Last update: Aug 24, 2025

1. Biographical Information

Wei Sun

Fred Hutchinson Cancer Center
Biostatistics Program
1100 Fairview Ave N, M3-B827, Seattle, WA 98109-1024

Phone: (206) 667-3188
Fax: (206) 667-7004
E-mail: wsun@fredhutch.org

2. Education

Peking University, Beijing, China, BS, Statistics, 06/2002
University of California at Los Angeles, Los Angeles, CA, MS, Statistics, 06/2004
University of California at Los Angeles, Los Angeles, CA, PhD, Statistics, 06/2007

3. Licensure Not applicable

4. Professional Positions

Assistant Professor, Department of Biostatistics, Genetics, University of North Carolina
July 2007 – May 2013

Associate Professor, Department of Biostatistics, Genetics, University of North Carolina
June 2013 – August 2016

Associate Member, Public Health Division, Fred Hutchison Cancer Center
September 2015 – December 2018

Professor, Public Health Division, Fred Hutchison Cancer Center
January 2019 – Present

Affiliated Associate Professor, Department of Biostatistics, University of North Carolina
September 2016 – August 2019

Affiliated Professor, Department of Biostatistics, University of North Carolina
September 2019 – Present

Affiliated Associate Professor, Department of Biostatistics, University of Washington
April 2017 – June 2020

Affiliated Professor, Department of Biostatistics, University of Washington
July 2020 – Present

5. Honors, Awards, Scholarhips

ASA Fellow, 2017
Junior Faculty Development Awards, UNC-Chapel Hill, 2010
University Research Council Award, UNC-Chapel Hill, 2008
University of California Fellowship, UCLA, 2002-2007

6. Professional Activities

Associate Editor
    JASA, Application and Case Studies, 2012-2022
    Statistics in Biosciences, 2023-

Service in professional organization:
    ENAR 2016 Program Chair, ENAR Regional Advisory Board for 2016-2018
    JSM 2013 Program Chair for Biometrics Section

Publication Referee:

Annual of Statistics / Annual of Applied Statistics / Bioinformatics / Biometrics /Biostatistics / BMC Bioinformatics / BMC Genomics / Genetic Epidemiology / Genetics / Genetics Research / International Journal of Biostatistics / JASA / Nucleic Acid Research (NAR) / Nature Genetics / Nature Methods / Nature Communication / Nature Biotechnology / PLOS Computational Biology / PLoS ONE / PNAS/ SAGMB / Society for Industrial and Applied Mathematics Book Series / Statistics and Its Interface / Statistics in Medicine / Statistics Surveys / Transactions on Computational Biology and Bioinformatics (IEEE)

Grant Reviewer:

Mail reviewer for NIH RC4 applications - Recovery Act Limited Competition

NIH Epidemiology of Cancer [EPIC] study section, June 2013

NCI Omnibus "Biomarkers" ZCA1 SRLB-5, December 2013

NIH Special Emphasis Panel, March 2015

NCI Special Emphasis Panel, April 2016

NIH Special Emphasis Panel, March 2017

NCATS Special Emphasis Panel, Feb 2019

NCATS Special Emphasis Panel, Feb 2020

NIGMS MIRA grant review special panel, March 2021

NHGRI Special Emphasis Panel, April 2023

NIH BDMA study section, April 2025, June 2025

External Reviewer:

NIEHS, Biostatistics and Computational Biology Branch, 2021

Memberships in Professional Societies:

2007-present: American Statistical Association
2008-present: International Biometric Society
2008-present: International Chinese Statistical Association
2018-present: American Association of Cancer Research

7. Bibliography

A. Refereed Research Articles

Peer Reviewed Publications (Statistical Methods)

1. Molstad AJ, Cai Y, Reiner AP, Kooperberg C, **Sun W**, Hsu L. (2024)
   Heterogeneity-aware integrative regression for ancestry-specific association studies.
   Biometrics. 2024 Oct 3;80(4):ujae109.

2. Li Z, Li R, Ganan-Gomez I, Abbas HA, Garcia-Manero G, **Sun W**. (2024)
   Accurate identification of locally aneuploid cells by incorporating cytogenetic information in single cell data analysis.
   Sci Rep. 2024 Oct 15;14(1):24152.

3. Lyu Y, Wu C, **Sun W**, Li Z. (2024)
   Regional analysis to delineate intrasample heterogeneity with RegionalST.
   Bioinformatics. 2024 Mar 29;40(4):btae186.

4. Liu, S., Bradley, P., & **Sun, W**. (2023).
   Neural network models for sequence-based TCR and HLA association prediction.
   PLOS Computational Biology, 19(11), e1011664.

5. Zhou, LY, Zou, F, and **Sun, W**. (2023)
   Prioritizing candidate peptides for cancer vaccines by PEPPRMINT: a statistical model to predict peptide presentation by HLA-I proteins.
   Biometrics. 2023 Sep;79(3):2664-2676.

6. Little P, Liu S, Zhabotynsky V, Li Y, Lin DY, **Sun W**. (2023)
   A computational method for cell type-specific expression quantitative trait loci mapping using bulk RNA-seq data.
   Nature Communications. 2023 May 25;14(1):3030.

7. Sheikh MT, Chen MH, Gelfond JA, **Sun W**, Ibrahim JG. (2023)
   New C-indices for assessing importance of longitudinal biomarkers in fitting competing risks survival data in the presence of partially masked causes.
   Stat Med. 2023 Apr 30;42(9):1308-1322

8. Little P, Hsu L, **Sun W**. (2022)

Associating somatic mutation with clinical outcomes through kernel regression and optimal transport.
Biometrics. 2023 Sep;79(3):2705-2718.

9. Miao Z, Kong W, Vinayak RK, **Sun W**, Han F. (2022)
   Fisher-Pitman permutation tests based on nonparametric Poisson mixtures with application to single cell genomics.
   Journal of the American Statistical Association. 119(545), 394–406.

10. Liu M, Goo J, Liu Y, **Sun W**, Wu MC, Hsu L, He Q. (2022)
    TCR-L: an analysis tool for evaluating the association between the T-cell receptor repertoire and clinical phenotypes.
    BMC bioinformatics. 2022 Dec;23(1):1-6.

11. Zhabotynsky V, Huang L, Little P, Hu YJ, Pardo-Manuel de Villena F, Zou F, **Sun W**. (2022)
    eQTL mapping using allele-specific count data is computationally feasible, powerful, and provides individual-specific estimates of genetic effects.
    PLoS Genet. 2022 Mar 14;18(3):e1010076.

12. Zhang M, Liu S, Miao Z, Han F, Gottardo R, **Sun W.** (2022)
    IDEAS: individual level differential expression analysis for single-cell RNA-seq data. *Genome Biol.* 2022 Jan 24;23(1):33.

13. Liu M, Goo J, Liu Y, **Sun W**, Wu M, Hsu L, He Q (2022)
    TCR-L: an analysis tool for evaluating the association between the T-cell receptor repertoire and clinical phenotypes
    BMC Genomics, in press

14. Heiling, H. M., Wilson, D. R., Rashid, N. U., Sun, W., & Ibrahim, J. G. (2021).
    Estimating cell type composition using isoform expression one gene at a time.
    Biometrics. 2023 Jun;79(2):854-865. Epub 2021 Dec 29

15. Liu, J., Wang, H., **Sun, W.,** & Liu, Y. (2021).
    Prioritizing Autism Risk Genes using Personalized Graphical Models Estimated from Single Cell RNA-seq Data.
    J Am Stat Assoc. 2022;117(537):38-51. Epub 2021 Jul 21.

16. Molstad, A. J., **Sun, W.,** & Hsu, L. (2021).
    A covariance-enhanced approach to multi-tissue joint eQTL mapping with application to transcriptome-wide association studies.
    The annals of applied statistics, 15(2), 998.

17. Jin C, Chen M, Lin DY, **Sun W** (2021)
    Cell-type-aware analysis of RNA-seq data.
    Nature Computational Science, 1:253–261

18. Zhang H, Cai R, Dai J, **Sun W** (2021).
    EMeth: An EM algorithm for cell type decomposition based on DNA methylation data.
    Sci Rep. 11(1):5717

19. Yang J, Liu Y, Liu Y, **Sun W** (2021).
    Model free estimation of graphical model using gene expression data.
    The Annals of Applied Statistics. 15(1):194-207.

20. Molstad AJ, Hsu L, **Sun W** (2021)
    Gaussian process regression for survival time prediction with genome-wide gene expression.
    Biostatistics, 22(1) 164-180.

21. Wilson DR, Ibrahim JG, **Sun W** (2020)
    Mapping Tumor-Specific Expression QTLs In Impure Tumor Samples.
    Journal of the American Statistical Association, 115(529), 79-89

22. Dai JY, Wang X, Wang B, **Sun W**, Jordahl KM, Kolb S, Nyame YA, Wright JL, Ostrander EA, Feng Z, Stanford JL.
    DNA methylation and cis-regulation of gene expression by prostate cancer risk SNPs.
    PLOS Genetics. 2020 Mar 30;16(3):e1008667.

23. Wilson DR, Ibrahim JG, **Sun W** (2020)
    Mapping Tumor-Specific Expression QTLs In Impure Tumor Samples.
    Journal of the American Statistical Association, 115(529), 79-89

24. Little P, Lin DY, **Sun W** (2019)
    Associating somatic mutations to clinical outcomes: a pan-cancer study of survival time.
    Genome medicine, 11(1):37.

25. **Sun W**, Jin C, Gelfond JA, Chen MH, Ibrahim JG (2019)
    Joint analysis of single-cell and bulk tissue sequencing data to infer intratumor heterogeneity.
    Biometrics, doi: 10.1111/biom.13198.

26. Liu J, **Sun W,** Liu Y (2019)
    Joint Skeleton Estimation of Multiple Directed Acyclic Graphs for Heterogeneous Population.
    Biometrics, 75 (1), 36-47.

27. Zhabotynsky V, Inoue K, Magnuson T, Mauro Calabrese J, **Sun W** (2019)
    A statistical method for joint estimation of cis-eQTLs and parent-of-origin effects under family trio design.
    Biometrics, 75 (3), 864-874.

28. Liu Y, **Sun W**, Reiner A, Kooperberg C, He Q (2019)
    Statistical inference of genetic pathway analysis in high dimensions.
    Biometrika, 106 (3). 651

29. Liu Y, He Q, **Sun W** (2018)
    Association analysis using somatic mutations. *PLOS Genetics*, 14(11), e1007746.

30. Ha MJ, **Sun W** (2018)
    Estimation of High-dimensional Directed Acyclic Graphs with Surrogate Intervention, Biostatistics, in press

31. **Sun W**, Bunn P, Jin C, Little P, Zhabotynsky V, Perou CM, Hayes DN, Chen M, Lin DY (2018).
    The association between copy number aberration, DNA methylation and gene expression in tumor samples.
    Nucleic acids research, 46(6), pp.3009-3018.

32. Wang W, **Sun W**, Wang W, Szatkiewicz J.
    A randomized approach to speed up the analysis of large-scale read-count data in the application of CNV detection.
    BMC bioinformatics. 2018 Dec;19(1):74.

33. Liu Y, Xiong S, **Sun W**, Zou F (2017)
    Joint Analysis of Strain and Parent-of-Origin Effects for Recombinant Inbred Intercrosses Generated from Multiparent Populations with the Collaborative Cross as an Example.
    G3: Genes, Genomes, Genetics. g3-300483.

34. Lee S, **Sun W**, Wright FA, Zou F (2017).
    An improved and explicit surrogate variable analysis procedure by coefficient adjustment.
    Biometrika, 104(2), 303-316.

35. Silva GO, Siegel MB, Mose LE, Parker JS, **Sun W**, Perou CM, Chen M (2017).
    SynthEx: a synthetic-normal-based DNA sequencing tool for copy number alteration detection and tumor heterogeneity profiling.
    Genome Biology, 18(1), 66.

36. Sylvia Richardson, George C. Tseng, and **Wei Sun** (2016) Statistical Methods in Integrative Genomics, *Annual Review of Statistics and Its Application*, 3, 181-209

37. Rashid, NU, **Sun, W***, and Ibrahim, JG (2016). A Statistical Model to Assess (Allele-Specific) Associations Between Gene Expression and Epigenetic Marks Using Sequencing Data, *Annals of Applied Statistics*, 10(4), 2254-2273

38. Chen TH, **Sun W***, and Fine JP (2016).
    Designing penalties for high dimensional regression.
    Electronic Journal of Statistics, 10(2), 2312-2328

39. Chen TH and **Sun W*** (2016).
    Prediction Of Cancer Drug Sensitivity Using High-Dimensional Genomic Features.
    Biostatistics, 18 (1): 1-14

40. Ha MJ, **Sun W***, & Xie, J (2016).

PenPC: A Two-step Approach to Estimate the Skeletons of High Dimensional Directed Acyclic Graphs.
Biometrics, 72(1), 146-155.

41. Zhang Y, Zhou H, Zhou J, and **Sun W**. (2016)
Regression models for multivariate count data, *Journal of Computational Graphical Statistics, in press*, http://dx.doi.org/10.1080/10618600.2016.1154063

42. Gelfond JA, Ibrahim JG, Chen MH, **Sun W**, Lewis K, Kinahan S, Hibbs M, Buffenstein R. (2015)
Homology cluster differential expression analysis for interspecies mRNA-Seq experiments. *Stat Appl Genet Mol Biol.* 14(6):507-16.

43. Wang, W., Wang, W., **Sun, W.**, Crowley, J. J., & Szatkiewicz, J. P. (2015).
Allele-specific copy-number discovery from whole-genome and whole-exome sequencing.
Nucleic acids research, *43*(14), e90-e90.

44. Hu YJ, **Sun W**, Tzeng JY, and Perou CM (2015)
Proper Use of Allele-Specific Expression Improves Statistical Power for *cis*-eQTL Mapping with RNA-Seq Data, *Journal of the American Statistical Association*, 110(511) 962-974

45. **Sun, W.**, Liu, Y., Crowley, J. J., Chen, T. H., Zhou, H., Chu, H., ... & de Villena, F. P. M. (2015).
IsoDOT Detects Differential RNA-isoform Usage with respect to a Categorical or Continuous Covariate with High Sensitivity and Specificity.
Journal of the American Statistical Association, 110(511) 975-986

46. Huang KC, **Sun W**, Wu Y, Chen M, Mohlke KL, Lange LA, Li Y. (2014)
Association studies with imputed variants using expectation-maximization likelihood-ratio tests.
PLoS One. 10;9(11):e110679.

47. Lin JA, Zhu H, Mihye A, **Sun W**, Ibrahim JG (2014)
Alzheimer's Neuroimaging Initiative. Functional-mixed effects models for candidate genetic mapping in imaging genetic studies.
Genet Epidemiol. 38(8):680-91

48. Wilkerson, M. D., Cabanski, C. R., **Sun, W.**, Hoadley, K. A., Walter, V., Mose, L. E., ... & Hayes, D. N. (2014).
Integrated RNA and DNA sequencing improves mutation detection in low purity tumors.
Nucleic acids research, 42(13), e107-e107.

49. Ha, MJ, and **Sun, W*** (2014)
Partial correlation matrix estimation using ridge penalty followed by thresholding and reestimation
Biometrics, 70(3), 762-770

50. Zou F, **Sun W**, Crowley JJ, Zhabotynsky V, Sullivan PF, Pardo-Manuel de Villena

FF (2014)
A Novel Statistical Approach for Jointly Analyzing RNA-seq Data from F1 Reciprocal Crosses and Inbred Lines
Genetics, 197(1), 389-399

51. Hu, YJ, Lin, DY, **Sun, W**, and Zeng, D (2014)
A Likelihood-Based Framework for Association Analysis of Allele-Specific Copy Numbers
Journal of the American Statistical Association, 109(508):1533-1545,

52. Rashid, NU, **Sun, W***, and Ibrahim, JG (2014).
Some Statistical Strategies for DAE-seq Data Analysis: Variable Selection and Modeling Dependencies among Observations.
Journal of the American Statistical Association, 109(505):78-94,

53. Szatkiewicz JP, Wang W, Sullivan PF, Wang W, **Sun W** (2013)
Improving detection of copy-number variation by simultaneous bias correction and read-depth segmentation.
Nucleic Acids Res. 1;41(3):1519-32. PMCID: PMC3561969

54. **Sun W**, Hu Y (2013)
eQTL Mapping Using RNA-seq Data.
Statistics in Bioscience, in 2013 May1;5(1):198-219. PMCID: PMC3650863

55. Lee W, Du Y, **Sun W**, Hayes D.N., Liu Y (2012)
Multiple Response Regression for Gaussian Mixture Models with Known Labels.
Statistical Analysis and Data Mining, 5(6), 493-508.

56. Zhang X, Huang S, **Sun W**, Wang W (2012)
Rapid and robust resampling-based multiple-testing correction with application in a genome-wide expression quantitative trait loci study.
Genetics. 2012 Apr;190(4):1511-20. PMCID: PMC3316660.

57. **Sun W** (2012)
A statistical framework for eQTL mapping using RNA-seq data.
Biometrics. 2012 Mar;68(1):1-11. Epub 2011 Aug 12. PMCID: PMC3218220.

58. **Sun W**, Li L (2012)
Multiple loci mapping via model-free variable selection.
Biometrics. 2012 Mar;68(1):12-22. Epub 2011 Aug 12. PMCID: PMC3218235.

59. Rashid NU, Giresi PG, Ibrahim JG, **Sun W***, Lieb JD* (2011)
ZINBA integrates local covariates with DNA-seq data to identify broad and narrow regions of enrichment, even within amplified genomic regions.
Genome Biology, 12(7):R67 (20 pages), PMCID: PMC3218829.
* co-corresponding author

60. Yu T, Peng H, **Sun W** (2011)
Incorporating Nonlinear Relationships in Microarray Missing Value Imputation.
IEEE/ACM Trans Comput Biol Bioinform. 2011 May-Jun;8(3):723-31.

61. Chu H, Nie L, Chen Y, Huang Y, **Sun W** (2010)
Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: Methods for the absolute risk difference and relative risk.
Stat Methods Med Res. 2010 Dec 21 (13 pages).

62. Van Loo P, Nordgard SH, Lingjærde OC, Russnes HG, Rye IH, **Sun W**, Weigman VJ, Marynen P, Zetterberg A, Naume B, Perou CM, Børresen-Dale AL, Kristensen VN (2010)
Allele-specific copy number analysis of tumors.
Proc Natl Acad Sci U S A. 2010 Sep 28;107(39):16910-5. PMCID: PMC2947907.

63. **Sun W**, Ibrahim JG, Zou F (2010)
Genome-wide multiple-loci mapping in experimental crosses by iterative adaptive penalized regression.
Genetics. 2010 May;185(1):349-59. Epub 2010 Feb 15. PMCID: PMC2870969.

64. Shen Y, **Sun W**, Li KC (2010)
Dynamically weighted clustering with noise set.
Bioinformatics. 2010 Feb 1;26(3):341-7. Epub 2009 Dec 9. PMCID: PMC2815660.

65. **Sun W** and Wright FA (2010)
A geometric interpretation of the permutation p-value and its application in eQTL studies.
Annals of Applied Statistics, 4(2), 1014-1033

66. **Sun W**, Wright FA, Tang Z, Nordgard SH, Van Loo P, Yu T, Kristensen VN, Perou
CM (2009)
Integrated study of copy number states and genotype calls using high-density SNP arrays. *Nucleic Acids Res*. 2009 Sep;37(16):5365-77. Epub 2009 Jul 6. PMCID: PMC2935461.

67. **Sun W**, Buck MJ, Patel M, Davis IJ (2009)
Improved ChIP-chip analysis by a mixture model approach.
BMC Bioinformatics. 2009 Jun 7;10:173 (13 pages). PMCID: PMC2700807

68. **Sun W***, Xie W*, Xu F, Grunstein M, Li KC (2009)
Dissecting nucleosome free regions by a segmental semi-Markov model.
PLoS One. 2009;4(3):e4721 (10 pages). Epub 2009 Mar 6. PMCID: PMC2648986.
* Co-first authors

69. Wu T*, **Sun W***, Yuan S, Chen CH, Li KC (2008)
A method for analyzing censored survival phenotype with gene expression data.
BMC Bioinformatics. 2008 Oct 6;9:417 (11 pages). PMCID: PMC2579309.
* Co-first authors

70. **Sun W***, Yuan S*, Li KC (2008)
Trait-trait dynamic interaction: 2D-trait eQTL mapping for genetic variation study.
BMC Genomics. 2008 May 23;9:242 (13 pages). PMCID: PMC2432080.

* Co-first authors

71. **Sun W**, Yu T, Li KC (2007)
Detection of eQTL modules mediated by activity levels of transcription factors.
Bioinformatics. 2007 Sep 1;23(17):2290-7.

72. Yu T, Ye H, **Sun W**, Li KC, Chen Z, Jacobs S, Bailey DK, Wong DT, Zhou X (2007)
A forward-backward fragment assembling algorithm for the identification of genomic amplification and deletion breakpoints using high-density single nucleotide polymorphism (SNP) array.
BMC Bioinformatics. 2007 May 3;8:145 (11 pages). PMCID: PMC1868765.

73. Yu T, **Sun W**, Yuan S, Li KC (2005)
Study of coordinative gene expression at the biological process level.
Bioinformatics, 2005 Sep 15;21(18):3651-7.

74. Li KC, Liu CT, **Sun W**, Yuan S, Yu T (2004)
A system for enhancing genome-wide coexpression dynamics study.
Proc Natl Acad Sci U S A. 2004 Nov 2;101(44):15561-6. PMCID: PMC524832.

Peer Reviewed Publications (Applied-Collaborative)

75. Cai Y, Johnson M, Haessler J, Molstad AJ, Hwang SJ, Joehanes R, Murabito JM, Tahir UA, Franceschini N, Gerszten RE, **Sun W**, Levy D, Raffield LM, Kooperberg C, Hsu L, Reiner AP.
Protein quantitative trait locus analysis in African American and non-Hispanic White individuals.
Genome Biol. 2025 Jul 10;26(1):200. doi: 10.1186/s13059-025-03671-x.

76. Adli M, Przybyla L, Burdett T, Burridge PW, Cacheiro P, Chang HY, Engreitz JM, Gilbert LA, Greenleaf WJ, Hsu L, Huangfu D, Hung LH, Kundaje A, Li S, Parkinson H, Qiu X, Robson P, Schürer SC, Shojaie A, Skarnes WC, Smedley D, Studer L, **Sun W**, Vidović D, Vierbuchen T, White BS, Yeung KY, Yue F, Zhou T; MorPhiC Consortium.
MorPhiC Consortium: towards functional characterization of all human genes.
Nature. 2025 Feb;638(8050):351-359. doi: 10.1038/s41586-024-08243-w.

77. Chalitsios CV, Markozannes G, Papagiannopoulos C, ... , **Sun W**, ..., Peters U, Phipps AI, Tsilidis KK (2025).
Waist Circumference, a Body Shape Index, and Molecular Subtypes of Colorectal Cancer: A Pooled Analysis of Four Cohort Studies.
Cancer Epidemiol Biomarkers Prev. 2025 Apr 3;34(4):568-577.

78. Adli M, Przybyla L, Burdett T, Burridge PW, Cacheiro P, Chang HY, Engreitz JM, Gilbert LA, Greenleaf WJ, Hsu L, Huangfu D, Hung LH, Kundaje A, Li S, Parkinson H, Qiu X, Robson P, Schürer SC, Shojaie A, Skarnes WC, Smedley D, Studer L, **Sun W**, Vidović D, Vierbuchen T, White BS, Yeung KY, Yue F, Zhou T (2025)
MorPhiC Consortium. MorPhiC Consortium: towards functional characterization of all human genes.

Nature. 2025 Feb;638(8050):351-359.

79. Aglago EK, Qu C, …, **Sun W**, …, Peters U. (2024)
    Folate intake and colorectal cancer risk according to genetic subtypes defined by targeted tumor sequencing.
    Am J Clin Nutr. 2024 Sep;120(3):664-673.

80. Thomas CE, ..., **Sun W**, ... Peters U, Phipps AI. (2024)
    Epidemiologic Factors in Relation to Colorectal Cancer Risk and Survival by Genotoxic Colibactin Mutational Signature.
    Cancer Epidemiol Biomarkers Prev. 2024 Apr 3;33(4):534-546

81. Ugai, T., Akimoto, N., Haruki, K., Harrison, T. A., Cao, Y., Qu, C., ..., **Sun W**, …, & Ogino, S. (2023).
    Prognostic role of detailed colorectal location and tumor molecular features: analyses of 13,101 colorectal cancer patients including 2994 early-onset cases.
    Journal of gastroenterology, 58(3), 229-245.

82. Lee, S. Y., Lee, D. H., **Sun, W**., Cervantes-Contreras, F., Basom, R. S., Wu, F., ... & Till, B. G. (2023). CD8+ chimeric antigen receptor T cells manufactured in absence of CD4+ cells exhibit hypofunctional phenotype. *Journal for Immunotherapy of Cancer*, 11(11).

83. Murphy N, Newton CC, Song M, Papadimitriou N, Hoffmeister M, Phipps AI, Harrison TA, Newcomb PA, …, **Sun W**, …, Peters U, Campbell PT (2023).
    Body mass index and molecular subtypes of colorectal cancer.
    JNCI: Journal of the National Cancer Institute. 2023 Feb;115(2):165-73.

84. Gelfond JA, Hernandez B, Goros M, Ibrahim JG, Chen MH, **Sun W**, Leach RJ, Kattan MW, Thompson IM, Ankerst DP, Liss M. (2022)
    Prediction of future risk of any and higher-grade prostate cancer based on the PLCO and SELECT trials.
    BMC urology. 2022 Dec;22(1):1-9.

85. Georgeson, P., Harrison, T. A., Pope, B. J., Zaidi, S. H., Qu, C., Steinfelder, R. S., ... **Sun W**, …, & Buchanan, D. D. (2022).
    Identifying colorectal cancer caused by biallelic MUTYH pathogenic variants using tumor mutational signatures.
    Nature Communications, 13(1), 1-12

86. Barfield, R., Qu, C., Steinfelder, R. S., Zeng, C., Harrison, T. A., Brezina, S., ..., **Sun W**, Zaidi SH & Peters, U. (2022).
    Association between germline variants and somatic mutations in colorectal cancer.
    Scientific reports, 12(1), 1-9.

87. Harlid S, … , Steinfelder RS, **Sun W**, Thibodeau SN, Toland AE, Ugai T, Um CY, Woods MO, Phipps AI, Harrison T, Peters U. (2022)
    Diabetes mellitus in relation to colorectal tumor molecular subtypes - a pooled analysis of more than 9,000 cases.
    Int J Cancer. doi: 10.1002/ijc.34015. Epub ahead of print.

88. Barfield R, Huyghe JR, …, Steinfelder RS, Sun W, Win AK, Zaidi SH, Campbell PT, Peters U, Hsu L. (2022)
    Genetic regulation of DNA methylation yields novel discoveries in GWAS of colorectal cancer.
    Cancer Epidemiol Biomarkers Prev. cebp.0724.2021. Epub ahead of print.

89. Borozan I, Zaidi SH, Harrison TA, Phipps AI, ..., **Sun W**, Hudson TJ, Ferretti V, Peters U. (2022)
    Molecular and Pathology Features of Colorectal Tumors and Patient Outcomes Are Associated with Fusobacterium nucleatum and Its Subspecies animalis.
    Cancer Epidemiol Biomarkers Prev. 31(1):210-220.

90. Labadie JD, Hua X, Harrison TA, Banbury BL, Huyghe JR, **Sun W**, Shi Q, Yothers G, Alberts SR, Sinicrope FA, Goldberg RM.
    Genetic Predictors of Severe Skin Toxicity in Patients with Stage III Colon Cancer Treated with Cetuximab: NCCTG N0147 (Alliance).
    Cancer Epidemiology and Prevention Biomarkers. 2021 Feb 1;30(2):404-11.

91. Zaidi SH, Harrison TA, Phipps AI, Steinfelder R, Trinh QM, Qu C, Banbury BL, Georgeson P, Grasso CS, Giannakis M, Adams JB, …, **Sun W**, Hudson TJ, Peters U.
    Landscape of somatic single nucleotide variants and indels in colorectal cancer and impact on survival.
    Nature communications. 2020 Jul 20;11(1):1-2.

92. Labadie JD, Harrison TA, Banbury B, Amtay EL, Bernd S, Brenner H, Buchanan DD, Campbell PT, Cao Y, Chan AT, … **Sun W**, …, Newcomb P.
    Postmenopausal Hormone Therapy and Colorectal Cancer Risk by Molecularly Defined Subtypes and Tumor Location.
    JNCI Cancer Spectrum. 2020 Aug;4(5).

93. Penney KL, Banbury BL, Bien S, Harrison TA, Hua X, Phipps AI, **Sun W**, Song M, Joshi AD, Alberts SR, Allegra CJ. (2019)
    Genetic Variant Associated with Survival of Patients with Stage II-III Colon Cancer.
    Clinical Gastroenterology and Hepatology. doi.org/10.1016/j.cgh.2019.11.046

94. Grasso CS, Giannakis M, … , Sun W, …, Gabriel SB, Lander ES, Garraway LA, Hudson TJ, Fuchs CS, Ribas A, Ogino S, Peters U. (2018)
    Genetic Mechanisms of Immune Evasion in Colorectal Cancer.
    Cancer Discov. 8(6):730-749.

95. Kim Y, Giusti-Rodriguez P, Crowley JJ, Bryois J, Nonneman RJ, Ryan AK, Quackenbush CR, Iglesias-Ussel MD, Lee PH, **Sun W**, de Villena FP, and PF Sullivan (2018)
    Comparative genomic evidence for the involvement of schizophrenia risk genes in antipsychotic effects.
    Molecular Psychiatry, 2018 Mar;23(3):708.

96. Mutlu E, Gao L, Collins LB, Walker NJ, Hartwell HJ, Olson JR, **Sun W**, Gold A, Ball LM, Swenberg JA. (2016)
Polychlorinated Biphenyls Induce Oxidative DNA Adducts in Female Sprague–Dawley Rats.
Chemical research in toxicology. 29(8):1335-44.

97. Sharma V, Collins LB, Chen TH, Herr N, Takeda S, **Sun W**, Swenberg JA, Nakamura J. (2016)
Oxidative stress at low levels can induce clustered DNA lesions leading to NHEJ mediated mutations.
Oncotarget. 2016 May 3;7(18):25377-90.

98. **Sun W***, Kechris K", Jacobson S*, …, O'Neal WK, Bowler RP (2016)
Common Genetic Polymorphisms Influence Blood Biomarker Measurements in COPD
PLoS Genetics, in press

99. Jansen R, Penninx BW, Madar V, Xia K, Milaneschi Y, Hottenga JJ, Hammerschlag
AR, Beekman A, van der Wee N, Smit JH, Brooks AI, Tischfield J, Posthuma D, Schoevers R, van Grootheest G, Willemsen G, de Geus EJ, Boomsma DI, Wright FA,
Zou F, Sun W, Sullivan PF. (2016)
Gene expression in major depressive disorder.
Mol Psychiatry. 2016 Mar;21(3):339-47.

100. O'Neal WK, Gallins P, Pace RG, Dang H, Wolf WE, Jones LC, Guo X, Zhou YH, Madar V, Huang J, Liang L, Moffatt MF, Cutting GR, Drumm ML, Rommens JM, Strug LJ, **Sun W**, Stonebraker JR, Wright FA, Knowles MR. (2015)
Gene Expression in Transformed Lymphocytes Reveals Variation in Endomembrane and HLA Pathways Modifying Cystic Fibrosis Pulmonary Phenotypes.
Am J Hum Genet. 96 (2), 318–328.

101. Crowley, J. J*., Zhabotynsky, V.*, **Sun, W.***, Huang, S., Pakatci, I. K., Kim, Y., ... & Pardo-Manuel, D. V. F. (2015).
Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance.
Nature Genetics, 47(4), 353.
* Co-first authors

102. Wright, F. A., Sullivan, P. F., Brooks, A. I., Zou, F., **Sun, W**., Xia, K., ... & Boomsma, D. I. (2014).
Heritability and genomics of gene expression in peripheral blood.
Nature Genetics, 46(5), 430-437.

103. Holley DW, Groh BS, Wozniak G, Donohoe DR, **Sun W**, Godfrey V, Bultman SJ. (2014)
The BRG1 chromatin remodeler regulates widespread changes in gene expression and cell proliferation during B cell activation.

J Cell Physiol. 2014 Jan;229(1):44-52.

104. Moeller BC, Recio L, Green A, **Sun W**, Wright FA, Bodnar WM, Swenberg JA (2013)
Biomarkers of Exposure and Effect in Human Lymphoblastoid TK6 Cells Following [13C2]-Acetaldehyde Exposure.
Toxicological Sciences, 2013 May;133(1):1-12 PMCID: PMC3627555

105. Calabrese JM, **Sun W**, Song L, Mugford J, Williams L, Yee D, Starmer J, Mieczkowski P, Crawford G, Magnuson T (2012)
Site-specific silencing of regulatory elements as a mechanism of X-inactivation.
Cell, 2012 Nov 21;151(5):951-63. PMCID: PMC3511858

106. Donohoe DR, Collins LB, Wali A, Bigler R, **Sun W**, Bultman SJ (2012)
The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation.
Mol Cell. 2012 Nov 30;48(4):612-26. PMCID: PMC3513569

107. Zhao N, Ang MK, Yin XY, Patel MR, Fritchie K, Thorne L, Muldrew KL, Hayward MC, **Sun W**, Wilkerson MD, Chera BS, Hackman T, Zanation AM, Grilley-Olson JE, Couch ME, Shockley WW, Weissler MC, Shores CG, Funkhouser WK, Olshan AF, Hayes DN (2012)
Different cellular p16(INK4a) localisation may signal different survival outcomes in head and neck cancer.
Br J Cancer, 2012 Jul 24;107(3):482-90. PMCID: PMC3405208

108. **Collaborative Cross Consortium** (2012)
The genome architecture of the Collaborative Cross mouse genetic reference population.
Genetics. 190(2):389-401. PMCID: PMC3276630

109. Xia K, Shabalin AA, Huang S, Madar V, Zhou YH, Wang W, Zou F, **Sun W**, Sullivan
PF, Wright FA (2012)
SeeQTL: a searchable database for human eQTLs.
Bioinformatics. 2012 Feb 1;28(3):451-2. PMCID: PMC3268245.

110. **Sun W**, Lee S, Zhabotynsky V, Zou F, Wright FA, Crowley JJ, Yun Z, Buus RJ, Miller DR, Wang J, McMillan L, Pardo-Manuel de Villena F, Sullivan PF (2012)
Transcriptome atlases of mouse brain reveals differential expression across brain regions and genetic backgrounds.
G3 (Bethesda). 2012 Feb;2(2):203-11. PMCID: PMC3284328

111. Donohoe DR, Garge N, Zhang X, **Sun W**, O'Connell TM, Bunger MK, Bultman SJ (2011)
The Microbiome and Butyrate Regulate Metabolism and Autophagy in the Mammalian Colon.
Cell Metabolism, Vol. 13, 517-526. PMCID: PMC3099420

112. Jones MD, **Sun W**, and Aitken MD (2011)

Multiple DNA extractions coupled to stable-isotope probing of anthracene-degrading bacteria in contaminated soil.
Applied and Environmental Microbiology 2011, Vol. 77, 2984-2991.

113. Campos M*, **Sun W***, Yu F*, Barbalic M, Tang W, Chambless LE, Wu KK, Ballantyne C, Folsom AR, Boerwinkle E, Dong JF (2011)
Genetic Determinants of Plasma von Willebrand Factor Antigen Levels: A Target Gene SNP and Haplotype Analysis of ARIC Cohort.
Blood 2011, Vol. 117, 5224-5230. PMCID: PMC3109544
* Co-first authors

114. Dehghan A, Dupuis J, …, **Sun W**., …, Chasman DI (2011)
Meta-analysis of genome-wide association studies in over 80,000 subjects identifies multiple loci for C-reactive protein levels.
Circulation 2011, Vol. 123, 731-738

115. Wright FA, Strug LJ, Doshi VK, ..., **Sun W**, ..., Knowles MR, Cutting GR (2011)
Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2.
Nat Genet 2011 Jun;43(6):539-46. PMCID: PMC3296486.

116. Gatti DM, Lu L, Williams RW, **Sun W**, Wright FA, Threadgill DW, Rusyn I (2011)
MicroRNA Expression in the Livers of Inbred Mice.
Mutation Research, 2011 Sep 1;714(1-2):126-33. Epub 2011 May 14

B. Other Refereed Scholarly Articles

C. Books and Book Chapters

Wei Sun and Yijuan Hu (2014),
Mapping of Expression Quantitative Trait Loci Using RNA-seq Data
Statistical Analysis of Next Generation Sequencing Data
Springer Series: Frontiers in Probability and the Statistical Sciences
Datta, Somnath, Nettleton, Dan (Eds.)

Wei Sun and Min Jin Ha (2015)
eQTL and Directed Graphical Model
Integrating omics data: statistical and computational methods
Springer
George Tseng, Debashis Ghosh and Xianghong Jasmine Zhou (Eds.)

8. Patents and Other Intellectual Property

“Modules for Identifying Cancer-Reactive T Cell Receptors and Mutations Thereof.”
Inventors: **Wei Sun**, Si Liu, Philip Bradley, and Marie Bleakley. U.S. Provisional Patent Application No. 63/831,763. Filed June 27, 2025.

9. Funding History

A. Funded Projects

Ongoing Research Support

U01 HG013177 “MorPhiC U01” (Sun, Hsu, Shojae) 9/22/2023-6/30/2028
NIH/NHGRI
Statistical Methods for Inferring Gene-Phenotype Associations Using Omic Data from Gene Knockout and Human Phenotype Studies
Role: PI

U24 CA288185 “LBC” (Zheng, Sun) 9/2023-8/2028
Precompetitive Collaboration on Liquid Biopsy for Early Cancer Assessment: Data Management and Coordinating Unit
Role: mPI

5 R01 GM105785-02 (Sun) 05/15/2014-7/31/2027
National Institutes of Health
Statistical Methods for RNA-seq Data Analysis
We propose to develop statistical methods and software for RNA-seq data analysis, with emphasis on dissecting the genetic and epigenetic basis of allele-specific and RNA-isoform-specific expression. Role: PI

U54 CA274374 (Grady) 9/1/2022 – 8/31/2027
NIH/NCI
Understanding Adenoma Progression: Interplay Among Tissue Microenvironment, Clonal Architecture, and Gut Microbiome
To study adenomas as early lesions and to determine the mechanisms involved in the formation and progression of early precancerous lesions. The proposed studies will integrate basic and translational cancer research projects to iteratively examine the direct causal relationships and interactions of adenomas, the colon “primed” microenvironment, and host-systemic factors as “co-organizers” of adenoma initiation and/or progression.
Status of Support: Active

W81XWH-21-BCRP-TBCCA (Li/Ghajar) 07/2022-6/2026
DOD
Prevention of lethal metastatic breast cancer by identifying and eradicating clinically relevant disseminated tumor cells
To comprehensively profile human disseminated breast cancer cells, their niches and systemic environments, and the defining features of patients who harbor them.

R01 CA266386-01A1 (Kooperberg/Buas) 09/2022-08/2027
NIH/NCI
Genetics, Epigenetics, and Risk Prediction for Esophageal Adenocarcinoma
Study underlying susceptibility to and develop prediction models for EAC.

10. Public Health Practice Activities

11. Conferences and Symposiums

Organizer for invited Sessions for JSM 2011, 2016, 2022 and ENAR 2011-2015, 2017

Invited Conference/Workshop presentation

1. Workshop on Statistical Frontiers, Institute of statistical science, Academia Sinica, Taipei, Taiwan, *December 2010*, Title: Statistical Methods for eQTL Mapping using RNA-seq Data
2. 2010 Annual Meeting of Chinese Statistical Society and International Statistical Conference, National Central University, Taiwan, *December 2010*, Title: Statistical Strategies for Some Variable Selection Problems in Genetic Studies
3. International Chinese Statistical Association 2011 Applied Statistics Symposium, *June 27, 2011*, New York City, Title: Statistical Methods for eQTL Mapping using RNA-seq Data
4. Peking University, School of Mathematical Sciences, Young Mathematician Forum in celebrating of 100th anniversary of the School of Mathematical Sciences (SMS) at Peking University, *June 18-21, 2013*, Title: Statistical Methods for Cancer Genomics
5. SAMSI, LDHD Transition Workshop, *May 05 2014*, Title: Estimation of High Dimensional Directed Acyclic Graphs using eQTL Data
6. Statistical Society of Canada 2014 Annual Meeting, *May 27 2014*, Toronto, Canada, Title: A study of RNA-seq data in Cancer Patients
7. Triangle Statistical Genetics Conference, SAS Campus, *Oct 31 2014*, Title: Statistical methods to exploit allele-specific and isoform-specific information from RNA seq data
8. 2015 IMS-China International Conference on Statistics and Probability, Kuming, China, *July 1-4*, Title: PenPC: A Two-step Approach to Estimate the Skeletons of High Dimensional Directed Acyclic Graphs
9. 2015 ICSA China Statistics Conference, Shanghai, China, *July 6-7*, Title: eQTL mapping and integrative analysis of cancer genomic data
10. 60th World Statistics Congress, *July 26-31, 2015*, Title: PenPC: A Two-step Approach to Estimate the Skeletons of High Dimensional Directed Acyclic Graphs
11. 2016 JSM, Chicago, *July 30-Aug 4, 2016*, Using Multiple Types of Omic Data to study ITH and beyond
12. 2017 ENAR, Washington DC, *March 12-15, 2017* Title: Studying Intra-Tumor Heterogeneity using DNA Sequencing Data from Single Cells and Bulk-Tissue Sample
13. Department of Biostatistics, University of Florida, *April 7-8, 2017*, Biostatistics workshop: “Statistical Inference for Biomedical Big Data”, Title: Mapping Gene Expression QTL of Impure Tumor Samples

14. 14th Graybill Conference on Statistical Genomics and Genetics, *June 5-7, 2017*, Title: Analyzing Cancer Omic Data as Compositional Data
15. 2018 ENAR, Atlanta, GA, *March 25-28, 2018* Title: Estimation of Intra-Tumor Heterogeneity and Assessing its Impact on Survival Time
16. iBright Conference 2019, MD Anderson, *November 11-13, 2019,* Title: CARseq: Cell type Aware analysis of RNA-seq data
17. 2021 JSM, virtual, *Aug 8 – Aug 12, 2021*, Analyzing Genomic Data Using Neural Networks
18. 2022 JSM, Washington DC, *Aug 6 – Aug 11, 2022*, Individual-Level Differential Expression Analysis for Single Cell RNA-Seq Data
19. Banff Workshop Mathematical Methods in Cancer Biology, Evolution and Therapy, 23w5084, May 14 - 19, 2023, A three-body problem
20. STATGEN 2024, Pittsburgh, PA, *May 1-3, 2024,* Title: Graph neural network for spatial-proteomic data
21. 2024 JSM, Portland, OR, *Aug 3–8, 2024*, Estimation of neoantigen burden and its association with patient response to immunotherapy
22. 2025 ENAR, New Orleans, LA, *March 23-26, 2025*, Estimation of Gene Regulatory Network Using Gene KO Data
23. STATGEN 2025, Minneapolis, MN, *May 21-23, 2025,* Title: Joint Analysis of Spatial Transcriptomics and Histopathology Image Data
24. Banff Workshop Novel Statistical Approaches for Studying Multi-omics Data, 25w5324, *July 13 - 18, 2025*, Integrating scRNA-seq and scATAC-seq to infer gene regulatory network.
25. 2025 JSM, Nashville, TN, *Aug 2–7, 2025*, Graph neural network for spatial-proteomic data.

Invited Seminars

1. Computational Biology and Statistics Workshop, in celebrating the 10th anniversary of UCLA Department of Statistics, *October 19-20, 2008*, Title: Genome-wide Multiple Loci Mapping in Experimental Crosses of Inbred Strains Using Dense Genetic Markers
2. Department of Biostatistics, Emory University, *September 3, 2009*, Title: Integrated Study of Copy Number States and Genotype Calls Using High Density SNP Arrays
3. Department of Biostatistics, University of Washington St Louis, *February 19, 2010*, Title: Integrated Dissection of Copy Number States and Genotype Calls
4. Department of Biostatistics, University of Texas at Houston, *April 4, 2011*, Title: Allele-specific eQTL Mapping, Title: Allele-specific eQTL Mapping
5. The Jackson Laboratory, *May 24, 2011*, Title: Allele-specific expression by RNA-seq
6. Department of Statistics, Duke University, *October 14, 2011*, Title: Multiple loci mapping by penalized regression

7. Department of Statistics, UCLA, *Nov 1st, 2011*, Title: Multiple loci mapping by penalized regression
8. Department of Biostatistics, University of Pittsburgh, *Oct 11, 2012*, Title: Statistical methods for RNA seq studies
9. Department of Statistics, University of Illinois at Urbana-Champaign, *Oct 18, 2012*, Title: Statistical methods for RNA seq studies
10. NIEHS, *Nov 20, 2012*, Title: Statistical methods for RNA seq studies
11. NC State University, Biostatistics Working Group, *Feb 14, 2013*, Title: PenPC: A Two-step Approach to Estimate the Skeletons of High Dimensional Directed Acyclic Graphs
12. Department of Biostatistics, Duke University, *May 15 2014*, Title: Statistical Methods for Cancer Genomics
13. Department of Applied and Computational Mathematics and Statistics, *Oct 27 2014*, University of Notre Dame, Title: Statistical Methods for Cancer Genomics
14. National Jewish Health, Denver, *Jan 12, 2015*, Title: Genetic basis of gene expression in human and mouse
15. Department of Biostatistics & Informatics Colorado School of Public Health, Statistical Methods for Cancer Genomics, *Jan 13, 2015*, Title: Statistical Methods for Cancer Genomics
16. Bioinformatics Research Center, NC state University, *Feb 10, 2015*, Title: Statistical Methods for Cancer Genomics
17. Department of Biostatistics, University of Michigan, *Sep 30, 2015*, Title: The interplay of somatic copy number aberration, DNA methylation, and gene expression
18. Department of Statistics, Oregon State University, *May 16, 2016*, Title: Statistical methods to exploit allele-specific and isoform-specific information from RNA seq data
19. Biostatistics Branch, DCEG, National Cancer Institute, *March 16, 2017*, Title: The interplay of somatic copy number aberration, DNA methylation, and gene expression
20. Department of Biostatistics, Ohio State University, *November 16, 2018*, Title: Intra-Tumor Heterogeneity and Its Associations with Survival Time and Immunotherapy Treatment
21. Department of Biostatistics, MD Anderson Cancer Research Center, *March 20, 2019*, Title: Intra-Tumor Heterogeneity and Its Associations with Survival Time and Immunotherapy Treatment
22. Department of Biostatistics, Duke University, *October 11th, 2019*, Title: eQTL Mapping Using RNA-seq Data, assumptions, implementations, and cautions
23. Department of Biostatistics, University of Chicago, *October 23rd, 2019*, Title: CARseq: Cell type Aware analysis of RNA-seq data

24. Penn State Statistics Colloquia, Department of Statistics, The Pennsylvania State University, Feb 11, 2021, Title: Cell type-aware analysis of RNA-seq data,
25. Biostatistics Research Seminar, Department of Biostatistics, St. Jude Children's Research Hospital, April 27th, 2021, Title: Cell type-aware analysis of RNA-seq data
26. Colloquium in MSU's Department of Computational Mathematics, Science and Engineering, October 04, 2021, Title: Cell type-aware analysis of RNA-seq data
27. Department of Biostatistics, University of Michigan, *November 3rd, 2022*, Two statistical methods for single cell RNA-seq data
28. Biostatistics and Research Design (BARD) Center, Harvard Medical School, *Dec 13th, 2022*, Statistical methods for single cell RNA-seq data
29. Centre for Quantitative Medicine (CQM), Duke-NUS Medical School, *Feb 21st, 2023*, Statistical methods for single cell RNA-seq data
30. Department of Biostatistics, Pittsburgh University, *March 30th, 2023*, A three-body problem
31. InterMEL: an international consortium to study stage II and III melanoma, Zoom semianr, *June 8th, 2023*, Computational methods to study the interactions between T cell receptors (TCRs), human leukocyte antigens (HLAs), and neoantigens
32. Department of Biostatistics, Memorial Sloan-Kettering Cancer Center, *May 8th, 2024,* Machine learning methods for somatic mutation data analysis

12. University Service

13. Professionally-Related Community Service

14. Diversity, Equity and Inclusion Activities

15. Teaching History

Formal Courses

Bios784 Introduction to computational biology
2009 spring, 2011 spring, 2013 spring

Bios735 Statistical Computing
2013 fall

Bios663 Intermediate linear models
2010 spring, 2012 spring, 2014 spring

16. Advising and Formal Mentoring

A. PhD dissertations, Chair or co-chair

John C. Schwarz (2006-2010), “Enhancing eQTL Analysis Techniques with Special Attention to the Transcript Dependency Structure”, co-advised with Dr. Fred Wright.

Current Position: Analytic Lead at Monsanto Company.

Naim Rashid (2007-2013), “Model-based approaches for the detection of biologically active genomic regions from next generation sequencing data”, co-advised with Dr. Joe Ibrahim.

Current Position: Research Assistant Professor at Department of Biostatistics, UNC Chapel Hill

Min-Jin Ha (2008-2013), “Estimation of directed and non-directed gene expression network using gene expression and genetic data”.

Current Position: Assistant Professor at Department of Biostatistics at MD Anderson Cancer Center

Ting-Huei Chen (2009-2014), “Penalized Estimation Methods and Their Applications in Genomics and Beyond”, co-advised with Dr. Jason Fine.

Current Position: Assistant Professor at Department of Mathematics, Laval University, Canada

Matt Psioda (2013-2016), “Statistical Methods for DNA Methylation Studies and for Bayesian Clinical Trial Design”, co-advised with Dr. Joe Ibrahim.

Current Position: Research Assistant Professor at Department of Biostatistics, UNC Chapel Hill

Doug Wilson (2014-2018), “Statistical Methods for The Deconvolution of Heterogeneous Cell Type Samples using RNA-seq”, co-advised with Dr. Joe Ibrahim

Jenny Yang (2014-2018), “Model-free estimation of graphic model”.

Paul Little (2014-2019), “Intra-tumor Heterogeneity” co-advised with Dr. Danyu Lin

Vasyl Zhabotynsky (2014-2020), “Statistical Model for RNA-seq data”, co-advised with Dr. Fei Zou

Chong Jin (2014-2020), “Intra-tumor Heterogeneity”, co-advised with Dr. Danyu Lin

Laura Zhou (2018-2022), “Deep learning methods to predict neoantigen”, co-advised with Dr. Fei Zou

B. Master Theses, Chair

Zhengzheng Tang (2009, UNC), Integrated study of copy number states and genotype calls using high-density SNP arrays

This work won Department of Biostatistics Regina Elandt-Johnson Award - Best Masters paper completed in 2009

Guanhua Chen (2010, UNC), Statistical methods for analyzing customized copy number variation array

Ni Zhao (2012, UNC), eQTL Mapping Using RNA-seq Data

Zhining Sui (2023, UW), A Deep Learning Approach to Infer Cellular Features from Pathology Imaging Data

C. Mentored Scientists and Postdocs

Yang Liu (2017-2018), Postdoctoral fellow at Fred Hutch Cancer Center

Aaron Molstad (2017-19), Postdoctoral fellow at Fred Hutch Cancer Center, co-mentored with Dr. Hsu Li

Mengqi Zhang (2019-2020), Postdoctoral fellow at Fred Hutch Cancer Center

Paul Little (2019-2022), Postdoctoral fellow at Fred Hutch Cancer Center

Si Liu (2020-2023) Postdoctoral fellow at Fred Hutch Cancer Center
Si Liu (2023-) Staff scientist at Fred Hutch Cancer Research Center

Bo Yu (2023-) Postdoctoral fellow at Fred Hutch Cancer Center

Saishi Cui (2024-) Postdoctoral fellow at Fred Hutch Cancer Center

Siqi Liu (2025-) Postdoctoral fellow at Fred Hutch Cancer Center, co-advised with Dr. Li Hsu

D. MS and PhD Committees in Non-Chair Role

	Student	**Department**	**Advisor**	**Graduation**
1	Ramon I. Garcia	Biostatistics	Joe Ibrahim & Hongtu Zhu	2009
2	Daniel Gatti	Environmental Sciences	Ivan Rusyn	2010 March
3	Seo Young Park	Statistics	Yufeng Liu	2010 April
4	Paul Giresi	Biology	Jason Lieb	2010 May
5	YuYing Xie	Genetics	David Threadgill	2010 July
6	Von Walter	Biostatistics	Fred Wright	2010 July
7	Yijuan Hu	Biostatistics	Danyu Lin	2011 May
8	Yi Gong	Biostatistics	Fei Zou	2011 June
9	Yihui Zhou	Biostatistics	Fred Wright	2011 Nov.
10	Eric Lock	Statistics	Andrew Nobel & Steve Marron	2012 April
11	Colin Lickwar	Biology	Jason Lieb	2012 July
12	Thomas Clarke	Biology	Vison Todd	2012 Oct

13	Ja-an Lin	Biostatistics	Hongtu Zhu & Joe Ibrahim	2013 July
14	Lan Liu	Biostatistics	Michael Hudgens	2013 July
15	Khondker Zakaria	Biostatistics	Hongtu Zhu & Joe Ibrahim	2013 Aug
16	Wonil Chung	Biostatistics	Fei Zou	2013 Nov
17	Gene Urrutia	Biostatistics	Michael Wu	2013 Nov
18	Andrea Byrnes	Biostatistics	Yun Li	2013 Nov
19	Ni Zhao	Biostatistics	Michael Wu	2013 Dec
20	Zhengzheng Tang	Biostatistics	Danyu Lin	2014 June
21	Guanhua Chen	Biostatistics	Michael Kosorok	2014 June
22	Weibo Wang	Computer Science	Wei Wang	2014 Dec
23	Shunping Huang	Computer Science	Wei Wang	2014 Dec
24	Zhaoyu Yin	Biostatistics	Fei Zou	2014
25	Matthew Weiser	Genetics	Terry Furey	2015
26	Yuying Xie	Genetics	William Valdar and Yufeng Liu	2015
27	Wei Cheng	Computer Science	Wei Wang	2015
28	Chen-Ping Fu	Computer Science	Leonard McMilan	2015

Pearl Liu (2021-), Department of Biostatistics, University of Washington
Kernel-based statistical methods for genomic and microbiome data
Advisor: Michael Wu

Angela Zhang (2021-2022) Department of Biostatistics, University of Washington
Statistical method development for microbiome data
Advisor: Michael Wu

Maio, Zhen (2020-2023), Department of Statistics, University of Washington
Advisor: Adrian Dobra

Sarah Huang (2023-) Molecular & Cellullar Biology, University of Washington
Advisor: Manu Setty

Liuye Huang (2025-), Epidemiology, University of Washington
Advisor: Christopher Li

E. Other mentors

Mary Cooter (2011), undergraduate student at UNC, A Two-step Approach for Accurate Detection of Copy Number Variations

F. Academic Advising

Yujia Cai (2021-2022), MS at UW biostatistics, analysis of TCR data.

Zhexiao Lin (2021-2022), MS at UW statistics, deep learning method for scRNA-seq data.

Zining Sui (2023-2024), MS at UW biostatistics, HE images and spatial transcriptomics.

Feng Ding (2023-2024), MS at UW statistics, TCR repertoire-wide association studies (TReWAS).

Xinyi Chen (2023-2025), MS at UW statistics, TCR predictor for severe COVID-19 disease.

Fumin Li (2023-2025), MS at UW statistics, prediction of TCR-HLA association. ScGPT model.

Zhaoheng Li (2023-2027), PhD at UW Biostatistics, gene regulator network and spatial omics.