

# HAO WU

*Last updated March 8, 2017*

---

<i>Mailing address</i>	Department of Biostatistics and Bioinformatics Rollins School of Public Health Emory University 1518 Clifton Rd., NE Atlanta, GA 30322
<i>Phone</i>	(404)727-8633
<i>Email</i>	<a href="mailto:hao.wu@emory.edu">hao.wu@emory.edu</a>
<i>Website</i>	<a href="http://web1.sph.emory.edu/users/hwu30/">http://web1.sph.emory.edu/users/hwu30/</a>

## Education

- **B.S. Electrical Engineering**, 1996, Tsinghua University, Beijing, China.
- **M.S. Electrical Engineering**, 2000, Iowa State University, Ames, Iowa.  
Minor in Economics. Advisor: Gerald B. Sheblé.  
Thesis: *Enhancing electricity auction mechanism with FACTS devices*.
- **MHS Bioinformatics**, 2010. Johns Hopkins University, Baltimore, MD.
- **Ph.D. Biostatistics**, 2010. Johns Hopkins University, Baltimore, MD. GPA 4.0.  
Advisors: Rafael A. Irizarry, Hongkai Ji.

## Professional Experience

- **July 1996–July 1998, Research Assistant**, Institute of Electrical Engineering, Chinese Academy of Science, Beijing, China.
- **February 2000–July 2001, Staff Engineer**, GE-Harris Energy Control Systems (Now GE Network Solution), Melbourne, Florida.
- **July 2001–August 2005, Scientific Software Engineer**, The Jackson Laboratory, Bar Harbor, Maine.
- **July 2010–August 2016, Assistant Professor**, Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia.
- **August 2011–present, Affiliated faculty**, Graduate Program in Population Biology, Ecology, & Evolution (PBEE), Emory University, Atlanta, Georgia
- **September 2016–present, Associate Professor (with tenure)**, Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia.
- **September 2016–present, Adjunct Associate Professor**, Department of Mathematics and Computer Science, Emory University, Atlanta, Georgia.

## Research Interests

- Statistical Methodology: Bioinformatics, computational biology, genomics, analyses of high-throughput genomics and epigenomics data (microarrays, RNA-seq, ChIP-seq, BS-seq), biomarker discovery, machine learning and large-scale data mining.
- Applications: Epigenetics, cell differentiation, development, and reprogramming; mental diseases (Alzheimer's disease, ataxia, autism); cancers.

## Publications

(\* corresponding author; §: equal contribution; underline: lab member, advisee, or student under my guidance.)

Google scholar citation statistics (as of March 2017): All citations: 5272. h-index: 24. i10-index: 34.

### Peer-Reviewed Journal - statistical methodology

1. Broman KW, **Wu H**, Sen S, Churchill GA (2003) *R/qtl: QTL mapping in experimental crosses*. **Bioinformatics** 19:889-890.
2. Yandell BS, Mehta T, Banerjee S, Shriner D, Venkataraman R, Moon JY, Neely WW, **Wu H**, von Smith R, Yi N (2007) *R/qtlbim: QTL with Bayesian Interval Mapping in experimental crosses*. **Bioinformatics** 23(5):641-643.
3. Frangakis CE, **Wu H** (2007) *The geometry of inadmissibility of independent observations for estimating a single parameter in two-parameter ordered symmetric problems*. **Metron** Vol. LXV n.3.
4. Irizarry RA, **Wu H**, Feinberg AP (2009) *A species-generalized probabilistic model-based definition of CpG islands*. **Mammalian Genome** 20:674–680.
5. **Wu H**, Jaffe HA, Feinberg AP, Irizarry RA (2010) *Redefining CpG Islands Using a Hidden Markov Model*. **Biostatistics**, 11(3): 499–514.
6. **Wu H**, Irizarry, R.A. and Bravo, H.C. (2010) *Intensity normalization improves color calling in SOLiD sequencing*, **Nature Method** 7:336–337.
7. **Wu H**, Ji H (2010) *JAMIE: Joint analysis of multiple ChIP-chip experiments*. **Bioinformatics** 26:1864–1870.
8. **Wu H**, Wang C, Wu Z (2012) *A new shrinkage estimator for dispersion improves differential expression detection in RNA-seq data*. **Biostatistics**, 14(2):232-43. doi: 10.1093/biostatistics/kxs033.
9. **Wu H\***, Qin ZS\* (2013) *Exploring the Cooccurrence Patterns of Multiple Sets of Genomic Intervals*. **BioMed Research International**. 2013:617545. doi: 10.1155/2013/617545.
10. **Wu H\***, Ji HK\* (2014) *PolyaPeak: detecting transcription factor binding sites from ChIP-seq using peak shape information*. **PLoS One**. 9(3): e89694. doi:10.1371/journal.pone.0089694.
11. Feng H, Conneely KN\*, **Wu H\*** (2014) *A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data*. **Nucleic Acids Research**. 42(8):e69. doi: 10.1093/nar/gku154.
12. **Wu H\***, Wang C, Wu Z\* (2014) *PROPER: Comprehensive Power Evaluation for Differential Expression using RNA-seq*. **Bioinformatics**. doi: 10.1093/bioinformatics/btu640.
13. Chen L, Wang C, Qin Z, **Wu H\*** (2015) *A novel statistical method for quantitative comparison of multiple ChIP-seq datasets*. **Bioinformatics**. doi: 10.1093/bioinformatics/btv094.
14. Xu T, Li B, Zhao M, Szulwach LE, Street RC, Lin L, Yao B, Jin P\*, **Wu H\***, Qin Z\* (2015) *Base-resolution methylation patterns accurately predict transcription factor bindings in vivo*. **Nucleic Acids Research** 43(5), 2757-2766. doi: 10.1093/nar/gkv151.
15. Zhang N, Wu H-J, Zhang W, Wang J, **Wu H\***, Zheng X\* (2015) *Predicting tumor purity from methylation microarray data*. **Bioinformatics** 31(21), 3401-3405. doi: 10.1093/bioinformatics/btv370.

16. **Wu H\***, Xu T, Feng H, Chen L, Li B, Yao B, Qin S, Jin P, Conneely KN (2015) *Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicate*. **Nucleic Acids Research**. doi: 10.1093/nar/gkv715.
17. Chong E, Huang Y, **Wu H**, Ghasemzadeh N, Uppal K, Quyyumi A, Jones D, Yu, T. (2015) *Local false discovery rate estimation using feature reliability in LC/MS metabolomics data*. **Scientific Report** 5:17221. doi: 10.1038/srep17221.
18. Park Y\*, **Wu H\*** (2016) *Differential methylation analysis for BS-seq data under general experimental design*. **Bioinformatics** 32 (10), 1446-1453. doi:10.1093/bioinformatics/btw026.
19. Wei Y\*, **Wu H\*** (2016) *Measuring the spatial correlations of protein binding sites*. **Bioinformatics**. doi:10.1093/bioinformatics/btw058.
20. Qin Z, Li B, Conneely KN, **Wu H**, Hu M, Ayyala D, Park Y, Jin VX, Zhang F, Zhang H, Li L, Lin S (2016) *Statistical challenges in analyzing methylation and long-range chromosomal interaction data*. **Statistics in Biosciences**. doi: 10.1007/s12561-016-9145-0.
21. Wang F, Zhang N, Wang J, **Wu H\***, Zheng X\* (2016) *Tumor purity and differential methylation in cancer epigenomics*. **Briefings in Functional Genomics**. doi: 10.1093/bfpg/elw016.
22. Liao P, **Wu H\***, Yu T\*. (2016) *ROC Curve Analysis in the Presence of Imperfect Reference Standards*. **Statistics in Biosciences**. DOI: 10.1007/s12561-016-9159-7.
23. Wang H, Horbinski C, **Wu H**, Liu Y, Sheng S, Liu J, Weiss H, Stromberg AJ, and Wang C. (2016) *NanoStringDiff: A Novel Statistical Method for Differential Expression Analysis Based on NanoString nCounter Data*. **Nucleic Acids Research**. doi: 10.1093/nar/gkw677.
24. Zheng X\*, Zhang N, Wu HJ, **Wu H\***. (2017) *Estimating and accounting for tumor purity in cancer methylation microarray analysis*. **Genome Biology** 18:17. doi: 10.1186/s13059-016-1143-5.
25. Hong C, Chen Y, Ning Y, Wang S, **Wu H**, Carroll RJ. (2017) *PLEMT: A novel pseudolikelihood based EM test for homogeneity in generalized exponential tilt mixture models*. In press. **J. Am. Statist. Assoc.**
26. Wang Y, **Wu H\***, Yu T\*. (2017) *Differential Gene Network Analysis from Single Cell RNA-Seq*. Accepted. **Journal of Genetics and Genomics**.

### **Peer-Reviewed Journal - project-targeted methodology**

Collaborative researches that require methodological development and extensive data analyses

27. Wen B<sup>§</sup>, **Wu H<sup>§</sup>**, Bjornsson H, Green RD, Irizarry RA, Feinberg AP (2008) *Overlapping euchromatin/heterochromatin marks are enriched in imprinted gene regions and predict allele-specific modification*. **Genome Research** 18(11):1806–1813.
28. Wen B, **Wu H**, Irizarry RA, Shinkai Y, Feinberg AP (2009) *Large histone H3 lysine-9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells*. **Nature Genetics** 41:246–250. Cover story.
29. McDonald OG, **Wu H**, Timp W, Doi A, Feinberg AP (2011), *Genome-scale epigenetic reprogramming during epithelial to mesenchymal transition*. **Nature Structural & Molecular Biology** 18(8):867-74.

30. **Wu H**, Wu MC, Zhi D, Santorico SA, Cui X (2012). *Statistical analysis for next generation sequencing – meeting report*. **Frontiers in Genetics**, 3:128.
31. Wen B<sup>§</sup>, **Wu H**<sup>§</sup>, Loh Y<sup>§</sup>, Briem E, Daley GQ, Feinberg AP (2012) *Euchromatin islands in large heterochromatin domains are enriched for CTCF binding and differentially DNA-methylated regions*. **BMC Genomics**, 13:566. doi: 10.1186/1471-2164-13-566.
32. Wang T, **Wu H**, Li Y, Szulwach KE, Lin L, Li X, Chen P, Goldlust IS, Chamberlain SJ, Dodd A, Gong H, Ananiev G, Han JW, Yoon Y, Rudd MK, Yu M, Song CX, He C, Chang Q, Warren ST, Jin P (2013) *Subtelomeric hotspots of aberrant 5-hydroxymethylcytosine-mediated epigenetic modifications during reprogramming to pluripotency*. **Nature Cell Biology**. 15(6):700-11. doi: 10.1038/ncb2748.
33. Tilghman J, **Wu H**, Sang Y, Shi X, Guerrero-Cazares H, Quinones-Hinojosa A, Eberhart C, Laterra J, Ying M (2014) *A HMMR maintains the stemness and tumorigenicity of glioblastoma stem-like cells*. **Cancer Research**. 74(11):3168-79. .

#### **Peer-Reviewed Journal - collaborative**

34. Irizarry RA, Ladd-Acosta C, Carvalho B, **Wu H**, Brandenburg SA, Wen B, Feinberg AP (2008) *Comprehensive high-throughput arrays for relative methylation (CHARM)*. **Genome Research** 18(5):780-790.
35. Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, Wen B, **Wu H**, Briem E, Irizarry RA, Feinberg AP (2011) *Generalized Loss of Stability of Epigenetic Domains Across Cancer Types*. **Nature Genetics** 26;43(8):768-75.
36. Szulwach KE, Li X, Li Y, Song CX, **Wu H**, Dai Q, Irier H, Upadhyay AK, Gearing M, Levey AI, Vasanthakumar A, Godley LA, Chang Q, Cheng X, He C, Jin P (2011) *5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging*. **Nature Neuroscience** 14:1607-1616, doi:10.1038/nn.2959.
37. Lu C, Lin L, Tan H, **Wu H**, Sherman SL , Gao F, Jin P, Chen D (2012) *Fragile X premutation RNA is sufficient to cause primary ovarian insufficiency in mice*. **Human Molecular Genetics**, 21(23):5039-47. doi: 10.1093/hmg/dds348.
38. Wang T, Pan Q, Lin L, Szulwach KE, Song C, He C **Wu H**, Warren ST, Jin P, Duan R, Li X (2012) *Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum*. **Human Molecular Genetics**, 21(26):5500-10. doi: 10.1093/hmg/dds394.
39. Song C, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, Street C, Li Y, Poidevin, M, **Wu H**, Gao J, Liu P, Li L, Xu GL, Jin P, He C (2013) *Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming*. **Cell**. 153(3):678-91. doi: 10.1016/j.cell.2013.04.001.
40. Kosters A, Sun D, **Wu H**, Tian F, Felix JC, Li W, Karpen SJ (2013) *Sexually Dimorphic Genome-Wide Binding of Retinoid X Receptor alpha (RXR $\alpha$ ) Determines Male-Female Differences in the Expression of Hepatic Lipid Processing Genes in Mice*. **PLoS One**. 8(8):e71538. doi: 10.1371/journal.pone.0071538.
41. Bai B, Hales CM, Chen PC, Gozal Y, Dammer EB, Fritz JJ, Wang X, Xia Q, Duong DM, Street RC, Cantero G, Cheng D, Jones DR, Wu Z, Li Y, Diner I, Heilman CJ, Rees HD, **Wu H**, Lin L, Szulwach KE, Gearing M, Mufson EJ, Bennett DA, Montine TJ, Seyfried NT, Wingo TS, Sun YE, Jin P, Hanfelt J, Willcock DM, Levey A, Lah JJ, Peng J (2013) *U1 Small Nuclear Ribonucleoprotein Complex and RNA Splicing Alterations in Alzheimer's Disease*. **Proc Natl Acad Sci**. 8;110(41):16562-7. doi: 10.1073/pnas.1310249110.

42. Singh AM, Chappell J, Trost R, Lin L, Wang T, Tang J, **Wu H**, Zhao S, Jin P, Dalton S (2013), *Cell-Cycle Control of Developmentally Regulated Transcription Factors Accounts for Heterogeneity in Human Pluripotent Cells*. **Stem Cell Reports**. **1**(6): 532–544.
43. Yao B, Lin L, Street RC, Zalewski ZA, Galloway JN, **Wu H**, Nelson DL, Jin P (2014) *Genome-wide alteration of 5-hydroxymethylcytosine in a mouse model of fragile X-associated tremor/ataxia syndrome*. **Human Molecular Genetics**. **23**(4):1095-107. doi: 10.1093/hmg/ddt504.
44. Potts LF, **Wu H**, Singh A, Marcilla I, Luquin MR, Papa SM (2014) *Modeling Parkinson's disease in monkeys for translational studies, a critical analysis*. **Experimental neurology**. **256**, 133-143.
45. Sun X, Fu X, Li J, Xing C, Frierson HF, **Wu H**, Ding X, Ju T, Cummings RD, Dong JT (2014) *Deletion of atbf1/zfhx3 in mouse prostate causes neoplastic lesions, likely by attenuation of membrane and secretory proteins and multiple signaling pathways*. **Neoplasia**. **16**(5):377-89. doi: 10.1016/j.neo.2014.05.001.
46. Lapp S, Mok S, Lei Z, **Wu H**, Preiser P, Bozdech ZB, Galinski M (2015) *Plasmodium knowlesi gene expression differs in ex vivo compared to in vitro blood-stage cultures*. **Malaria Journal**. **14** (1), 110.
47. Halliley JL, Tipton C, Liesved J, Rosenber AF, Darce J, Gregoret IV, Papova L, Kaminiski D, Fucile C, Albizua-Santin I, Kyu S, Chiang KY, Bradley K, Burack R, Slifka M, Hammarlund E, **Wu H**, Zhao L, Walsh EE, Flasey AR, Randall TD Cheun WC, Sanz I, Lee FE (2015) *Identification of Long-lived Plasma Cells in Human Bone Marrow*. **Immunity**. doi:10.1016/j.immuni.2015.06.016.
48. Archer KJ, Dobbin K, Biswas S, Day RS, Wheeler DC, **Wu H** (2015) Computer Simulation, Bioinformatics, and Statistical Analysis of Cancer Data and Processes. **Cancer Informatics** **14**(Suppl 2):247-51.
49. Fu Y, Lv P, Yan G, Fan H, Cheng L, Zhang F, Dang Y, **Wu H**, Wen B. (2015) *MacroH2A1 associates with nuclear lamina and maintains chromatin architecture in mouse liver cells*. **Scientific Report** **5**:17186. doi:10.1038/srep17186.
50. Zeng Y, Yao B, Shin J, Lin L, Kim N, Song Q, Liu S, Su Y, Guo JU, Huang L, Wan J, **Wu H**, Qian J, Cheng X, Zhu H, Ming GL, Jin P, Song H. (2015) *Lin28A Binds Active Promoters and Recruits Tet1 to Regulate Gene Expression*. **Molecular Cell** **61** (1), 153-160 (cover story). doi: <http://dx.doi.org/10.1016/j.molcel.2015.11.020>.
51. Qian X, Nguyen HN, Song MM, Hadiono C, Ogden SC, Hammack C, Yao B, Hamersky G, Jacob F, Zhong C, Yoon K, Jeang W, Li L, Li Y, Thakor J, Berg D, Zhang C, Kang E, Chickering M, Naeun D, Wen Z, Christian KM, Ho CY, Shi PY, Maher B, **Wu H**, Jin P, Tang H, Song H and Ming GL. (2016) *Brain Region-specific Organoids using a Miniaturized Spinning Bioreactor and Modeling of Zika Virus Exposure*. **Cell** **165** (5), 1238-1254.
52. 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 Z, 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 Research**. doi: 10.1093/nar/gkw765.
53. Rao A, Kosters A, Mells JE, Zhang W, Setchell KDR, Amanso AM, Xu T, Wynn GM, Keller BT, Yin H, Banton S, Jones DP, **Wu H**, Dawson PA, Karpen SJ. (2016) *Inhibition of*

*Ileal Bile Acid Uptake Protects Against Non-alcoholic Fatty Liver Disease in High Fat Diet-fed Mice.* **Science Translational Medicine.** Vol. 8, Issue 357, pp. 357ra122. doi: 10.1126/scitranslmed.aaf4823.

54. McDonald OG, Li X, Saunders T, Tryggvadottir R, Mentch SJ, Warmoes MO, Word AE, Carrer A, Salz TH, Natsume S, Stauffer KM, Makohon-Moore A, Zhong QY, **Wu H**, Wellen KE, Locasale JW, Iacobuzio-Donahue C, Feinberg AP. (2017) *Large-scale epigenetic reprogramming is punctuated late during the evolution of pancreatic cancer progression.* Accepted. **Nature Genetics.**

## **Book Chapters**

1. **Wu H**, Kerr K and Churchill GA(2002), *MAANOVA: A Software Package for the Analysis of Spotted cDNA Microarray Experiments*, **The analysis of gene expression data: methods and software**, Springer.
2. **Wu H**, Ji H (2011) *JAMIE: A Software Tool for Jointly Analyzing Multiple ChIP-chip Experiments.* **Next Generation Microarray Bioinformatics, Methods Mol Biol.** Humana Press, Springer.
3. Wu Z , **Wu H** (2016) *Experimental Design and Power Calculation for RNA-seq Experiments.* **Methods in Statistical Genomics, Methods Mol Biol.** Humana Press, Springer

## **Thesis**

1. **Wu H** (2000) *Enhancing Electricity Auction Mechanism with FACTS devices*, Master's thesis, Iowa State University.
2. **Wu H** (2010) *Three Novel Statistical Applications in Genomics: Redefining CpG Island, Peak Detection from Multiple ChIP-chip Experiments, and Data Pre-Processing for ABI/SOLiD Second Generation Sequencing Technology*, Ph.D. thesis, Johns Hopkins University.

## **Talks and lectures** (Since July 2010)

### **Talks at professional conferences or workshops**

1. Modeling intensity data from ABI SOLiD sequencing. *ENAR meeting, Miami, FL.* March 22, 2011.
2. Modeling intensity data from ABI SOLiD sequencing. *ICSA meeting, New York, NY.* June 28, 2011.
3. Differential expression in RNA-seq data. *Workshop on Biostatistics and Bioinformatics. Department of Statistics, Georgia State University, Atlanta, GA.* May 5th, 2012.
4. A new shrinkage method for detecting differential expression in RNA-seq data. *ICSA meeting, Boston, MA.* June 24th, 2012.
5. Evaluating the correlations of protein binding sites. *Joint Statistical Meeting, San Diego, CA.* August 2, 2012.
6. Dispersion shrinkage methods for differential analysis in secondgeneration sequencing data. *BIT's 4th World DNA and Genome Day, Nanjing, China.* April 26, 2013.

7. A Bayesian Hierarchical Model to Detect Differentially Methylated Loci from Single Nucleotide Resolution Sequencing Data. *ICSA and KISS Joint Applied Statistics Symposium, Portland, OR.* June 16, 2014.
8. A Bayesian Hierarchical Model to Detect Differentially Methylated Loci from Single Nucleotide Resolution Sequencing Data. *Third Joint Biostatistics symposium, Chengdu, China.* June 28, 2014.
9. Differential methylation analysis from whole-genome bisulfite sequencing: a matter of spatial correlation, coverage depth and biological variance, *SAMSI epigenetics workshop, Duke University, Durham, NC.* March 10, 2015.
10. Differential methylation analysis for BS-seq data, *Cancer Genetics & Epigenetics Program Retreat 2015, Winship Cancer Institute, Emory University, Atlanta, GA.* May 21, 2015.
11. Differential methylation analysis for BS-seq data, *2015 Tsinghua Summer Workshop on Modern Statistics, Tsinghua University, Beijing, China.* June 26, 2015.
12. Differential methylation analysis for BS-seq data, *Bioinformatics workshop, Tongji University, Shanghai, China.* July 3, 2015.
13. Differential methylation analysis for BS-seq data under general experimental design, *Nankai statistics forum, Nankai University, Tianjin, China.* July 7, 2015.
14. Base-resolution methylation patterns accurately predict transcription factor bindings in vivo, *Joint Statistical Meeting, Seattle, WA.* August 9, 2015.
15. Estimating and accounting for tumor purity in DNA methylation data analysis, *The 4th Institute of Mathematical Statistics Asia Pacific Rim Meeting, Hongkong.* June 29, 2016.
16. Estimating and accounting for tumor purity in DNA methylation data analysis, *Joint Statistical Meeting, Chicago, IL.* August 3, 2016.
17. Redefining differential expression with single cell RNA-seq, *Dahshu 2017: Data Science & Computational Precision Health, San Francisco, CA.* February 21, 2017.

#### **Invited seminars at other academic institutes**

18. Genomic bump finding. *Center for Bioinformatics and Computational Genomics, School of Biology, Georgia Institute of Technology, Atlanta, GA.* December 12, 2010.
19. Intensity pre-processing and quality assessment for ABI/SOLiD second generation sequencing data. *Computational Systems Biology Laboratory, Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA.* February 17, 2011.
20. Genomic bump finding. *Department of Statistics, University of Georgia, Athens, GA.* March 3, 2011.
21. Genomic bump finding. *Department of Statistics, Georgia State University, Atlanta, GA.* March 10, 2011.
22. Modeling intensity data from ABI SOLiD sequencing. *Department of Statistics, Pennsylvania State University, State college, PA.* October 27, 2011.
23. Differential expression in RNA-seq data. *The Jackson Laboratory, Bar Harbor, ME.* March 2nd, 2012.
24. Differential expression in RNA-seq data. *Department of Biostatistics, University of Kentucky, Lexington, KY.* April 13th, 2012.

25. On the detection of differentially methylated loci. *Department of Bioinformatics, Tongji University, Shanghai, China.* April 22, 2013.
26. A novel statistical method for quantitative comparison of multiple ChIP-seq datasets. *Department of Biostatistics, Brown University, Providence, RI.* October 21, 2013.
27. Differential methylation analysis from whole-genome bisulfite sequencing: a matter of spatial correlation, coverage depth and biological variance, *Institute of Bioinformatics, University of Georgia, Athens, GA.* November 7, 2014.
28. Base-resolution methylation patterns accurately predict transcription factor bindings in vivo, *Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA.* February 25, 2015.
29. Differential methylation analysis for BS-seq data, *Department of Statistics, The Chinese University of Hong Kong, Hong Kong.* June 23, 2015.
30. Base-resolution methylation patterns accurately predict transcription factor bindings *in vivo*, *Shenzhen Institutes of Advanced Technology, Chinese Academy of Science, Shen Zhen, China.* June 23, 2015.
31. Differential methylation analysis for BS-seq data, *Department of Mathematics, Hong Kong Baptist University, Hong Kong.* June 24, 2015.
32. Differential methylation analysis for BS-seq data, *Department of Mathematics, Shanghai Normal University, Shanghai, China.* June 30, 2015.
33. Differential methylation analysis for BS-seq data, *Department of Biostatistics and Epidemiology, Georgia Regents University, Augusta, GA.* October 20, 2015.
34. Base-resolution methylation patterns accurately predict transcription factor bindings in vivo, *Department of Mathematics, Shanghai Normal University, Shanghai, China.* June 1, 2016.
35. Base-resolution methylation patterns accurately predict transcription factor bindings in vivo, *Department of Mathematics, Tianjin University, Tianjin, China.* Oct 13, 2016.

### Invited lectures and short courses

36. Analysis of ChIP-seq Data. **Third Annual Short Course on Statistical Genetics and Genomics**, *University of Alabama at Birmingham, Birmingham, AL.* July 11, 2013.
37. Introduction to the analysis of ChIP-seq Data. **Third Annual Short Course on Next-Generation Sequencing: Technology and Statistical Methods**, *University of Alabama at Birmingham, Birmingham, AL.* December 18, 2013.
38. Introduction to next-generation sequencing data analysis: RNA-seq and ChIP-seq, **7th Annual Bayesian Biostatistics and Bioinformatics Conference**, *The University of Texas MD Anderson Cancer Center, Houston, TX.* February 12, 2014.
39. Statistical methods for genome wide regional analysis with next generation sequencing data. **ENAR 2014 spring meeting.** *Baltimore, MD.* March 16, 2014.
40. Introduction to R and Bioconductor. Fudan University, School of Medicine. *Shanghai, China.* June 20-26, 2014.
41. Introduction to statistical epigenomics. **Fourth Annual Short Course on Statistical Genetics and Genomics**, *University of Alabama at Birmingham, Birmingham, AL.* July 10, 2014.



42. Introduction to the analysis of ChIP-seq Data. **Fourth Annual Short Course on Next-Generation Sequencing: Technology and Statistical Methods**, *University of Alabama at Birmingham, Birmingham, AL*. December 17, 2014.
43. Introduction to statistical epigenomics. **Fifth Annual Short Course on Statistical Genetics and Genomics**, *University of Alabama at Birmingham, Birmingham, AL*. July 30, 2015.
44. Introduction to the analysis of ChIP-seq Data. **Fifth Annual Short Course on Next-Generation Sequencing: Technology and Statistical Methods**, *University of Alabama at Birmingham, Birmingham, AL*. December 17, 2015.
45. Introduction to statistical epigenomics. **Sixth Annual Short Course on Statistical Genetics and Genomics**, *University of Alabama at Birmingham, Birmingham, AL*. July 13, 2016.
46. Introduction to the analysis of ChIP-seq Data. **Sixth Annual Short Course on Next-Generation Sequencing: Technology and Statistical Methods**, *University of Alabama at Birmingham, Birmingham, AL*. December 14, 2016.
47. Introduction to single-cell sequencing. **Sixth Annual Short Course on Next-Generation Sequencing: Technology and Statistical Methods**, *University of Alabama at Birmingham, Birmingham, AL*. December 15, 2016.

## Research Funding

### Curent external funding

1R01GM122083-01 (PI: Wu) Statistical Methods for Single-Cell RNA-Seq Role: PI	08/15/16 – 07/31/21
2 P50 AG025688-11 (PI: Levey) Alzheimer's Disease Research Centers Role: Biostatistician	04/01/05 – 04/30/20
5 UL1 TR000454-09 (PI: Stephens) Atlanta Clinical and Translational Science Institute (ACTSI) Role: Biostatistician	09/17/07 – 05/31/17
5 P30 AI050409-17 (PI: Del Rio) Emory/Atlanta Center for AIDS Research (CFAR) Role: Co-Investigator	08/01/12 – 07/31/17
R01 DK056239 (PI: Karpen) Molecular Regulation of Hepatic Transporters Role: Co-Investigator	09/01/12 – 08/31/18
R01 AI111948-01 (PI: Day) NK Cell-mediated Regulation of T Cell Immunity in TB/HIV Co-infection Role: Co-Investigator	04/01/14 – 03/31/19

U19 AI110483-01 (PI: Sanz) 05/01/14 – 04/30/19  
B cell Autoimmunity in Human SLE  
Role: Co-Investigator

U19 AI111211-02 (PI: Blumberg) 08/11/14 – 07/31/21  
Role of Antigen Specific T Cell Responses in the Control of TB  
Role: Co-Investigator

5 R01 MH102690-03 (PI: Jin) 04/01/14–02/28/18  
Tet-mediated Epigenetic Modulation in Autism  
Role: Biostatistician

1 R01 HL130471-01 (PI: Lewis) 01/15/16–12/31/19  
Expectations of Discrimination and CVD risk in African-American Women  
Role: Co-Investigator

### **Past external funding**

2 R21 HD073162 (PI: Jin) 08/01/10 – 07/31/14  
Mapping 5-Hydroxymethylcytosine at Single-Base Resolution in Human ES Cells  
Role: Co-Investigator

VA ORD (PI: Brandes) 01/01/14 – 12/31/15  
VA IPA for Hao Wu: HDAC inhibition for the prevention of Lung Cancer  
Role: PI

Boeing RC187-S1 (PI: Weiss, Hertzberg) 07/30/13 – 07/30/15  
Assessing the Rate of Infectious Disease Transmission in an Aircraft Cabin  
Role: Co-Investigator

Simons Foundation 239320 (PI: Jin) 07/01/13 – 06/30/15  
5-Hydroxymethylcytosine-Mediated Epigenetic Regulation in Autism  
Role: Co-Investigator

### **Emory funding**

Emory University Research Committee (URC) Grant (PI: Wu) 5/1/2012 – 4/30/2013  
Statistical methods for data integration in genetic association studies  
Role: PI

Emory University Alzheimer Disease Research Center (ADRC) pilot Grant (mPI: Wu & Jin) 5/1/2016 – 4/30/2017  
Methylation of Cell-free DNA in Alzheimer's Disease  
Role: PI

## Teaching at Emory University

- Spring 2013–2017: BIOS 545 *Introduction to R programming*, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University.
- Fall 2012, 2014, 2016: BIOS 731 *Advanced Statistical Computing*, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University.
- Fall 2011–2015: BIOS 555 *High-throughput data analysis using R and bioConductor*, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University.
- Fall 2015: BIOS 516 *Introduction to Large-Scale Biomedical Data Analysis*, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University.
- Spring 2012, BIOS 590R *Seminar in Biostatistics*, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University.
- Spring 2012, BIOS 790R *Advanced Seminar in Biostatistics*, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University.

## Advising at Emory University

### Thesis advisees

- Hao Feng, Ph.D. candidate, Department of Biostatistics and Bioinformatics.
- Ben Li, Ph.D. candidate, Department of Biostatistics and Bioinformatics. Co-advise with Zhaohui Qin.
- Li Chen, Ph.D. candidate, Department of Computer Science. Co-advise with Zhaohui Qin.
- Tianlei Xu, Ph.D. candidate, Department of Computer Science. Co-advise with Zhaohui Qin.
- Haonan Feng, MSPH 2017, Department of Biostatistics and Bioinformatics.
- Hao Feng, MSPH 2012, Department of Biostatistics and Bioinformatics. Co-advise with Karen Conneely. One of the six finalist of the *Charles C. Shepard Awards* for the best master's thesis of the Rollins School of Public Health. Currently a Ph.D. candidate.

### Academic advisees

- Yingtian Hu, Ph.D. student, Department of Biostatistics and Bioinformatics.
- Ran Shi, Ph.D. candidate, Department of Biostatistics and Bioinformatics.
- Yunxuan Jiang, Ph.D. candidate, Department of Biostatistics and Bioinformatics.
- Jay Soh, Ph.D. candidate, Department of Biostatistics and Bioinformatics.
- Hao Feng, Ph.D. candidate, Department of Biostatistics and Bioinformatics.
- Ruizhe Wu, MSPH, Department of Biostatistics and Bioinformatics.
- Chenchen Yu, MSPH, Department of Biostatistics and Bioinformatics.
- Shiwei Tang, MPH, Department of Biostatistics and Bioinformatics.
- Ganzhong Tian, MSPH, Department of Biostatistics and Bioinformatics.

### Practicum advisees

- Haonan Feng, MSPH candidate, Department of Biostatistics and Bioinformatics.
- Yufeng Huang, MSPH 2015, Department of Epidemiology.

### Thesis committees

- Crystal Grant, Ph.D. candidate, Genetics and Molecular Biology Graduate Program, Emory University.
- Robert Petit, Ph.D. candidate, Population Biology, Ecology and Evolution Graduate Program, Emory University.
- Jiani Hu, MSPH 2016, Department of Biostatistics and Bioinformatics.
- Jessica Chen, MSPH 2016, Department of Biostatistics and Bioinformatics.
- Tao Wang, Ph.D. 2014, Department of Human Genetics, Emory University.

## Service at Emory University

- 2011–present, Affiliated faculty, Graduate Program in Population Biology, Ecology, & Evolution (PBEE).
- 2011–2012, BIOS seminar chair. Department of Biostatistics and Bioinformatics.
- 2011, 2013, Faculty Search Committee, Department of Biostatistics and Bioinformatics.
- 2011–2014, Ph.D. admissions Committee, Department of Biostatistics and Bioinformatics.
- 2014–present, Master admissions Committee, Department of Biostatistics and Bioinformatics.
- 2011–present, Curriculum Committee, Department of Biostatistics and Bioinformatics.
- 2013–present, BIOS web site committee, Department of Biostatistics and Bioinformatics.
- 2011–present, High performance computing committee, Emory University.
- 2012–2014, Woodruff Natural Science Fellowship selection committee, Emory University.

## Other professional Activities and Services

### Editorial activities

- Associate editor: *Journal of Applied Statistics*.
- Review editor: *Frontiers in Epigenomics*.
- Guest editor: a supplement issue of *Cancer Informatics*.
- Article reviewer: *Annals of Applied Statistics*, *Bayesian Analysis*, *Bioinformatics*, *Biometrics*, *Biostatistics*, *Biosensors*, *BMC Bioinformatics*, *BMC Genomics*, *BMC Research Notes*, *Briefings in Bioinformatics*, *Cancer Research*, *CSH Protocols*, *G3: Genes/Genomes/Genetics*, *Genetic Epidemiology*, *Genome Biology*, *Genome Research*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *Journal of Genetics and Genomics*, *Plant Cell*, *PLoS One*, *PLoS Computational Biology*, *Nature Genetics*, *Nature Protocols*, *Nucleic Acids Research*, *Scientific Report*, *Statistical Analysis and Data Mining*, *Statistics and Its Interface*, *Statistical Application in Genetics and Molecular Biology*, *Statistics in Medicine*.

## Conference Program

- Organizer and Chair: invited session “Statistical Modeling and data mining of High-throughput Genomics Data”. *Third Joint Biostatistics symposium, Chengdu, China*. June 28, 2014.
- Chair: invited session “Modeling of long range interaction II”. *SAMSI epigenetics workshop*, Duke University, Durham, NC. March 10, 2015.

## Review Services

- Member: NSF DMS/NIGMS grant review panel. December 2014, 2016.

## Software developed

- **DSS**: Dispersion shrinkage for Sequencing. An R/Bioconductor package for differential analysis from high-throughput sequencing data, including differential expression for RNA-seq and differential methylation for bisulfite-sequencing data. Available at <http://www.bioconductor.org/packages/release/bioc/html/DSS.html>. Total downloads over the last 12 months (according to Bioconductor download statistics): **5095**.
- **Rmaanova**: An R package for analyzing gene expression microarray data using mixed effect model. <http://www.bioconductor.org/packages/release/bioc/html/maanova.html>. Total downloads over the last 12 months (according to Bioconductor download statistics): **2941**.
- **PROPER**: PROspective Power Evaluation for RNA-seq. An R/Bioconductor package for power assessment and sample size determination for RNA-seq. Available at <http://www.bioconductor.org/packages/devel/bioc/html/PROPER.html>. Total downloads over the last 12 months (according to Bioconductor download statistics): **2042**.
- **ChIPComp**: Quantitative comparison of multiple ChIP-seq datasets. An R/Bioconductor package for differential protein binding or histone modification analyses from ChIP-seq. Available at <http://bioconductor.org/packages/devel/bioc/html/ChIPComp.html>. Total downloads over the last 12 months (according to Bioconductor download statistics): **2050**.
- **RSOLiD**: An R package for intensity pre-processing of ABI/SOLiD second generation sequencing data. <http://rafalab.jhsph.edu/Rsolid/>.
- **makeCGI**: An R package for defining CpG islands from DNA sequences using Hidden Markov Models. <http://web1.sph.emory.edu/users/hwu30/software/makeCGI/>.
- **JAMIE**: An R package for jointly analyzing multiple ChIP-chip data sets and detecting transcription factor binding sites. <http://www.biostat.jhsph.edu/~hji/jamie/>.
- **J/qlt**: A Java graphical user interface for R/qlt, a popular R software for analyzing QTL data. <http://churchill.jax.org/software/jqlt.shtml>.

## Honors and Awards

- Member of the Phi Beta Kappa Society, 2011
- 2010, Margaret Merrell Award for excellence in research, Johns Hopkins University

## Professional membership

- American Statistical Association (ASA)
- Institute of Mathematical Statistics (IMS)
- International Biometric Society (ENAR)
- International Society for Computational Biology (ISCB)
- International Chinese Statistical Association (ICSA)