

W. Evan Johnson, Ph.D.

Division of Infectious Disease
Center for Data Science
Rutgers New Jersey Medical School
225 Warren Street
Newark, NJ 07103

email: w.evan.johnson@rutgers.edu
phone: (801) 472-6951
fax: (973) 972-3644
twitter: @wejlab

EDUCATION

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

ACADEMIC APPOINTMENTS

- 2022-Present **Professor of Medicine**, Division of Infectious Disease, Department of Medicine, Rutgers New Jersey Medical School
- 2022-Present **Professor of Biostatistics**, Department of Biostatistics and Epidemiology, Rutgers University School of Public Health
- 2021-Present **Adjunct Professor**, Department of Mathematics, Southern Utah University
- 2020-2022 **Founding Faculty**, Faculty of Computing and Data Sciences, Boston University
- 2015-2022 **Associate Professor of Medicine**, Division of Computational Biomedicine, Department of Medicine, Boston University School of Medicine
- 2015-2022 **Associate Professor of Biostatistics**, Department of Biostatistics, Boston University School of Public Health
- 2011-2015 **Assistant Professor of Medicine**, Division of Computational Biomedicine, Department of Medicine, Boston University School of Medicine
- 2011-2015 **Assistant Professor of Biostatistics**, Department of Biostatistics, Boston University School of Public Health
- 2008-2018 **Adjunct Assistant Professor**, Department of Oncological Sciences, Huntsman Cancer Institute and University of Utah
- 2007-2011 **Assistant Professor of Statistics**, Department of Statistics, Brigham Young University
- 2004-2007 **Research Assistant**, Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute
- 2003-2007 **Teaching Assistant**, Department of Biostatistics, Harvard School of Public Health
- 2002-2003 **Teaching Assistant**, Department of Statistics, Brigham Young University, 2002-2003

LEADERSHIP EXPERIENCE

- 2022-Present **Founding Director**, Center for Data Science, Rutgers New Jersey Medical School
- 2021-2022 **Co-director**, Predictors of Efficacy and Resistance to Cancer Checkpoint Inhibition, Affinity Research Collaborative, BU Interdisciplinary Biomedical Research Office
- 2018-2022 **Co-scientific Director**, Single Cell Sequencing Resource, Department of Medicine, Boston University School of Medicine
- 2017-2020 **Co-director**, Systems Biology Approaches to Microbiome Research, Affinity Research Collaborative, BU Interdisciplinary Biomedical Research Office
- 2015-2022 **Associate Division Chief**, Division of Computational Biomedicine, Department of Medicine, Boston University School of Medicine

RESEARCH INTERESTS

Applications: Applications in precision genomic medicine, clinical metagenomics and infectious disease research (tuberculosis & HIV), bulk and single cell RNA sequencing data analysis, and cancer research

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

HONORS

Institutional

| | |
|-----------|---|
| 2006 | Presidential Instructional Fellowship, Award in Instructional Technology, Harvard University, Cambridge, MA |
| 2004-2006 | Certificate of Distinction in Teaching, Biostatistics Department, Harvard University, Cambridge, MA |
| 2003-2007 | NIH Pre-doctoral Traineeship in Quantitative Cancer Research, Harvard School of Public Health, Boston MA |
| 2002 | Outstanding Scholar Award Finalist, Mathematics Department, Southern Utah University, Cedar City, UT |
| 2002 | Summa Cum Laude, Southern Utah University, Cedar City, UT |

National

| | |
|------|---|
| 2006 | Prior Alumni Fellowship, Alpha Chi Honor Society, Little Rock, AK |
| 2001 | Name published in The National Deans List |

EXTERNAL FUNDING

Ongoing Research Support

| | | |
|---|---------------------------|--------------|
| 1R01AI175315-01 (Johnson, Salgame, Song) | 02/01/2023 – 01/31/2028 | 1.2 calendar |
| NIH/NIAID | \$5,288,988 (total award) | |
| Title: Biomarker Signatures of TB Infection in Young Children with and Without HIV | | |
| Description: i) Confirm that biomarkers which have shown promise in adult populations can identify children ≥5 years of age at risk of progression to TB; ii) Evaluate and develop novel biomarkers that can identify children at risk of progression to TB detecting Mycobacterium tuberculosis bacterial products secreted in blood; iii) Evaluate and develop novel biomarkers that can identify children at risk of progression to TB by detecting host biomarkers in non-sputum body fluids; iv) Identify combinations of biomarkers that applied together could be used to identify children at risk of progression to TB. | | |
| 1R01 GM127430-05 (Johnson) | 05/01/2018-04/30/2023 | 3.0 calendar |
| NIH/NIGMS | \$1,303,335 (total award) | |
| Title: Removing batch effects in genomic and epigenomic studies | | |
| Description: We will develop algorithms and software for integrating data from genomic or epigenomic sequencing studies, creating reference standards for batch adjustment, accounting for the impacts of batch adjustment, identification and visualization of batch effects. | | |
| 1 R21 CA260382-01A1 (Johnson, Segre) | 09/01/2021-08/31/2023 | 1.0 calendar |
| NIH/NCI | \$424,257 (total award) | |
| Title: Microbiome-based biomarkers and models of lung cancer development and treatment | | |
| Description: This proposal will use RNA-sequencing data from nasal samples of current and former smokers in combination with state-of-the art computational modeling to develop effective biomarkers and explore the role of microbe-microbe and microbe-host interactions in lung cancer. | | |

- G-OISE-20-66940-1 (Johnson) 5/1/22- 4/30/2023 0.6 calendar
 NIH/NIAID via CRDF Global \$46,941 (total award)
Title: Dynamics and Immune Mechanisms of QFN Response in Close Contacts of TB cases
Description: Using a prospective observational cohort of household contacts (HHC) of TB patients at two sites in India, we plan to characterize baseline and dynamic changes in the innate and adaptive immune responses for sustained IGRA negatives, sustained IGRA positives, IGRA converters, and IGRA reverters for risk of progression as well as development of incident TB.
- DAA3-19-65672-1 (Ellner) 10/1/2019 – 09/30/2024 1.2 calendar
 NIH/NIAID via CRDF Global \$235,767 (per year)
Title: Biomarkers for Risk of Tuberculosis and for Tuberculosis Treatment Failure and Relapse
Description: The major goal of this project is to collaborate with partners in India in clinical studies of tuberculosis infection and disease.
- U19 AI162598-02 (Alland, Salgame, Ellner) 07/01/2021- 06/30/2026 1.2 calendar
 NIH/NIAID \$14,285,110 (total award)
Title: Bacterial and Host Heterogeneity in TB latency, persistence and progression
Description: The overall goals are i) to determine the effects of bacterial and host heterogeneity on the manifestations, progression and consequences of close exposure to TB in the household and of active TB; ii) to uncover the immunological mechanisms underlying the diverse clinical outcomes in hosts infected with high and low transmission strains of Mycobacterium tuberculosis; iii) to define the host immune pathways that induce drug tolerance and identify potential routes to therapeutic intervention; and iv) to define bacterial factors that contribute to the heterogeneous expression of drug tolerance and characterize links with adverse treatment outcomes.

Completed Projects

- 5R01CA221260-03 (Bryant) 02/07/2018-01/31/2022 0.6 calendar
 NIH/NIDA \$368,197 (per year)
Title: Genetic Basis of Chemotherapy-Induced Neuropathy in a Reduced Complexity Cross
Description: The primary objective of this proposal is to identify novel genetic factors that contribute to paclitaxel- induced neuropathy in mice. In Aim 1, we will use the RCC to map genomic regions, or QTLs, that are causally associated with susceptibility versus resilience to multiple measures of CIPN. In Aim 2, we will conduct transcriptome analysis via mRNA sequencing (RNA-seq) of spinal and peripheral neuronal regions in control mice and paclitaxel-treated mice from the parental male and female C57BL/6J and C57BL/6NJ substrains.
- Research Agreement (Kulke) 09/13/2020-09/12/2021 0.6 calendar
 Bristol-Myers Squibb \$103,110 (Johnson support)
Title: Multidimensional IO markers of Response and Resistance to Checkpoint Inhibition
Description: In this project we will explore the role of the microbiome as well as tissue and circulating biomarkers in the treatment sensitivity and resistance in patients receiving treatment with checkpoint inhibitors.
- 1R01LM013154-01 (Campbell) 08/01/2019-07/31/2022 0.6 calendar
 NIH/NIEHS \$217,500 (per year)
Title: Integrative clustering of cells and samples using multi-modal single-cell data
Description: In this proposal, we will develop novel discrete Bayesian hierarchical models that can cluster cells into subpopulations using multiple data types and cluster patients into subgroups using both single-cell data and patient-level characteristics.
- SAC180086 (Palmer) 06/19/2019 - 06/18/2021 1.2 calendar
 Susan G. Komen Foundation \$200,000 (per year)
Title: Family History, Genetics, and Environment: Breast Cancer Risk in U.S. Black Women
Description: This proposal will develop a risk prediction tool that will consider the different risk factors for

ER+ and ER- breast cancer, along with age-related incidence patterns in African-American women. This tool will help identify African-American women who would benefit from earlier and more frequent screening or alternative modes of screening.

1R21 AI154387-01 (Johnson) 07/10/2020 – 06/30/2022 1.8 calendar
NIH/NIAID \$467,564 (total award)

Title: Signature of Profiling and Staging the Progression of TB from Infection to Disease

Description: We will compile existing TB gene expression data and develop a wholistic map of TB progression from infection to active disease. We will provide curated data for dozens of existing TB RNA-sequencing datasets and allow others to access and explore these data through a user-friendly software toolkit and platform.

1U01 CA220413-01 (Johnson) 08/01/2017 - 07/31/2021 1.2 calendar
NIH/NCI \$241,877 (direct per year)

Title: An interactive analysis toolkit for single cell RNA-seq in cancer research

Description: We plan to develop a comprehensive and interactive R-software framework for complete data processing and analysis of single cell RNA-sequencing data from heterogeneous tumor samples. We will develop an R/Shiny user interface that will enable interactive analysis and visualization in the data.

5U19 AI111276-04 (Ellner) 08/01/2014-07/31/2021 1.2 calendar
NIH/NIAID \$2,823,545 (per year)

Title: Biomarkers and Mechanisms of Paucibacillary and Latent Tuberculosis

Description: The proposed studies will discover biomarkers to stratify risk of individuals with latent TB infection and promote their use in targeting preventive therapy, allowing individualized short course treatment regimens, and as surrogate endpoints in clinical trials. Clinical studies and the rabbit model will establish comparability in rates of replication and mutation rates, PET-CT findings and sites of reactivation, validating the use of the model to study new drugs, regimens and immunotherapies, which can be rapidly translated into clinical trials.

Research Agreement (Kulke) 09/13/2019-09/12/2020 0.72 calendar
Bristol-Myers Squibb \$103,110 (Johnson support only)

Title: Multidimensional IO markers of Response and Resistance to Checkpoint Inhibition

Description: In this project we will explore the role of the microbiome as well as tissue and circulating biomarkers in the treatment sensitivity and resistance in patients receiving treatment with checkpoint inhibitors.

R01 DA039168-01A1 (Bryant) 07/01/15 - 06/30/20 0.6 calendar
NIH/NIDA \$378,670 (per year)

Title: Bridging genetic variation with behavior

Description: Molecular and functional mechanisms of quantitative trait gene regulation of the stimulant and addictive properties of methamphetamine in mice.

R21 CA226188-01A1 (Campbell) 01/01/2019-12/31/2020 0.30 calendar
NIH/NCI \$150,000 (direct per year)

Title: Enhanced deconvolution and prediction of mutational signatures

Description: In this proposal, we will develop a novel discrete Bayesian hierarchical model to characterize mutational signatures in tumor sequencing data that overcomes the limitations of NMF.

2R01 GM083084-01 (Irizarry) 09/01/2016 - 06/30/2020 1.2 calendar
NIH/NIGMS \$337,725 (per year)

Title: Preprocessing and Analysis Tools for High-Throughput Technologies

Description: We propose providing a precise and accurate single-sample processing method measures to facilitate clinical application of RNA-Seq, developing statistical methodology that accounts for the problem of detection bias in high-throughput single cell data, and developing a framework for statistical inference for region detection

- 5U01 CA187508-03 (Palmer) 09/01/2015-08/31/2019 1.2 calendar
 NIH/NIEHS \$544,404 (per year)
Title: A Prospective Investigation of the Oral Microbiome and Risk of Pancreatic Cancer
Description: Leveraging the existing resources of two prospective, epidemiologic studies, the Black Women's Health Study and the Southern Community Cohort Study, we propose to substantially advance this emerging line of enquiry by conducting a comprehensive investigation of the oral microbiome and pancreatic cancer.
- 1R01 ES025002-01 (Johnson) 09/18/14 - 08/31/16 3.0 calendar
 NIH/NIEHS \$219,600 (per year)
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.
- 1R01 AI113321-01 (Fearnly) 06/01/14 - 04/30/15 0.3 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.
- UC7 AI095321-01 (Murphy) 06/01/14 - 04/30/15 0.6 calendar
 NIH/NIAID \$1,836,198
Title: NEIDL Operations Grant (BMP Core)
Description: The National Emerging Infectious Diseases Laboratories (NEIDL) Institute's mission is to perform cutting-edge basic and clinical research on emerging infectious diseases, including category A, B, and C agents.
- 5R01 HL118542-02 (Lenburg) 08/15/13 - 05/31/17 0.6 calendar
 NIH/NHLBI \$500,221 (per year)
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.
- Janssen Pharmaceuticals, Inc. (Spira) 03/01/2013 - 02/28/2016 1.2 calendar
 Janssen Pharmaceuticals, Inc. \$748,210 (total award)
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 pre-malignant airway lesions to identify predictors of lesion progression and of lung cancer.
- 1U01 CA164720-03 (Johnson, Bild, Gray) 08/01/12 - 07/31/17 1.92 calendar
 NIH/NCI \$88,024 (Johnson 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.
- 5R01 HG005692-05 (Johnson) 06/01/10 - 05/31/15 2.64 calendar
 NIH/NHGRI \$252,056 (per year)
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.
- 2 P01 CA073992-11 (Burt) 12/01/09 - 11/30/14 1.2 calendar
 NIH/NCI \$1,867,447 (per year)
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

ORCID: <https://orcid.org/0000-0002-6247-6595>

MyNCBI: <https://www.ncbi.nlm.nih.gov/myncbi/william.johnson.1/bibliography/public/>

Google Scholar: <https://scholar.google.com/citations?user=PI6bhY4AAAAJ&hl=en>

As of 2/2023: 16,831 citations, h-index of 40, and mean Relative Citation Ratio (RCR) of 3.7

Peer-Reviewed Articles

1. Leibler JH, Abdelgadir A, Seidel J, White RF, **Johnson WE**, Reynolds SJ, Gray GC, Schaeffer JW. Influenza D virus exposure among US cattle workers: A call for surveillance. *Zoonoses Public Health*. 2023 Mar;70(2):166-170. doi: 10.1111/zph.13008. Epub 2022 Nov 12. PubMed PMID: 36370131.
2. Odom-Mabey AR, Gill CJ, Pieciak R, Ismail A, Thea D, MacLeod WB, **Johnson WE**, Lapidot R. Characterization of longitudinal nasopharyngeal microbiome patterns in maternally HIV-exposed Zambian infants [version 1; peer review: awaiting peer review]. *Gates Open Research* 2022, 6:143 (<https://doi.org/10.12688/gatesopenres.14041.1>)
3. Lapidot R, Faits T, Ismail A, Khumalo Z, MacLeod W, Kwenda G, Mupila Z, Nakazwe R, Segrè D, **Johnson WE**, Thea DM, Mwananyanda L, Gill CJ. Nasopharyngeal Dysbiosis Precedes the Development of Lower Respiratory Tract Infections in Young Infants, a Longitudinal Infant Cohort Study [version 1; peer review: 1 not approved]. *Gates Open Research* 2022, 6:48 (<https://doi.org/10.12688/gatesopenres.13561.1>)
4. Chitale P, Lemenze AD, Fogarty EC, Shah A, Grady C, Odom-Mabey AR, **Johnson WE**, Yang JH, Eren AM, Brosch R, Kumar P, Alland D. A comprehensive update to the Mycobacterium tuberculosis H37Rv reference genome. *Nature Communications*. 2022 Nov 18;13(1):7068. doi: 10.1038/s41467-022-34853-x. PubMed PMID: 36400796; PubMed Central PMCID: PMC9673877.
5. Kaipilyawar V, Zhao Y, Wang X, Joseph NM, Knudsen S, Prakash Babu S, Muthaiah M, Hochberg NS, Sarkar S, Horsburgh CR, Ellner JJ, **Johnson WE**, Salgame P. Development and Validation of a Parsimonious Tuberculosis Gene Signature Using the digital NanoString nCounter Platform. *Clinical Infectious Disease*. 2022 Sep 29;75(6):1022-1030. doi: 10.1093/cid/ciac010. PubMed PMID: 35015839; PubMed Central PMCID: PMC9522394. (Joint senior author)
6. VanValkenburg A, Kaipilyawar V, Sarkar S, Lakshminarayanan S, Cintron C, Prakash Babu S, Knudsen S, Joseph NM, Horsburgh CR, Sinha P, Ellner JJ, Narasimhan PB, **Johnson WE**, Hochberg NS, Salgame P. Malnutrition leads to increased inflammation and expression of tuberculosis risk signatures in recently exposed household contacts of pulmonary tuberculosis. *Frontiers in Immunology*. 2022;13:1011166. doi: 10.3389/fimmu.2022.1011166. eCollection 2022. PubMed PMID: 36248906; PubMed Central PMCID: PMC9554585.
7. Wang Z, Yang S, Koga Y, Corbett SE, Shea CV, **Johnson WE**, Yajima M, Campbell JD. Celda: a Bayesian model to perform co-clustering of genes into modules and cells into subpopulations using single-cell RNA-seq data. *NAR Genomics and Bioinformatics*. 2022 Sep;4(3):lqac066. doi: 10.1093/nargab/lqac066. eCollection 2022 Sep. PubMed PMID: 36110899; PubMed Central PMCID: PMC9469931.
8. Krishnamoorthy Y, Ezhumalai K, Murali S, Rajaa S, Majella MG, Sarkar S, Lakshminarayanan S, Joseph NM, Soundappan G, Prakash Babu S, Horsburgh C, Hochberg N, **Johnson WE**, Knudsen S, Pentakota SR, Salgame P, Roy G, Ellner J. Development of prognostic scoring system for predicting 1-year mortality among pulmonary tuberculosis patients in South India. *Journal Public Health (Oxf)*. 2022 Aug 27;. doi: 10.1093/pubmed/fdac087. [Epub ahead of print] PubMed PMID: 36038507.

9. Sheerin D, Abhimanyu, Peton N, Vo W, Allison CC, Wang X, **Johnson WE**, Coussens AK. Immunopathogenic overlap between COVID-19 and tuberculosis identified from transcriptomic meta-analysis and human macrophage infection. *iScience*. 2022 Jun 17;25(6):104464. doi: 10.1016/j.isci.2022.104464. eCollection 2022 Jun 17. PubMed PMID: 35634577; PubMed Central PMCID: PMC9130411.
10. Barathi A, Krishnamoorthy Y, Sinha P, Horsburgh C, Hochberg N, **Johnson WE**, Salgame P, Govindarajan S, Senbagavalli PB, Lakshinarayanan S, Roy G, Ellner J, Sarkar S. Effect of treatment adherence on the association between sex and unfavourable treatment outcomes among tuberculosis patients in Puducherry, India: a mediation analysis. *Journal of Public Health (Oxf)*. 2022 Jun 11; doi: 10.1093/pubmed/fdac062. [Epub ahead of print] PubMed PMID: 35692180.
11. Hong R, Koga Y, Bandyadka S, Leshchik A, Wang Y, Akavoor V, Cao X, Sarfraz I, Wang Z, Alabdulatif S, Jansen F, Yajima M, **Johnson WE**, Campbell JD. Comprehensive generation, visualization, and reporting of quality control metrics for single-cell RNA sequencing data. *Nature Communications*. 2022 Mar 30;13(1):1688. doi: 10.1038/s41467-022-29212-9. PubMed PMID: 35354805; PubMed Central PMCID: PMC8967915
12. Petrick JL, Wilkinson JE, Michaud DS, Cai Q, Gerlovin H, Signorello LB, Wolpin BM, Ruiz-Narváez EA, Long J, Yang Y, **Johnson WE**, Shu XO, Huttenhower C, Palmer JR. The oral microbiome in relation to pancreatic cancer risk in African Americans. *British Journal of Cancer*. 2022 Feb;126(2):287-296. doi: 10.1038/s41416-021-01578-5. Epub 2021 Oct 30. PubMed PMID: 34718358; PubMed Central PMCID: PMC8770575.
13. Li T, Zhang Y, Patil P, **Johnson WE**. Overcoming the impacts of two-step batch effect correction on gene expression estimation and inference. *Biostatistics*. 2021 Dec 10;. doi: 10.1093/biostatistics/kxab039. [Epub ahead of print] PubMed PMID: 34893807.
14. Krishnamoorthy Y, Ezhumalai K, Murali S, Rajaa S, Jose M, Sathishkumar A, Soundappan G, Horsburgh C, Hochberg N, **Johnson WE**, Knudsen S, Salgame P, Ellner J, Prakash Babu S, Sarkar S. Prevalence and risk factors associated with latent tuberculosis infection among household contacts of smear positive pulmonary tuberculosis patients in South India. *Tropical Medicine & International Health*. 2021 Oct 15;. doi: 10.1111/tmi.13693. [Epub ahead of print] PubMed PMID: 34654061.
15. Goldberg LR, Yao EJ, Kelliher JC, Reed ER, Wu Cox J, Parks C, Kirkpatrick SL, Beierle JA, Chen MM, **Johnson WE**, Homanics GE, Williams RW, Bryant CD, Mulligan MK. A quantitative trait variant in *Gabra2* underlies increased methamphetamine stimulant sensitivity. *Genes, Brain and Behavior* 2021 Nov;20(8):e12774. doi: 10.1111/gbb.12774. Epub 2021 Oct 22. PubMed PMID: 34677900.
16. Huang X, Qiao Y, Brady SW, Factor RE, Downs-Kelly E, Farrell A, McQuerry JA, Shrestha G, Jenkins D, **Johnson WE**, Cohen AL, Bild AH, Marth GT. Novel temporal and spatial patterns of metastatic colonization from breast cancer rapid-autopsy tumor biopsies. *Genome Med*. 2021 Oct 28;13(1):170. doi: 10.1186/s13073-021-00989-6. PubMed PMID: 34711268; PubMed Central PMCID: PMC8555066.
17. Cintron C, Narasimhan PB, Locks L, Babu S, Sinha P, Rajkumari N, Kaipilyawar V, Bhargava A, Maloomian K, Chandrasekaran P, Verma S, Joseph N, **Johnson WE**, Wanke C, Horsburgh CR Jr, Ellner JJ, Sarkar S, Salgame P, Lakshminarayanan S, Hochberg NS. Tuberculosis-Learning the Impact of Nutrition (TB LION): protocol for an interventional study to decrease TB risk in household contacts. *BMC Infectious Diseases* 2021 Oct 12;21(1):1058. doi: 10.1186/s12879-021-06734-z. PubMed PMID: 34641820; PubMed Central PMCID: PMC8506078.
18. Borrelli KN, Yao EJ, Yen WW, Phadke RA, Ruan QT, Chen MM, Kelliher JC, Langan CR, Scotellaro JL, Babbs RK, Beierle JC, Logan RW, **Johnson WE**, Wachman EM, Cruz-Martín A, Bryant CD. Sex Differences in Behavioral and Brainstem Transcriptomic Neuroadaptations following Neonatal Opioid Exposure in Outbred Mice. *eNeuro*. 2021 Sep-Oct;8(5). doi: 10.1523/ENEURO.0143-21.2021. Print 2021 Sep-Oct. PubMed PMID: 34479978; PubMed Central PMCID: PMC8454922.

19. Delpech JC, Pathak D, Varghese M, Kalavai SV, Hays EC, Hof PR, **Johnson WE**, Ikezu S, Medalla M, Luebke JI, Ikezu T. Wolframin-1-expressing neurons in the entorhinal cortex propagate tau to CA1 neurons and impair hippocampal memory in mice. *Science Translational Medicine*. 2021 Sep 15;13(611):eabe8455. doi: 10.1126/scitranslmed.abe8455. Epub 2021 Sep 15. PubMed PMID: 34524859.
20. Zhang Y, Patil P, **Johnson WE**, Parmigiani G. Robustifying genomic classifiers to batch effects via ensemble learning. *Bioinformatics*. 2021 Jul 12;37(11):1521-1527. doi: 10.1093/bioinformatics/btaa986. PubMed PMID: 33245114. (Joint corresponding author)
21. Ikezu S, Yeh H, Delpech JC, Woodbury ME, Van Enoo AA, Ruan Z, Sivakumaran S, You Y, Holland C, Guillaumon-Vivancos T, Yoshii-Kitahara A, Botros MB, Madore C, Chao PH, Desani A, Manimaran S, Kalavai SV, **Johnson WE**, Butovsky O, Medalla M, Luebke JI, Ikezu T. Inhibition of colony stimulating factor 1 receptor corrects maternal inflammation-induced microglial and synaptic dysfunction and behavioral abnormalities. *Mol Psychiatry*. 2021 Jun;26(6):1808-1831. doi: 10.1038/s41380-020-0671-2. Epub 2020 Feb 18. PubMed PMID: 32071385; PubMed Central PMCID: PMC7431382.
22. Braun MR, Noton SL, Blanchard EL, Shareef A, Santangelo PJ, **Johnson WE**, Fearn R. Respiratory syncytial virus M2-1 protein associates non-specifically with viral messenger RNA and with specific cellular messenger RNA transcripts. *PLoS Pathogens*. 2021 May;17(5):e1009589. doi: 10.1371/journal.ppat.1009589. eCollection 2021 May. PubMed PMID: 34003848; PubMed Central PMCID: PMC8162694.
23. Zhao Y, Federico A, Faits T, Manimaran S, Segrè D, Monti S, **Johnson WE**. animalcules: interactive microbiome analytics and visualization in R. *Microbiome*. 2021 Mar 28;9(1):76. doi: 10.1186/s40168-021-01013-0. PubMed PMID: 33775256; PubMed Central PMCID: PMC8006385.
24. **Johnson WE**, Odom A, Cintron C, Muthaiah M, Knudsen S, Joseph N, Babu S, Lakshminarayanan S, Jenkins DF, Zhao Y, Nankya E, Horsburgh CR, Roy G, Ellner J, Sarkar S, Salgame P, Hochberg NS. Comparing tuberculosis gene signatures in malnourished individuals using the TBSignatureProfiler. *BMC Infectious Diseases*. 2021 Jan 22;21(1):106. doi: 10.1186/s12879-020-05598-z. PubMed PMID: 33482742; PubMed Central PMCID: PMC7821401.
25. Manabe YC, Andrade BB, Gupte N, Leong S, Kintali M, Matoga M, Riviere C, Samaneka W, Lama JR, Naidoo K, Zhao Y, **Johnson WE**, Ellner JJ, Hosseinipour MC, Bisson GP, Salgame P, Gupta A. A Parsimonious Host Inflammatory Biomarker Signature Predicts Incident Tuberculosis and Mortality in Advanced Human Immunodeficiency Virus. *Clin Infect Dis*. 2020 Dec 17;71(10):2645-2654. doi: 10.1093/cid/ciz1147. PubMed PMID: 31761933; PubMed Central PMCID: PMC7744990.
26. Faits T, Walker ME, Rodriguez-Morato J, Meng H, Gervis JE, Galluccio JM, Lichtenstein AH, **Johnson WE**, Matthan NR. Exploring changes in the human gut microbiota and microbial-derived metabolites in response to diets enriched in simple, refined, or unrefined carbohydrate-containing foods: a post hoc analysis of a randomized clinical trial. *American Journal of Clinical Nutrition* 2020 Dec 10;112(6):1631-1641. doi: 10.1093/ajcn/nqaa254. PubMed PMID: 32936872; PubMed Central PMCID: PMC7727488.
27. Zhang Y, Parmigiani G, **Johnson WE**. ComBat-seq: batch effect adjustment for RNA-seq count data. *Nucleic Acids Research Genomics and Bioinformatics*. 2020 Sep;2(3):lqaa078. doi: 10.1093/nargab/lqaa078. Epub 2020 Sep 21. PubMed PMID: 33015620; PubMed Central PMCID: PMC7518324.
28. Miossec MJ, Valenzuela SL, Pérez-Losada M, **Johnson WE**, Crandall KA, Castro-Nallar E. Evaluation of computational methods for human microbiome analysis using simulated data. *PeerJ*. 2020;8:e9688. doi: 10.7717/peerj.9688. eCollection 2020. PubMed PMID: 32864214; PubMed Central PMCID: PMC7427543.
29. Babbs RK, Beierle JA, Yao EJ, Kelliher JC, Medeiros AR, Anandakumar J, Shah AA, Chen MM, **Johnson WE**, Bryant CD. The effect of the demyelinating agent cuprizone on binge-like eating of sweetened palatable food in female and male C57BL/6 substrains. *Appetite*. 2020 Jul 1;150:104678. doi: 10.1016/j.appet.2020.104678. Epub 2020 Mar 21. PubMed PMID: 32209386; PubMed Central PMCID: PMC7206526.

30. Ruan QT, Yazdani N, Reed ER, Beierle JA, Peterson LP, Luttki KP, Szumlinski KK, **Johnson WE**, Ash PEA, Wolozin B, Bryant CD. 5' UTR variants in the quantitative trait gene *Hnrnp1* support reduced 5' UTR usage and hnRNP H protein as a molecular mechanism underlying reduced methamphetamine sensitivity. *FASEB J*. 2020 Jul;34(7):9223-9244. doi: 10.1096/fj.202000092R. Epub 2020 May 13. PubMed PMID: 32401417.
31. Yang S, Corbett SE, Koga Y, Wang Z, **Johnson WE**, Yajima M, Campbell JD. Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biology*. 2020 Mar 5;21(1):57. doi: 10.1186/s13059-020-1950-6. PubMed PMID: 32138770; PubMed Central PMCID: PMC7059395.
32. Xavier JB, Young VB, Skufca J, Ginty F, Testerman T, Pearson AT, Macklin P, Mitchell A, Shmulevich I, Xie L, Caporaso JG, Crandall KA, Simone NL, Godoy-Vitorino F, Griffin TJ, Whiteson KL, Gustafson HH, Slade DJ, Schmidt TM, Walther-Antonio MRS, Korem T, Webb-Robertson BM, Styczynski MP, **Johnson WE**, Jobin C, Ridlon JM, Koh AY, Yu M, Kelly L, Wargo JA. The Cancer Microbiome: Distinguishing Direct and Indirect Effects Requires a Systemic View. *Trends in Cancer*. 2020 Mar;6(3):192-204. doi: 10.1016/j.trecan.2020.01.004. Epub 2020 Feb 7. Review. PubMed PMID: 32101723.
33. Leong S, Zhao Y, Ribeiro-Rodrigues R, Jones-López EC, Acuña-Villaorduña C, Rodrigues PM, Palaci M, Alland D, Dietze R, Ellner JJ, **Johnson WE**, Salgame P. Cross-validation of existing signatures and derivation of a novel 29-gene transcriptomic signature predictive of progression to TB in a Brazilian cohort of household contacts of pulmonary TB. *Tuberculosis (Edinb)*. 2020 Jan;120:101898. doi: 10.1016/j.tube.2020.101898. Epub 2020 Jan 7. PubMed PMID: 32090859. (Joint corresponding author).
34. Buwembo W, Munabi IG, Kaddumukasa M, Kiryowa H, Mbabali M, Nankya E, **Johnson WE**, Okello E, Sewankambo NK. Non-surgical oral hygiene interventions on disease activity of Rheumatoid arthritis patients with periodontitis: A randomized controlled trial. *J Dent Res Dent Clin Dent Prospects*. 2020 Winter;14(1):26-36. doi: 10.34172/jodddd.2020.004. PubMed PMID: 32454955; PubMed Central PMCID: PMC7235694.
35. Buwembo W, Munabi IG, Kaddumukasa M, Kiryowa H, Nankya E, **Johnson WE**, Okello E, Sewankambo N. Periodontitis and Rheumatoid Arthritis in sub-Saharan Africa, gaps and way forward: a systematic review and meta-analysis. *Open Journal of Stomatology*. 2019 Oct;9(10):215-226. doi: 10.4236/ojst.2019.910023. Epub 2019 Oct 10. PubMed PMID: 31695961; PubMed Central PMCID: PMC6834353.
36. McQuerry JA, Jenkins DF, Yost SE, Zhang Y, Schmolze D, **Johnson WE**, Yuan Y, Bild AH. Pathway activity profiling of growth factor receptor network and stemness pathways differentiates metaplastic breast cancer histological subtypes. *BMC Cancer*. 2019 Sep 5;19(1):881. doi: 10.1186/s12885-019-6052-z. PubMed PMID: 31488082; PubMed Central PMCID: PMC6727561.
37. Wang Z, Hu J, **Johnson WE**, Campbell JD. scruff: an R/Bioconductor package for preprocessing single-cell RNA-sequencing data. *BMC Bioinformatics*. 2019 May 2;20(1):222. doi: 10.1186/s12859-019-2797-2. PubMed PMID: 31046658; PubMed Central PMCID: PMC6498700.
38. Walker ME, Matthan NR, Lamón-Fava S, Solano-Aguilar G, Jang S, Lakshman S, Molokin A, Urban JF Jr, Faits T, **Johnson WE**, Lichtenstein AH. A Western-Type Dietary Pattern Induces an Atherogenic Gene Expression Profile in the Coronary Arteries of the Ossabaw Pig. *Curr Dev Nutr*. 2019 May;3(5):nzz023. doi: 10.1093/cdn/nzz023. eCollection 2019 May. PubMed PMID: 31049488; PubMed Central PMCID: PMC6488722.
39. Walker ME, Matthan NR, Solano-Aguilar G, Jang S, Lakshman S, Molokin A, Faits T, Urban JF Jr, **Johnson WE**, Lamón-Fava S, Lichtenstein AH. A Western-type dietary pattern and atorvastatin induce epicardial adipose tissue interferon signaling in the Ossabaw pig. *J Nutr Biochem*. 2019 Feb 20;67:212-218. doi: 10.1016/j.jnutbio.2019.02.003. [Epub ahead of print] PubMed PMID: 30981985.
40. Bryant CD, Bagdas D, Goldberg LR, Khalefa T, Reed ER, Kirkpatrick SL, Kelliher JC, Chen MM, **Johnson WE**, Mulligan MK, Imad Damaj M. C57BL/6 substrain differences in inflammatory and

- neuropathic nociception and genetic mapping of a major quantitative trait locus underlying acute thermal nociception. *Molecular Pain*. 2019 Jan-Dec;15:1744806918825046. doi: 10.1177/1744806918825046. PubMed PMID: 30632432; PubMed Central PMCID: PMC6365993.
41. Griffin PJ, Zhang Y, **Johnson WE**, Kolaczyk ED. Detection of multiple perturbations in multi-omics biological networks. *Biometrics*. 2018 Dec;74(4):1351-1361. doi: 10.1111/biom.12893. Epub 2018 May 17. PubMed PMID: 29772079.
 42. Zhao Y, **Johnson WE**. Exploring Host-Microbe Interactions in Lung Cancer. *Am J Respir Crit Care Med*. 2018 Nov 1;198(9):1116-1118. doi: 10.1164/rccm.201807-1225ED. PubMed PMID: 30063838.
 43. Zhang Y, Jenkins DF, Manimaran S, **Johnson WE**. Alternative empirical Bayes models for adjusting for batch effects in genomic studies. *BMC Bioinformatics*. 2018 Jul 13;19(1):262. doi: 10.1186/s12859-018-2263-6. PubMed PMID: 30001694; PubMed Central PMCID: PMC6044013.
 44. Leong S, Zhao Y, Joseph NM, Hochberg NS, Sarkar S, Pleskunas J, Hom D, Lakshminarayanan S, Horsburgh CR Jr., Roy G, Ellner JJ, **Johnson WE**, Salgame P. Existing blood transcriptional classifiers accurately discriminate active tuberculosis from latent infection in individuals from south India. *Tuberculosis (Edinb)*. 2018 Mar;109:41-51. doi: 10.1016/j.tube.2018.01.002. Epub 2018 Jan 31. PubMed PMID: 29559120. (Joint corresponding author).
 45. Brady SW, McQuerry JA, Qiao Y, Piccolo SR, Shrestha G, Jenkins DF, Layer RM, Pedersen BS, Miller RH, Esch A, Selitsky SR, Parker JS, Anderson LA, Dalley BK, Factor RE, Reddy CB, Boltax JP, Li DY, Moos PJ, Gray JW, Heiser LM, Buys SS, Cohen AL, **Johnson WE**, Quinlan AR, Marth G, Werner TL, Bild AH. Combating subclonal evolution of resistant cancer phenotypes. *Nature Communications* 2017 Nov 1;8(1):1231. doi: 10.1038/s41467-017-01174-3. PubMed PMID: 29093439; PubMed Central PMCID: PMC5666005.
 46. Goldberg LR, Kirkpatrick SL, Yazdani N, Luttik KP, Lacki OA, Babbs RK, Jenkins DF, **Johnson WE**, Bryant CD. Casein kinase 1-epsilon deletion increases mu opioid receptor-dependent behaviors and binge eating1. *Genes Brain Behavior* 2017 Sep;16(7):725-738. doi: 10.1111/gbb.12397. Epub 2017 Jul 6. PubMed PMID: 28594147.
 47. Perez-Rogers JF, Gerrein J, Anderlind C, Liu G, Zhang S, Alekseyev Y, Smith KP, Whitney D, **Johnson WE**, Elashoff DA, Dubinett SM, Brody J, Spira A, Lenburg ME; AEGIS Study Team. Shared Gene Expression Alterations in Nasal and Bronchial Epithelium for Lung Cancer Detection. *Journal of the National Cancer Institute*. 2017 Jul 1;109(7). doi: 10.1093/jnci/djw327. PubMed PMID: 28376173.
 48. Kirkpatrick SL, Goldberg LR, Yazdani N, Babbs RK, Wu J, Reed ER, Jenkins DF, Bolgioni AF, Landaverde KI, Luttik KP, Mitchell KS, Kumar V, **Johnson WE**, Mulligan MK, Cottone P, Bryant CD. Cytoplasmic FMR1-Interacting Protein 2 is a Major Genetic Factor Underlying Binge Eating. *Biological Psychiatry*. 2017 May 1;81(9):757-769. doi: 10.1016/j.biopsych.2016.10.021. Epub 2016 Oct 25. PubMed PMID: 27914629; PubMed Central PMCID: PMC5386810.
 49. Rahman M, MacNeil SM, Jenkins DF, Shrestha G, Wyatt SR, McQuerry JA, Piccolo SR, Heiser LM, Gray JW, **Johnson WE**, Bild AH. Activity of distinct growth factor receptor network components in breast tumors uncovers two biologically relevant subtypes. *Genome Medicine*. 2017 Apr 26;9(1):40. doi: 10.1186/s13073-017-0429-x. PubMed PMID: 28446242; PubMed Central PMCID: PMC5406893.
 50. Manimaran S, Selby HM, Okrah K, Ruberman C, Leek JT, Quackenbush J, Haibe-Kains B, Bravo HC, **Johnson WE**. BatchQC: interactive software for evaluating sample and batch effects in genomic data. *Bioinformatics*. 2016 Dec 15;32(24):3836-3838. Epub 2016 Aug 18. PubMed PMID: 27540268; PubMed Central PMCID: PMC5167063.
 51. Lam S, Mandrekar SJ, Gesthalter Y, Allen Ziegler KL, Seisler DK, Midthun DE, Mao JT, Aubry MC, McWilliams A, Sin DD, Shaipanich T, Liu G, **Johnson WE**, Bild A, Lenburg ME, Ionescu DN, Mayo J, Yi JE, Tazelaar H, Harmsen WS, Smith J, Spira AE, Beane J, Limburg PJ, Szabo E; Cancer Prevention Network. A Randomized Phase IIb Trial of myo-Inositol in Smokers with Bronchial Dysplasia. *Cancer Prevention Research (Phila)*. 2016 Dec;9(12):906-914. Epub 2016 Sep 22. PubMed PMID: 27658890; PubMed Central PMCID: PMC5136333.

52. Yazdani N, Shen Y, **Johnson WE**, Bryant CD. Striatal transcriptome analysis of a congenic mouse line (chromosome 11: 50-60Mb) exhibiting reduced methamphetamine sensitivity. *Genomic Data*. 2016 Jun;8:77-80. PubMed PMID: 27222804; PubMed Central PMCID: PMC4856813.
53. Hilton SK, Castro-Nallar E, Pérez-Losada M, Toma I, McCaffrey TA, Hoffman EP, Siegel MO, Simon GL, **Johnson WE**, Crandall KA. Metataxonomic and Metagenomic Approaches vs. Culture-Based Techniques for Clinical Pathology. *Frontiers in Microbiology*. 2016 Apr 7;7:484. doi: 10.3389/fmicb.2016.00484. PubMed PMID: 27092134; PubMed Central PMCID: PMC4823605.
54. Piccolo SR, Hoffman LM, Conner T, Shrestha G, Cohen AL, Marks JR, Neumayer LA, Agarwal CA, Beckerle MC, Andrulis IL, Spira AE, Moos PJ, Buys SS, **Johnson WE**, Bild AH. Integrative analyses reveal signaling pathways underlying familial breast cancer susceptibility. *Molecular Systems Biology*. 2016 Mar 10;12(3):860. PubMed PMID: 26969729; PubMed Central PMCID: PMC4812528. (Joint corresponding author).
55. Shah DJ, Rohlfing F, Anand S, **Johnson WE**, Bench Alvarez MT, Cobell J, Jackson King J, Young SA, Kauwe JSK, Graves SW. Discovery and confirmation of novel serum biomarkers diagnosing Alzheimer's Disease. *Journal of Alzheimer's Disease*. 2015 Oct 10;49(2):317-27. PubMed PMID: 26484917.
56. Yazdani N, Parker CC, Shen Y, Reed ER, Guido MA, Kole LA, Kirkpatrick SL, Lim JE, Sokoloff G, Cheng R, **Johnson WE**, Palmer AA, Bryant CD. Hnrnp1 Is A Quantitative Trait Gene for Methamphetamine Sensitivity. *PLOS Genetics*. 2015 Dec 10; 11(12):e1005713. doi:10.1371/journal.pgen.1005713
57. Rahman M, Jackson LK, **Johnson WE**, Li DY, Bild AH, Piccolo SR. Alternative preprocessing of RNA-Sequencing data in The Cancer Genome Atlas leads to improved analysis results. *Bioinformatics*. 2015 Nov 15;31(22):3666-72. PubMed PMID: 26209429; PubMed Central PMCID: PMC4804769.
58. Piccolo SR, Andrulis IL, Cohen AL, Conner T, Moos PJ, Spira AE, Buys SS, **Johnson WE**, Bild AH. Gene-expression patterns in peripheral blood classify familial breast cancer susceptibility. *BMC Medical Genomics*. 2015 Nov 4;8(1):72. doi: 10.1186/s12920-015-0145-6. PubMed PMID: 26538066; PubMed Central PMCID: PMC4634735. (Joint corresponding author).
59. Castro-Nallar E, Shen Y, Freishtat RJ, Pérez-Losada M, Manimaran S, Liu G, **Johnson WE**, Crandall KA. Integrating microbial and host transcriptomics to characterize asthma-associated microbial communities. *BMC Medical Genomics*. 2015 Aug 16;8:50. PubMed PMID: 26277095; PubMed Central PMCID: PMC4537781.
60. MacNeil SM, **Johnson WE**, Li DY, Piccolo SR, Bild AH. Inferring pathway dysregulation in cancers from multiple types of omic data. *Genome Med*. 2015;7(1):61. doi: 10.1186/s13073-015-0189-4. eCollection 2015. PubMed PMID: 26170901; PubMed Central PMCID: PMC4499940.
61. 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 Jun 1;31(11):1745-53. PubMed PMID: 25617415; PubMed Central PMCID: PMC4443674.
62. Mortenson JB, Heppler LN, Banks CJ, Weerasekara VK, Whited MD, Piccolo SR, **Johnson WE**, Thompson JW, Andersen JL. Histone Deacetylase 6 (HDAC6) Promotes the Pro-survival Activity of 14-3-3 ζ via Deacetylation of Lysines within the 14-3-3 ζ Binding Pocket. *Journal of Biological Chemistry*. 2015 May 15;290(20):12487-96. PubMed PMID: 25770209; PubMed Central PMCID: PMC4432270.
63. Hong C, Manimaran S, **Johnson WE**. PathoQC: Computationally Efficient Read Preprocessing and Quality Control for High-Throughput Sequencing Data Sets. *Cancer Informatics*. 2015 May 12;13(Suppl 1):167-76. doi: 10.4137/CIN.S13890. eCollection 2014. Review. PubMed PMID: 25983538; PubMed Central PMCID: PMC4429651.
64. Whipple JM, Youssef OA, Aruscavage PJ, Nix DA, Hong C, **Johnson WE**, Bass BL. Genome-wide profiling of the C elegans dsRNAome. *RNA*. 2015 May;21(5):786-800. PubMed PMID: 25805852; PubMed Central PMCID: PMC4408787.

65. 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.
66. 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*. 2015;3:e761. PubMed PMID: 25737810; PubMed Central PMCID: PMC4338773.
67. 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, Sciruba 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.
68. Anand S, Bench Alvarez TM, **Johnson WE**, Esplin MS, Merrell K, Porter TF, Graves SW. Serum biomarkers predictive of pre-eclampsia. *Biomark Med*. 2015;9(6):563-75. doi: 10.2217/bmm.15.21. PubMed PMID: 26079961.
69. 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.
70. 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.
71. 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.
72. 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).
73. 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.
74. 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.
75. 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.
76. 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.
77. 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.

78. 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.
79. 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.
80. 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.
81. 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.
82. **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.
83. 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.
84. 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.
85. 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.
86. Warf MB, **Johnson WE**, Bass BL. Improved annotation of *C elegans* microRNAs by deep sequencing reveals structures associated with processing by Droscha and Dicer. *RNA*. 2011 Apr;17(4):563-77. PubMed PMID: 21307183; PubMed Central PMCID: PMC3062169.
87. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, **Johnson WE**, Geman D, Baggerly K, Izarray 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.
88. 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.
89. Thyagarajan B, Blaszcak 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.
90. 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.

91. 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.
92. **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.
93. 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.
94. **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.
95. 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)
96. **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.

Preprints and Technical Reports

1. Wang X, VanValkenberg A, Odom-Mabey AR, Ellner JJ, Hochberg NS, Salgame P, Patil P, **Johnson WE**. Comparison of gene set scoring methods for reproducible evaluation of multiple tuberculosis gene signatures. *bioRxiv*. 2023 Jan 30;. doi: 10.1101/2023.01.19.520627. PubMed PMID: 36711818; PubMed Central PMCID: PMC9882404. <https://www.biorxiv.org/content/10.1101/2023.01.19.520627v1>
2. McClintock J, Odom-Mabey AR, Keber N, Ismail A, Mwananyanda L, Gill CJ, MacLeod WB, Pieciak RC, Lapidot R, **Johnson WE** (2022). Post-mortem Nasopharyngeal Microbiome Analysis of Zambian Infants with and without Respiratory Syncytial Virus Disease: A Nested Case Control Study. *medRxiv* 2022.12.23.22283745. <https://doi.org/10.1101/2022.12.23.22283745>
3. Hawkins DY, Zuch DT, Huth J, Rodriguez-Sastre N, McCutcheon KR, Glick A, Lion AT, Thomas CF, Descoteaux AE, **Johnson WE**, Bradham CA (2022). ICAT: A Novel Algorithm to Robustly Identify Cell States Following Perturbations in Single Cell Transcriptomes. *bioRxiv* 2022.05.26.493603; doi: <https://doi.org/10.1101/2022.05.26.493603>

4. Faits T, Odom-Mabey AR, Castro-Nallar E, Crandall KA, **Johnson WE** (2022). Metagenomic profiling pipelines improve taxonomic classification for 16S amplicon sequencing data. *bioRxiv* 2022.07.27.501757; doi: <https://doi.org/10.1101/2022.07.27.501757>
5. Chitale P, Lemenze AD, Fogarty EC, Shah A, Grady C, Odom-Mabey AR, **Johnson WE**, Yang JH, Eren AM, Brosch R, Kumar P, Alland D (2022). A comprehensive update to the Mycobacterium tuberculosis H37Rv reference genome. *bioRxiv* 2022.07.15.500236; doi: <https://doi.org/10.1101/2022.07.15.500236>
6. Wang Y, Sarfraz I, Hong R, Koga Y, Akavoor V, Cao X, Alabdullatif S, Pervaiz N, Zaib SA, Wang Z, Jansen F, Yajima M, **Johnson WE**, Campbell JD. (2022). Interactive analysis of single-cell data using flexible workflows with SCTK2.0. *bioRxiv* 2022.07.13.499900; doi: <https://doi.org/10.1101/2022.07.13.499900>
7. Lapidot R, Faits T, Ismail A, Khumalo Z, MacLeod W, Kwenda G, Mupila Z, Nakazwe R, Segrè D, **Johnson WE**, Thea DM, Mwananyanda L, Gill CJ. Nasopharyngeal Dysbiosis Precedes the Development of Lower Respiratory Tract Infections in Young Infants, a Longitudinal Infant Cohort Study. *medRxiv* 2021.10.13.21264939. <https://doi.org/10.1101/2021.10.13.21264939>
8. Li T, Zhang Y, Patil P, **Johnson WE** (2021). Overcoming the impacts of two-step batch effect correction on gene expression estimation and inference. *bioRxiv*, 2021.01.24.428009. <https://doi.org/10.1101/2021.01.24.428009>.
9. Sheerin D, Abhimanyu, Wang X, bf Johnson WE, Coussens A. Systematic evaluation of transcriptomic disease risk and diagnostic biomarker overlap between COVID-19 and tuberculosis: a patient-level meta-analysis. *medRxiv*. 2020 Nov 26;. <https://doi: 10.1101/2020.11.25.20236646>. PubMed PMID: 33269371; PubMed Central PMCID: PMC7709192.
10. Zhang, Y, Parmigiani, G, **Johnson, WE** (2020). ComBat-Seq: batch effect adjustment for RNA-Seq count data. *bioRxiv*, 2020.01.13.904730. <https://doi.org/10.1101/2020.01.13.904730>.
11. Wang Z, Yang S, Koga Y, Corbett SE, **Johnson WE**, Yajima M, Campbell JD (2020). Celda: A Bayesian model to perform bi-clustering of genes into modules and cells into subpopulations using single-cell RNA-seq data. *bioRxiv* 2020.11.16.373274; doi: <https://doi.org/10.1101/2020.11.16.373274>
12. Hong R, Koga Y, Bandyadka S, Leshchuk A, Wang Z, Alabdullatif S, Wang Y, Akavoor V, Cao X, Sarfraz I, Jansen F, **Johnson WE**, Yajima M, Campbell JD (2020). Comprehensive generation, visualization, and reporting of quality control metrics for single-cell RNA sequencing data. *bioRxiv* 2020.11.16.385328; doi: <https://doi.org/10.1101/2020.11.16.385328>
13. Zhao Y, Federico A, Faits T, Manimaran S, Monti S, **Johnson WE** (2020). animalcules: Interactive Microbiome Analytics and Visualization in R. *bioRxiv* 2020.05.29.123760; doi: <https://doi.org/10.1101/2020.05.29.123760>
14. Ruan, QT, Yazdani, N, Reed, ER, Beierle, JA, Peterson, LP, Luttik, KP, Szumlinski, KK, **Johnson, WE**, Ash, PE, Wolozin, B, Bryant, CD (2020). Functional variants within the 5' UTR support reduced exon usage as a molecular mechanism underlying the quantitative trait gene Hnrnp1 for reduced methamphetamine sensitivity. *bioRxiv*, 2020.01.11.902908. <https://doi.org/10.1101/2020.01.11.902908>.
15. Zhang, Y., **Johnson, WE**, Parmigiani, G (2019). Robustifying genomic classifiers to batch effects via ensemble learning. *bioRxiv*, 703587. <https://doi.org/10.1101/703587>.
16. Yang, S, Corbett, SE, Koga, Y, Wang, Z., **Johnson, WE**, Yajima, M, Campbell, JD (2019). Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *bioRxiv*, 704015. <https://doi.org/10.1101/704015>.
17. Wang, Z, Hu, J, **Johnson, WE**, Campbell, JD (2019). scruff: An R/Bioconductor package for preprocessing single-cell RNA-sequencing data. *bioRxiv*, 522037. <https://doi.org/10.1101/522037>.

18. Bryant, CD, Bagdas, D, Goldberg, LR, Khalefa, MT, Reed, MR, Kirkpatrick, SL, Kelliher, JC, Chen, MM, **Johnson, WE**, Mulligan, MK, Damaj, MI (2018). C57BL/6 substrain differences in inflammatory and neuropathic nociception and genetic mapping of a major quantitative trait locus underlying acute thermal nociception. **bioRxiv**, 413856. <https://doi.org/10.1101/413856>.
19. Jenkins, DF, Faits, T, Briars, E, Carrasco Pro, S, Cunningham, S, Campbell, JD, Yajima, M, **Johnson, WE**. (2018). Interactive single cell RNA-Seq analysis with the Single Cell Toolkit (SCTK). *bioRxiv*, 329755. <https://doi.org/10.1101/329755>.
20. 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.
21. 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.
22. 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

Personal GitHub Repository (29 projects): <https://github.com/wevanjohnson>

Lab/Section GitHub Repository: <https://github.com/compbioem>

1. **ComBat, Combat-Seq, BatchQC**: *Combining Multiple Batches of Microarray Genomic Data*. ComBat is a 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-throughput studies. ComBat-Seq uses a Negative Binomial error model to extend ComBat to sequencing-based experiments. BatchQC is a user-interface for interactive evaluation of batch effects in -omic data. These tools are highly cited (>8,000 citations), widely-used, and commonly downloaded (107,477 package downloads in 2020; top 3% of all packages in R/Bioconductor).
2. **PathoScope, animalcules**: *Metagenomic profiling and multi-sample analysis*. 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. animalcules is new a graphical user interface for multi-sample metagenomic data analysis.
3. **singleCellTK**: *Tools and Workflow for Single Cell RNA-seq Analysis*. The singleCellTK is an NCI-funded project to construct a comprehensive and interactive R-software framework for complete data processing and analysis of single cell RNA-sequencing data from heterogeneous tumor samples. We have developed the singleCellTK with an R/Shiny user interface that enables interactive analysis and visualization in the data.
4. **ASSIGN, TBSignatureProfiler, curatedTBData**: *Signature scoring, curation, and validation in TB research*. ASSIGN utilizes Bayesian factor regression model to identify genomic biomarkers for applications in pathway profiling, drug responsiveness, environmental exposure, and infectious disease diagnosis. The TBSignatureProfiler is a compilation of more than 50 signatures of TB outcomes and includes package tools for scoring samples and visualizing the results. The curatedTBData is an effort to compile and harmonize data from more than 45 datasets with more than 4,000 samples. The curatedTBData can be combined with the TBSignatureProfiler to generate or validate new signatures, evaluate existing signatures on subsets or to provide data for other projects.

5. **SCAN-UPC, MAT, MA2C:** *Single and Multi-Channel Array Normalization and barcoding.* 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 barcodes that estimate gene activity in data from microarray and RNA-sequencing platforms. MAT is designed for the analysis of data from Affymetrix tiling microarrays. The MA2C software is a similar approach but designed to analyze data from two-color tiling arrays.
6. **GNUMAP:** *Genomic Next-generation Universal Mapper.* GNUMAP 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.

PRESENTATIONS

Invited Presentations: International

1. Novartis Institutes for BioMedical Research (2021), Basel, Switzerland (presented remotely)
Talk Title: *Methods for Metatranscriptomic Profiling: Applications in Pulmonary Diseases*
2. International Chinese Statistical Association, (2020; presented virtually due to COVID19)
Talk Title: *Addressing unwanted heterogeneity in genomic data: applications in RNA-sequencing and microbiome research*
3. RePORT India 9th Annual Joint Leadership Meeting, PD Hinduja Hospital (2020), Mumbai, India
Talk Title: *Evaluation of gene signatures for predicting the progression of TB from infection to disease*
4. Department of Preventive and Social Medicine, Jawaharlal Institute of Postgraduate Medical Education & Research (2019), Pondicherry, India
Talk Title: *Parsimonious Gene Signatures for TB Outcomes*
5. 50th Union World Conference on Lung Health, (2019), Hyderabad, India
Talk Title: *Parsimonious Gene Signatures for TB Outcomes*
6. RePORT India 8th Annual Joint Leadership Meeting, Raintree Anna Salai (2019), Chennai, India
Talk Title: *Pathway signature profiling of tuberculosis RNA-Seq data*
7. Department of Preventive and Social Medicine, Jawaharlal Institute of Postgraduate Medical Education & Research (2018), Pondicherry, India
Talk Title: *Pathway signature profiling of tuberculosis RNA-Seq data*
8. RePORT India 7th Annual Joint Leadership Meeting, International Centre for Genetic Engineering and Biotechnology (ICGEB) (2017), New Delhi, India
Talk Title: *Pathway signature profiling of tuberculosis RNA-Seq data*
9. 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*

Invited Presentations: National

1. Liu Lab Reunion, (2022), Boston MA
Talk Title: *Profiling host-microbe interactions in pulmonary diseases*
2. 17th Coordinated Science Lab (CSL) Student Conference, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL (**Keynote**, presented remotely)
Talk Title: *Profiling host-microbe interactions in pulmonary diseases*
3. Clinical Trials Network (CTN), VA Boston Healthcare System (2022), Boston MA (presented remotely)
Talk Title: *Profiling host-microbe interactions in pulmonary diseases*

4. Center for Inflammation and Immunity, (2021), Rutgers University, NJ. (presented remotely)
Talk Title: *Profiling host-microbe interactions in pulmonary diseases*
5. Tuberculosis Research Units Network (TBRU-N) Annual Meeting, (2021), Washington, D.C. (presented remotely)
Talk Title: *Methods for Metatranscriptomic Profiling: Applications in Pulmonary Diseases*
6. Data Science Workshop, National Cancer Institute (2019), Washington, D.C. (presented remotely)
Talk Title: *Data Science and Single Cell RNA-sequencing*
7. Cancer Systems Biology Consortium Workshop, (2019), Portland, OR. (presented remotely)
Talk Title: *Interactive scRNA-Seq analysis with the Single Cell Toolkit (SCTK)*
8. Informatics Technology for Cancer Research Annual Meeting, (2019), Salt Lake City, UT
Talk Title: *Interactive scRNA-Seq analysis with the Single Cell Toolkit (SCTK)*
9. Tuberculosis Research Units Network (TBRU-N) Annual Meeting, (2019), Washington, D.C.
Talk Title: *Signature profiling using nanostring technology: the pan TB107 profiler*
10. Cancer Systems Biology Consortium Workshop, (2018), Park City, UT
Talk Title: *Interactive scRNA-Seq analysis with the Single Cell Toolkit (SCTK)*
11. Informatics Technology for Cancer Research Annual Meeting, (2018), Washington, D.C.
Talk Title: *Interactive scRNA-Seq analysis with the Single Cell Toolkit (SCTK)*
12. NCI Tumor Heterogeneity Workshop, Oregon Health and Sciences University (2015), Portland, OR
Talk Title: *Modeling functional tumor evolution over time or across multiple metastatic sites*
13. International Chinese Statistical Association (2015), Fort Collins, CO
Talk Title: *PathoScope: Species identification and strain attribution with unassembled sequencing data*
14. Department of Biology, Brigham Young University (2015), Provo, UT
Talk Title: *Longitudinal profiling of functional pathways in cancer cell subpopulations*
15. Computational Biology Program, Oregon Health and Sciences University (2015), Portland, OR
Talk Title: *Longitudinal profiling of functional pathways in cancer cell subpopulations*
16. Biological Sciences Graduate Program, University of Maryland (2015), College Park, MD
Talk Title: *Longitudinal profiling of functional pathways in cancer cell subpopulations*
17. Computational Biology Institute, George Washington University (2015), Washington, D.C.
Talk Title: *Longitudinal profiling of functional pathways in cancer cell subpopulations*
18. Tumor Heterogeneity Workshop, Oregon Health and Sciences University (2014), Portland, OR
Talk Title: *Adaptive models for accounting for heterogeneity in biomarker development*
19. 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*
20. Cancer Biomarker Informatics Workshop, Early Detection Research Network (2013), Pasadena, CA
Talk Title: *Data preprocessing and integration for reproducible multiomic biomarker discovery and validation*
21. Joint Statistical Meetings, (2012), San Diego, CA
Talk Title: *Latent Variable Models for Predicting Drug Sensitivity and Hereditary Disease Risk*
22. International Biometrics Society, Eastern North American Spring Meeting (2011), Miami, FL
Talk Title: *Estimation of Genome-Wide Methylation Levels Using Next-generation Sequencing Data*
23. 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*

24. Division of Computational Biomedicine, Boston University (2011), Boston, MA
Talk Title: *Statistical and Computational Tools for Assessing Drug Sensitivity and Predicting Hereditary Disease Risk*
25. 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*
26. International Biometrics Society, Western North American Spring Meeting (2010), Seattle, WA
Talk Title: *SLAM: GausSian Dynamic Linear Analysis of Methylated Chip Data*
27. 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*
28. Joint Statistical Meetings (2010), Vancouver, CA
Talk Title: *The GNUMAP Project: Probabilistic Mapping of Next-Generation Sequencing Data*
29. 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*
30. Joint Statistical Meetings (2010), Vancouver, CA
Talk Title: *Impact of Statistics on Next-Generation Sequencing*
31. 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*
32. Joint Statistical Meetings (2009), Washington, D.C.
Poster Title: *Predicting Breast Cancer Using Exon Array Data*
33. Joint Statistical Meetings (2008), Denver, CO
Talk Title: *Normalizing and comparing multiple genome tiling microarrays*
34. Joint Statistical Meetings (2006), Seattle, WA
Talk Title: *Model-based analysis of tiling arrays for ChIP-chip*
35. 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*
36. Department of Statistics, Brigham Young University (2007), Provo, UT
Talk Title: *Doubly stochastic latent variable analysis of ChIP-chip experiments on genome tiling arrays*
37. 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*
38. Joslin Diabetes Center (2006), Boston, MA
Talk Title: *Adjusting batch effects in microarray data: An Empirical Bayes approach*
39. Mathematical Biosciences Institute (2004), The Ohio State University
Talk Title: *Empirical Bayes methods for adjusting for batch effects in microarray experiments*

Invited Presentations: Regional

1. St. George Academy (2022), St. George UT
Talk Title: *Data Science: Profiling host-microbe interactions in pulmonary diseases*
2. Department of Mathematics, Southern Utah University (2021), Cedar City, UT
Talk Title: *Data Science in Biomedical Research: Profiling host-microbe interactions in pulmonary diseases*
3. Early Oncology Group, AstraZeneca (2020), Waltham, MA (presented remotely)
Talk Title: *Profiling tumor evolution and drug response in cancer cell subpopulations*

4. Department of Statistics, University of Connecticut (2019), Storrs, CT
Talk Title: *Addressing unwanted heterogeneity in genomic data: applications in RNA-sequencing and prediction*
5. Illumina Single Cell Workshop, (2018), Waltham, MA
Talk Title: *Interactive scRNA-Seq analysis with the Single Cell Toolkit (SCTK)*
6. Behavior Genetics Association Annual Meeting (2018), Cambridge, MA
Talk Title: *The Genomics of Neonatal Abstinence Syndrome*
7. Early Oncology Group, AstraZeneca (2018), Waltham, MA
Talk Title: *Growth factor receptor network profiling in bulk and single cells*
8. Boston Bioconductor Meetup (2017), Boston, MA
Talk Title: *Introduction to R/Shiny*
9. Single Cell Congress, (2018), Boston, MA
Talk Title: *Interactive scRNA-Seq analysis with the Single Cell Toolkit (SCTK)*
10. New England TB Day, (2017) Boston, MA
Talk Title: *Transcriptional Profiling for Biomarkers of TB resistance*
11. Biostats & Comp Bio, Dana Farber Cancer Inst, Harvard School of Public Health (2015), Boston, MA
Talk Title: *Longitudinal profiling of functional pathways in cancer cell subpopulations*
12. 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*
13. Department of Biostatistics, Harvard University (2012), Boston, MA
Talk Title: *Predicting hereditary breast-cancer development via peripheral-blood transcription profiling*
14. 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*
15. Department of Biostatistics, Boston University (2011), Boston, MA
Talk Title: *Dynamic Latent Models for Profiling Methylation Levels in Genome-wide Experiments*
16. 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*
17. Timpview High School (2011), Provo, UT
Talk Title: *Personalizing the Cure: Improving the Odds of Cancer Recovery using Statistics*
18. Southern Utah University (2010), Cedar City, UT
Talk Title: *Predicting Breast Cancer Risk based on Gene Expression Profiles*
19. Department of Oncological Sciences, Huntsman Cancer Institute (2007), Salt Lake City, UT
Talk Title: *Normalizing and comparing multiple genome tiling microarrays*
20. 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*
21. Department of Oncological Sciences, Huntsman Cancer Institute (2007), Salt Lake City, UT
Talk Title: *Topics in microarray analysis: ChIP-chip experiments and Batch Effects*
22. 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*
23. Department of Mathematics, Southern Utah University (2006), Cedar City, UT
Talk Title: *Methods for tiling array analysis*
24. National Science Teachers Association Regional Conference (2001), Salt Lake City, UT
Talk Title: *Let's Go on a Rainbow Hunt*

TEACHING AND MENTORING

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)

Department of Mathematics, Southern Utah University, 2021-2022

- *MATH 1040: Introduction to Statistical Inference*
- *MATH 3190: Foundations of Data Science*

Guest Lectures

- *BF 831: Translational Bioinformatics*, Department of Medicine, Boston University, Fall 2022
- *EP 755: Infectious Disease Epidemiology*, Dept. of Epidemiology, Boston University, Fall 2022
- *EP 755: Infectious Disease Epidemiology*, Dept. of Epidemiology, Boston University, Spring 2021
- *EP 755: Infectious Disease Epidemiology*, Dept. of Epidemiology, Boston University, Fall 2020
- *EP 755: Infectious Disease Epidemiology*, Dept. of Epidemiology, Boston University, Spring 2019
- *EP 755: Infectious Disease Epidemiology*, Dept. of Epidemiology, Boston University, Spring 2018
- *BF 831: Translational Bioinformatics*, Department of Medicine, Boston University, Fall 2018
- *BS 803: Statistical Computing*, Department of Biostatistics, Boston University, Spring 2018
- *BS 860: Applied Genetic Analysis*, Department of Biostatistics, Boston University, Spring 2013
- *BF 831: Translational Bioinformatics*, Department of Medicine, Boston University, Fall 2017
- *BF 831: Translational Bioinformatics*, Department of Medicine, Boston University, Fall 2015
- *BF 690: Bioinformatics Challenge Project*, Program in Bioinformatics, Boston University, 2012-2014
- *BF 831: Translational Bioinformatics*, Department of Medicine, Boston University, Fall 2014
- *BE 768: Biological Database Analysis*, Program in Bioinformatics, 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, Symposia, and Workshops

- *Data Science: Profiling host-microbe interactions in pulmonary diseases*. Makerere University (2022), Kampala, Uganda.
- *The analysis of RNA-sequencing data*. Tufts University (2019), Boston, MA.
- *The analysis of single cell RNA-sequencing data*. Boston University (2019), Boston, MA.
- *The analysis of RNA-sequencing data*. Tufts University (2018), Boston, MA.
- *Analyzing microbiome sequencing data*. Makerere University (2018), Kampala, Uganda.
- *Generating genomic biomarkers*. Makerere University (2018), Kampala, Uganda.
- *MicroBU*, BU Microbiome Day, Boston University (2017), Boston MA
- *Analyzing microbiome sequencing data*. Boston University SPH (2017), Boston, MA.
- *Processing and Visualization of Genomic Sequencing Data*. Tufts University (2017), Boston, MA.
- *Metagenomic profiling and rapid infectious disease diagnostics*. George Washington University (2017), Washington D.C.
- *BU Day of Data Science*, Boston University (2017), Boston MA
- *Software Carpentry Workshop*, Division of Computational Biomedicine, Boston University (2015), Boston, MA.
- *Low level Processing and Visualization of Genomic 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)

Junior Faculty Mentoring

- Shariq Muhammed, PhD, Assistant Professor, Biostatistics, BU School of Public Health (2021-2022)
- Ribka Berhanu, MD, Assistant Professor, Global Health, BU School of Public Health (K award Mentor; 2019-2022)
- Lindsey Locks, MPH, ScD, Assistant Professor, Department of Health Sciences, Sargent College (K award pending; 2019-2022)
- Prasad Patil, PhD, Assistant Professor, Biostatistics, BU School of Public Health (2019-2022)
- Jai Marathe, MD, Assistant Professor, Infectious Diseases, BU School of Medicine (CFAR Developmental Grant; 2018-2022)
- Kaku So-Armah, PhD, Research Assistant Professor, Internal Medicine, Boston University School of Medicine (CFAR Developmental Grant; 2017-2021)
- Jessica Leibler, DrPH, Assistant Professor, Environmental Health, BU School of Public Health (K award mentor; 2016-2021)

Postdoctoral Fellows

- Lucas Schiffer, Ph.D. (Rutgers University, 2022-present)
- Arthur VanValkeburg, Ph.D. (Boston University, 2020-2022)
- Tyler Faits, Ph.D. (Boston University, 2020-2021)
- Tenglong Li, Ph.D. (Boston University, 2018-2021)
- Ethel Nankya (Boston University, 2017-2018)
- Ying Shen, Ph.D. (Boston University, 2012-2014)
- Changjin Hong, Ph.D. (Boston University, 2011-2014)
- Stephen Piccolo, Ph.D. (Boston University, 2010-2014)
- 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

- Nisreen Khambati (Pediatrics, MD-PhD, Oxford University, 2022-present)
- Aubrey Odom (Bioinformatics, Boston University, 2021-present)
- Ethel Nankya (Bioinformatics, Boston University, 2020-present)
- Xutao Wang (Biostatistics, Boston University, 2019-present)
- Kiloni Quiles-Franco (Molecular Medicine, Boston University, 2019-present)
- Howard Fan (Bioinformatics, Boston University, 2019-present)
- Lucas Schiffer (Bioinformatics, Boston University, 2019-2022)
- Yue (Jason) Zhao (Bioinformatics, Boston University, 2017-2020)
- Yuqing Zhang (Bioinformatics, Boston University, 2016-2020)
- David Jenkins (Bioinformatics, Boston University, 2015-2019)
- Tyler Faits (Bioinformatics, Boston University, 2015-2020)
- Supriya Sharma (Molecular Medicine, Boston University, 2013-2017)
- Solaiappan Manimaran (Biostatistics, Boston University, 2012-2016)
- Michelle Withers (Statistics, Brigham Young University, 2008-2011)
- Rachel Poulsen (Statistics, Brigham Young University, 2008-2009)

Graduate Committees

- Devlin Moyer (Bioinformatics, Boston University, 2020-2022)
- Jacquelyn Turcinovic (Bioinformatics, Boston University, 2020-2022)
- Dominick Leone (Epidemiology, Boston University, 2020-present)
- Zachary Baucom (Biostatistics, Boston University, 2019-2022)
- Dakota Hawkins (Bioinformatics, Boston University, 2017-present)
- Joey Orofino (Biochemistry, Boston University, 2017-2021)
- Brian Haas (Bioinformatics, Boston University, 2018-2020)
- Zhe Wang (Bioinformatics, Boston University, 2018-2021)
- Eric Reed (Bioinformatics, Boston University, 2017-2020)
- Katherine Norwood (Bioinformatics, Boston University, 2016-2020)
- Maura Walker (Tufts University, 2016-2018)
- R Cooper Baer (Bioengineering, 2014-2020)
- Heather Selby (Bioinformatics, Boston University, 2014-2019)
- Allyson Byrd (Bioinformatics, Boston University, 2013-2018)
- Vinay Kartha (Bioinformatics, Boston University, 2013-2017)
- Akshaya Ramesh (Boston University, 2013-2017)
- Joseph Perez-Rogers (Bioinformatics, Boston University 2013-2018)
- Ana Pavel (Bioinformatics, Boston University, 2013-2017)
- Neema Yazdani (Boston University, 2013-2018)

- Greg Wasserman (Boston University, 2013-2017)
- Joshua Loving (Bioinformatics, Boston University, 2013-2018)
- Anna Lyubetskaya (Bioinformatics, Boston University, 2013-2017)
- Paula Griffin (Biostatistics, Boston University, 2013-2018)
- Kevin Tuttle (Biology, Brigham Young University, 2012-2016)
- Nacho Caballero (Bioinformatics, Boston University, 2012-2015)
- Yuxiang Tan (Bioinformatics, Boston University, 2012-2015)
- Teresa Wang (Bioinformatics, Boston University, 2012-2015)
- John Farrell (Bioinformatics, Boston University, 2012-2015)
- Matthew Bendall (Biology, Brigham Young University, 2011)
- Mark Nielson (Statistics, Brigham Young University, 2010-2011)
- Scott Morris (Statistics, Brigham Young University, 2009-2010)
- Rozaura Vivas Hall (Biology, Brigham Young University, 2008-2009)
- Rick Smith (Biology, Brigham Young University, 2008-2009)
- Scott Howard (Statistics, Brigham Young University, 2008)

Undergraduate Research Mentoring

- Anika Reuppell (University of Toronto, 2021-2023)
- Rahul Varki (Boston University, 2020-2022)
- Aubrey Odom (Boston University, 2019)
- Nebiyou Kebede (Boston University, 2019)
- Sarah Hilton (Boston University, 2015)
- Nicholas Dalton (Boston University, 2015)
- 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)

PROFESSIONAL MEMBERSHIPS AND SERVICE

Service to the Discipline

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

Study Sections

Special Emphasis Panel, ZCA1 SRB-1 (A1)1, Research Specialist Award (R50), NCI, 2021
 Special Emphasis Panel, ZCA1-SRB-P(01), National Cancer Institute, 2020
 Special Emphasis Panel, ZCA1-TCRB-T-M1, National Cancer Institute, 2020
 GCAT Study Section, National Institutes of Health, 2010, 2015, 2019
 Special Emphasis Panel, ZRG1-BST-R-03, National Cancer Institute, 2019
 Special Emphasis Panel, ZCA1 SRB-P(01), National Cancer Institute, 2019
 Special Emphasis Panel, ZCA1 TCRB-D M1, National Cancer Institute, 2019
 Special Emphasis Panel, ZCA1 SRB-P (O1), National Cancer Institute, 2018
 Space Biology Computational Biology Panel, National Aeronautics and Space Administration, 2018
 Special Emphasis Panel, ZCA1 SRB-C (J2), National Cancer Institute, 2017
 COBRE Phase I review, National Institute of General Medical Sciences, 2017
 TRI Research Topics Review, National Aeronautics and Space Administration, 2017
 Program Project (P01) Review, National Institute of Allergy and Infectious Diseases, 2017
 Physical Sciences-Oncology Projects (PS-OP) (U01), National Cancer Institute, 2016
 Modeling-Bioinformatics Panel, National Aeronautics and Space Administration, 2016
 COBRE Phase II review, National Institute of General Medical Sciences, 2016
 Special Emphasis Panel, ZCA1 TCRB-T (J1), National Cancer Institute, 2015
 Special Emphasis Panel, ZCA1 TCRB-T J2 S, National Cancer Institute, 2015
 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, Journal of the American Statistical Association, 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

Rutgers New Jersey Medical School Committees and Leadership

- Search Committee, Center for Emerging and Re-emerging Pathogens, Department of Medicine, 2022-Present
- Search Committee, Director for Biomedical Informatics, Rutgers Biomedical and Health Sciences, 2022-Present

Boston University Committees and Leadership

- Search Committee, Division of Biomedical Genetics, Department of Medicine, 2021
- Search Committee, Library Director of the Alumni Medical Library, BUMC, 2020-2021
- Steering Committee, Rafik B. Hariri Institute for Computing and Computational Science & Engineering, 2020-2022
- Search Committee, Faculty of Computing and Data Sciences, 2020
- Curriculum Committee, Faculty of Computing and Data Sciences, 2019-2022
- Research Computing Governance Committee, 2018-2022
- President's Task Force to Envision Data Science at Boston University, 2018-2019
- Co-scientific director of the Single Cell Sequencing Resource, 2018-2022
- Co-director of the BU campus-wide microbiome affinity research collaborative, 2017-2021
- Search Committee, ID Section Chief, Department of Medicine, 2016-2019
- Biostatistics PhD Program review task force, 2018
- Bioinformatics Academic Program review committee, 2018
- Organizing Committee Co-Chair, BU Day of Data Science, 2016-2017
- Bioinformatics Graduate Admission Committee, 2012-present

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, Committee 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 (2008-2011)

INNOVATION, ENTREPRENEURSHIP, EXPERT WITNESS

- **Co-Founder, Aperionics, Inc, Sterling VA.** Aperionics identifies every known bacteria, virus, fungus, and parasite through Deep Shotgun Metagenomic Sequencing, using our world renown database containing over 41,000 microbes. We help doctors and their patients identify the causes of infections that other tests cannot identify, thus streamlining the path to a positive clinical outcome. Many patients that we have helped were sick for years with no understanding of why. **Acquired by LabCorp in 2021.**
- **Provisional Patent: 61/693,014** Blood-based Biomarker of Breast Cancer Risk
- **Provisional Patent: 62/864,049,** Multi-Pathway Gene Expression Biomarkers for Cancer Outcomes.
- NSF/SBIR Grant #1534469: Rapid Pathogen Diagnostics and Biosurveillance using Multiplexed High-throughput Sequencing
- **Expert Witness:** in the case of Yanick Paquet (plaintiff) v. Life Technologies Corporation (Defendant), Suffolk Superior Court, Civil No. 14-01423-G