SHAOBO (SEBASTIAN) LI, Ph.D.

Los Angeles, CA, 90404 • 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)   Aug 2018 – May 2022
• Dissertation: Perinatal epigenetic and genetic analyses in childhood cancers

Fudan University, Shanghai Medical College   Shanghai, China
Bachelor of Medicine (Forensic Medicine)   Sept 2009 – June 2014
• Thesis: Expression of CPEB4 in invasive ductal breast carcinoma and its prognostic significance

WORK EXPERIENCE

Kite Pharma, a Gilead Company   Santa Monica, CA
Research Scientist, Computational Biology   Nov 2022 – Present

University of Southern California, Center for Genetic Epidemiology   Los Angeles, CA
Postdoctoral Fellow   May 2022 – Nov 2022
• 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

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

University of California, Berkeley, School of Public Health   Remote
Research Affiliate   July 2019 – Nov 2022
• 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

Zhongshan Hospital   Shanghai, China
Clinical Researcher (Volunteer)   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, Machine Learning, Deep Learning, Statistical Genetics, Ancestry Analyses, Clinical Analyses; **Programming:** R, Python, Linux/Unix, AWS, SQLITE, C, JAVA, SAS, STATA, SPSS, Git; **Bioinformatics:** EWAS, GWAS, PRS, GATK, Dock/Singularity, Seurat; **Media productions**: Adobe Suite (Photoshop, Illustrator, After Effect), Apple Final Cut Pro, Apple Motion

CERTIFICATES

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

PUBLICATIONS

23 peer-reviewed publications (5 first author), 3 conference presentations. Full list available on ORCID:
https://orcid.org/0000-0002-0544-5338