MENGJIE CHEN

www.mengjiechen.com
5841 S. Maryland Ave ◇ N417A ◇ Chicago, IL 60637 USA
mengjiechen@uchicago.edu

Updated on Nov, 2018

EDUCATION

Yale University, New Haven, CT, USA *May 2014*
PhD in Computational Biology & Bioinformatics
Advisors: Hongyu Zhao, Haifan Lin

Huazhong University of Science and Technology, Wuhan, China *June 2009*
B.S in Biotechnology
Minor in Computer Science

PROFESSIONAL EXPERIENCE

June 2017 - now, Assistant Professor, Department of Human Genetics, University of Chicago

December 2016 - now, Assistant Professor, Section of Genetic Medicine, Department of Medicine, University of Chicago

August 2014 - December 2016, Assistant Professor, Department of Biostatistics, UNC-Chapel Hill

August 2014 - December 2016, Assistant Professor, Department of Genetics, UNC-Chapel Hill

June 2010 - May 2014, Research Assistant, Yale Center for Statistical Genomics and Proteomics

June 2010 - May 2014, Research Assistant, Yale Stem Cell Center

SELECTED HONORS AND AWARDS

- DMS/NIGMS interface of mathematics and biology (R01) 2017-2022
- UNC Junior Faculty Development Award 2015
- Student Marshal, Yale Graduate School of Arts and Sciences 2014
- China Scholarship Council-Yale World Scholarship 2009-2012
- National Scholarship in China 2007, 2008

JOURNAL PUBLICATIONS

* Joint first author

1. Chen M, Zhou X: **Variability-Preserving Imputation for Accurate Gene Expression Recovery in Single Cell RNA Sequencing Studies**. *Genome Biology (Accepted)* 2018
2. Yang C, Wan X, Lin X, Chen M, Zhou X, Liu J: **CoMM: a collaborative mixed model to dissecting genetic contributions to complex traits by leveraging regulatory information**. *Bioinformatics (Accepted)* 2018
3. Sun S, Zhu J, Mozaffari S, Ober C, Chen M, Zhou X: **Heritability Estimation and Differential Analysis with Generalized Linear Mixed Models in Genomic Sequencing Studies**. *Bioinformatics* 2018, :359265

4. Sun W, Bunn P, Jin C, Little P, Zhabotynsky V, Perou CM, Hayes DN, Chen M, Lin DY: **The association between copy number aberration, DNA methylation and gene expression in tumor samples**. *Nucleic Acids Research* 2018, :gky131, URL [+http://dx.doi.org/10.1093/nar/gky131]

5. Siegel MB, He X, Hoadley KA, Hoyle A, Pearce JB, Garrett AL, Kumar S, Moylan VJ, Brady CM, Van Swearingen AE, Marron D, Gupta GP, Thorne LB, Kieran N, Livasy C, Mardis ER, Parker JS, Chen M, Anders CK, Carey LA, Perou CM: **Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer**. *The Journal of Clinical Investigation* 2018, **128**(4)

6. Butler EN, Bensen JT, Chen M, Conway K, Richardson DB, Sun X, Geradts J, Olshan AF, Troester MA: **Pre-diagnostic smoking is associated with binary and quantitative measures of ER protein and ESR1 mRNA expression in breast tumors**. *Cancer Epidemiology and Prevention Biomarkers* 2017, :cebp–0404

7. Chen M, Zhou X: **Controlling for Confounding Effects in Single Cell RNA Sequencing Studies Using both Control and Target Genes**. *Scientific reports* 2017, **7**:13587

8. Chen M, Gao C, Ren Z: **Robust Covariance Matrix Estimation via Matrix Depth**. *the Annals of Statistics (Accepted)* 2017

9. *Silva GO, *Siegel MB, Mose LE, Parker JS, Sun W, Perou CM, Chen M: **SynthEx: a synthetic-normal-based DNA sequencing tool for copy number alteration detection and tumor heterogeneity profiling**. *Genome Biology* 2017, **18**:66, URL [https://doi.org/10.1186/s13059-017-1193-3]

10. *Chang J, *Tan W, *Lin Z, *Xi R, *Shao M, *Chen M, Luo Y, Zhao Y, Liu Y, Huang X, Xia Y, Hu J, Parker J, Marron D, Cui Q, Peng L, Chu J, Li H, Du Z, Han Y, Tan W, Liu Z, Zhan Q, Li YL, Mao W, Wu C, Lin D: **Comprehensive analysis of esophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and novel genomic alterations.** *Nature communications* 2017, **8**:15290

11. Dinh TA, Vitucci EC, Wauthier E, Graham RP, Pitman WA, Oikawa T, Chen M, Silva GO, Greene KG, Torbenson MS, et al.: **Comprehensive analysis of The Cancer Genome Atlas reveals a unique gene and non-coding RNA signature of fibrolamellar carcinoma**. *Scientific reports* 2017, **7**:44653

12. Chen M, Gao C, Ren Z: **A General Decision Theory for Huber's -Contamination Model**. *Electronic Journal of Statistics (To appear)* 2016

13. Wang T, Chen M, Zhao H, Zhu L: **Model-free dimension reduction and variable selection for general regression in high dimensions.** *Statistics and computing (Accepted)* 2016

14. Allott EH, Geradts J, Sun X, Cohen SM, Zirpoli GR, Khoury T, Bshara W, Chen M, Sherman ME, Palmer JR, Ambrosone CB, Olshan AF, Troester MA: **Intratumoral heterogeneity as a source of discordance in breast cancer biomarker classification**. *Breast Cancer Research* 2016, **18**:68

15. Chen M, Lin H, Zhao H: **Change point analysis of histone modifications reveals epigenetic blocks with distinct regulatory activity and biological functions**. *Annals of Applied Statistics* 2016, **10**:506–526

16. Wang X, Chen M, Yu X, Pornputtapong N, Chen H, Zhang NR, Powers S, Krauthammer M: **Global copy number profiling of cancer genomes**. *Bioinformatics* 2016, **32(6)**:926–928

17. Wang T, Chen M, Zhao H: **Estimating DNA methylation levels by joint modeling of multiple methylation profiles from microarray data**. *Biometrics* 2016, **72**(2):354–363

18. Chen M, Gao C, Zhao H: **Posterior Contraction Rates of the Phylogenetic Indian Buffet Processes.** *Bayesian Analysis* 2016, **11**(2):477–497

19. Chen M, Ren Z, Zhao H, Zhou H: **Asymptotic normal and efficient estimation of covariate-adjusted gaussian graphical model**. *Journal of the American Statistical Association (Theory and Methods)* 2016, **111**(513):394–406

20. Lin H, Chen M, Kundaje A, Valouev A, Yin H, Liu N, Neuenkirchen N, Zhong M, Snyder M: **Reassessment of Piwi Binding to the Genome and Piwi Impact on RNA Polymerase II Distribution**. *Developmental Cell* 2015, **32**(6):772–774

21. Xu Z, Zhang G, Jin F, Chen M, Furey TS, Sullivan PF, Qin Z, Hu M, Li Y: **A hidden Markov random field based Bayesian method for the detection of long-range chromosomal interactions in Hi-C Data**. *Bioinformatics* 2015, **32**(5):650–656

22. Huang KC, Sun W, Wu Y, Chen M, Mohlke KL, Lange LA, Li Y: **Association Studies with Imputed Variants Using Expectation-Maximization Likelihood-Ratio Tests**. *PloS one* 2014, **9**(11):e110679

23. Zhu Y, Watson J, Chen M, Shen DR, Yarde M, Agler M, Burford N, Alt A, Jayachandra S, Cvijic ME, et al.: **Integrating High-Content Analysis into a Multiplexed Screening Approach to Identify and Characterize GPCR Agonists**. *Journal of biomolecular screening* 2014, :1087057114533146

24. Brownstein CA, Beggs AH, Homer N, Merriman B, Yu TW, Flannery KC, Dechene ET, Towne MC, Savage SK, Price EN, et al.: **Dworzy Ski P, Fairbrother W: An international effort towards developing standards for best practices in analysis, interpretation and reporting of clinical genome sequencing results in the CLARITY Challenge**. *Genome Biol* 2014, **15**:R53

25. Chen M, Gunel M, Zhao H: **SomatiCA: identifying, characterizing and quantifying somatic copy number aberrations from cancer genome sequencing data**. *PloS one* 2013, **8**(11):e78143

26. Yang C, Li C, Chen M, Chen X, Hou L, Zhao H: **A penalized linear mixed model for genomic prediction using pedigree structures**. In *BMC Proceedings*, *Volume 8*, BioMed Central Ltd 2014:S67

27. Li C, Yang C, Chen M, Chen X, Hou L, Zhao H: **Adjustment of familial relatedness in association test for rare variants**. In *BMC Proceedings*, *Volume 8*, BioMed Central Ltd 2014:S39

28. Chen M, Yang C, Li C, Hou L, Chen X, Zhao H: **Admixture mapping analysis in the context of GWAS with GAW18 data**. In *BMC Proceedings*, *Volume 8*, BioMed Central Ltd 2014:S3

29. Chen M, Svicher V, Artese A, Costa G, Alteri C, Ortuso F, Parrotta L, Liu Y, Liu C, Perno CF, et al.: **Detecting and understanding genetic and structural features in HIV-1 B subtype V3 underlying HIV-1 co-receptor usage**. *Bioinformatics* 2013, **29**(4):451–460

30. Saxe JP, Chen M, Zhao H, Lin H: **Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline**. *The EMBO journal* 2013, **32**:1869–1885

31. Svicher V, Cento V, Bernassola M, Neumann-Fraune M, Hemert FV, Chen M, Salpini R, Liu C, Longo R, Visca M, et al.: **Novel HBsAg markers tightly correlate with occult HBV infection and strongly affect HBsAg detection**. *Antiviral research* 2012, **93**:86–93

32. Svicher V, Chen M, Alteri C, Costa G, Dimonte S, Chang L, Parrotta L, Dimaio C, Carta S, Surdo M, et al.: **Key-genetic elements in HIV-1 gp120 V1, V2, and C4 domains tightly and differentially modulate gp120 interaction with the CCR5 and CXCR4 N-terminus and HIV-1 antigenic potential**. *Antiviral Therapy* 2011, **16**(Suppl 1):A14

SELECTED SUBMITTED MANUSCRIPTS

33. Chen M, Gao C, Ren Z, Zhou H: **Sparse CCA via adjusted iterative thresholding**. *Submitted* 2013.arXiv:1311.6186

34. Zhang Z, Eckret MA, Zhu A, Chryplewicz A, De Jesus DF, Ren D, Kulkarni R, Lengyel E, He C, Chen M: **RADAR: Differential analysis of MeRIP-seq with a random effect model**. *Submitted* 2018

BOOK CHAPTERS

1. Yang C, Li C, Chung D, Chen M, Zhao H: **Introduction to statistical methods in genome-wide association studies**. In *Genome-Wide Association Studies: From Polymorphism to Personalized Medicine*. Edited by Appasani K, Cambridge University Press 2015

2. Chen M, Yang C, Li C, Zhao H: **eQTL mapping**. In *Genome-Wide Association Studies: From Polymorphism to Personalized Medicine*. Edited by Appasani K, Cambridge University Press 2015

3. Chen M, Hou L, Zhao H: **Statistical Methods for the Analysis of Next Generation Sequencing Data from Paired Tumor-Normal Samples**. In *Statistical Analysis of Next Generation Sequencing Data*, Springer 2014:379–404

INVITED PRESENTATIONS

- **Variability-Preserving Imputation for Accurate Gene Expression Recovery (VIPER) in Single Cell RNA Sequencing Studies.** *JSM, Vancouver, July 2018*
- **Variability-Preserving Imputation for Accurate Gene Expression Recovery (VIPER) in Single Cell RNA Sequencing Studies.** *IMS-APRM, Singapore, June 2018*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *Department of Mathematics, The Hong Kong University of Science and Technology, Hong Kong, Dec 2017*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *Genome Institute of Singapore, Singapore, Dec 2017*
- **SynthEx: A Synthetic- normal based DNA sequencing tool for copy number alteration detection and tumor heterogeneity profiling.** *ICSA, Chicago, June 2017*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *14th Graybill Conference on Statistical Genomics and Genetics, Fort Collins, June 2017*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *Department of Statistics, University of Virginia, March 2017*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *Department of Medicine Grand Round, University of Chicago, March 2017*
- **Statistical methods for single cell RNA sequencing data.** *ICSA-China, Shanghai, China, Dec 2016*
- **Construct clonal history from whole genome sequencing data by incorporating phase information.** *JSM, Chicago, USA, August 2016*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *Department of Automation, Tsinghua University, July 2016*

- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *The Third Taihu International Statistics Forum, Shanghai, China, July 2016*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *ICSA Conference on Data Science, Dali, China, July 2016*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *Department of Human Genetics, University of Chicago, May 2016*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *Department of Biomedical Data Science, Stanford University, March 2016*
- **Removing unwanted variation using both control and target genes in single cell RNA sequencing studies.** *Genetics Research Colloquium, UNC-Chapel Hill, Feb 2016*
- **De-noising and clustering single cell RNA sequencing data.** *Bioinformatics and Computational Biology Colloquium, UNC-Chapel Hill, Nov 2015*
- **A synthetic normal strategy for copy number calling using cancer exome sequencing data.** *ICSA, Fort Collins, May 2015*
- **Two high dimensional statistical methods for data integration challenges in biology.** *Department of Applied Math and Statistics, Johns Hopkins University, April 2014*
- **Statistical analysis of deep sequencing data from tumor samples.** *ENAR, Baltimore, USA, March 2014*
- **Profiling and accounting for heterogeneity in the analysis of cancer sequencing data.** *Department of Biostatistics, UNC-Chapel Hill, March 2014*
- **Profiling and accounting for heterogeneity in the analysis of cancer sequencing data.** *Center for imaging sciences, School of Engineering, Johns Hopkins University, Feb 2014*
- **Profiling and accounting for heterogeneity in the analysis of cancer sequencing data.** *Department of Biostatistics and Medical Informatics, UW-Madison, Feb 2014*
- **Phylogenetic indian buffet process: theory and applications in integrative analysis of cancer genomics.** *Statistical Data Integration Challenges in Computational Biology: Regulatory Networks and Personalized Medicine, Banff International Research Station, Canada, August 2013* (**poster**)
- **SomatiCA: identifying, characterizing and quantifying somatic copy number aberrations from cancer genome sequencing.** *Yale Stem Cell Center, New Haven, USA, June 13th, 2013*
- **Admixture rate estimation and post-call correction for complete genomics matched tumor-normal whole genome sequencing data.** *The American Society of Human Genetics 62nd Annual Meeting, San Francisco, USA, November 2012* (**poster**)
- **Change point analysis of histone modifications reveals epigenetic blocks with distinct regulatory activity and biological functions.** *Yale Stem Cell Center, New Haven, USA, September 27th, 2012*

SOFTWARE

SynthEx: A Synthetic- normal based DNA sequencing tool for copy number alteration detection and tumor heterogeneity profiling.

SomatiCA: identifying, characterizing, and quantifying somatic copy number aberrations from cancer genome sequencing.

ANTAC: Asymptotic normal estimation of covariate-adjusted gaussian graphical model

CAPIT: Canonical correlation analysis via precision adjusted iterative thresholding

MBAmethyl: Model-based analysis of DNA methylation data

Citrus: Toolkit for single cell sequencing data analysis

SERVICE ON UNIVERSITY/DEPARTMENTAL COMMITTEES AT U CHICAGO

Chicago Fellows Review Committee 2018

U Chicago / Argonne National Laboratory Collaborative Research Proposal Review Committee 2017

SERVICE ON DEPARTMENTAL COMMITTEES AT UNC

Chair for Seminar Committee 2014-2016

Faculty Search Committee 2015-2016

SCIENTIFIC REVIEW GROUPS, EXPERT PANEL

Organizing and planning committee, 2017 Graybill Conference on Statistical Genomics 2017

Study Section Member, DMS/NIGMS interface of mathematics and biology 2015, 2017

OTHER PROFESSIONAL ACTIVITIES

Reviewer for peer-reviewed journals, including *JASA*, *Annals of Applied Statistics*, *Biometrics*, *PLoS One*, *Statistics in Biosciences*, *Molecular and Cellular Proteomics*, *BMC Bioinformatics*, *Bioinformatics*, *Statistical Applications in Genetics and Molecular Biology*, *Genome Biology*, *Genetics in Medicine*, *Journal of Genetics and Genomics* and *Nucleic Acid Research*.

TEACHING EXPERIENCE

Instructor Department of Human Genetics, U Chicago

HGEN 48800: Fundamentals of Computational Biology: Algorithms and Applications Spring 17, 18

Instructor Department of Biostatistics, UNC-Chapel Hill

BIOS 735: Statistical Computing Fall 2015

Teaching Assistant Department of Statistics, Yale University

STAT 645: Statistical Methods in Genetics and Bioinformatics Spring 2011

STAT 100: Introductory Statistics Spring 2013

MEMBERSHIP

Member, American Association for the Advancement of Science

Member, American Statistical Association

Member, American Society of Human Genetics

Member, ENAR