SHAOBO (SEBASTIAN) LI, Ph.D.

Los Angeles, CA, 90017 • (213)-270-4222 • sebastian.li@outlook.com

linkedin.com/in/sebastian-shaobo-li-22864217 • https://sebastian-li.github.io/

EDUCATION

University of Southern California, Keck School of Medicine   Los Angeles, CA
PhD in Cancer Biology and Genomics (Bioinformatics)   2018.08 – 2022.05
• Dissertation: Perinatal epigenetic and genetic analyses in childhood cancers

Fudan University, Shanghai Medical College   Shanghai, China
Bachelor of Medicine (Forensic Medicine)   2009.09 – 2014.06
• Thesis: Expression of CPEB4 in invasive ductal breast carcinoma and its prognostic significance

WORK EXPERIENCE

University of Southern California, Center for Genetic Epidemiology   Los Angeles, CA
Postdoctoral Fellow   May 2022 – Present
• Led multiple large scale GWAS and EWAS projects on childhood cancers
• Designed deep learning pipelines to predict childhood leukemia and eye disease risks with DNA methylation
• Built bioinformatics pipelines for germline/somatic variant discovery with WGS data
• Trained new PhD students and pediatrician research fellows

University of California, Berkeley, School of Public Health   Remote
Research Affiliate   July 2019 - Present
• Carried out childhood leukemia GWAS analyses; pesticide exposure EWAS analyses
• Assessed telomere length and aging changes in children with childhood leukemia
• Managed and monitored proper use of UC Berkeley HPC and delivered routine reports on research progress

University of Southern California, Center for Genetic Epidemiology   Los Angeles, CA
Predoctoral Research Fellow   June 2019 – May 2022
• Conducted high-throughput association models to investigate associations between disease risks and neonatal genetics, epigenetics, cytokines, T-cell receptors, and nucleated cell proportions
• Performed ancestry analysis to investigate contribution of European ancestry to Latino population disease risks

Zhongshan Hospital   Shanghai, China
Clinical Researcher   Sept 2014 – June 2018
• Applied machine learning methods to predict portal pressure and patient outcome from CT-based radiomics
• Collaborated with surgeons and internal doctors on data collection, experiment design and manuscript writing

SKILLS

Data Analysis: Data Cleaning, Visualization, Modelling, Machine Learning, Deep Learning, Statistical Genetics, Ancestry Analyses, Clinical Analyses; **Programming:** R, Python, Linux/Unix (HPC), Bash, SQLITE, C, JAVA, SAS, STATA, SPSS, Git**; Bioinformatics:** EWAS, GWAS (PLINK, SNPTEST), PRS, GATK, Dock/Singularity, Seurat; **Graphic and video productions**: Adobe Suite (Photoshop, Illustrator, Lightroom, After Effect), Apple Final Cut Pro, and Apple Motion

CERTIFICATES

HarvardX	High-Dimensional Data Analysis	May 2017
HarvardX	Statistical Inference and Modeling for High-throughput Experiments	Aug 2017
StanfordOnline	Machine Learning	Oct 2021

PUBLICATIONS

22 peer-reviewed publications (4 first author), 1 first author under review, 2 first author in submission, 3 conference presentations. Full list available on ORCID: https://orcid.org/0000-0002-0544-5338