

---

# ZHIGUANG (CALEB) HUO

4600 Bayard ST, APT 308, Pittsburgh, PA 15213

(412)-979-0592 ♦ zhh18@pitt.edu

<http://www.pitt.edu/~zhh18/dev/>

## EDUCATION

---

University of Pittsburgh,

Pittsburgh, PA, US

- Ph.D. in Biostatistics, *Expected:* Summer 2017
  - Dissertation: *Statistical integrative omics methods for disease subtype discovery and methods for single cell methylation*
  - 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 both statistical methodology and application on genomics and bioinformatics. I am particularly interested in genomic data integration, modeling and variable selection in high-dimensional data, graphical model, Bayesian methods, optimization and statistical computing. I have collaborated with biologists in the fields of cancer and psychiatry to analysis a broad range of genomic data, which motivates me to develop practical methodology and user-friendly software.

## PUBLICATIONS

---

### Methydology

1. **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.
2. **Zhiguang Huo**, Chi Song, George C. Tseng. (2016) 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 second round of review).
3. **Zhiguang Huo**, George C. Tseng. (2016) Integrative Sparse  $K$ -means for disease subtype discovery using multi-level omics data. Submitted to *Annals of Applied Statistics* (under second round of review).
4. Li Zhu, Ying Ding, Cho-Yi Chen, Lin Wang, **Zhiguang Huo**, SungHwan Kim, Christos Sotiriou, Steffi Oesterreich and George C. Tseng. (2016) MetaDCN: meta-analysis framework for differential coexpression network detection with an application in breast cancer. *Bioinformatics* (accepted).
5. SungHwan Kim, Dongwan Kang, **Zhiguang Huo**, Yongseok Park, George C. Tseng. (2016) Meta-analytic principal component analysis in integrative omics application. (Submitted).
6. 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.

- 
7. 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.

## Application

8. Dominique Arion, **Zhiguang Huo**, John F. Enwright, John P. Corradi, George Tseng and David A. Lewis. (2016) Transcriptome alterations in prefrontal pyramidal neurons distinguish schizophrenia from bipolar and major depressive disorders. Submitted to *Biological Psychiatry*, (under second round of review).
9. 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.
10. 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.

## Manuscript in preparation

### • Methydology

11. **Zhiguang Huo**, Shaowu Tang, YongSeok Park and George Tseng. (2017) Biomarker categorization and fast computing of adaptively weighted Fisher's method for meta-analysis in omics applications.
12. Li Zhu, **Zhiguang Huo**, Tianzhou Ma, George Tseng. (2017) Bayesian indicator variable selection model with multi-layer overlapping groups.
13. Tianzhou Ma, **Zhiguang Huo**, . . . , George Tseng. MetaOmics - a Comprehensive Software Suite with Interactive Visualization for Transcriptomic Meta-Analysis.
14. AW theory

### • Application

15. Enwright, John, Dominique Arion, John Corradi, Aiqing He, **Zhiguang Huo**, George Tseng, and David Lewis. Transcriptome Profiling of Layer 3 Parvalbumin Neurons from the Dorsolateral Prefrontal Cortex of Schizophrenia Subjects.
16. RNAseq96 from Marianne.
17. iPSC Ryan.

## Book

18. George C. Tseng, **Zhiguang Huo** and Tianzhou Ma. (2017) Foundations for High-Throughput Omics Data Analysis: Methods, Theories and Applications. Chapman & Hall/CRC. (Under preparation).

## Book Chapter

19. 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.

---

## AWARD

### Student Awards

- 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.
  - 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 Sep 2014 - Dec 2014
- PHYS 0212 - Introduction to Laboratory Physics Jan 2012 - Apr 2012
- PHYS 0212 - Introduction to Laboratory Physics Aug 2011 - Dec 2011

## PRESENTATIONS

### Poster and oral presentation

- 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

## REFERENCES

---

### George C. Tseng

Professor Phone: 412-624-5318  
 Department of Biostatistics (primary appointment)  
 Department of Human Genetics  
 Department of Computational & Systems Biology E-mail: ctseng@pitt.edu  
 University of Pittsburgh

### Yong Seok Park

Assistant Professor Phone: 412-624-3028  
 Department of Biostatistics E-mail: yongpark@pitt.edu  
 University of Pittsburgh

### David A. Lewis, MD

Distinguished Professor of Psychiatry and Neuroscience,  
 Thomas Detre Professor of Academic Psychiatry, Chair Phone: 412-246-6010  
 Department of Psychiatry E-mail: lewisda@upmc.edu  
 University of Pittsburgh

## HOBBIES

---

Running, skiing, endurance activities.

### Running record

- Bank of America Chicago Marathon, Chicago, IL. TIME 04:08:27 10/09/2016
- First National Bank Pittsburgh Triathlon (Sprint), Pittsburgh, PA 08/14/2016  
 TIME 1:10:30 (only biking + running)
- Dick’s Sporting Goods Pittsburgh Marathon, Pittsburgh, PA. TIME 3:58:46 05/01/2016
- First National Bank Pittsburgh Triathlon (Olympic standard), Pittsburgh, PA 08/09/2015  
 Time 3:29:24
- Dick’s Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/03/2015  
 TIME 3:49:38
- Dick’s Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/04/2014  
 TIME 4:07:44
- Dick’s Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/05/2013  
 TIME 4:48:49