Zhang NR** (2017) DNA copy number profiling using single-cell sequencing, *Briefings in Bioinformatics*, bbx004, https://doi.org/10.1093/bib/bbx004.
37. Jiang Y, **Zhang NR***, Li M* (2017) SCALE: modeling allele-specific gene expression by single-cell RNA-sequencing, *Genome Biology* 18, 74.
38. Chen H, Jiang Y, Maxwell K, Nathanson K, **Zhang NR*** (2017) Allele-specific copy number estimation by whole exome sequencing, *Annals of Applied Statistics* 11, 1169.
39. Jia C, Hu Y, Kelly D, Kim J, Li M*, **Zhang NR*** (2017) Accounting for technical noise in differential expression analysis of single-cell RNA sequencing data, *Nucleic Acids Research*, 45, 10978.
40. Maxwell KN, Wubbenhorst B, Wenz BM, Sloover DD, Pluta J, Emery L, Barrett A, Kraya AA, Anastopoulos IN, Yu S, Jiang Y, Chen H, **Zhang NR**, Hackman N, D'Andrea K, Daber R, Morrissette JJ, Mitra N, Feldman M, Domchek SM, Nathanson KL (2017) BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers, *Nature Communications* 8, 319.

41. Xia LC, Bell JM, Wood-Bouwens C, Chen JJ, **Zhang NR***, Ji HP* (2017) Single molecule-based discovery of complex genomic rearrangements, *Nucleic Acids Research* 46, e19.
42. Garman B, Anastopoulos IN, Krepler C, Brafford P, Sproesser K, Jiang Y, Wubbenhorst B, Amaravadi R, Bennett J, Beqiri M, Elder D, Flaherty KT, Frederick DT, Gangadhar TC, Guarino M, Hoon D, Karakousis G, Liu Q, Mitra N, Petrelli NJ, Schuchter L, Shannan B, Sheilds CL, Wargo J, Wenz B, Wilson MA, Xiao M, Xu W, Xu X, Yin X, **Zhang NR**, Davies MA, Herlyn M, Nathanson KL (2017) Genetic and genomic characterization of 462 melanoma patient-derived xenografts, tumor biopsies and cell lines, *Cell Reports* 21, 1936.
43. Huang M, Wang J, Torre E, Dueck H, Shaffer S, Bonasio R, Murray J, Raj A, Li M, **Zhang NR*** (2018) SAVER: Gene expression recovery for single cell RNA sequencing, *Nature Methods* 15, 539.
44. Zhou Z, Wang W, Wang L-S, **Zhang NR*** (2018) Integrative DNA copy number detection and genotyping from sequencing and array-based platforms, *Bioinformatics* 34, 2349.
45. Wang X, Jiang Y, **Zhang NR**, Small D (2018) Sensitivity analysis and power for instrumental variable studies, *Biometrics* doi: 10.1111/biom.12873.
46. Urrutia E, Chen H, Zhou Z, **Zhang NR***, Jiang Y* (2018) Integrative pipeline for profiling DNA copy number and inferring tumor phylogeny, *Bioinformatics* 34, 2126.
47. Zhang H, **Zhang NR**, Li M, Reilly MP (2018) First giant steps towards a cell atlas of atherosclerosis, *Circulation Research* 122, 1632.
48. Wang J, Huang M, Torre E, Dueck H, Shaffer S, Murray J, Raj A, Li M, **Zhang NR*** (2018) Gene expression distribution deconvolution in single cell RNA sequencing, *Proceedings of the National Academy of Sciences* 115, E6437.
49. Jiang Y, Nathanson KL, **Zhang NR*** (2018) CODEX2: full-spectrum copy number variation detection by high-throughput DNA sequencing, accepted by *Genome Biology* 19, 202.
50. Wang X, Park J, Susztak K, **Zhang NR***, Li M* (2019) Bulk Tissue Cell Type Deconvolution with Multi-Subject Single-Cell Expression Reference, accepted by *Nature Communications* 10, 380.
51. Wang J, Agarwal D, Huang M, Hu G, Zhou Z, Ye C, **Zhang NR*** (2019) Data denoising with transfer learning in single-cell transcriptomics, *Nature Methods* 16, 875.
52. Benci JL et al. (2019) Opposing Functions of Interferon Coordinate Adaptive and Innate Immune Responses to Cancer Immune Checkpoint Blockade, *Cell* 178 (4), 933-948. e14.
53. Nguyen S et al. (2019) Elite control of HIV is associated with distinct functional and transcriptional signatures in lymphoid tissue CD8+ T cells, *Science Translational Medicine* 11(523).
54. Agarwal D, **Zhang NR*** (2019) A rank-based semblance kernel on probability spaces, *Science Advances* 5 (12), eaau9630.
55. Zhou Z, Xu B, Minn A, Zhang NR* (2020) Dendro: genetic heterogeneity profiling by single cell RNA sequencing, Genome Biology 21, 10. (R package: DENDRO)
56. Zhou Z, Ye C, Wang J, **Zhang NR*** (2020) Surface protein imputation from single cell transcriptomes by deep neural networks, *Nature Communications* 11, Article number: 651
57. Agarwal D, Wang J, **Zhang NR*** (2020) Data denoising and post-denoising corrections in single cell RNA sequencing, Statistical Science 35 (1), 112-128.
58. Rozenblatt-Rosen et al. (2020) The Human Tumor Atlas Network (HTAN): charting tumor transitions across space and time at single-cell resolution. Cell 181, 236. (Consortia paper as part of HTAN)
59. Mukherjee S, Agarwal D, **Zhang NR** & Bhattacharya BB (2020) Distribution-Free Multisample Tests Based on Optimal Matchings with Applications to Single Cell Genomics, Journal of the American Statistical Association, DOI: 10.1080/01621459.2020.1791131
60. Wu C-Y, Lau BT, Kim H, Sathe A, Grimes SM, Ji HP, **Zhang NR*** (2021) Integrative single-cell analysis of allele-specific copy number alterations and chromatin accessibility in cancer. *Nature Biotechnology,* doi: 10.1038/s41587-021-00911-w

PUBLISHED BOOK CHAPTERS

61. Chan HP, Tu I-P, **Zhang NR** (2009) Boundary crossing probability computations in the analysis of scan statistics, in *Scan Statistics - Methods and Applications.* Birkhauser, Boston.
62. **Zhang NR** (2010) DNA copy number profiling in normal and tumor genomes, In *Frontiers in Computational and Systems Biology*, ed. Jianfeng Feng, Wenjiang Fu and Fengzhu Sun.

GRANT AWARDS

*(**NSF** = National science foundation, **DMS** = Division of Mathematical Sciences, **NIH** = National Institutes of Health, **NHGRI** = National Human Genome Research Institute, **DoJ**=Department of Justice, **NIA** = National Institutes of Aging, **NCI** = National Cancer Institute, **NHLBI** = National Heart, Lung and Blood Institute)*

Period	Agency, *Mechanism*	Role	Title	Direct Cost ($)
2009 - 2012	NSF (DMS)	PI	Change-point Problems in Genomic Profiling	100,000
2010- 2013	NSF (DMS)	co-PI	Statistical Methods for Threat Detection	711
7/6/11 - 6/30/17	NIH (NHGRI) *R01*	PI	Statistical Models and Analysis of Complex Variation in Clonal Mixtures	577,971
9/15/11 - 9/15/13	Alfred P. Sloan Foundation	PI	Statistical Methods for Genome Profiling	50,000
5/1/12 - 4/30/14	NIH (NHGRI) *R01*	PI (Subcontract)	Statistical Models for Genome Sequencing and Association	81,110
1/1/14 - 12/31/16	DoJ	PI (Subcontract)	Highly Parallel Analysis of Complex Genetic Mixtures	72,731
6/15/14 – 5/31/18	NIH (NIA) *UF1*	Co-Investigator	Consortium for Alzheimers Sequence Analysis (CASA)	11,063,917
4/15/16 – 2/28/21	NIH (NIA) *U54*	Co-Investigator	Coordinating Center for Genetics and Genomics of Alzheimers Disease (CGAD)	10,801,796
5/1/16 – 4/30/20	NSF (DMS)	Co-PI	Statistical Methods for High- Resolution Multiscale Analysis 3D DNA	948,742
9/30/16 – 6/30/21	NIH *U54*	Co-Investigator	Identifying Genes and Pathways that Impact Tau Toxicity in FTD	1,250,000
4/1/17 – 3/31/22	NIH (NIA) *U24*	Co-Investigator	The NIA Genetics of Alzheimer's Disease Data Storage Site (NIAGADS)	4,802,337
8/1/17- 7/31/22	NIH (NCI) *P01*	Co-Investigator	Radiation and Checkpoint Blockade for Cancer Immune Therapy	8,795,373
9/1/17 – 8/31/21	NIH (NIGMS) *R01*	PI (Multiple PI Grant)	Statistical Methods for Single- Cell Transcriptomics	948,000
9/14/17 – 6/30/20	NIH (NHGRI) *R01*	PI (Multiple PI Grant)	Genomic and Cellular Variation from Single Molecules to Single Cells	917,539
7/1/18- 6/30/23	NIH (NHLBI) *R01*	Co-Investigator	Elucidation of Tissue-Specific Transcriptomic Profiles in Cardiometabolic Disease	2,222,228
9/01/18- 8/31/23	NIH (NCI) *U2C*	Data Analysis Unit Co-lead	Center for Pediatric Tumor Cell Atlas	13,553,635

COURSES TAUGHT

Stanford University STATISTICS 191 – Applied Statistics	2007, 2008
Stanford University STATISTICS 203 – Introduction to ANOVA	2009, 2010
Stanford University STATISTICS 205 – Nonparametric Statistics	2007, 2008
Stanford University STATISTICS 215 – Stochastics Processes with Applications in Biology	2008, 2009, 2010

Stanford University STATISTICS 345/GEN245 – Computational Algorithms in Statistical Genetics	2009
Stanford University STATISTICS 366 – Statistical Methods in Genetics	2010
Wharton School STAT 102 – Introductory to Business Statistics	2012, 2015
Wharton School STAT 431 – Introductory Statistics	2012
Wharton School STAT 471/701 – Intermediate Statistics	2013
Wharton School STAT 405/705 – Statistical Computing with R	2016, 2017, 2019

MENTORING

I served (or am serving) as doctoral dissertation advisor for:

Yunting Sun (Joint with Art Owen), Department of Statistics, Stanford University (Joined Google Inc.)	2012
Jeremy Shen, Department of Statistics, Stanford University (Joined Two Sigma Investments.)	2012
Hao Chen (Joint with David Siegmund), Department of Statistics, Stanford University (Joined Department of Statistics as Assistant Professor, University of California Davis.)	2014
Yuchao Jiang, Graduate Program in Genomics and Computational Biology, University of Pennsylvania (Joined Departments of Biostatistics and Genetics as Assistant Professor, University of North Carolina.)	2017
Yang Jiang (Joint with Dylan Small) Department of Statistics, University of Pennsylvania	2017
Xuran Wang, Graduate Program in Applied Mathematics and Computational Sciences, University of Pennsylvania (Joined Carnegie Mellon University as Postdoctoral Researcher)	2019
Mo Huang, Department of Statistics, The Wharton School, University of Pennsylvania (Joined Merck Pharmaceuticals)	2020
Zilu Zhou, Graduate Program in Genomics and Computational Biology, University of Pennsylvania (Joined Google Research)	2020
Divyansh Agarwal, Graduate Program in Genomics and Computational Biology, University of Pennsylvania	2020
Chi-Yun Wu, Graduate Program in Genomics and Computational Biology, University of Pennsylvania	Current
Kaishu Mason, Department of Statistics, University of Pennsylvania	Current

I served (or am serving) as postdoc mentor for:

Charlie Xia (Joint with Hanlee Ji, joined Albert Einstein College of Medicine as Assistant Professor.)	2017
Jingshu Wang (Joined University of Chicago Department of Statistics as Assistant Profesor.)	2019
Kevin Lin	Current
Paul Hess	Current

Since arriving at Penn, I've served on the Thesis Advising Committees of:

Jun Chen, Graduate Group in Genomics and Computational Biology	2012
Jonathan Toung, Graduate Group in Genomics and Computational Biology	2013
Joseph Glassner, Graduate Group in Genomics and Computational Biology	2014
Vicky Wu, Department of Biostatistics, Epidemiology and Informatics	2014
Scott Sherrill-Mix, Graduate Group in Genomics and Computational Biology	2015
Hannah Dueck, Graduate Group in Genomics and Computational Biology	2015
Yih-Chii Hwang, Graduate Group in Genomics and Computational Biology	2015
Hyunseung Kang, Department of Statistics	2017
Ying Chen, Graduate Group in Genomics and Computational Biology	2017
Xiao Ji, Graduate Group in Genomics and Computational Biology	2017
Xinyao Ji, Department of Statistics	2017
Cheng Jia, Department of Biostatistics, Epidemiology and Informatics	2017
Yu Hu, Department of Biostatistics, Epidemiology and Informatics	2018
Gemma Moran, Department of Statistics	2019
Benjamin Emert, Graduate Group in Genomics and Computational Biology	2021
Katerina Gawronski, Graduate Group in Genomics and Computational Biology	2021
Gregory Way, Graduate Group in Genomics and Computational Biology	2019
Eric Sanford, Graduate Group in Genomics and Computational Biology	2021
Sammy Klasfeld, Graduate Group in Genomics and Computational Biology	Current
Kathy Huang, Graduate Group in Genomics and Computational Biology	Current
Jingya Qiu, Graduate Group in Genomics and Computational Biology	Current
Yang Xu, Graduate Group in Genomics and Computational Biology	Current
Jason Xu, Graduate Group in Genomics and Computational Biology	Current

SELECT NOTABLE SERVICE ACTIVITIES

EDITORIAL BOARDS

Associate Editor, Annals of Applied Statistics	2015-2018
Editorial Board, Briefings in Bioinformatics	2017-Current
Guest Editor, Genome Research Special Issue on Single Cell Biology	2020-2021

GRANT REVIEW PANEL AND STUDY SECTIONS

National Science Foundation – National Institute of General Medical Sciences Joint Study Section	2011
National Institutes of Health – Genomics, Computational Biology and Technology (GCAT)	2012, 2015, 2017
National Institutes of Health – Advanced Genomic Technology Development Panel	2017

ACADEMIC SERVICE

Chair-elect, American Statistical Association Section on Statistical Genetics and Genomics	2022

(At Stanford)

Masters student advisor, Department of Statistics	2010
Undergraduate advisor, Computational Mathematics Major	2007-2011

VPUE Undergraduate Summer Research Program Coordinator	2007
(At Penn)	
Doctoral Program Co-Director, Department of Statistics, The Wharton School	2012-2017
Doctoral Program Advisory Committee, Graduate Group in Genomics and Computational Biology	2014-2019
Center for Neurodegeneration Faculty Search Committee	2016-2017
Director of Admissions, Department of Statistics, The Wharton School	2017-2019

DIVERSITY, EQUITY, and INCLUSION

	2019
Undergraduate Mentor, Leadership Alliance Program	2021
Founder, Wharton Directed Reading Program	

SELECT NOTABLE INVITED TALKS (SINCE 2013)

Department of Statistics, Stanford University	2013
Department of Statistics, Harvard University	2013
IMS-China Meeting, Chengdu, China	2013
Department of Biostatistics, Johns Hopkins University	2014
ENAR Spring Meeting, Baltimore, MD	2014
iBright Conference, Houston, TX	2015
Department of Statistics, Georgia Institute of Technology	2016
Center for Statistics and Machine Learning, Princeton University	2016
Cornell Day of Statistics, Ithaca, NY	2016
ICSA Applied Statistics Symposium, Atlanta, GA	2016
Department of Biostatistics, Brown University	2016
Department of Statistics, Stanford University	2016
Graybill Conference, Fort Collins, CO	2017
Department of Biostatistics, University of Michigan	2017
ENAR Spring Meeting, Washington, DC	2017
Joint Statistical Meetings, Baltimore, MD	2017
ICSA Applied Statistics Symposium, Chicago, IL	2017
Department of Statistics, Pennsylvania State University	2017
Machine Learning Seminar Series, Duke University	2017
DahShu Virtual Journal Club	2017
Biostatistics Branch, National Cancer Institute	2018
ENAR Spring Meeting, Atlanta, GA	2018
Joint Statistical Meetings, Vancouver, Canada	2018
Department of Statistics, University of Chicago	2018
Department of Biostatistics, University of Washington	2018
Department of Statistics and Data Science, Carnegie Mellon University	2018
The Australian Bioinformatics and Computational Biology Society Annual Conference (keynote speaker)	2018
Department of Biostatistics, University of North Carolina	2019
Frontiers in Single-cell Technology, Applications and Data Analysis (Banff Workshop)	2019
New York University Genomics Symposium	2019
Mathematical Biosciences Institute, Ohio State University	2019

Institute for Advanced Studies (Missing Data Challenges in Computation, Statistics, and Applications), Princeton University	2020
Department of Statistics, Yale University	2020
Department of Statistics, Stanford University	2020
Department of Statistics, UC Berkeley	2020
Joint Statistical Meetings	2020
Gordon Conference on Single Cell Cancer Biology	2020
Keystone Symposia on Single Cell Biology	2021
American Society of Nephrology Kidney Week	2021
Joint Statistical Meetings (Medallion Lecture)	2021

SOFTWARE PACKAGES DEVELOPED BY MY GROUP

For single cell data analysis:

TASC (Toolkit for noise modeling in single cell RNA-seq with spike-ins)
https://github.com/scrna-seq/TASC

SCALE (Single cell allele-specific expression analysis)
https://github.com/yuchaojiang/SCALE

DESCEND (Expression distribution deconvolution for single cell RNA-seq)
https://github.com/jingshuw/descend

MUSIC (Bulk expression deconvolution with scRNA-seq reference)
https://github.com/xuranw/MuSiC

DENDRO (Genetic heterogeneity profiling by scRNA-seq)
https://github.com/zhouzilu/DENDRO

SAVER (Gene expression imputation and denoising for single cell RNA sequencing)
https://github.com/mohuangx/SAVER

SAVER-X (SAVER harnessing external data)
https://singlecell.wharton.upenn.edu/saver-x/

cTP-Net (single cell Transcriptome to Protein prediction with deep neural network)
https://github.com/zhouzilu/cTPnet/

Alleloscope (allele-specific copy number estimation for scDNA and scATAC sequencing)
https://github.com/seasoncloud/Alleloscope

For copy number profiling and tumor heterogeneity analysis:

CANOPY (Tumor phylogeny reconstruction by spatial and temporal bulk RNA sequencing)
https://cran.r-project.org/web/packages/Canopy/

MARATHON (Comprehensive pipeline for copy number profiling in normal and tumor samples)
https://github.com/yuchaojiang/MARATHON

SWAN (Structural variant profiling using paired-end genome sequencing data)
https://bitbucket.org/charade/swan/overview

CODEX/CODEX2 (statistical framework for full-spectrum CNV profiling in whole genome, whole exome, and targeted DNA sequencing)
https://github.com/yuchaojiang/CODEX2

iCNV (Integration across array and sequencing platforms for copy number detection)
https://github.com/zhouzilu/iCNV

FALCON (Allele-specific copy number estimation using whole genome sequencing data)
https://cran.r-project.org/web/packages/falcon/index.html

FALCON-X (Allele-specific copy number estimation using whole exome sequencing data)
https://cran.r-project.org/web/packages/falconx/index.html

General statistical tools:

SEMBLANCE (rank-semblance kernel for data compression, niche detection, and feature extraction)
https://cran.r-project.org/web/packages/Semblance/index.html

GSEG (Change-point detection for multivariate data through a similarity graph on the observations)
https://cran.r-project.org/web/packages/gSeg/index.html

GCAT (Two-sample tests for categorical data utilizing similarity information among the categories)
https://cran.r-project.org/web/packages/gCat/index.html

SEQCBS (Segmentation and Bayesian confidence interval calculation for matched case/control point processes)
https://cran.r-project.org/web/packages/seqCBS/index.html

LEAPP (Latent factor ("batch effect") adjustment in multiple hypothesis testing)
https://cran.r-project.org/web/packages/leapp/index.html

MEMBERSHIPS

American Statistical Association