Kathie Y. Sun

Statistical Geneticist · Bioinformatician · Computational Biologist

Metuchen, NJ

(858) 761-7615 | kathie.y.sun@gmail.com | kathiesun.github.io/ | kathiesun | kathieysun

Education

University of North Carolina, Chapel Hill (UNC-CH) — *Chapel Hill, NC*

Ph.D. in Bioinformatics and Computational Biology (BCB) — *2016 – 2021*

- Minor: Biostatistics

UNC-CH School of Public Health — *Chapel Hill, NC*

M.S.P.H. in Environmental Health Sciences — *2014 – 2016*

Duke University — *Durham, NC*

B.S. in Environmental Sciences and Biology — *2008 – 2012*

Work Experience

Regeneron Genetics Center — *Tarrytown, NY*

Associate Scientist – Bioinformatics — *2021 – present*

- Improve genetic variant annotation to better understand rare human variation in large cohort studies and elucidate complex traits
- Systematically catalogue and analyze carriers of homozygous protein-truncating variants and deleterious mouse knock-out mutations to characterize gene essentiality
- Implement solutions at scale using modern cluster and cloud computing environments
- Lead and contribute to projects with dual goals of 1) broad dissemination to the scientific community as papers and resources, alongside 2) identification and prioritization of potential drug targets using human genetics at scale

Valdar Lab, UNC-CH — *Chapel Hill, NC*

Graduate Research Assistant — *2016 – 2021*

- Successfully defended thesis: *Statistical modeling of parent-of-origin effects and allelic imbalance in genetically-diverse mouse crosses*
- Developed Bayesian hierarchical models and novel computational pipelines to analyze large genomic (DNA- & RNA-seq) data sets
- Customized new methods for variant detection (SNP & CNV), and to estimate genetic and environmental effects on complex behavioral phenotypes and gene expression in genetically-heterogeneous mice
- Presented at international conferences and published work in *Genetics*

Cardno ChemRisk — *Aliso Viejo, CA*

Summer Intern — *2015*

- Conducted epidemiological meta-analysis of human clinical data to evaluate the profile of common or therapy-targetable mutations in lung cancers of never and ever smokers
- Designed study methodology, conducted statistical analysis, and drafted work for manuscript published in *Lung Cancer*

Nylander-French Lab, UNC-CH — *Chapel Hill, NC*

Graduate Research Assistant — *2014 – 2016*

- Designed linear mixed model in SAS containing whole-genome markers – selected from human GWAS (PLINK) and functional pathway analysis – along with environmental exposure predictors to determine the contribution of genetic variants to observed biomarker levels among occupationally-exposed workers
- Presented work at the 2016 Society of Toxicology annual meeting in New Orleans, LA

National Human Genome Research Institute — *Rockville, MD*

Scientific Program Analyst — *2012 - 2014*

- Facilitated cooperation and collaboration between NHGRI Program Directors and academic researchers on extramural precision medicine grant portfolios
- Synthesized information generated from meetings and coordinated 2-3 consortia meetings per year
- Conducted literature reviews resulting in 2 manuscripts

Skills

Statistics	Probability & statistical inference, Bayesian methods, machine learning, data analysis & visualization
Programming & software	R, Python, Stan & JAGS, SQL, SAS, Unix, command line, LaTeX, Vim, Git / version control, Adobe Creative Suite, Microsoft Office
Bioinformatics	NGS data (WGS, RNA-seq, genotyping), GWAS, QTL mapping, polygenic risk scores, pathway annotation and gene set analysis, quantifying heritability

Teaching Experience

UNC-CH School of Public Health and Graduate School — *Chapel Hill, NC*

Teaching Assistant — *2014 – 2018*

- Helped teach two graduate-level core courses in (1) environmental health sciences and (2) statistical modeling
- Guided students toward understanding of difficult concepts, facilitated weekly discussions, and graded tests & homework

How To Learn To Code — *Chapel Hill, NC*

Instructor — *2018*

- Developed curriculum and materials for a introductory summer course on data analysis in R for biological scientists at UNC-CH
- Designed and provided feedback on homework assignments, assessments, and capstone projects

Varsity Tutors and ColPrep, Inc. — *Chapel Hill, NC*

Tutor — *2015 – 2021*

- Private tutoring for all levels of mathematics up through college-level calculus, including SAT & ACT prep and test-taking strategies
- Volunteered as a tutor for students in under-privileged communities around the Durham, NC area

Publications

Bayesian modeling of skewed X inactivation in genetically diverse mice identifies a novel *Xce* allele associated with copy number changes
Kathie Y Sun, Daniel Oreper, Sarah A Schoenrock, Rachel McMullan, Paola Giusti-Rodríguez, Vasyl Zhabotynsky, Darla R Miller, Lisa M Tarantino, Fernando Pardo-Manuel Villena, William Valdar
Genetics 218.1 (2021) iyab034

Lung cancer mutation profile of EGFR, ALK, and KRAS: Meta-analysis and comparison of never and ever smokers
Aaron M Chapman, **Kathie Y Sun**, Peter Ruestow, Dallas M Cowan, Amy K Madl
Lung Cancer 102 (2016) pp. 122–134

POLG2 disease variants: analyses reveal a dominant negative heterodimer, altered mitochondrial localization and impaired respiratory capacity
Matthew J Young, Margaret M Humble, Karen L DeBalsi, **Kathie Y Sun**, William C Copeland
Human Molecular Genetics 24.18 (2015) pp. 5184–5197

The ethical, legal, and social implications program of the National Human Genome Research Institute: reflections on an ongoing experiment
Jean E McEwen, Joy T Boyer, **Kathie Y Sun**, Karen H Rothenberg, Nicole C Lockhart, Mark S Guyer
Annual Review of Genomics and Human Genetics 15 (2014) pp. 481–505

Evolving approaches to the ethical management of genomic data
Jean E McEwen, Joy T Boyer, **Kathie Y Sun**
Trends in Genetics 29.6 (2013) pp. 375–382

Awards

2016 – 2021 **Caroline H. & Thomas S. Royster Fellowship**, UNC-CH Graduate School
2017 – 2018 **T32 Research Fellowship**, National Institute of General Medical Sciences
2016 **Student Scholarship**, American Industrial Hygiene Foundation
2014 – 2016 **T32 Research Fellowship**, National Institute of Occupational Safety and Health
2014 – 2015 **David Allison Fraser scholarship**, UNC-CH School of Public Health
2012 ***Cum laude* and Graduation with High Distinction in Biology**, Duke University
2008 – 2012 **Academic Scholarship**, Becton Dickinson Biosciences

Leadership Activities

2020 - 2021 **Committee Member**, Royster Distinguished Professor for Graduate Education Search Committee — *Chapel Hill, NC*
2019 - 2021 **Senator**, UNC-CH Graduate and Professional Students Federation — *Chapel Hill, NC*
2019 - 2021 **Committee Member**, BCB Students' Invited Speaker Seminar Series Committee — *Chapel Hill, NC*
2019 **Scientist Ambassador**, *DNA Day* teaching genetics in high school classrooms — *Wilkesboro, NC*
2017 **Conference Speaker and Participant**, Royster Global Conference at King's College London — *London, UK*
2015 - 2016 **Vice-President**, Environmental Sciences and Engineering Student Organization — *Chapel Hill, NC*
2008 - 2019 **Team Leader & Coxswain**, Duke Varsity Women's Rowing and Carolina Masters Club Crew — *Durham, NC*