Zhaohui (Steve) Qin

Department of Biostatistics and Bioinformatics
Rollins School of Public Health
Emory University
1518 Clifton Rd NE
Atlanta, GA 30322
Tel: (404) 712-9576
Fax: (404) 727-1370
Email: zhaohui.qin@emory.edu
URL:http://sph.emory.edu/faculty/profile/#ZQIN4
Twitter: @SteveQinEmory
ORCID: 0000-0002-1583-146X
LinkedIn: https://www.linkedin.com/in/steveqin/
Google Scholar: https://scholar.google.com/citations?user=9F-Jk74AAAAJ&hl=en
My NCBI: https://www.ncbi.nlm.nih.gov/myncbi/zhaohui.qin.1/bibliography/public/
Publons: https://publons.com/researcher/4470669/zhaohui-qin/
Kudos: https://www.growkudos.com/profiles/151858/
Scopus: https://www-scopus-com.proxy.library.emory.edu/authid/detail.uri?authorId=7202822915

Current position

Professor	May 2020 – present
Director of Graduate Studies	June 2016 – present
Department of Biostatistics and Bioinformatics	
Rollins School of Public Health	
Emory University	Atlanta, GA

Additional affiliation
Department of Biomedical Informatics, Emory University School of Medicine
Department of Computer Science, Emory University
Biostatistics and Bioinformatics Shared Resource, Winship Cancer Institute

Education

B.S. in Probability and Statistics	Sep 1990 – Jun 1994
Peking University	Beijing, China
M.S. in Statistics	Sep 1995 – Jun 1997
University of Michigan	Ann Arbor, MI
Ph.D. in Statistics	Sep 1995 – Jun 2000
University of Michigan	Ann Arbor, MI

Professional experience

Graduate Student Instructor	Sep 1995 – May 1999
Department of Statistics	
University of Michigan	Ann Arbor, MI
Graduate Student Research Assistant	Sep 1998 – May 1999
Department of Statistics and Management Science	
University of Michigan Business School	Ann Arbor, MI
Visiting Research Assistant	Aug 1999 – Sep 1999
Department of Statistics	
Stanford University	Stanford, CA
Graduate Student Research Assistant	Sep 1999 – Aug 2000
Center for Statistics Consulting and Research	
University of Michigan	Ann Arbor, MI
Postdoctoral fellow	Sep 2000 – Aug 2003
Department of Statistics, Advisor: Jun S. Liu	
Harvard University	Cambridge, MA
Assistant Professor	Sep 2003 – Jul 2010
Center for Statistical Genetics, Department of Biostatistics	
University of Michigan	Ann Arbor, MI
Associate Professor	Aug 2010 – May 2020
Department of Biostatistics and Bioinformatics	
Emory University	Atlanta, GA
Core Member of Long Program	Sep 2011 – Dec 2011
Workshop on Mathematical and Computational Approaches on High Throughput Genomics	
Institute of Pure and Applied Mathematics	
University of California at Los Angeles,	Los Angeles, CA

Research Funding

Active grants:

R01 ES033603-01 “Gene-Environment Interactions in Autism”
PI: Corces, Sloan, Qin
02/01/2022 – 01/31/2026

P50AG025688-15S1 “Emory Alzheimer’s Disease Research Center (Supplement)”
PI: Levy
06/01/2015-04/30/2025

U01MH116441 “Dynamic RNA Modifications in Human Brain Development and Neuropsychiatric Disorder”
PI: Jin
08/01/2018-04/30/2023

P30 ES019776 “HERCULES: Health and Exposome Research Center at Emory”
PI: Marsit
06/01/2017-03/31/2022

1R01AI145231-01A1 “Using donor dendritic cells to optimize GvHD and GvL in allogeneic stem cell transplantation”
PI: Waller
06/01/2020 – 05/31/2025

R01AG070937-01 “The Emory Healthy Brain Study: Discovering Predictive Biomarkers for Alzheimer’s Disease”
PI: Lah
02/01/2021 - 1/31/2026

Completed PI or co-PI grants:

University of Michigan Rackham Graduate School Interdisciplinary workshop. “*New Statistical Methods in Molecular Biology*”. 2004-2008.

University of Michigan Center for Computational Medicine and Biology 2006 pilot grant “*Joint Modeling of OMICS Data to Prioritize Drug Target Genes in B. anthracis*”.
Multiple PIs: Qin, Bergman

The Genetics and Genomics 2008 Pilot Feasibility Grant. University of Michigan Center for Genomics Health and Medicine “*Targeted Resequencing Study of Seven New Psoriasis Candidate Genes*”.

Career Development Research Project. The University of Michigan Comprehensive Cancer Center Prostate SPORE “*Model-based Methods to Analyze Massive Parallel Sequencing Data with Application to Prostate Cancer Progression*”. 2008 – 2010

The Genetics and Genomics 2008 Pilot Feasibility Grant. University of Michigan Center for Genomics Health and Medicine “*Use of ChIP-Seq to identify target genes in a novel, regulated transcriptional repressor-activator relationship*”.
Multiple PIs: Schwartz, Qin

The Genetics and Genomics 2010 Pilot Feasibility Grant. University of Michigan Center for Genomics Health and Medicine “*ChIP-Seq Analysis of Transcriptional Coactivator PGC-1α in the Control of Circadian Energy Metabolism in the Liver*”.
Multiple PIs: Lin, Qin

R01HG005119-01 “*Model-based Methods for Analyzing ChIP Sequencing Data*”
PI: Qin
07/22/09 – 06/30/13

R21HG004751-01A2 “*Statistical Methods for Analyzing Resequencing Data*”
PI: Qin
09/21/10 – 07/31/12

DMS1000617 "*Spatial Model-Based Methods for RNA-seq Data Analysis*"
Multiple PIs: Zhu (Purdue University), Qin
10/01/10 – 09/30/13

P01 GM085354 "Understanding Mechanisms of hESC Self-Renewal and Cell Fate Commitment"
PI: Dalton (U of Georgia),
07/01/14-02/28/19
Role: Bioinformatics Core director

R56 AG060757 "Computational Prediction and Functional Validation of Novel Risk Loci of Alzheimer's Disease"
MPI: Jin, Qin, Wingo
09/30/2018-08/31/2019

R56 AG062256 "Computational Prediction and Functional Validation of Novel Epigenetic Risk Loci in Alzheimer's Disease"
MPI: Jin, Qin, Wingo
09/30/2018-08/31/2019

Woodruff Health Science Center Synergy Award "Towards federated data-driven precision medicine approaches for rare conditions: Down syndrome as a use case"
MPI: Sherman, Zwick, Qin, Sunderam.
04/01/2018-03/31/2019

Alzheimer Disease Research Center Pilot grant "Use machine learning to discover new genetic variants associated with AD"
PI: Qin
01/01/2018-12/31/2019

Completed co-I grants (selected)

Department of Defense W81XWH-08-1-0110 "*A Search for Gene Fusions/ Translocations in Breast Cancer*" (PI: Chinnaiyan).
Role: Co-investigator.

American Diabetes Association 7-09-BS-168 "*Role of transcriptional repressor Bcl-6 in GH-induced insulin resistance*". (PI: Schwartz).
Role: Co-investigator.

1 R56 AI091827-01A1 "*Genetics of Vancomycin Intermediate Resistance in S. Aureus*"
PI: Read
08/01/11 – 07/31/12
Role: Co-investigator

R03 AI111396-01 "Epigenome-wide Association Study of Asthma in Populations of African Descent"
PI: Hu
08/01/14-07/31/16
Role: Co-Investigator

R01 HL104608-01A1 "*New Approaches for empowering Studies of Asthma in Populations of African Descent*"
PI: Barnes (Johns Hopkins),
09/28/11 – 09/30/16
Role: Emory Subcontract PI

R01 HL104608-01S1 "*New Approaches for empowering Studies of Asthma in Populations of African Descent*" (Administrative Supplement)
PI: Barnes (Johns Hopkins),
09/29/14 – 09/30/15
Role: Emory Subcontract PI

R01 CA 077337-11 "Genesis and Consequences of Aberrant DNA Methylation"
PI: Vertino
04/01/12 - 03/31/17
Role: Co-investigator

U54 NS091859 "Modifiers of FMR1-associated Disorders: Application of High Throughput Technologies"
PI: Warren
09/22/14-05/31/18

R01 MD011698-01 "Influence of prenatal psychosocial stressors on maternal and fetal circulating miRNAs"
PI: Breton (USC) / Marsit (Emory)
07/26/17-02/28/22

Peer reviewed journal, conference and book chapter publications

* trainee.
# joint first author.

Chen L, **Qin ZS**, Liu JS. (2001) Exploring Hybrid Monte Carlo in Bayesian Computation. *Bayesian Methods with Application to Science, Policy and Official Statistics.* 71-80.

Qin ZS, Liu JS. (2001) Multi-Point Metropolis Method with Application to Hybrid Monte Carlo. *J. Comp. Phys.* **172** 827-840.

#Niu T, #**Qin ZS**, Xu X, Liu JS. (2002) Bayesian Haplotype Inference for Multiple Linked Single Nucleotide Polymorphisms. *Am. J. Hum. Genet.* **70** 157-169.

#Qin ZS, #Niu T, Liu JS. (2002) Partition-Ligation EM Algorithm for Haplotype Inference with Single Nucleotide Polymorphisms. *Am. J. Hum. Genet.* **71** 1242-1247.

Qin ZS, McCue LA, Thompson W, Mayerhoffer L, Lawrence CE, Liu JS. (2003) Identification of Co-regulated Genes Through Bayesian Clustering of Predicted Regulatory Binding Sites. *Nat Biotechnol.* **21** 435-439.

Lu X, Zhang W, **Qin ZS,** Kwast KE, Liu JS. (2004) Statistical Resynchronization and Bayesian Detection of Periodically Expressed Genes. *Nucleic Acids Res.* **32** 447-455.

Kang H, **Qin ZS**, Niu T, Liu JS. (2004) Incorporating Genotyping Uncertainty in Haplotype Inference for Single-Nucleotide Polymorphisms. *Am. J. Hum. Genet.* **74** 495-510.

Zhang K, **Qin ZS**, Liu JS, Chen T, Waterman MS, Sun F. (2004) Haplotype Block partitioning and Tag SNP Selection Using Genotype Data and Their Applications to Association Studies. *Genome Res.* **14** 908-916.

Zhang K, **Qin ZS**, Chen T, Liu JS, Waterman MS, Sun F. (2004) HapBlock: Haplotype Block Partitioning and Tag SNP Selection Software Using a Set of Dynamic Programming Algorithms. *Bioinformatics.* **21** 131-134.

Niu T, Lu X, Kang H, **Qin ZS**, Liu JS. (2004) Haplotype Inference and Its Application in Linkage Disequilibrium Mapping. Computational Methods in SNPs and Haplotype Inference. DIMACS/RECOMB Satellite Workshop, Piscataway, NJ, USA, November 21-22, 2002, Editors: Sorin Istrail, Michael Waterman, Andrew Clark. Springer-Verlag. 48-61.

Zhang K, Chen T, Waterman MS, **Qin ZS**, Liu JS, Sun F. (2004) Dynamic Programming Algorithms for Haplotype Block Partitioning and Tag SNP Selection Using Haplotype Data or Genotype Data DIMACS/RECOMB Satellite Workshop, Piscataway, NJ, USA, November 21-22, 2002. Editors: Sorin Istrail, Michael Waterman, Andrew Clark. Springer-Verlag. 96-112.

*Zhu D, **Qin ZS** (2005) Structure Comparison of Metabolic Networks in Selected Single Cell Organisms. *BMC Bioinformatics.* **6** 8.

Cluster 17 Collaboration (2005) Fine Mapping of the Psoriasis Susceptibility Gene PSORS1: A Reassessment of Risk Associated With a Putative Risk Haplotype Lacking HLA-Cw6. *J Invest Dermatol.* **124** 921-930.

*Zhu D, Hero AO, **Qin ZS,** Swaroop A. (2005) High Throughput Screening of Co-expressed Gene Pairs with Controlled False Discovery Rate (FDR) and Minimum Acceptable Strength (MAS). *J. Comput. Biol.* **12** 1029-1045.

Nistor I, Nair RP, Stuart P, Abecasis GR, Hiremagalore R, Thompson RA, Jenisch S, Weichenthal M, Abecasis GR, **Qin ZS**, Christophers E, Lim HW, Voorhees JJ, Elder JT. (2005) Protein Tyrosine Phosphatase Gene PTPN22 Polymorphism in Psoriasis: Lack of Evidence for Association. *J Invest Dermatol.* **125** 395-396.

Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P; International HapMap Consortium. (2005) A Haplotype Map of the Human Genome. *Nature.* **437** 1299-1320.

Stuart P, Nair RP, Abecasis GR, Nistor I, Hiremagalore R, Chia NV, **Qin ZS**, Thompson RA, Jenisch S, Weichenthal M, Janiga J, Lim HW, Christophers E, Voorhees JJ, Elder JT. (2006) Analysis of RUNX1 Binding Site and RAPTOR Polymorphisms in Psoriasis: No Evidence for Association Despite Adequate Power and Evidence for Linkage. *J Med Genet.* **43** 12-17.

Qin ZS, *Gopalakrishnan S, Abecasis GR. (2006) An Efficient Comprehensive Search Algorithm for TagSNP Selection Using Linkage Disequilibrium Criteria. *Bioinformatics.* **22** 220-225.

*Ulintz PJ, Zhu J, **Qin ZS**, Andrews PC. (2006) Improved Classification of Mass Spectrometry Database Search Results Using Newer Machine Learning Approaches. *Mol Cell Proteomics.* **5** 497-509.

Marchini J, Cutler D, Patterson N, Stephens M, Eskin E, Halperin E, Lin S, **Qin ZS**, Munro HM, Abecasis GR, Donnelly P; International HapMap Consortium. (2006) A Comparison of Phasing Algorithms for Trios and Unrelated Individuals. *Am J Hum Genet.* **78** 437-450.

*Gopalakrishnan S, **Qin ZS.** (2006) TagSNP Selection Based on Pairwise LD Criteria and Power Analysis in Association Studies. *Pac Symp Biocomput.* **11** 511-522.

Qin ZS. (2006) Clustering Microarray Gene Expression Data Using Weighted Chinese Restaurant Process. *Bioinformatics*. **22** 1988-1997.

Bergman NH, Passalacqua KD, Hanna PC, **Qin ZS**. (2007) Operon Prediction in Sequenced Bacterial Genomes without Experimental Information. *Appl Environ Microbiol*. **73** 846-854.

Xiang Z, **Qin ZS**, He Y. (2007) CRCView: A Web Server for Analyzing and Visualizing Microarray Gene Expression Data Using Model-based Clustering. *Bioinformatics.* **23** 1843-1845.

*Ma J, **Qin ZS.** (2007) Different Normalization Strategies for Microarray Gene Expression Traits Affect the Heritability Estimation. *BMC Proc*. **1 (Suppl 1)** S154.

Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE, Zhai Y, Giordano TJ, **Qin ZS**, Moore BB, MacDougald OA, Cho KR, Fearon ER. (2007) p53-mediated Activation of miRNA34 Candidate Tumor-suppressor Genes. *Curr Biol.* **17** 1298-1307.

The International HapMap Consortium. (2007) A Second Generation Human Haplotype Map of Over 3.1 Million SNPs. *Nature*. **449** 851-861.

Sabeti PC and the International HapMap Consortium. (2007) Genome-wide Detection and Characterization of Positive Selection in Human Populations. *Nature*. **449** 913-918.

Qin ZS, Zhu J, Ulintz PJ, Andrews PC. (2008) Machine Learning Approaches for Peptide Identification based on Mass Spectrometry Database Search Results. *Introduction to Machine Learning and Bioinformatics.* Editors: Sushmita Mitra, Sujay Datta, Theodore Perkins, George Michailidis. CRC Press.

Zandi PP, Zöllner S, Avramopoulos D, Willour VL, Chen Y, **Qin ZS**, Burmeister M, Miao K, *Gopalakrishnan S, McEachin R, Potash JB, DePaulo JR, McInnis MG. (2008) Family-based SNP Association Study on 8q24 in Bipolar Disorder. *Am J Med Genet B: Neuropsychiatr Genet.* **147B** 612-618.

Sandstedt SA, Zhang L, Patel M, McCrea KM, **Qin ZS**, Marrs CF, Gilsdorf JR. (2008) Comparison of Laboratory-based and Phylogenetic Methods to Distinguish Between *Haemophilus influenzae* and *H. haemolyticus.* *J Microbiol Methods.* **75** 369-371.

*Zhou J, **Qin ZS**, Quinney SK, Kim S, Wang Z, Yu M, Chien JY, Lucksiri A, Hall SD, Li L. (2009) A New Probabilistic Rule for Drug-Drug Interaction Prediction. *J Pharmacokinet and Pharmacodyn.* **36** 1-18.

*Hu M, **Qin ZS.** (2009) Query large scale microarray compendium datasets using a model-based Bayesian approach with variable selection. *PLoS ONE.* 4 e4495.

Moroi SE, Raoof DA, Reed DM, Zöllner S, **Qin ZS**, Richards JE. (2009) Progress toward Personalized Medicine for Glaucoma. *Expert Review of Ophthalmology.* **4** 145-161.

*Zhou J, **Qin ZS**, Sara QK, Kim S Wang Z, Hall SD, Li L. (2009) Drug-Drug Interactions Prediction Assessment. *J Biopharm Stat*. **19** 641-657.

*Chen Y, Lin G, Huo JS, Barney D, Wang ZN, Livshiz T, States DJ, **Qin ZS**, Schwartz J. (2009) Computational and Functional Analysis of Growth Hormone-regulated Genes Identifies the Transcriptional Repressor Bcl6 as a Participant in GH-regulated Transcription. *Endocrinology*. **150** 1645-1654.

Gudjonsson JE, Ding J, Li X, Nair R, Tejasvi T, **Qin ZS**, Ghosh D, Aphale A, Gumucio DL, Voorhees JJ, Abecasis G, Elder JT. (2009) Global Gene Expression Analysis Reveals Evidence for Decreased Lipid Biosynthesis and Increased Innate Immunity in Uninvolved Psoriatic Skin. *J Invest Dermatol.* **129** 2795-2804.

*Choi H, Nesvizhskii A, Ghosh D, **Qin ZS.** (2009) Hierarchical Hidden Markov Model with Application to Joint Analysis of ChIP-chip and ChIP-seq Data. *Bioinformatics* **25** 1715-1721.

*Choi H, **Qin ZS**, Ghosh D. (2010) A double-layered mixture model for the joint analysis of DNA copy number and gene expression data. *J. Comput. Biol.* 17 121-137.

*Hu M, Yu J, Taylor, JM, Chinnaiyan AM, **Qin ZS.** (2010) On the Detection and Refinement of Transcription Factor Binding Sites Using ChIP-Seq Data. *Nucleic Acids Res.* **38** 2154-2167.

Yu J, Yu J, Mani R, Cao X, Cao Q, Brenner CJ, Cao X, Wang GX, Wu L, Li J, *Hu M, Gong Y, Cheng H, Laxman B, Vellaichamy A, Shankar S, Li Y, Dhanasekaran SM, Morey R, Barrette T, Lonigro RJ, Tomlins SA, Varambally S, **Qin ZS**, Chinnaiyan, AM. (2010) An Integrated Network of Androgen Receptor, Polycomb and TMPRSS2-ERG Gene Fusion in Prostate Cancer Progression. *Cancer Cell* **17** 443-454. (From the cover with preview article Chen and Sawyers. Coordinate transcriptional regulation by ERG and androgen receptor in fusion-positive prostate cancers (2010). *Cancer Cell* **17** 415-416).

Breitkreutz A, Choi H, Sharom J, Boucher L, Neduva V, Larsen B, Lin Z, Breitkreutz B, Stark C, Liu G , Ahn J, Dewar-Darch D, **Qin ZS**, Pawson T, Gingras A, Nesvizhskii AI, Tyers M. (2010) A global Protein Kinase and Phosphatase Interaction Network in Yeast. *Science* **328** 1043-1046 (with preview article in the same issue Levy SD, Landry CR, Michnick SW (2010). Cell signaling. Signaling through cooperation. *Science* **328** 983-984).

Qin ZS, Yu J, *Shen J, Maher CA, *Hu M, Kalyana-Sundaram S, Yu J, Chinnaiyan AM. (2010) HPeak: An HMM-based Algorithm for Defining Read-enriched Regions in ChIP-Seq Data. *BMC Bioinformatics.* **11** 369.

Ghosh D, **Qin ZS**. (2010) Statistical Issues in the Analysis of ChIP-Seq and RNA-Seq Data. *Genes* **1** 317-334.

Choi H, Larsen B, Lin ZY, Breitkreutz A, Mellacheruvu D, Fermin D, **Qin ZS**, Tyers M, Gingras AC, Nesvizhskii AI. (2011) SAINT: Probabilistic Scoring of Affinity Purification-Mass Spectrometry Data. *Nat Methods*. **8** 70-73.

Cui TX, Lin G, LaPensee CR, Calinescu A, Rathore M, Streeter C, Pwien-Pilipuk G, Lanning N, Jin H, Carter-Su C, **Qin ZS**, Schwartz J. (2011) C/EBPb mediates Growth Hormone-regulated expression of multiple target genes. *Molecular Endocrinology* **25** 681-693.

Kim JH, Dhanasekaran SM, Prensner JR, Cao X, Robinson D, Kalyana-Sundaram S, Shankar S, Jing X, Iyer M, *Hu M, Sam L, Grasso C, Maher CA, Palanisamy N, Mehra R, Huang C, Siddiqui J, Yu Y, **Qin ZS**, Chinnaiyan AM. (2011) Deep Sequencing Reveals Distinct Patterns of DNA Methylation in Prostate Cancer. *Genome Res*. **21** 1028-1041.

Cao Q, Mani R, Ateeq B, Dhanasekaren SM, Asangani IA, Prensner JR, Kim JH, Brenner JC, Jing X, Cao X, Wang R, Li Y, Dahiya A, Wang L, Pandhi M, Lonigro RJ, Wu Y, Tomlins SA, Palanisamy N, **Qin ZS**, Yu J, Maher CA, Varambally S, Chinnaiyan AM. (2011) Coordinated Regulation of Polycomb Group Complexes through microRNAs in Cancer. *Cancer Cell*. **20** 187-199.

Katoh H, **Qin ZS**, Liu R, Wang L, Li W, Li X, Wu L, Du Z, Lyons R, Liu CG, Liu X, Dou Y, Zheng P, Liu Y. (2011) FOXP3 Orchestrates H4K16 Acetylation and H3K4 Trimethylation for Activation of Multiple Genes by Recruiting MOF and Causing Displacement of PLU-1. *Mol Cell* **44** 770-784.

*Hu M, Zhu Y, Taylor JMG, Liu JS, **Qin ZS** (2012). Using Poisson Mixed-effects Model to Quantify Exon-level Gene Expression in RNA-seq. *Bioinformatics* **28** 63-68.

Zhao JC, Yu J, Runkle C, Wu L, Hu M, Wu D, Liu JS, Wang Q, **Qin ZS**, Yu J. (2012). Cooperation Between Polycomb and Androgen Receptor during Oncogenic Transformation. *Genome Res.* **22** 322-331.

*Zandevakili P, *Hu M, **Qin ZS** (2012) GPUmotif: An Ultra Fast and Energy-efficient Motif Analysis Program Using Graphics Processing Units. *PLoS ONE*. **7** e36865.

#Li X, #*Li L, Pandey R, Byun JS, Gardner K, **Qin ZS**, and Dou Y (2012) The Histone Acetyltransferase MOF Is a Key Regulator of the Embryonic Stem Cell Core Transcriptional Network. *Cell Stem Cell* **11** 163-178. (with preview article of Kloc A and Ivanova N (2012) Chromatin and pluripotency: the MYSTerious connection. *Cell Stem Cell.* **11** 139-140).

#Hou C, #*Li L, **#Qin ZS**, Corces, VG. (2012) Gene Density, Transcription and Insulators Contribute to the Partition of the Drosophila Genome into Physical Domains. *Mol Cell*. **48** 471-484 (with preview article of Xu and Felsenfeld (2012) Order from Chaos in the Nucleus. *Mol Cell.* **48** 327-328).

Hu M, Deng K, Selvaraj S, **Qin ZS**, Ren B, Liu JS. (2012) HiCNorm: removing biases in Hi-C data via Poisson regression. *Bioinformatics.* **28** 3131-3133.

Qin ZS, Bilenly M, *Su G, Jones S (2013) MotifOrganizer: A Scalable Two-stage Model-based Clustering Approach for Grouping Conserved Non-coding Elements in Mammalian Genomes. *Frontiers in Biosciences.* **5** 785-797.

*Choi H, Fermin D, Nesvizhskii AI, Ghosh D, **Qin ZS**. (2013). Sparsely-correlated Hidden Markov Models with Application to Genome-wide Location Studies. *Bioinformatics.* **29** 533-541.

Hu M, Deng K, **Qin ZS**, Dixon J, Selvaraj S, Fang J, Ren B, Liu JS. (2013) Bayesian inference of three-dimensional chromosomal organization. *PLoS Comput Biol.* **9** e1002893.

Wu H, **Qin ZS**, Zhu Y. (2013). PM-Seq: Using Finite Poisson Mixture Models for RNA-Seq Data Analysis and Transcript Expression Level Quantification. *Stat. Biosci.* **5.** 71-87.

*Yuan S. **Qin ZS.** (2013) Read-mapping using personalized diploid reference genome for RNA sequencing data reduced bias for detecting allele-specific expression. *BIBM 2012 Workshop on Data-Mining of Next Generation Sequencing.*

Sun Z, Wu H, **Qin ZS**, Zhu Y. (2013) Model-Based Methods for Transcript Expression Level Quantification in RNA-Seq. In *Advances in Statistical Bioinformatics: Models and Integrative Inference for High-Throughput Data.* Edited by Vannucci M, Do K and Qin ZS. Cambridge University Press.

Wu H, **Qin ZS.** (2013) Exploring the co-occurrence patterns of multiple sets of genomic intervals. Biomed. *Res. Int.* **2013:** 617545.

Hu M, Deng K, **Qin ZS**, Liu JS. (2013) Understanding spatial organizations of chromosomes via quantitative analysis of Hi-C data. *Quantitative Biology*. **1**. 156-174.

Sun Y, Tawara I, *Zhao M, **Qin ZS**, Toubai T, Mathewson N, Tamaki H, Nieves E, Chinnaiyan AM, Reddy P. (2013). Allogeneic T cell responses are regulated by a specific miRNA-mRNA network. *Journal of Clinical Investigation*. **123**:4739-54.

Xiang Z, Qin T, **Qin ZS**, He Y. (2013) A genome-wide literature mining system predicts implicit gene-to-gene relationships and networks. *BMC Syst. Biol.* **7.** Suppl **3**:S9.

Cao F, Townsend EC, Karatas H, Xu J, *Li L, Lee S, Liu L, Chen Y, Ouillette P, Zhu J, Hess JL, Atadja P, Lei M, **Qin ZS**, Malek S, Wang S, Dou Y. (2014) Targeting MLL1 H3K4 Methyltransferase Activity in Mixed-Lineage Leukemia. *Molecular Cell*. **53**. 247-261.

Cao Q, Wang X, *Zhao M, *Yang R, Malik R, Qiao Y, Poliakov A, Yocum AK, Li Y, Chen W, Cao X, Jiang X, Dahiya A, Harris C, Feng FY, Kalantry S, **Qin ZS**, Dhanasekaran SM, Chinnaiyan AM. (2014) The central role of EED in the orchestration of polycomb group complexes. *Nature Communications*. **5**:3127.

*Yang R, Bai Y, **Qin ZS**, Yu T. (2014) EgoNet: identification of human disease ego-network modules. *BMC Genomics*. **15**:314.

Asangani IA, Dommeti VL, Wang X, Malik R, Cieslik M, *Yang R, Escara-Wilke J, Wilder-Romans K, Dhanireddy S, Engelke C, Iyer MK, Jing X, Wu YM, Cao X, **Qin ZS**, Wang S, Feng FY, Chinnaiyan AM. (2014) Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature*. **510** 278-282.

Van Bortle K, Nichols M, *Li L, Ong C, Takenaka N, **Qin ZS.** Corces V. (2014) Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol.* **15** R82.

Joseph SJ, *Li B, Ghonasgi T, Haase CP, **Qin ZS**, Dean D, Read TD. (2014) Direct amplification, sequencing and profiling of *Chlamydia trachomatis* strains in single and mixed infection clinical samples. *PLoS ONE* **9** e99290.

*Yang R, *Chen L, Newman S, Gandhi K, Doho G, Moreno CS, Vertino PM, Bernal-Mizarchi L, Lonial S, Boise LH, Rossi M, Kowalski J, **Qin ZS**. (2014) Integrated analysis of whole genome paired-end and mate-pair sequencing data for identifying genomic structural variations in multiple myeloma. *Cancer Informatics*. **13** (Suppl 2):49-53.

Lin G, LaPensee C, **Qin ZS**, Schwartz J. (2014) Reciprocal occupancy of BCL6 and STAT5 on Growth Hormone target genes: contrasting transcriptional outcomes and promoter-specific roles of p300 and HDAC3. *Molecular and Cellular Endocrinology*. **395**: 19-31.

#Wu L, **#***Li L, Zhou B, **Qin ZS**, Dou Y. (2014) H2B ubiquitylation promotes RNA Pol II processivity via PAF1 and pTEFb. *Mol Cell*. **54.** 920-31.

*Chen L, Wang C, **Qin ZS**, Wu H (2015) A novel statistical method for quantitative comparison of multiple ChIP-seq datasets. *Bioinformatics*. **31.** 1889-1896.

#*Xu T, #*Li B, *Zhao M, Szulwach KE, Street RC, Lin L, Yao B, Zhang F, Jin P, Wu H, **Qin ZS** (2015) Base-resolution methylation patterns accurately predict transcription factor bindings in vivo. *Nucleic Acids Research.* **43.** 2757-2766.

*Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong CT, Cubeñas-Potts C, Hu M, Lei EP, Bosc G, **Qin ZS**, Corces VG. (2015) Widespread rearrangement of 3D chromatin organization underlies Polycomb-mediated stress-induced silencing. *Mol. Cell*. **58**. 216-231.

Xia Y, Wang F, Qian M, **Qin ZS**, Deng M. (2015) PDEGEM: Modeling non-uniform read distribution in RNA-seq data. *BMC Med Genomics.* **8 Suppl 2**:S14.

Cui J, Yin Y, Ma Q, Wang G, Olman V, Zhang Y, Chou WC, Hong CS, Zhang C, Cao S, Mao X, Li Y, **Qin ZS**, Zhao S, Jiang J, Hastings P, Li F, Xu Y. (2015) Comprehensive characterization of the genomic alterations in human gastric cancer. *Int J Cancer* **137.** 86-95.

*Yuan S, *Johnston HR, Zhang G, Li Y, Hu Y-J, **Qin ZS.** (2015) One Size Doesn't Fit All - RefEditor: Building Personalized Diploid Reference Genome to Improve Read Mapping and Genotype Calling in Next Generation Sequencing Studies. *PLoS Comput Biol* **11**. e1004448.

Singh AM, Sun Y, *Li L, Zhang W, Wu T, Zhao S, **Qin ZS**, Dalton S. (2015) Cell-Cycle Control of Bivalent Epigenetic Domains Regulates the Exit from Pluripotency. *Stem Cell Reports.* **5**. 323-336.

#Zhao Z, #*Chen L, Dawlaty MM, Pan F, Weeks O, Zhou Y, Cao Z, Shi H, Wang J, Lin L, Chen S, Yuan W, **Qin ZS**, Ni H, Nimer SD, Yang FC, Jaenisch R, Jin P, Xu M. (2015). Combined loss pf Tet1 and Tet2 promotes B cell, but not myeloid malignancies, in mice. *Cell Reports.* **13.** 1-13.

Wu H, *Xu T, Feng H, Li C, *Li B, Yao B, **Qin ZS,** Jin P, Conneely KN. (2015) Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. *Nucleic Acids Research.* **43.** e141.

Xu Z, Zhang G, Jin F, Chen M, Furey TS, Sullivan PF, **Qin ZS**, Hu M, Li Y. (2015) A hidden Markov random field based Bayesian method for the detection of long-range chromosomal interactions in Hi-C Data. *Bioinformatics*. **32.** 650-656.

*Li B, Sun Z, He Q, Sun Z, Zhu Y, **Qin ZS** (2015) Bayesian inference with historical data-based informative priors improves detection of differentially expressed genes. *Bioinformatics*. **32.** 682-689.

*Chen L, **Qin ZS.** (2015) traseR: an R package for performing trait-associated SNP enrichment analysis in genomic intervals. *Bioinformatics.* **32.** 1214-1216.

Kowalski J, Dwivedi B, Newman S, Switchenko J, Pauly R, Gutman David, Arora J, Gandhi K, Ainslie Kylie, Doho G, **Qin ZS**, Moreno, C, Rossi, M, Vertino P, Lonial S, Bernal-Mizrachi L, Boise, L. (2016) Gene Integrated Set Profile Analysis: A Context-Based Approach for Inferring Biological Endpoints. *Nucleic Acids Research.* **44.** e69.

Qin ZS, *Li B, Conneely KN, Wu H, Hu M, Ayyala D, Park Y, Jin VX, Zhang F, Zhang H, *Li Li, Lin S. (2016) Statistical Challenges in Analyzing Methylation and Long-Range Chromosomal Interaction Data. *Statistics in Biosciences.* **8.** 284-309.

Xu J, *Li L, Xiong J, denDekker A, Ye A, Karatas H, Liu L, Wang H, **Qin ZS**, Wang S, Dou Y. (2016) MLL1 and MLL1 fusion proteins have distinct functions in regulating leukemic transcription program. *Cell Discovery.* **2.** 16008.

Chung C, Elrick MJ, Dell'Orco JM, **Qin ZS**, Kalyana-Sundaram S, Chinnaiyan AM, Shakkottai VG, Lieberman AP. (2016) Heat Shock Protein Beta-1 Modifies Anterior to Posterior Purkinje Cell Vulnerability in a Mouse Model of Niemann-Pick Type C Disease. *PLoS Genet.* **12.** e1006042.

Mathias R, Taub M, Gignoux C, Fu W, Musharoff S, O'Connor T, Vergara C, Torgerson D, Pino-Yanes M, Shringarpure S, Huang L, Rafaels N, Boorgula MP, *Johnston H, Ortega V, Levin A, Song W, Torres R, Padhukasahasram B, Eng C, Mejia-Mejia D, Ferguson T, **Qin ZS**, Scott A, Yazdanbakhsh M, Wilson J, Marrugo J, Lange L, Kumar R, Avila P, Williams LK, Watson H, Ware L, Olopade C, Olopade O, Oliveira R, Ober C, Nicolae D, Meyers H, Mayorga A, Knight-Madden J, Hartert T, Hansel N, Foreman M, Ford J, Faruque M, Dunston G, Caraballo L, Burchard E, Bleecker E, Araujo M, Herrera-Paz E, Gietzen K, Grus W, Bamshad M, Bustamante C, Kenny C, Hernandez R, Beaty T, Ruczinski I, Akey J, Barnes K. (2016) A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. *Nature Communications*. **7.** 12522.

Zhang F, Hammack C, Ogden SC, Cheng Y, Lee EM, Wen Z, Qian X, Nguyen HN, Li Y, Yao B, Xu M, Xu T, Chen L, Wang Z, Feng H, Huang WK, Yoon KJ, Shan C, Huang L, **Qin ZS.** Christian KM, Shi PY, Xu M, Xia M, Zheng W, Wu H, Song H, Tang H, Ming GL, Jin P. (2016) Molecular Signatures Associated with ZIKV Exposure in Human Cortical Neural Progenitors. *Nucleic Acids Res.* **44.** 8610-8620.

Joseph S, *Li B, Petit R, **Qin ZS**, Darrow L, Read T. (2016) The single-species metagenome: subtyping Staphylococcus aureus core genome sequences from shotgun metagenomic data. *PeerJ* **4.** e2571.

*Chen L, Jin P, **Qin ZS.** (2016) DIVAN: accurate identification of non-coding disease-specific risk variants using multi-omics profiles. *Genome Biol.* **17.** 252.

Choi H, Ghosh D, **Qin ZS.** (2017) Computationally tractable Multivariate HMM in genome-wide mapping studies. *Methods Mol Biol.* **1552.** 135-148.

Liu L, Xie J, Sun X, Luo K, **Qin ZS**, Liu H. (2017) An approach of identifying differential nucleosome regions in multiple samples. *BMC Genomics.* **18.** 135.

Sun Y, Iyer M, McEachin R, Zhao M, Wu Y-M, Cao X, Oravecz-Wilson K, Zajac C, Wu S-R, Mathewson N, Rossi C, Toubai T, **Qin ZS**, Chinnaiyan AM, Reddy P. (2017) Genome-Wide STAT3 Binding Analysis after Histone Deacetylase Inhibition Reveals Novel Target Genes in Dendritic Cells. *J Innate Immun.* **9.** 126-144.

*Johnston HR, Hu YJ, Gao J, O'Connor TD, Abecasis G, Wojcik GL, Gignoux CR, Gourraud P, Lizee A, Hansen M, Genuario R, Bullis D, Lawley C, Kenny EE, Bustamante C, Beatty TH, Mathias RA, Barnes KC, **Qin ZS**, on behalf of the CAAPA Consortium**.** (2017) Identifying tagging SNPs for African specific genetic variation from the African Diaspora Genome. *Sci Rep.* **7**:46398.

*Li B, *Li Y, **Qin ZS**. (2017) Improving hierarchical models using historical data with applications in high throughput genomics data analysis. *Statistics in Biosciences*. **9.** 73-90.

*Sun X, Pittard WS, *Xu T, *Chen L, Zwick ME, Jiang X, Wang F, **Qin ZS.** (2017) Omicseq: A web-based search engine for exploring omics datasets. *Nucleic Acids Research.* **45.** W445-W452.

Li X, Yao B, *Chen L, Kang Y, Li Y, Cheng Y, Li L, Lin L, Wang Z, Wang M, Pan F, Dai Q, Zhang W, Wu H, Shu Q, **Qin ZS**, He C, Xu M, Jin P. (2017) Ten-eleven translocation 2 interacts with forkhead box O3 and regulates adult neurogenesis. *Nat Commun.* **8:**15903.

Yang S, Chang R, Yang H, Zhao T, Hong Y, Kong HE, *Sun X, **Qin ZS**, Jin P, Li S, Li X. (2017) CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease. *J Clin Invest.* **127.** 2719-2724.

*Sun X, Wang F, **Qin ZS.** (2017) High Performance Merging of Massive Data from Genome-Wide Association Studies. In *Data Management and Analytics for Medicine and Healthcare (DMAH 2017).* Begoli et al. (Eds.) Springer International Publishing AG.

*Chen L and **Qin ZS.** (2017) Using DIVAN to assess disease/trait-associated single nucleotide variants in genome-wide scale. *BMC Res Notes*. **10.** 530.

Karisani P, **Qin ZS**, Agichtein E. (2018) Probabilistic and Machine Learning-based Retrieval Approaches for Biomedical Dataset Retrieval. *Database*. **2018**. 1-12.

*Sun X, Gao J, Jin P, Eng C, Burchard EG, Beaty TH, Ruczinski I, Mathias RA, Barnes KC, Wang F, **Qin ZS** on behalf of CAAPA consortium (2018) Optimized Distributed Systems Achieve Significant Performance Improvement on Sorted Merging of Massive VCF Files. *Gigascience*. **7.** giy052.

Gollavilli PN, Pawar A, Wilder-Romans K, Ramakrishnan N, Engelke CG, Dommeti VL, Krishnamurthy PM, Nallasivam A, Apel I, *Xu T, **Qin ZS**, Feng FY, Asangani IA. (2018) EWS/ETS-driven Ewing Sarcoma requires BET bromodomain proteins**.** *Cancer Res*. **78.** 4760-4773.

Zekavat SM, Ruotsalainen S, Handsaker RE, Alver M, Bloom J, Poterba T, Seed C, Ernst J, Chaffin M, Engreitz J, Peloso GM, Manichaikul A, Yang C, Ryan KA, Fu M, Johnson WC, Tsai M, Budoff M, Vasan RS, Cupples LA, Rotter JI, Rich SS, Post W, Mitchell BD, Correa A, Metspalu A, Wilson JG, Salomaa V, Kellis M, Daly MJ, Neale BM, McCarroll S, Surakka I, Esko T, Ganna A, Ripatti S, Kathiresan S, Natarajan P; NHLBI TOPMed Lipids Working Group. (2018) Deep coverage whole genome sequences and plasma lipoprotein(a) in individuals of European and African ancestries. *Nat Commun.* **9:**2606.

Natarajan P, Peloso GM, Zekavat SM, Montasser M, Ganna A, Chaffin M, Khera AV, Zhou W, Bloom JM, Engreitz JM, Ernst J, O'Connell JR, Ruotsalainen SE, Alver M, Manichaikul A, Johnson WC, Perry JA, Poterba T, Seed C, Surakka IL, Esko T, Ripatti S, Salomaa V, Correa A, Vasan RS, Kellis M, Neale BM, Lander ES, Abecasis G, Mitchell B, Rich SS, Wilson JG, Cupples LA, Rotter JI, Willer CJ, Kathiresan S; NHLBI TOPMed Lipids Working Group. (2018) Deep-coverage whole genome sequences and blood lipids among 16,324 individuals. *Nat Commun.* **9:**3391.

Daya M, Barnes K, Boorgula MP, Campbell M, Chavan S, Lange L, Lange E, Shetty A, Rafaels N, Brunetti T, Konigsberg I, Gignoux C, Beaty T, Ford J, Foster C, Gao L, Hansel N, Mathias R, Vergara C, Ruczinski I, Scott A, Taub M, Salzberg S, Gao J, Hu YJ, **Qin ZS**, *Johnston H, Levin A, Williams LK et al. (2019) Association study in African-admixed populations across the Americas recapitulates asthma risk loci in non-African populations. *Nat Commun.* **10.** 880.

Rowson SA, Bekhbat M, Kelly SD, Binder EB, Hyer MM, Shaw G, Bent MA, Hodes G, Tharp G, Weinshenker D, **Qin ZS**, Neigh GN. (2019) Chronic adolescent stress sex-specifically alters the hippocampal transcriptome in adulthood. *Neuropsychopharmacology.* **44.** 1207-1215.

*Poulet A, *Li B, Dubos T, Rivera-Mulia JC, Gilbert DM, **Qin ZS**. (2019) RT States: systematic annotation of the human genome using cell type-specific replication timing programs. *Bioinformatics.* **35.** 2167-2176.

Dileep V, Wilson KA, Marchal C, Lyu X, Zhao PA, *Li B, *Poulet A, Bartlett DA, Rivera-Mulia JC, **Qin ZS**, Robins AJ, Schulz TC, Kulik MJ, Dalton S, Corces VG, Gilbert DM. (2019) Rapid Irreversible Transcriptional Reprogramming in Human Stem Cells Accompanied by Discordance between Replication Timing and Chromatin Compartment. *Stem Cell Reports.* **13.** 193-206.

*Yang G, Ma A, **Qin ZS.** (2019) An integrated system biology approach yields drug repositioning candidates for the treatment of heart failure. *Frontiers in Genetics.* **10.** 916.

Hou Y, Li F, Zhang R, Li S, Liu H, **Qin ZS**, Sun X. (2019) Integrative characterization of G-Quadruplexes in the three-dimensional chromatin structure. *Epigenetics.* **14.** 894-911.

Zhang R, Hu M, Zhu YM, **Qin ZS**, Deng K, Liu JS. (2020) Inferring Spatial Organization of Individual Topologically Associated Domains via Piecewise Helical Model. *IEEE/ACM Transactions on Computational Biology and Bioinformatics.* **17.** 647-656.

*Xu T, Zheng X, *Li B, Jin P, **Qin ZS**, Wu H. (2020) A comprehensive review of computational prediction of genome-wide features. *Brief Bioinform.* **21.** 120-134.

*Xu T, Jin P, **Qin ZS.** (2020) Regulatory annotation of genomic intervals based on tissue-specific expression QTLs. *Bioinformatics.* **36.** 690-697.

*Yang G, Ma A, **Qin ZS**. *Chen L. (2020) Application of Topic Models to a Compendium of ChIP-Seq Datasets Uncovers Recurrent Transcriptional Regulatory Modules. *Bioinformatics.* **36.** 2352-2358.

Yang S, Yang H, Huang L, *Chen L, **Qin ZS**, Li S, Li X. (2020) Lack of RAN-mediated toxicity in Huntington's disease knock-in mice. *PNAS.* **117.** 4411-4417.

Yang H, Yang S, Jing L, Huang L, *****Chen L, Zhao X, Yang W, Pan Y, Yin P, **Qin ZS**, Li S, Li XJ. (2020) Truncation of mutant Huntingtin in knock-in mice demonstrates exon1 Huntingtin is a key pathogenic form. *Nat Commun.* **11.** 2582.

Liu H, Li H, Sharma A, Luo K, **Qin ZS,** Sun X. (2020) DeconPeaker, a deconvolution model to identify cell types based on chromatin accessibility in ATAC-Seq data of mixture samples. *Frontiers in Genetics.* **11.** 392.

Saba N, Dinasarapu A, Magliocca K, Dwivedi B, Seby S, **Qin ZS**, Patel M, Griffith C, Wang X, El-Deiry M, Steuer C, Kowalski J, Shin D, Zwick M, Chen Z. (2020) Signatures of somatic mutations and gene expression from $p16^{INK4A}$ positive head and neck squamous cell carcinomas (HNSCC). *PLoS ONE*. **15.** e0238497.

*Guo Z, *Cui Y, Shi X, Birchler JA, Albizua I, Sherman SL, **Qin ZS,** Ji T. (2020) An Empirical Bayesian Approach for Testing Gene Expression Fold Change and Its Application in Detecting Global Dosage Effects. *Nucleic Acids Research Genomics and Bioinformatics.* **2.** lqaa072.

He I, Jiang T, Thakur A, Bai Y, **Qin ZS.** Systematic and Comprehensive Survey of Genomic Loci Associated with Complex Diseases and Traits. in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Seoul, Korea (South), 2020 pp. 1447-1452.

Yang J, Zhang D, Motojima M, Kume T, Hou Q, Pan Y, Duan A, Zhang M, Jiang S, Hou J, Shi J, **Qin ZS**, Jin P, Liu Z. (2021) Super-enhancer associated transcription factors maintain transcriptional regulation in mature podocytes. *J Am Soc Nephrol.* **32.** 1323-1337.

Duan A, Wang H, Zhu Y, Wang Q, Zhang J, Hou Q, Xing Y, Shi J, Hou J, **Qin ZS**, Chen Z, Liu Z, Yang J. (2021) Chromatin architecture reveals cell type-specific target genes for kidney disease risk variants. *BMC Biol.* **19.** 38.

Bekhbat M, Mukhara D, Dozmorov M, Stansfield J, Benusa S, Hyer M, Rowson S, Kelly S, **Qin ZS**, Dupree J, Tharp G, Tansey M, Neigh G. (2021) Adolescent Stress Sensitizes the Adult Neuroimmune Transcriptome and Leads to Sex-specific Microglial and Behavioral Phenotypes. *Neuropsychopharmacology.* **46.** 949-958.

Taliun D et al. (2021) Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature.* **590.** 290-299.

Wang H, Duan A, Zhang J, Wang Q, Xing Y, **Qin ZS**, Liu Z, Yang J. (2021) Glucocorticoid receptor wields chromatin interactions to tune transcription for cytoskeleton stabilization in podocyte. *Commun Biol.* **4.** 675.

*Huang Y, *Sun X, *Jiang H, *Yu S, Robins C, Armstrong JM, Li R, Mei Z, Shi X, Gerasimov ES, De Jager PL, Bennett DA, Wingo AP, Jin P, Wingo TS, **Qin ZS** (2021) A machine learning approach to brain epigenetics analysis reveals kinases associated with Alzheimer's disease. *Nat Commun.* **12.** 4472.

*Su K, *Yu Q, Shen R, Sun SY, Moreno CS, Li X, **Qin ZS** (2021) Pan-cancer analysis of pathway-based gene expression pattern at the individual level reveals novel biomarkers of clinical prognosis. *Cell Reports Methods.* **1.** 100050.

*Jin Y, *Jiang J, Wang R, **Qin ZS** (2021) Systematic Evaluation of DNA Sequence Variations on *in vivo* Transcription Factor Binding Affinity. *Front. Genet*. **12.** 667866.

Eldridge R, Uppal K, Hayes DN, Smith MR, Hu X, **Qin ZS**, Beitler J, Miller A, Wommack E, Higgins K, Shin D, Ulrich B, Qian D, Saba N, Bruner D, Jones DP, Xiao C. (2021) Plasma metabolic phenotypes of HPV-associated versus smoking-associated head and neck cancer and patient survival. *Cancer Epidemiology Biomarkers & Prevention.* **10.** 1858-1866.

*Wang B, *Yang J, *Qiu S, Bai Y, **Qin ZS**. (2021) Systematic Exploration in Tissue-pathway Association of Complex Diseases with Comprehensive eQTLs Catalog. *Frontiers in Big Data*. **4.** 719737.

Wang Q, Zhang Y, Zhang B, Fu R, Zhao X, Zhang J, Zuo K, Xing Y, **Qin ZS**, Li E, Guo H, Liu Z, Yang J. (2022). Single-cell chromatin accessibility landscape in kidney identifies additional cell-of-origin in heterogenous papillary renal cell carcinoma. *Nat. Commun*. **13.** 31.

Cao Z, *#Huang Y, Duan R, Jin P, #**Qin ZS**, Zhang S. (2022) Disease Category-specific Annotation of Variants Using an Ensemble Learning Framework. *Brief. Bioinform*. **23.** bbab438.

Eldridge RC, Uppal K, Shokouhi M, Smith MR, Hu X, **Qin ZS**, Jones DP, Hajjar I. (2022) Multiomics analysis of structural MRI brain imaging and cerebrospinal fluid metabolomics in cognitively normal and impaired adults. *Front. Aging Neurosci*. **13.** 796067.

Chandrashekar DS, Karthikeyan SK, Korla PK, Patel H, Shovon AS, Athar M, Netto GJ, **Qin ZS**, Kumar S, Manne U, Crieghton CJ, Varambally S. (2022) UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia*. **25.** 18-27.

Lang L, Loveless R, Dou J, Lam T, Chen A, Wang F, Sun L, Juarez J, **Qin ZS**, Saba NF, Shay C, Teng Y. (2022) ATAD3A mediates activation of RAS-independent mitochondrial ERK1/2 signaling, favoring head and neck cancer development. *J Exp Clin Cancer Res.* **41.** 43.

Liu C, Qian L, Vallega KA, Ma G, Zong D, *Chen L, Wang S, Ramalingam SR, **Qin ZS**, Sun SY. (2022) The novel BET degrader, QCA570, is highly active against the growth of human NSCLC cells and synergizes with osimertinib in suppressing osimertinib-resistant EGFR-mutant NSCLC cells. *Am J Cancer Res.* **12.** 779-792.

*Ma W, Sharma S, Jin P, Gourley SL, **Qin ZS**. (2022) LRcell: detecting the source of differential expression at the sub-cell type level from bulk RNA-seq data. *Brief. Bioinform.* Accepted.

Kong HE, Lim J, Linsalata A, Kang Y, Malik I, Allen EG, Cao Y, Shubeck L, Johnston R, *Huang Y, Gu Y, Guo X, Zwick ME, **Qin ZS**, Wingo TS, Juncos J, Nelson DL, Epstein M, Cutler DJ, Todd PK, Sherman SL, Warren ST, Jin P (2022). Identification of PSMB5 as a genetic modifier of Fragile X-associated Tremor/Ataxia Syndrome. *PNAS*. Accepted.

Submitted

Li RY*, Ma W*, **Qin ZS.** (2022) Approaches to marker gene identification from single-cell RNA-sequencing data.

Sahoo A, Jones AT, Cheedarla N, Gangadhara S, Roy V, Styles TT, Shiferaw A, Walter KL, Williams LD, Shen X, Ozorowski G, Lee WH, Burton S, **Qin ZS**, Derdeyn CA, Ward AB, Clements JD, Varadarajan R, Tomaras GD, Kozlowski PA, Alter G, Amara RR. (2022) MVA/Protein HIV-1 vaccine protects against heterologous SHIV 1 infection by modulating IgG glycosylation and T helper response in macaques.

Welbourn S, Chakraborty S, Yang J, Gleinich A, Gangadhara S, Khan S, Ferrebee C, Yagnik B, Burton S, Charles T, Smith SA, Williams D, Mopuri R, Upadhyay A, Thompson J, Price N, Wang S, **Qin ZS**, Shen X, Williams L, Eisel N, Peters T, Zhang L, Kilembe W, Karita E, Tomaras G, Bosinger S, Amara R, Azadi P, Wright E, Gnanakaran S, Derdeyn C. (2022) A neutralizing antibody target in early HIV-1 infection was recapitulated in rhesus macaques immunized with the transmitted/founder envelope sequence.

Edited book

Advances in Statistical Bioinformatics: Models and Integrative Inference for High-Throughput Data. (2013) Edited by Vannucci M, Do K and **Qin ZS**. Cambridge University Press.

Other publications

Qin ZS, Damien P, Walker S. (2003) Scale Mixture Models with Applications to Bayesian Inference. The Monte Carlo Methods in the Physical Sciences: Celebrating the 50th Anniversary of the Metropolis Algorithm, Los Alamos, New Mexico 2003. Gubernatis JE (Ed) American Institute of Physics. 394-395.

Qin ZS. (2008) Book Review: Data Analysis and Graphics Using R: An Example-Based Approach, Second Edition. John Maindonald and John Braun. *Brief Bioinform*. **9.** 258-259.

Zhu D, **Qin ZS** (2014) Editorial: Systems biology and metagenomics: a showcase of Chinese bioinformatics researchers and their work. *Science China Life Science*. **57.** 1051-1053.

Qin ZS. (2017) Editorial: Special collection of Bioinformatics in the era of precision medicine. *Quant. Biol*. **5.** 277-279.

Wren JD, Bai Y, **Qin ZS**, Yan D, Homayouni R. (2020) Proceedings of the 2019 MidSouth Computational Biology and Bioinformatics Society (MCBIOS) Conference. *BMC Bioinformatics*. **21.** 254.

He I, **Qin ZS**, Bai Y. (2021) Identification of Genetic Variants for Prioritized miRNA-targeted Genes Associated with Complex Traits. *Adv. Sci. Technol. Eng. Syst. J.* **6**. 418-423.

Qin ZS. (2021) A Wonderful time – exciting progress made in the past 20 years in genetics powered by the Human Genome Project. *Quant. Biol*. **9.** 366-370.

Oral presentations (within 10 years)

Keynote lecture

Model-based Methods for Analyzing NGS Data. Mini Conference—Research and Collaboration Forum for Southeastern Researchers in Mathematical Modeling of Biological Systems. Augusta, GA. March 28-29, 2012.

Invited/selected conference presentations

Bayesian Model-based Methods for Analyzing ChIP Sequencing Data. *ENAR 201*0 Spring meeting. New Orleans, LA. March 20-24, 2010.

Bayesian Model-based Methods for Analyzing RNA-Seq Data. *ENAR 2011* Spring meeting. Miami, FL. March 20-23, 2011.

Do chromosomes have "consensus" 3D structures? -- Statistical clues from Hi-C data. Cold Spring Harbor Asia Conferences series 2011: High Throughput Biology. Suzhou, China, April 19-23, 2011.

Bayesian Model-based Methods for Analyzing RNA-seq Data. Sixth International Conference on Dynamic Systems and Applications, Atlanta, GA. May, 25-28, 2011

Bayesian Model-based Methods for Analyzing RNA-seq Data. 2011 International Research Symposium on Frontiers of Statistics. Hefei, China. July 13-18, 2011.

Inference of correlated hidden Markov models with application to genome-wide studies. Banff International Research Station Workshop on Current challenges in statistical learning. Banff, Alberta, Canada. December 11-16, 2011.

Statistical Models for Analyzing Sequencing Applications. *ENAR 2012* Spring meeting. Washington D C., April 1-4, 2012.

Inference of correlated hidden Markov models with application to genome-wide studies. Conference on Statistical Learning and Data Mining. Ann Arbor, MI. June 5 - 7, 2012.

Inference of correlated hidden Markov models with application to genome-wide studies. Second Joint Biostatistics Symposium, Beijing China, July 7-9, 2012.

Towards the understanding of the three-dimensional genome organization, statistical challenges and opportunities for analyzing Hi-C data. Barcelona BioMed Conference on Bayesian methods in Biostatistics and Bioinformatics, Barcelona, Spain, December 17-19, 2012.

Towards the understanding of the three-dimensional genome organization, statistical challenges and opportunities for analyzing Hi-C data. The Second Workshop on Biostatistics and Bioinformatics. Atlanta, GA. May 10 - 12, 2013.

An Alternative to the Bayesian Hierarchical Model and its Application to Microarray Gene Expression Data. The Ninth ICSA International Conference: Challenges of Statistical Methods for Interdisciplinary Research and Big Data. December 20-23, 2013. Hong Kong Baptist University.

Extensions to Hidden Markov Models and Their Application to Integrated Analysis of Multiple Chromatin Immunoprecipitation Data. ENAR 2014, March 16-19, 2014. Baltimore, MD.

Inference of Correlated Hidden Markov Models with Application to ChIP-Seq Compendium Data. The 2014 Workshop on Genomic Signal Processing and Statistics (GENSIPS'14), December 3-5, 2014. Atlanta, GA.

Base-resolution methylation patterns accurately predict transcription factor bindings in vivo. The 2014 International Conference on Intelligent Biology and Medicine (ICIBM 2014), December 4-6. 2014. San Antonio, TX.

Omicseq—A genomics BigData search engine and knowledge discovery system. Tsinghua Sanya International Mathematical Forum on BigData. December 26-29, 2014. Sanya, Hainan, China.

Webinar presentations: A genomics BigData search engine and knowledge discovery system. bioCADDIE (NIH BD2K funded Center of Excellence) December 11, 2014.

Invited workshop tutorial: A genomics BigData search engine and knowledge discovery system. BIO-IT 2015, Boston MA. April 21-23, 2015.

Statistical challenges in analyzing methylation and long-range interaction data. Beyond Bioinformatics Transition Workshop. Research Triangle Park, NC. May 11-13, 2015.

An Alternative to the Bayesian Hierarchical Model and its Application to Microarray Gene Expression Data. International Statistics Workshop. Beijing, China. June 22-26, 2015.

A Machine learning approach for predicting transcription factor binding. Joint Statistical Meeting 2015. Seattle, WA, August 7-13, 2015.

Accurate identification of disease-specific non-coding risk variants based on multi-omics profiles. ENCODE Research Applications and Users Meeting. Stanford, CA. June 6-9, 2016.

traseR: linking genomic intervals to phenotypes using trait-associated GWAS SNPs. International Conference of Translational Bioinformatics (ICBTI-2016). June 23-25, 2016. Guiyang, China.

Improving Hierarchical models using historical data. The Second International Symposium on Data Driven Health and Medicine. June 29-July2, 2016. Shanghai, China.

RefEditor: Building Personalized Diploid Reference Genome to Improve Read Mapping and Genotype Calling in Next Generation Sequencing Studies. 2016 International Workshop on Interdisciplinary Research between Mathematics and Biology. July 15-17, 2016. Peking University, Beijing, China.

Chromosomal interaction patterns across multiple human cell types. Joint Statistical Meeting 2016. July 30 – August 4, Chicago, IL.

Improving hierarchical models using historical data. The 10th ICSA International Conference: Global Growth of Modern Statistics in the 21st Century. December 19-22, 2016. Shanghai, China.

Statistical inference of high throughput data with low sample size. The 5th Workshop on Biostatistics and Bioinformatics. May 6-7, 2017. Atlanta, GA.

Bayesian Hierarchical Model in the Big Data Era –Old Wine in a New Bottle. The Third Kliakhandler Conference International Conference on Bayesian Inference in Statistics and Statistical Genetics. Aug. 16-20, 2017, Michigan Tech. University, Houghton, MI.

Statistical inference of high throughput data with low sample size. Georgia Statistics Day. October, 9 2017. Emory University. Atlanta, GA.

Opportunities and challenges in omics Bigdata. Annual Bioinformatics symposium of the Jiangsu Province. Nanjing, China, November 16, 2017.

Using collections of GWAS variants for data mining and machine learning. Symposium on Neurodegenerative Diseases Changsha, China. December 15-16, 2017.

Using collections of GWAS variants for data mining and machine learning. SRCOS 2018, Virginia Beach, VA. June 3-6, 2018.

Apply machine learning methods to predict new disease variants genome-wide. Machine Learning in Science and Engineering 2018. Carnegie Mellon University, Pittsburgh, PA. June 6-8, 2018.

Predicting the impact of genomic variants on functions and diseases at the Bioinformatics cloud forum. October 23, 2020.

Invited departmental seminars (selected)

Model-based Methods for Mining Large Scale Genomic Data, Department of Statistics, University of Illinois at Urbana Champion, Urbana-Champion, IL, November 8, 2007.

HPeak: A Model-Based Algorithm for Defining Read-Enriched regions. Department of Human Genetics, University of Michigan. Next-Generation Sequencing Seminar Series. Ann Arbor, MI. July 14, 2008.

Detection and Refinement of Transcription Factor Binding Sites Using Hybrid Monte Carlo Method. Center for Clinical Epidemiology and Biostatistics, Department of Biostatistics and Epidemiology. University of Pennsylvania. Philadelphia, PA. March 2, 2010.

Detection and Refinement of Transcription Factor Binding Sites Using Hybrid Monte Carlo Method. Department of Computer Science, University of New Orleans. New Orleans, LA. March 22, 2010.

Detection and Refinement of Transcription Factor Binding Sites Using Hybrid Monte Carlo Method. Institute of Computational Biology, Weill Cornell Medical College. New York, NY. March 25, 2010.

Detection and Refinement of Transcription Factor Binding Sites Using Hybrid Monte Carlo Method. Statistical Bioinformatics Center, Department of Statistics, Purdue University. West Lafayette, IN. March 30, 2010.

Bayesian model-based methods for analyzing ChIP sequencing data. Computational Systems Biology Laboratory, Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA. October 14, 2010.

Statistical Modeling of RNA-seq Data. Division of Biostatistics, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL. October 25, 2010.

Bayesian model-based methods for analyzing ChIP sequencing data. Center for Bioinformatics and Computational Genomics, School of Biology, Georgia Institute of Technology, Atlanta, GA. November 23, 2010.

Bayesian model-based methods for analyzing RNA-Seq data. Department of Mathematics and Statistics, Georgia State University, Atlanta, GA. March 18, 2011.

Using model-based methods to quantify exon-level gene expression from RNA-seq data. Department of Biomathematics, UCLA, Los Angeles, CA. September 29, 2011.

Model-based methods for analyzing NGS data. Department of Statistics, Yale University, New Haven, CT. October 17, 2011.

Model-based methods for analyzing NGS data. Department of Molecular and Computational Biology, University of Southern California, Los Angeles, CA. November 3, 2011.

Model-based methods for analyzing NGS data. Department of Biostatistics, the University of Texas M.D. Anderson Cancer Center, Houston, TX. December 7, 2011.

Inference of correlated hidden Markov models with application to genome-wide studies. Virginia Tech Research Center, Arlington, VA. April 4, 2012.

Towards the understanding of the three-dimensional genome organization, statistical challenges and opportunities for analyzing Hi-C data. Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ. October 24, 2012.

Towards the understanding of the three-dimensional genome organization, statistical challenges and opportunities for analyzing Hi-C data. Computational Systems Biology Laboratory, Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA. November 8, 2012.

How chromosomes fold? What sequencing and Bayesian modeling can tell us about the three-dimensional organization of mammalian genomes. H. Milton Stewart School of Industrial and Systems Engineering , Georgia Institute of Technology, Atlanta, GA. November 27, 2012.

How chromosomes fold? What sequencing and Bayesian modeling can tell us about the three-dimensional organization of mammalian genomes. CAS-MPG Partner institute for Computational Biology, Shanghai Institute for Biological Sciences, Chinese Academy of Sciences, December 27, 2012.

Improving hierarchical model for analyzing high throughput genomics data. Department of Bioinformatics Biostatistics and Bioinformatics, Georgetown University Medical Center, Washington DC. March 22, 2013.

How does the chromosome fold? Bayesian model-based inference of three dimensional mammalian genome organization. Seminar given at computer science seminar series of Middle Tennessee State University. January 24, 2014.

What to do when you have a million genomic datasets? Department of Biomedical Informatics, UC San Diego, May 22, 2015.

Improving hierarchical model using historical data. Department of Biostatistics and medical Informatics. University of Wisconsin, Madison. November 20, 2015.

Improving hierarchical model using historical data. School of Life Sciences. Peking University. Beijing, China. December 23, 2015.

Base-resolution methylation patterns accurately predict transcription factor bindings in vivo. Institute of Bioinformatics. Tsinghua University. Beijing, China. December 24, 2015.

DIVAN: Accurate identification of non-coding disease-specific risk variants based on multi-omics profiles. Finnish Institute of Molecular Medicine. Helsinki, Finland. December 1, 2016.

Utilizing Big Data to solve small data inference problem – Alternatives to hierarchical models with applications to genomics data. Department of Statistics, Tel Aviv University, Tel Aviv, June 6, 2017.

Utilizing historical data to aid statistical inference of high throughput data with low sample size. Department of Mathematics and Statistics. Auburn University. November 10, 2017.

Improving Hierarchical Models Using Historical Data with Applications to High-Throughput Genomics Data. Department of Statistics, University of Missouri. Columbia, MO, March 11, 2019.

Harnessing machine learning and omics big data to predict novel risk loci of complex human diseases. Department of Biomedical Informatics. The Ohio State University College of Medicine. Columbus, OH, March 22, 2019.

Harnessing machine learning and omics big data to discover novel risk loci of complex human diseases. Distinguished Lecture Series of The Molecular Basis of Disease Area of Focus (MBDAF). Georgia State University. Atlanta, GA, April 18, 2019.

Harnessing machine learning and omics big data to discover novel risk loci of complex human diseases. School of Biomedical Informatics. The University of Texas Health Science Center at Houston. October 15, 2019.

Harnessing machine learning and omics big data to better understand complex human diseases. Center for Computational Biology and Bioinformatics Indiana University School of Medicine. Indianapolis, IN. October 28, 2019.

Harnessing Public Genomics Big Data to Gain Functional Insights on Complex Diseases. PowerTalk Seminars 2021-2022. Informatics Institute, University of Alabama at Birmingham School of Medicine. September 24, 2021.

Poster presentations

MotifOrganizer: A Scalable Multi-stage Model-based Clustering Approach for Grouping Conserved Noncoding Elements in Mammalian Genomes. Presented at the Conference on Emerging Design and Analysis Issues in genomics Studies in Population Sciences. Boston, MA. October 11-12, 2007.

Molecular Cross-talk of Androgen Receptor and TMPRSS2:ERG Gene Fusion Product in Prostate Cancer, co-presented with Dr. Jindan Yu. 2008 Cold Spring Harbor Meeting on System Biology: Global Regulation of Gene Expression. Cold Spring Harbor, NY. March 27-30, 2008.

Detection and Refinement of Transcription Factor Binding Sites Using Hybrid Monte Carlo Method. Pacific Symposium on Biocomputing (PSB) 2010. Hawaii, HI. January 4-8, 2010.

Detection and Refinement of Transcription Factor Binding Sites Using Hybrid Monte Carlo Method. 2010 Cold Spring Harbor Meeting on System Biology: Global Regulation of Gene Expression. Cold Spring Harbor, NY. March 23-27, 2010.

Detection and Refinement of Transcription Factor Binding Sites Using Hybrid Monte Carlo Method. Frontiers in Mathematical Biology: NSF-NIH PIs Meeting 2010. College Park, MD. April 26-27, 2010

Discovering combinatorial patterns of chromatin marks using advanced Hidden Markov models. Cold Spring Harbor Asia Meeting on Epigenetics, Chromatin & Transcription. Su Zhou, China. April 23-27, 2012.

Omicseq—A novel framework for mining genomics big data. Biological Data Science, Cold Spring Harbor Laboratory, November 6-9, 2014.

A novel framework for mining genomics big data. Big Data Symposium, Athens, GA, October 12, 2015.

RefEditor: building personalized diploid reference genome to improve read mapping and genotype calling in next generation sequencing studies. Cold Spring Harbor Asia Conference on Big Data, Computation and Systems Biology in *Cancer*. Suzhou, China. December 2-5, 2015.

Software developed

- **EM-DeCODER** (haplotype phasing using EM algorithm).
- **HAPLOTYPER** (haplotype phasing using Partition-Ligation and Gibbs Sampler).
- **PL-EM** (haplotype phasing using Partition-Ligation and EM algorithm).
- **BMC** (putative binding motif clustering using Gibbs Sampler).
- **FESTA** (tagSNP selection based on pairwise LD) With Shyam Gopalakrishnan.
- **tripleM** (PL-EM for phasing trios).
- **CRC** (model-based clustering algorithm for gene expression microarray data).

- **OperonHMM** (a hidden Markov model for predicting operon structure using only sequence data). Co-developed with Dr. Nick Bergman.
- **CRCView** (point-and-click web server for clustering, visualizing and interpreting microarray gene expression data). Co-developed with Zuoshuang Xiang and Yongqun He.
- **BEST** (query large scale microarray compendium datasets using model-based Bayesian approach with variable selection). Developed by Ming Hu.
- **HPeak** (A hidden Markov model-based algorithm for calling ChIP-enriched peaks using ChIP-seq data).
- **ChIP-meta** (hierarchical hidden Markov model-based algorithm for inferring ChIP-enriched regions using both ChIP-seq and ChIP-chip data). Developed by Hyung Won Choi.
- **HMS** (a hybrid Monte Carlo-based *de nono* motif discovery tool designed specifically for analyzing ChIP-Seq data). Developed by Ming Hu.
- **POME** (Poisson mixed-effects model for quantification expression in RNA-Seq data). Developed by Ming Hu in collaboration with Michael Yu Zhu.
- **GPUmotif** (*de novo* motif finding and motif scan using GPU). Developed by Pooya Zandevakili.
- **Methylphet** (predict TF binding in vivo from base-level methylome profiling data). Developed by Tianlei Xu, Ben Li and Meng Zhao.
- **RefEditor** (customize reference genome using individual's known mutation profile). Developed by Shuai Yuan.
- **IPBT** (detect differentially expressed genes with informative prior obtained from historical data). Developed by Ben Li.
- **traseR** (to perform trait associated SNP enrichment analysis). Developed by Li Chen.
- **DIVAN** (provide disease-specific annotation on noncoding variants using epigenomics profiles). Developed by Li Chen.
- **Loci2path** (to perform tissue-specific eQTL enrichment analysis for functional annotation). Developed by Tianlei Xu.
- **EWASplus** (extend EWAS coverage from methylation array to all CpGs in the genome). Developed by Yanting Huang.
- **iPath** (identify pathway/gene sets for which their expression pattern is associated with clinical prognostic). Developed by Kenong Su.
- **CASAVA** (provide disease category-specific annotation of non-coding variants using genomics and epigenomics features). Developed by Zhen Cao in Dr. Shihua Zhang's group.
- **LRcell** (detecting the source of differential expression at the sub-cell type level from bulk RNA-seq data). developed by Wenjing Ma.

Professional services

Board of Directors

MidSouth Computational Biology and Bioinformatics Society (MCBIOS) 2019 – present

Editorial board

Plant Cell. 2011 – 2015
Statistics in BioScience, 2015 – present
PLoS ONE. 2016 – present
PeerJ. 2017 – present.
Quantitative Biology 2018- present.

Review editor board

Frontier in Statistical Genetics and Methodology (a specialty of Frontiers in Genetics). 2011 -present

Statistical Advisor

PLoS ONE. 2013 – present

Guest editor

Science China Life Science. A special issue on Bioinformatics. 2014

Quantitative Biology. A special issue on Bioinformatics in the era of Precision Medicine. 2017.

Reviewer for journals (selected):

American Journal of Human Genetics, Annals of Applied Statistics, Annals of Human Genetics, Bayesian Analysis, Bioinformatics, Biostatistics, BMC Bioinformatics Biometrics, Biometrika, BioTechniques, Briefings in Bioinformatics, Cancer Epidemiology, Biomarkers & Prevention, Briefings in Bioinformatics, Cancer Cell, Cancer Informatics, Cancer Research, Circulation, Computational Statistics, Diabetes, F1000Research, Frontiers in Biosciences, Frontiers in Genetics, Frontiers in Plant Sciences, Genetics, Genetic Epidemiology, Genome Biology, Genome Research, Genomics, Human Genetics, Human Molecular Genetics, Human Heredity, IEEE IEEE/ACM Transactions on Computational Biology and Bioinformatics, The International Journal of Bioinformatics Research and Applications, Journal of the American Statistical Association, Journal of Bioinformatics and Computational Biology, Journal of Clinical Oncology, Journal of Computational and Graphical Statistics, Mammalian Genome, Nature, Nature Communications, Nature Methods, Nucleic Acids Research, Pharmacogenomics, PLoS Computational Biology, PLoS Genetics, PLoS ONE, PLoS Physiology, Proceedings of National Academy of Science, Scandinavian Journal of Statistics, Science Advances, Scientific Reports, Statistica Sinica, Statistical Application in Genetics and Molecular Biology, Statistical Analysis and Data Mining, Statistics in Biosciences.

Reviewer for conference:

RECOMB 2007, RECOMB-Regulatory Genomics 2007, The International Conference on Bioinformatics (InCoB) 2007, Life System Modeling and Simulation (LSMS) 2007, Bioinformatics Research and Development (BIRD) 2008, ISMB/ECCB 2009, ISMB 2010, BIBM 2011, ECCB 2012. ICSA 2016. BioKDD 2021

Reviewer for book proposal:

Springer.

Grant review:

External:

Ad hoc biostatistics reviewer, NIH study section CICS, 2005-2011, 2013, 2014.
Ad hoc reviewer NIH study section BDMA, 2011.
Reviewer for
NIH special emphasis panel HEMT-D, 2008.
NIH special emphasis panel VH-E 90S, 2009.
NIH special emphasis panel ZRG1 HDM-S (02) 2011.
NIH special emphasis panel ZCA1 RPRB-B (O1) 2013.
NIH special emphasis panel for R01, 2014.
NIH special emphasis panel ZCA1 SRB-X (J1) 2015.
NIH special emphasis panel ZRG1 GGG-R (50) 2015,
NIH special emphasis panel ZCA1 RPRB-B (M1) 2016, 2019.
NIH special emphasis panel ZRG1 CVRS-E (02) 2016.
NIH special emphasis panel for U24 2016.
NIH special emphasis panel ZHG1 HGR-M (O1) 2016.

NIH special emphasis panel ZCA1 TCRB-T (O2) 2019.
NIH special emphasis panel ZCA1 SRB-2 (M1) 2020.
NIH special emphasis panel ZCA1 TCRB-Q (M1) 2020.

Mail reviewer, NIH RC1 and RC4 grants 2009, 2010.
Reviewer, Wellcome Trust Research Career Development Fellowship 2004-2005.
Reviewer, NSF 2005, 2006.
Reviewer, South Carolina EPSCoR/IDeA Program 2012.
Reviewer, Blue Cross Blue Shield of Kansas City Proposal 2016.
Reviewer, Human Frontier Science Program 2012.
Reviewer for US-Israel Binational Science Foundation Proposal 2017.

Internal:

Reviewer, University of Michigan CCMB Pilot Grant. 2006 - 2008.
Reviewer, Georgia CTSA Biostatistics, Epidemiology & Research Design (BERD) Program 2018.
Reviewer, Emory Synergy II/Nexus Award 2019.
Reviewer, Rollins School of Public health Dean's Pilot Award 2020.
Reviewer, Emory Synergy Award Round 9 2020.
Reviewer, HERCULES Pilot Award 2020 - 2021.
Reviewer Office of SVPR grant 2021

Conferences organization committee

The First Wuxi International Statistics Forum, July 17-19, 2011, Wuxi, China.

Ga Tech/Emory Bioinformatics Conference. November 10-12, 2011, Atlanta, GA.

BIBM 2011 Workshop on Next-generation Sequencing Analysis. November 12, 2012, Atlanta, GA.

BIBM 2012 Workshop on Data Mining of Next-generation Sequencing. October 4-7, 2012, Philadelphia, PA.

The Second Taihu International Statistics Forum, July 6-8, 2013, Su Zhou, China.

ACM-BCB 2015. September 9-12. Ga Tech, Atlanta, GA.

The Third Taihu International Statistics Forum, July 9-11, 2016, Shanghai, China.

BioKDD 2021: 20th International Workshop on Data Mining in Bioinformatics, Virtual, August 15, 2021.

Conference program participation

Organized an invited session "Bayesian methods applied to genomics and genetics" in Eastern North American Region/International Biometric Society (ENAR) 2012, April 1-4, 2012, Washington D.C.

Organized an invited session "Statistical method in high throughput biology" for the International Chinese Statistical Association (ICSA) Applied Statistics Symposium, June 23-26, 2012, Boston MA.

Organized an invited session "Recent advanced in statistical genetics and genomics" in IMS-China 2013, June 30-July 4, 2013, Chengdu, China.

Participated in the planning committee of SAMSI Bioinformatics planning meeting. May 31 2013, Research Triangle Park, NC.

Served on the student travel award committee of the ACM-BCB 2015. September 9-12. Ga Tech, Atlanta, GA.

Organized and chaired an invited session "New informatics methods for precision medicine" for the Mid-South Computational Biology and Bioinformatics Society (MCBIOS) 2019 annual conference, March 28-30, 2019, Birmingham, AL.

AI against Cancer Hackathon 2021. Virtual, August 10-13, 2021.

University-wide committee:

Computational Life Sciences faculty search committee 2011.
Big Data/Data Science Working Group 2015.

Teaching

At Emory University

BIOS 710 Probability Theory II
Fall 2016 (enrollment: 13).

BIOS 760R Special Topics in Biostatistics: Bayesian Nonparametrics.
Spring 2012 (enrollment: 7).

BIOS 738 Bayesian and Empirical Bayes Methods
Spring 2013 (enrollment 18).
Fall 2015 (enrollment 15).
Spring 2017 (enrollment 23).
Spring 2018 (enrollment 18).
Spring 2019 (enrollment 22).

BIOS 516 Introduction to Large-scale Biomedical Data Analysis.
Co-taught with Hao Wu, Tianwei Yu and Yijuan Hu.
Fall 2014 (Enrollment: 10).
Fall 2017 (Enrollment: 6).
Fall 2019 (Enrollment: 10).
Fall 2021 (Enrollment: 41).

BIOS 731 Advanced Statistical Computing
Co-taught with Hao Wu.
Fall 2016 (Enrollment: 18).
Fall 2018 (Enrollment: 18).
Fall 2020 (Enrollment: 16).

BIOS 506 Advanced Statistical Computing Methods.
Co-taught with Hao Wu, Tianwei Yu, Howard Chang and Lance Waller.
Fall 2018 (Enrollment: 23).

BIOS 534 / CS 534 Machine Learning.
Spring 2020 Co-taught with Tianwei Yu, Steve Pittard. (Enrollment: 45).
Spring 2021 Co-taught with Steve Pittard. (Enrollment: 44).
Spring 2022 Co-taught with Steve Pittard. (Enrollment: 52).

BIOS 540 Introduction to Bioinformatics.
Fall 2020 (Enrollment: 15).

At University of Michigan

BIOSTAT 601 Probability and Distribution Theory.
Fall 2006 (enrollment: 57 (2 sessions)),
Fall 2008 (enrollment: 41),
Fall 2009 (enrollment: 40).

BIOSTAT 646 / STAT 545 / BIOINFO 545 Data Analysis in Molecular Biology (inaugural class designed by myself):
Winter 2004 (enrollment: 25),
Winter 2005 (enrollment: 25),
Winter 2008 (enrollment: 14),
Winter 2009 (enrollment: 18).

BIOSTAT 666 Statistical Models and Numerical Methods in Human Genetics.
Winter 2008 (enrollment: 14).

BIOSTAT 682 Applied Bayesian Inference.
Winter 2005 (enrollment: 16),
Winter 2006 (enrollment: 27).

BIOSTAT 830 Special Topics in Statistical Genetics.
inaugural class designed by myself:
Winter 2006 (enrollment: 9).

Postdoctoral fellows supervised

Meng Zhao (2011 - 2013)
Rendong Yang (2011 - 2014)
Li Li (2011- 2016)
Rich Johnston (2012 – 2016)
Axel Poulet (2016-2019)

Doctoral committee service

At Emory University

Committee chair or co-chair:

Shuai Yuan (2014)
Li Chen (2017)
Ben Li (2017)
Tianlei Xu (2018)

Xiaobo Sun (2018)
Yanting Huang (2021)
Ronnie Li (2023)

Committee member:

Hesen Peng (2012)
Brooke Weckbelblatt, Human Genetics (2015)
Yunxuan Jiang (2017)
Jung-Ting Chien, MathCS (2017)
Peizhou Liao (2017)
Qing He (2017)
Zhuxuan Jin (2017)
Hao Feng (2018)
Qingpo Cai (2018)
Yunxiao Li (2019)
Yunchuan Kong (2020)
Kenong Su (2021)
Zhengyi Zhu (2021)
Sumeet Sharma, Medical School (2019)
Bonggun Shin, CSI (2020)
Kaitlyn Roman, GDBBS (2021)
Danielle Clarkson-Townsend, EHS (2021)
Zhenxing Guo (2022)
Ziyue Wu (2022)
Cynthia Perez, GMB (2024)
Taylor Head (2024)
Luxiao Chen (2024)
Hannah Waddel (2025)

MPH/MSPH committee service

Committee chair:

Quran Wu (2017)
Yutong Jin (2018)
Wanqi Chen (2019)
Xiaozhu Zhang (2019)
Chen Zhao (2020)
Qi Yu (2020)
Yifei Han (2020)
Qingyu Wang (2020)
Lidan Zhang (2020)
Yanlong Yu (2020)
Nanxi Guo (2021)
Jiahui Jiang (2021)
Anqi Dai (2021)
Yawei Wang (2021)
Runyuan Zhang (2021)

Committee member:

Luxiao Chen (2018), Bokai Zhao (2019), Mengda Yu (2020), Can Li (2020), Weishan Song (2020), Xin Wei (2021)

At University of Michigan and other institutes

Committee chair or co-chair:

Jihao Zhou, Biostatistics (co-chair with Lang Li 2007)
Hyung Won Choi, Biostatistics (co-chair with Debashis Ghosh 2008)
Ming Hu, Biostatistics (2010)
Shyam Gopalakrishnan, Biostatistics (2012)

Committee member:

Lei Xu, Biostatistics (2007)
Xiaoxi Zhang, Biostatistics (2007)
Dongxiao Zhu, Bioinformatics (2006)
Ji Chen, Bioinformatics (2006)
Yili Chen, Bioinformatics (2007)
Xing Li, Bioinformatics (2008)
Shankar Subramanian, Bioinformatics (2008)
Ben-Yang Liao, EEB (2008)
Hokeun Sun, Statistics (2008)
Victorya Strumba, Bioinformatics (2009)
Jinyao Zhang, Physics (2009)
Jin Zheng, Biostatistics (2009)
Zhi Wang, Ecology and Evolutionary Biology (2010)
Grace Lin, Cellular and Molecular Biology (2012)
Douglas D. Baumann (Department of Statistics, Purdue University, 2012)

Bioinformatics doctoral preliminary examination committee service

Jung Hei Kim (2004), Andrew Hodges (2006), Jenna Vanliere (2007), Sirarat Sarntivijai (2007), Gang Su (2008), Junguk Hur (2008), Youngsheng Huang (2008), Yu-Hsuan Lin (2008), Ryan Welch (2008), Arun Manoharan (2008), Yasin Senbabaoglu (2008), Erin R. Shellman (2009), Chunchao Zhang (2009), Lee Sam (2009). Yan Zhang (2009)