

Date of Preparation: Jan 21, 2021

## **Mengjie Chen, Ph.D.**

The University of Chicago  
Department of Medicine  
Section of Genetic Medicine  
N417A  
5841 S. Maryland Ave  
Chicago, IL 60637  
Office: (773)-834-3175  
Email: mengjiechen@uchicago.edu  
Web page: <http://www.mengjiechen.com/>

### **ACADEMIC APPOINTMENTS**

2014-2016 Assistant Professor, Department of Biostatistics, University of North Carolina-Chapel Hill  
2014-2016 Assistant Professor, Department of Genetics, University of North Carolina-Chapel Hill  
2016- Assistant Professor, Section of Genetic Medicine, Department of Medicine, the University of Chicago  
2017- Assistant Professor, Department of Human Genetics, the University of Chicago

### *Ph.D.-Granting Committee, Program, Institute, and Center Appointments*

2017- Committee on Genetics, Genomics and Systems Biology  
2017- Trainor, Biophysics Training Grant  
2018- Trainor, NIH T32 Genetics Training Grant  
2019- Member, the University of Chicago Medicine Comprehensive Cancer Center (UCCCC)  
2019- Committee on Cancer Biology  
2019-2020 Interim Director, Bioinformatics Core, Center for Research Informatics, Biological Science Division  
2020- Director, Bioinformatics Core, Biological Science Division

### **ACADEMIC TRAINING**

2004-2009 B.S., Biotechnology, Huazhong University of Science and Technology  
2009-2014 Ph.D., Computational Biology and Bioinformatics, Yale University

### **PROFESSIONAL SOCIETIES**

American Association for the Advancement of Science  
American Statistics Society  
International Chinese Statistical Association  
Eastern North American Region, International Biometric Society

### **HONORS, PRIZES, AND AWARDS**

2009-2012 China Scholarship Council-Yale World Scholarship  
2014 Student Marshal, Yale Graduate School of Arts and Sciences  
2015 UNC Junior Faculty Development Award  
2017 Student Paper Award by the ASA Section of Statistics Genomics and Genetics  
2019 Alfred P. Sloan Research fellowship in Computational and Molecular Evolutionary Biology

## SCHOLARSHIP

### (a) Peer-reviewed publications in the primary literature, exclusive of abstracts:

1. Svicher V, Cento V, Bernassola M, Neumann-Fraune M, Van Hemert F, **Chen M**, Salpini R, Liu C, Longo R, Visca M, Romano S. Novel HBsAg markers tightly correlate with occult HBV infection and strongly affect HBsAg detection. *Antiviral research*. 2012 Jan 1;93(1):86-93.
2. 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* 32, no. 13 (2013): 1869-1885.
3. **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* 29, no. 4 (2013): 451-460.
4. **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
5. Brownstein CA, Beggs AH, Homer N, Merriman B, Yu TW, Flannery KC, Dechene ET, Towne MC, Savage SK, Price EN, ..., **Chen M**, ..., 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 Biology* 2014, 15:R53.
6. Yang C, Li C, **Chen M**, Chen X, Hou L, Zhao H: A penalized linear mixed model for genomic prediction using pedigree structures. *BMC Proceedings*, Volume 8, BioMed Central Ltd 2014:S67
7. Li C, Yang C, **Chen M**, Chen X, Hou L, Zhao H: Adjustment of familial relatedness in association test for rare variants. *BMC Proceedings*, Volume 8, BioMed Central Ltd 2014:S39
8. **Chen M**, Yang C, Li C, Hou L, Chen X, Zhao H: Admixture mapping analysis in the context of GWAS with GAW18 data. *BMC proceedings* 2014 Jun (Vol. 8, No. 1, p. S3). BioMed Central.
9. 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
10. 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
11. 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 Mar 23;32(6):772-4.
12. 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 Nov 4;32(5):650-6.
13. Wang T, **Chen M**, Zhao H: Estimating DNA methylation levels by joint modeling of multiple methylation proles from microarray data. *Biometrics* 2016, 72(2):354-363
14. **Chen M**, Gao C, Zhao H: Posterior Contraction Rates of the Phylogenetic Indian Buffet Processes. *Bayesian Analysis* 2016, 11(2):477-497
15. **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

16. **Chen M**, Gao C, Ren Z: A General Decision Theory for Huber's  $\epsilon$ -Contamination Model. *Electronic Journal of Statistics* 2016
17. Wang T, **Chen M**, Zhao H, Zhu L: Model-free dimension reduction and variable selection for general regression in high dimensions. *Statistics and computing* 2016
18. 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
19. **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
20. Wang X, **Chen M**, Yu X, Pomputtpong N, Chen H, Zhang NR, Powers S, Krauthammer M: Global copy number profiling of cancer genomes. *Bioinformatics* 2016, 32(6):926-928
21. **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
22. **Chen M**, Gao C, Ren Z: Robust Covariance Matrix Estimation via Matrix Depth. *The Annals of Statistics* 2017
23. \*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.
24. \*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
25. 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 brolamellar carcinoma. *Scientific reports* 2017, 7:44653
26. 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* 2018 Jan 1;27(1):67-74.
27. 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 Feb 26;46(6):3009-18.
28. **Chen M**, Gao C, Ren Z, Zhou H: Sparse CCA via adjusted iterative thresholding. *International Congress of Chinese Mathematicians Proceedings* 2018.
29. 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)
30. **Chen M**, Zhou X: VIPER: variability-preserving imputation for accurate gene expression recovery in single-cell RNA sequencing studies. *Genome biology* 2018, 19:196
31. 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* 2018

32. Sun S, Zhu J, Mozaari S, Ober C, **Chen M**, Zhou X: Heritability Estimation and Differential Analysis with Generalized Linear Mixed Models in Genomic Sequencing Studies. *Bioinformatics* 2018, 359265
33. Puvanesarajah S, Nyante SJ, Kuzmiak CM, **Chen M**, Tse CK, Sun X, Allott EH, Kirk EL, Carey LA, Perou CM, Olshan AF. PAM50 and risk of recurrence scores for interval breast cancers. *Cancer Prevention Research*. 2018 Jun 1;11(6):327-36.
34. Allott EH, Shan Y, **Chen M**, Sun X, Garcia-Recio S, Kirk EL, Olshan AF, Geradts J, Earp HS, Carey LA, et al.: Bimodal age distribution at diagnosis in breast cancer persists across molecular and genomic classifications. *Breast cancer research and treatment* 2019, 1–11
35. Chen X, Cao G, Wu J, Wang X, Gao J, Xu L, Tian Q, **Chen M**, Sun B, Zhu B, Huang J, and Pan Z: The histone methyltransferase EZH2 primes the early differentiation of follicular helper T cells during acute viral infection. *Cellular & Molecular Immunology* 2019, 1
36. De Jesus DF, Zhang Z, Kahraman S, Brown NK, **Chen M**, Hu J, Gupta MK, He C, Kulkarni RN. m 6 A mRNA methylation regulates human  $\beta$ -cell biology in physiological states and in type 2 diabetes. *Nature Metabolism*. 2019 Aug;1(8):765-74.
37. Schnepf P, **Chen M**, Keller E and Zhou X: SNP Identification from Single-Cell RNA Sequencing Data. *Human Molecular Genetics*. 2019 Nov 1;28(21):3569-3583
38. 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. *Genome Biology*. 2019 Dec 23;20(1):294
39. Sun S, Ying M, Shang X, Keller E, **Chen M** and Zhou X: Integrative Differential Expression and Gene Set Enrichment Analysis for Single Cell RNA-seq Studies. *Nature Communications*. 2020 Mar 27;11(1):1585.
40. Liu S, He C and **Chen M**: REPIC: A database for exploring N6-methyladenosine methylome. *Genome Biology*. 2020 Apr 28;21(1):100
41. Zhang S, Zhang H, Zhou Y, Qiao M, Zhao S, Kozlova A, Shi J, Sanders A, Wang G, Luo K, Sengupta S, West S, Qian S, Strait M, Avramopoulos D, Cowan C, **Chen M**, Pang Z, Gejman P, He X, and Duan J. Allele-specific open chromatin in human iPSC neurons elucidates functional noncoding disease variants. *Science*. 2020, Jul 31, 369(6503): 561-565
42. Kim T, Zhou X and **Chen M**: Demystifying "drop-outs" in single cell UMI data. *Genome Biology*. 2020 Aug 06, 21(196)
43. **Chen M**, Zhan Q, Mu Z, Wang L, Zhu P and Li Y: Alignment of single-cell RNA-seq samples without overcorrection using kernel density matching. *Genome Research*.
44. Jin C, **Chen M**, Lin D and Sun W: Cell Type Aware analysis of RNA-seq data (CARseq) reveals difference and similarities of the molecular mechanisms of Schizophrenia and Autism. *Nature Computational Science*.

(b) Book chapters:

45. 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
46. **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

47. **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

(c) Works in review, in preparation, etc. not yet publically available [list ONLY if available for BSD review]

48. Jin C, Sun W and **Chen M**: Inferring Intra-Tumor Heterogeneity by Jointly Modeling Copy Number Alterations and Somatic Point Mutations. Manuscript.
49. Sobreira D, Joslin AC, Zhan Q, ..., **Chen M**, Aneas I, Nóbrega MA: Extensive pleiotropism and allelic heterogeneity mediate metabolic effects of IRX3 and IRX5. Revised at *Science*.
50. Hu L, Liu S, Peng Y, ..., **Chen M\***, Chen J\*, Chuan He\*: Quantitative RNA m6A sequencing reveals dynamic human epitranscriptome. Submitted.
51. He PC, Harada BT, Wang S, Zhang Z, Lyu R, **Chen M**, Chuan He: Splicing and the exon junction complex control m6A epitranscriptome specificity. Submitted to *Cell*.
52. Vistain L\*, Phan V, Jordi C, **Chen M**, Reddy ST, and Tay S: Quantification of proteins, protein-complexes and mRNA in single-cells by proximity-sequencing. Submitted.
53. Liu S, Zhang S, He C, **Chen M**: Recent development on computational methods for epitranscriptomics. Invited Chapter In Handbook of Statistical Bioinformatics II, Springer, 2022
54. Hu J, **Chen M\***, Zhou X\*: Statistical analysis for single cell transcriptomics data. Invited Chapter In Handbook of Statistical Bioinformatics II, Springer, 2022

## FUNDING

### (a) Current:

1. NIH/NIGMS R01 GM126553. PI: Chen M. My role: PI. Title: "Advanced statistical methods for single cell RNA sequencing studies". Annual salary recovery or effort: 20%. Project period: 08/01/17 - 05/31/22.
2. Alfred P. Sloan Foundation Research Fellowship. Project period: 09/01/19 - 08/31/21
3. Chan Zuckerberg Initiative Human Seed Network Grant. PI: Basu A, Chen M, Lengyel E and Ecket. My role: co-PI. Title: "A Female Reproductive Cell Atlas". Annual salary recovery or effort: 10%. Project period: 09/01/19 - 08/31/22. In this project, my group will analyze all the generated single cell sequencing data and integrate multiple-omics to characterize the female reproductive system.
4. NIH/NIMDH R01 MD013452. PI: Olopade F. My role: Co-investigator. Title: "Etiology and genomics of breast cancer progression in women of African ancestry". Annual salary recovery or effort: 10%. Project period: 01/01/20 - 12/30/24. In this project, I will oversee the analysis of RNA-seq and exome-seq data and help interpret the associations of genomic features with clinical outcomes.
5. NSF DMS 2016307. PI: Chen M. My role: PI. Title: "Statistical Methods for Intra-tumor Heterogeneity Studies Using Sequencing Data". Annual salary recovery or effort: 9%. Project period: 09/01/20 - 08/31/23.
6. NIH R35CA264619. PI: Lengyel E. My role: Co-investigator. Title: "Metabolic reprogramming of the tumor microenvironment and therapy resistance". Annual salary recovery or effort: 10%. Project period: 07/01/21-06/30/28. In this project, I will oversee the analysis of functional interrogation novel pathways with a focus on understanding how metabolism influence ovarian cancer progression and the response of ovarian cancer to chemotherapy.

7. NIH/NHGRI RM1. PI: He C. My role: co-Investigator. Title: "Center for Dynamic RNA Epitranscriptomics". Annual salary recovery or effort: 20%. Project period: 07/01/21 - 06/31/26.

(b) Completed:

1. UChicago Cancer Epigenome and Epitranscriptome Pilot Grants. PI: Chen M and He C. My role: co-PI. Title: "Profiling of m6A modifications in prostate cancers". Annual salary recovery or effort: 0% (not allowed). Project period: 01/01/19 - 12/31/19.
2. Breast Cancer SPORE Developmental Research Program Projects. PI: Chen M. My role: PI. Title: "Single cell transcriptomics profiling of breast cancer women in NBCS". Annual salary recovery or effort: 0% (not allowed). Project period: 01/15/20 - 12/31/20.

**INVITED SPEAKING**

2012	Research Seminar	Yale Stem Cell Center, New Haven, USA,
2014	Research Seminar	Department of Biostatistics and Medical Informatics, UW-Madison, Madison, USA
2014	Research Seminar	Center for imaging sciences, School of Engineering, Johns Hopkins University, Baltimore, USA
2014	Research Seminar	Department of Biostatistics, UNC-Chapel Hill, Chapel Hill, USA
2014	Research Seminar	Department of Applied Math and Statistics, Johns Hopkins University, Baltimore, USA
2014	Invited speaker	Eastern North American Region, International Biometric Society Meeting, Baltimore, USA
2015	Invited speaker	International Chinese Statistical Association Meeting, Fort Collins, USA
2015	Research Seminar	Bioinformatics and Computational Biology Colloquium, UNC-Chapel Hill, Chapel Hill, USA
2015	Research Seminar	Genetics Research Colloquium, UNC-Chapel Hill, Chapel Hill, USA
2016	Research Seminar	Department of Biomedical Data Science, Stanford University, Stanford, USA
2016	Research Seminar	Department of Human Genetics, the University of Chicago, Chicago, USA
2016	Invited speaker	International Chinese Statistical Association Conference on Data Science, Dali, China
2016	Invited speaker	The Third Taihu International Statistics Forum, Shanghai, China
2016	Research Seminar	Department of Automation, Tsinghua University, China
2016	Invited speaker	International Chinese Statistical Association-China Meeting, Shanghai, China
2017	Grand Round Speaker	Department of Medicine, the University of Chicago, USA
2017	Research Seminar	Department of Statistics, University of Virginia, Charlottesville, USA
2017	Invited speaker	14th Graybill Conference on Statistical Genomics and Genetics, Fort Collins, USA
2017	Invited speaker	International Chinese Statistical Association Meeting, Chicago, USA

2017	Research Seminar	Duke-Nanyang University of Singapore Medical School, Singapore
2017	Research Seminar	Genome Institute of Singapore, Singapore
2017	Research Seminar	Department of Mathematics, The Hong Kong University of Science and Technology, Hong Kong, China
2018	Invited speaker	Eastern North American Region, International Biometric Society Meeting, Atlanta, USA
2018	Invited speaker	Institute of Mathematical Statistics Asia Pacific Rim Meeting, Singapore
2018	Invited speaker	Joint Statistical Meetings, Vancouver, Canada
2019	Invited speaker	BIRS workshop - Frontiers in Single-cell Technology Applications and Data Analysis, Banff, Canada
2019	Invited speaker	International Chinese Statistical Association Meeting, Durham, USA
2019	Invited speaker	The 3rd International Conference on Econometrics and Statistics, Taiwan
2019	Invited speaker	International Society for Bayesian Analysis Meeting, Kobe, Japan
2019	Invited speaker	International Conference of Cancer and Aging, Wuhan, China
2019	Research Seminar	School of Biological Science, Huazhong University of Science and Technology, Wuhan, China
2019	Research Seminar	Tongji School of Medicine, Huazhong University of Science and Technology, Wuhan, China
2019	Research Seminar	Shanghai Jiaotong University - Yale Joint Center of Biostatistics and Data Science Center, Shanghai, China
2020	Research Seminar	Department of Genetics, Emory University, Atlanta, USA
2020	Invited speaker	NorthShore University HealthSystem, Chicago, USA
2020	Invited speaker	Eastern North American Region, International Biometric Society Meeting, Nashville, USA (Virtual)
2020	Invited speaker	Dahshu online forum for Data Science
2021	Invited speaker	Western North American Region, International Biometric Society Meeting, Anchorage, USA
2021	Invited speaker	Eastern North American Region, International Biometric Society Meeting

#### **INVITED, ELECTED, OR APPOINTED EXTRAMURAL SERVICE**

2015	Study Section Member (ad hoc)	DMS/NIGMS interface of mathematics and biology, NSF
2016	Committee Member	Best student paper award, American Statistics Society
2016	Organizing and planning committee	Graybill Conference on Statistical Genomics
2017	Study Section Member (ad hoc)	DMS/NIGMS interface of mathematics and biology, NSF
2019	Committee Member	Best student paper award, American Statistics Society
2020	Co-editor	A special issue entitled 'Methods for Single-cell and Microbiome Sequencing Data' in Frontiers in Genetics
2020	Committee Member	Best student paper award, American Statistics Society
2020	Study Section Member (ad hoc)	The Cancer Informatics Technology [ZCA1 TCRB-T (M1)] , NIH/NCI

2020 Study Section Member (ad hoc) BDMA, NIH  
 2020 Study Section Member (ad hoc) DMS/NIGMS interface of mathematics and biology, NSF  
 2020 Committee Member C3ai.DTI COVID-19 grant application  
 2021 Review Committee Member Chan Zuckerberg Biohub Investigator Awards  
 Various Reviewer for peer-reviewed journals, including *Nature Methods*, *PNAS*, *Nature Communications*, *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*, *Genome Research*, *Genetics in Medicine*, *Journal of Genetics and Genomics* and *Nucleic Acid Research*. Guest Editor for *PLoS Genetics*.

## SERVICE

### University of Chicago

#### Committee membership:

2018 Chicago Fellows Review Committee  
 2019 Chicago Fellows Review Committee  
 2019 The Department of Medicine's 2020 Research Day Organizing Committee  
 2020 Chicago Fellows Review Committee  
 2020 The Department of Medicine's 2020 Research Day Organizing Committee  
 2020 GGSB/HG Admission Committee  
 2021 Chicago Fellows Review Committee

#### Other:

2017 U Chicago / Argonne National Laboratory Collaborative Research Proposal Review Committee  
 Various Interviewer of graduate school applicants, Human Genetics, GGSB, Biophysics  
 Various Interviewer for Physician Scientist Development Program

## TEACHING

#### For The College (B.A., B.S.):

2019- BIOS 26121, I am a part of a bioinformatics sequence offered by the Biological Sciences Collegiate Division started from Winter Quarter 2019. The sequence will contain three new computational biology courses. I am mainly teaching one, "Introduction to transcriptomics", and participate in the preparation of all three. These courses are intended for biology majors and can be counted towards a BA or BS in Biology. The key objective is to get students informed of the modern genomics technologies and their applications in biological and biomedical studies. My course covers new statistics topics emerging from transcriptomics studies in last two decades, including batch effect correction, false discovery rate, bi-clustering, sparse PCA, empirical Bayes methods, etc. I gave 10 lectures, 4 discussion sections and laboratories. I prepared 6 homework assignments, 4 quizzes and 1 final project. The class had 4 students in 2019, 6 students in 2020 and 5 students in 2021.

#### For Graduate programs (Ph.D.):

2017- HGEN 48800. Starting from Spring Quarter 2017, my colleague Dr. Xin He from Genetics and I designed and co-taught a new course, Fundamentals of Computational Biology: Algorithms and Applications. This course covered principles of discrete and numerical algorithms that have broad applications in computational biology. The specific topics include basic data structure, dynamic



programming, numerical optimization, singular value decomposition, linear regression with shrinkage, and classification. We also discussed their applications, including variant calling, transcriptome inference, and methods for reconstructing gene regulatory networks. This course has been offered four times in 2017, 2018, 2019 and 2020 to students with diverse backgrounds and received positive feedbacks. Each time, I gave 10 lectures and prepared 5 homework assignments and 1 final project. The class had 5 students in 2017, 6 students in 2018, 10 students in 2019 and 19 students in 2020.

2020- Quantitative Analysis Bootcamp. Bootcamp is one-week workshop that makes new matriculants at U Chicago benefit from establishing a core level of familiarity with computational workflows and the programming language R; and by participating in a range of professional development activities and discussions designed to build camaraderie, to be prepared for the first year of graduate school regardless of the instructional format, and to be introduced more broadly to the University of Chicago community. I gave four half-day workshops to four groups of students about RNA-seq data analysis in 2020.

Research trainees:

(a) Undergraduate (B.A., B.S.)

2017-2018 Sam Baureis, Statistics.

2020 summer Shuwen Fang

2020 summer Allison Rodas

(b) Graduate (M.S., Ph.D.)

2017- 2019 Kaiqian Zhang. Master student in Financial Mathematics. Now a PhD student in Computational Biology at Princeton University.

2017- 2020 Tae Kim. PhD student in Statistics. Winner of David Wallace Award for Applied Statistics in Department of Statistics. Now a data scientist at Microsoft.

2017- Yifan Zhou. PhD student in Biophysics. Joint with Dr. Xin He in Department of Human Genetics.

2017- 2020 Zijie Zhang. PhD student in GGSB. Now a Professor at Yunnan University in China.

2019- 2021 Qi Zhan. Summer Intern and then a predoctoral Researcher. Now a PhD student at Mercedes Pascual lab.

2019- Chih-Hsuan Wu. PhD student in Statistics.

2019- Yuguan Wang. Master student and then PhD student in Statistics.

2019- Jovian Yu. Hematology/oncology fellow physician. PhD student in GGSB. Won an ASH Abstract Achievement Award in 2020.

2021- Michiko Ryu. Master student in Physical Science.

(c) Postdoctoral

2018- Shun Liu. Joint with Dr. Chuan He in Department of Chemistry.

2019- Xiaoyang Dou. Joint with Dr. Chuan He in Department of Chemistry.

2020- Yong Peng. Joint with Dr. Chuan He in Department of Chemistry.

(d) Other

2018-2019 Qi Zhan. Summer intern from St. Johns College (undergraduate majored in Math and Philosophy) and then a pre-doc student.

2018- Cody He. MSTP fellow in Chuan He lab.

2019- Sheila Rajagopal. Hematology/oncology fellow physician. Postdoc in Olopade lab.  
2020 Arpit Panda. Rotation student. MSTP MD-PhD program.  
2021- Ira Kraft. Med/Peds resident at UCMC.