

# ZHIGUANG (CALEB) HUO

2004 Mowry Road, 5th Floor CTRB, P.O. Box 117450, Gainesville, FL 32611-7450

(352)-294-5929  $\diamond$  [zhuo@ufl.edu](mailto:zhuo@ufl.edu)

<https://caleb-huo.github.io>

## EDUCATION

---

**University of Pittsburgh,**

Pittsburgh, PA, US

- Ph.D. in [Biostatistics](#), *April 2017*
  - Dissertation: *Statistical integrative omics methods for disease subtype discovery*
  - GPA: *3.93/4.00*
  - Advisors: [George C. Tseng](#), ScD and [Yong Seok Park](#), PhD
- M.S. in [Physics](#), *Apr 2012*
  - GPA: *3.86/4.00*

**Harbin Institute of Technology,**

Harbin, Heilongjiang, China

- B.S. in [Physics](#), *June 2011*
  - GPA: *90.43/100*

## RESEARCH INTEREST

---

My research interest lies in the intersection between statistical methodology and its applications to genomics and bioinformatics. I am particularly interested in genomic data integration, models and variable selection in high-dimensional data, graphical models, Bayesian methods, optimization and statistical computing. I have collaborated with biologists in the fields of cancer and psychiatry, analyzing a broad range of genomic data. These experiences motivate me to develop methodology and software that are practical, user-friendly and easy to use.

## PROFESSIONAL EXPERIENCE

---

- Assistant Professor *July 2017 ~ Now*
  - Department of Biostatistics, University of Florida

## PUBLICATIONS

---

<sup>†</sup>: co-first author; \*: corresponding author.

### Submitted

1. **Zhiguang Huo**, Yun Zhu, Zanhua Huang, Lei Yu, Jingyun Yang, David A. Bennett, Jinying Zhao. Brain and blood metabolome for Alzheimer's dementia: Findings from a targeted metabolomics analysis.
2. Marianne L. Seney, Kelly Cahill, John F. Enwright III, **Zhiguang Huo**, George Tseng, and Colleen A. McClung. Novel Diurnal Rhythms in the Prefrontal Cortex in Schizophrenia Drive Differential Gene Expression.
3. William Paden, Kelly Barko, Rachel Puralewski, Kelly Cahill, **Zhiguang Huo**, Micah Shelton, George Tseng, Ryan Logan, Marianne Seney. Developmental origin of sex differences in adult mood and in stress-induced transcriptional coherence across mesocorticolimbic circuitry.
4. Nader Shaikh, ..., George Tseng, ..., **Zhiguang Huo**, ..., Hans Pohl. Host and bacterial markers that differ in children with cystitis and pyelonephritis.

---

<sup>0</sup>Last modified: January 22, 2019

5. Li Zhu, **Zhiguang Huo**, Tianzhou Ma, George Tseng. Bayesian indicator variable selection to incorporate multi-layer overlapping group structure in multi-omics applications.
6. **Zhiguang Huo\***, Li Zhu, Tianzhou Ma, Hongcheng Liu, Song Han, Daiqing Liao, Jinying Zhao and George Tseng\*. Two-way Horizontal and Vertical Omics Integration for Disease Subtype Discovery.
7. **Zhiguang Huo**, Shaowu Tang, Yongseok Park, George Tseng. P-value evaluation, variability index and biomarker categorization for adaptively weighted Fisher's meta-analysis method in omics applications.

## 2019

1. **Zhiguang Huo**, Yun Zhu, Lei Yu, Jingyun Yang, Philip De Jager, David A. Bennett, Jinying Zhao (2019). DNA methylation variability in Alzheimer's Disease. *Neurobiology of Aging*.
2. **Zhiguang Huo**, Chi Song\*, George C. Tseng\*. (2019) Bayesian latent hierarchical model for transcriptomic meta-analysis to detect biomarkers with clustered meta-patterns of differential expression signals. *Annals of Applied Statistics*.
3. Cynthia R. Johnson, Kimberly Brown, Susan Hyman, Maria M. Brooks, Courtney Aponte, Lynne Levato, Brianna Schmidt, Victoria Evans, **Zhiguang Huo**, Roxanna Bendixen, Heather Eng, Theresa Sax, Tristram Smith (2019). Parent Training for Feeding Problems in Children with Autism Spectrum Disorder: Initial Randomized Trial. *Journal of pediatric psychology*.
4. Tianzhou Ma<sup>†</sup>, **Zhiguang Huo**<sup>†</sup>, Anche Kuo<sup>†</sup>, Li Zhu, Fang Zhou, Xiangrui Zeng, Chien-Wei Lin, Silvia Liu, Lin Wang, Tanbin Rahman, Lun-Ching Chang, Sunghwan Kim, Jia Li, Yongseok Park, Chi Song, Steffi Oesterreich, Etienne Sibille and George C. Tseng. (2019). MetaOmics - Comprehensive Analysis Pipeline and Web-based Software Suite for Transcriptomic Meta-Analysis. *Bioinformatics*
5. Marianne Seney, **Zhiguang Huo**, Kelly Cahill, Leon French, Rachel Puralewski, Joyce Zhang, Ryan W. Logan, George Tseng, David A Lewis, Etienne Sibille. (2019) Opposite molecular signatures of depression in men and women. *Biological Psychiatry*

## 2018

1. John Enwright, **Zhiguang Huo**, Dominique Arion, John P. Corradi, George Tseng, and David A. Lewis. Transcriptome alterations of prefrontal cortical parvalbumin neurons in schizophrenia. *Molecular psychiatry* 23, no. 7 (2018): 1606.
2. Kelly Cahill<sup>†</sup>, **Zhiguang Huo**<sup>†</sup>, George Tseng, Ryan W. Logan\*, Marianne L. Seney\* (2018), Improved identification of concordant and discordant gene expression signatures using an updated rank-rank hypergeometric overlap approach. *Scientific Reports* 8.1 (2018): 9588.
3. Wang, Kai Wen, Xiangrui Zeng, Xiaodan Liang, **Zhiguang Huo**, Eric P. Xing, and Min Xu (2018). Image-derived generative modeling of pseudo-macromolecular structures-towards the statistical assessment of Electron CryoTomography template matching. *The British Machine Vision Conference*

## 2017

4. **Zhiguang Huo**, George C. Tseng. (2017) Integrative Sparse  $K$ -means with overlapping group lasso in genomic applications for disease subtype discovery. *The Annals of Applied Statistics*, 11(2), 1011-1039.
5. SungHwan Kim, Dongwan Kang, **Zhiguang Huo**, Yongseok Park, and George C. Tseng. Meta-analytic principal component analysis in integrative omics application. *Bioinformatics* 34, no. 8 (2017): 1321-1328.
6. Dominique Arion, **Zhiguang Huo**, John F. Enwright, John P. Corradi, George Tseng, and David A. Lewis. Transcriptome alterations in prefrontal pyramidal cells distinguish schizophrenia from bipolar and major depressive disorders. *Biological psychiatry* 82, no. 8 (2017): 594-600.

## 2016

7. **Zhiguang Huo**, Ying Ding, Silvia Liu, Steffi Oesterreich, and George Tseng. Meta-Analytic Framework for Sparse  $K$ -Means to Identify Disease Subtypes in Multiple Transcriptomic Studies. *Journal of the American Statistical Association*, 111, no. 513 (2016): 27-42.
8. Zhu, Li, Ying Ding, Cho-Yi Chen, Lin Wang, **Zhiguang Huo**, SungHwan Kim, Christos Sotiriou, Steffi Oesterreich, and George C. Tseng. "MetaDCN: meta-analysis framework for differential co-expression network detection with an application in breast cancer." *Bioinformatics* 33, no. 8 (2016): 1121-1129.

## 2015 and before

9. Silvia Liu, Wei-Hsiang Tsai, Ying Ding, Rui Chen, Zhou Fang, **Zhiguang Huo**, SungHwan Kim, Tianzhou Ma, Ting-Yu Chang, Nolan Michael Friedigkeit, Adrian V. Lee, Jianhua Luo, Hsei-Wei Wang, I-Fang Chung, George C. Tseng. (2015). Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Research*, 10.1093/nar/gkv1234.
10. Tiffany A. Katz, Serena G. Liao, Vincent J. Palmieri, Robert K. Dearth, Thushangi Pathiraja, **Zhiguang Huo**, Patricia Shaw, Sarah Small, Nancy E. Davidson, David G. Peters, George C. Tseng, Steffi Oesterreich, Adrian V. Lee. (2015) Targeted DNA Methylation Screen in the Mouse Mammary Genome Reveals a Parity-Induced Hypermethylation of IGF1R That Persists Long after Parturition. *Cancer Prevention Research* 8, no. 10 (2015): 1000-1009.
11. Yan P. Yu, Silvia Liu, **Zhiguang Huo**, Amantha Martin, Joel B. Nelson, George C. Tseng and Jian-Hua Luo. (2015) Genomic copy number variations in the genomes of leukocytes predict prostate cancer clinical outcomes. *PloS one*, 10(8):e0135982.
12. SungHwan Kim, **Zhiguang Huo**, YongSeok Park and George Tseng. (2015) MetaOmics: transcriptomic meta-analysis methods for biomarker detection, pathway analysis and other exploratory purposes. Book chapter in Integrating omics data: statistical and computational methods. Edited by George C. Tseng, Debashis Ghosh, Xianghong Jasmine Zhou. *Cambridge University Press*. Page 39-67.
13. Xingbin Wang, Dongwan Kang, Kui Shen, Chi Song, Shuya Lu, Lunching Chang, Serena G. Liao, **Zhiguang Huo**, Naftali Kaminski, Etienne Sibille, Yan Lin, Jia Li and George C. Tseng. (2012) A Suite of R Packages for Quality Control, Differentially Expressed Gene and Enriched Pathway Detection in Microarray Meta-analysis. *Bioinformatics*, 28:2534-2536.

## AWARD

### Student Awards

- Delta Omega Membership April 2017
- American Statistics Association (ASA) Pittsburgh chapter March 2016
  - Student of the year
- Department of Physics, Harbin Institute of Technology May 2009
  - National Scholarship of P.R. China.
  - (Awarded to the top 2 students in my Bachelors degree.)

### Travel Awards

- 2018 ASA Biometrics Section JSM Travel Award, Vancouver, BC, Canada. Aug 2018
- Objective Bayes meeting 2017 travel award, Austin, TX. Dec 2017
- SAMSI, Research Triangle Park, NC.
  - Interface of Statistics and Optimization Feb 2017
  - Optimization Summer School Aug 2016
  - Epigenetics Workshop Mar 2015
  - Beyond Bioinformatics Workshop June 2014

## TEACHING EXPERIENCE

---

### Lecturer, University of Florida

- PHC6068 - Biostatistical computing Fall 2018
- PHC6937 - Frontiers in Biostatistics Feb 2018
- PHC6068 - Biostatistical computing Fall 2017

### Lecturer, University of Pittsburgh

- BOST2094 - Advanced R Computing – (with Tianzhou Ma) Spring 2017
- BOST2025 - Special Studies in Bayesian Data Analysis Fall 2016
  - (with George Tseng, Tianzhou Ma and Li Zhu)

### Guest Lecturer, University of Pittsburgh

- BOST2055 - Introductory high-throughput genomic data analysis I: data mining and applications Mar 2016
  - Differential and isoform analysis of RNA-seq data
- BOST2078 - Introductory high-throughput genomic data analysis II: theories and algorithms Dec 2015
  - Reproducible research and parallel computing in R
- BOST2078 - Introductory high-throughput genomic data analysis II: theories and algorithms Dec 2014
  - Reproducible research

### Teaching Assistant, University of Pittsburgh

- BOST 2078 - Introductory high-throughput genomic data analysis II: theories and algorithms Fall 2014
- PHYS 0212 - Introduction to Laboratory Physics Spring 2012
- PHYS 0212 - Introduction to Laboratory Physics Fall 2011

## PRESENTATIONS

---

### Invited Oral Presentation

- Department of Physiology and Functional Genomics at UF, Gainesville, FL October 2018
  - Statistical methods on omics data meta-analysis and integration, for disease subtype discovery and differential expression analysis.
- Joint Statistical Meetings, Vancouver, BC, Canada August 2018
  - Bayesian meta-analysis for biomarkers of meta-patterns.
- International Chinese Statistical Association conference, New Brunswick, NJ June 2018
  - Two-way Horizontal and Vertical Omics Integration for Disease Subtype Discovery.
- Rush Alzheimer's Disease Center ROSMAP Investigator's Meeting, Chicago, IL May 2018
  - Identifying variably methylated regions (VMRs) associated with AD.
- International Indian Statistical Association conference, Gainesville, FL May 2018
  - Two-way Horizontal and Vertical Omics Integration for Disease Subtype Discovery.
- Department of Biostatistics, University of Florida, Gainesville, FL Feb 2017
  - Meta-analytic and integrative framework for sparse  $K$ -means to identify disease subtypes.

### Contributed Poster and Oral Presentation

- Poster, ENAR conference, Atlanta, GA Mar 2018
  - p-value evaluation, variability index and biomarker categorization for adaptively weighted Fisher's meta-analysis method in omics applications
- Poster, 2018 UF Stats Winter Workshop, University of Florida, Gainesville, FL Jan 2018

- Bayesian latent hierarchical model for transcriptomic meta-analysis to detect biomarkers with clustered meta-patterns of differential expression signals
- Poster, Objective Bayes meeting 2017, University of Texas, Austin, TX Dec 2017
  - Bayesian latent hierarchical model for transcriptomic meta-analysis to detect biomarkers with clustered meta-patterns of differential expression signals
- Poster, Dean’s Day’s competition, GSPH, University of Pittsburgh April 2017
  - Circadian rhythms of gene expression in the human prefrontal cortex reveal distinct pattern between schizophrenia and control subjects
- Oral Presentation, ENAR, Washington, DC Mar 2017
  - Bayesian latent hierarchical model for transcriptomic meta-analysis to detect biomarkers with clustered meta-patterns of differential expression signals.
- Poster, SAMSI optimization summer school, Research Triangle Park, NC Aug 2016
  - Integrative Sparse  $K$ -means for disease subtype discovery using multi-level omics data.
- Poster, Pittsburgh ASA banquet, Pittsburgh, PA Mar 2016
  - Integrative Sparse  $K$ -means for disease subtype discovery using multi-level omics data.
- Department of Biostatistics, University of Pittsburgh Nov 2015
  - How to use Latex to make slides
- Oral Presentation, JSM, Seattle, WA Aug 2015
  - Meta-analytic framework for sparse  $K$ -means to identify disease subtypes in multiple transcriptomic studies.
- Poster, Pittsburgh ASA banquet, Pittsburgh, PA Apr 2015
  - Meta-analytic framework for sparse  $K$ -means to identify disease subtypes in multiple transcriptomic studies.
- Oral Presentation, ENAR Conference, Miami, FL Mar 2015
  - Meta-analytic framework for sparse  $K$ -means to identify disease subtypes in multiple transcriptomic studies.
- Poster, Dean’s Day’s competition, GSPH, University of Pittsburgh Mar 2015
  - Discover and Characterize Invasive Lobular Breast Carcinoma Subtypes.
- Oral Presentation, ENAR Conference, Baltimore, MA Mar 2014
  - Meta-analytic framework for sparse  $K$ -means to identify disease subtypes in multiple transcriptomic studies.
- Poster, Dean’s Day’s competition, GSPH, University of Pittsburgh Mar 2014
  - Meta-analytic framework for sparse  $K$ -means to identify disease subtypes in multiple transcriptomic studies.

## CONFERENCE SERVICE

---

- ICSA, New Brunswick, NJ June 2018
  - Chair for Session New developments in microbiome sequencing data modeling and analysis.
  - Chair for Session Modern Statistical Development for Biomedical Big data.

## REVIEWER SERVICE

---

- Journal of the American Statistical Association (1)
- Journal of the Royal Statistical Society Series C (1)
- Biometrics (1)
- Bioinformatics (1)
- BMC Bioinformatics (1)
- PLOS ONE (2)
- Scientific Reports (2)
- Journal of Alzheimer’s Disease (1)

## DEPARTMENT SERVICE

---

- Member of PhD Qualifying Exam committee *Jan 2018 ~ Now*
  - Department of Biostatistics, University of Florida
- Director of the Biostatistics Consulting Lab *Sep 2017 ~ Now*
  - Department of Biostatistics, University of Florida

## MEMBERSHIP

---

- Member of International Chinese Statistical Association *Mar 2015 ~ Now*
- Member of Eastern North American Region International Biometric Society *Oct 2013 ~ Now*
- Member of American Statistical Association *Nov 2013 ~ Now*

## HOBBIES

---

Running, skiing, other endurance activities.

### Running Record

- Las Olas International Triathlon, Fort Lauderdale, FL. 03/10/2019, Coming soon!.
- Five Points of Life Half Marathon, Gainesville, FL. 02/16/2019, Coming soon!.
- The Southernmost Marathon, Key West, FL. 10/06/2018
- Dick's Sporting Goods Pittsburgh Half Marathon, Pittsburgh, PA. 05/06/2018
- Five Points of Life Race Marathon, Gainesville, FL. 02/18/2018
- Bank of America Chicago Marathon, Chicago, IL. 10/09/2016
- First National Bank Pittsburgh Triathlon (Sprint), Pittsburgh, PA 08/14/2016
- Dick's Sporting Goods Pittsburgh Marathon, Pittsburgh, PA. 05/01/2016
- First National Bank Pittsburgh Triathlon (Olympic standard), Pittsburgh, PA 08/09/2015
- Dick's Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/03/2015
- Dick's Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/04/2014
- Dick's Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/05/2013