

## Curriculum Vitae

### Zhijin Wu, Ph.D.

Associate Professor  
Department of Biostatistics  
Brown University

### ADDRESS

Department of Biostatistics  
Brown University, Box G-S121-7  
Providence, RI 02912

Tel. 401 863 1230

Email: [zhijin\\_wu@brown.edu](mailto:zhijin_wu@brown.edu)

Fax. 401 863 9182

Web: <http://www.stat.brown.edu/zwu>

### EDUCATION

1997 B.S. Biochemistry and Molecular Biology Peking University  
2000 M.S. Molecular Biology University of Southern California  
2005 Ph.D. Biostatistics Johns Hopkins University  
Dissertation: *Probe Level Models for DNA Microarrays*  
Advisor: Rafael A. Irizarry

### ACADEMIC APPOINTMENTS

July, 2012 – present **Associate Professor**, Department of Biostatistics (Tenured),  
Brown University  
July, 2005 – June, 2012 **Assistant Professor**, Department of Biostatistics (Tenure  
Track), Brown University

### OTHER ACADEMIC AND PROFESSIONAL POSITIONS

July, 2012 – present Director, PhD Program in Biostatistics,  
Department of Biostatistics, Brown University  
Dec, 2011 – June, 2012 Director, Masters Program in Biostatistics,  
Department of Biostatistics, Brown University

### ACADEMIC HONORS & AWARDS

2011 CAREER Award, National Science Foundation  
2007 ADVANCE Career Development Award  
Brown University  
2007 CCMB Scholarship Innovator Award  
Center for Computational Molecular Biology  
Brown University  
2005 Margaret Merrell Award for outstanding research  
Department of Biostatistics  
Johns Hopkins Bloomberg School of Public Health  
2005 Third Place, Delta Omega Poster Competition  
Johns Hopkins Bloomberg School of Public Health

## MEMBERSHIP IN PROFESSIONAL SOCIETIES

American Statistical Association  
International Biometric Society (ENAR)  
International Chinese Statistics Association

## PUBLICATIONS (students marked with \* and corresponding author indicated with †)

### A. Articles in peer-reviewed journals

1. Viswanath V, **Wu Z**, Fonck C, Wei Q, Boonplueang R and Andersen JK<sup>†</sup>. (2000) Transgenic mice neuronally expressing baculoviral p35 are resistant to diverse types of induced apoptosis, including seizure – associated neurodegeneration. *Proc Natl Acad Sci U. S. A.* 97(5):2270 –5.
2. Peng J, **Wu Z**, Wu Y, Hsu M, Stevenson FF, Boonplueang R, Roffler–Tarlov SK and Andersen JK<sup>†</sup> (2002). Inhibition of caspases protects cerebellar granule cells of the weaver mouse from apoptosis and improves behavioral phenotype. *J Biol Chem* 277(46):44285–91.
3. Wu, Z<sup>†</sup>, **Wu, Z**, Wiegandt, DA and Nassar, CR (2003) High – performance 64 – QAM OFDM via carrier interferometry spreading codes. *Vehicular Technology Conference, 2003 IEEE 58th*
4. Irizarry RA<sup>†</sup>, Ooi SL, **Wu Z** and Boeke J (2003) Use of Mixture Models in a Microarray – Based Screening Procedure for Detecting Differentially Represented Yeast Mutants. *Statistical Applications in Genetics and Molecular Biology* 2(1).
5. Cope LM, Irizarry RA, Jaffee H, **Wu Z** and Speed TP<sup>†</sup> (2003). A Benchmark for Affymetrix GeneChip Expression Measures. *Bioinformatics* 20: 323–331.
6. **Wu Z**, Irizarry RA<sup>†</sup>, Gentleman R, Martinez Murillo F and Spencer F (2004) A Model Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association*, 99(468):909–917.
7. **Wu Z** and Irizarry RA<sup>†</sup> (2004). Preprocessing of oligonucleotide array data. *Nature Biotechnology* 22: 656–658. PMID:15175677
8. **Wu Z** and Irizarry RA<sup>†</sup> (2005). Stochastic Models Inspired by Hybridization Theory for Short Oligonucleotide Arrays. *Journal of Computational Biology*, 12(6) 882–893.
9. Irizarry RA<sup>†</sup>, Cope L, **Wu Z** (2006). Feature – level exploration of the Choe et al. Affymetrix GeneChip control dataset. *Genome Biology*, 2006, 7:404
10. Irizarry RA<sup>†</sup>, **Wu Z** and Jaffee H (2006). Comparison of Affymetrix GeneChip Expression Measures. *Bioinformatics* 22(7):789–794
11. **Wu Z**<sup>†</sup> and Irizarry RA (2007). A Statistical Frame Work for the Analysis of Microarray Probe – level Data. *Annals of Applied Statistics*. 1(2) 333–357
12. Sui Y\* and **Wu Z**<sup>†</sup> (2007). Alternative Statistical Parameter for High – Throughput Screening Assay Quality Assessment, *Journal of Biomolecular Screening* 12(2):229–234.

13. Yegnasubramanian S, Haffner MC, Zhang Y, Gurel B, Cornish TC, **Wu Z**, Irizarry RA, Morgan J, Hicks J, DeWeese TL, Isaacs WB, Bova GS, De Marzo AM, Nelson WG (2008). DNA hypomethylation arises later in prostate cancer progression than CpG island hypermethylation and contributes to metastatic tumor heterogeneity. *Cancer Research*. 68(21):8954–67. PMID:[18974140](#) PMCID: [PMC2577392](#)
14. Santiago Jr RB, **Wu Z**, Zhang L, and Widmer G (2008). Identification of Growth Inhibiting Compounds in A Giardia Lamblia High-Throughput Screen. *Molecular & Biochemical Parasitology*. 162(2):149–54
15. Miselis N, **Wu Z**, Rooijen NV and Agnes KB (2008). Targeting tumor-associated macrophages in an orthotopic murine model of diffuse malignant mesothelioma. *Molecular Cancer Therapeutics* 7(4):788–799.
16. **Wu Z**<sup>†</sup>, Liu D and Sui Y\* (2008). Quantitative Assessment of Hit Detection and Confirmation in Single and Duplicate HTS Screenings. *Journal of Biomolecular Screening*. 13(2):159–167.
17. **Wu Z**<sup>†</sup> A review of statistical methods for preprocessing oligonucleotide microarrays (2009). *Statistical Methods in Medical Research*, 18 (6), 533–541
18. Paruthiyil S, Cvorov A, Zhao X, **Wu Z**, Sui Y and others (2009). Drug and Cell Type-Specific Regulation of Genes with Different Classes of Estrogen Receptor  $\beta$ -Selective Agonists. *Public Library of Science One*. 4(7): e6271.
19. Sui Y\*, Zhao X, Speed TP and **Wu Z**<sup>†</sup>. Background Adjustment for DNA Microarrays Using a Database of Microarray Experiments (2009). *Journal of Computational Biology*. 16(11): 1501–1515
20. Irizarry RA, Ladd-Acosta C, Wen B, **Wu Z**, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics*. 41(2):178–86. PMID: 19151715
21. **Wu Z**, Li X, Husnay R, Chakravarthy V, Wang B and Wu Z (2009). A novel highly accurate log skew normal approximation method to lognormal sum distributions. *WCNC 2009. IEEE* 1–6
22. **Wu Z**<sup>†</sup>, Jenkins BD, Rynears TA and others (2010). Empirical Bayes Analysis of Sequencing-based Transcriptional Profiling without Replicates. *BMC Bioinformatics*. 11:564
23. Miselis N, Lau BW, **Wu Z**, Kane AB (2010). Kinetics of Host Cell Recruitment During Dissemination of Diffuse Malignant Peritoneal Mesothelioma *Cancer Microenvironment* 4(1): 39–50 PMID:21505561
24. **Wu Z**<sup>†</sup> and Aryee M. Subset Quantile Normalization using Negative Control Features (2010). *Journal of Computational Biology*, 17(10):1385-1395
25. Champion SN, Houseman EA, Sandrof MA, Hensley JB, Sui Y, Gaido KW, **Wu Z** and Boekelheide K (2010). Suppression of Radiation-induced Testicular Germ Cell Apoptosis by 2,5-Hexanedione Pretreatment. II. Gene Array Analysis Reveals Adaptive Changes in Cell Cycle and Cell Death Pathways. *Toxicological Sciences* 117(2):457–465 PMID:20616210

26. Ghosh M, Fahey JV, Shen Z, Lahey T, Cu-Uvin S, **Wu Z**, Mayer K, Wright PF, Kappes JC, Ochsenauber C and Wira CR\* (2010). Anti-HIV Activity in Cervical-Vaginal Secretions from HIV-Positive and -Negative Women Correlate with Innate Antimicrobial Levels and IgG Antibodies. *Public Library of Science One*. 5(6):e11366. PMID:20614007
27. Hansen KD, **Wu Z**, Irizarry RA and Leek JT (2011). Sequencing technology does not eliminate biological variability. *Nature Biotechnology* 29:572–573 PMID:21747377
28. Yegnasubramanian S, **Wu Z**, Haffner MC, Esopi D, Aryee MJ, Badrinath R and others (2011). Chromosome-wide mapping of DNA methylation patterns in normal and malignant prostate cells reveals pervasive methylation of gene-associated and conserved intergenic sequences. *BMC Genomics* 12(1):313–332 PMCID: PMC3124442
29. Aryee, M, **Wu Z**, Ladd-Acosta C, Herb B, Feinberg Ap, Yegnasubramanian S and Irizarry R (2011). Accurate genome-scale percentage DNA methylation estimates from microarray data. *Biostatistics* 12(2):197–210 PMID:20858772
30. Faghiri Z, Bonilla Santiago R, **Wu Z**, and Widmer G. High-Throughput Screening In Sub-Optimal Growth Conditions Identifies Agonists Of Giardia Lamblia Proliferation (2011). *Parasitology* 138(2):194–200
31. Gong J\*, Campos H, McGarvey S, **Wu Z**, Goldberg R and Baylin A (2011). Adipose tissue palmitoleic acid and obesity in humans: does it behave as a lipokine? *The American Journal of Clinical Nutrition* 93(1):186–191 PMID:22049297
32. Gong J\*, Campos H, McGarvey S, **Wu Z**, Goldberg R and Baylin A (2011). Genetic variation in stearoyl-CoA desaturase 1 is associated with metabolic syndrome prevalence in Costa Rican adults. *The Journal of nutrition*:141(12) 2211-2218. PMID: 22049297
33. DeLong AK, Wu M, Bennett D, Parkin N, **Wu Z**, Hogan JW and Kantor R (2011). SQUAT (Sequence Quality Analysis Tool) for protease and reverse transcriptase. *AIDS Research and Human Retroviruses* 27 PMID:21916749
34. Li X, **Wu Z**, Chakravarthy V and Wu Z† (2011). A Low Complexity Approximation to Lognormal Sum Distributions via Transformed Log Skew Normal Distribution. *IEEE Transactions on Vehicular Technology*. 60(8):4040-4045
35. Hansen KD, Irizarry RA and **Wu Z**† (2012). Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics*, 13(2):204-216 PMID:22285995
36. Champion SN, Catlin N, Houseman EA, Hensley J, Sui Y, Caido KW, **Wu Z** and Boekelheide K. (2012) Molecular alterations underlying the enhanced disruption of spermatogenesis by 2,5-hexanedione and carbendazim co-exposure. *Reproductive Toxicology* 33 (3): 382-389 PMID:22382377
37. Dyhrman ST, Jenkins BD, Rynearson TA, Saito MA, Mercier ML, Harriet Alexander H, Whitney LP, Drzewianowski A, Bulygin VV, Bertrand EM, **Wu Z**, Benitez-Nelson C, Heithoff A. The Transcriptome and Proteome of the Diatom *Thalassiosira pseudonana* Reveal a Diverse Phosphorus Stress Response. *PLoS ONE*, 7.3 (2012) :e33768 PMID:22479440
38. Wu H, Wang C and **Wu Z**† (2013). A new shrinkage estimator for dispersion improves differential expression detection in RNA-seq data. *Biostatistics* 14 (2), 232-243 PMID:23001152
39. Austin A\*, Linkletter C and **Wu Z** (2013) "Covariate-defined latent space random effects model." *Social Networks*, 35(3):338-346

40. CW Dunn, X Luo, **Z Wu** (2013) Phylogenetic analysis of gene expression. *Integrative and comparative biology* 53 (5), 847-856 PMID:23748631
41. Srivatsan A, Jenkins S, Jeon M, **Wu Z**, Kim C, Chen J and Pandey R (2014) Gold Nanocage-Photosensitizer Conjugates for Dual-Modal Image-Guided Enhanced Photodynamic Therapy. *Theranostics* 2014; 4(2):163-174. doi:10.7150/thno.7064 PMID:25273110
42. Wu H, Chi Wang and **Wu Z**† (2015) PROPER: comprehensive power evaluation for differential expression using RNA-seq. *Bioinformatics* 31 (2): 233-241. PMID:25273110
43. Mossman, J. A., Tross, J. G., Li, N., **Wu, Z.**, & Rand, D. M. (2016). Mitochondrial-Nuclear Interactions Mediate Sex-Specific Transcriptional Profiles in *Drosophila*. *Genetics*, genetics-116.
44. Li, N., McCall, M. N., & **Wu, Z**†. (2016). Establishing Informative Prior for Gene Expression Variance from Public Databases. *Statistics in Biosciences*, 1-18.
45. Mossman, J. A., Tross, J. G., Jourjine, N. A., Li, N., **Wu, Z.**, & Rand, D. M. (2016). Mitonuclear interactions mediate transcriptional responses to hypoxia in *Drosophila*. *Molecular Biology and Evolution*, msw246.
46. Huang Y-T, Zhang Y, Wu Z and Michaud DS (2016) Genotype-based gene signature of glioma risk identified by integrative genome-wide association study. *Neuro-Oncology* (accepted)

## B. Book Chapters

1. Bolstad, B. M., Irizarry, R. A., Gautier, L. and **Wu, Z.** Low – level analysis and Pre – processing of Affymetrix GeneChip data. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer 2005*
2. **Wu Z**† and Wu Z. Exploration, Visualization and Preprocessing of High- Dimensional Data. *Statistical Methods in Molecular Biology* Humana Press, 2010
3. **Wu Z**† and Wu H. Experimental Design and Power Calculation for RNA-seq Experiments. *Statistical Genomics: Methods and Protocols, Methods in Molecular Biology, Springer 2016*

## Invited Lectures

1. Apr 2005. A Model – based Background Adjustment for Oligonucleotide Expression Arrays. National Cancer Institute, Biostatistics Branch, Division of Cancer Epidemiology and Genetics
2. May 2005. Unifying model for probe level microarray data National Cancer Institute, Biometric Research Branch, Division of Cancer Treatment and Diagnosis
3. Feb 2006. A unified model for Microarray Probe Level Data. Wright State University
4. May 2006. A Statistical Framework for the Analysis of Microarray Probe Level Data. University of Delaware
5. June 2006. A unified model for Microarray Probe Level Data. The 4th International Bioinformatics Workshop. Changsha, China
6. June 2006. Preprocessing of Microarray Data: Computational and Statistical Aspects of Microarray Analysis. Bioconductor workshop. Brixen, Italy.

7. May 2007. A Statistical Framework for the Analysis of Microarray Probe Level Data. National Institute of Standards and Technology
8. Apr 2008. Improving the Detection of Gene Expression in DNA Microarrays. University of California, Berkeley
9. Apr 2008. Background estimation using database of DNA Microarrays. Yale University
10. March 2010. Normalization using negative controls. ENAR Annual meetings
11. Dec 2010. Empirical Bayes Analysis of Sequencing-based Transcriptional Profiling without Replicates. The 8<sup>th</sup> ICSA conference, Guangzhou, China
12. March 2011. Empirical Bayes Analysis of RNA-seq data without Replicates. University of Connecticut
13. March 2011. Empirical Bayes Analysis of RNA-seq data without Replicates. ENAR Annual meetings, Miami
14. April 2011. A Nonlinear Normalization Method for RNA-seq data, NESS 2011 Connecticut
15. Sep 2011. Removing technical variability in RNA-seq data using conditional quantile normalization. NHGRI Special Conference: Statistical Analyses for Next Generation Sequencing. Birmingham, AL
16. June 2012. Normalizing RNA-seq data using conditional quantile normalization. ICSA Applied Statistical Symposium. Boston, MA
17. Nov 2012. A new shrinkage estimator for dispersion improves differential expression detection in RNA-seq data. University of Texas MD Anderson Cancer Center.
18. Dec 2012. Differential expression in RNA-seq data. Correlated and high dimensional data seminar. Harvard University, Department of Biostatistics
19. June 2013. Simulation based power computation in RNA-seq. ICSA Joint Statistics Conference. Bethesda, MD
20. July 21 2015 Gene Set Analysis for large and sparse effects BioC 2015, Bioconductor Annual Conference, Seattle, WA.
21. Aug 5. 2015 Gene Set Analysis for large and sparse effects, JSM. Seattle, WA.
22. Nov 20. 2015 Gene Set Analysis for large and sparse effects. Boston University.
23. June, 2016 Normalization issues in scRNAseq data. ICSA 2016 Symposium. Atlanta, Georgia.
24. Sept, 2016. Single cell RNA sequencing. Penn State University.

**F. Contributed Papers and Abstracts Presented (abstracts presented by others not listed)**

1. Using Sequence Information to Improve RNA Expression Measures. Joint Statistical Meetings 2003, San Francisco

2. Stochastic models inspired by hybridization theory for short oligonucleotide arrays.  
RECOMB 2004, San Diego

## **RESEARCH GRANTS AND CONTRACTS: ACTIVE**

### **Statistical Methods**

1. **Statistical and Computational Methods for RNA-seq data**  
Role: Principal Investigator  
Agency: NSF
2. **Statistical Methods for Single-Cell RNA-Seq (R01GM122083-01)**  
Role subcontract PI  
Agency NIH/NIGMS

### **Interdisciplinary Collaborations**

3. **COBRE Center for Central Nervous System Function**  
Role: Core director  
PI: Sanes  
Agency: NIGMS
4. **COBRE Center for Computational Biology of Human Disease**  
Role: Core B co-director  
PI: Rand  
Agency: NIGMS

## **RESEARCH GRANTS AND CONTRACTS: COMPLETED**

1. **CFAR Developmental Award: Novel statistical methodology for the design and analysis of High Throughput Screening (HTS) experiments**  
Role: Principal Investigator  
Agency: NIH/The Miriam Hospital Subcontract
2. **Effects of Tissue Heterogeneity on Gene Expression Measures in Tumor Samples**  
Role: Principal Investigator  
Agency: RI Foundation
3. **Gene Networks in Peri – pubertal Sertoli Cell Injury**  
Role: Co – Investigator  
Project PI: Mary Hixon  
Agency: NIH(R01 ES01 5704 – 01)
4. **HCV in Alcoholics**  
Role: Co – Investigator  
Project PI: Jack Wands  
Agency: NIH/Rhode IslandHospital subcontract
5. **High-throughput screening for new inhibitors of Giardia lamblia**  
Role: Co – Investigator

- Project PI: Widmer  
Agency: NIH/Tufts University Subcontract
- 6. Manufactured Nanomaterials: Physico-chemical Principles of Biocompatibility and Toxicity**  
Role: Co – Investigator  
Project PI: Agnes Kane  
Agency: NIEHS
- 7. Preprocessing and Analysis Tools for Contemporary Microarray Applications**  
Role: Co – PI  
Project PI: Rafael Irizarry  
Agency: NIH/Johns Hopkins University Subcontract
- 8. Reuse In RI: A State-Based Approach To Complex Exposures**  
Role: Co-Director of Analytical Core  
Project PI: K. Boekelheide  
Agency: NIH
- 9. COBRE for Perinatal Biology**  
Role: Consultant  
Project PI: James Padbury  
Agency: NIH
- 10. Lifespan/Tufts/Brown Center for AIDS Research (Biostatistics Core)**  
Role: Co – Investigator, core faculty member  
Project PI: Charles C. J. Carpenter  
Agency: NIH/NIAID (P30 AI 42853)
- 11. Nuclear-Mitochondrial Fitness Interactions in Drosophila (multiple PI with David Rand)**  
Role: co-PI  
Agency: NIH/NIGMS

## SERVICE ACTIVITIES

### Professional Service

#### NSF Grant Review Panelist 2009/2011 /2013/2015/2016

**Journal Reviewer** *Nucleic Acid Research, Nature Biotechnology, Journal of Bioinformatics and Computational Biology, Proceedings of the National Academy of Sciences, Bioinformatics, Statistical Applications in Genetics and Molecular Biology, Journal of Computational Biology, Biometrical Journal, Science Translational Medicine, BMC bioinformatics, PLOS*

#### PhD Thesis Directed

**Yunxia Sui**, Ph.D. in Biostatistics 2010

*Robust Gene Expression Measure using Databases of Microarrays*

Current position: Statistical Leader, General Electric Global Research Center

**Andrea Austin**, Ph.D. in Biostatistics 2013

*Latent class social network models and application in genomics*

Current position: Research Scientist at Dartmouth College

#### Postdoctoral advisees

**Nan Li**, postdoctoral fellow



Statistical and computational methods for RNAseq data  
Current position: Research Scientist at MD Anderson Cancer Center

**Thesis Committee**

1. Devin Koestler, PhD Biostatistics, Aug 2011
2. Chui Sun Yap, PhD Molecular, Cellular Biology and Biochemistry, 2008
3. Jian Gong, PhD Epidemiology, 2010
4. Hsin-ta Wu. PhD Computer Science, May 2016
5. Bahar Erar, PhD Biostatistics, Aug 2016