Naihui Zhou

✉ nzhou@iastate.edu   🌐 ashleyzhou972.github.io

🐙 github.com/ashleyzhou972   🎓 Naihui Zhou

Education

PhD | Bioinformatics & Statistics (co-major)
Iowa State University | 2015-2020
- Advisors: Iddo Friedberg and Mark S. Kaiser
- Lora and Russ Talbot Graduate Fellowship in Veterinary Medicine (2019)
- Miller Scholarship (2015 - 2018)

MA | Statistics
Columbia University | 2013-2014
- Relevant coursework: generalized linear models, sample surveys and survival analysis
- Grader for STAT W1111, Intro to Statistics

BS | Information and Computing Science
Beijing Jiaotong University | 2009-2013
- Thesis topic: Bayesian inference on time series data

Selected Publications

- **Zhou, N.**, *et al.* (2019). The CAFA challenge reports improved protein function prediction and new functional annotations for hundreds of genes through experimental screens. Genome Biology. *in press.*
- **Zhou, N.**, *et al.* (2018). Crowdsourcing image analysis for plant phenomics to generate ground truth data for machine learning. PLoS computational biology. 🔗

Software Developed

- **PhiMRF**: R package for Poisson Hierarchical Markov Random Field models for capturing spatial dependence in count data observed on irregular lattices.
- **CAFA Assessment Tool**: Assessment tool for automated protein function prediction using precision-recall.

Selected Talks

- Exploring the 3D spatial dependency of gene expression using Markov Random Fields. **Genome Informatics**, Cold Spring Harbor, NY, 2019
- The Critical Assessment of Protein Function Annotation: Preliminary Assessment. **Intelligent Systems for Molecular Biology**, Prague, Czech Republic, 2017
- Effects of Beta Blockers on Hospital Admission Rates and Cost. **Joint Statistical Meeting**, Seattle, WA, 2015

Experience

ISU Friedberg Lab | Graduate Research Assistant
2015 - Present
- Developed a Poisson hierarchical Markov Random Field model to analyze spatial dependency in RNA-seq data
- Developed an R package for Bayesian inference in a hierarchical autoregressive model
- Implemented algorithms to assess accuracy of structured-output prediction

New York City Health and Hospitals Corporation | Research Intern
2014 - 2015
- Developed a zero-inflated Poisson regression model to analyze patient diagnoses and drug prescriptions from hospital admission data
- Cleaned and organized electronic medical records using SAS

Pfizer China | Intern
2013
- Supported the External Medical Affairs team in organizing academic seminars and conferences

Skills

Statistics
- Hierarchical models, Markov random fields, Bayesian methods, Markov chain Monte Carlo
- Data mining, machine Learning, statistical computing
- Generalized linear mixed models
- Survival analysis, sample surveys

Programming and scientific computing
- Proficient in R, python, C and Linux/Unix shell scripting
- R package development
- Experience in SAS and Matlab
- Parallel computing, OpenMP, LAPACK, BLAS

Large-scale genomics
- RNA-seq data analysis
- Gene Ontology and Gene Ontology Annotation Database

Leadership Experience

- Senator, **Graduate and Professional Students Senate**, Iowa State University, 2017-2018
- Director of Outreach, **BCB Graduate Student Organization**, Iowa State University, 2016-2017
- Volunteer, **MUMO Education and Orphanage Center**, Nairobi, Republic of Kenya, 2011