

# 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

---

- Director of the Biostatistics Consulting Lab *Sep 2017 ~ Now*
  - Department of Biostatistics, University of Florida
- Assistant Professor *July 2017 ~ Now*
  - Department of Biostatistics, University of Florida

## PUBLICATIONS

---

**2017**

1. Enwright, John, **Zhiguang Huo**, Dominique Arion, John Corradi, Aiqing He, George Tseng, and David Lewis. (2017) Transcriptome alterations of prefrontal cortical parvalbumin neurons in schizophrenia. *Molecular Psychiatry*.
2. Tianzhou Ma\*, **Zhiguang Huo\***, Anche Kuo\*, 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 and George C. Tseng. (2017). MetaOmics - Comprehensive Analysis Pipeline and Web-based Software Suite for Transcriptomic Meta-Analysis. (Submitted) (\*: co-first author).
3. **Zhiguang Huo**, Shaowu Tang, Yongseok Park, George Tseng. (2017) P-value evaluation, variability index and biomarker categorization for adaptively weighted Fisher's meta-analysis method in omics applications. (Submitted)

---

<sup>0</sup>Last modified: November 10, 2017

4. Marianne Seney, **Zhiguang Huo**, Leon French, Rachel Puralewski, Joyce Zhang, David A Lewis, George Tseng, Etienne Sibille. (2017) Distinct molecular signatures of depression in men and women. (Submitted)
5. **Zhiguang Huo**, Chi Song, George C. Tseng. (2017) Bayesian latent hierarchical model for transcriptomic meta-analysis to detect biomarkers with clustered meta-patterns of differential expression signals. Submitted to *Annals of Applied Statistics* (under revision).
6. **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.
7. Dominique Arion, **Zhiguang Huo**, John F. Enwright, John P. Corradi, George Tseng and David A. Lewis. (2017) Transcriptome alterations in prefrontal pyramidal neurons distinguish schizophrenia from bipolar and major depressive disorders. *Biological Psychiatry*,
8. SungHwan Kim, Dongwan Kang, **Zhiguang Huo**, Yongseok Park, George C. Tseng. (2017) Meta-analytic principal component analysis in integrative omics application. (minor revision).

## 2016

9. **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.
10. 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* (2016): btw788.

## 2015 and before

11. Silvia Liu, Wei-Hsiang Tsai, Ying Ding, Rui Chen, Zhou Fang, **Zhiguang Huo**, SungHwan Kim, Tianzhou Ma, Ting-Yu Chang, Nolan Michael Priedigkeit, 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.
12. 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.
13. 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.
14. 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.
15. Xingbin Wang, Dongwan Kang, Kui Shen, Chi Song, Shuya Lu, Luning 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.

## Under preparation

- YongSeok Park, **Zhiguang Huo**, Shaowu Tang and George Tseng. (2017) Asymptotic properties of adaptive weighted Fisher's method.

- Li Zhu, **Zhiguang Huo**, Tianzhou Ma, George Tseng. (2017) Bayesian indicator variable selection model with multi-layer overlapping groups.
- George C. Tseng, **Zhiguang Huo** and Tianzhou Ma. (2017) Foundations for High-Throughput Omics Data Analysis: Methods, Theories and Applications. *Chapman & Hall/CRC*.

## 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

- 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 (UNIVERSITY OF PITTSBURGH)

---

### Main Lecturer (teaching fellow)

- BIOST2094 - Advanced R Computing – (with Tianzhou Ma) Spring 2017
- BIOST2025 - Special Studies in Bayesian Data Analysis Fall 2016
  - (with George Tseng, Tianzhou Ma and Li Zhu)

### Guest Lecturer

- BIOST2055 - Introductory high-throughput genomic data analysis I: data mining and applications Mar 2016
  - Differential and isoform analysis of RNA-seq data
- BIOST2078 - Introductory high-throughput genomic data analysis II: theories and algorithms Dec 2015
  - Reproducible research and parallel computing in R
- BIOST2078 - Introductory high-throughput genomic data analysis II: theories and algorithms Dec 2014
  - Reproducible research

### Teaching Assistant

- BIOST 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

---

### Poster and Oral Presentation

- 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

- Invited talk, University of Florida, Gainesville, FL Feb 2017  
– Meta-analytic and integrative framework for sparse  $\ell_1/\ell_2$ -means to identify disease subtypes.
- 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.
- 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.

### Seminar Talk

- Department of Biostatistics, University of Pittsburgh Nov 2015  
– How to use Latex to make slides

## REVIEWER SERVICE

---

- PLOS ONE
- Scientific Reports
- Journal of the Royal Statistical Society Series C

## 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

- 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