

Michael C. Wu

1100 Fairview Avenue North
M3-C102
Seattle, WA 98109-1024
<http://research.fredhutch.org/wu/>

Phone: 206-667-6603
FAX: 206-667-4408
mcwu@fredhutch.org

PERSONAL INFORMATION

Birthplace: Howard County, Maryland, U.S.A.
Birthdate: July 1982
Citizenship: United States of America

EDUCATION

Ph.D., Biostatistics, June 2009
Minor: Computational Molecular Biology
Harvard University, Cambridge, MA.
Committee: Drs. Xihong Lin (Advisor), Tianxi Cai (Co-Advisor), and Jun Liu

A.M., Biostatistics, June 2006
Harvard School of Public Health, Boston, MA.

B.S., Mathematical and Computational Science, Biological Sciences Track, April 2004
Stanford University, Stanford, CA.

PROFESSIONAL EXPERIENCE

Division of Public Health Sciences, Fred Hutchinson Cancer Center
Seattle, Washington USA

<i>Professor</i>	2021 –
<i>Associate Professor</i>	2016 – 2021
<i>Assistant Professor</i>	2013 – 2016

Department of Biostatistics, University of Washington
Seattle, Washington USA

<i>Affiliate Professor</i>	2022 –
<i>Affiliate Associate Professor</i>	2017 – 2022
<i>Affiliate Assistant Professor</i>	2015 – 2017

SWOG Statistical Center, Fred Hutchinson Cancer Center
Seattle, Washington USA

<i>Faculty Statistician, Melanoma & Translational Medicine</i>	2013 –
--	--------

Fred Hutchinson/University of Washington Cancer Consortium
Seattle, Washington USA

<i>Member</i>	2013 –
---------------	--------

Department of Biostatistics, The University of North Carolina at Chapel Hill
Chapel Hill, North Carolina USA

<i>Assistant Professor (Adjunct)</i>	2014 – 2017
<i>Assistant Professor (On Leave)</i>	2013 – 2014
<i>Assistant Professor (Tenure Track)</i>	2009 – 2013

Carolina Center for Genome Sciences, and Center for Environmental Health and Susceptibility, The University of North Carolina at Chapel Hill
Chapel Hill, North Carolina USA

Member 2009 – 2013

Department of Biostatistics, Harvard University
Boston, Massachusetts USA

Graduate Research Assistant 2006 – 2009

Department of Biostatistics & Computational Biology, Dana Farber Cancer Institute
Boston, Massachusetts USA

Graduate Research Assistant, X. Shirley Liu Lab Spring 2006

Graduate Research Assistant Summer 2005

National Human Genome Research Institute, NIH
Bethesda, Maryland USA

Summer Intern, Cancer Genetics Branch Summer 2003

National Institute of Dental and Craniofacial Research, NIH
Bethesda, Maryland USA

Summer Intern, Craniofacial Developmental Biology and Regeneration Branch Summer 2000, Summer 2001

HONORS & AWARDS

- Robert H. Riffenburgh Award, American Statistical Association (ASA): 2023
“For his contributions transferring and extending knowledge from machine learning to biological applications. In recognition of the significant impact of his work on research in human genetics and microbiology.”
- W.J. Youden Award in Interlaboratory Testing, American Statistical Association: 2023*
- Elected Fellow, American Statistical Association: 2022
“For transformational methodologic innovation in high-dimensional genomics and microbiome analysis, outstanding leadership in applied multi-omics studies, generosity in education and mentoring, and exemplary service to the ASA community.”
- Elected Chair, ASA Section on Statistics in Genomics and Genetics: 2020
- Invited Keynote, Japanese Society of Computational Statistics: 2016
- Lead author, *American Journal of Epidemiology/Society for Epidemiologic Research* Paper of the Year: 2014[†]
- Lead author, *American Journal of Epidemiology* “Editor’s Choice” Article: 2014[†]
- Lead author, *American Journal of Human Genetics* “Best of” Paper: 2012[‡]
- UW Biostatistics Retreat, Best Faculty Presentation (as voted by students): 2013
- ASA Statistical Computing & Statistical Graphics Sections Student Paper Award: 2008[§]
- ICSA Applied Statistics Symposium Student Paper Award: 2008[§]
- ENAR Distinguished Student Paper Award: 2008[¶]
- Mu Sigma Rho National Statistics Honorary Society: 2007
- Certificate of Distinction in Teaching, Department of Biostatistics: 2006
- NIH Environmental Health Statistics Predoctoral Fellowship: 2008 – 2009
- NIH Interdisciplinary Biostatistics Predoctoral Fellowship: 2005 – 2008
- NIH Cancer Predoctoral Fellowship: 2004 – 2005

*For Publication # 93

†For Publication # 30

‡For Publication # 13

§For Publication # 6 and Tech Report # 5

¶For Tech Report # 6

PROFESSIONAL
ACTIVITIES

Grant Review (Extramural)

- Reviewer, NIH, NIDDK Special Emphasis Panel: ZDK1 GRB-1 (J3) (PA 12-265): 2013.
- Reviewer, NIH, NIEHS: CHEAR Grants: 2015
- External Reviewer, Health Effects Institute: 2012, 2013.
- Biostatistical Reviewer, Susan G. Komen Foundation: 2012, 2014, 2015.
- *Ad hoc* Reviewer, NIH: DDK-C Study Section: 2015, 2016.
- *Ad hoc* Reviewer, NIH: BDMA Study Section: 2020
- *Ad hoc* Reviewer, NIH: IRAP Study Section: 2017, 2019, 2021.
- Reviewer, NIH: BMRD Special Emphasis Panel: 2021
- Reviewer, NIH: PSE Special Emphasis Panel: 2021
- Site Reviewer, NIH/NICHD Division of Intramural Public Health Research (DIPHR): 2021
- Reviewer, NIH, NEI: U24, Characterization of the Resident Ocular Microbiome: 2022
- Reviewer, NIH NIEHS/NTP Intramural Contracts: 2023

Editorial Roles

- *Clinical Cancer Research*, Member, Editorial Board: 2014 – Present.

Paper, Conference & Book Review

1. *American Journal of Bioethics*
2. *American Journal of Epidemiology*
3. *American Journal of Human Genetics*
4. *Annals of Applied Statistics*
5. *Annals of Human Genetics*
6. *Annals of Statistics*
7. *Applied and Environmental Microbiology*
8. *Bioinformatics*
9. *Biometrical Journal*
10. *Biometrics*
11. *Biometrika*
12. *Biostatistics*
13. *BMC Bioinformatics*
14. *BMC Genetics*
15. *BMC Genomics*
16. *BMC Medical Genomics*
17. *BMC Medical Research Methodology*
18. *Briefings in Bioinformatics*
19. *Canadian Journal of Statistics*
20. *Cancer*
21. *Cancer Communications*
22. *Cancer Research*
23. *Carcinogenesis*
24. *Chapman & Hall/CRC Press*
25. *Circulation Research*
26. *Clinical Cancer Research*
27. *Clinical Infectious Diseases*
28. *Computational and Mathematical Methods in Medicine*
29. *Computational Statistics and Data Analysis*
30. *Environmental Health Perspectives*
31. *Environment International*
32. *Epigenetics*
33. *Epigenomics*
34. *European Journal of Human Genetics*
35. *Frontiers in Genetics*
36. *Genetic Epidemiology*
37. *Genetics*
38. *Genetics in Medicine*
39. *Genome Biology*
40. *Human Genetics*
41. *Human Heredity*
42. *Intelligent Systems for Molecular Biology*
43. *Journal of the American Medical Association*
44. *Journal of the American Statistical Association*
45. *Journal of Applied Statistics*
46. *Microbial Ecology in Health & Disease*
47. *Microbiome*

- | | |
|---|---|
| 48. <i>Molecular Psychiatry</i> | 58. <i>PLOS Computational Biology</i> |
| 49. <i>mSystems</i> | 59. <i>PLOS Genetics</i> |
| 50. <i>Mucosal Immunology</i> | 60. <i>PLOS One</i> |
| 51. <i>Nature Communications</i> | 61. <i>Proceedings of the National Academy of Sciences</i> |
| 52. <i>Nature Genetics</i> | 62. <i>Scientific Reports</i> |
| 53. <i>Nature Medicine</i> | 63. <i>Statistical Applications in Genetics and Molecular Biology</i> |
| 54. <i>Nature Microbiology</i> | 64. <i>Statistics in Medicine</i> |
| 55. <i>The Oncologist</i> | 65. <i>Technometrics</i> |
| 56. <i>Oncotarget</i> | |
| 57. <i>The Pharmacogenomics Journal</i> | |

Committee Appointments

- Chair, Planning Committee, ENAR Junior Researcher Workshop, 2012 – 2013.
- Member, Planning Committee, ENAR Junior Researcher Workshop, 2011 – 2015.
- Member, Organizing Committee, BIRS Workshop on *Emerging Statistical Challenges and Methods For Analysis of Massive Genomic Data in Complex Human Disease Studies*, 2014.
- JSM Program Chair, ASA Section on Statistics in Genomics and Genetics, JSM 2017.
- Member, Organizing Committee, BIRS Workshop on *Emerging Statistical Challenges and Methods for Analysis of Human Microbiome Data*, 2019.
- Chair, Student Paper Competition, ASA Section on Statistics in Genomics and Genetics, JSM 2018, 2019, 2020, 2021.
- Chair-Elect/Chair/Past-Chair, ASA Section on Statistics in Genomics and Genetics. 2021 – 2023.

Invited Conference Sessions Organized and Chaired

- Organizer, Session on “Getting to the Gut of the Matter: New Statistics for Studying the Microbiome,” WNAR, 2023, Anchorage, AK.
- Organizer, Session on “Cutting Edge Statistics for Keeping your Microbiome Analyses out of the Toilet,” ENAR, 2022, Houston, TX.
- Organizer, Session on “What to do When Your Data are Poop: Current Methods for Microbiome Analysis,” ENAR, 2021, Baltimore, MD.
- Session Chair, Session on “Bugs, Bugs Everywhere - the Statistics Behind Our Microbiome ,” JSM 2018, Vancouver, BC.
- Organizer and Chair, Session on “Addressing Emerging Statistical Challenges in Microbiome Studies,” JSM 2018, Vancouver, BC.
- Organizer, Session on “Statistical Methods for Addressing Challenges in Microbiome and Metagenomic Analysis,” ENAR 2016, Austin, TX.
- Organizer/Chair, Session on “New Statistical Methods for Analysis of High Dimensional Genomic Data,” ICSA and KISS Joint Applied Statistical Symposium, 2014, Portland, OR.
- Organizer, Session on “Statistical Methods for High Dimensional Genomic Data,” IMS Asia Pacific Rim Meeting, 2012, Tsukuba, Japan.
- Organizer, Session on “Statistical Methods and Applications in Rare Variant Sequencing Studies,” ENAR, 2012, Washington D.C.
- Session Chair, First Joint Biostatistics Symposium, 2010, Beijing, China.
- Session Chair, ENAR, 2009, San Antonio, TX.
- Session Chair, ICSA Applied Statistics Symposium, 2008, Piscataway, NJ.

Data Safety Monitoring Board (DSMB) Membership

- SIMCAP Trial (NCT 03456843)

Professional Organization Memberships:

- American Statistical Association (2003 – Present); Fellow, 2022.
- Institute of Mathematical Statistics (2003 – Present).
- International Society for Computational Biology (2003 – Present).
- Eastern North American Region, International Biometric Society (2007 – Present).
- International Chinese Statistical Association (2008 – Present).
- American Society of Human Genetics (2009 – Present).
- International Genetic Epidemiology Society (2009 – Present).

GRANTS AND CONTRACTS

Active

PI OR SUB-PI GRANTS

- Principal Investigator, “Statistical Methods for Large Scale Microbiome Studies of Cardiovascular Disease Risk”, NIH/NHLBI, R01 HL155417. 2021-2025. \$1,767,647 (total award costs). (Received Impact score of 20; Percentile = 1%)
- Principal Investigator, “Statistical Methods for Enhanced Mapping of Microbiome Relationships”, NIH/NIGMS, R01 GM151301. 2023-2027.
- Subcontract PI and Co-Investigator, “Development of an Integrated Risk Prediction Model of Taxane-induced Peripheral Neuropathy”, NIH/NCI, R37 CA277043. PI: Dan Hertz (Michigan/SWOG). 2023-2027.

NON-PI GRANTS

- Faculty Biostatistician, “SWOG Statistics and Data Management Center (SDMC),” NIH/NCI, U10 CA180819, PI: Michael LeBlanc (Public Health Sciences, FHCC), 2014-2025.
- Faculty Biostatistician, “Lung-MAP: S1400 Biomarker-Targeted Second-Line Therapy in Treating Patients With Recurrent Stage IIIB-IV Squamous Cell Lung Cancer”, SWOG-CTI, PI: Michael LeBlanc (Public Health Sciences, FHCC), 2014-202X.
- Co-Investigator, “Fecal Microbiota Transplantation and Fiber for the Treatment of Graft-versus-host Disease After Hematopoietic Cell Transplantation”, NIH/NHLBI, R01 HL166107. MPIs: David Fredricks (VIDD, FHCC) & David Hockenberry (CRD, FHCC). 2023-2027.

Submitted (Pending Review, Council or Resubmission)

- Multiple Principal Investigator, “Developing Data-Driven Cancer Researchers Training Grant”, NIH/NCI, T32 CA009168, MPIs: Michael Wu and Steve Schwartz (Public Health Sciences, FHCC)
- Subcontract PI and Co-Investigator, “Deciphering Cellular and Genetic Features that Give Rise to Osteonecrosis of the Jaw”, NIH, R01, PI: Lauren Surface (Michigan/SWOG).
- Subcontract PI and Co-Investigator, “Robust Statistical Methods for Longitudinal Microbiome Studies”, NIH, R01, PI: Wodan Ling (Cornell)
- Co-Investigator, “Translational Research Program in Colorectal Disparities (SPORE)”, NIH/NCI. 2024-2029. MPIs: Ulrike Peters (PHS, FHCC), Christopher Li (PHS, FHCC), Timothy Thomas

Completed

PI OR SUB-PI GRANTS

- Principal Investigator, “New Statistical Faculty Translational Medicine Development Award”, The Hope Foundation, 2013-2018 (NCE). \$250,000 (total direct costs).
- Subcontract PI and Co-Investigator, “Overcoming Primary Resistance to PD1 Inhibitors with Combined Immunotherapy,” SU2C AACR/BMS, Catalyst Grant, PI: Toni Ribas (UCLA/SWOG). 2017-2020. \$67,065 (subcontract total costs).
- Subcontract PI and Co-Investigator, “Enhanced Gene Identification in Complex Traits using Kernel Machines,” NIH/NHGRI, R01 HG007508, PI: Michael Epstein (Human Genetics, Emory University), 2013-2017. \$221,577 (total subcontract direct costs).
- Subcontract PI and Co-Investigator, “Investigation of Genetic and Immune Mechanisms of Response to BCG for Non-Muscle Invasive Bladder Cancer: A Translational Study of S1602,” DoD CRCRP-TTSA, CA170270, PIs: Joshua Meeks (Lead PI, Northwestern/SWOG), David McConkey (JHU/SWOG), Robert Svatek (UTHSCSA/SWOG). 2018-2021. \$93,537 (subcontract total costs).
- Subcontract PI and Co-Investigator, “The Vaginal Microbiome and Racial Disparity in Preterm Delivery,” NIH/NIMHD, R01 MD011504, PI: Stephanie Engel (Epidemiology, UNC), 2017-2023 (NCE). \$225,000 (subcontract total costs to date).
- Principal Investigator, “Joint Analysis of Microbiome and Other Genomic Data Types,” NIH/NIGMS, R01 GM129512. 2018-2023 (NCE). \$1,584,000 (total award costs).

NON-PI GRANTS

- Co-Investigator, “The Gut Microbiota and Graft versus Host Disease (GVHD),” NIH/NIAID, R01 AI134808. PI: David Fredricks (VIDD, FHCC). 2017-2022.
- Biostatistician, “MsFLASH: Living a Healthy Menopause”, NIH/NIA, R01 AG048209, MPIs: Andrea LaCroix (Epidemiology, UCSD), Katherine Guthrie (Public Health Sciences, FHCRC), Susan Reed (Women’s Health, UW). 2017-2020 (NCE).
- Co-Investigator, “Estimation and Association Analysis of Biomarkers for Tumor Immune Microenvironment”, NIH/NCI, R21 CA224026. PI: Wei Sun (Public Health Sciences, FHCRC). 2018-2020.
- Co-Investigator, “Mucus and Hypoxia in Heterogeneous and Progressive CF Lung Disease,” NIH/NHLBI, R01 HL116228, PI: Marianne Muhlebach (Pediatrics, UNC), 2012-2013.
- Co-Investigator, “Influence of Genetic Markers on Exposure Assessment Models,” CDC/NIOSH, R21 OH010203, PI: Leena Nylander-French (Environmental Sciences and Engineering, UNC), 2012 – 2013.
- Co-Investigator, “Design and Analysis of Sequencing-based Studies for Complex Human Traits,” NIH/NHGRI, R01 HG006292, PI: Yun Li (Genetics, UNC), 2011–2013.
- Co-Investigator, “Prenatal Smoking, Maternal and Fetal Genetic Variation and Risk of Preeclampsia,” NIH/NICHD, R01 HD058008, PI: Stephanie Engel (Epidemiology, UNC), 2011-2013.
- Co-Investigator, “Elucidating Risks: From Exposure and Mechanism to Outcome - Core D,” NIH/NIEHS, P42 ES005948, PI: James Swenberg (Environmental Sciences and Engineering, UNC), 2011–2013.
- Biostatistician, “UNC Clinical Translation Science Award - Biostatistics Core,” NIH/NCRR, UL1 TR000083, PI: Marschall Runge (Medicine, UNC), 2009 – 2013.
- Biostatistician, “UNC-CH Center for Environmental Health & Susceptibility - Facility Core 2: Biostatistics & Bioinformatics,” NIH/NIEHS, P30 ES010126, PI: James Swenberg (Environmental Sciences and Engineering, UNC), 2009 – 2012.

Published Papers:

1. Koblinski, J.E., **Wu, M.**, Demeler, B., Kleinman, H.K. (2005). "Matrix cell adhesion activation by non-adhesion proteins". *Journal of Cell Science*, 118(Pt 13):2965-74. PMID: 15976454
2. Koblinski, J.E., Kaplan-Singer, B.R., VanOsdol, S.J., **Wu, M.**, Engbring, J.A., Goldsmith, C.M., Piper, J.T., Vostal, J.G., Harms, J.F., Welch, D.R., Kleinman, H.K. (2005). "Endogenous osteonectin/SPARC/BM-40 expression inhibits MDA-MB-231 breast cancer cell metastasis". *Cancer Research*, 65(16):7370-7. PMID: 16103089
3. Ji, H., Ramsey, M.R., Hayes, D.N., Fan, C., McNamara, K., Kozlowski, P., Torrice, C., **Wu, M.C.**, Shimamura, T., Perera, S., Liang, M-C., Cai, D., Naumov, G.N., Bao, L., Contreras, C., Li, D., Chen, L., Krishnamurthy, J., Koivunen, J., Chirieac, L.R., Padera, R., Bronson, R.T., Lindeman, N.I., Christiani, D.C., Lin, X., Shapiro, G.I., Janne, P.A., Johnson, B., Meyerson, M., Kwiatkowski, D.J., Castrillon, D.H., Badeesy, N., Sharpless, N.E., Wong, K-K. (2007). "LKB1 modulates lung cancer differentiation and metastasis". *Nature*, 448(7155):807-10. PMID: 17676035
4. Harezlak, J.#, **Wu, M.C.#**, Wang, M., Schwartzman, A., Christiani D.C., Lin, X. (2008). "Biomarker discovery for arsenic exposure using functional data analysis and feature learning of mass spectrometry proteomic data". *Journal of Proteome Research*, 7(1):217-224. PMID: 18173220
[# Equal Contribution]
5. Engbring, J.A., Hossain, R., VanOsdol, S.J., Kaplan-Singer, B., **Wu, M.**, Hibino, S., Koblinski, J.E. (2008). "The laminin alpha-1 chain derived peptide, AG73, increases fibronectin levels in breast and melanoma cancer cells". *Clinical and Experimental Metastasis*, 25(3):241-252. PMID: 18185912
- ◇6. **Wu, M.C.**, Zhang, L., Wang, Z., Christiani, D.C., Lin, X. (2009). "Sparse linear discriminant analysis for simultaneous testing for the significance of a gene set/pathway and gene selection". *Bioinformatics*, 25(9):1145-1151. PMC2732305
7. Huang, Y.T., Heist, R.S., Chirieac, L.R., Lin, X., Skaug, V., Haugen, A., **Wu, M.C.**, Wang, Z., Su, L., Asomaning, K., Christiani, D.C. (2009). "Genome-wide analysis of survival in early-stage non-small cell lung cancer". *Journal of Clinical Oncology*, 27(16):2660-2667. PMC2690391
8. **Wu, M.C.**, Lin, X. (2009). "Prior biological knowledge based approaches for the analysis of genome-wide expression profiles using gene sets and pathways". *Statistical Methods in Medical Research*, 18(6): 577-593. PMC2827341
9. **Wu, M.C.**, Kraft, P., Epstein, M.P., Taylor, D.M., Chanock, S.J., Hunter, D.J., Lin, X. (2010). "Powerful SNP-set analysis for case-control genome-wide association studies". *The American Journal of Human Genetics*, 86(6): 929-942. PMC3032061
10. Liu, C.Y., **Wu, M.C.**, Chen, F., Ter-Minassian, M., Asomaning, K., Zhai, R., Wang, Z., Su, L., Heist, R., Kulke, M.H., Lin, X., Liu, G., Christiani, D.C. (2010). "A large scale genetic association study of esophageal adenocarcinoma risk." *Carcinogenesis*, 31(7):1259-63. PMC2893800
11. Locke, A.E., Dooley, K.J., Tinker, S.W., Cheong, S.Y., Feingold, E., Allen, E.G., Freeman, S.B., Torfs, C.P., Cua, C., Epstein, M.P., **Wu, M.C.**, Lin, X., Capone, G., Sherman, S.L., Bean, L.J.H. (2010). "Variation in folate pathway genes contributes to risk of congenital heart defects among individuals with Down syndrome". *Genetic Epidemiology*, 34:613-23. PMC3378053
12. Ter-Minassian, M., Wang, Z., Asomaning, K., **Wu, M.C.**, Liu, C.Y., Paulus, J., Liu, G., Bradbury, P., Zhai, R., Su, L., Fraumeni, C., Hooshmand, S.M., DeVivo, I., Lin, X.,

◇ AN EARLIER VERSION WON A 2008 ICSA APPLIED STATISTICS SYMPOSIUM STUDENT AWARD AND A 2008 ASA STATISTICAL COMPUTING & GRAPHICS SECTION STUDENT PAPER AWARD

- Christiani, D.C., Kulke, M.H. (2011). "Genetic associations with sporadic neuroendocrine tumor risk". *Carcinogenesis*, 32(8):1216-22. PMC3149206
- ♣13. **Wu, M.C.**, Lee, S., Cai, T., Li, Y., Boehnke, M., Lin, X. (2011). "Rare variant association testing for sequencing data with the sequence kernel association test (SKAT)". *The American Journal of Human Genetics*, 89(1): 82-93. PMC3135811 [# Equal Contribution]
14. Lin, Xinyi, Cai, T., **Wu, M.C.**, Zhou, Q., Liu, G., Christiani, D.C., Lin, X. (2011). "Kernel machine SNP-set analysis for censored survival outcomes in genome-wide association studies". *Genetic Epidemiology*, 35(7):620-31. PMC3373190
15. Glickman, S.W., Shofer, F.S., **Wu, M.C.**, Scholer, M.J., Ndubizu, A., Cairns, C.B., Glickman, L.T. (2012). "Development and validation of a prioritization rule for obtaining an immediate 12-lead electrocardiogram in the emergency department to identify ST-elevation myocardial infarction". *The American Heart Journal*, 163(3): 372-382. PMID: 22424007
16. Theken, K.N., Schuck, R.N., Edin, M.L., Tran, B., Ellis, K., Bass, A., Lih, F.B., Tomer, K.B., Poloyac, S.M., **Wu, M.C.**, Hinderliter, A.L., Zeldin, D.C., Stouffer, G.A., Lee, C.R. (2012). "Evaluation of cytochrome P450-derived eicosanoids in humans with stable atherosclerotic cardiovascular disease". *Atherosclerosis*, 222(2): 530-6. PMC3361525
17. Lee, S., **Wu, M.C.**, Lin, X. (2012). "Optimal tests for rare variant effects in sequencing association studies". *Biostatistics*, 13(4): 762-75. PMC3440237
18. Wu, H., **Wu, M.C.**, Zhi, D., Santorico, S.A., Cui, X. (2012). "Statistical analysis for next generation sequencing – meeting report". *Frontiers in Genetics*, 3:128. PMC3395795
- ♠19. Joubert, B.R., Håberg, S.E., Nilsen, R.M., Wang, X., Vollset, S.E., Murphey, S.K., Huang, Z., Hoyo, C., Midttun, Ø., Cupul-Uicab, L.A., Ueland, P.M., **Wu, M.C.**, Nystad, W., Bell, D.A., Peddada, S.D., London, S.J. (2012). "450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy". *Environmental Health Perspectives*, 120: 1425-31. PMC3491949
20. Bailey, K.M., **Wu, M.C.**, Ward, W.O., Smeester, L., Rager, J.E., Garcia-Varga, G., Del Razo, L., Drobna, Z., Styblo, M., Fry, R.C. (2013). "Arsenic and the epigenome: interindividual differences in arsenic metabolism related to distinct patterns of DNA methylation". *Journal of Biochemical and Molecular Toxicology*, 27(2): 106-115. PMC3892431
21. Schuck, R.N., Theken, K.N., Edin, M.L., Caughey, M., Bass, A., Ellis, K., Tran, B., Steele, S., Simmons, B.P., Lih, B.P., Tomer, K.B., **Wu, M.C.**, Hinderliter, A.L., Stouffer, G.A., Zeldin, D.R., Lee, C.R. (2013). "Cytochrome P450 derived eicosanoids and vascular function in coronary artery disease patients". *Atherosclerosis*, 227: 442-448. PMC3638946
22. **Wu, M.C.**, Maity, A., Lee, S., Simmons, E.M., Harmon, Q.E., Lin, Xinyi, Engel, S.M., Moll-drem, J.J., Armistead, P.M. (2013). "Kernel machine SNP-set testing under multiple candidate kernels". *Genetic Epidemiology*, 37(3): 267-75. PMC3769109
23. Bencharit, S., Baxter, S.S., Carlson, J., Byrd, W.C., Mayo, M.V., Border, M.B., Kohlfarber, H., Urrutia, E., Saldarriaga, I., Howard-Williams, E.L., Offenbacher, S., **Wu, M.C.**, Buse, J.B. (2013). "Salivary proteins associated with hyperglycemia in diabetes: a proteomic analysis". *Molecular Biosystems*, 9: 2785-97. PMC3888809
24. Harmon, Q.E., Engel, S.M., Olshan, A., Moran, T., Stuebe, A., Luo, J., **Wu, M.C.**, Avery, C. (2013). "Association of polymorphisms in natural killer cell-related genes with preterm birth". *American Journal of Epidemiology*, 178(8): 1208-18. PMC3792727
25. Carroll, I.M., Ringel-Kulka, T., Ferrier, L., **Wu, M.C.**, Siddle, J.P., Bueno, L., Ringel, Y. (2013). "Fecal protease activity is associated with compositional alterations in the intestinal microbiota". *PLOS One*, 8(10): e78017. PMC3798377

♣ THIS PAPER WAS SELECTED BY *AJHG* AS THE "BEST OF" 2011-2012

♠ THIS PAPER WAS SELECTED BY *EHP* AS THE "2013 PAPER OF THE YEAR"

26. Byrnes, A.E., **Wu, M.C.**, Wright, F.A., Li, M., Li, Y. (2013). "The value of statistical or bioinformatics annotation for rare variant association with quantitative trait". *Genetic Epidemiology*, 37(7): 267-75. PMC4083762
27. Cleary, S.P., Jeck, W.R., Zhao, X., Chen, K., Selitsky, S.R., Savich, G.L., Tan, T.X, **Wu, M.C.**, Getz, G., Lawrence, M.S., Parker, J.S., Li, J, Powers, S., Kim, H., Fischer, S., Giundi, M., Ghanekar, A., Chiang, D.Y. (2013). "Identification of driver genes in hepatocellular carcinoma by exome sequencing". *Hepatology*, 58(5): 1693-1702. PMC3830584
28. Sanders, A.P., Smeester, L., Rojas, D., DeBussycher, T., **Wu, M.C.**, Wright, F.A., Zhou, Y., Laine, J.E., Rager, J.E., Swamy, G.K., Ashley-Koch, A., Miranda, M.L., Fry, R.C. (2014). "Cadmium exposure and the epigenome: exposure-associated patterns of DNA methylation in leukocytes from mother-baby pairs". *Epigenetics*, 9(2): 212-221. PMC3962531
29. **Wu, M.C.**, Joubert, B.R., Kuan, P.F., Håberg, S.E., Nystad, W., Peddada, S.D., London, S.J. (2014). "A systematic assessment of normalization approaches for the Infinium 450k methylation platform". *Epigenetics*, 9(2): 318-329. PMC3962542
- [#]30. Engel, S.M.[#], Joubert, B.R.[#], **Wu, M.C.[#]**, Olshan, A.F., Håberg, S.E., Ueland, P.M., Nystad, W., Nilsen, R.M., Vollset, S.E., Peddada, S.D., London, S.J. (2014). "Neonatal genome-wide methylation patterns in relation to birthweight in the Norwegian Mother and Child Birth Cohort". *American Journal of Epidemiology*, 179(7): 834-42. PMC3969535
[[#] Equal Contribution]
31. Ebert, C.S. Jr., McKinney, K.A., Urrutia, E., **Wu, M.**, Rose, A.S., Fleischman, G.M., Thorp, B., Senior, B.A., Zanation, A.M. (2014). "Expression of protease-activated receptors in allergic fungal rhinosinusitis". *International Forum of Allergy & Rhinology*, 4(4): 266-71. PMID:24500840
32. Harmon, Q.E., Engel, S.M., **Wu, M.C.**, Moran, T., Luo, J., Stuebe, A., Avery, C., Olshan, A. (2014). "Polymorphisms in inflammatory genes are associated with term small for gestational age and preeclampsia". *American Journal of Reproductive Immunology*, 71(5): 472-84. PMC4040534
33. Joubert, B.R., Håberg, S.E., Bell, D.A., Nilsen, R.M., Vollset, S.E., Midttun, Ø., Ueland, P.M., **Wu, M.C.**, Nystad, W., Peddada, S.D., London, S.J. (2014). "Maternal smoking and DNA methylation in newborns: *In utero* effects or epigenetic inheritance?". *Cancer Epidemiology, Biomarkers, & Prevention*, 23(6): 1007-17. PMC4140220
34. Nylander-French, L.A., **Wu, M.C.**, Boyner, J., Smeester, L., Sanders, A.P., French, J.E., Fry, R.C. (2014). "DNA methylation modifies urine biomarker levels in 1,6-hexamethylene diisocyanate (HDI) exposed workers: a pilot study". *Toxicology Letters*, 231(2): 217-26. PMID:25445006.
35. Zhao, N., Bell, D.A., Maity, A., Staicu, A.-M., Joubert, B.R., London, S.J., **Wu, M.C.** (2015). "Global analysis of methylation profiles from high resolution CpG data". *Genetic Epidemiology*, 39(2): 53-64. PMC4314375
36. Family, L., Bensen, J.T., Troester, M.A., **Wu, M.C.**, Anders, C.K., Olshan, A.F. (2015). "Single nucleotide polymorphisms in DNA bypass polymerase genes and association with breast cancer and breast cancer subtypes among African Americans and Whites". *Breast Cancer Research and Treatment*, 149(1):181-90. PMC4498665.
37. Hair, B.Y., Troester, M.A., Edmiston, S.N., Parrish, E.A., Robinson, W.R., **Wu, M.C.**, Olshan, A.F., Swift-Scanlan, T., Conway, K. (2015). "Body mass index is associated with gene methylation in estrogen receptor-positive breast tumors". *Cancer Epidemiology, Biomarkers, & Prevention*, 24(3):580-86. PMC4355173
38. Zhao, N., Chen, J., Carroll, I.M., Ringel-Kulka, T., Epstein, M.P., Zhou, H., Zhou, J.J., Ringel, Y., Li, H., **Wu, M.C.** (2015). "Testing in microbiome profiling studies with MiRKAT,

[#]THIS PAPER WAS SELECTED AS AN *AJE* "EDITOR'S CHOICE" ARTICLE AND WAS ALSO SELECTED AS AN *AJE* AND *SER* "2014 PAPER OF THE YEAR"

- the Microbiome Regression-based Kernel Association Test”. *The American Journal of Human Genetics*, 96(5): 797-807. PMC4570290
39. Hair, B.Y., Xu, Z., Kirk, E., Harlid, S., Sandhu, R., Robinson, W.R., **Wu, M.C.**, Olshan, A.F., Conway, K., Taylor, J., Troester, M.A. (2015). “Body mass index associated with genome-wide methylation in breast tissue”. *Breast Cancer Research and Treatment*, 151(2): 453-63. PMC4474159.
 40. Bohlin, J., Andreassen, B.K., Joubert, B.R., Magnus, M., **Wu, M.C.**, Parr, C.L., Håberg, S.E., Magnus, P., Reese, S.E., Stoltenberg, C., London, S.J., Nystad, W. (2015). “Effect of maternal gestational weight gain on offspring DNA methylation: a follow-up to the ALSPAC cohort study”. *BMC Research Notes*, 8:321. PMC4518864
 - ^b41. Urrutia, E., Lee, S., Maity, A., Zhao, N., Shen, J., Li, Y., **Wu, M.C.** (2015). “Rare variant testing across methods and thresholds using the multi-kernel sequence kernel association test (MK-SKAT)”. *Statistics and Its Interface*, 8: 495-505. PMC4698916
 42. Geybels, M.S., Zhao, S., Wong, C., Bibikova, M., Klotzle, B., **Wu, M.**, Ostrander, E.A., Fan, J., Feng, Z., Stanford, J.L. (2015). “Epigenome-wide profiling of DNA methylation in paired prostate cancer versus adjacent benign tissue”. *The Prostate*, 75(16): 1941-50. PMC4928710
 43. Chen, J., Chen, W., Zhao, N., **Wu, M.C.**, Schaid, D.J., (2016). “Small sample kernel association test for human genetic and microbiome association studies”. *Genetic Epidemiology*, 40(1): 5-19. PMC4679685
 44. Joubert, B.R., den Dekker, H.T., Felix, J.F., Bohlin, J., Beckett, E., Tiemeier, H., van Meurs, J.B., Uitterlinden, A.G., Hofman, A., Håberg, S.E., Reese, S.E., Andreassen, B.K., Nilsen, R.M., Vollset, S.E., Midtun, Ø., Ueland, P.M., **Wu, M.C.**, Wang, T., Peddada, S.D., Jad- doe, V.W.V., Nystad, W., Dujits, L., London, S.J. (2016). “Maternal plasma folate during pregnancy impacts differential DNA methylation in an epigenome-wide meta-analysis of new- borns”. *Nature Communications*, 7:10577. PMC474995
 45. Suderman, M., Stene, L.C., Bohlin, J., Page, C.M., Holvik, K., Parr, C.L., Magnus, M.C., Håberg, S.E., Joubert, B.R., **Wu, M.C.**, London, S.J., Relton, C., Nystad, W. (2016). “25-hydroxyvitamin D in pregnancy and genome wide cord blood DNA methylation in two pregnancy cohorts (MoBa and ALSPAC)”. *The Journal of Steroid Biochemistry and Molecular Biology*, 159:102-109. PMC4829940
 46. Lin, Xinyi, Lee, S., **Wu, M.C.**, Wang, C., Chen, H., Li, Z., Lin, X. (2016). “Test for rare variants by environment interactions in sequencing association studies”. *Biometrics*, 72: 156-164. PMC4733434
 - [%]47. Joubert, B.R., Felix, J.F., Yousefi, P., Bakulski, K.M., Just, A.C., Breton, C., Reese, S., Markunas, C., ..., **Wu, M.C.**, ... , London, S.J. (2016). “DNA methylation in newborns and maternal smoking in pregnancy: Genome-wide consortium meta-analysis”. *The American Journal of Human Genetics*, 98(4): 680-696. PMC4833289
 48. Dai, H., Wu, G., **Wu, M.**, Zhi, D. (2016). “An optimal bahadur-efficient method in detection of sparse signals with applications to pathway analysis in sequencing association studies”. *PLOS One*, 11(7):e0152667. PMC4933358
 49. He, Q.#, Cai, T., Liu, Y., Zhao, N., Harmon, Q.E., Almli, L.M., Binder, E.B., Engel, S.M., Ressler, K.J., Conneely, K.N., Lin, X., **Wu, M.C.#** (2016). “Prioritizing individual variants after kernel machine testing using variable selection”. *Genetic Epidemiology*, 40(8):722-731. PMC5118060 [# Joint Corresponding Authors]
 50. Rotroff, D.M., Joubert, B.R., Marvel, S.W., Håberg, S.E., **Wu, M.C.**, Nilsen, R.M., Ueland, P.M., Nystad, W., London, S.J., Motsinger-Reif, A. (2016). “Maternal smoking impacts key biological pathways in newborns through epigenetic modification in utero”. *BMC Genomics*, 17(1):976. PMC5124223

^bAN EARLIER VERSION WON AN ASHG PREDOCTORAL TRAINEE RESEARCH SEMIFINALIST AWARD

[%]THIS PAPER WAS SELECTED BY *AJHG* AS THE “BEST OF” 2015-2016

51. Zhan, X., Girirajan, S., Zhao, N., **Wu, M.C.**, Ghosh, D. (2016). “A novel copy number variants kernel association test with application to autism spectrum disorders studies”. *Bioinformatics*, 32 (23): 3603-3610. PMC5181531
52. Plantinga, A., Zhan, X., Zhao, N., Chen, J., Jenq, R.R., **Wu, M.C.** (2017). “MiRKAT-S: A community-level test of association between the microbiota and survival times”. *Microbiome* 5(2): 17. PMC5299808
53. Zhan, X., Tong, X., Zhao, N., Maity, A., **Wu, M.C.**, Chen, J. (2017). “A small-sample multivariate kernel machine test for microbiome association studies”. *Genetic Epidemiology*, 41(3):210-220. PMC5560787 [**# Joint Corresponding Authors**]
54. Reese, S.E., Zhao, S., **Wu, M.C.**, Joubert, B.R., Parr, C.L., Håberg, S.E., Ueland, P.M., Nilsen, R.M., Midttun, Ø., Vellset, S.E., Peddada, S., Nystad, W., London, S.J. (2017). “DNA methylation score as a biomarker in newborns for sustained maternal smoking during pregnancy”. *Environmental Health Perspectives*, 125(4): 760-766. PMC5381987
55. Zhan, X., Zhao, N., Plantinga, A., Thornton, T.A., Conneely, K.N., Epstein, M.P., **Wu, M.C.** (2017). “Powerful genetic association analysis for common or rare variants with high-dimensional structured traits”. *Genetics*, 206(4): 1779-1790. PMC5560787
56. Mitchell, C.M., Srinivasan, S., Zhan, X., **Wu, M.C.**, Reed, S.D., Guthrie, K.A., LaCroix, A.Z., Fiedler, T., Munch, M., Liu, C., Hoffman, N., Blair, I.A., Newton, K., Freeman, E.W., Joffe, H., Cohen, L., Fredricks, D.N. (2017). “Vaginal microbiota and genitourinary menopausal symptoms: a cross sectional analysis”. *Menopause*, 24(10):1160-1166. PMC5607086
57. Golob, J.L., Pergam, S.A., Srinivasan, S., Fiedler, T.L., Liu, C., Garcia, K., Mielcarak, M., Ko, D., Aker, S., Marquis, S., Loeffelholz, T., Plantinga, A., **Wu, M.C.**, Celustka, K., Morrison, A., Woodfield, M., Fredricks, D.N. (2017). “The stool microbiota at neutrophil recovery is predictive for severe acute graft versus host disease after hematopoietic cell transplantation.” *Clinical Infectious Diseases*, 65(12):1984-1991. PMC5850019
- [&]58. Zhan, X., Plantinga, A., Zhao, N., **Wu, M.C.** (2017). “A fast small-sample kernel independence test for microbiome community-level association analysis”. *Biometrics*, 73(4):1453-1463. PMC5592124
59. Bauer, A.E., Avery, C.L., Shi, M., Weinberg, C.R., Olshan, A.F., Harmon, Q.E., Luo, J., Yang, J., Manuck, T., **Wu, M.C.**, Williams, N., McGinnis, R., Morgan, L., K’ungsøyr, K., Trostad, L., Magnus, P., Engel, S.M. (2018). “A family-based study of carbon monoxide and nitric oxide signaling genes and preeclampsia”. *Paediatric and Perinatal Epidemiology*, 32:1-12. PMC5771849
- [△]60. Zhao, N., Zhan, X., Huang, Y.-T., Almli, L.M., Smith, A., Epstein, M.P., Conneely, K.N., **Wu, M.C.** (2018). “Kernel machine methods for integrative analysis of genome-wide methylation and genotyping studies”. *Genetic Epidemiology*, 42(2):156-167.
61. Teran Hidalgo, S.J., **Wu, M.C.**, Engel, S.M., Kosorok, M.R. (2018). “Goodness-of-fit test for nonparametric regression models: smoothing splines ANOVA models as example”. *Computational Statistics and Data Analysis*, 122: 135-155. PMC5983390
62. Fujii, R., Hishida, A., **Wu, M.C.**, Kondo, T., Hattori, Y., Naito, M., Endoh, K., Nakatochi, M., Hamajima, N., Kubo, M., Kuriki, K., Wakai, K. (2018). “Genome-wide association study for pollinosis identified two novel loci in interleukin (IL)-1B in a Japanese population”. *Nagoya Journal of Medical Science*, 80: 109-120. PMC5857507
63. Luo, Y., Maity, A., **Wu, M.C.**, Smith, C., Duan, Q., Li, Y., Tzeng, J.-Y. (2018). “On the sub-structure controls in rare variant analysis — principal components or variance components?”. *Genetic Epidemiology*, 42(3):276-287. PMC5851819

[&]SELECTED AS A TOP DOWNLOADED ARTICLE IN *Biometrics* 2017-2018; AND AN EARLIER VERSION WON A 2017 ASA BIOMETRICS SECTION PAPER AWARD

[△]SELECTED AS THE IGES HIGHLIGHT/PAPER OF THE MONTH

64. Mitchell, C.M., Srinivasan, S., Plantinga, A., **Wu, M.C.**, Reed, S.D., Guthrie, K.A., LaCroix, A.Z., Fiedler, T., Munch, M., Liu, C., Hoffman, N.G., Blair, I.A., Newton, K., Freedman, E.W., Joffe, H., Cohen, L., Fredricks, D.N. (2018). "Associations between improvement in genitourinary symptoms of menopause and changes in the vaginal ecosystem". *Menopause*, 25(5): 500-507. PMC5898977
65. Manuzak, J.A., Gott, T.M., Kirkwood, J.S., Coronado, E., Hensley-McBain, T., Miller, C., Cheu, R.K., Collier, A.C., Funderburg, N.T., Martin, J.N., **Wu, M.C.**, Isoherranen, N., Hunt, P.W., Klatt, N.R. (2018). "Heavy cannabis use associated with reduction in activated and inflammatory immune cell frequencies in antiretroviral therapy-treated Human Immunodeficiency Virus-infected individuals." *Clinical Infectious Diseases*, 66(12): 1872-1882. PMC6248381
66. Stewart, T.G., Zeng, D., **Wu, M.C.** (2018). "Constructing support vector machines with missing data". *WIREs Computational Statistics*. 10(4): e1430.
67. Zhan, X., **Wu, M.C.** (2018). "A note on testing and estimation in marker-set association study using semiparametric quantile regression kernel machine". *Biometrics*, 74(2): 764-766. PMC5932287
68. Zhan, X., Xue, L., Zheng, H., Plantinga, A., **Wu, M.C.**, Schaid, D.J., Zhao, N., Chen, J. (2018). "Small sample kernel association tests for correlated data with application to microbiome data analysis". *Genetic Epidemiology*, 42(8):772-782.
69. Marceau-West, R., Lu, W., Rotroff, D.M., Kuenemann, M., Chang, S., **Wu, M.C.**, Wagner, M.J., Buse, J.B., Motsinger-Reif, A., Fourches, D., Tzeng, J. (2019). "Identifying individual risk rare variants using protein structure-guided local tests (POINT)". *PLoS Computational Biology*, 15(2):e1006722. PMC6396946
- [∇]70. Sun, S., Lulla, A., Sioda, M., Winglee, K., **Wu, M.C.**, Jacobs, D.R., Shikany, J.M., Lloyd-Jones, D.M., Launer, L.J., Fodor, A.A., Meyer, K.A. (2019). "Gut microbiota composition and blood pressure: Coronary Artery Risk Development in Young Adults (CARDIA) Study". *Hypertension*, 73(5):998-1006. PMC6458072
71. Hensley-McBain, T., **Wu, M.C.**, Manuzak, J., Cheu, R., Gustin, A., Driscoll, C., Zevin, A.S., Miller, C., Coronado, E., Smith, E., Law, L., Gale Jr., M., Somsouk, M., Burgener, A.D., Hunt, P.W., Hope, T., Collier, A., Klatt, N.R. (2019). "Increased mucosal neutrophil survival is associated with altered microbiota in HIV infection". *PLoS Pathogens*, 15(4):e1007672. PMC6459500
72. Hishida, A., Ugai, T., Fujii, R., Nakatochi, M., **Wu, M.C.**, Ito, H., Oze, I., Masahiro, T., Yasumasa, N., Nishiyama, T., Nakagawa-Senda, H., Suzuki, S., Koyama, T., Matsui, D., Watanabe, Y., Kawaguchi, T., Matsuda, F., Momozawa, Y., Kubo, M., Naito, M., Matsuo, K., Wakai, K. (2019). "GWAS and genome-wide SNP-set analysis for the risk of H. pylori-induced gastric atrophy revealed the significant contribution of PSCA to its risk". *Carcinogenesis*, 40(5):661-668. PMID: 30753327
73. Zhao, N., Zhang, H., Clark, J.J., Maity, A., **Wu, M.C.** (2019). "Composite kernel machine regression based on likelihood ratio test for joint testing of genetic and gene-environment interaction effect". *Biometrics*, 75(2): 363-717. PMID: 30430548
74. Plantinga, A., Chen, J., Jenq, R.R., **Wu, M.C.** (2019). "pldist: Ecological dissimilarities for paired and longitudinal microbiome association analysis". *Bioinformatics*, 35(19):3567-3575. PMC6761933
75. Golob, J.L., DeMeules, M.M., Loeffelholz, T., Quinn, Z.Z., Dame, M., **Wu, M.C.**, Schmidt, T., Fiedler, T.L., Pergam, S.A., Fredricks, D.N. (2019). "Butyrogenic bacteria after severe acute graft versus host disease associate with the development of chronic gut GVHD and steroid refractory GVHD." *Blood Advances*, 3(19): 2866-2869. PMC6784520

76. Kaplan, R.C., Wang, Z., Usyk, M., Sotres-Alvarez, D., Daviglius, M.L., Schneiderman, N., Talavera, G.A., Gellman, M.D., Thyagarajan, B., Moon, J.Y., Vazquez-Baeza, Y., McDonald, D., Williams-Nguyen, J.S., **Wu, M.C.**, North, K.E., Shaffer, J., Sollecito, C.C., Qi, Q., Isasi, C.R., Wang, T., Knight, R., Burk, R.D. (2019). "Gut microbiome composition in the Hispanic Community Health Study/Study of Latinos is shaped by geographic relocation, environmental factors, and obesity". *Genome Biology*, 20(1):219. PMC6824043
77. Liu, M., Liu, Y., **Wu, M.C.**, Hsu, L., He, Q. (2021). "A method for subtype analysis with somatic mutations". *Bioinformatics*, 37(1):50-56. PMID:33416828
78. Mitchell, C.M., Srinivasan, S., Ma, N., Reed, S.D., **Wu, M.C.**, Hoffman, N.G., Valint, D., Proll, S., Fiedler, T., Agnew, K., Guthrie, K.A., Fredricks, D.N. (2021). "Bacterial communities associated with abnormal Nugent score are different in pre- vs. post-menopausal women". *Journal of Infectious Diseases*, 223(12):2048-2052. PMC8350750
79. Wilson, N., Zhao, N., Zhan, X., Fu, W., Chen, J., Li, H., **Wu, M.C.**, Plantinga, A. (2021). "MiRKAT: Kernel machine regression-based global association tests for the microbiome". *Bioinformatics*, 37(11):1595-1597. PMC8495888
80. Mitchell, C.M., Ma, N., Mitchell, A., **Wu, M.C.**, Valint, D.J., Proll, S., Reed, S.D., Guthrie, K.A., Larson, J., Raftery, D., Fredricks, D.N., Srinivasan, S. (2021). "Association between postmenopausal vulvovaginal discomfort, vaginal microbiota and mucosal inflammation: Vaginal microbiome and treatment response in postmenopausal women". *American Journal of Obstetrics & Gynecology*, 225(2):159.e1-159.e15. PMC8328873
81. Ling, W., Zhao, N., Plantinga, A., Launer, L., Fodor, A., Meyer, K., **Wu, M.C.** (2021). "Powerful and robust non-parametric association testing for microbiome data via a zero-inflated quantile approach (ZINQ)". *Microbiome*, 9(1):181. PMC8414689
82. He, Q., Liu, Y., Liu, M., **Wu, M.C.**, Hsu, L. (2021). "Random effect based tests for multinomial logistic regression in genetic association studies". *Genetic Epidemiology*, 45(7):736-40. PMC9209005
83. Liu, H., Plantinga, A.M., Xiang, Y., **Wu, M.C.** (2021). "A kernel-based test of independence for cluster-correlated data". *Neural Information Processing Systems (NeurIPS)*, 34: 9869-9881. PMC9801702
84. Ling, W., Qi, Y., Hua, X., **Wu, M.C.** (2021). "Deep ensemble learning over the microbial phylogenetic tree (DeepEn-Phy)". *IEEE Bioinformatics and Biomedicine (BIBM)*, 470-477. PMC9875567
85. Hudson, P.L., Ling, W., **Wu, M.C.**, Hayward, M.R., Mitchell, A.J., Larson, J., Guthrie, K.A., Reed, S., Kwon, D.S., Mitchell, C. (2021). "Comparison of the vaginal microbiota in postmenopausal Black and White women". *Journal of Infectious Diseases*, 224(11): 1945-1949. PMC8825215
86. Kimmel, M., Jin, W., Xia, K., Lun, K., Azcarate-Peril, M.A., Plantinga, A., **Wu, M.**, Ataei, S., Rackers, H., Carroll, I., Meltzer-Brody, S., Fransson, E., Knickmeyer, R. (2022). "Metabolite trajectories across the perinatal period and mental health: A preliminary study of tryptophan-related metabolites, bile acids and microbial composition". *Behavioural Brain Research*, 418:113635. PMID:34755640
87. Wang, T., Ling, W., Plantinga, A.M., **Wu, M.C.**, Zhan, X. (2022). "Testing microbiome association using integrated quantile regression models". *Bioinformatics*, 38(2): 419-425. PMID:34554223
88. Rosen, E.M., Martin, C.L., Siega-Riz, A.M., Dole, N., Ludwig, K., Basta, P.V., Serrano, M., Fettweis, J., **Wu, M.**, Sun, S., Thorp, J.M., Buck, G., Fodor, A.A., Engel, S.M. (2022). "Is prenatal diet associated with the composition of the vaginal microbiome?". *Paediatric and Perinatal Epidemiology*, 36(2):243-253. PMC8881389
89. Fredricks, D.N., Plantinga, A., Srinivasan, S., Oot, A., Wisner, A., Fiedler, T.L., Proll, S., **Wu, M.C.**, Marrazzo, J.M. (2022). "Extra-vaginal bacterial colonization and risk for incident

- bacterial vaginosis in a population of women who have sex with men”. *Journal of Infectious Diseases*, 225(7): 1261-1265. PMC8974833
90. Srinivasan, S., Hua, X., **Wu, M.C.**, Prohl, S., Valint, D.J., Reed, S.D., Guthrie, K.A., Lacroix, A.Z., Larson, J.C., Pepin, R., Raftery, D., Fredricks, D.N., Mitchell, C.M. (2022). “Impact of topical interventions on the vaginal microbiota and metabolome in postmenopausal women: A secondary analysis of a randomized clinical trial”. *JAMA Network Open*, 5(3), e225032. PMC8968546
 91. Goo, J., Liu, Y., Sun, W., Hsu, L., **Wu, M.C.**, He, Q. (2022). “TCR-L: an analysis tool for evaluating the link between the T-cell receptor repertoire and clinical phenotypes”. *BMC Bioinformatics*, 23(1):152. PMC9052542
 92. Sun, S., Serrano, M., Fetweiss, J., Basta, P., Rosen, E., Ludwig, K., Sorgen, A., Blakley, I., **Wu, M.C.**, Dole, N., Thorp, J., Siega-Riz, A., Buck, G.A., Fodor, A.A., Engel, S.M. (2022). “Race, the vaginal microbiome, and spontaneous preterm birth”. *mSystems*, e00017-22. PMC9238383
 - **93. Ling, W., Lu, J., Zhao, N., Lulla, A., Plantinga, A.M., Fu, W., Zhang, A., Liu, H., Song, H., Li, Z., Chen, J., Randolph, T., Koay, W.L.A., White, J.R., Launer, L.J., Fodor, A.A., Meyer, K.A., **Wu, M.C.** (2022). “Batch effects removal for microbiome data via conditional quantile regression (ConQuR)”. *Nature Communications*, 13(1):5418. PMC9477887
 94. Huang, C., Callahan, B., **Wu, M.C.**, McLaren, M., Brochu, H., Lu, W., Peng, X., Tzeng, J.Y. (2022). “Phylogeny-guided microbiome OTU-specific association test (POST)”. *Microbiome*, 10(1):86. PMC9171974
 95. Song, H., Liu, H., **Wu, M.C.** (2022). “A fast kernel independence test for cluster-correlated data”. *Scientific Reports*, 12(1): 21659. PMC9755291
 96. Song, H., Ling, W., Zhao, N., Plantinga, A.M., Broedlow, C.A., Klatt, N.R., Hensley-McBain, T., **Wu, M.C.** (2023). “Accommodating multiple potential normalizations in microbiome associations studies”. *BMC Bioinformatics*, 24, 22. PMC9850542
 97. Carter, K.A., Fodor, A.A., Balkus, J.E., Zhang, A., Serrano, M.G., Buck, G.A., Engel, S.M., **Wu, M.C.**, Sun, S. (2023). “Vaginal microbiome metagenome inference accuracy: differential measurement error according to community composition.” *mSystems*, 8(2):e0100322. PMC10134888
 98. Liu, H., Ling, W., Hua, X., Moon, J., Williams-Nguyen, J., Zhan, X., Plantinga, A.M., Zhao, N., Zhang, A., Knight, R., Qi, Q., Burk, R.D., Kaplan, R.C., **Wu, M.C.** (2023). “Kernel-based genetic association analysis for microbiome phenotypes identifies host genetic drivers of beta-diversity”. *Microbiome*, 8(2):e0100322. PMC10134888
 99. Byrd, D.A., Fan, W., Greathouse, L., **Wu, M.C.**, Xie, H., Wang, X. (2023). “The intratumor microbiome is associated with microsatellite instability.” *Journal of the National Cancer Institute*, 115(8):989-993. PMC10407713
 100. VanderWalde, A., Bellasea, S., Kendra, K., Khushalani, N.I., Campbell, K.M., Scumpia, P.O., Kuklinski, L.F., Collichio, F., Sosman, J.A., Ikeguchi, A., Victor, A.I., Truong, T., Chmielowski, B., Portnoy, D.C., Chen, Y., Margolin, K., Bane, C., Dasanu, C.A., Johnson, D.B., Eroglu, Z., Chandra, S., Medina, E., Gonzalez, C.R., Baselga Carretero, I., Vega-Crespo, A., Perez Garcilazo, I., Sharon, E., Hu-Lieskovan, S., Patel, S.P., Grossmann, K.F., Moon, J., **Wu, M.C.**, Ribas, A. (2023). “Ipilimumab with or without nivolumab in PD1/PDL1 blockade refractory metastatic melanoma: a randomized phase 2 trial”. *Nature Medicine*, 29(9): 2278-2285.
 101. Bauer, A.E., Avery, C.L., Shi, M., Weinberg, C.R., Olshan, A.F., Harmon, Q.E., Luo, J., Yang, J., Manuck, T., **Wu, M.C.**, Klungsøyr, K., Trostad, L., Magnus, P., Engel, S.M. “Do genetic variants modify the effect of smoking on risk of preeclampsia in pregnancy?”. *American Journal of Perinatology*, In Press. PMC10127527

**THIS PAPER RECEIVED THE W.J. YODEN AWARD IN INTERLABORATORY TESTING FROM THE AMERICAN STATISTICAL ASSOCIATION

102. Song, H., **Wu, M.C.** “Limitation of permutation-based differential correlation analysis”. *Genetic Epidemiology*, In Press.

Published Book Chapters:

103. **Wu, M.C.**, Kuan, P.F. (2018). “A Guide to Illumina BeadChip Data Analysis”. In: Tost, J. (ed): *DNA Methylation Protocols, 3rd Edition*. Springer Methods in Molecular Biology, vol 1708; 303-330. Humana Press, New York, NY. PMID: 29224151
104. Plantinga, A. **Wu, M.C.** (2021). “Beta Diversity and Distance-based Analysis of Microbiome Data”. In: Datta, S., Guha, S. (eds): *Statistical Analysis of Microbiome Data*; 101-127. Springer Frontiers in Probability and Statistical Sciences, Springer Nature, Switzerland.

Submitted:

- ^{††}105. Stewart, T.G., **Wu, M.C.**, Zeng, D. “Augmented weighted support vector machines for missing covariates”.
106. Pete, D., Lampe, J.W., **Wu, M.C.**, Salama, N.R., Phipps, A.I. “Prevalence of Helicobacter pylori, cagA virulence gene carriage, and drivers for Helicobacter pylori infections in adults in the Navajo Nation”.
107. Sachdeva, A., **Wu, M.C.**, Guha, S. “Bayesian predictive models for high-dimensional microbiome data”.
108. Plantinga, A., **Wu, M.C.** “Multilevel compositional lasso”.
109. Kozono, D., Hua, X., **Wu, M.C.**, Tolba, K.A., Waqar, S.N., Dragnev, K.H., Cheng, H., Hirsch, F.R., Mack, P.C., Gray, J., Kelly, K., Borghaei, H., Herbst, R.S., Gandara, D.R., Redman, M.W. “Lung-MAP next generation sequencing analysis of advanced squamous cell lung cancers (SWOG S1400)”.
110. Fu, W., Plantinga, A., Ma, N., Zhan, X., Zhao, N., Martinez, K., Maity, A., Tzeng, J.-Y., **Wu, M.C.** “Screening for associations between microbiome community profiles and a large number of individual genomic outcomes using robust kernel machine methods”.
111. Ma, N., Plantinga, A., **Wu, M.C.** “Identifying genomic features related to microbiome community composition using a sparse kernel RV coefficient”.

Technical Reports:

All Reports Available Upon Request:

1. **Wu, M.C.**, Neuberg, D.S., Manola, J., Vail, S. (2005). “Phase III randomized trial of interferon alfa-2b alone versus interferon alfa-2b plus thalidomide in patients with previously untreated metastatic or unresectable renal cell carcinoma”. Technical Report 1128E of the *Eastern Cooperative Oncology Group*.
2. **Wu, M.C.**, Lee, J.S. (2006). “Gaussian mixture modeling of polarimetric synthetic aperture radar images for unsupervised terrain classification”. Technical Report.
3. **Wu, M.C.**, Lin, X. (2007). “Adaptively penalized logistic regression for testing the significance of gene pathways”. Technical Report.
4. Harezlak, J., Jiang, S., **Wu, M.C.**, Wang, M., Christiani, D.C., Lin, X. (2007). “Joint modeling of functional SELDI-TOF mass spectrometry proteomic data and disease risk”. Technical Report.
- ^{††}5. **Wu, M.C.**, Zhang, L., Lin, X. (2008). “Two group classification using sparse linear discriminant analysis”. Technical Report.

^{††}EARLIER VERSIONS OF THIS WORK WON BOTH THE ENAR POSTER COMPETITION AND THE ENAR DISTINGUISHED STUDENT PAPER AWARD

^{††}THIS PAPER WON A 2008 ASA STATISTICAL COMPUTING & GRAPHICS SECTION STUDENT PAPER AWARD

- ^{§§}6. **Wu, M.C.**, Cai, T., Lin, X. (2009). “Testing for regression coefficients in LASSO regularized regression”. Technical Report.
7. **Wu, M.C.** (2009). “Statistical methods for high-dimensional genomic data”. Ph.D. Dissertation, Harvard University.
8. Wang, Z., Peterson, M.S., Zhang, L., **Wu, M.C.**, Su, L., Lin, X., Grandjean, P., Christiani, D.C. (2009). “Plasma SELDI-TOF protein profiling in a methylmercury and polychlorinated biphenyl exposed population”. Technical Report.
9. Lee, S., **Wu, M.C.**, Cai, T., Li, Y., Boehnke, M., Lin, X. (2011). “Power and sample size calculations for designing rare variant sequencing association studies”. Technical Report.
10. Li, L., **Wu, M.C.**, Lin, X. (2012). “Prioritizing phenotype-associated uncommon variants in sequencing association studies”. Technical Report.
11. Family, L., Bensen, J.T., Troester, M.A., **Wu, M.C.**, Anders, C.K., Olshan, A.F. (2015). “Single nucleotide polymorphisms in base excision repair pathway genes and association with breast cancer and breast cancer subtypes among African Americans and Whites”. Technical Report.
12. **Wu, M.C.**, Dai, J., LeBlanc, M. (2017). “Design and Analysis Guidelines for Translational Medicine Studies (v. 1.0)”. Technical Report, *SWOG SDMC*.

EXTRAMURAL
TALKS &
PRESENTATIONS

1. TBD, Department of Biostatistics, UCLA, Los Angeles, CA. (Apr 2024). {Invited}
2. TBD, Division of Biostatistics, University of Pennsylvania, Philadelphia, PA. (Mar 2024). {Invited}
3. TBD, CANSSI STAGE International Speaker Seminar Series, University of Toronto, Toronto, Canada. (Feb 2024). {Invited}
4. “Assessing the Conditional Correlation Between Individual Genomic Features and Microbial Taxa”, JSM, Toronto, Canada. (August 2023). {Invited}
5. “Considerations for Keeping your Microbiome Studies out of the Toilet”, Nicholas Vogelzang Symposium: “Microbiome as a New Frontier in SWOG Clinical Trials”, SWOG Group Meeting, San Francisco, CA. (May 2023). {Invited}
6. “Keeping your Microbiome Analyses out of the Toilet: Batch Effects and Robust Testing”, Department of Biostatistics, Memorial Sloan Kettering Cancer Center. (April 2023). {Invited}
7. “Assessing the Conditional Correlation Between Individual Genomic Features and Microbial Taxa”, ENAR, Nashville, TN. (March 2023). {Invited}
8. “Assessing the Conditional Correlation Between Individual Genomic Features and Microbial Taxa”, Advances in Statistical and Computational Methods for Analysis of Biomedical, Genetic, and Omics Data, University of Texas, Dallas, TX. (March 2023). {Invited}
9. “Integrative Analysis of Microbiome and Other Omics Data”, CHARGE Investigator Meeting, Seattle, WA. (October 2022). {Invited}
10. “Keeping your Microbiome Analyses out of the Toilet: Batch Effects and Robust Testing”, Department of Biostatistics, University of Michigan, Ann Arbor, MI. (September 2022). {Invited}
11. “Kernel-based genetic association analysis for microbiome phenotypes identifies host genetic drivers of beta-diversity”, ICSA-Canada, Banff, Canada. (July 2022). {Invited}
12. “Statistical Analysis of Large-Scale Microbiome Profiling Studies: Batch Effects and Robust Testing”, Channing Division of Network Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA. (April 2022). {Invited}
13. “Kernel Methods for Integrating Microbiome and Other Omics Data”, ENAR, Houston, TX. (March 2022). {Invited}

^{§§}AN EARLIER VERSION WON AN ENAR DISTINGUISHED STUDENT PAPER AWARD

14. “Statistical Analysis of Large-Scale Microbiome Profiling Studies: Batch Effects and Robust Testing”, Institute for Mathematical and Statistical Innovation (IMSI) workshop on *Multiscale Microbial Communities*. Chicago, IL. (February 2022). {Invited}
15. “Kernel Methods for Identifying of Genetic Variants Associated with Microbiome Composition”, JSM, Seattle, WA. (August 2021). {Invited}
16. “ConQuR-ing Batch Effects in Microbiome Studies”, WNAR, Virtual, Anchorage, AK. (June 2021). {Invited}
17. “ConQuR-ing Batch Effects in Microbiome Studies”, ENAR, Virtual, Baltimore, MD. (March 2021). {Invited}
18. “Robust Analysis of Large Scale Microbiome Studies”, Division of Cancer Epidemiology and Genetics, NCI, NIH, Virtual, Bethesda, MD. (March 2021). {Invited}
19. “Robust Analysis of Large Scale Microbiome Studies”, Department of Biostatistics, Virginia Commonwealth University, Virtual, Richmond, VA. (November 2020). {Invited}
20. “Statistics for Poop Data: Keeping Your Microbiome Analyses out of the Toilet”, JSM, Virtual, Philadelphia, PA. (August 2020). {Invited}
21. “Incorporating Auxiliary Information to Improve Microbiome-based Prediction Models”, ENAR, Virtual, Nashville, TN. (March 2020). {Invited}
22. “Statistical Integration of Microbiome and Host Genomic Data”, Human Medical Genetics and Genomics Program, University of Colorado, Denver, CO. (January 2020). {Invited}
23. “Integrative Analysis of Microbiome and other Genomic Data Types”, iBRIGHT, Houston, TX. (November 2019). {Invited}
24. “Testing Associations Between Microbiome and Other Omics Data Types”, Workshop on Emerging Statistical Challenges and Methods For Analysis Microbiome Data, Banff International Research Station, Banff, Canada. (September 2019). {Invited}
25. “Statistical Approaches for Jointly Analyzing Microbiome and other -Omics Data Types”, JSM, Denver, CO. (August 2019). {Topic Contributed}
26. “Development of Microbiome-based Prediction Models using Co-informative Information”, ICSA China Conference, Tianjin, China. (July 2019). {Invited}
27. “Development of Microbiome-based Prediction Models using Co-informative Information”, EcoSta, Taichung, Taiwan. (June 2019). {Invited}
28. “Statistical Approaches for Paired and Longitudinal Microbiome Association Studies”, Harvard Chan Center for the Microbiome in Public Health, Harvard University, Boston, MA. (May 2019). {Invited}
29. “Testing Statistical Interactions between Microbiome Community Profiles and Covariates”, ENAR, Philadelphia, PA. (March 2019). {Invited}
30. “Testing Statistical Interactions between Microbiome Community Profiles and Covariates”, 11th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics), Pisa, Italy. (December 2018). {Invited}
31. “Statistical Methods for Correlating Microbiome and other -Omics Data”, Department of Biostatistics, Columbia University, New York, NY. (November 2018) {Invited}
32. “Identifying Associations Between Genetic Variants and Microbiome Community Diversity using Kernel Methods”, ASHG, San Diego, CA. (October 2018) {Poster}
33. “Statistical Methods for Associating Genetic Markers and Complex, High-dimensional Phenotypes”, CHARGE Investigator Meeting, Baltimore, MD. (October 2018). {Invited}
34. “Microbiome in Cancer Clinical Trials”, NCTN Statistical Group Meeting, Portland, OR. (May 2018). {Invited}
35. “Joint Modeling and Analysis of Microbiome with Other Omics Data”, ENAR, Atlanta, GA. (March 2018). {Invited}

36. “Statistical Methods for Correlating Microbiome and other –Omics Data”, Department of Biostatistics, MD Anderson Cancer Center, Houston, TX. (October 2017). {Invited}
37. “Joint Modeling and Analysis of Microbiome with Other Omics Data”, JSM, Baltimore, MD. (August 2017). {Invited}
38. “Kernel Methods for Analysis of High-Dimensional Traits”, ICSA Applied Statistics Symposium, Chicago, IL. (June 2017). {Invited}
39. “Statistical Methods for Microbiome Data Analysis”, Genomics Seminar Series, Penn State University, University Park, PA. (May 2017). {Invited}
40. “Kernel Methods for Correlating Genetic Data with Microbiome and Other -Omic Outcomes”, TopMed Analysis Working Group, University of Washington, Seattle, WA. (March 2017).
41. “Kernel Machine Based Methods for Analyzing Genome Data from Different Sources”, ICSA International Conference, Shanghai, China. (December 2016). {Invited}
42. “Next Generation Kernel Methods for Now Generation Sequencing Data”, Japanese Society of Computational Statistics. (October 2016). {Invited Keynote}
43. “Statistical Methods for Microbiome Data Analysis”, Department of Lab Medicine, University of Washington, Seattle, WA. (October 2016). {Invited}
44. “Kernel Machine Methods for Genetic Studies with High-Dimensional and Complex Outcomes”, JSM, Chicago, IL. (August 2016). {Invited}
45. “Kernel Machine Methods for Community Level Analysis of Microbiome Data”, ENAR, San Antonio, TX. (March 2016). {Invited}
46. “Kernel Machine Methods for Complex Genomic Data: Integrative Analysis and Accommodating Interactions”, Department of Biomedical Informatics, CHOP/U. Penn., Philadelphia, PA. (October 2015). {Invited}
47. “Next-Generation Statistical Approaches for Now-Generation Sequencing Studies”, Non-Clinical Biostatistics Conference, Philadelphia, PA. (October 2015). {Invited}
48. “Semi-Nonparametric Interaction Analysis via Brownian Distance Covariance”, JSM, Seattle, WA. (August 2015). {Topic Contributed}
49. “Joint Analysis of Genomic Data from Different Sources using Kernel Machine Regression with Multiple Kernels”, ICSA Applied Statistics Symposium, Fort Collins, CO. (June 2015). {Invited}
50. “Flexible Testing Approaches for Microbiome and Multi-omics Analysis”, Department of Biostatistics, University of Washington, Seattle, WA. (October 2014). {Invited}
51. Discussant, “Statistical Methods for Genomics and High Dimensional Data”, JSM, Boston, MA. (August 2014). {Invited Discussant}
52. “Analysis of Complex Genomic Data using Recent Advances in Statistical Science”, West Virginia Clinical and Translational Science Institute, University of West Virginia, Morgantown, WV. (July 2014). {Invited}
53. “Statistical Methods for Integrative Analysis of Different Types of Genomic Data”, IMS Asia Pacific Rim, Taipei, Taiwan. (July 2014). {Invited}
54. “Kernel Machine Methods for Gene-Gene Interaction Testing”, Workshop on Emerging Statistical Challenges and Methods For Analysis of Massive Genomic Data in Complex Human Disease Studies, Banff International Research Station, Banff, Canada. (June 2014). {Invited}
55. “Integrative Genomic Analysis and Detecting Epistasis using Flexible Statistical Methods”, Department of Statistics, Oregon State University, Corvallis, OR. (April 2014). {Invited}
56. “Kernel Machine Methods for Gene-Gene Interaction Testing and for Integrative Genomic Analysis”, Cardiovascular Health Research Unit, University of Washington, Seattle, WA. (January 2014). {Invited}

57. "Normalization and Related Issues in Analysis of Methylation Data", National Institute of Environmental Health Sciences, Research Triangle Park, NC. (June 2013). {Invited}
58. "Inferential Strategies for High Dimensional -omics Data", Division of Public Health Sciences, Fred Hutchinson Cancer Research Institute, Seattle, WA. (March 2013). {Invited}
59. "Flexible Methods for Testing Interactions in Genetic Association Studies", EITC-Bio 2012: Recent Advances in Biomedical Research, Princeton, NJ. (October 2012). {Invited}
60. "Kernel Machine Based Methods for Gene-Environment Interaction Testing In Genetic Association Studies", International Conference on Advances in Interdisciplinary Statistics and Combinatorics, Greensboro, NC. (October 2012). {Invited}
61. "Resampling Based Variable Selection Methods for Next Generation Sequencing Studies", JSM, San Diego, CA. (August 2012). {Topic Contributed}
62. "Estimation and Testing in Genetic Association Studies under the Additive Kernel Machine Model", IMS Asia Pacific Rim, Tsukuba, Japan. (July 2012). {Invited}
63. "Kernel Machine Methods for Assessing Interactions in Genomic Studies", ICSA Applied Statistics Symposium, Boston, MA. (June 2012). {Invited}
64. "Flexible Modeling and Powerful Testing of Interaction in Genome-, Epigenome and Environment-Wide Association Studies in WHI", Annual WHI Investigator Meeting, Washington D.C. (May 2012). {Invited}
65. "Kernel Machine Based Testing of Rare Variant by Environment Interactions", ENAR, Washington D.C. (April 2012). {Invited}
66. "Statistical Issues in Genome Wide Methylation Profiling Studies", Department of Epidemiology and Biostatistics, University of South Carolina, Charleston, SC. (March 2012). {Invited}
67. "Kernel Machine Based Testing of Rare Variant by Environment Interactions", Biostatistics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC. (February 2012). {Invited}
68. "Rare Variant Analysis Using the Multi-Kernel Sequence Kernel Association Test", Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC. (February 2012). {Invited}
69. "Rare Variant Analysis Using the Multi-Kernel Sequence Kernel Association Test", Research Triangle Statistical Genetics Conference, Research Triangle Park, NC. (October 2011). {Invited}
70. "Rare Variant Analysis Using the Multi-Kernel Sequence Kernel Association Test", Statistical Analyses for Next Generation Sequencing, University of Alabama at Birmingham, Birmingham, AL. (September 2011). {Abstract Selected for Oral Presentation}
71. "Statistical Issues in the Analysis of Genome Wide Methylation Data", Workshop on Epigenetic Approaches to Studying Prenatal Influences on Childhood Health and Disease, Oslo, Norway. (September 2011). {Invited}
72. "Analysis of High-throughput Sequencing Data via the Sequence Kernel Association Test", Emerging Information and Technology Conferences (EITC-2011), Chicago, IL. (July 2011). {Invited}
73. "Kernel Machine Tests for Rare Genetic Variants in Sequencing Studies", ICSA, New York, NY. (June 2011). {Contributed}
74. "Kernel Machine Based Analysis of High-throughput Sequencing Data", Division of Public Health Sciences, Fred Hutchinson Cancer Research Institute, Seattle, WA. (April 2011). {Invited}
75. "Rare Variant Association Testing Using the Sequence Kernel Association Test (SKAT)", Statistical Genetics Group, GlaxoSmithKline, Research Triangle Park, NC. (December 2010). {Invited}
76. "Rare Variant Association Testing Using the Sequence Kernel Association Test (SKAT)", Department of Statistics, NC State University, Raleigh, NC. (November 2010). {Invited}

77. "Variable Selection in the Kernel Machine Framework via the Garrote Kernel Selector", JSM, Vancouver, BC Canada. (August 2010). {Invited}
78. "Kernel Machine Based Analysis of Genetic Association Studies", First Joint Biostatistics Symposium, Beijing, China. (July 2010). {Invited}
79. "Variable Selection in the Kernel Machine Framework via the Garrote Kernel Selector", Statistical Analysis of Complex Data, Kunming, Yunnan China. (July 2010). {Contributed}
80. "Kernel Machine Methods for the Analysis of Large Scale Genetic Association Studies", ENAR, New Orleans, LA. (March 2010). {Contributed}
81. "Powerful SNP Set Analysis for Case-Control Genome Wide Association Studies", Conference on Human Genetic Variation, Health and Disease: New Knowledge, New Quantitative Challenges. (November 2009). {Poster}
82. "Kernel Machine Approaches for the Analysis of Genome Wide Association Studies", Emerging Information and Technology Conferences (EITC-2009), Cambridge, MA. (August 2009). {Invited}
83. "Kernel Based Variable Selection via the Garrote Kernel Selector", ICSA Applied Statistics Symposium, San Francisco, CA. (June 2009). {Invited}
84. "Variable Selection in the Kernel Machine Framework", ENAR, San Antonio, TX. (March 2009). {Contributed}
85. "Regularization and Variable Selection in Classification and Regression Problems using Omics Data", Department of Statistics, Cornell University, Ithaca, NY. (March 2009). {Invited}
86. "Regularization and Variable Selection in Classification and Regression Problems using Omics Data", Division of Public Health Sciences, Fred Hutchinson Cancer Research Institute, Seattle, WA. (February 2009). {Invited}
87. "Inferential Strategies for High-dimensional 'Omics' Data", Department of Biostatistics, University of North Carolina - Chapel Hill, Chapel Hill, NC. (February 2009). {Invited}
88. "Inferential Strategies for High-dimensional 'Omics' Data", Department of Epidemiology and Biostatistics, University of Maryland, College Park, MD. (February 2009). {Invited}
89. "Regularization and Variable Selection in Classification and Regression Problems using Omics Data", Department of Statistics, University of Virginia, Charlottesville, VA. (February 2009). {Invited}
90. "Regularization and Variable Selection in Predictive Modeling using Omics Data", Department of Statistics, North Carolina State University, Raleigh, NC. (February 2009). {Invited}
91. "Regularization and Variable Selection in Predictive Modelling using Omics Data", Department of Biostatistics, Johns Hopkins University, Baltimore, MD. (January 2009). {Invited}
92. "Inferential Strategies for High-dimensional 'Omics' Data", Department of Biostatistics, Yale University, New Haven, CT. (January 2009). {Invited}
93. "Regularization and Variable Selection in Predictive Modelling using Omics Data", Department of Statistics, U.C. Davis, Davis, CA. (January 2009). {Invited}
94. "Genomic Feature Based Analysis of Genome Wide Association Studies via Logistic Kernel Machines", Racebrook Environmental Statistics Symposium, Sheffield, MA. (November 2008).
95. "Identification of Genes and Gene Pathways Associated with Metal Particulate Exposure using Sparse Linear Discriminant Analysis", GENI, Boston, MA. (October 2008). {Contributed Poster}
96. "Methods for Detection of Prostate Cancer Auto-antibody Response from High Density Self-Assembling Protein Microarrays", Harvard Institute of Proteomics, HMS, Cambridge, MA. (September 2008). {Invited}
97. "Sparse Linear Discriminant Analysis for Classification and Testing Gene Pathways", JSM, Denver, CO. (August 2008). {Topic Contributed Session for ASA Stat. Computing/Graphics Student Paper Competition Winners}

98. "Use of Sparse Linear Discriminant Analysis in Testing Gene Pathways", ICSA Applied Statistics Symposium, Piscataway, NJ. (June 2008). {Invited}
99. "A Parametric Permutation Test for Regression Coefficients in LASSO Regularized Regression for High Dimensional Data", ENAR, Arlington, VA. (March 2008). {Contributed}
100. "Use of Variable Selection in Testing the Significance of Gene Sets", ENAR, Atlanta, GA. (March 2007). {Contributed}
101. "Prior Biological Knowledge Based Approach for Analysis of Genome-Wide Expression Profiles", Racebrook Environmental Statistics Symposium, Sheffield, MA. (November 2006). – Joint Presentation with Dr. Zhaoxi (Mike) Wang

INTRAMURAL TALKS & PRESENTATIONS

1. "Keeping your Microbiome Studies out of the Toilet", Public Health Sciences Faculty Meeting, Fred Hutchinson Cancer Research Center, Seattle, WA. (March 2022).
2. "Microbiome in SWOG Studies", SWOG Statistical Center All-Staff Meeting, SWOG SDMC, Seattle, WA. (November 2021).
3. "Integrative Analysis of Microbiome and Other Omic Data Types", Fred Hutch/CMiST Microbiome Research Initiative Symposium, Fred Hutchinson Cancer Research Center, Seattle, WA. (April 2020).
4. "Fundamentals of Genomics in Cancer Research", Cancer Research and Biostatistics, Seattle, WA. (July 2019).
5. "Statistical Approaches for Paired and Longitudinal Microbiome Association Studies", Symposium on Colorectal Cancer and the Gut Microbiome, Fred Hutchinson Cancer Research Center, Seattle, WA. (March 2019).
6. "Community Level Analysis in Modern Microbiome Studies", Microbiome Research Initiative, Fred Hutchinson Cancer Research Center, Seattle, WA. (February 2018).
7. "Kernel Methods for Correlating Microbiome and Other -Omics Data". CCSG Biostat & Comp Bio Retreat, Fred Hutchinson Cancer Research Center, Seattle, WA. (October 2017).
8. "Next Generation Methods for Analyzing Next Generation Sequencing Data", Public Health Sciences All Staff Meeting, Fred Hutchinson Cancer Research Center, Seattle, WA. (April 2017).
9. "Statistical Methods for Correlating Microbiome and Other Genomic Data Types", Clinical Research Division Seminar Series, Fred Hutchinson Cancer Research Center, Seattle, WA. (March 2017).
10. Panelist, "No Fishing Allowed! P-Hacking and other Misuses of Statistical Methods in Research", UW/Hutch Summer Bioethics Series, Fred Hutchinson Cancer Research Center, Seattle, WA. (August 2016).
11. "Statistical (Non?) Issues in Microbiome Association Studies", VIDD Scientific Seminar Series, Fred Hutchinson Cancer Research Center, Seattle, WA. (March 2016).
12. "Statistical Issues in Analyzing DNA Methylation Data", Analysis Tools and Methods for High-throughput Genomic Data (ATME) Affinity Group, Fred Hutchinson Cancer Research Center, Seattle, WA. (January 2015).
13. "Kernel Machine Methods for Analysis of -Omics Data from Different Sources", Computational Biology Program Seminar Series, Fred Hutchinson Cancer Research Center, Seattle, WA. (August 2014).
14. "SKAT for Scat: Analysis of Fecal Microbiome Profiles", UW Biostatistics Department Retreat, Leavenworth, WA. (September 2013).
15. "Statistical Issues in the Analysis of Genome Wide Methylation Data", NC TRaCS Biostatistics Grand Rounds, UNC, Chapel Hill, NC. (April 2013).

14 WON BEST FACULTY PRESENTATION (VOTED BY STUDENTS)

16. “Career Development for Students and How to Properly do Crap Research”, Biostatistics Students Association Seminar, UNC, Chapel Hill, NC. (March 2013).
17. “Statistical Methods for the Analysis of Sequencing Data”, NC TRaCS Biostatistics Grand Rounds, UNC, Chapel Hill, NC. (January 2012).
18. “Environmentally Friendly Approaches for the Analysis of Genomic Data”, Department of Biostatistics Colloquium, UNC, Chapel Hill, NC. (January 2012).
19. “Analysis of Genome Wide Expression Profile Experiments”, NC TRaCS Biostatistics Seminar Series, UNC, Chapel Hill, NC. (September 2010).
20. “A Direct and Powerful Approach for Testing Rare Variants”, Mini-Symposium on Next-Generation Sequencing and Analysis, UNC, Chapel Hill, NC. (May 2010).
21. “Prior biological knowledge based approaches for the analysis of genome-wide expression profiles using gene sets and pathways”, Lineberger Comprehensive Cancer Center, UNC, Chapel Hill, NC. (March 2010).
22. “HD-TVs: An overview of High Dimensional Testing and Variable Selection”, (First) Biostatistics Student Invited Seminar, UNC, Chapel Hill, NC. (January 2010).
23. “Gene Set Analysis using Sparse Linear Discriminant Analysis”, NC TRaCS Biostatistics Grand Rounds, UNC, Chapel Hill, NC. (November 2009).
24. “Sparse Linear Discriminant Analysis for Simultaneous Testing for the Significance of a Gene Set/Pathway and Gene Selection”, Correlated and High-Dimensional Data Seminar Series, HSPH, Boston, MA. (October 2008).
25. “Sparse Linear Discriminant Analysis for Testing Differential Gene Pathway Activity Induced by Metal Particulate Exposure”, Environmental Statistics Seminar Series, HSPH, Boston, MA. (May 2008).
26. “An Overview of Variable Selection Procedures for High Dimensional Data”, Correlated and High-Dimensional Data Seminar Series, HSPH, Boston, MA. (October 2006). – Joint Presentation with Dr. Tianxi Cai

STUDENT
ADVISING

Doctoral Dissertations Directed

1. Jennifer J. Clark, Ph.D., Biostatistics (UNC). 2009 – 2013.
Dissertation: *Estimation and Hypothesis Testing with Additive Kernel Machines for High-Dimensional Data*
Placement/Current Position: Mathematical Statistician, Food and Drug Administration
2. Eugene Urrutia, Ph.D., Biostatistics (UNC). 2010 – 2013.
Dissertation: *Statistical Analysis of Genetic Sequencing and Rare Variant Association Studies*
His dissertation won the *2011 American Society of Human Genetics Trainee Research Semifinalist Award*.
Placement: Post-Doctoral Fellow, UNC
Current Position: Senior Data Scientist, Hill-Rom
3. Ni Zhao, Ph.D., Biostatistics (UNC). 2012 – 2013 (Student) 2014–2016 (Post Doc).
Dissertation: *Kernel Machine Methods for Analysis of Genomic Data from Different Sources*
Placement: Assistant Professor, The Johns Hopkins University
Current Position: Associate Professor, The Johns Hopkins University
4. Thomas G. Stewart, Ph.D., Biostatistics (UNC). 2011 – 2015.
Dissertation: *Statistical Learning with Missing Data*
His dissertation won the *2014 ENAR Poster Competition* and the *2015 ENAR Distinguished Student Paper Award*.
Placement: Assistant Professor, Vanderbilt University

Current Position: Associate Professor, University of Virginia

5. Sebastian Teran Hidalgo, Ph.D., Biostatistics (UNC). 2012 – 2016.
Dissertation: *Applications of Independence Statistics to Goodness-of-Fit, Clustering of Variables and Multivariate Change-Point Estimation*
Placement: Post-Doctoral Fellow, Yale University
Current Position: Data Scientist, Delfi Diagnostics
6. Anna Plantinga, Ph.D., Biostatistics (UW). 2015 – 2018.
Dissertation: *Statistical Methods for Analysis of Microbiome Data*
Her dissertation work was awarded an *NIH F31 grant* (received nearly perfect score) and the *Thomas R. Fleming Excellence in Biostatistics Award*.
Placement/Current Position: Assistant Professor, Williams College
7. Nanxun Ma, Ph.D., Biostatistics (UW). 2017 – 2021.
Dissertation: *Kernel Methods for Data Integration in Microbiome-Genomics Studies*
Placement/Current Position: Data Scientist, Google
8. Angela Zhang, Ph.D., Biostatistics (UW). 2019 – 2022.
Dissertation: *Statistical Tools for the Multiomics Analysis of Microbiome Data*
Placement/Current Position: Illumina
9. Amarise Little, Ph.D., Biostatistics (UW). 2021 – 2023 (Student) 2023– Present (Post Doc).
Dissertation: *Generalization of Kernel Machine Methods for Association Testing of Multi-Omics Data*
10. Hongjiao (Pearl) Liu, Ph.D., Biostatistics (UW). 2020 – 2023.
Dissertation: *Statistical Methods for Association Analysis of Microbiome Data*
Placement/Current Position: Analysis Group

Post Doctoral Fellows Mentored:

1. Ni Zhao, 2014 – 2016. (FHCRC)
2. Xiang Zhan, 2015 – 2017. (FHCRC)
His post-doc work won a *2017 ASA Biometrics Section Paper Award*.
Placement: Assistant Professor, Penn State University
Current Position: Associate Professor, Peking University
3. Wodan Ling, 2019 – 2022. (FHCC)
Her post-doc work won the *2023 ASA W.J. Youden Award in Interlaborator Testing*.
Placement/Current Position: Assistant Professor, Cornell University
4. Hoseung Song, 2021 – Present. (FHCC)
5. Amarise Little, 2023 – Present. (FHCC)
6. Sarah Samorodnitsky, 2023 – Present. (FHCC)

Doctoral Scientific Staff Mentored:

1. Xing Hua, 2019 – 2022. Research Associate. (FHCC)

Undergraduate Research Mentor

1. Elizabeth Simmons, 2009 – 2011. (UNC)

2. Teresa Lee, 2021. (UW)
3. Kai Marin, 2021. (UW)

Rotational and Project Research Assistants

1. Yatong Li, 2015 – 2016. (UW)
2. Yezi (Zora) Yang, 2017. (UW)
3. Yunhua Xiang, 2019 – 2021. (UW)
4. Walter Chen, 2022 – 2023. (UW)
5. Yichen Wang, 2023 – Present. (UW)

Biostatistics Doctoral Committee Membership

1. Qianchuan Chad He, 2011 – 2012. (Advisor: Danyu Lin, UNC)
2. Andrea Byrnes, 2011 – 2013. (Advisor: Yun Li, UNC)
3. Xiaoxi Liu, 2012 – 2014. (Advisor: Donglin Zeng, UNC)
4. Jennifer Kirk, 2015 – 2016. (Advisor: Tim Thornton, UW)
5. Yuxiang Xie, 2018 – 2019. (Advisor: Gary Chan, UW)
6. Yunhua Xiang, 2020 – 2021. (Advisor: Noah Simon, UW)
7. Anya Mikhaylova, 2016 – 2022. (Advisor: Tim Thornton, UW)

Non-Biostatistics Doctoral Committee Membership

1. Christopher Cabanski, 2010 – 2012. (Statistics, UNC; Advisor: Steve Marron)
2. Quaker Harmon, 2011 – 2012. (Epidemiology, UNC; Advisor: Stephanie Engel)
3. Jeremy Sabourin, 2011 – 2013. (Statistics, UNC; Advisor: Will Valdar & Andrew Nobel)
4. Leila Family, 2012 – 2014. (Epidemiology, UNC; Advisor: Andy Olshan)
5. Brionna Hair, 2012 – 2014. (Epidemiology, UNC; Advisor: Melissa Troester)
6. Kristina Jordahl, 2016 – 2018. (Epidemiology, UW; Advisor: Parveen Bhatti)
7. Xinwei Hua, 2018 – 2020. (Epidemiology, UW; Advisor: Polly Newcomb)
8. Cameron Haas, 2018 – 2021. (Epidemiology, UW; Advisor: Sara Lindstroem)
9. Dornell Pete, 2019 – 2022. (Epidemiology, UW; Advisor: Amanda Phipps)
10. Kayla Carter, 2020 – 2022. (Epidemiology, UW; Advisor: Jen Balkus)
11. Susana Lozano, 2022 – Present. (Epidemiology, UW; Advisor: Amanda Phipps & Holly Harris)
12. Hang Yin, 2023 – Present. (Public Health Genetics, UW; Advisor: Ulrike Peters)
13. Austin Sugar, 2023 – Present. (Epidemiology, UW; Advisor: Sara Lindstroem)

Masters Dissertations Chaired

1. Weijia Fu, 2018 – 2019. (Biostatistics, UW)

Masters Dissertation Committee Membership

1. Yichen Jia, 2016 – 2017. (Biostatistics, UW; Advisor: Tim Thornton)

Academic/Course Advisor

- 2010: Jung In Kim, Maria Reynolds
- 2012: Brian Barkley

SERVICE

Fred Hutchinson Cancer Center:

- Biostat. & Biomath. Faculty Search Committee: 2014 – 2015.
- Biostat. & Biomath. Seminar Organizer: 2015 – 2016.
- Computational Biology Faculty Search Committee: 2016 – 2017.
- VIDD Faculty Search Committee: 2021 – 2022.

Department of Biostatistics, University of Washington:

- Faculty Search Committee: 2022 – 2023.

Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill:

- Lineberger Data Warehouse and Biorepository Data Sharing Committee: 2012 – 2013.

Department of Biostatistics, The University of North Carolina at Chapel Hill:

- Awards Committee: 2009 – 2012.
- Masters Examinations Committee: 2010 – 2013.
- Information Technology Committee: 2012 – 2013.

Department of Biostatistics, Harvard University:

- Student Admissions Committee: 2004 – 2005
- Departmental Representative, Graduate Student Council: 2006 – 2009

Other Service:

- Tenure/Promotion Review Service: Duke University, Duke-NUS, Georgetown University, Emory, M.D. Anderson Cancer Center, Michigan, Moffitt Cancer Center, Mount Sinai, New York University, UCLA, Virginia Commonwealth University.
- Judge, ENAR Regional Advisory Board Poster Competition 2012

Community:

- Volunteer martial arts and self-defense instructor

TEACHING EXPERIENCE

Department of Biostatistics, University of Washington

Seattle, Washington USA

Instructor

Biostat/Stat 571: Advanced Regression Methods for Dependent Data

Winter 2021, Winter 2022, Winter 2023: Core course for Biostat Ph.D. students.

Guest Lecturer

2020-2021: Biostat 111; Biostat 572

Department of Biostatistics, The University of North Carolina at Chapel Hill

Chapel Hill, North Carolina USA

Instructor

Biostat 660: Probability and Statistical Inference I

Fall 2010, Fall 2012: Core course for MS, DrPH, and beginning Ph.D. students.

Biostat 735: Statistical Computing

Fall 2011: Ph.D. level elective course.

Guest Instructor

2011-2012: BIOS 740, BCB 725

2012-2013: *PHYI 703*

2013-2014: *BIOS 740*

Department of Biostatistics, Harvard University

Boston, Massachusetts USA

Derek Bok Center Departmental Teaching Fellow

Co-Head Teaching Fellow, Biostatistics Department

2008-2009: Developed, organized, and led training programs for graduate TAs.

Instructor

Summer Course in Probability

Summer 2008 & 2007: Required course for incoming biostatistics Ph.D. students.

Teaching Assistant

Biostat 277: Computational Biology (Advanced)

Fall 2008: Advanced elective course for biostatistics Ph.D. students.

Biostat 251: Statistical Inference II

Fall 2007: Advanced core course for biostatistics Ph.D. students.

Biostat 230: Probability Theory and Applications I

Fall 2006: Core course for first year biostatistics Ph.D. students.

Biostat 205: Biostatistics for Health Policy

Fall 2005: Introductory biostatistics course for social science MPH students.

Biostat 200: Introduction to Biostatistics

Fall 2004: Introductory biostatistics course for biomedical MPH students.

Other Teaching

Short Courses

- Instructor, “Analysis of Genome-wide Sequencing Association Studies”, JSM, Boston, MA. — August 2014.
- Instructor, “Human Association Mapping”, *Taipei 2015 Summer Institute in Statistical Genetics*, Taipei, Taiwan — August 2015.
- Instructor, “Rare Variant Analysis”, *NIGMS/NIH short course on Statistical Genetics and Genomics*, University of Alabama at Birmingham, Birmingham, AL. — July 2015, 2016, & 2017.
- Instructor, “Rare Variant Analysis”, *NHGRI/NIH short course on Next Generation Sequencing: Technology and Statistical Methods*, University of Alabama at Birmingham, Birmingham, AL. — December 2013, 2014, 2015, 2016, & 2017.
- Instructor, “Association Mapping: GWAS and Sequencing Data”, *Summer Institute in Statistical Genetics*, University of Washington, Seattle, WA — July 2015, 2016, 2017, 2018, 2019, 2020, & 2021.
- Instructor, “Microbiome Data Analysis”, *Summer Institute in Statistical And Modeling in Infectious Diseases*, University of Washington, Seattle, WA — July 2023.

Guest Instructor (Extramural)

- Guest Lecturer, BIOS 775, “Statistical Methods for Analysis of Large Scale Epigenetic Profiling Studies”, Department of Biostatistics, University of South Carolina, Columbia, SC. — March 2012.

TECHNICAL SKILLS

- Programming Languages: C, C++(STL), JAVA, PYTHON, PASCAL, BASIC, knowledge of LISP and PERL.
- Software: R/S-PLUS, MATLAB, PV-WAVE/IDL, SAS, STATA.
- Operating Systems: Unix and Windows.
- Languages: Mandarin Chinese, knowledge of Spanish.