W. Evan Johnson, Ph.D.

Division of Computational Biomedicine
Department of Medicine
Boston University School of Medicine
72 East Concord St., E-645
Boston, MA 02118

http://jlab.bu.edu/
phone: (617) 638-2541
fax: (617) 414-6947
email: wej@bu.edu
twitter: @wejlab

PROFESSIONAL EXPERIENCE

Assistant Professor of Medicine, Biostatistics, & Bioinformatics, Division of Computational Biomedicine, Department of Medicine, Boston University School of Medicine, 2011-Present
Adjunct Assistant Professor, Department of Oncological Sciences, University of Utah, 2008-Present
Assistant Professor, Department of Statistics, Brigham Young University, 2007-2011
Research Assistant, Biostatistics & Computational Biology, Dana Farber Cancer Institute, 2004-2007
Teaching Assistant, Department of Biostatistics, Harvard University, 2003-2007
Teaching Assistant, Department of Statistics, Brigham Young University, 2002-2003

EDUCATION

Ph.D., Biostatistics, Harvard University, Cambridge, Massachusetts; June 2007
Dissertation: *Statistical models for removing microarray batch effects and analyzing genome tiling arrays*
Advisors: X. Shirley Liu and Jun S. Liu
M.A., Biostatistics, Harvard University, Cambridge, Massachusetts, June 2006
M.S., Statistics, Brigham Young University, Provo, Utah; August 2003
B.S., Summa Cum Laude, Mathematics, Physics Minor, Southern Utah Univ., Cedar City, Utah; May 2002
A.A., Summa Cum Laude, Dixie College, St. George, Utah; May 2000

RESEARCH INTERESTS

Methodology: Bayesian methods, factor analysis and structural equations models, Hidden Markov models, dynamic programming, nonparametric regression, mixture models, high-performance and parallel computing, Bayesian networks

Applications: Applications in precision genomic medicine, clinical metagenomics, epigenomics, transcription regulation, next-generation sequencing and microarray analysis, and cancer research

EXTERNAL FUNDING

Ongoing Research Support

1R01 ES025002-01 (Johnson)	09/18/14 - 08/31/16	4.0 calendar
NIH/NIEHS	$219,600	

Title: Integrative analyses of reference epigenomic maps and applications
Description: We will develop approaches for data standardization across tissue types and profiling platforms and develop multiomic drug efficacy biomarkers for epigenetic drugs across multiple tissue or cell types.

5R01 HG005692-05 (Johnson)	06/01/10 - 05/31/15	2.64 calendar
NIH/NHGRI	$252,056	

Title: Statistical tools and methods for next-generation sequencing in epigenomics
Description: The goal of this study is to develop of statistical and computational tools for the analysis of second generation sequencing technologies with applications in epigenomics.

1U01 CA164720-03 (Johnson, Bild, Gray) 08/01/12 - 07/31/17 2.16 calendar
NIH/NCI $88,024 (sub only)
Title: Integrative growth signaling models to decipher complex cancer phenotypes
Description: This research aims to systematically profile the multi-tiered growth factor receptor networks (GFRNs) to investigate novel breast cancer phenotypes and the drugs that will be most effective against each subtype.

Janssen Pharmaceuticals, Inc. (Spira) 03/01/2013 - 02/28/2016 1.2 calendar
Title: The Premalignant Cancer Genome Atlas (PCGA) for Squamous Cell Lung Carcinoma
Description: The major goals of this project are to characterize genomic alterations associated with premalignant airway lesions to identify predictors of lesion progression and of lung cancer.

5R01 HL118542-02 (Lenburg) 08/15/13 - 05/31/17 0.6 calendar
NIH/NHLBI $500,221
Title: Integrative Omics to Discover Molecular Determinants of COPD
Description: We will discover molecular determinants of and novel therapies for COPD by integrating multiple levels of genomic data together with a multi-tiered in vitro validation strategy.

UC7 AI095321-01 (Murphy) 06/01/14 - 04/30/15 0.6 calendar
NIH/NIAID $10,775,850
Title: NEIDL Operations Grant
Description: The National Emerging Infectious Diseases Laboratories (NEIDL) Institutes mission is to perform cutting-edge basic and clinical research on emerging infectious diseases, including category A, B, and C agents.

1R01 AI113321-01 (Fearns) 06/01/14 - 04/30/15 0.03 calendar
NIH/NIAID $260,500
Title: Initiation and regulation of RSV mRNA transcription and genome replication.
Description: The aim of this project is to elucidate the functional and structural properties of the viral polymerase as a step towards designing antiviral drugs to control RSV disease.

Completed Projects

2 P01 CA073992-11 (Burt) 12/01/09 - 11/30/14 1.2 calendar
NIH/NCI $1,867,447
Title: Molecular and Clinical Approaches to Colon Cancer Precursors
Description: The overall objective of this Program Project Grant is to identify and test new ways to prevent, detect, and treat colon cancer through an increased understanding of the genetics, cell biology and pathogenesis of this malignancy and its precursor lesion, the adenomatous polyp.

PUBLICATIONS

Peer-Reviewed Articles

1. Hong C, Manimaran S, **Johnson WE**. PathoQC: Computationally efficient read pre-processing and quality control for high throughput sequencing datasets. *Cancer Informatics* (to appear).
2. Grubaugh ND, Sharma S, Krajacich BJ, Fakoli LS, Bolay FK, Diclaro JW, **Johnson WE**, Ebel GD, Foy BD, Brackney DE. Xenosurveillance: a novel mosquito-based approach for examining the human-pathogen landscape. *PLoS Neglected Tropical Diseases* 2015 Mar;9(3):e0003628. PubMed PMID: 25775236.
3. Castro-Nallar E, Hasan NA, Cebula TA, Colwell RR, Robison RA, **Johnson WE**, Crandall KA. Concordance and discordance of sequence survey methods for molecular epidemiology. *PeerJ* 3:e761.
4. Campbell JD, Liu G, Luo L, Xiao J, Gerrein J, Juan-Guardela B, Tedrow J, Aleksyev YO, Yang IV, Correll M, Geraci M, Quackenbush J, Sciurba F, Schwartz DA, Kaminski N, **Johnson WE**, Monti

S, Spira A, Beane J, Lenburg ME. Assessment of microRNA differential expression and detection in multiplexed small RNA sequencing data. *RNA*. 2015 Feb;21(2):164-71. PubMed PMID: 25519487.

5. Shen Y, Rahman M, Piccolo SR, Gusenleitner D, Ei-Chaar NN, Cheng L, Monti S, Bild AH, **Johnson WE**. ASSIGN: Context-specific Genomic Profiling of Multiple Heterogeneous Biological Pathways. *Bioinformatics*. 2015 Jan 22;PubMed PMID: 25617415.

6. Shamsaddini A, Pan Y, **Johnson WE**, Krampis K, Shcheglovitova M, Simonyan V, Zanne A, Mazumder R. Census-based rapid and accurate metagenome taxonomic profiling. *BMC Genomics*. 2014 Oct 21;15:918. PubMed PMID: 25336203; PubMed Central PMCID: PMC4218995.

7. Byrd AL, Perez-Rogers JF, Manimaran S, Castro-Nallar E, Toma I, McCaffrey T, Siegel M, Benson G, Crandall KA, **Johnson WE**. Clinical PathoScope: rapid alignment and filtration for accurate pathogen identification in clinical samples using unassembled sequencing data. *BMC Bioinformatics*. 2014 Aug 4;15:262. PubMed PMID: 25091138; PubMed Central PMCID: PMC4131054.

8. Hong C, Manimaran S, Shen Y, Perez-Rogers JF, Byrd AL, Castro-Nallar E, Crandall KA, **Johnson WE**. PathoScope 2.0: a complete computational framework for strain identification in environmental or clinical sequencing samples. *Microbiome*. 2014; 2:33. PubMed PMID: 25225611; PubMed Central PMCID: PMC4164323.

9. Bild AH, Chang JT, **Johnson WE**, Piccolo SR. A field guide to genomics research. *PLOS Biology*. 2014 Jan;12(1):e1001744. PubMed PMID: 24409093; PubMed Central PMCID: PMC3883637. (Joint corresponding author).

10. Hong C, Clement NL, Clement S, Hammoud SS, Carrell DT, Cairns BR, Snell Q, Clement MJ, **Johnson WE**. Probabilistic alignment leads to improved accuracy and read coverage for bisulfite sequencing data. *BMC Bioinformatics*. 2013 Nov 21;14:337. PubMed PMID: 24261665; PubMed Central PMCID: PMC3924334.

11. Piccolo SR, Withers MR, Francis OE, Bild AH, **Johnson WE**. Multiplatform single-sample estimates of transcriptional activation. *Proceedings of the National Academy of Sciences*. 2013 Oct 29;110(44):17778-83. PubMed PMID: 24128763; PubMed Central PMCID: PMC3816418.

12. Francis OE, Bendall M, Manimaran S, Hong C, Clement NL, Castro-Nallar E, Snell Q, Schaalje GB, Clement MJ, Crandall KA, **Johnson WE**. Pathoscope: species identification and strain attribution with unassembled sequencing data. *Genome Research*. 2013 Oct;23(10):1721-9. PubMed PMID: 23843222; PubMed Central PMCID: PMC3787268.

13. Cohen AL, Piccolo SR, Cheng L, Soldi R, Han B, **Johnson WE**, Bild AH. Genomic pathway analysis reveals that EZH2 and HDAC4 represent mutually exclusive epigenetic pathways across human cancers. *BMC Medical Genomics*. 2013 Sep 30;6:35. PubMed PMID: 24079712; PubMed Central PMCID: PMC3850967.

14. Woolstenhulme CJ, Parajuli S, Healey DW, Valverde DP, Petersen EN, Starosta AL, Guydosh NR, **Johnson WE**, Wilson DN, Buskirk AR. Nascent peptides that block protein synthesis in bacteria. *Proceedings of the National Academy of Sciences*. 2013 Mar 5;110(10):E878-87. PubMed PMID: 23431150; PubMed Central PMCID: PMC3593848.

15. O'Rawe J, Jiang T, Sun G, Wu Y, Wang W, Hu J, Bodily P, Tian L, Hakonarson H, **Johnson WE**, Wei Z, Wang K, Lyon GJ. Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. *Genome Medicine*. 2013;5(3):28. PubMed PMID: 23537139; PubMed Central PMCID: PMC3706896.

16. Piccolo SR, Sun Y, Campbell JD, Lenburg ME, Bild AH, **Johnson WE**. A single-sample microarray normalization method to facilitate personalized-medicine workflows. *Genomics*. 2012 Dec;100(6):337-44. PubMed PMID: 22959562; PubMed Central PMCID: PMC3508193.

17. Warf MB, Shepherd BA, **Johnson WE**, Bass BL. Effects of ADARs on small RNA processing pathways in C elegans. *Genome Research*. 2012 Aug;22(8):1488-98. PubMed PMID: 22673872; PubMed Central PMCID: PMC3409262.

18. Leek JT, **Johnson WE**, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012 Mar 15;28(6):882-3. PubMed PMID: 22257669; PubMed Central PMCID: PMC3307112.

19. **Johnson WE**, Welker NC, Bass BL. Dynamic linear model for the identification of miRNAs in next-generation sequencing data. *Biometrics*. 2011 Dec;67(4):1206-14. PubMed PMID: 21385162; PubMed Central PMCID: PMC3116054.

20. Cohen AL, Soldi R, Zhang H, Gustafson AM, Wilcox R, Welm BE, Chang JT, **Johnson WE**, Spira A, Jeffrey SS, Bild AH. A pharmacogenomic method for individualized prediction of drug sensitivity. *Molecular Systems Biology*. 2011 Jul 19;7:513. PubMed PMID: 21772261; PubMed Central PMCID: PMC3159972.

21. Rope AF, Wang K, Evjenth R, Xing J, Johnston JJ, Swensen JJ, **Johnson WE**, Moore B, Huff CD, Bird LM, Carey JC, Opitz JM, Stevens CA, Jiang T, Schank C, Fain HD, Robison R, Dalley B, Chin S, South ST, Pysher TJ, Jorde LB, Hakonarson H, Lillehaug JR, Biesecker LG, Yandell M, Arnesen T, Lyon GJ. Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. *American Journal of Human Genetics*. 2011 Jul 15;89(1):28-43. PubMed PMID: 21700266; PubMed Central PMCID: PMC3135802.

22. Lyon GJ, Jiang T, Van Wijk R, Wang W, Bodily PM, Xing J, Tian L, Robison RJ, Clement M, Lin Y, Zhang P, Liu Y, Moore B, Glessner JT, Elia J, Reimherr F, van Solinge WW, Yandell M, Hakonarson H, Wang J, **Johnson WE**, Wei Z, Wang K. Exome sequencing and unrelated findings in the context of complex disease research: ethical and clinical implications. *Discovery Medicine*. 2011 Jul;12(62):41-55. PubMed PMID: 21794208; PubMed Central PMCID: PMC3544941.

23. Warf MB, **Johnson WE**, Bass BL. Improved annotation of C elegans microRNAs by deep sequencing reveals structures associated with processing by Drosha and Dicer. *RNA*. 2011 Apr;17(4):563-77. PubMed PMID: 21307183; PubMed Central PMCID: PMC3062169.

24. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, **Johnson WE**, Geman D, Baggerly K, Irizarry RA. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics*. 2010 Oct;11(10):733-9. PubMed PMID: 20838408; PubMed Central PMCID: PMC3880143.

25. Rai K, Sarkar S, Broadbent TJ, Voas M, Grossmann KF, Nadauld LD, Dehghanizadeh S, Hagos FT, Li Y, Toth RK, Chidester S, Bahr TM, **Johnson WE**, Sklow B, Burt R, Cairns BR, Jones DA. DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. *Cell*. 2010 Sep 17;142(6):930-42. PubMed PMID: 20850014; PubMed Central PMCID: PMC2943938.

26. Thyagarajan B, Blaszczak AG, Chandler KJ, Watts JL, **Johnson WE**, Graves BJ. ETS-4 is a transcriptional regulator of life span in Caenorhabditis elegans. *PLOS Genetics*. 2010 Sep 16;6(9):e1001125. PubMed PMID: 20862312; PubMed Central PMCID: PMC2940738.

27. Clement NL, Snell Q, Clement MJ, Hollenhorst PC, Purwar J, Graves BJ, Cairns BR, **Johnson WE**. The GNUMAP algorithm: unbiased probabilistic mapping of oligonucleotides from next-generation sequencing. *Bioinformatics*. 2010 Jan 1;26(1):38-45. PubMed PMID: 19861355.

28. Hollenhorst PC, Chandler KJ, Poulsen RL, **Johnson WE**, Speck NA, Graves BJ. DNA specificity determinants associate with distinct transcription factor functions. *PLOS Genetics*. 2009 Dec;5(12):e1000778. PubMed PMID: 20019798; PubMed Central PMCID: PMC2787013.

29. **Johnson WE**, Liu JS, Liu XS. Doubly-stochastic Continuous-Time Hidden Markov Analysis of Genome Tiling Arrays. *Annals of Applied Statistics* 2009; 3:1183-1203.

30. Gottardo R, Li W, **Johnson WE**, Liu XS. A flexible and powerful bayesian hierarchical model for ChIP-Chip experiments. *Biometrics*. 2008 Jun;64(2):468-78. PubMed PMID: 17888037.

31. Song JS, **Johnson WE**, Zhu X, Zhang X, Li W, Manrai AK, Liu JS, Chen R, Liu XS. Model-based analysis of two-color arrays (MA2C). *Genome Biology*. 2007;8(8):R178. PubMed PMID: 17727723; PubMed Central PMCID: PMC2375008. (Joint first author)

32. **Johnson WE**, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007 Jan;8(1):118-27. PubMed PMID: 16632515.

33. **Johnson WE**, Li W, Meyer CA, Gottardo R, Carroll JS, Brown M, Liu XS. Model-based analysis of tiling-arrays for ChIP-chip. *Proceedings of the National Academy of Sciences*. 2006 Aug 15;103(33):12457-62. PubMed PMID: 16895995; PubMed Central PMCID: PMC1567901.

Peer-Reviewed Conference Proceedings

1. Price J, Clement M, Snell Q, **Johnson WE**. Identification and correction of substitution errors in Moleculo reads. *Proceedings of the 13th IEEE International Conference on Bioinformatics and Bioengineering (IEEE BIBE 2013)*. Chania, Greece 2013.

2. Clement NL, Shepherd B, Bodily P, Tumur-Ochir S, Gim Y, Snell Q, Clement M, **Johnson WE**. Parallel Pair-HMM SNP Detection. *IEEE International Workshop on High Performance Computational Biology*. Shanghai, China 2009.

3. Clement NL, Clement MJ, Snell Q, **Johnson WE**. Parallel Mapping Approaches for GNUMAP. *HICOMB 2011: Tenth IEEE International Workshop on High Performance Computational Biology*, Anchorage, AK. 2011;PubMed PMID: 23396612; PubMed Central PMCID: PMC3565456.

Book Chapters

1. **Johnson WE**, Li C . Adjusting for batch effects in small sample size. In "Batch effects and experimental noise in microarray studies: sources and solutions". Wiley 2009.

2. **Johnson WE**, Liu JS, Liu XS. Analysis of ChIP-chip Data on Genome Tiling Arrays. "New developments in Biostatistics and Bioinformatics." World Scientific Publishing Company 2009.

Technical Reports

1. Francis OE, Bendall M, Clement NL, Snell Q, Schaalje GB, Clement MJ, Crandall KA, **Johnson WE**. Bayesian assignment of unassembled sequencing data to a reference database of known genomes. *Department of Statistics Technical Report Series*. 2009 **TR12-116**, Brigham Young University, Provo, Utah.

2. Bahr TM, Gustafsen N, **Johnson WE**. Unconfounding the Confounded: Bayesian Mixture Model to Adjust for Batch Effects in Confounded Microarray Experiments. *Department of Statistics Technical Report Series*. 2011 **TR11-115**, Brigham Young University, Provo, Utah.

3. Piccolo S, Withers M, Sun Y, Crowther B, Bild A, **Johnson WE**. Gene expression barcoding approach for the integration of expression data from multiple technology platforms. *Department of Statistics Technical Report Series*. 2009 **TR11-114**, Brigham Young University, Provo, Utah.

SOFTWARE

1. **ComBat:** *Combining Batches of Microarray Data*. ComBat is a widely-used software for reducing batch effects when combining microarray data from different labs, experiments, or hybridization batches, or technology platforms. It utilizes an empirical Bayesian linear modeling approach to robustly account for technical variability across multiple high-thoughput studies.

2. **MAT, MA2C:** *Model-based Analysis of Tiling Arrays for ChIP-chip*. MAT is a developed for the analysis of data form Affymetrix tiling microarrays. It removes bias in microarray data attributable to probe and sample on each array individually. It also facilitates genomic profiling and identification of significantly enriched genome regions. The MA2C software is a similar approach, but designed to analyze data from two-color tiling arrays.

3. **GNUMAP:** *Genomic Next-generation Universal Mapper.* GUNUMAP is a software suite for aligning next sequencing data from DNA-seq, BS-seq, and RNA-seq (including small RNAs, RNA editing) experiments. It uses a highly accurate probabilistic alignment approach that incorporates base uncertainty into the alignment algorithm.
4. **SCAN-UPC:** *Single Channel Array Normalization and Universal Probability Codes.* SCAN is a microarray normalization method that removes background noise using only data from within each array individually, therefore facilitating applications in precision medicine. UPC utilizes a similar modeling approach to produce barcode that estimates gene activity in data from microarray and RNA-sequencing platforms. UPC can be applied to any expression profiling platform, thus promoting data integration and platform agnostic downstream analysis tools.
5. **ASSIGN:** *Adaptive Signature Selection and Identification in GeNome-wide profiling data.* ASSIGN utilizes Bayesian factor regression model to identify genomic biomarkers for applications in pathway profiling, drug responsiveness, environmental exposure, and infectious disease diagnosis. ASSIGN can estimate background signal and adapt biomarker signatures into different biological contexts.
6. **PathoScope:** *Rapid and Accurate Sequence-based Infectious Disease Diagnostics.* PathoScope is a complete bioinformatics framework for the metagenomic analysis of data from clinical or environmental sequencing samples. PathoScope includes modules for reference genome library extraction and indexing, read quality control and alignment, strain identification, and annotation of results.

PRESENTATIONS

Invited Presentations

1. Tumor Heterogeneity Workshop. Oregon Health and Sciences University (2014), Portland, OR
   Talk Title: *Adaptive models for accounting for heterogeneity in biomarker development*
2. Stanford Center for Cancer Systems Biology, Stanford (2014), Palo Alto, CA
   Talk Title: *Adaptive Models for Assessing Drug Sensitivity and Pathway Activation in Individual Patient Samples*
3. Cancer Biomarker Informatics Workshop, Early Detection Research Network (2013), Pasadena, CA
   Talk Title: *Data preprocessing and integration for reproducible multiomic biomarker discovery and validation*
4. Joint Statistical Meetings, (2012), San Diego, CA
   Talk Title: *Latent Variable Models for Predicting Drug Sensitivity and Hereditary Disease Risk*
5. Center for Functional Cancer Epigenetics, Dana Farber Cancer Institute (2012), Boston, MA
   Talk Title: *Probabilistic unbiased estimation of genome-wide methylation levels from BS-seq data*
6. Department of Biostatistics, Harvard University (2012), Boston, MA
   Talk Title: *Predicting hereditary breast-cancer development via peripheral-blood transcription profiling*
7. Harvard Program in Quantitative Genomics (2012), Boston, MA
   Talk Title: *Universal Probability of Expression Codes: Platform-independent preprocessing of expression profiling data for personalized medicine work flows*
8. Department of Biostatistics, Boston University (2011), Boston, MA
   Talk Title: *Dynamic Latent Models for Profiling Methylation Levels in Genome-wide Experiments*
9. International Biometrics Society, Eastern North American Spring Meeting (2011), Miami, FL
   Talk Title: *Estimation of Genome-Wide Methylation Levels Using Next-generation Sequencing Data*
10. Huntsman Cancer Institute, University of Utah (2011), Salt Lake City, UT
    Talk Title: *Statistics and Personalized Medicine: Latent Variable Models for Predicting Drug Sensitivity and Hereditary Disease Risk*

11. Department of Biostatistics, Johns Hopkins School of Public Health (2011), Baltimore, MD
    Talk Title: *Personalized Genomic Medicine: Latent Variable Models for Predicting Drug Sensitivity and Hereditary Disease Risk*
12. Division of Computational Biomedicine, Boston University (2011), Boston, MA
    Talk Title: *Statistical and Computational Tools for Assessing Drug Sensitivity and Predicting Hereditary Disease Risk*
13. Joint Statistical Meetings (2010), Vancouver, CA
    Talk Title: *Impact of Statistics on Next-Generation Sequencing*
14. Southern Utah University (2010), Cedar City, UT
    Talk Title: *Predicting Breast Cancer Risk based on Gene Expression Profiles*
15. Department of Oncological Sciences, Huntsman Cancer Institute (2007), Salt Lake City, UT
    Talk Title: *Normalizing and comparing multiple genome tiling microarrays*
16. Department of Biomedical Informatics, University of Utah (2007), Salt Lake City, UT
    Talk Title: *Doubly stochastic latent variable analysis of ChIP-chip experiments on genome tiling arrays*
17. Department of Oncological Sciences, Huntsman Cancer Institute (2007), Salt Lake City, UT
    Talk Title: *Topics in microarray analysis: ChIP-chip experiments and Batch Effects*
18. Department of Mathematics, Massachusetts Institute of Technology (2007), Cambridge, MA
    Talk Title: *Doubly stochastic latent variable analysis of ChIP-chip experiments on genome tiling arrays*
19. Department of Statistics, Texas A&M University (2007), College Station, TX
    Talk Title: *Doubly stochastic latent variable analysis of ChIP-chip experiments on genome tiling arrays*
20. Department of Statistics, Brigham Young University (2007), Provo, UT
    Talk Title: *Doubly stochastic latent variable analysis of ChIP-chip experiments on genome tiling arrays*
21. Department of Biostatistics, UNC-Chapel Hill (2007), Chapel Hill, NC
    Talk Title: *Doubly stochastic latent variable analysis of ChIP-chip experiments on genome tiling arrays*
22. Joslin Diabetes Center (2006), Boston, MA
    Talk Title: *Adjusting batch effects in microarray data: An Empirical Bayes approach*
23. Mathematical Biosciences Institute (2004), The Ohio State University
    Talk Title: *Empirical Bayes methods for adjusting for batch effects in microarray experiments*

Other Presentations

1. International Biometrics Society, Eastern North American Spring Meeting (2011), Orlando, FL
   Talk Title: *Unconfounding the Confounded: Adjusting for Batch Effects in Completely Confounded Designs in Genomic Studies*
2. Timpview High School (2011), Provo, UT
   Talk Title: *Personalizing the Cure: Improving the Odds of Cancer Recovery using Statistics*
3. Joint Statistical Meetings (2010), Vancouver, CA
   Talk Title: *The GNUMAP Project: Probabilistic Mapping of Next-Generation Sequencing Data*
4. Intelligent Systems for Molecular Biology (2010), Boston, MA
   Talk Title: *GNUMAP-SNP: Probabilistic PHMM for SNP Detection in Next-Gen Sequencing Data*
   Poster Title: *SLAM: GausSian Dynamic Linear Analysis of Methylated Chip Data*
5. International Biometrics Society, Western North American Spring Meeting (2010), Seattle, WA
   Talk Title: *SLAM: GausSian Dynamic Linear Analysis of Methylated Chip Data*
6. International Biometrics Society, Eastern North American Spring Meeting (2010), New Orleans, LA
   Talk Title: *Dynamic Linear Model for the Identification of miRNAs in Next-Gen Sequencing Data*

7. Joint Statistical Meetings (2009), Washington, D.C.
   Poster Title: *Predicting Breast Cancer Using Exon Array Data*
8. Intelligent Systems for Molecular Biology (2009), Stockholm, Sweden
   Talk Title: *The GNUMAP Algorithm: Unbiased Probabilistic Mapping of Next-Gen Sequencing Data*
   Poster Title: *Promatch: Identifying conserved transcription factor binding sites*
   Poster Title: *Predicting Breast Cancer Using Exon Array Data*
9. International Biometrics Society, Western North American Spring Meeting (2009), Portland, OR
   Talk Title: *Dynamic Linear Model for the Identification of miRNAs in Next-Gen Sequencing Data*
10. Joint Statistical Meetings (2008), Denver, CO
    Talk Title: *Normalizing and comparing multiple genome tiling microarrays*
11. Department of Mathematics, Southern Utah University (2006), Cedar City, UT
    Talk Title: *Methods for tiling array analysis*
12. Joint Statistical Meetings (2006), Seattle, WA
    Talk Title: *Model-based analysis of tiling arrays for ChIP-chip*
13. National Science Teachers Association Regional Conference (2001), Salt Lake City, UT
    Talk Title: *Let's Go on a Rainbow Hunt*

TEACHING EXPERIENCE

Courses Taught

Department of Statistics, Brigham Young University, 2003, 2007-2011

- *STAT 121: Principles of Statistics*
- *STAT 221H: Honors Principles of Statistics*
- *STAT 301: Statistics and Probability for Secondary Educators*
- *STAT 322: Matrix Computations in Statistics*
- *STAT 337: Statistical Methods 2*
- *STAT 642: Probability Theory and Mathematical Statistics 2* (graduate level)

Guest Lectures

- *BF 690: Bioinformatics Challenge Project*, Program in Bionformatics, Boston University, 2012-2014
- *BF 831: Translational Bioinformatics*, Department of Medicine, Boston University, Fall 2014
- *BE 768: Biological Database Analysis*, Program in Bionformatics, Boston University, Spring 2014
- *FC 702: Foundations in Biomedical Sciences*, Graduate Medical Sciences, Boston University, Fall 2013
- *BS 859: Applied Genetic Analysis*, Department of Biostatistics, Boston University, Spring 2013
- *BS 859: Applied Genetic Analysis*, Department of Biostatistics, Boston University, Spring 2012
- *PHTX7777: Applied genomics*, Genome Sciences Program, University of Utah, Spring 2012
- *BIO 465: Genomics*, Department of Biology, Brigham Young University, Spring 2009
- *BIO 465: Genomics*, Department of Biology, Brigham Young University, Spring 2008

Tutorials and Symposia

- *Low level Processing and Visualization of Whole Exome Sequencing Data.* Harvard Program in Quantitative Genomics (2011), Boston, MA.
- *Computational analysis of tiling arrays for ChIP-chip on mammalian genomes.* Computational Systems Bioinformatics (2006), Stanford, CA.

Courses as Teaching Assistant

Department of Biostatistics, Harvard University 2003-2007

- *BIO 201: Introduction to Statistical Methods* (graduate level)
- *BIO 230: Probability I* (graduate level)
- *BIO 280: Introduction to Computational Molecular Biology* (graduate level; Presidential Instructional Technology Fellow)

Department of Statistics, Brigham Young University, 2002-2003

- *STAT 105: Introduction to Statistics*
- *STAT 212: Statistical Computing 1*
- *STAT 301: Statistics and Probability for Secondary Educators*
- *STAT 321: Elements of Mathematical Statistics*
- *STAT 441: Statistical Theory 1*
- *STAT 510: Introduction to Statistics* (graduate level)
- *STAT 512: Statistical Methods for Research 2* (graduate level)
- *STAT 525: Statistical Inference* (graduate level)
- *STAT 545: Stochastic Processes* (graduate level)

PROFESSIONAL MEMBERSHIPS AND SERVICE

Service to the Discipline

Treasurer, Utah Chapter of the American Statistical Association, 2009-2011

Study Sections

GCAT Study Section, National Institutes of Health, 2010
Special Emphasis Panel, ZRG1 SBIB-Z (59), National Institutes of Health, 2014
Special Emphasis Panel, ZCA1 SRLB-4 (M1) R, National Institutes of Health, 2014
External Advisory Committee, NHLBI Progenitor Cell Biology Consortium, 2014

Served as Referee for

Annals of Applied Statistics, Biostatistics, Bioinformatics, BMC Bioinformatics, BMC Medical Genomics, Breast Cancer Research, Cancer Prevention Research, Epigenetics and Chromatin, Evolutionary Bioinformatics, International Conference on Intelligent Systems for Molecular Biology, Journal of Clinical Bioinformatics, Journal of Immunological Methods, Plant Methods, Proceedings of the National Academy of Sciences, Nucleic Acids Research, PeerJ, PLoS Computational Biology, PLoS One, Science Translational Medicine, Statistical Applications in Genetics and Molecular Biology, Statistics in Biosciences

Memberships

- International Society for Computational Biology, 2009-Present
- Institute of Mathematical Statistics, 2006-Present
- International Biometrics Society, Eastern North American Region (WNAR/ENAR), 2006-Present
- American Statistical Association, Biometrics Section, Boston/Utah Chapter, 2004-Present
- Mu Sigma Ro, National Statistics Honors Society, 2002-2003, 2006-Present
- Alpha Chi, National College Honors Society, 2001-Present

UNIVERSITY SERVICE

Boston University Committees

- Bioinformatics Graduate Admission Committee
- Early Stage Faculty Advisory Committee

University of Utah Committees

- Search Committee, Biostatistics Search, Huntsman Cancer Institute, 2008-2009

Brigham Young University Committees

- Computational Biology Search, Department of Biology (2010-2011)
- Methods Sub-Committee, Commitee for Teaching and Learning (2009-2011)
- Department Seminar Series Coordinator (2009-2010)
- Statistics Search Committee (2009-2010)
- Spring Research Conference Department Coordinator (2009)
- Comprehensive Exam Committee (2008-Present)

Postdoctoral Fellows

- Ying Shen, Ph.D. (Boston University, 2012-present)
- Changjin Hong, Ph.D. (Boston University, 2011-present)
- Stephen Piccolo, Ph.D. (Boston University, 2010-present)
- Bing Han, Ph.D. (Boston University, 2012-2013)
- Yumei Li, Ph.D. (Huntsman Cancer Insititute, 2009-2010)
- Ying Sun, Ph.D. (University of Utah, 2008-2010)

Graduate Students

- Supriya Sharma (Boston University, 2013-present)
- Solaiappan Manimaran (Boston University, 2012-present)
- Michelle Withers (Brigham Young University, 2008-2011)
- Rachel Poulsen (Brigham Young University, 2008-2009)

Graduate Committees

- Joshua Loving (Boston University, 2013-present)
- Anna Lyubetskaya (Boston University, 2013-present)
- Daniel Gusenleitner (Boston University, 2013-present)
- Nacho Caballero (Boston University, 2012-present)
- Yuxiang Tan (Boston University, 2012-present)
- Teresa Wang (Boston University, 2012-present)
- John Farrell (Boston University, 2012-present)

- Kevin Tuttle (Brigham Young University, 2012-present)
- Matthew Bendall (Brigham Young University, 2011)
- Mark Nielson (Brigham Young University, 2010-2011)
- Scott Morris (Brigham Young University, 2009-2010)
- Rozaura Vivas Hall (Brigham Young University, 2008-2009)
- Rick Smith (Brigham Young University, 2008-2009)
- Scott Howard (Brigham Young University, 2008)

Undergraduate Research Mentoring

- Spencer Clement (Brigham Young University, 2010-2011)
- Ariana Hedges (Brigham Young University, 2010-2011)
- Kathleen Finlinson (Brigham Young University, 2009-2011)
- Owen Francis (Brigham Young University, 2009-2011)
- Brent Shepherd (Brigham Young University, 2009-2011)
- Christa Schank (Brigham Young University, 2009-2011)
- Nathan Clement (Brigham Young University, 2008-2010)
- Brandon Crowther (Brigham Young University, 2008-2009)
- Jamie Lapierre (Brigham Young University, 2007-2010)
- Colin Rogerson (Brigham Young University, 2007-2010)
- Tim Bahr (Brigham Young University, 2007-2009)

AWARDS AND FELLOWSHIPS

- NIH Pre-doctoral Traineeship in Cancer Research, 2003-2007
- Prior Alumni Fellowship, Alpha Chi Honor Society, 2006
- Honorarium to travel and present at Computational Systems Biology, 2006
- Award for Achievement in Instructional Technology, Harvard PITF, 2006
- Certificate of Distinction in Teaching, Harvard Biostatistics Dept, 2004-2006
- Nominated for the Departmental Outstanding Scholar Award, SUU Math Dept, 2002
- Name published in The National Deans List, 2001