

# Ben Li

Email: [ben.li@emory.edu](mailto:ben.li@emory.edu) • Phone: +1(678)-899-1859  
1518 Clifton Rd., Atlanta, Georgia 30322. USA

## EDUCATION

---

<b>Emory University</b>	Atlanta, GA
<i>Ph.D., Biostatistics and Bioinformatics</i>	Sept. 2012 – May 2017 (Expected)
<i>M.S., Biostatistics and Bioinformatics</i>	Sept. 2012 – Aug. 2016
<b>Nanjing University</b>	Nanjing, China
<i>B.S., Computational and Applied Mathematics (Minor in Statistics)</i>	Sept. 2008 – June 2012
<b>University of California, Davis</b>	Davis, CA
<i>Visiting Student</i>	Sept. 2011 – Dec. 2011

## RESEARCH INTERESTS

---

- Bioinformatics
- Bayesian Statistics
- Machine Learning on Biomedical and Healthcare Data Analysis
- Next Generation Sequencing (NGS) Data Analysis
- Transfer Learning on Genomics

## SELECTED PUBLICATIONS

---

\* Co-first author

1. Aly SS, Zhao J, **Li B**, Jiang J. (2014) Reliability of environmental sampling culture results using the negative binomial intraclass correlation coefficient. *Springerplus*. 3:40.
2. Joseph SJ, **Li B**, Ghonasgi T, Haase CP, Qin ZS, Dean D, Read TD. (2014) Direct amplification, sequencing and profiling of Chlamydia trachomatis strains in single and mixed infection clinical samples. *PLoS ONE*. 9(6).
3. Xu T\*, **Li B\***, Zhao M, Szulwach KE, Street RC, Lin L, Yao B, Zhang F, Jin P, Wu H, Qin ZS. (2015) Base-resolution methylation patterns accurately predict transcription factor bindings in vivo. *Nucleic Acids Research*. 43(5):2757-66.
4. 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 Res*. 43(21):e141.
5. **Li B**, Sun Z, He Q, Zhu Y, Qin ZS. (2015) Bayesian inference with historical data-based informative priors improves detection of differentially expressed genes. *Bioinformatics*. 32(5):682-689.
6. Joseph SJ, **Li B**, Petit RA, Qin Z, Darrow L, Read TD. (2016) The single-species metagenome: subtyping Staphylococcus aureus core genome sequences from shotgun metagenomic data. *PeerJ*. To Appear.
7. Qin Z, **Li B**, Conneely KN, Wu H, Hu M, Ayyala D, Park Y, Jin VX, Zhang F, Zhang H, Li L, Lin S. (2016) Statistical challenges in analyzing methylation and long-range chromosomal interaction data. *Statistics in Biosciences*. 1-26.
8. **Li B**, Li Y, Qin Z. (2016) Improving hierarchical models using historical data with applications in high throughput genomics data analysis. *Statistics in Biosciences*. 1-18.

## WORKING PAPERS

---

1. **Li B**, Qin ZS, Wu H. A novel statistical method for differential expression analysis of single cell RNA-Seq data.
2. **Li B**, Qin ZS. IPBTSeq: Using historical data to improve the analysis of RNA-Seq Data
3. **Li B**, Qin ZS. Comparative analyses of whole-genome bisulfite sequencing data across multiple cell types and species.
4. **Li B**, Dileep V, Gilbert DM, Qin ZS. RepliHMM: A Hidden Markov Model based method on the discovery of replication timing states.

## SOFTWARE DEVELOPED

---

1. BinStrain: Statistical modeling for SNP based genotyping of mixed bacterial strains. Available from: <https://github.com/benliemory/BinStrain>
2. Methylphet: Base-resolution methylation patterns accurately predict transcription factor bindings in vivo. Available from: <https://github.com/benliemory/Methylphet>
3. IPBT: Bayesian inference with historical data-based informative priors improves detection of differentially expressed genes. Available from: <https://github.com/benliemory/IPBT>
4. adaptiveHM: Historical rank-based adaptive hierarchical model on genomics data analysis. Available from: <https://github.com/benliemory/adaptiveHM>

## CONFERENCE TALKS & POSTERS

---

1. Eastern North American Region Meetings, Baltimore, MD, 2014. **Using informative priors obtained from historical data significantly improves detection of differentially expressed genes using microarray data**
2. (Invited) SAMSI Bioinformatics: Transition Workshop, Research Triangle Park, NC, 2015. **Comparative analyses of whole-genome bisulfite sequencing data across multiple cell types and species**
3. (Poster) SAMSI Bioinformatics: Statistical Modeling and Analysis of Whole Genome Methylation and Chromatin Interaction (Epigenetics) Workshop, Research Triangle Park, NC, 2015. **Base-resolution methylation patterns accurately predict transcription factor bindings in vivo**
4. (Invited) P01 PPG Basic Biology of Human Pluripotent Stem Cells Scientific Retreat, Lake Oconee, GA, 2016. **Analyses for Replicate Timing and HiC Data**

## PROFESSIONAL EXPERIENCES

---

**Emory University** Atlanta, GA  
*Research Assistant* Sept. 2012 – Present  
- Supervisor: Prof. Zhaohui (Steve) Qin, Prof. Hao Wu

**University of California, Davis** Davis, CA  
*Research Assistant* Sept. 2011 – Dec. 2011  
- Supervisor: Prof. Jiming Jiang

## TEACHING EXPERIENCES

---

**Emory University** Atlanta, GA  
*Teaching Associate/Guest Lecturer*  
- BIOS 731: Advanced Statistical Computing Fall 2016

- BIOS 511: Statistical Inference Spring 2015
- BIOS 731: Advanced Statistical Computing Fall 2014
- BIOS 510: Introduction to Probability Theory Fall 2014

*Teaching Assistant*

- BIOS 591: Biostatistics Methods II Spring 2014
- BIOS 500: Statistical Methods I Fall 2013

**COMPUTING SKILLS**

---

- Extensive experiences in R, Python, C++/C, Matlab, LaTeX
- Knowledgeable in SAS, SPSS, SQL, JAVA
- Familiar with both UNIX and Windows operating systems

**HONORS & AWARDS**

---

- Michael Kutner Distinguished Doctoral Student Award, Emory University 2016
- Ph.D. Fellowship, Emory University 2012-2015
- Outstanding Student Award, Nanjing University 2011
- Renmin Scholarship, Nanjing University 2011, 2010, 2009

**MEMBERSHIPS**

---

- American Statistical Association (ASA)
- International Biometric Society (IBS)