December 20, 2022

CURRICULUM VITAE

HONGKAI JI, Ph.D., M.A., M.E.

PERSONAL DATA

Mailing Address:	Department of Biostatistics Johns Hopkins Bloomberg School of Public Health 615 North Wolfe Street, Room E3638 Baltimore, MD 21205
Phone:	410-955-3517
Fax:	410-955-0958
E-Mail:	hji@jhu.edu
Web Page:	http://www.biostat.jhsph.edu/~hji

EDUCATION AND TRAINING

Ph.D.	2007	Harvard University, Cambridge, MA, USA (Thesis Advisor: Professor Wing Hung Wong)	Statistics
	2004-2007	Stanford University, Stanford, CA, USA (Visiting Student)	Statistics
M.A.	2004	Harvard University, Cambridge, MA, USA	Statistics
M.E.	2002	Tsinghua University, Beijing, P.R. China	Automation
B.E.	1999	Tsinghua University, Beijing, P.R. China	Automation

PROFESSIONAL EXPERIENCE

Johns Hopkins University (JHU)

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health (JHSPH)

Professor	05/2018 – present
Associate Chair (Education)	08/2018 – present
Graduate Program Director	07/2017 – present
Graduate Program Co-director	08/2015 – 07/2017
Associate Professor	09/2013 – 05/2018
Assistant Professor	09/2007 – 08/2013

The Institute for Data Intensive Engineering and Science, Johns Hopkins University

Affiliated Faculty	12/2013 – present

Center for Computational Biology, Johns Hopkins University

Affiliated Faculty	03/2012 – present

High Throughput Biology HiT Center, Johns Hopkins School of Medicine

Affiliated Faculty	01/2011 – present

Other Non-JHU Professional Experience

Department of Statistics, University of California, Berkeley, CA, USA

Visiting Scholar	04/2011 – 07/2011

Beijing Institute of Genomics Chinese Academy of Sciences, Beijing, China

Visiting Scholar	07/2009 – 07/2009

PROFESSIONAL ACTIVITIES

Society Membership and Leadership

American Statistical Association (ASA), Fellow
ASA Section on Statistics in Genomics and Genetics, Program Chair (2019)
Institute of Mathematical Statistics (IMS)
International Chinese Statistical Association (ICSA)
International Society for Computational Biology (ISCB)

Program or Project Development

2022	Session Chair, ICSA Applied Statistics Symposium
2018	Scientific Program Committee, The 11th ICSA International Conference
2017	Editor, Special Issue on Statistics and Genomics for Statistics in Biosciences
2017	Session Organizer and Session Chair, ENAR
2016	Session Organizer and Session Chair, The 10th ICSA International Conference
2016	Session Chair, JSM
2016	Session Chair, IMS Asia Pacific Rim Meeting
2016	Session Chair, ICSA Applied Statistics Symposium
2016	Session Organizer and Session Chair, ENAR
2014	ASA White Paper for “Discovery with Data: Leveraging Statistics and Computer Science to Transform Science and Society”
2014	Session Organizer and Session Chair, ENAR
2011	Program Director for the Johns Hopkins – Nanjing University Exchange Program in Statistical and Data Sciences
2011	Organizing Committee, Statistical Methods for Very Large Datasets Conference
2011	Session Organizer and Session Chair, ENAR
2010	Session Chair, ENAR
2010	Session Organizer, ICSA
2009	Session Organizer, ENAR

EDITORIAL AND OTHER PEER REVIEW ACTIVITIES

Journal Peer Review Activities

Annals of Applied Statistics (AOAS)
Bioinformatics
Biometrics
Biostatistics
BMC Bioinformatics
BMC Genomics
Cancer Cell
Cancer Research
Cell Research
Cell Stem Cell
Cell Systems
Communications Biology
Computational Statistics
Computational Statistics and Data Analysis (CSDA)
F1000Research
Genome Biology
Genome Research
Genomics
IEEE/ACM Transactions on Computational Biology and Bioinformatics
IEEE Transactions on NanoBioscience
ISMB
Journal of Computational Biology (JCB)
Journal of the American Statistical Association (JASA)
Journal of the Royal Statistical Society, Series B (JRSSB)
Nature
Nature Biotechnology
Nature Cell Biology
Nature Communications
Nature Machine Intelligence
Nature Medicine
Nature Methods
Nature Review Genetics
Nucleic Acids Research (NAR)
Pacific Symposium on Biocomputing (PSB)
PLoS Computational Biology
Proceedings of the National Academy of Sciences of the USA (PNAS)
Quantitative Biology
Scientific Reports
Statistica Sinica
Statistical Analysis and Data Mining (SADM)
Statistical Applications in Genetics and Molecular Biology (SAGMB)

Statistics in Biosciences (SIBS)
Wiley Interdisciplinary Reviews: Systems Biology and Medicine

Journal or Other Editorial Board Membership

2022-2024	Co-Editor-in-Chief, Statistics in Biosciences (SIBS)
2015-2021	Associate Editor, Statistics in Biosciences (SIBS)
2011-present	Associate Editor, Statistical Applications in Genetics and Molecular Biology (SAGMB)
2019	Guest Editor, PLoS Computational Biology

Proposal Reviews

06/2022	Canada Research Chairs Program, Canada
12/2021	Dutch Research Council, NWO Talent Programme Vidi, The Netherlands
06/2020	NIH Study Section, MABS, US
02/2020	JHU IDIES Seed Grant Review, US
10/2019	NIH/NIDDK Review Panel, US
06/2019	NIH Study Section, MABS, US
03/2019	NIH/NIGMS Special Emphasis Panel for MIRA/R35, ZGM1 TRN-7 (MR), US
10/2018	NIH Study Section, MABS, US
02/2018	NIH/NHLBI Special Emphasis Panel for SBIR Contract Solicitation Topic 105, US
10/2017	NIH Study Section, MABS, US
07/2017	NIH NIGMS Review Panel for the Centers of Biomedical Research Excellence (COBRE) Program, US
08/2016	MRC Grant Review, UK
02/2016	NIH Study Section, BDMA, US
12/2014	NIH Member Conflict Review Panel, Development, Differentiation and Modulation of Immune Responses, US
10/2014	NIH Member Conflict Review Panel, Genes, Genomes, and Genetics IRG, US
10/2014	NIH Study Section, BMRD, US
07/2014	NIH Special Emphasis Review Panel, Endocrinology, Metabolism, Nutrition and Reproductive Sciences IRG, US
06/2014	NIH Study Section, GCAT, US
11/2011	NIH Special Emphasis Review Panel, ZHG1 HGR-M (J2), US
05/2010	NIH Challenge Grant Review, US
08/2007	BBSRC Grant Review, UK

Review of Reports and Other Documents

10/2009	Book Review, “Numerical Analysis for Statisticians (Second Edition)”, Springer

HONORS AND AWARDS

Honors

2021	Excellence in Teaching, Johns Hopkins Bloomberg School of Public Health
2020	Excellence in Teaching, Johns Hopkins Bloomberg School of Public Health
2019	Fellow, American Statistical Association
2018	Excellence in Teaching, Johns Hopkins Bloomberg School of Public Health
2017	Excellence in Teaching, Johns Hopkins Bloomberg School of Public Health
2016	Excellence in Teaching, Johns Hopkins Bloomberg School of Public Health
2011	Excellence in Teaching, Johns Hopkins Bloomberg School of Public Health
2010	Excellence in Teaching, Johns Hopkins Bloomberg School of Public Health
2004	Certificate of Distinction in Teaching, Harvard University
2002	Outstanding Master’s Thesis, Tsinghua University
2000	Outstanding Graduate of Tsinghua University
1999	First Place of "Challenge Cup" Student Extra-Curricular Science and Technology Project Competition, Tsinghua University
1999	Third Place of NHK Robocon' 99 (Robosoccer) International Robotics Design Contest for Students, Japan

Awards

2020	Johns Hopkins Discovery Award
2016	Johns Hopkins Catalyst Award
2007	Faculty Innovation Award, Johns Hopkins Bloomberg School of Public Health
2005	Chinese Medical Science and Technology Award (No. 20050319P0802), Chinese Medical Association
2005	Beijing Science and Technology Award (No. 2005Med-2-024-02), Beijing Municipal People's Government, China
2002-2004	Student Fellowship, Harvard University
1995-2002	Tsinghua University Excellent Student Scholarship

Named Lectureships

06/03/2010	Constance van Eeden Distinguished Speaker Seminar, BC, Canada

PUBLICATIONS

(* corresponding author; § equal contribution; blue: advisees and lab members)

Journal Articles (Peer Reviewed)

1. Fang WX, Bell CM, Sapirstein A, Asami S, Leeper K, Zack DJ, **Ji HK***, Kalhor R* (2022) Quantitative fate mapping: A general framework for analyzing progenitor state dynamics via retrospective lineage barcoding. *Cell*. 185:4604–4620

2. Ramachandran J, Zhou WQ, Bardenhagen AE, Nasr T, Yates ER, Zorn AM, **Ji HK**, Vokes SA (2022) Hedgehog regulation of epithelial cell state and morphogenesis in the larynx. *eLife*. 11:e77055

3. Hou WP, Zhang MY, Ji YL, Hong XM, Wang GY, Xu R, Liang LM, Saria S, **Ji HK**. (2022) In utero exposure to cigarette smoking on child long-term risk of obesity: concordance of self-report, maternal and cord blood biomarkers. *Precision Nutrition*. 1(3):e00017

4. Liu QB, Miao HY, Li S, Zhang P, Gerber GF, Follmann D, **Ji HK**, Zeger SL, Chertowa DS, Quinna TC, Robinson ML, Kickler TS, Rothman RE, Fenstermacher KZJ, Braunstein EM, Cox AL, Farci P, Fauci AS, Lusso P (2022) Anti-PF4 Antibodies Associated with Disease Severity in Covid-19. *Proc. Natl. Acad. Sci. USA*. 119 (47):e2213361119

5. Frumento N, Figueroa A, Wang TC, Zahid MN, Wang SY, Massaccesi G, Stavrakis G, Crowe Jr. JE, Flyak AI, **Ji HK**, Ray SC, Shaw GM, Cox AL, Bailey JR (2022) Repeated exposure to heterologous hepatitis C viruses associates with enhanced neutralizing antibody breadth and potency. *Journal of Clinical Investigation*. 132(15):e160058

6. Dykema AG, Zhang BY, Woldemeskel BA, Garliss CC, Rashid R, Westlake T, Zhang L, Zhang JJ, Cheung LS, Caushi JX, Pardoll DM, Cox AL, **Ji HK**, Smith KN, Blankson JN (2022) SARS-CoV-2 vaccination diversifies the CD4+ spike-reactive T cell repertoire in patients with prior SARS-CoV-2 infection. *eBioMedicine*. 80:104048

7. Shu C, Jaffe AE, Sabunciyan S, **Ji HK**, Astemborski J, Sun J, Bakulski KM, Sosnowski DW, Mehta SH, Kirk GD, Maher BS (2002) Epigenome-wide association analyses of active injection drug use. *Drug and Alcohol Dependence*. 235(1): 109431

8. Wei X, Li ZY, **Ji HK**, Wu H (2022) EDClust: An EM-MM hybrid method for cell clustering in multiple-subject single-cell RNA sequencing. *Bioinformatics*. 38(10):2692-2699

9. Lex RK, Zhou WQ, Ji ZC, Falkenstein KN, Schuler KE, Windsor KE, Kim JD, **Ji HK**, Vokes SA (2022) GLI transcriptional repression is inert prior to Hedgehog pathway activation. *Nature Communications*. 13:808

10. Hong XM, Liang LM, **Ji HK**, Guerrerio P, Wang GY, Pearson C, Stampfer MJ, Hu F, Wang XB (2022) Fetal Lipidome and Incident Risk of Food Allergy: A prospective birth cohort study. *Pediatric Allergy and Immunology*. 33(2):e13722

11. Sharrock AV, Mulligan TS, Hall KR, Williams EM, White DT, Zhang L, Emmerich K, Matthews F, Nimmagadda S, Washington S, Le KD, Meir-Levi D, Cox OL, Saxena MT, Calof AL, Lopez-Burks ME, Lander AD, Ding D, **Ji HK**, Ackerley DF, Mumm JS (2022) NTR 2.0: a rationally-engineered prodrug converting enzyme with substantially enhanced efficacy for targeted cell ablation. *Nature Methods*. 19:205–215

12. Turnbull AE, **Ji HK**, Dinglas VD, Mendez-Tellez PA, Dennison Himmelfarb C, Shanholtz CB, Hosey MM, Hopkins RO, Needham DM. (2022) Understanding patients' perceived health after critical illness: analysis of two prospective, longitudinal studies of ARDS survivors. *Chest*. 161(2):407-417

13. Wang GY, Foney D, DiBari J, Hong XM, Showell N, Kim KS, **Ji HK**, Pearson C, Mirolli G, Rusk S, Sharfstein J, Cheng T, Zuckerman B, Wang XB (2021) A prospective cohort study on the intersectionality of obesity, chronic disease, social factors, and incident risk of COVID-19 in US low-income minority middle-age mothers. *International Journal of Obesity*. 45:2577–2584

14. Stephens KE, Zhou WQ, Renfro Z, Ji ZC, **Ji HK**, Guan Y, Taverna SD (2021) Global gene expression and chromatin accessibility of the peripheral nervous system in animal models of persistent pain. *Journal of Neuroinflammation*. 18(1):185

15. Zhang LY, Chen C, Fu J, Lilley B, Berlinicke C, Hansen B, Ding D, Wang G, Wang T, Shou D, Ye Y, Mulligan T, Emmerich K, Saxena MT, Hall KR, Sharrock AV, Brandon C, Park H, Kam TI, Dawson VL, Dawson TM, Shim JS, Hanes J, **Ji HK**, Liu JO, Qian J, Ackerley DF, Rohrer B, Zack DJ, Mumm JS (2021) Large-scale Phenotypic Drug Screen Identifies Neuroprotectants in Zebrafish and Mouse Models of Retinitis Pigmentosa. *eLife*. 10:e57245

16. Caushi J§, Zhang JJ§, Ji ZC, Vaghasia A, Zhang BY, Hsiue EHC, Mog BJ, Hou WP, Justesen S, Blosser R, Tam A, Anagnostou V, Cottrell TR, Guo HD, Chan HY, Singh D, Thapa S, Dykema AG, Burman P, Choudhury B, Aparicio L, Cheung LS, Lanis M, Belcaid Z, El Asmar M, Illei PB, Wang RL, Meyers J, Schuebel K, Gupta A, Skaist A, Wheelan S, Naidoo J, Marrone KA, Brock M, Ha J, Bush EL, Park BJ, Bott M, Jones DR, Reuss JE, Velculescu VE, Chaft JE, Kinzler KW, Zhou SB, Vogelstein B, Taube JM, Hellmann MD, Brahmer JR, Merghoub T, Forde PM, Yegnasubramanian S*, **Ji HK***, Pardoll DM*, Smith KN* (2021) Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. *Nature*. 596: 126-132

17. Cottrell TR§, Zhang J§, Zhang BY, Kaunitz GJ, Chan HY, Verde F, Bell JM, Hooper JE, Hammers HJ, Allaf ME, **Ji HK**, Taube JM, Smith KN (2021) Evaluating T cell cross-reactivity between tumors and immune-related adverse events with TCR sequencing: pitfalls in interpretations of functional relevance. *Journal for ImmunoTherapy of Cancer*. 9(7):e002642

18. Akindele NP, Kouo T, Gordon O, Karaba AH, Fenstermacher KZJ, Rubens JH, Beaudry J, Atik CC, Zhou WQ, **Ji HK**, Tao XT, Vaidya D, Mostafa H, Caturegli P, Blair P, Sauer L, Cox AL, Persaud D (2021) Distinct cytokine and chemokine dysregulation in hospitalized children with acute COVID-19 and multisystem inflammatory syndrome with similar levels of aasopharyngeal SARS-CoV-2. *The Journal of Infectious Diseases*. 224(4):606-615.

19. Karaba AH[§], Zhou WQ[§], Hsieh LL, Figueroa A, Massaccesi G, Rothman RE, Fenstermacher KZJ, Sauer L, Shaw-Saliba K, Blair PW, Robinson ML, Leung S, Wesson R, Alachkar N, El-Diwany R, **Ji HK**, Cox AL (2022) Differential cytokine signatures of SARS-CoV-2 and influenza infection highlight key differences in pathobiology. *Clinical Infectious Diseases*. 74(2):254-262

20. Gao ML, Fu Y, Zhou WQ, Gui GG, Lal B, Li YQ, Xia SL, **Ji HK**, Eberhart CG, Laterra J, Ying MY (2021) EGFR activates a TAZ-driven oncogenic program in glioblastoma. *Cancer Research*. 81:3580-3592.

21. Dykema AG[§], Zhang BY[§], Woldemeskel BA, Garliss CC, Cheung LS, Choudhury D, Zhang JJ, Aparicio L, Bom S, Rashid R, Caushi JX, Hsiue EHC, Cascino K, Thompson EA, Kwaa AK, Singh D, Thapa S, Ordonez AA, Pekosz A, D'Alessio FR, Powell JD, Yegnasubramanian S, Zhou SB, Pardoll DM, **Ji HK**, Cox AL, Blankson JN, Smith KN (2021) Functional characterization of CD4+ T-cell receptors cross-reactive for SARS-CoV-2 and endemic coronaviruses. *Journal of Clinical Investigation*. 131(10):e146922

22. Thompson EA, Cascino K, Ordonez AA, Zhou WQ, Vaghasia A, Hamacher-Brady A, Brady NR, Sun IH, Wang R, Rosenberg AZ, Delannoy M, Rothman R, Fenstermacher K, Sauer L, Shaw-Saliba K, Bloch EM, Redd AD, Tobian AAR, Horton M, Smith K, Pekosz A, D'Alessio FR, Yegnasubramanian S, **Ji HK**, Cox AL, Powell JD (2021) Metabolic programs define dysfunctional immune responses in severe COVID-19 patients. *Cell Reports*. 34(11): 108863

23. Koldobskiy MA, Jenkinson G, Abante J, Rodriguez DiBlasi VA, Zhou WQ, Pujadas E, Idrizi1 A, Tryggvadottir R, Callahan C, Bonifant CL, Rabin KR, Brown PA, **Ji HK**, Goutsias J, Feinberg AP (2021) Converging genetic and epigenetic drivers of paediatric acute lymphoblastic leukaemia identified by an information-theoretic analysis. *Nature Biomedical Engineering*. 5: 360–376

24. Xu R, Hong XM, Zhang BY, Huang WY, Hou WP, Wang GY, Wang XB, Igusa T, Liang LM, **Ji HK** (2021) DNA methylation mediates the effect of maternal smoking on offspring birthweight: A birth cohort study of multi-ethnic US mother-newborn pairs. *Clinical Epigenetics*. 13: 47

25. Ji ZC, **Ji HK*** (2021) Discussion of "Exponential-family Embedding with Application to Cell Developmental Trajectories for Single-cell RNA-seq Data". *Journal of the American Statistical Association*. 116(534): 471-474

26. Chamling X, Kallman A, Fang WX, Berlinicke C, Mertz J, Devkota P, Pantoja IM, Smith M, Ji ZC, Chang C, Kaushik A, Chen L, Whartenby K, Calabresi P, Mao HQ, **Ji HK**, Wang TH, Zack DJ (2021) Single-cell transcriptomic reveals molecular diversity and developmental heterogeneity of human stem cell-derived oligodendrocyte lineage cells. *Nature Communications*. 12:652

27. Zhou WQ, **Ji HK** (2021) Genome-wide prediction of chromatin accessibility based on gene expression. *WIREs Computational Statistics*. 13:e1544

28. Sinha S, Satpathy AT, Zhou WQ, **Ji HK**, Stratton JA, Jaffer A, Bahlis N, Morrissy S, Biernaskie JA (2021) Profiling chromatin accessibility at single-cell resolution. *Genomics, Proteomics & Bioinformatics*. 19(2):172-190

29. Wang GY, Tang WY, **Ji HK**, Wang XB (2020) Prenatal exposure to mercury and precocious puberty: a prospective birth cohort study. *Human Reproduction*. 36(3): 712–720

30. Wang GY, Tang WY, Wills-Karp M, **Ji HK**, Bartell TR, Ji YL, Hong XM, Pearson C, Cheng TL, Wang XB (2020) A nonlinear relationship between maternal red blood cell manganese concentrations and child's blood pressure at age 6-12 y: A prospective birth cohort study. *The Journal of Nutrition*. 151(3): 570–578

31. Shu C, Jaffe AE, Sabunciyan S, **Ji HK**, Astemborski J, Sun J, Bakulski KM, Mehta SH, Kirk GD, Maher BS (2020) Epigenome-wide association scan identifies methylation sites associated with HIV infection. *Epigenomics*. 12(21):1917-1927

32. Hong XM, Zhang BY, Liang LM, Zhang Y, Ji YL, Wang GY, **Ji HK**, Clish C, Burd I, Pearson C, Zuckerman B, Hu FB, Wang XB (2020) Postpartum plasma metabolomic profile among women with preeclampsia and preterm delivery: implications for long-term health. *BMC Medicine*. 18(1):277

33. Hou WP, Ji ZC, **Ji HK***, Hicks S* (2020) A systematic evaluation of single-cell RNA-sequencing imputation methods. *Genome Biology*. 21:218

34. Hong XM, Surkan P, Zhang BY, Keiser A, Ji YL, **Ji HK**, Burd I, Bustamante-Helfrich B, Ogunwole SM, Tang WY, Liu L, Pearson C, Cerda S, Zuckerman B, Hao LX, Wang XB (2020) Genome-wide association study identifies a novel maternal gene × stress interaction associated with spontaneous preterm birth. *Pediatric Research*. 89(6):1549-1556

35. Qiu X, Ma F, Zhao MM, Cao Y, Shipp L, Liu A, Dutta A, Singh A, Braikia FZ, De S, Wood III WH, Becker KG, Zhou WQ, **Ji HK**, Zhao KJ, Atchison ML, Sen R (2020) Altered 3D chromatin structure permits inversional recombination at the IgH locus. *Science Advances*. 6(33):eaaz8850

36. Ji ZC, Zhou WQ, Hou WP, **Ji HK*** (2020) Single-cell ATAC-seq signal extraction and enhancement with SCATE. *Genome Biology*. 21: 161.

37. Lex RL§, Ji ZC§, Falkenstein KN§, Zhou WQ, Henry JL, **Ji HK**, Vokes SA (2020) GLI transcriptional repression regulates tissue-specific enhancer activity in response to Hedgehog signaling. *eLife*. 9:e50670

38. Surkan PJ, Hong XM, Zhang BY, Nawa N, **Ji HK**, Tang WY, Ji YL, Kimmel MC, Wang GY, Pearson C, Wang XB (2020) Can social support during pregnancy affect maternal DNA methylation? Findings from a cohort of African-Americans. *Pediatric Research*. 88:131-138

39. Zhang JJ§, Ji ZC§, Caushi JX§, Asmar ME§, Anagnostou V, Cottrell TR, Chan HY, Suri P, Guo HD, Merghoub T, Chaft JE, Reuss JE, Tam A, Blosser R, Abu-Akeel M, Sidhom JW, Zhao N, Ha JS, Jones DR, Marrone KA, Naidoo J, Gabrielson E, Taube JM, Velculescu VE, Brahmer JR, Housseau F, Hellmann MD, Forde PM, Pardoll DM, **Ji HK***, Smith KN* (2020) Compartmental analysis of T cell clonal dynamics as a function of pathologic response to neoadjuvant PD-1 blockade in resectable non-small cell lung cancer. *Clinical Cancer Research*. 26(6):1327-1337

40. Chen ZY§, Ji ZC§, Ngiow SF, Manne S, Cai ZY, Huang AC, Johnson J, Staupe RP, Bengsch B, Xu CY, Yu SX, Kurachi M, Herati RS, Vella LA, Baxter AE, Wu JE, Khan O, Beltra JC, Giles JR, Stelekati E, McLane LM, Lau CW, Yang XL, Berger SL, Vahedi G, **Ji HK**, Wherry EJ (2019) TCF-1-Centered Transcriptional Network Drives an Effector versus Exhausted CD8 T Cell-Fate Decision. *Immunity*. 51(5):840-855.e5

41. Wang GY, DiBari J, Bind E, Steffens AM, Mukherjee J, Azuine RE, Singh GK, Hong XM, Ji YL, **Ji HK**, Pearson C, Zuckerman BS, Cheng TL, Wang XB (2019) Association Between Maternal Exposure to Lead, Maternal Folate Status, and Intergenerational Risk of Childhood Overweight and Obesity. *JAMA Network Open*. 2(10):e1912343.

42. Zhou WQ, Ji ZC, Fang WX, **Ji HK*** (2019) Global prediction of chromatin accessibility using small-cell-number and single-cell RNA-seq. *Nucleic Acids Research*. 47(19):e121

43. Azuine RE, Ji YL, Chang HY, Kim Y, **Ji HK**, DiBari J, Hong XM, Wang GY, Singh GK, Pearson C, Zuckerman B, Surkan PJ, Wang XB (2019) Prenatal Risk Factors and Perinatal and Postnatal Outcomes Associated With Maternal Opioid Exposure in an Urban, Low-Income, Multiethnic US Population. *JAMA Network Open*. 2(6):e196405.

44. Hong XM, Liang L, Sun Q, Keet CA, Tsai HJ, Ji YL, Wang GY, **Ji HK**, Clish C, Pearson C, Wang Y, Wood RA, Hu FB, Wang X. (2019) Maternal triacylglycerol signature and risk of food allergy in offspring. *Journal of Allergy and Clinical Immunology*. 144(3): 729–737.

45. Stephens KE, Zhou WQ, Ji ZC, Chen Z, He S, **Ji HK**, Guan Y, Taverna SD (2019) Sex differences in gene regulation in the dorsal root ganglion after nerve injury. *BMC Genomics*. 20(1):147

46. Olapeju B, Ahmed S, Wang GY, Ji YL, Hong XM, Raghavan R, Summers A, Keiser A, **Ji HK**, Zuckerman B, Yarrington C, Hao LX, Surkan P, Cheng T, Wang XB (2019) Maternal postpartum plasma folate status and preterm birth in a high-risk US population. *Public Health Nutrition*. 22(7):1281-1291.

47. Zhao MM, Joy J, Zhou WQ, De S, Wood III WH, Becker KG, **Ji HK**, Sen R (2018) Transcriptional outcomes and kinetic patterning of gene expression in response to NF-κB activation. *PLoS Biology*. 16(9):e2006347

48. Tsai HJ, Wang GY, Hong XM, Yao TC, Ji YL, Radovick S, **Ji HK**, Cheng TL, Wang XB (2018) Early life weight gain and development of childhood asthma in a prospective birth cohort. *Annals of the American Thoracic Society*. 15(10):1197-1204

49. Ji YL, Riley A, Lee LC, Hong XM, Wang GY, Tsai HJ, Mueller N, Pearson C, Thermitus J, Panjwani A, **Ji HK**, Bartell T, Burd I, Fallin MD, Wang XB (2018) Maternal biomarkers of acetaminophen use and offspring attention deficit hyperactivity disorder. *Brain Sciences*. 8(7). pii: E127

50. Zhang BY, Hong XM, **Ji HK**, Tang WY, Kimmel M, Ji YL, Wang GY, Pearson C, Zuckerman B, Surkan PJ, Wang XB (2018) Maternal smoking during pregnancy and cord blood DNA methylation: new insight on sex differences and effect modification by maternal folate. *Epigenetics*. 13(5):505-518

51. Qiu X, Kumari G, Gerasimova, T, Du H, Labaran L, Singh A, De S, Wood III WH, Becker KG, Zhou WQ, **Ji HK**, Sen R (2018) Sequential enhancer sequestration dysregulates recombination center formation at the IgH locus. *Molecular Cell*. 70:21-33

52. Wang J, Zibetti C, Shang P, Sripathi SR, Zhang PW, Cano M, Hoang T, Xia SL, **Ji HK**, Merbs SL, Zack DJ, Handa JT, Sinha D, Blackshaw S, Qian J. (2018) ATAC-seq analysis reveals a widespread decrease of chromatin accessibility in age-related macular degeneration. *Nature Communications*. 9:1364

53. Kuang Z, **Ji HK**, Boeke JD. (2018) Stress response factors drive regrowth of quiescent cells. *Current Genetics*. 64(4):807-810

54. Kuang Z, Ji ZC, Boeke JD*, **Ji HK***. (2018) Dynamic motif occupancy (DynaMO) analysis identifies transcription factors and their binding sites driving dynamic biological processes. *Nucleic Acids Research*. 46(1): e2

55. Ji YL, Riley AW, Lee LC, Volk H, Hong XM, Wang GY, Angomas R, Stivers T, Wahl A, **Ji HK**, Bartell TR, Burd I, Paige D, Fallin MD, Zuckerman B, Wang XB. (2017) A prospective birth cohort study on maternal cholesterol levels and offspring attention deficit hyperactivity disorder: new insight on sex differences. *Brain Sciences*. 8(1). pii: E3

56. Surkan PJ, Dong LM, Ji YL, Hong XM, **Ji HK**, Kimmel M, Tang WY, Wang XB. (2017) Paternal involvement and support and risk of preterm birth: findings from the Boston Birth Cohort. *Journal of Psychosomatic Obstetrics & Gynecology*. 40(1):48-56

57. Han F, **Ji HK**, Ji ZC, Wang HL. (2017) A provable smoothing approach for high dimensional generalized regression with an application in genomics. *Electronic Journal of Statistics*. 11(2):4347-4403.

58. Kuang Z, Pinglay S, **Ji HK***, Boeke JD*. (2017) Msn2/4 regulate expression of glycolytic enzymes and control transition from quiescence to growth. *eLife*. 6:e29938

59. Martos SN§, Li T§, Ramos RB, Lou D, Dai HZ, Xu JC, Gao GL, Gao Y, Wang QL, An C, Zhang XL, Jia YK, Dawson VL, Dawson TM, **Ji HK**, Wang ZB. (2017) Two approaches reveal a new paradigm of 'switchable or genetics-influenced allele-specific DNA methylation (ASM) with potential in human disease. *Cell Discovery*. 3: 17038

60. Zhou WQ, Sherwood B, Ji ZC, Xue Y, Du F, Bai JW, Ying MY, **Ji HK***. (2017) Genome-wide prediction of DNase I hypersensitivity using gene expression. *Nature Communications.* 8:1038.

61. Ji ZC§, Zhou WQ§, **Ji HK***. (2017) Single cell regulome data analysis by SCRAT. *Bioinformatics*. 33(18): 2930-2932

62. White DT, Sengupta S, Saxena MT, Xu Q, Hanes J, Ding D, **Ji HK**, Mumm JS (2017) Immunomodulation-accelerated neuronal regeneration following selective rod photoreceptor cell ablation in the zebrafish retina. *Proc. Natl. Acad. Sci. USA*. 114:E3719-E3728.

63. Hong XM, Hao K, **Ji HK**, Peng SN, Sherwood B, Narzo AD, Tsai HJ, Liu X, Burd I, Wang GY, Ji YL, Caruso D, Mao GY, Bartell TR, Zhang ZY, Pearson C, Heffener L, Cerda S, Beaty T, Fallin MD, Lee-Parritz A, Zuckerman B, Weeks DE, Wang XB (2017) Genome-wide approach identified a novel gene-maternal pre-pregnancy BMI interaction on preterm birth. *Nature Communications*. 8:15608

64. Hong XM, Sherwood B, Ladd-Acosta C, Peng S, **Ji H**, Hao K, Burd I, Bartell TR, Wang G, Tsai HJ, Liu X, Ji Y, Wahl A, Caruso D, Lee-Parritz A, Zuckerman B, Wang X (2017) Genome-wide DNA methylation associations with spontaneous

preterm birth in US blacks: findings in maternal and cord blood samples. *Epigenetics* 13(2):163-172.

65. White DT, Eroglu AU, Wang G, Zhang L, Sengupta S, Ding D, Rajpurohit SK, Walker SL, **Ji HK**, Qian J, Mumm JS (2016) ARQiv-HTS, a versatile whole-organism screening platform enabling in vivo drug discovery at high-throughput rates. *Nature Protocols* 11:2432–2453

66. Norrie JL, Li Q, Co S, Huang BL, Ding D, Uy JC, Ji ZC, Mackem S, Bedford MT, Galli A, **Ji HK** and Vokes SA (2016) PRMT5 is essential for the maintenance of chondrogenic progenitor cells in the limb bud. *Development*. 143(24): 4608-4619.

67. Zhou WQ, Sherwood B, **Ji HK*** (2016) Global prediction of the functional genomic landscape: applications, methods and challenges. *Human Heredity*. 81(2): 88-105.

68. Mao G, Nachman RM, Sun Q, Zhang X, Koehler K, Chen Z, Hong XM, Wang GY, Caruso D, Zong G, Pearson C, **Ji HK**, Biswal S, Zuckerman B, Wills-Karp M, Wang XB (2016) Individual and Joint Effects of Early-Life Ambient PM2.5 Exposure and Maternal Pre-Pregnancy Obesity on Childhood Overweight or Obesity. *Environmental Health Perspectives*. 125(6):067005

69. Ji ZC, **Ji HK*** (2016) TSCAN: pseudo-time reconstruction and evaluation in single-cell RNA-seq analysis. *Nucleic Acids Research*. 44(13):e117.

70. Wang J, Xia SL, Arand B, Zhu H, Machiraju R, Huang K, **Ji HK**, Qian J (2016) Single-cell co-expression analysis reveals distinct functional modules, co-regulation mechanisms and clinical outcomes. *PLoS Computational Biology*. 12(4): e1004892.

71. Wang GY, Hu FB, Mistry KB, Zhang CL, Ren FZ, Huo Y, Paige D, Bartell T, Hong XM, Caruso D, Ji ZC, Chen Z, Ji YL, Pearson C, **Ji HK**, Zuckerman B, Cheng TL, Wang XB (2016) Associations of maternal prepregnancy BMI and plasma folate concentrations with child metabolic health. *JAMA Pediatrics*. 170(8):e160845

72. Hong XM, Ladd-Acosta C, Hao K, Sherwood B, **Ji HK**, Keete CA, Kumarf R, Carusoa D, Liu X, Wang GY, Chen Z, Ji YL, Mao GY, Walker SO, Bartell TR, Ji ZC, Sun YF, Tsai HJ, Pongracic JA, Weeks DE, Wang XB (2016) Epigenome-wide association study links site-specific DNA methylation changes with cow's milk allergy. *Journal of Allergy and Clinical Immunology*. 138 (3): 908 - 911.e9

73. Li Q, Lex RK, Chung HW, Giovanetti SM, Ji ZC, **Ji HK**, Person MD, Kim J, Vokes SA (2016) NANOG binds to GLI proteins and represses hedgehog-mediated transcription. *Journal of Biological Chemistry*. 291:7171-7182.

74. Ji ZC, Vokes SA, Dang CV, **Ji HK*** (2016) Turning publicly available gene expression data into discoveries using gene set context analysis. *Nucleic Acids Research*. 44(1): e8.

75. Lewandowski JP, Du F, Zhang SL, Powell MB, Falkenstein KN, **Ji HK**, Vokes SA (2015) Spatiotemporal regulation of GLI target genes in the mammalian limb bud. *Developmental Biology*. 406: 92-103.

76. Jin K, Park S, Teo WW, Korangath P, Cho SS, Yoshida T, Győrffy B, Goswami CP, Nakshatri H, Cruz LA, Zhou WQ, **Ji HK**, Su Y, Ekram M, Wu Z, Zhu T, Polyak K, Sukumar S (2015) HOXB7 is an ERα cofactor in the activation of HER2 and multiple ER target genes leading to endocrine resistance. *Cancer Discovery*. 5: 944-959.

77. Wamaitha SE, del Valle I, Cho LTY, Wei YY, Fogarty NME, Blakeley P, Sherwood RI, **Ji HK**, Niakan KK (2015) Gata6 potently initiates reprogramming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes & Development*. 29: 1239-1255

78. Zhang SL, Du F, **Ji HK*** (2015) A novel DNA sequence motif in human and mouse genomes. *Scientific Reports*. 5: 10444.

79. Wei YY, Tenzen T, **Ji HK*** (2015) Joint analysis of differential gene expression in multiple studies using correlation motifs. *Biostatistics*. 16: 31-46.

80. Ying MY, Tilghman J, Wei YY, Guerrero-Cazares H, Quinones-Hinojosa A, **Ji HK**, Laterra J (2014) KLF9 inhibits glioblastoma stemness through global transcription repression and integrin-α6 inhibition. *Journal of Biological Chemistry*. 289: 32742-32756.

81. Kuang Z, Cai L, Zhang XK, **Ji HK**, Tu BP, Boeke JD (2014) High-temporal-resolution view of transcription and chromatin states across distinct metabolic states in budding yeast. *Nature Structural & Molecular Biology*. 21: 854-863.

82. Wu H*, **Ji HK*** (2014) PolyaPeak: detecting transcription factor binding sites from ChIP-seq using peak shape information. *PLoS ONE*. 9:e89694.

83. Jaffe AE, Storey JD, **Ji HK**, Leek JT (2013) Gene set bagging for estimating the probability a statistically significant result will replicate. *BMC Bioinformatics*. 14:360.

84. Wu G, **Ji HK*** (2013) ChIPXpress: improved ChIP-seq and ChIP-chip target gene ranking using publicly available gene expression data. *BMC Bioinformatics*. 14:188.

85. Wang JY, Park JS, Wei YY, Rajurkar M, Cotton JL, Fan Q, Lewis BC, **Ji HK**, Mao JH (2013) TRIB2 acts downstream of Wnt/TCF in liver cancer cells to regulate YAP and C/EBPalpha function. *Molecular Cell*. 51:211-225.

86. **Ji HK***, Li X, Wang QF, Ning Y (2013) Differential principal component analysis of ChIP-seq. *Proc. Natl. Acad. Sci. USA*. 110: 6789-6794. PMCID: PMC3637734

87. Wu G[§], Yustein JT[§], McCall MN, Zilliox M, Irizarry RA, Zeller K, Dang CV, **Ji HK*** (2013) ChIP-PED enhances the analysis of ChIP-seq and ChIP-chip data. *Bioinformatics*. 29:1182-1189.

88. Newman RH, Hu JF, Rho HS, Xie Z, Neiswinger J, Cooper C, Shirley M, Clark H, Hu SH, Hwang W, Jeong J, Wu G, Lin J, Gao XX, Ni Q, Goel R, Xia SL, **Ji HK**, Dalby K, Birnbaum MJ, Cole PA, Knapp S, Ryazanov A, Zack D, Blackshaw S, Pawson T, Gingras AC, Desiderio S, Pandey A, Turk B, Zhang J, Zhu H, Qian J. (2013) Construction of human activity-based phosphorylation networks. *Molecular Systems Biology*. 9:655.

89. Wei YY, Wu G, **Ji HK*** (2013) Global mapping of transcription factor binding sites by sequencing chromatin surrogates: a perspective on experimental design, data analysis, and open problems. *Statistics in Biosciences*. 5:156-178.

90. Kano S, Colantuoni C, Han F, Zhou Z, Yuan Q, Wilson A, Takayanagi Y, Lee Y, Rapoport J, Eaton W, Cascella N, **Ji HK**, Goldman D, Sawa A. (2013) Genome-wide profiling of multiple histone methylations in olfactory cells: further implications for cellular susceptibility to oxidative stress in schizophrenia. *Molecular Psychiatry*. 18(7):740-2.

91. Wei YY[§], Li X[§], Wang QF, **Ji HK*** (2012) iASeq: integrating multiple ChIP-seq datasets for detecting allele-specific binding. *BMC Genomics*. 13:681. PMCID: PMC3576346

92. Hwang W, Hackler Jr. L, Wu G, **Ji HK**, Zack DJ, Qian J. (2012) Dynamics of regulatory networks in the developing mouse retina. *PLoS ONE*. 7:e46521. PMCID: PMC3463606

93. **Ji HK***, Wu G, Zhan X, Nolan A, Koh C, De Marzo A, Doan, HM, Fan JS, Cheadle C, Fallahi M, Cleveland JL, Dang CV*, Zeller K*. (2011) Cell-type independent MYC target genes reveal a primordial signature involved in biomass accumulation. *PLoS ONE*. 6:e26057. PMCID: PMC3198433

94. Wang QF, Wu G, Mi SL, He FH, Wu J, Dong JF, Luo R, Mattison R, Kaberlein J, Prabhakar S, **Ji HK**, Thirman MJ. (2011) MLL fusion proteins preferentially regulate a subset of wild type MLL target genes in the leukemic genome. *Blood*. 117: 6895-6905. PMCID:PMC3128481

95. Chen L, Wu G, **Ji HK*** (2011) hmChIP: a database and web server for exploring publicly available human and mouse ChIP-seq and ChIP-chip data. *Bioinformatics*. 27: 1447-1448. PMCID:PMC3087956

96. Niakan KK, **Ji HK**[§], Maehr R[§], Vokes SA[§], Rodolfa KT, Sherwood RI, Yamaki M, Dimos JT, Chen AE, Melton DA, McMahon AP and Eggan K (2010) Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating

extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes & Development*. 24: 312-326.

97. Wu H and **Ji HK*** (2010) JAMIE: joint analysis of multiple ChIP-chip experiments. *Bioinformatics*. 26: 1864-1870.

98. Huang CR, Schneider AM, Lu Y, Niranjan T, Shen P, Robinson MA, Steranka JP, Valle D, Civin CI, Wang T, Wheelan SJ, **Ji HK**, Boeke JD, Burns KH (2010) Mobile interspersed repeats are major structural variants in the human genome. *Cell*. 141: 1171-1182.

    Comment in Lupski JR, *Cell*, 141: 1110-1112 (2010).

    Huang CR is a PhD student in human genetics (advisor: Boeke JD) enrolled concurrently in the biostatistics MHS program. I am her advisor for the MHS degree. We developed the computational tool for TIP-chip data analysis in this paper.

99. Lee EY, **Ji HK**, Ouyang ZQ, Zhou BY, Ma WX, Vokes SA, McMahon AP, Wong WH, Scott MP (2010) Hedgehog pathway-regulated gene networks in cerebellum development and tumorigenesis. *Proc. Natl. Acad. Sci. USA*. 107: 9736-9741.

100. **Ji HK*** and Liu XS* (2010) Analyzing 'omics data using hierarchical models. *Nature Biotechnology*. 28: 337-340.

101. Sun Y, Fan XY, Cao DM, Tang W, He K, Zhu JY, He JX, Bai MY, Zhu S, Oh E, Patil S, Kim TW, **Ji HK**, Wong WH, Rhee SY, Wang ZY. (2010) Integration of brassinosteroid signal transduction with the transcription network for plant growth regulation in Arabidopsis. *Developmental Cell*. 19: 765-777.

102. Judy JT and **Ji HK*** (2009) TileProbe: modeling tiling array probe effects using publicly available data. *Bioinformatics*. 25: 2369-2375.

103. Nishi Y, **Ji HK**, Wong WH, McMahon AP and Vokes SA (2009) Modeling the spatio-temporal network that drives patterning in the vertebrate central nervous system. *BBA - Gene Regulatory Mechanisms*. 1789: 299-305.

104. **Ji HK**, Jiang H, Ma W, Johnson DS, Myers RM and Wong WH (2008) An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nature Biotechnology*. 26: 1293-1300.

    Research Highlights in *Nature Methods* 6: 9 (2009).
    Highlighted in Park PJ, *Nature Reviews Genetics* 10: 669-680 (2009).

105. Vokes SA, **Ji HK**, Wong WH and McMahon AP (2008) A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog mediated patterning of the mammalian limb. *Genes & Development*. 22: 2651-2663.

106. Vokes SA, **Ji HK**, McCuine S, Tenzen T, Giles S, Zhong S, Longabaugh WJ, Davidson EH, Wong WH and McMahon AP (2007) Genomic characterization of Gli-activator targets in sonic hedgehog-mediated neural patterning. *Development*, 134: 1977-1989.

107. Paik JH, Kollipara R, Chu G, **Ji HK**, Xiao Y, Ding Z, Miao L, Tothova Z, Horner JW, Carrasco DR, Jiang S, Gilliland DG, Chin L, Wong WH, Castrillon DH and DePinho RA (2007) FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. *Cell*. 128 (2): 309-323.

108. **Ji HK**, Vokes SA and Wong WH (2006) A comparative analysis of genome-wide chromatin immunoprecipitation data for mammalian transcription factors. *Nucleic Acids Research*, 34: e146.

109. **Ji HK** and Wong WH (2006) Computational biology: toward deciphering gene regulatory information in mammalian genomes. *Biometrics*, 62: 645-663.

110. Kim RS, **Ji HK** and Wong WH (2006) An improved distance measure between the expression profiles linking co-expression and co-regulation in mouse. *BMC Bioinformatics*, 7:44.

111. **Ji HK** and Wong WH (2005) TileMap: create chromosomal map of tiling array hybridizations. *Bioinformatics*, 21: 3629-3636.

112. **The Chimpanzee Sequencing and Analysis Consortium** (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, 437: 69-87. (Listed among 67 authors).

Comment in News and Views: Li WH and Saunders MA**,** *Nature*, 437: 50-51 (2005).

113. Hellmann I, Prüfer K, **Ji HK**, Zody MC, Pääbo S and Ptak SE (2005) Why do human diversity levels vary at a megabase scale? *Genome Research*, 15: 1222-1231.

114. **Chinese Schizophrenia Consortium** (2004) Association study of an SNP combination pattern in the dopaminergic pathway in paranoid schizophrenia: a novel strategy for complex disorders. *Molecular Psychiatry*, 9: 510-521. **(Joint First Author with Xu Q, Yuan YB and Yu YQ)**

Highlighted in News & Commentary: Irizarry KJL and Galbraith SJ, *Mol Psychiatry* 9: 431-432 (2004).

115. **Ji HK**[§], Zhou Q[§], Wen F, Xia HY, Lu X and Li YD (2001) AsMamDB: an alternative splice database of mammals. *Nucleic Acids Research*, 29: 260-263.

116. **Ji HK**, Zhou Q, Wen F and Ji L (2001) Research on multiple alignments for alternative splicing. _Journal of Tsinghua University (Science and Technology)_, 41(9): 111-114.

Books, Monographs, Thesis

117. **Ji HK** (2007) Decoding mammalian gene regulatory programs through efficient microarray, ChIP-chip and sequence analysis. PhD thesis, Harvard University

118. **Ji HK** (2002) Analysis of alternative splicing and complex disease based on population comparisons. Master's Dissertation. Tsinghua University.

Chapters

119. Ji ZC, **Ji HK** (2019) Pseudotime Reconstruction Using TSCAN. _Computational Methods for Single-Cell Data Analysis_, pp 115-124, Edited by Guo-Cheng Yuan, Humana Press, Springer

120. Wei YY, Tenzen T and **Ji HK** (2016) Cormotif: An R package for jointly detecting differential gene expression in multiple studies. _Computational Biology and Bioinformatics_, pp 28-47. Edited by Ka-Chun Wong, CRC Press

121. **Ji HK** and Wei YY (2015) Integrative analysis of multiple ChIP-X data sets using correlation motifs. _Integrating Omics Data_, pp 110-134. Edited by George C. Tseng, Debashis Ghosh and Xianghong J. Zhou. Cambridge University Press.

122. Wu H and **Ji HK** (2012) JAMIE: A Software Tool for Jointly Analyzing Multiple ChIP-chip Experiments. In the *Next Generation Microarray Bioinformatics*, _Methods Mol Biol._, 802: 363-75, Humana Press, Springer

123. **Ji HK** (2011) Computational Analysis of ChIP-chip Data. _Handbook of Computational Statistics_, edited by Henry Horng-Shing Lu, Bernhard Schölkopf and Hongyu Zhao. 2011, Part 2, 257-282, DOI: 10.1007/978-3-642-16345-6_12. Springer

124. **Ji HK**, Jiang H, Ma WX and Wong WH (2011) Using CisGenome to Analyze ChIP-chip and ChIP-seq Data. _Current Protocols in Bioinformatics_. 2011 Mar, Chapter 2: Unit2.13. Wiley

125. **Ji HK** (2010) Computational Analysis of ChIP-seq Data. In the *Computational Biology of Transcription Factor Binding*, _Methods in Molecular Biology_, 674: 143-159. Edited by Istvan (Steve) Ladunga, Humana Press, Springer

126. Cullen P et al. including **Ji HK** (2004) Computational methods and bioinformatic tools. _Analysing Gene Expression_, edited by Lorkowski S and Cullen P, pp 769-904, Wiley-VCH Verlag GmbH & Co. KGaA

Articles, Editorials and Other Publications Not Peer Reviewed

127. Rudin C, Dunson D, Irizarry R, **Ji HK**, Laber E, Leek J, McCormick T, Rose S, Schafer C, van der Laan M, Wasserman L, Xue L. (2014) Discovery with data: Leveraging statistics with computer science to transform science and society. American Statistical Association White Paper

Book Translation

128. Translated the *Bioinformatics (The Machine Learning Approach)* by Pierre Baldi and Saren Brunak into Chinese, joint work with Zhang DH, Huang Y, Cai J, Sun YF, Xia HY, Hu CF, Zhu ZH.

Conferences

129. Mi SL, He FH, Wu J, Zhou J, Wu G, Thirman MJ, **Ji HK**, Wang QF (2010) Genome wide location analysis reveals deregulated microRNA genes in MLL rearragned leukemic genome. In the 52nd Annual Meeting of the American Society of Hematology (ASH), Orlando, FL. *Blood*, 116(21): 1039

130. Wang QJ, Wu G, Mi SL, He FH, Dong JF, Mattison RJ, Kaberlein J, Wang QB, Prabhakar S, **Ji HK**, Thirman MJ (2009) MLL fusion proteins directly regulate a small set of wild type MLL target genes. In the 51st Annual Meeting of the American Society of Hematology, New Orleans, LA. *Blood*, 114(22): 525

131. Vokes SA, **Ji HK**, Wong WH, McMahon AP (2006) A tale of two morphogen gradients: identifying Gli targets of hedgehog signaling. In the 65th Annual Meeting of the Society for Developmental Biology, Ann Arbor, MI. *Developmental Biology*, 295(1): 423

132. Vokes SA, **Ji HK**, Wong WH, McMahon AP (2005) Identification of Gli target genes using chromatin immuno-precipitation with a genetically inducible system on genomic arrays. In the 64th Annual Meeting of the Society for Development Biology, San Francisco, CA. *Developmental Biology*, 283(2): 666

Technical Reports

133. **Ji HK** and Wong WH (2005) Increasing power of microarray gene selection: an empirical Bayes approach. Harvard Qualifying Paper

PRACTICE ACTIVITIES

Software and Database

BIRD: Big data regression for predicting DNase I hypersensitivity using gene expression data.
https://github.com/WeiqiangZhou/BIRD

CisGenome: An integrated software system with graphic user interface for analyzing ChIP-seq and ChIP-chip data, DNA motif discovery, and genome annotations.
http://www.biostat.jhsph.edu/~hji/cisgenome/
110,000+ downloads, 700+ citations according to Google Scholar since 2008

ChIP-PED: Using public gene expression data in GEO to boost the value of ChIP-seq and ChIP-chip experiments.
http://www.biostat.jhsph.edu/~gewu/ChIPPED/

ChIPXpress [R/Bioconductor package]: Improved ChIP-seq and ChIP-chip target gene ranking using publicly available gene expression data.
http://www.biostat.jhsph.edu/~gewu/ChIPXpress/index.html
5,600+ downloads since 2012

Cormotif [R/Bioconductor package]: Integrative analysis of multiple gene expression datasets for simultaneously detecting differentially expressed genes and patterns.
http://www.bioconductor.org/packages/release/bioc/html/Cormotif.html
13,800+ downloads since 2012

dPCA: Software for analyzing differential protein-DNA interactions.
http://www.biostat.jhsph.edu/~hji/dpca/

GSCA: Gene set context analysis. Using publicly available gene expression data to find new biological contexts of gene set activities.
https://github.com/zji90/GSCA
6,900+ downloads since 2014

hmChIP: A database of public human and mouse ChIP-seq/ChIP-chip data.
http://jilab.biostat.jhsph.edu/database/cgi-bin/hmChIP.pl

iASeq [R/Bioconductor package]: Integrative analysis of allele-specificity in multiple ChIP-seq experiments.
http://bioconductor.org/packages/2.11/bioc/html/iASeq.html
12,700+ downloads since 2012

JAMIE: Joint analysis of multiple ChIP-chip experiments.
http://www.biostat.jhsph.edu/~hji/jamie/

Lamian: Differential pseudotime analysis with multiple single-cell RNA-seq samples.
https://github.com/Winnie09/Lamian

PDDB: Predicted DNase I hypersensitivity database.

http://jilab.biostat.jhsph.edu/~bsherwo2/bird/index.php

PolyaPeak: Detecting ChIP-seq peaks using peak shape information.
http://web1.sph.emory.edu/users/hwu30/software/polyaPeak.html

QFM: Quantitative fate mapping via retrospective lineage barcodes
https://github.com/Kalhor-Lab/QFM/

SCATE: Single-cell ATAC-seq signal extraction and enhancement
https://github.com/zji90/SCATE/

SCRAT: Single-cell regulome data analysis (single cell ATAC-seq, ChIP-seq, DNase-seq)
https://github.com/zji90/SCRAT

TileMap: Software for analyzing whole genome tiling array data.
http://jilab.biostat.jhsph.edu/software/tilemap/index.htm
270+ citations according to Google Scholar since 2005

TileProbe: Software for removing tiling array probe effects using public data in GEO
http://www.biostat.jhsph.edu/~hji/cisgenome/index_files/tileprobe.htm

TreeCorTreat: Tree-based correlation screen and visualization for exploring phenotype-cell type association with multiple single-cell RNA-seq samples
https://github.com/byzhang23/TreeCorTreat

TSCAN [R/Bioconductor package]: Software tool for single cell RNA-seq analysis
http://www.bioconductor.org/packages/release/bioc/html/TSCAN.html
5,600+ downloads since 2014

Patent

CN200310122497.6: Reagents and primers for predicting susceptibility to paranoid schizophrenia, by Xu Q, Shen Y, **Ji HK**, Ji L, Yuan YB, Shen YC (China).

PART II

TEACHING

Academic Advisees

(Postdoctoral fellows)

Wenpin Hou, PhD Co-mentor: Stephanie Hicks Current Position: Assistant Professor, Columbia University	2019-2022	Biostatistics
Benjamin Sherwood, PhD Current Position: Assistant Professor, University of Kansas	2014-2016	Biostatistics
Weiqiang Zhou, PhD Current Position: Assistant Scientist, Johns Hopkins University	2013-2018	Biostatistics
Fang Du, PhD Current Position: Senior Data Scientist at Covance/LabCorp	2013-2014	Biostatistics
Xuekui Zhang, PhD Current Position: Assistant Professor, University of Victoria, Canada	2011-2013	Biostatistics

(PhD students)

Shuai Li, PhD candidate Co-advisor: Ni Zhao	2022-present	Biostatistics
Yi Man, PhD candidate	2022-present	Biostatistics
Wanlu Chen, PhD candidate	2022-present	Biostatistics
Wentao Zhan, PhD candidate Co-advisor: Abhirup Datta	2021-present	Biostatistics
Jordan Wilson, PhD candidate Co-advisor: Jennifer Elisseeff	2020-present	XDBio
Yi Wang, PhD candidate Co-advisor: Kasper Hansen	2020-present	Biostatistics
Ruzhang Zhao, PhD candidate Co-advisor: Nilanjan Chatterjee	2020-present	Biostatistics
Runzhe Li, PhD candidate Co-advisor: Ni Zhao	2019-present	Biostatistics

Boyang Zhang, PhD 2017-2022 Biostatistics
Thesis: Statistical and computational methods for analyzing single-cell RNA-seq and immune profiling data
Award: 2016 Kocherlakota Award
Current Position: Postdoctoral fellow at Kundaje and Engelhardt Labs, Stanford University

Weixiang Fang, PhD 2016-2022 Biostatistics
Thesis: Quantitative methods for genomics and lineage tracing
Award: 2015 Kocherlakota Award, 2017 award for outstanding performance in PhD comprehensive exam
Current Position: Postdoctoral fellow at Kalhor Lab, Johns Hopkins University

Zhicheng Ji, PhD 2015-2020 Biostatistics
Thesis: Statistical methods for decoding gene regulation in single cells
Awards: 2014 Kocherlakota Award, 2016 ASA SGG Distinguished Student Paper Award, 2019 Margaret Merrell Award
Current Position: Assistant Professor, Duke University

Dan Jiang, PhD 2013-2017 Biostatistics
Thesis: Adjustment procedure to permutation tests in epigenomic differential analysis
Current Position: Machine Learning Scientist, LinkedIn

Bing He, PhD 2013-2017 Biostatistics
Thesis: FCAT: flexible classification-based signal detection in high-throughput sequencing data
Award: 2012 Kocherlakota Award
Current Position: Quantitative Researcher, Vatic Labs

Yingying Wei, PhD 2011-2014 Biostatistics
Thesis: Integrative statistical models for genomic signal detection
Awards: 2012 Culley Award, 2013 ENAR Student Travel Award, 2013 Margaret Merrell Award
Current Position: Associate Professor, Chinese University of Hong Kong

George Wu, PhD 2007-2013 Biostatistics
Thesis: Statistical methods for the integrative analysis of high-throughput genomic data
Current Position: Associate Director, Gilead Sciences

Yang Ning, PhD 2011-2012 Biostatistics
Thesis: Nonstandard likelihood based inference
Co-advisor: Kung-Yee Liang
Current Position: Assistant Professor, Cornell University

Hao Wu, PhD 2007-2010 Biostatistics
Thesis: Three statistical applications in genomics: redefining CpG islands, peak detection from multiple ChIP-chips, and data normalization for second generation sequencing.
Co-advisor: Rafael Irizarry

Award: 2010 Margaret Merrell Award
Current Position: Associate Professor, Emory University

Yuting Xu, PhD (Academic advisor) 2012-2013 Biostatistics

Fang Han, PhD (Academic advisor) 2010-2011 Biostatistics

Jeongyong Kim, PhD (Academic advisor) 2009-2011 Biostatistics

(Master's students)

Zhirui Fu, ScM candidate 2022-present Biostatistics

Ziqi Fu, ScM candidate 2022-present Biostatistics
Co-advisor: Weiqiang Zhou

Richard Xu, ScM candidate 2021-present Biostatistics

Meihan Wei, MSE candidate 2021-2022 BME

Brian Guo, ScM candidate 2021-2022 Biostatistics
Thesis: Subset-based cross-tissue transcriptome-wide association study
Co-advisor: Nilanjan Chatterjee, Eneda Toska
Current Position: PhD student, University of South California

Tingchang Wang, ScM candidate 2021-2022 Biostatistics
Thesis: CUT&RUN data analysis: gene expression discovery in ciliopathy mice

Shuai Li, ScM transferred to PhD 2020-2022 Biostatistics

Yimin Zhao, ScM candidate 2020-2022 Biostatistics
Thesis: A statistical framework for analyzing single cell multi-modal omics data: imputation and prediction
Current Position: PhD student, Biostatistics, University of Washington

Chaoran Chen, MSE 2020-2021 BME
Current Position: WuXi Biologics

Wanlu Chen, ScM transferred to PhD 2019-2021 Biostatistics

Yueyi Li, ScM 2019-2020 Biostatistics
Thesis: scBIRD, A Single-Cell Regulome Prediction Package with R
Current Position: Foundation Medicine

Luqin Gan, ScM 2017-2019 Biostatistics
Thesis: Integrated DNase I hypersensitivity prediction using RNA-seq and unmatched public DNase-seq
Current Position: PhD student, Statistics, Rice University

Chang (April) Shu, MHS 2016-2018 Biostatistics
Current Position: Postdoc, Columbia University

Gege Gui, ScM 2016-2018 Biostatistics
Thesis: High-throughput sequencing data analysis for amyotrophic lateral sclerosis patient
Current Position: NIH/PhD student, Biostatistics, Johns Hopkins University

Yifan Zhou, ScM 2015-2017 Biostatistics
Thesis: Statistical methods to detect hierarchical topological domains in chromatin
Current Position: PhD student, Statistics, George Washington University

Boyang Zhang, ScM transferred to PhD 2015-2017 Biostatistics

Yun Qing, MHS 2015-2016 Public Health
Current Position: Biostatistician, MD Anderson Cancer Center

Ding Ding, ScM 2015-2017 Biostatistics
Thesis: Spanning tree representation of high dimensional data
Current Position: Residency (medicine), Northwell Health

Weixiang Fang, ScM transferred to PhD 2014-2016 Biostatistics

Zhicheng Ji, ScM transferred to PhD 2013-2015 Biostatistics

Shilu Zhang, ScM 2013-2014 Biostatistics
Thesis: Transcription factor binding motif analyses in two biological systems
Award: 2013 Kocherlakota Award
Current Position: PhD student, Systems Biology, Univ. of Wisconsin at Madison

Zheng Kuang, MHS 2012-2014 Public Health
Current Position: Assistant Professor, Carnegie Mellon University

Wei Wang, MHS 2011-2013 Public Health

Jiehuan Sun, ScM 2011-2012 Biostatistics
Thesis: Systematic annotation and exploration of Gene Expression Omnibus
Award: 2011 Kocherlakota Award
Current Position: Assistant Professor, University of Illinois at Chicago

Cheng Ran Huang, MHS 2010-2013 Public Health
Current Position: Director of Corporate Development, Praxis Precision Medicines

Li Chen, MHS-Bioinformatics 2009-2011 Biostatistics
Project: hmChIP: a database and web server for exploring publicly available human and mouse ChIP-seq and ChIP-chip data
Current Position: Assistant Professor, Indiana University

(Undergraduate students)

Kyle Reese	2022-present	CS
Zeqing Li	2021-present	Biophysics
Richard Xu	2020-2021	CS

Capstone Advisees

Fatima Zaman, MPH candidate	2021-2022	Public Health

Other Students Worked on Projects

Yu-chi Chang	2016	Computer Science, Undergraduate
Ashwini Patil	2013-2014	Bioinformatics
Jiawei Bai	2012	Biostatistics
Luolan Li	2012	Visiting student from Nanjing University
Xueying Ji	2009	Geography and Environmental Engineering
Jennifer T. Judy	2008-2009	Mental Health
Thai Le	2007-2008	Biomedical Engineering and Mathematics

Master's Thesis Reader

Lai Wei	2022	Biochemistry and Molecular Biology
Mengyu He	2021	Biostatistics
Jingyi Lu	2015	Biomedical Engineering
Michael Ayars	2010	Molecular Microbiology and Immunology

Preliminary Oral Exam Participation
(*: I am an alternate)

Jordan Wilson	04/05/2022	XDBio
Yi Wang	01/10/2022	Biostatistics
Ruzhang Zhao	12/01/2021	Biostatistics
Runzhe Li	12/18/2020	Biostatistics
Jingning Zhang*	12/10/2020	Biostatistics
Jingyi Zhai	11/17/2020	Biostatistics, University of Michigan
Yuan He	11/06/2020	Biomedical Engineering
Taibo Li	08/18/2020	MD/PhD, School of Medicine
Boyang Zhang	05/01/2020	Biostatistics
Roshan V. Chikarmane	04/22/2020	Pharmacology and Molecular Sciences
Athena Chen*	11/05/2019	Biostatistics
Weixiang Fang	05/29/2019	Biostatistics
Zhicheng Ji	02/02/2018	Biostatistics
Guanghao Qi*	11/14/2017	Biostatistics
Yuelong Ji	06/07/2017	Population, Family and Reproductive

Chang (April) Shu	05/02/2017	Mental Health
Stephen Cristiano	03/27/2017	Biostatistics
Dan Jiang	05/05/2015	Biostatistics
Bing He	04/24/2015	Biostatistics
Leonardo Collado Torres	05/01/2014	Biostatistics
Suzanne Martos	04/23/2014	Environmental Health Sciences
Man Li	12/20/2013	Epidemiology
Taeyoung Hwang	11/09/2012	Biomedical Engineering
Yingying Wei	12/14/2011	Biostatistics
Yiyi Zhang*	01/28/2011	Epidemiology
George Wu	09/03/2010	Biostatistics
Simina Boca	12/10/2008	Biostatistics
Ji Li	06/09/2008	International Health
Hao Wu	01/10/2008	Biostatistics

Doctoral Thesis Committee

Sydney Connor	2022-present	Biomedical Engineering
Andy Kwan Pui Chan	2022-present	Molecular Microbiology and Immunology
Eric Kernfeld	2022-present	Biomedical Engineering
Elizabeth Christine Lee	2022-present	Molecular Microbiology and Immunology
Yi Wang	2021-present	Biostatistics
Ruzhang Zhao	2021-present	Biostatistics
Benjamin H. Nguyen	2021-present	Molecular Microbiology and Immunology
Yuqi Fang	2020-2022	Biomedical Engineering
Jingyi Zhai	2020-2022	Biostatistics (Univ. of Michigan)
Boyang Zhang	2020-2022	Biostatistics
Roshan Chikarmane	2020-2022	Pharmacology and Molecular Sciences
Yuan He	2020-2021	Biomedical Engineering
Vy Tran	2020-2020	Environmental Health and Engineering
Eduardo Martinez-Montes	2019-present	Biochemistry, Cellular & Molecular Biol.
Weixiang Fang	2019-2022	Biostatistics
Danielle Bouchard	2019-2020	Biochemistry and Molecular Biology
Claire Bell	2018-2022	Human Genetics and Molecular Biology
Zhicheng Ji	2018-2020	Biostatistics
Yuqi Tan	2017-2021	Biochemistry, Cellular & Molecular Biol.
Chang (April) Shu	2017-2018	Mental Health
Kipper Fletez-Brant	2016-2018	Human Genetics
Dan Jiang	2015-2017	Biostatistics
Bing He	2015-2017	Biostatistics
Suzanne Martos	2015-2017	Environmental Health Sciences
Amanda Price	2014-2018	Human Genetics
Alexandra Maertens	2013-2014	Environmental Health Sciences
Zheng Kuang	2012-2014	Human Genetics
Yingying Wei	2011-2014	Biostatistics
David Gorkin	2011-2013	Human Genetics

Rahul Karnik	2011-2012	Biomedical Engineering
Yang Ning	2011-2012	Biostatistics
George Wu	2010-2013	Biostatistics
Simina Boca	2010-2011	Biostatistics
Hao Wu	2008-2010	Biostatistics

Final Oral Exam Participation

(*: I am an alternate)

Yuqi Fang	12/07/2022	Biomedical Engineering
Jingyi Zhai	10/11/2022	Biostatistics (Univ. of Michigan)
Boyang Zhang	06/29/2022	Biostatistics
Weixiang Fang	04/13/2022	Biostatistics
Yuan He	07/15/2021	Biomedical Engineering
Danielle Bouchard	11/19/2020	Biochemistry and Molecular Biology
Zhicheng Ji	03/27/2020	Biostatistics
Anita Panjwani	05/15/2019	International Health
Stephen Cristiano	05/15/2019	Biostatistics
Ying Tong	02/03/2019	Bioinformatics, University of Hong Kong
Lu Li	11/27/2018	Biostatistics
Chang (April) Shu	03/06/2018	Mental Health
Dan Jiang	08/25/2017	Biostatistics
Suzanne Martos	08/09/2017	Environmental Health Sciences
Bing He	02/08/2017	Biostatistics
Leonardo Collado Torres*	06/21/2016	Biostatistics
Hailun Liang*	03/01/2016	Health Policy and Management
Jean-Philippe Fortin*	02/22/2016	Biostatistics
Alexandra Maertens	09/02/2014	Environmental Health Sciences
Yingying Wei	03/14/2014	Biostatistics
George Wu	05/10/2013	Biostatistics
Rahul Karnik	08/08/2012	Biomedical Engineering
Yang Ning	04/10/2012	Biostatistics
Simina Boca*	03/08/2011	Biostatistics
Tao Wu*	11/24/2010	Epidemiology
Hao Wu	05/19/2010	Biostatistics

PhD Thesis Reader

Jiehuan Sun	03/2017	Biostatistics, Yale University

Classroom Instruction

Advanced Methods in Biostatistics III (140.753) Principal Instructor, 2015-present

2022 Enrollment	13
2021 Enrollment	9

2020 Enrollment	11	
2019 Enrollment	11	
2018 Enrollment	10	
2017 Enrollment	10	
2016 Enrollment	13	
2015 Enrollment	9	
Advanced Methods in Biostatistics IV (140.754)		Principal Instructor, 2015-present
2022 Enrollment	13	
2021 Enrollment	9	
2020 Enrollment	11	
2019 Enrollment	11	
2018 Enrollment	10	
2017 Enrollment	10	
2016 Enrollment	13	
2015 Enrollment	8	
Methods in Biostatistics III (140.653)		Principal Instructor, 2013-2014
2014 Enrollment	41	
2013 Enrollment	32	
Methods in Biostatistics IV (140.654)		Principal Instructor, 2013-2014
2014 Enrollment	36	
2013 Enrollment	35	
Advanced Statistical Computing (140.778)		Principal Instructor, 2008-2011
2011 Enrollment	10	
2010 Enrollment	20	
2009 Enrollment	5	
2008 Enrollment	5	
Special Topics in Genetics and Genomics (140.668)		Principal Instructor, 2009
2008 Enrollment	6	
Statistical Computing (140.776)		Principal Instructor, 2009-2011
2011 Enrollment	38	
2010 Enrollment	41	
2009 Enrollment	34	
Advanced Special Topics in Biostatistics (140.850)		Principal Instructor, 2020
2020 Enrollment	3	
Advanced Statistical Theory II (140.772)		Co-instructor, 2007
Statistics for Genomics (140.688)		Guest Lectures, 2010,2012,2014

Genomics (260.605)	Guest Lectures, 2009

Visiting Scholar

Weiqing Huang	02/2019-08/2019	Tsinghua University
Hongqiang Lyu	11/2016-10/2017	Xi’an Jiaotong Univ.
Yu Gu	06/2017-09/2017	University of Science and Technology of China
Runzhe Li	06/2017-09/2017	Tsinghua University
Zhanhao Peng	07/2015-09/2015	Tsinghua University
Haomin Peng	07/2015-09/2015	Tsinghua University

Others

1. Advised Steffen Cornwell, a high school student, for the Center Scholar Program, Center for Talented Youth 06/2014 - 07/2014

2. Advised Alexander Smith, a high school student, for the Center Scholar Program, Center for Talented Youth 07/2012 - 08/2012

3. Taught a five-day summer course *Statistical Computing* for the Hopkins-Nanjing Exchange Program in Statistical and Data Sciences 07/2011

4. Advised Rachel Napthal from Weston High School, MA for the Center Scholar Program, Center for Talented Youth 06/2009 - 08/2009

5. Offered online tutorial session “Introduction to CisGenome” at the *Affymetrix Workshop at EMBL*. 11/2007

RESEARCH GRANT PARTICIPATION

Current

(Principal Investigator (PI))

Title of Grant: Computational Tools for Regulome Mapping Using Single-cell Genomic Data
Principal Investigator: Ji, HK
Dates: 08/22/2019-06/30/2023
Sponsoring Agency: NIH/NHGRI
Grant No.: R01HG010889
Award Amount: $1,637,500 (Direct cost: $1,000,000)

Main Grant Objective: The objective of this proposal is to develop computational and statistical methods and software tools for mapping and analyzing gene regulatory landscape using single-cell genomic data. In Aim 1, we will develop a method to accurately measure CRE activities in single cells using single-cell regulome data. In Aim 2, we will develop a method that uses single-cell RNA sequencing data to predict cells' regulatory landscape. In Aim 3, we will develop a method to align single-cell RNA-seq and single-cell regulome data to generate an integrated map of transcriptome and regulome.
Effort Allocation: 3.6 calendar months (30%)

Title of Grant: Big Data Methods for Decoding Gene Regulation
Principal Investigator: Ji, HK
Dates: 08/10/2018-05/31/2022
Sponsoring Agency: NIH/NHGRI
Grant No.: R01HG009518
Award Amount: $1,700,624 (Direct cost: $1,000,000)
Main Grant Objective: The goal of this study is to develop analytical methods that use massive amounts of publicly available gene expression microarray, exon array and bulk RNA-seq data to map and analyze human gene regulatory programs. We will develop a tree-guided big data regression approach for inferring a biological sample’s bulk regulatory profile using its bulk gene expression data (Aim 1). We will apply our method to 290,000+ human bulk gene expression samples in GEO to create a comprehensive regulatory map in diverse sample types (Aim 2). We will also develop a regularized spanning tree approach to help researchers explore this massive dataset by projecting samples to a low-dimensional lineage structure (Aim 3).
Effort Allocation: 1.8 calendar months (15%)

Title of Grant: Immune Development Across the Life Course: Integrating Exposures and Multi-Omics in the Boston Birth Cohort
Principal Investigators (MPI): Wang, XB; Ji, HK; Larman, HB
Dates: 09/15/2022-06/30/2027
Sponsoring Agency: NIH/NIEHS
Grant No.: U01ES034983
Award Amount: $822,244/year (Direct cost: $547,891/year)
Main Grant Objective: Our overarching goal is to investigate the impact of early life immune response to a broad array of pathogenic and commensal microbes and exposure to multiple environmental pollutants on the development and prognosis of allergic diseases from birth up to age 18 years and their underlying molecular pathways. By including 1,000 mother-child dyads of the BBC with key longitudinal data elements and biospecimens from birth up to age 18 years and harnessing cutting-edge antibody profiling technology (PhIP-Seq) to profile IgG and IgE antibodies in 1,000 BBC children at three important developmental windows (birth, 1-2 years, and 15-18 years), we aim to investigate: (1) effects of early life immune response to microbes on child allergic phenotypes; (2) effects of early life exposure to environmental pollutants (e.g., air pollution, toxic metals) on child allergic phenotypes; and (3) molecular pathways underlying the link between early life environmental exposures and allergic diseases.
Effort Allocation: 1.8 calendar months (15%)

(Subcontract PI)

Title of Grant: Cis-regulatory Circuitry Underlying Hedgehog Mediated Limb Development
Principal Investigator: Vokes, SA
Dates: 07/01/2012- 06/30/2023
Sponsoring Agency: NIH/NICHD
Grant No.: R01HD073151
Main Grant Objective: Combine high-throughput genomic technologies with transgenic experiments to study gene regulatory programs in the hedgehog signaling pathway and limb development
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 0.36 calendar months (3%)

Title of Grant: Molecular Genetics of Mammalian Larynx and Vocal Fold Development
Principal Investigator: Wallingford, JB/Vokes, SA
Dates: 07/10/2017- 04/30/2022
Sponsoring Agency: NIH/NICHD
Grant No.: R01HD090163
Main Grant Objective: This proposal seeks to fill a fundamental gap in our understanding of mammalian developmental biology by exploring the molecular genetics and morphogenesis of the mammalian larynx and vocal folds.
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 0.36 calendar months (3%)

(Co-investigator)

Title of Grant: Mechanisms of spontaneous and vaccine mediated hepatitis C virus control to direct rational development of a novel HCV vaccine
Principal Investigator: Cox, AL
Dates: 05/01/2021 - 04/30/2026
Sponsoring Agency: NIH/NIAID
Grant No.: U19AI159822
Main Grant Objective: Persistent HCV is a common cause of liver failure and death. There was limited effect of the only vaccine tested for prevention of HCV infection. We will define the successful immune responses observed in people infected with HCV many times who never become persistently infected and compare them to those seen in the only human vaccine efficacy trial and in vaccine models in order to design a more effective HCV vaccine.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.24 calendar months (2%)

Title of Grant: Immunogenomic determinants of response and resistance to neoadjuvant anti-PD-1 in resectable NSCLC

Principal Investigator: Smith, KN
Dates: 03/01/2021 - 02/28/2026
Sponsoring Agency: NIH/NCI
Grant No.: R37CA251447
Main Grant Objective: The major goal of this project is to leverage biospecimens obtained from an existing clinical trial evaluating safety and feasibility of neoadjuvant PD-1 blockade in resectable lung cancer to identify biomarkers for response and targetable molecules/pathways that are associated with non-response as a guide for development of future therapies.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Deciphering Mechanisms of HIV Latency Reversal in Perinatal Infections
Principal Investigator: Persaud, D
Dates: 12/01/2020 - 08/31/2024
Sponsoring Agency: NIH/NIAID
Grant No.: R01AI150412
Main Grant Objective: Following up on our preliminary observations that the LR may be more resistant to reactivation in pediatric populations, this application will examine, ex vivo, the latency reversal potential of the different classes of latency reversal agents in pediatric compared with adult HIV infections. We will examine effects of HIV subtype and geographic region of residence, along with assessment of baseline CD4+ T cells transcriptomic profiles that may reflect permissivity to virus reactivation with LRAs. We will also conduct exploratory studies of novel single-cell RNA-seq approaches to enhance resolution of various CD4+ T cell subsets to distinct classes of LRAs, with direct implications for clinical trial design and potential for biomarker discovery for ART-free remission and cure in pediatric populations.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Understanding Antitumor T cell Immunity in High Risk Localized Prostate Cancer after Neoadjuvant Treatment with an Antibody Targeting the B7-H3 Checkpoint
Principal Investigator: Shenderov, E
Dates: 08/01/2019 - 07/31/2023
Sponsoring Agency: DOD
Grant No.: W81XWH-19-1-0511
Main Grant Objective: To perform biological and immunological characterization of how B7-H3 targeted therapy affects outcomes in men with high-risk prostate cancer using samples from a phase II clinical study (NCT02923180).
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.12 calendar months (1%)

Title of Grant: Data Center for Acute to Chronic Pain Biosignatures
Principal Investigator: Lindquist, M
Dates: 09/01/2019 - 07/31/2023
Sponsoring Agency: NIH/NIDA

Grant No.: U54DA049110
Main Grant Objective: As part of the NIH A2CPS Program, we will establish a Data Integration and Resource Center (DIRC) to integrate imaging, peripheral physiology, omics, behavior, and clinical measures to study chronic pain.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.48 calendar months (4%)

Title of Grant: Epigenetic Drivers of Cancer Progression
Principal Investigator: Feinberg, AP
Dates: 07/15/2018 – 06/30/2023
Sponsoring Agency: NIH/NCI
Grant No.: 2R01CA054358
Main Grant Objective: The goal of the project is to identify epigenetic drivers of tumor progression, focusing on pancreatic cancer, identifying disruptions in DNA sequences and trans-acting regulatory factors that cause epigenetic instability.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 0.48 calendar months (4%)

Title of Grant: Johns Hopkins Excellence in Pathogenesis and Immunity Center for SARS-CoV-2 (JH-EPICS)
Principal Investigator: Klein, S; Cox AL
Dates: 09/30/2020 - 08/31/2022
Sponsoring Agency: NIH/NCI
Grant No.: U54CA260492
Main Grant Objective: We propose development of a Serological Sciences Center of Excellence: the Johns Hopkins Excellence in Pathogenesis and Immunity Center for SARS-CoV-2 (JH-EPICS). The overarching goal of JH-EPICS is to distinguish immune responses that protect from those that cause pathology during infection.
Principal Responsibilities of Individual: co-lead Analysis Resource Core
Effort Allocation: No salary effort

Title of Grant: Maternal and Cord Blood Metabolome, Infant Feeding, and Development of Food Allergy in Prospective Birth Cohort
Principal Investigator: Hong, X
Dates: 07/01/2020 - 06/30/2022
Sponsoring Agency: NIH/NIAID
Grant No.: R21AI154233-01
Main Grant Objective: Food Allergy (FA) has emerged as a major clinical and public health problem worldwide. This project aims to investigate maternal metabolic profile (to reflect the in-utero metabolic environment) and cord blood metabolic profile (to reflect the fetal metabolic environment) in relation to FA development in the offspring, as well as the potential modification effects by infant feeding pattern and fetal genetic susceptibility, in a well-established U.S. urban low-income minority prospective birth cohort.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.60 calendar months (5%)

Title of Grant: Expanding Therapeutic Options for Lung Cancer (EXTOL) in memory of Waun Ki Hong
Principal Investigator: Pardoll, A
Dates: 08/01/2019 - 07/31/2022
Sponsoring Agency: MSKCC(Mark Foundation- Prime sponsor)
Grant No.: 19-028-ASP
Main Grant Objective: The major goal of subproject 2 is to leverage existing data and expand integrated research using biopsy specimens from NSCLC patients receiving commercial PD-1/PD-L1 inhibitors. Specifically, we will use new technologies to enhance understanding of tumor-immune cell interactions.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.22 calendar months (1.8%)

Completed

(Principal Investigator (PI))

Title of Grant: Harnessing antibody profiles and multi-omics data to delineate drivers of the allergic march
Principal Investigators: Larman B, Hong XM, Ji HK, Wang XB
Dates: 12/31/2020 - 12/31/2021
Sponsoring Agency: Johns Hopkins University
Grant No.: Johns Hopkins Discovery Award
Main Grant Objective: The project aims to use high-throughput antibody profiling to investigate immune system evolvement related to allergy development in early childhood
Principal Responsibilities of Individual: Co-PI
Effort Allocation: No salary effort

Title of Grant: Computational Tools for Mining Large Amounts of ChIP and Gene Expression Data
Principal Investigator: Ji, HK
Dates: 07/25/2012- 04/30/2018
Sponsoring Agency: NIH/NHGRI
Grant No.: R01HG006282
Award Amount: $ 2,035,000 (Direct cost: 1,250,000)
Main Grant Objective: Develop computational and statistical tools that integrate large amounts of publicly available gene expression data with ChIP-seq and ChIP-chip data to systematically map functional contexts of gene regulatory pathways.
Effort Allocation: 3.0 calendar months (25%)

Title of Grant: Comparative Functional Genomics Based on Massive Amounts of Publicly Available Omics Data
Principal Investigator: Ji, HK
Dates: 08/2016-12/2017
Sponsoring Agency: Johns Hopkins Catalyst Award

Award Amount: $75,000
Main Grant Objective: Build a computational framework for comparative functional genomics.
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Statistical and Computational Tools for Next-generation ChIP-seq Applications
Principal Investigator: Ji, HK
Dates: 09/12/2012- 05/31/2016
Sponsoring Agency: NIH/NHGRI
Grant No.: R01HG006841
Award Amount: $972,000 (Direct cost: $600,000)
Main Grant Objective: Develop computational and statistical tools for analyzing genome-wide gene regulation dynamics and allele-specific binding using ChIP-seq data.
Effort Allocation: 1.8 calendar months (15%)

Title of Grant: Genome-wide Prediction of DNase I Hypersensitivity and Transcription Factor Binding Sites Based on Gene Expression
Principal Investigator: Ji, HK
Dates: 05/01/2015- 04/30/2016
Sponsoring Agency: JHU/Institute for Data Intensive Engineering and Science
Grant No.: NA
Award Amount: $ 25,000
Main Grant Objective: The objective of this proposal is to develop a data science approach to predict genome-wide DNase I hypersensitivity (DHS) and transcription factor binding sites (TFBSs) based on gene expression data.

Title of Grant: Global Prediction of Transcription Factor Binding Sites in Lineage Specific Neuronal Differentiation
Principal Investigator: Ji, HK
Dates: 07/01/2012- 06/29/2014
Sponsoring Agency: TEDCO/Maryland Stem Cell Research Fund
Grant No.: 2012-MSCRFE-0135-00
Award Amount: $ 230,000 (Direct cost: $200,000)
Main Grant Objective: We propose to develop computational methods to predict genome-wide transcription factor (TF) binding sites for many TFs and many biological conditions simultaneously based on exon array data. Applying this approach, we will study key regulatory programs in differentiation programs of human stem cell derived neuron progenitor cells.
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Decoding transcriptional regulatory programs in tumors dependent on hedgehog signaling
Principal Investigator: Ji, HK
Dates: 08/01/2007-08/01/2008

Sponsoring Agency: The Richard L. Gelb Cancer Research Fund, Johns Hopkins Bloomberg School of Public Health
Award Amount: $ 10,000
Main Grant Objective: Identify genes that are direct targets of Gli proteins in Hh induced medulloblastoma, BCC and RMS, by utilizing genome-wide chromatin immunoprecipitation (ChIP) analysis coupled with genomic tiling (chip) arrays in a novel mouse model of Hh-related tumors. Develop statistical methods to analyze tiling array data and to compare TF-DNA binding activities across tumor samples.

Title of Grant: Johns Hopkins – Nanjing Exchange Program in Statistical and Data Sciences
Principal Investigator: Ji, HK
Dates: 06/01/2011- 06/30/2013
Sponsoring Agency: JHU Benjamin and Rhea Yeung Center for Collaborative Chinese Studies
Award Amount: $ 150,000
Main Grant Objective: We propose to establish the Johns Hopkins (JHU) – Nanjing University (NJU) exchange program in statistical and data sciences. The program involves bi-directional educational and research exchanges between JHU and NJU, and we aim to promote new cross-institutional and interdisciplinary collaborations.

Title of Grant: Characterization of Novel Sequence Motifs Important for Stem Cell Biology
Principal Investigator: Zeller, KI; Ji, HK
Dates: 09/01/2010 – 08/30/2012
Sponsoring Agency: TEDCO/Maryland Stem Cell Research Fund
Grant No.: 2010-MSCRFE-0050-00
Main Grant Objective: Develop computational and experimental approaches to identify proteins that bind to novel DNA motifs in human ESC.
Principal Responsibilities of Individual: Co-PI
Effort Allocation: 1.2 calendar months (10%)

(Subcontract PI)

Title of Grant: Chromatin Modifications in GBM-propagating Cells
Principal Investigator: Laterra, J
Dates: 03/01/2017 – 02/28/2022
Sponsoring Agency: NIH
Grant No.: R01NS096754
Main Grant Objective: The goals of this proposal are to identify mechanisms that induce GBM cells to express a GSC state and to develop novel approaches to target these mechanisms in order to inhibit GBM growth and improve survival of GBM- bearing animals.
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 0.8 calendar months (6.7%)

Title of Grant: Highly Efficient Conversion of Human iPS Cells to Dopaminergic Neurons by Synthetic Modified mRNAs

Principal Investigator: Ying, MY
Dates: 07/01/2015 - 06/30/2017
Grant No.: KKI subcontract 613044
Main Grant Objective: Analyze high-throughput genomic data for studying efficient conversion of iPS cells to dopaminergic neurons.
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Brain Cancer Stem Cell Reprogramming by c-Met
Principal Investigator: Laterra, J
Dates: 05/15/2012 – 04/30/2017
Sponsoring Agency: NIH
Grant No.: R01NS073611
Main Grant Objective: This research will study how the c-Met tyrosine kinase and reprogramming transcription factors regulate glioblastoma stem cells.
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Modeling Aggressive Medulloblastoma using Human Induced Pluripotent Stem Cells
Principal Investigator: Ying, MY
Dates: 09/30/2015 - 09/29/2016
Grant No.: KKI subcontract 110866
Main Grant Objective: Computational analyses of the high-throughput genomic data for studying medulloblastoma using human iPS cells.
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Suppression of Glioblastoma Stem Cells by Kruppel-Like Factor 9
Principal Investigator: Laterra, J
Dates: 09/15/2011 – 07/31/2016
Sponsoring Agency: NIH
Grant No.: R01NS076759
Main Grant Objective: The research focuses on understanding glioblastoma stem cell regulation by the KLF9 transcription factor and to develop KLF9-based Brain cancer treatment strategies.
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Regulation of Neural and Neoplastic Stem Cells by Kruppel-Like Transcription Factors
Principal Investigator: Laterra, J
Dates: 07/01/2011 – 06/30/2014
Sponsoring Agency: TEDCO/Maryland Stem Cell Research Fund
Grant No.: 2011-MSCRFII-0073-00

Main Grant Objective: We will analyze ChIP-seq data to identify a KLF target genes in neural and neoplastic stem cells and will participate in follow-up genomic analyses.
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Highly efficient conversion of human stem cells to dopaminergic neurons by proneural transcription factor Atoh1
Principal Investigator: Ying, MY
Dates: 07/01/2012- 06/30/2014
Sponsoring Agency: TEDCO/Maryland Stem Cell Research Fund
Grant No.: TBA
Main Grant Objective: This project will use high-throughput genomic technologies to study reprogramming programs in human stem cells.
Principal Responsibilities of Individual: Subcontract PI
Effort Allocation: 0.6 calendar months (5%)

(Co-investigator)

Title of Grant: In Utero Exposure to Metals and Vitamins B on Placenta and Child Cardiometabolic Outcomes
Principal Investigator: Wang, GY
Dates: 07/15/2019 - 06/30/2021
Sponsoring Agency: NIH/NIEHS
Grant No.: R03ES029594
Main Grant Objective: This proposal investigates whether and to what extent maternal exposures to heavy metals and maternal B vitamin status affect placenta and child cardiometabolic outcomes during childhood.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.24 calendar months (2%)

Title of Grant: Preterm Birth, Maternal and Cord Blood Metabolome, and Child Metabolic Risk
Principal Investigator: Wang, XB/Hu, F
Dates: 05/03/2016 - 03/31/2021
Sponsoring Agency: NIH/NICHD
Grant No.: R01HD041702
Main Grant Objective: This proposal aims to investigate whether prematurity and maternal and fetal metabolic characteristics can jointly affect the future development of child adverse metabolic outcomes in a well-established U.S. urban low- income minority prospective birth cohort.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Early Life Determinants of Obesity in U.S. Urban Low Income Minority Birth Cohort

Principal Investigator: Wang, XB
Dates: 01/01/2016 - 12/31/2020
Sponsoring Agency: NIH
Grant No.: 1R01HD086013
Main Grant Objective: This project aims to identify early life risk and protective factors for obesity development.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Maternal Stress and Preterm Birth: Role of Genome and Epigenome
Principal Investigator: Hong, XM
Dates: 07/20/2018 – 06/30/2020
Sponsoring Agency: NIH/NICHD
Grant No.: R03HD096136
Main Grant Objective: The major goal of this project is to investigate 1) the interactions of maternal stress and maternal genome on risk of preterm birth (PTB); 2) the role of the maternal epigenome in mediating gene – stress interaction on PTB risk.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 0.24 calendar months (2%)

Title of Grant: Neoadjuvant Nivolumab, or Nivolumab in Combination with Ipilimumab, in Resectable Non-Small Cell Lung Cancer
Principal Investigator: Forde, P
Dates: 05/03/2017-05/02/2020
Sponsoring Agency: Bristol Myers Squibb Co
Grant No.: J1414
Main Grant Objective: The major goal of this project is to evaluate the safety, efficacy and potential biomarkers associated with neoadjuvant immune checkpoint blockade in early stage lung cancer.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 3.0 calendar months (25%)

Title of Grant: Exposure-Altered Gene Expression in Five Candidate Imprinted Loci for Adult Disease
Principal Investigator: Wang, ZB
Dates: 07/01/2016 - 03/31/2020
Sponsoring Agency: NIH/NIEHS
Grant No.: R01ES025761
Main Grant Objective: The goal of this project is to determine how toxicant Bisphenol A changes expression of genes within 17 novel candidate imprinted loci during mouse early embryonic development, therefore contributing to the pathogenesis of diseases in adulthood.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.24 calendar months (2%)

Title of Grant: Expression profiling of neoantigen-specific T cells in resectable NSCLC
Principal Investigator: Smith, KN

Dates: 07/01/2019 - 12/31/2019
Sponsoring Agency: SKCCC
Grant No.: Swim Across America
Main Grant Objective: The major goal of this project is to evaluate the pan-cancer tumor immune microenvironment and antigen recognition in patients with early stage disease.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.24 calendar months (2%)

Title of Grant: Center for Systems Biology of Retrotransposition
Principal Investigator: Boeke, JD
Dates: 09/05/2013 – 05/31/2019
Sponsoring Agency: NIH
Grant No.: 1P50GM107632
Main Grant Objective: Develop systems approaches to understand and manipulate the vital, highly coordinated relationship between host and transposons in targeted ways.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.12 calendar months (1%)

Title of Grant: Prenatal Multi-level Stressors and Alterations in Maternal and Fetal Epigenomes
Principal Investigator: Surkan, PJ/ Wang, XB
Dates: 07/01/2016 - 06/30/2018
Sponsoring Agency: NIH/NICHD
Grant No.: R21HD085556
Main Grant Objective: This proposed study will investigate a broad spectrum of maternal psychosocial stressors in relation to both mothers' and children's DNA methylation profiles at birth.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Genome-wide association study of preterm birth
Principal Investigator: Wang, XB
Dates: 01/01/2012 - 12/31/2015
Sponsoring Agency: NIH
Grant No.: R01HD041702
Main Grant Objective: The goal is to conduct genome-wide association study to identify susceptibility genes of preterm delivery in the Boston Birth Cohort.
Principal Responsibilities of Individual: Co-investigator
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: c-Myc Targets in the Pathogenesis of Human Cancers
Principal Investigator: Dang, CV
Dates: 07/01/2010 – 12/31/2014
Sponsoring Agency: NIH
Grant No.: R01CA57341
Main Grant Objective: Study roles of MYC oncogene in human cancers.

Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Oral Clefts: Moving from Genome Wide Studies toward Functional Genomics
Principal Investigator: Beaty, T
Dates: 09/01/2009-08/31/2014
Sponsoring Agency: NIH/NIDCR
Grant No.: U01-DE-092003
Main Grant Objective: This study aims to investigate genes/regions yielding evidence of containing causal genes from a genome wide study of oral clefts as part of the FaceBase consortium.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Analysis Tools and Software for Second Generation Sequencing Data
Principal Investigator: Irizarry, RA
Dates: 12/01/2009 – 11/30/2013
Sponsoring Agency: NIH
Grant No.: R01HG005220
Main Grant Objective: Develop statistical methods and bioinformatics tools for analyzing high-throughput sequencing data.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 2.4 calendar months (20%)

Title of Grant: Flow Cytometry Analysis of Lyme Diseases
Principal Investigator: Soloski, M; Aucott J
Dates: 09/01/2011-08/31/2013
Sponsoring Agency: Non-sponsored
Main Grant Objective: This study aims to use flow cytometry and high-throughput genomics technologies to investigate immune response to lyme diseases.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 0.6 calendar months (5%)

Title of Grant: Johns Hopkins Center in Urban Environmental Health
Principal Investigator: Groopman. J
Dates: 04/01/2008-03/31/2013
Sponsoring Agency: NIH/NIEHS
Grant No.: 2P30ES003819
Main Grant Objective: The Biostatistics Center consults with environmental health scientists including molecular toxicologists, physiologists, environmental engineers and epidemiologists.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Preprocessing and Analysis Tools for Contemporary Microarray Applications

Principal Investigator: Irizarry, RA
Dates: 09/24/2007 – 08/31/2012
Sponsoring Agency: NIH
Grant No.: R01GM083084
Main Grant Objective: Develop statistical methods and bioinformatics tools for analyzing high-density tiling array data.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Characterizing Blood Progenitor Cells Differentiated from Human iPS and ES cells
Principal Investigator: Cheng, LZ
Dates: 09/30/2009-08/31/2011
Sponsoring Agency: NIH
Grant No.: RC2 GO grant HL101582 - 01
Main Grant Objective: Using genomic and epigenetic analysis tools to characterize similarities and differences between embryonic stem cells and iPS cells.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Microarray Detection of Retrotransposons
Principal Investigator: Boeke, JD
Dates: 09/22/2008-09/21/2010
Sponsoring Agency: JHMI Microarray Core
Main Grant Objective: Develop statistical models for detecting active transposon elements in human genome.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 1.2 calendar months (10%)

Title of Grant: Bioinformatics and Resequencing in Mood Disorders
Principal Investigator: Potash, JB
Dates: 08/01/2008-07/31/2010
Sponsoring Agency: Buerger Fund for Bipolar Disorder
Main Grant Objective: Develop bioinformatics tools for analyzing massively parallel sequencing data for mood disorders.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 1.8 calendar months (15%)

Title of Grant: Center for the Epigenetics of Common Human Diseases
Principal Investigator: Feinberg, AP
Dates: 05/14/2007-04/30/2010
Sponsoring Agency: NIH/NHGRI
Grant No.: 5P50HG003233
Main Grant Objective: Develop high throughput tools for epigenome analysis. Develop s tatistical methods to analyze allele specific expression data.
Principal Responsibilities of Individual: Co-Investigator

Effort Allocation: 2.4 calendar months (20%)

Title of Grant: Epigenetic Variation and Its Determinants in Depression
Principal Investigator: Potash, JB
Dates: 08/01/2008-07/31/2009
Sponsoring Agency: NIH/NIMH
Grant No.: R01MH074131
Main Grant Objective: The goal of this project is to find and measure epigenetic marks in subjects with major depressive disorder and to test potential genetic and environmental factors influencing epigenetic variation.
Principal Responsibilities of Individual: Co-Investigator
Effort Allocation: 1.2 calendar months (10%)

ACADEMIC SERVICE

Department of Biostatistics

Director, Graduate Program & Admission Committee (2017 - present)
Co-director, Graduate Program & Admission Committee (2015 - 2017)
Member, Graduate Program & Admission Committee (2010 - present)
Member, Statistical Genomics Bloomberg Professor Search Committee (2013-2014)
Member, Merrell Award Committee (2011)
Member, Faculty Search Committee (2008)
Member, Postdoctoral Fellow Committee (2007)

Bloomberg School of Public Health

Faculty Senate (2011 - 2013)
Committee on Information Technology (2010)

University

Internal Advisory Committee (IAC) for Johns Hopkins Rheumatic Diseases Resource-based Core Center (2022 - present)

PRESENTATIONS

Scientific Meetings

(Invited Oral Presentations)

Multi-sample single-cell RNA-seq data analysis and visualization - methods, software, and benchmark. *2022 ICSA Applied Statistics Symposium, Gainesville, FL,* June 20, 2022

Cross-modality prediction and imputation of functional genome on single cells. *Banff workshop on "Deep Learning for Genetics, Genomics and Metagenomics: Latest developments and New Directions"*. June 7, 2022 [Online due to COVID-19 pandemic]

A statistical framework for differential pseudotime analysis with multiple single-cell RNA-seq samples. *Statistical Methods in Genetic/Genomic Studies Workshop, Singapore*, Jan 6, 2022 [Online due to COVID-19 pandemic]

A statistical framework for differential pseudotime analysis with multiple single-cell RNA-seq samples with an application to COVID-19. *Society of Chinese Bioscientists in America DC-Baltimore Chapter 2021 Annual Symposium*, Dec 4, 2021 [Online due to COVID-19 pandemic]

Statistical analysis of multi-sample single-cell RNA-sequencing data with applications to COVID-19. *JSM*, Aug 12, 2021 [Online due to COVID-19 pandemic]

An integrative analysis of multi-study and multi-sample COVID-19 single-cell RNA-seq data. *ENAR*, Mar 16, 2021 [Online due to COVID-19 pandemic]

Statistical analysis of coupled single-cell RNA-seq and immune profiling data. *JSM*, Aug 6, 2020 [Online due to COVID-19 pandemic]

RAISIN: Regression Analysis of Single Cell RNA-seq Data with Multiple Samples. *ENAR*, Mar 25, 2020 [Online due to COVID-19 pandemic]

RAISIN: Regression Analysis of Single Cell RNA-seq Data with Multiple Samples. *The 11th ICSA International Conference, Hangzhou, Zhejiang, China*, Dec 22, 2019

Alignment and integrative analysis of single-cell RNA-seq and single-cell ATAC-seq data. *JSM, Denver, CO,* July 30, 2019

Reconstructing gene regulatory dynamics along pseudotemporal trajectories using single-cell RNA-seq. *Banff workshop on "Frontiers in Single-cell Technology, Applications and Data Analysis", Banff, Alberta, Canada*, Feb 26, 2019

Single-cell ATAC-seq signal extraction and enhancement. *JSM, Vancouver, BC, Canada*, Aug 2, 2018

Single-cell ATAC-seq signal extraction and enhancement. *The 5th IMS-APRM meeting, Singapore*, June 26, 2018

Single-cell ATAC-seq signal extraction and enhancement. *The 6th Workshop on Biostatistics and Bioinformatics, Atlanta, GA*, May 5, 2018

Single-cell ATAC-seq signal extraction and enhancement. *ENAR, Atlanta, GA,* Mar 27, 2018

Global prediction of chromatin accessibility using RNA-Seq from small number of cells. *JSM, Baltimore, MD,* Aug 2, 2017

Global prediction of chromatin accessibility using RNA-Seq from small number of cells. *2017 IMS China International Conference on Statistics and Probability, Nanning, Guangxi, China,* June 30, 2017

Global prediction of chromatin accessibility using RNA-Seq from single cell and small number of cells. *The 26th ICSA Applied Statistics Symposium, Chicago, IL,* June 27, 2017

Decoding gene regulation using single cell genomic data. *The Graybill Conference XIV on Statistical Genomics and Genetics, Fort Collins CO,* June 6, 2017

Single cell RNA-seq analysis by spanning trees. *The 31st New England Statistics Symposium, Storrs, CT,* Apr 22, 2017

Global prediction of chromatin accessibility using RNA-Seq from small number of cells. *ENAR, Washington DC,* Mar 13, 2017

Pseudo-time Reconstruction and Evaluation in Single-cell RNA-seq Analysis. *The 10th ICSA International Conference, Shanghai, China,* Dec 19, 2016

Genome-wide prediction of DNase I hypersensitivity using gene expression. *The 10th Annual Symposium and Poster Session on Genomics and Bioinformatics, JHU, Baltimore, MD,* Oct 13, 2016

Epigenome isoform analysis with applications. *JSM, Chicago, IL,* Aug 3, 2016

Big data regression and prediction in functional genomics. *The 4th IMS Asia Pacific Rim Meeting, Hong Kong,* Jun 29, 2016

Big data regression for predicting genome-wide functional genomic signals. *The 25th ICSA Applied Statistics Symposium, Atlanta, GA,* Jun 13, 2016

Big data regression for predicting genome-wide regulatory element activities. *Johns Hopkins Institute for Data Intensive Engineering and Science (IDIES) Annual Symposium, Baltimore, MD,* Oct 16, 2015

A big data approach for integrating different functional genomic data types. *JSM, Seattle, WA,* Aug 9, 2015

Big data methods for dissecting variations in high-throughput genomic data. *Symposium on Advances in Genomics, Epidemiology, and Statistics (SAGES), Philadelphia, PA,* May 29, 2015

A big data method for decomposing variations in high-throughput genomic data. *ENAR, Miami, FL,* Mar 18, 2015

A big data method for decomposing variations in high-throughput genomic data. *The Third Joint Biostatistics Symposium, Chengdu, China*, June 28, 2014

A model-based approach for characterizing correlation in big genomic data. *The Annual Meeting of Statistical Society of Canada, Toronto, ON, Canada*, May 27, 2014

Big data methods for dissecting variations in high-throughput genomic data. *Tsinghua Chair Professors' Bioinformatics Workshop, Beijing, China*, May 23, 2014

Statistical analysis of time-course ChIP-seq data. *ENAR, Baltimore, MD*, Mar 17, 2014

The Future of Statistics in Biology. *The Simply Statistics Unconference on the Future of Statistics*, Oct 30, 2013.

ChIP-seq: moving from single dataset peak calling to multiple experiment data integration. *ENAR, Orlando, FL*, Mar 13, 2013

Predicting functional targets of transcription factors by coupling ChIP-seq with GEO. *JSM, San Diego, CA*, Aug 2, 2012

Differential Principal Component Analysis of ChIP-seq. *ENAR, Washington DC,* Apr 2, 2012

Integrating ChIP-seq data with other high dimension data types. Statistical analyses for next generation sequencing. *Birmingham, AL*, Sep 26-27, 2011

Differential Principal Component Analysis of ChIP-seq. *JSM, Miami, FL*, July 31, 2011

Differential Principal Component Analysis of ChIP-seq. *WNAR, San Luis Obispo, CA*, June 21, 2011

Functional interpretation of ChIP-seq using publicly available gene expression data. *ENAR, Miami, FL*, Mar 21, 2011

A latent mixture model for analyzing multiple related ChIP-chip and gene expression data sets. *ENAR, New Orleans, LA*, Mar 24, 2010

FlexModule: a flexible cis-regulatory module sampler. *IMS-China International Conference on Statistics and Probability, Weihai, Shandong, China,* July 4, 2009

A correlation motif based hidden Markov model for pooling information from multiple ChIP-chip experiments. *ENAR, San Antonio, TX,* Mar 17, 2009

CisGenome: an integrated tool for the analysis of ChIP-chip tiling array data. *The first North American ChIP-on-chip Scientific Meeting, Affymetrix, Boston, MA*, Sept 14, 2007

(Contributed Oral Presentations)

Efficient exploration of multiple ChIP-seq and ChIP-chip data sets. *JSM, Vancouver, BC, Canada*, Aug 4, 2010

(Posters)

TSCAN: pseudo-time reconstruction and evaluation in single cell RNA-seq data. *ENAR, Austin, TX,* Mar 6, 2016

Global prediction of chromatin accessibility using RNA-seq in samples with small number of cells. *CSHL Systems Biology: Global Regulation of Gene Expression Meeting, Cold Spring Harbor, NY,* Mar 16, 2016

Invited Seminars

Statistical analysis of single cell genomic data with multiple patient samples. *The Bloomberg~Kimmel Institute for Cancer Immunotherapy (BKI) Scientific Retreat, Baltimore, MD,* Sep 30, 2022

Identifying and Preventing Artifacts in High Dimensional Data: Computational Science in Immuno-Oncology Webinar. *Society for Immunotherapy of Cancer*, May 26, 2022

A statistical framework for differential pseudotime analysis with multiple single-cell RNA-seq samples. *The Division of Biostatistics at the Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN*, Mar 21, 2022 [Online due to COVID-19 pandemic]

Statistical analysis of single-cell RNA-seq data with multiple samples. *Dahshu Virtual Journal Club*, Jan 21, 2022

Unleash the potential of single-cell genomic analyses using publicly available bulk data. *Department of Biostatistics and Bioinformatics, Duke University, Durham, NC,* Dec 3, 2021 [Online due to COVID-19 pandemic]

Single cell genomic data science methods for mapping gene regulatory landscape. *Department of Statistics, University of Chicago, Chicago, IL*, Oct 26, 2020 [Online due to COVID-19 pandemic]

Single cell genomics, data science and public health. *The Johns Hopkins Bloomberg School of Public Health Dean's Lecture, Baltimore, MD,* Feb 19, 2020

Global prediction of gene regulatory landscape using bulk and single-cell RNA-seq, *Tsinghua University, Beijing, China*, Dec 16, 2019

Computational methods for integrating single-cell transcriptome and regulome data. Otto Warburg International Summer School and Research Symposium on Cell-type Heterogeneity and Single-cell Analysis, *CAS-MPG Partner Institute for Computational Biology (PICB), Shanghai, China*, Aug 30, 2019

Statistical and computational analysis of single-cell ATAC-seq data. Otto Warburg International Summer School and Research Symposium on Cell-type Heterogeneity and Single-cell Analysis, *CAS-MPG Partner Institute for Computational Biology (PICB), Shanghai, China*, Aug 29, 2019

Global prediction of gene regulatory landscape using bulk and single-cell RNA-seq. *Dahshu Virtual Journal Club*, May 3, 2019

Genome-wide prediction of gene regulatory landscape using gene expression. *School of Medicine and Dentistry, University of Rochester, NY*, Mar 22, 2019

Genome-wide prediction of DNase I hypersensitivity using gene expression. *Department of Biostatistics, University of Michigan, Ann Arbor, MI*, Oct 4, 2018

Genome-wide prediction of DNase I hypersensitivity using gene expression. *Department of Biostatistics, University of North Carolina at Chapel Hill, NC*, Aug 23, 2018

Big data regression for predicting functional genomic signals using gene expression. *University of Notre Dame, Notre Dame, IN*, Mar 20, 2017

Genome-wide prediction of DNase I hypersensitivity using gene expression. *Department of Statistics, University of Virginia, Charlottesville, VA*, Nov 11, 2016

Genome-wide prediction of DNase I hypersensitivity using gene expression. *MOFFITT Cancer Center, Tampa, FL*, Sept 8, 2016

Genome-wide prediction of DNase I hypersensitivity using gene expression. *UCLA, Los Angeles, CA*, May 16, 2016

Computational analysis of genome-wide chromatin immunoprecipitation data. *Brady Urological Institute, Johns Hopkins University, Baltimore, MD*, July 13, 2015

ChIP-seq: Unleashing its full potential through data integration. *University of South California, Los Angeles*, Sept 18, 2014

Crowd intelligence: a solution to new challenges in education? *The Institute of Economics of Education, Peking University, Beijing, China*, July 18, 2014

Differential principal component analysis of ChIP-seq. *Department of Statistics, Pennsylvania State University*, *State College, PA*, Oct 10, 2013

iASeq: integrative analysis of allele-specificity of protein-DNA interactions in multiple ChIP-seq datasets. *Beijing Institute of Genomics Chinese Academy of Sciences, Beijing, China,* July 18, 2013

ChIP-seq: Unleashing its full potential through data integration. *University of Maryland, College Park, MD,* June 6, 2013

ChIP-seq: Unleashing its full potential through data integration. *Department of Biostatistics & Epidemiology, University of Pennsylvania*, *Philadelphia, PA*, Apr 24, 2013

Differential Principal Component Analysis of ChIP-seq. *Department of Statistics, Yale University, New Haven CT,* Apr 16, 2012.

Differential Principal Component Analysis of ChIP-seq. *Department of Biostatistics, Bioinformatics, Georgetown University, Washington DC*, Mar 23, 2012.

Differential Principal Component Analysis of ChIP-seq. *Workshop II: Transcriptomics and Epigenomics, Institute for Pure & Applied Mathematics*, *University of California, Los Angeles, CA*, Oct 25-28, 2011.

Differential Principal Component Analysis of ChIP-seq. *Beijing Institute of Genomics Chinese Academy of Sciences, Beijing, China,* July 27, 2011.

Integrative analysis of many ChIP-seq and ChIP-chip experiments. *Department of Statistics, University of California, Berkeley, CA,* May 5, 2011.

Integrative modeling and analysis of multiple ChIP-chip experiments. *National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC*, Jan 18, 2011.

Integrating ChIP-seq and ChIP-chip with publicly available microarray data provides a new way to make discoveries. *Beijing Institute of Genomics Chinese Academy of Sciences, Beijing, China,* Jun 10, 2010.

Joint analysis of multiple genome-wide chromatin immunoprecipitation experiments. *Department of Statistics, University of British Columbia, Vancouver, BC, Canada*, Jun 3, 2010. (**Constance van Eeden Distinguished Speaker Seminar**)

Integrative analysis of multiple chromatin immunoprecipitation data from public domains. *Department of Biostatistics & Medical Informatics, University of Wisconsin, Madison, WI*, Apr 23, 2010.

Improving high-throughput data analysis by using gene expression omnibus. *Department of Biostatistics, Harvard University School of Public Health, Boston, MA*, Oct 19, 2009.

Decoding gene regulatory Program using ChIP-chip and ChIP-seq. *Department of Biology, The Johns Hopkins University, Baltimore, MD*, Oct 15, 2009.

Decoding gene regulatory Program using ChIP-chip and ChIP-seq. *Bioinformatics Division, Tsinghua National Lab of Information Science and Technology, Beijing, China*, July 27, 2009.

Decoding gene regulatory Program using ChIP-chip and ChIP-seq. *Beijing Institute of Genomics Chinese Academy of Sciences, Beijing, China,* July 8, 2009.

Decoding gene regulatory Program using ChIP-chip and ChIP-seq. *National Heart Lung and Blood Institute, NIH, Bethesda, MD*, Feb 5, 2009.

CisGenome: an integrated system for analyzing ChIP-chip and ChIP-seq data. *Department of Bioinformatics and Computational Biology, George Mason University, Manassas, VA*, Oct 21, 2008.

Detecting gene regulatory signals from genome-wide ChIP-chip and sequence data. *Department of Biostatistics & Epidemiology, University of Pennsylvania*, Oct 23, 2007.

Analysis of ChIP-chip tiling array data for mammalian transcription factors. *Affymetrix Northern California ChIP-on-chip Workshop, Berkeley, CA*, Jul 24, 2007.

Detecting gene regulatory signals from genome-wide ChIP-chip and sequence data. *Department of Statistics*, *Purdue University*, Mar 7, 2007.

Detecting gene regulatory signals from genome-wide ChIP-chip and sequence data. *Department of Biostatistics, University of North Carolina at Chapel Hill*, Mar 5, 2007.

Detecting gene regulatory signals from genome-wide ChIP-chip and sequence data. *Department of Statistics, North Carolina State University*, Feb 23, 2007.

Detecting gene regulatory signals from genome-wide ChIP-chip and sequence data. *Department of Biostatistics, Johns Hopkins University*, Feb 19, 2007.

Detecting gene regulatory signals from genome-wide ChIP-chip and sequence data. *Department of Biostatistics, University of Washington*, Feb 15, 2007.

Detecting gene regulatory signals from genome-wide ChIP-chip experiments and sequence data: towards developing a strategy with increased statistical power, *Department of Statistics, Pennsylvania State University*, Feb 6, 2007.

Detecting gene regulatory signals from genome-wide ChIP-chip and sequence data. *Department of Statistics, Texas A&M University*, Jan 30, 2007.

Statistical analysis of genome-wide chromatin immunoprecipitation data for mammalian transcription factors. *The Stanford Workshop in Biostatistics, Stanford, CA,* Nov 16, 2006.

CisGenome – an *in silico* lab for the analysis of ChIP-chip tiling array data. *The Affymetrix GeneChip Exon and Tiling Arrays Data Analysis Workshop, Cambridge, MA*, Mar 23, 2006.

ADDITIONAL INFORMATION

Personal statement of research and research objectives

I am interested in developing statistical and computational methods for analyzing big and complex data, particularly high-throughput genomic data and data in biomedical and health studies. I apply these tools to investigate basic biology, human disease, and public health. My current research topics include:

- ***Analytical methods and software tools for high-throughput genomic data*** such as transcriptome (RNA-seq, microarray), epigenome and regulome (ChIP-seq, DNase-seq, ATAC-seq) data. Examples of our work include CisGenome, TileMap, hmChIP, and BIRD.
- ***Single-cell genomics***. Examples include our methods for cells' pseudotemporal trajectory analysis (TSCAN, Lamian), single-cell ATAC-seq (SCATE, SCRAT), multi-sample differential expression analysis (TreeCorTreat, Lamian), and cross-data modality prediction (BIRD).
- ***Statistical and machine learning methods for big data and scalable data integration***. Examples include dPCA, CorMotif, and BIRD.
- ***Gene regulation***. Examples include our work on transcription factors (e.g. GLI, MYC, GATA6) and signaling pathways (e.g. sonic hedgehog pathway) in stem cell differentiation, development, and diseases.
- ***Immunoinformatics in cancer and infectious diseases***. Examples include our work on immune repertoire (T cell receptor sequencing) and single-cell omics analysis in cancers treated with immunotherapy, and in infectious diseases such as COVID-19.
- ***Public health data science***. Examples include our investigation of the impacts of genetic, epigenetic and environmental factors on human disease risks using the prospective Boston Birth Cohort (BBC).

My long-term goal is to build effective and efficient strategies, technologies, methods, and software tools for data generation, collection, analysis, integration and interpretation, and to transform data into knowledge that can improve human health.

Keywords

Big data, data science, statistics, machine learning, modeling, computing, integration, prediction, visualization, Bayesian, empirical Bayes, hierarchical model, latent variable, Markov Chain Monte Carlo, tree

Genomics, computational biology, bioinformatics, gene regulation, gene expression, single-cell genomics, functional genomics, epigenomics, immunoinformatics, high-throughput sequencing, ChIP-seq, DNase-seq, ATAC-seq, RNA-seq, TCR-seq, microarray, tiling array, ChIP-chip, sequence motif, stem cell, differentiation, development, cancer, infectious disease, COVID-19, birth cohort