Helyaneh Ziaei Jam

949.201.5332 | hziaeija@ucsd.edu | https://heliziii.github.io/
3869 Miramar St, 3923, La Jolla, CA, 92092

EDUCATION

University of California, San Diego, *GPA: 3.93/4* Jan. 2021 – Present
Ph.D. in Computer Science and Engineering. Advisor: Prof. Melissa Gymrek San Diego, CA

Sharif University of Technology, *GPA: 3.87/4* Sep. 2015 – July. 2020
B.Sc. in Software Engineering Tehran, Iran

RESEARCH INTERESTS

- Population genetics
- Statistical genetics
- Computational Biology

PUBLICATIONS

- **Ziaei Jam H**, et al. "A deep population reference panel of tandem repeat variation." **Nature Communications** 2023. PMID: 36945429
- **Ziaei Jam H**, et al. "Genome-wide profiling of genetic variation at tandem repeat from long reads." **Genome Biology** 2024. PMID: 38328152
- English A, Dolzhenko E, **Ziaei Jam H**, et al. "Analysis and benchmarking of small and large genomic variants across tandem repeats." **Nature Biotechnology** 2024. PMID: 38671154
- Lundström OS, Verbiest MA, Xia F, **Ziaei Jam H**, et al. "WebSTR: A Population-wide Database of Short Tandem Repeat Variation in Humans." **Journal of Molecular Biology** 2023. PMID: 37678708
- Sehgal A, **Ziaei Jam H**, et al. "Genome-wide detection of somatic mosaicism at short tandem repeats." **Bioinformatics** 2024.
- Huang B, Durvasula A, Mousavi N, **Ziaei Jam H**, et al. "Genome-wide selection inference at short tandem repeats." preprint on biorxiv

EXPERIENCE

Uber Technologies Inc. June 2024 – September 2024
Ph.D. Software Engineer Intern Sunnyvale, CA

- **Embedding analytics:** Developed a Residual-Quantized Variational AutoEncoder (RQ-VAE) model to learn semantic IDs from text embeddings, enhancing performance in downstream tasks such as hierarchical classification and recommender systems.

Illumina, Inc. AI Lab June. 2022 – September 2022
Data Scientist Intern Foster City, CA

- **Non-linear covariate correction for phenotype prediction:** Trained and tested Boosted Trees and Multi-layer perceptron models to assess their performance in predicting phenotype values based on environmental factors and common associated variants. Compared the performance against a typical linear model with various methods including rare variant and enrichment analysis. Increased the averaged variation explained by 4.16% and found at least one significant false positive rare variant.

University of California San Diego April. 2021 – Present
Graduate Research Assistant at Gymrek lab, selected projects San Diego, CA

- **Improved Multi-ancestry Polygenic Risk Score estimation through parameter sharing:** Introduced a novel Bayesian Shrinkage Model for simultaneous parameter learning across multiple ancestries, enabling inter-population knowledge sharing to enhance the estimation of causal variables for specific traits.
- **A global reference for human genetic variations in tandem repeats:** Developed a method, EnsembleTR, that takes tandem repeat calls from various tools as input and outputs a consensus TR callset, leading to the largest tandem repeat catalog to date. Conference talk at **ASHG 2022**, Los Angeles, CA, **SVAR 2023**, Stanford, CA.
- **Genotyping tandem repeats using long reads:** Developed LongTR, a tool written in C++ for quantifying tandem repeats using long reads sequencing data. Our tool shows superior performance compared to the state-of-the-art. Presented as an oral talk at **RECOMB-seq 2022**, San Diego, CA.

Sharif University of Technology November. 2018 – July. 2020
Research Assistant Tehran, Iran

- **Assessing potential of Stem Cells to generate new vessels:** Trained a **U-net** neural network using Keras for vessels detection in human tissue images to assess the potential of various types of stem cells in forming new vessels.

Max Plank Institute Summer 2019 – Winter. 2020
Research Intern Under supervision of Prof. Tobias Marschall Saarbrücken, Germany

- Designed and tested a **classifier** for partitioning long reads to haplotype-specific sets based on their alignments to the De Bruijn graph.

University of Tartu Summer 2018
Software Engineer Intern Tartu, Estonia

- Design and implementation of a high-level and object-oriented quantum programming language in C++ that translates to different quantum assembly languages such as QASM.

MENTORSHIP AND OUTREACH

- **Workshop organizer and tutor** Organized and led a DS-I Africa workshop on tandem repeat expansion analysis, featuring coding sessions, short talks, and lectures. The workshop was attended by 17 selected attendants, including PhD candidates and postdoctoral researchers, from various African countries.
- **Graduate mentor, Graduate Women in Computing at UCSD, GradWIC** Mentored an international first-year PhD student in her academic and professional path. 2023-2024 academic year.
- **Graduate research mentor** Provided mentorship to three students at various academic levels from Nigeria and Uganda, guiding them in conducting research within the field of tandem repeats. Since 2023.

HONORS AND AWARDS

- Selected for the 2025 class of the **Siebel Scholars** program at UC San Diego.
- Recipient of the 2024-25 **Friends of the International Center Fellowship**, associated with the UC San Diego University.
- Recipient of UC San Diego University **Doctoral Fellowship**, United States, 2021.
- Recipient of the fellowship for a research internship at Max Planck Institute for Informatics, Germany, Summer 2019, Winter 2020.
- Recipient of the scholarship for a research internship at the University of Tartu, Estonia, 2018
- **Silver Medal** in the National Astronomy and Astrophysics Olympiad Iran, 2014
- **198th** amongst More Than 180,000 Participants in Iranian Nation-wide University Entrance Exam, Mathematics and Physics Discipline, 2015.