RICHARD SHUAI

richardshuai98@gmail.com | 818-398-1883 | www.linkedin.com/in/richard-shuai | www.github.com/richardshuai

EDUCATION

University of California, Berkeley **2018-2022**
B.A. Computer Science & B.A. Statistics GPA: 4.0 / 4.0

RESEARCH EXPERIENCE

Gray Lab, *NSF REU Research Fellow*, Johns Hopkins University **Jun. 2021 – Current**

Using generative language modeling to design synthetic antibody libraries. Trained a ~12.8M parameter GPT-based transformer on ~558M antibody sequences. Sampled novel antibody sequences, measured properties for therapeutic developability, and visualized antibody structural predictions with PyMOL. Patent pending.

Ioannidis Lab, *Research Assistant*, UC Berkeley **Jan. 2021 – Current**

Using CNN-based deep learning models with scATAC-seq data to prioritize variants affecting chromatin accessibility in kidney tissue. Trained models and evaluated performance on predicting allelic imbalance. Analyzed first-layer convolutional filters and in silico saturation mutagenesis outputs to discover transcription factor binding motifs.

Waller Lab, *Research Assistant*, UC Berkeley **Nov. 2019 – Mar. 2021**

Used deep learning for real-time image reconstruction from convolutions with a spatially-variant point spread function. Proposed and trained a U-Net with learnable Wiener deconvolutions, achieving ~1000x speedup against existing iterative algorithms with similar reconstruction quality.

Abbasi Lab, *UC Berkeley URAP Apprentice*, UCSF **Sept. 2020 – Dec. 2020**

Worked on statistical pipelines to associate multiple sclerosis patient genotypes and RNA-seq data with their phenotypes. Performed exploratory analysis of gene expression data using DESeq2 and clustering algorithms.

Shuai Lab, *Research Assistant*, UCLA **May 2019 – Sept. 2019**

Discovered supporting evidence for a novel epigenetic regulatory pathway in chemokine production with RNA-seq analysis. Aligned and counted reads using STAR, detected differentially expressed genes with edgeR and DESeq2, and created visualizations of significant results.

EMPLOYMENT

Amazon, *Software Engineering Intern* **May 2020 – Aug. 2020**

Developed pricing configuration page for Amazon Pricing Team and external business users for competitive price analysis with Agile development. Expanded backend functionality within Java Spring Framework to add key features. Revamped previous UI with React Redux, supporting 8 month phase-out of previous interface.

Abiocode, Inc., *Lab Technician* **May 2018 – Aug. 2018**

Prepared Arabidopsis sample protein extracts for gel electrophoresis. Performed SDS-PAGE and Western Blotting experiments, using basic lab techniques such as pipetting and centrifugation.

WORKSHOPS & PRESENTATIONS

- **Richard W. Shuai**, Jeffrey A. Ruffolo, Jeffrey J. Gray. "Generative Language Modeling for Antibody Design." *NeurIPS Workshop on Machine Learning for Structural Biology*, poster (2021).

- **Richard W. Shuai**, Jeffrey A. Ruffolo, Jeffrey J. Gray. "Generative Language Modeling for Antibody Design." *RosettaCON,* poster (2021).

PUBLICATIONS & PRE-PRINTS

- **Richard W. Shuai**, Jeffrey A. Ruffolo, Jeffrey J. Gray. "Generative Language Modeling for Antibody Design." *bioRxiv* (2021).

- Kyrollos Yanny*, Kristina Monakhova*, **Richard W. Shuai**, Laura Waller. "Deep learning for fast spatially-varying deconvolution." *Optica* 9*, 96-99 (2022)*

- **Richard W. Shuai***, Kyrollos Yanny*, Kristina Monakhova, Laura Waller. "MultiWienerNet: Deep Learning for Fast Shift-Variant Deconvolution." *Computational Optical Sensing and Imaging*, pages CTh5A–5. Optical Society of America (2021).

- Kathleen M. Yee, **Richard W. Shuai**, Bin Liu, Christian A. Huynh, Chao Niu, Hailey R. Lee, Min S. Lee, Jirui Wen, Jian Zou, Jiang Wu, Ke Shuai. "TET1 controls Cxcl1 induction by DNA demethylation and promotes neutrophil recruitment during acute lung injury." *bioRxiv* (2021).

SKILLS

Programming languages: Python, Java, R, C, Javascript
Libraries: PyTorch, Tensorflow, Keras, HuggingFace, NumPy, pandas, SciPy
Molecular modeling: PyRosetta, PyMOL
Relevant coursework: Machine Learning, Deep Learning, Computational Methods in Genomics and Medicine, Artificial Intelligence, Optimization Models, Probability and Random Processes, Data Structures, Linear Algebra, Linear Modeling