

## Li Chen

---

**CONTACT INFORMATION** Department of Biostatistics  
University of Florida  
2004 Mowry Road, 5th Floor CTRB, 5211 RM  
Gainesville, FL 32611-7450  
*Tel:* (352) 294-5909  
Email: li.chen1@ufl.edu  
Homepage: <https://lichen-lab.github.io/>

**RESEARCH INTERESTS** My current research focuses on developing statistical and machine learning methods for analyzing large-scale omics data with special focus on transcriptomics, epigenomics, genetics, metagenomics and single-cell genomics.

**CURRENT POSITION** Associate Professor  
Department of Biostatistics  
University of Florida

**EDUCATION** **Emory University**, Atlanta, GA  
Ph.D., Computer Science and Informatics Sep 2011-Aug 2017  
**Johns Hopkins University**, Baltimore, MD  
M.S.E., Computer Science Sep 2010-August 2011  
M.H.S., Biostatistics Sep 2009-May 2011

**HONORS AND AWARDS** Maximizing Investigators' Research Award (MIRA) NIGMS R35, 2021  
Showalter Research Trust Fund Award, 2020  
American Association of Colleges of Pharmacy New Investigator Award, 2019

**PUBLICATIONS** **Peer-reviewed journal articles** \*: corresponding author; §: equal contribution  
**Google scholar citation statistics (as of Nov 2022): All citations: 1896**

### Peer-Reviewed Journal - methodology

1. Wang Y, **Chen L\*** (2022) *DeepPerVar: a multimodal deep learning framework for functional interpretation of genetic variants in personal genome*. **Bioinformatics**, btac696
2. **Chen L\***, Saykin AJ, Yao B, Zhao F (2022) *Multi-task deep autoencoder improves the prediction for Alzheimer's disease progression using temporal DNA methylation data in peripheral blood*. **Computational and Structural Biotechnology Journal**
3. **Chen L\***, Wang Y, Zhao F (2022) *Exploiting deep transfer learning for the prediction of functional noncoding variants using genomic sequence*. **Bioinformatics**, btac214
4. Wang Y, Jiang Y, Yao B, Huang K, Liu Y, Wang Y, Qin X, Saykin AJ, **Chen L\*** (2021) *WEVar: a novel statistical learning framework for predicting noncoding regulatory variants*. **Briefings in Bioinformatics**, bbab189

5. Wang Y, Bhattacharya T, Jiang Y, Qin X, Wang Y, Liu Y, Saykin AJ, **Chen L\*** (2021) *A novel deep learning method for predictive modeling of microbiome data.* **Briefings in Bioinformatics**, bbaa073
6. **Chen L\*** (2020). *powmic: an R package for power assessment in microbiome case-control studies.* **Bioinformatics**, btaa197
7. Yang G, Ma A, Qin ZS\*, **Chen L\*** (2020). *Application of topic models to a compendium of ChIP-Seq datasets uncovers recurrent transcriptional regulatory modules.* **Bioinformatics**, btz975
8. **Chen L\***, Wang F, Bruggeman E, Li C, Yao B\* (2020). *circMeta: a unified computational framework for genomic feature annotation and differential expression analysis of circular RNAs.* **Bioinformatics**, btz606
9. Sun Z§, **Chen L§**, Xin H, Huang Q, Cillo A, Tabib T, Ding Y, Kolls J, Bruno T, Lafyatis R, Vignali D, Chen K, Hu M, Chen W (2019). *A Bayesian mixture model for clustering droplet-based single cell transcriptomic data from population studies.* **Nature Communications**, 10:1649
10. Urrutia E, **Chen L**, Zhou H, Jiang Y (2019). *Destin: toolkit for single-cell analysis of chromatin accessibility.* **Bioinformatics**, 35(19):3818-3820
11. Xiao J, **Chen L\***, Johnson S, Zhang X, Chen J\* (2018). *A phylogeny-regularized sparse regression model for predictive modeling of microbial community data.* **Frontiers in Microbiology**, 9:3112
12. **Chen L\***, Wang Y, Yao B, Mitra A, Wang X, Qin X (2018). *TIVAN: Tissue-specific cis-eQTL single nucleotide variant annotation and prediction.* **Bioinformatics**, 35(9):1573-1575
13. Xiao J, **Chen L\***, Johnson S, Zhang X, Chen J\* (2018). *Predictive modeling of microbiome data using a phylogeny-regularized generalized linear mixed model.* **Frontiers in Microbiology**, 9:1391
14. **Chen L**, Reeve J, Zhang L, Huang S, Wang X, Chen J (2018). *GMPR: A robust normalization method for zero-inflated count data with application to microbiome sequencing data.* **PeerJ**, 6:e4600
15. **Chen L\***, Qin ZS\* (2017). *Using DIVAN to assess disease/trait-associated single nucleotide variants in genome-wide scale.* **BMC Research Notes**, 10(1):530
16. Sun X, Pittard WS, Xu T, **Chen L**, Zwick ME, Jiang X, Wang F, Qin ZS (2017). *OmicSeq: A web-based search engine for exploring omics datasets.* **Nucleic Acids Research**, 45(W1):W445-W452
17. **Chen L**, Jin P, Qin ZS (2016). *DIVAN: Accurate identification of non-coding disease-specific risk variants based on multi-omics profiles.* **Genome Biology**, 17:252
18. **Chen L**, Qin ZS (2016). *traseR: an R package for performing trait-associated SNP enrichment analysis in genomic intervals.* **Bioinformatics**, 32(8):1214-1216
19. **Chen L**, Wang C, Qin ZS, Wu H (2015). *A novel statistical method for quantitative*

- comparison of multiple ChIP-seq datasets. **Bioinformatics**, 31(12):1889-1896*
20. **Chen L**, Liu H, Kocher JP, Li H, Chen J (2015). *glmgraph: an R package for variable selection and predictive modeling of structured genomic data. **Bioinformatics**, 31(24):3991-3993*
  21. Wu H, Xu T, Feng H, **Chen L**, Li B, Yao B, Qin Z, Jin P, Conneely KN (2015). *Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. **Nucleic Acids Research**, 43(21):e141*
  22. **Chen L**, Wu G, Ji H (2011). *hmChIP: a web server for exploring publicly available human and mouse ChIP-chip and ChIP-seq data. **Bioinformatics**, 27(10):1447-1448*

### Peer-Reviewed Journal - collaborative

1. Tang PC, Chen L, Singh S, Groves AK, Koehler KR, Liu XZ, Nelson RF (2022) *Early Wnt signaling activation promotes inner ear differentiation via cell caudalization in mouse stem cell-derived organoids. **Stem Cells**, doi: 10.1093/stmcls/sxac071*
2. Li Y, Wang F, Teng P, Ku L, **Chen L**, Feng Y, Yao B (2022) *Accurate identification of circRNA landscape and complexity reveals their pivotal roles in human oligodendroglia differentiation. **Genome Biology**, 23:48*
3. Li C, **Chen L**, Chou C, Ngorsuraches S, Qian J (2021) *Using Machine Learning Approaches to Predict Short-Term Risk of Cardiotoxicity Among Patients with Colorectal Cancer after Starting Fluoropyrimidine-based Chemotherapy. **Cardiovascular Toxicology**, 22(2):130-140*
4. **Chen L**, Li MJ (2021) *Editorial Deciphering Noncoding Regulatory Variants Computational and Functional Validation. **Frontiers in Bioengineering and Biotechnology** )*
5. Kim H, Kang Y, Li Y, **Chen L**, Lin L, Johnson ND, Zhu D, Robinson MH, McSwain L, Barwick BG, Yuan X, Liao X, Zhao J, Zhang Z, Shu Q, Chen J, Allen EG, Kenney AM, Castellino RC, Van Meir EG, Conneely KN, Vertino PM, Jin P, Li J (2021) *Ten-eleven translocation protein 1 modulates medulloblastoma progression. **Genome Biology** 22(1):125*
6. Li C, Ngorsuraches S, Chou C, **Chen L**, Qian J (2021) *Risk Factors of Fluoropyrimidine Induced Cardiotoxicity Among Cancer Patients: A Systematic Review and Meta-analysis. **Critical Reviews in Oncology/Hematology** 162:103346*
7. Kuehner JN, Chen J, Bruggeman EC, Wang F, Li Y, Xu C, McEachin ZT, Li Z, **Chen L**, Hales CM, Wen Z, Yang J, Yao B (2021) *5-hydroxymethylcytosine is dynamically regulated during forebrain organoid development and aberrantly altered in Alzheimer's disease. **Cell Reports** 35(4):109042*
8. Brothwell J, Griesenauer B, **Chen L** and Spinola SM (2021) *Interactions of the Skin Pathogen *Haemophilus ducreyi* With the Human Host. **Frontiers in Immunology**, <https://doi.org/10.3389/fimmu.2020.615402>*
9. Mishuk AU, **Chen L**, Li C, Huo N, Hansen RA, Harris I, Kiptanui Z, Qian J (2021). *Patient factors associated with oral generic olanzapine initiation and*

*substitution among Medicaid beneficiaries: a new user cohort study.* **Current Medical Research and Opinion**, DOI: 10.1080/03007995.2021.1882413

10. Mishuk AU, **Chen L**, Gaillard P, Westrick S, Hansen RA, Qian J (2020). *National Trends in Prescription Proton Pump Inhibitors (PPI) Use and Expenditure in the U.S. in 2002-2017.* **Journal of the American Pharmacists Association**, <https://doi.org/10.1016/j.japh.2020.09.015>
11. Mei W, Jiang Z, Chen Y, **Chen L**, Sancar A, Jiang Y (2020). *Genome-wide circadian rhythm detection methods: systematic evaluations and practical guidelines.* **Briefings in Bioinformatics**, bbaa135
12. Mitra AK, Kumar H, Ramakrishnan V, **Chen L**, Kumar S, Rajkumar SV, Van Ness B (2020). *In vitro and ex vivo gene expression profiling reveals differential kinetic response of HSPs and UPR genes are associated with PI-resistance in multiple myeloma.* **Blood Cancer Journal**, 10:78
13. **Chen L**, Li C, Huo N, Mishuk AU, Hansen RA, Harris I, Kiptanui Z, Qian J (2020). *Oral Generic Tacrolimus Initiation and Substitution in The Medicaid Population: A New User Cohort Study.* **Current Medical Research and Opinion**, 36:9,1533-1540
14. Li C, **Chen L**, Huo N, Mishuk AU, Hansen RA, Harris I, Kiptanui Z, Wang Z, Qian J (2020). *Generic escitalopram initiation and substitution among Medicare beneficiaries: A new user cohort study.* **Plos One**, e0232226
15. Huo N, **Chen L**, Ullah Mishuk A, Li C, Hansen RA, Harris I, Kiptanui Z, Wang Z, Dutcher SK, Qian J (2020). *Generic levothyroxine initiation and substitution among Medicare and Medicaid populations: a new user cohort study.* **Endocrine**, pp1-13
16. Qin L, Xu Q, Li Z, **Chen L**, Li Y, Yang N, Liu Z, Guo J, Shen L, Allen GE, Chen C, Ma C, Wu H, Zhu X, Jin P, Tang B (2020). *Ethnicity-specific and overlapping alterations of brain hydroxymethylome in Alzheimer's disease.* **Human Molecular Genetics**, 29(1):149-158
17. Hohmann L, Hastings T, Ha D, Garza K, Huston S, **Chen L**, Westrick S (2019). *Impact of a multi-component immunization intervention on pneumococcal and herpes zoster vaccinations: A randomized controlled trial of community pharmacies in 2 states.* **Research in Social and Administrative Pharmacy**, S1551-7411(18)30747-2
18. Cheng Y, Sun M, **Chen L**, Li Y, Lin L, Yao B, Li Z, Wang Z, Chen J, Miao Z, Xin N, Huang L, Allen E, Wu H, Xu X, Jin P (2018). *Ten-eleven translocation proteins modulate the response to environmental stress in mice.* **Cell Reports** 25(11):3194-3203
19. Yao B, Li Y, Wang Z, **Chen L**, Poidevin M, Zhang C, Lin L, Wang F, Bao H, Jiao B, Lim J, Cheng Y, Huang L, Phillips BL, Xu T, Duan R, Moberg K, Wu H, Jin P (2018). *Active N<sup>6</sup>-Methyladenine Demethylation by DMAD Regulates Gene Expression by Coordinating with Polycomb Protein in Neurons.* **Molecular Cell**, 71(5):848-857
20. Yao B, Cheng Y, Wang Z, Li Y, **Chen L**, Huang L, Zhang W, Wu H, Tang B,

- Jin P (2017). *DNA N6-methyladenine is dynamically regulated in the mouse brain following environmental stress*. **Nature Communications**, 8:1122
21. Pan F, Wingo TS, Zhao Z, Gao R, Makishima H, Qu G, Lin L, Yu M, Ortega JR, Wang J, Nazha A, **Chen L**, Yao B, Liu C, Chen S, Weeks O, Ni H, Phillips BL, Huang S, Wang J, He C, Li GM, Radivoyevitch T, Aifantis I, Maciejewski JP, Yang FC, Jin P, Xu M (2017). *Tet2 loss leads to hypermutagenicity in hematopoietic stem/progenitor cells*. **Nature Communications**, 8:15102
  22. Li X, Yao B, **Chen L**, Kang Y, Li Y, Cheng Y, Li L, Lin L, Wang Z, Wang M, Pan F, Dai Q, Zhang W, Wu H, Shu Q, Qin Z, He C, Xu M, Jin P (2017). *Ten-eleven translocation 2 interacts with Forkhead box O3 and regulates adult neurogenesis*. **Nature Communications**, 8:15903
  23. Zhang F, Hammack C, Ogden SC, Cheng Y, Lee EM, Wen Z, Qian X, Nguyen HN, Li Y, Yao B, Xu M, Xu T, **Chen L**, Wang Z, Feng H, Huang WK, Yoon KJ, Shan C, Huang L, Qin Z, Christian KM, Shi PY, Xu M, Xia M, Zheng W, Wu H, Song H, Tang H, Ming GL, Jin P (2016). *Molecular Signatures Associated with ZIKV Exposure in Human Cortical Neural Progenitors and Astrocytes*. **Nucleic Acids Research**, 44(18):8610-8620
  24. Papale LA, Li S, Madrid A, Zhang Q, **Chen L**, Chopra P, Jin P, Keles S, Alisch RS (2016). *Gender-specific hippocampal 5-hydroxymethylcytosine is disrupted in response to acute stress*. **Neurobiology of Disease**, 96:54-66
  25. Li S, Papale LA, Zhang Q, Madrid A, **Chen L**, Chopra P, Keles S, Jin P, Alisch RS (2016). *Genome-wide alterations in hippocampal 5-hydroxymethylcytosine links plasticity genes to acute stress*. **Neurobiology of Disease**, 86:99-108
  26. Sun M, Song MM, Wei B, Gao Q, Li L, Yao B, **Chen L**, Lin L, Dai Q, Zhou X, Tao J, Chen J, He C, Jin P, Xu Z (2016). *5-Hydroxymethylcytosine-mediated alteration of transposon activity associated with the exposure to adverse in utero environments in human*. **Human Molecular Genetics**, 25(11):2208-2219
  27. Miao Z, He Y, Xin N, Sun M, **Chen L**, Lin L, Li J, Kong J, Jin P, Xu X (2015). *Altering 5-hydroxymethylcytosine modification impacts ischemic brain injury*. **Human Molecular Genetics**, 24(20):5855-5866
  28. Zhao Z§, **Chen L**§, Dawlaty MM§, Pan F, Weeks O, Zhou Y, Cao Z, Shi H, Wang J, Lin L, Chen S, Yuan W, Qin Z, Ni H, Nimer SD, Yang FC, Jaenisch R, Jin P, Xu M (2015). *Combined loss of Tet1 and Tet2 promotes B-cell, but not myeloid malignancies in mice*. **Cell Reports**, 13(8):1692-1704
  29. Zhang L, Yin JC, Yeh H, Ma NX, Lee G, Chen XA, Wang Y, Lin L, **Chen L**, Jin P, Wu GY, Chen G (2015). *Small molecules efficiently reprogram human astroglial cells into functional neurons*. **Cell Stem Cell**, 17(6):735-747
  30. Yang R, **Chen L**, Newman S, Gandhi K, Doho G, Moreno CS, Vertino PM, Bernal-Mizarchi L, Lonial S, Boise LH, Rossi M, Kowalski J, Qin ZS (2014). *Integrated Analysis of Whole-Genome Paired-End and Mate-Pair Sequencing Data for Identifying Genomic Structural Variations in Multiple Myeloma*. **Cancer Informatics**, 13(Suppl2):49-53

PROFESSIONAL  
ACTIVITIES

**Journal Referee**

Bioinformatics, BMC Bioinformatics, BMC Genomics, Journal of Computational Biology, Cancer Informatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Bioinformatics and Biology Insights, Journal of Zhejiang University-Science B, Quantitative Biology, Plos One, Journal of Applied Statistics, Applied Sciences, Scientific Reports, PeerJ, Plos Computational Biology, Molecular Therapy, Cellular Physiology and Biochemistry, Nucleic Acids Research, Frontiers in Genetics, Genetics, Briefings in Bioinformatics, Processes, Sustainability, Information, Genes, Genetic Testing and Molecular Biomarkers, Science Advances, Genome Research, Biometrics, Microbiome, iScience, BMC Medical Genomics, NAR Genomics and Bioinformatics, Nature Communications

**Editorial activities**

- Plos One. 2018~Now
- PeerJ. 2019~Now
- Guest editor, a special issue “Deciphering non-coding regulatory variant: computation and functional validation” of Frontiers in Bioengineering and Biotechnology

**Grant review services**

- American Association of Colleges of Pharmacy New Investigator Award (2019)
- NIH BCHI (Feb 2020)
- NIH BMRD (Nov 2021)
- NIH ZAG1 ZIJ-5 (O2) (June 2022)
- NIH AGCD-1 (Jan 2022)
- NIH AGCD-1 (June 2022)
- NIH ZAG1 ZIJ-Y (J1) (Oct 2022)
- NSF (Jan, Feb, Nov 2020 and Jan 2021)
- Indiana CTSI COVID-19 pilot grant (June 2020)
- Indiana Alzheimer’s Disease Research Center Developmental Projects (Oct 2021)

SOFTWARE  
DEVELOPED

- **hmChIP** (a web server and database for large-scale genomics data query and computing)  
<http://jilab.biostat.jhsph.edu/database/cgi-bin/hmChIP.pl>
- **glmgraph** (an R package for graph-constrained regularization for sparse generalized linear models)  
<http://cran.r-project.org/web/packages/glmgraph/index.html>
- **SICS** (an R package for phylogeny-constrained regularization for sparse generalized linear models with encouraging local smoothing)  
<https://github.com/lichen-lab/SICS>
- **ChIPComp** (an R package for quantitative comparison of multiple ChIP-seq datasets)  
<http://bioconductor.org/packages/devel/bioc/html/ChIPComp.html>
- **traseR** (an R package for GWAS trait-associated SNP enrichment analyses in genomic intervals)  
<http://bioconductor.org/packages/devel/bioc/html/traseR.html>
- **DIVAN** (a pipeline for prioritizing disease-specific noncoding risk variants)  
<https://sites.google.com/site/emorydivan/>
- **glmmTree** (a software for predictive modeling of microbiome data using a phylogenetic tree-based generalized mixed effects model)  
<https://github.com/lichen-lab/glmmTree>
- **GMPR** (a software for normalizing zero-inflated count data particular microbiome sequencing data )  
<https://github.com/lichen-lab/GMPR>

- **TIVAN** (Tissue-specific cis-eQTL single nucleotide variant Annotation and prediction)  
<https://github.com/lichen-lab/TIVAN>
- **BAMMSC** (A Bayesian mixture model for clustering droplet-based single cell transcriptomic data from population studies)  
<https://github.com/lichen-lab/BAMMSC>
- **circMeta** (a unified computational framework for genomic feature annotation and differential expression analysis of circular RNAs)  
<https://github.com/lichen-lab/circMeta>
- **tfLDA** (an R package for applying state-of-the-art topic models to multiple ChIP-Seq datasets to decipher the combinatorial binding events of multiple TFs)  
<https://github.com/lichen-lab/tfLDA>
- **powmic** (an R package for power assessment in microbiome case-control studies)  
<https://github.com/lichen-lab/powmic>
- **MDeep** (an Python package for predictive modeling of microbiome data using deep learning approach)  
<https://github.com/lichen-lab/MDeep>
- **WEVar** (an Python package for predicting regulatory variants)  
<https://github.com/lichen-lab/WEVar>
- **TLVar** (an Python package for predicting noncoding functional variants)  
<https://github.com/lichen-lab/TLVar>
- **DeepPerVar** (an Python package for predicting individual-level noncoding functional variants)  
<https://github.com/lichen-lab/DeepPerVar>
- **MTAE** (an Python package for predicting AD progression using longitudinal methylation data)  
<https://github.com/lichen-lab/MTAE>