Sujit Silas Armstrong Suthahar

sujitsilas@gmail.com | https://www.linkedin.com/in/sujitsilasarmstrong/ | 951-756-0678 | 840 Hilgard Ave, Los Angeles, CA, 90024

WORK EXPERIENCE

UCLA Institute for Society and Genetics Los Angeles, CA
Graduate Student Researcher 09/2023 - Present

- Spearheaded alignment and analysis of RRBS, WGBS, ATAC-seq, and RNA-seq data reducing turnaround times by 25%; analyzed bisulfite sequencing data and RNA-seq data using matrix factorization methods identifying key regulative elements
- Devised NextFlow pipelines for both local and distributed computing environments, enhancing reproducibility while reducing data preprocessing times by 30% of NGS data analysis workflows
- Conducted workshops on Nextflow nf-core pipelines, enabling non-expert users to efficiently set up environment, submit, and manage jobs on high-performance computing clusters

GRAIL Menlo Park, CA
Computational Biology & Machine Learning Team 06/2024 - 09/2024

- Optimized classifier features and Galleri assay probes by implementing a novel false positive investigation toolkit increasing cell-free DNA detection sensitivity for a few early-stage cancer types by 8%
- Deployed a genome annotation pipeline reducing non-specific probe selection by 10% through DMR analysis, resulting in a revamped targeted panel and improved classifier features
- Refactored functions, updated pipelines, and landed two merge requests related to development of two vignettes and false positive analysis workflow

UCLA Biomedical AI Research Lab Los Angeles, CA
Research Scientist 09/2023 - 06/2024

- Implemented an ensemble-based deep learning model using multiple instance learning, achieving 76% diagnostic accuracy for indeterminate thyroid nodules—a 25% improvement over traditional methods and decreasing unnecessary biopsies
- Leveraged CUDA architecture to accelerate deep learning model training on GPUs within a High-Performance Computing (HPC) environment, cutting down processing time and improving model iteration speed
- Developed an advanced U-Net architecture featuring enhanced skip connections and integrated image preprocessing techniques achieving a sensitivity of 60% in biopsy needle segmentation problems for medical imaging

La Jolla Institute for Immunology San Diego, CA
Bioinformatics Tech 06/2021 - 07/2023

- Wrote custom bash scripts for automating pipelines for sequencing data alignment and preprocessing (scRNA-seq, CITE-seq, TCR-seq, RNA-seq WGS, WES), reducing processing time by 40% while standardizing alignment and preprocessing protocols
- Led end-to-end RNA-seq and scRNA-seq analysis workflows using DESeq2, EdgeR, Seurat, and other packages performing comprehensive data exploration and visualization to derive biological insights from bulk and single-cell datasets
- Developed and applied a computational method to validate and optimize antibody titration concentrations for a panel of 124 antibodies used in CITE-seq experiments, improving accuracy and reliability of multi-omics data integration
- Designed a statistical framework for analyzing 348,000+ TCR sequences from CVD PBMC samples, identifying 672 APOB-reactive T-cell clones and 532 conserved motifs, tracking disease-specific T-cells across donors
- Developed and deployed machine learning solutions for characterizing new cell types in single cell RNA-seq data, including implementing a SVM classifier to identify novel ex-T Regulatory cell populations in CD4T sub-clusters
- Spearheaded cross-functional research projects leveraging Python/R for statistical analysis, establishing streamlined workflows for data-driven discoveries contributing to over 12+ publications, including 4 first-authored papers in top journals

University of California, Riverside Riverside, CA
Research Assistant (RA) 10/2019 - 06/2021

- Conducted in-depth mRNA transcriptome analysis using R and Python, spearheading an independent research project comparing hESC transcriptomes of differentiating neural crest and ectodermal cells
- Employed advanced protein modeling tools including Pymol, I-Tasser, PSIPRED, Robetta, and SAVES 6.0 to elucidate DNA-transcription factor (TF) interactions in early development, integrating molecular biology knowledge with computational approaches to gain novel insights
- Gained proficiency in molecular biology techniques, including cell culturing, western blotting, immunofluorescence imaging, and ELISA, to support experimental workflows
- Managed and maintained hESC cell lines and other cell lines, performing knockdown treatments to investigate gene function and cellular responses

SELECTED PUBLICATIONS

- **Armstrong, S. S.**, Chen, D. G., Kumar, S., Heath, J. R., Feinstein, M. J., Greenland, J. R., Calabrese, D. R., Lanier, L. L., Ley, K., & Shemesh, A. (2025). CITE-Seq Analysis Reveals a Differential Natural Killer Cell SPON2 Expression in Cardiovascular Disease Patients Impacted by Human-Cytomegalovirus Serostatus and Diabetes. International Journal of Molecular Sciences, 26(3), 1369. https://doi.org/10.3390/ijms26031369
- **Armstrong Suthahar SS**, Nettersheim FS, Alimadadi A, Wang E, Billitti M, Resto-Trujillo N, Roy P, Hedrick CC, Ley K, Orecchioni M. Olfr2-positive macrophages originate from monocytes proliferate in situ and present a pro-inflammatory foamy-like phenotype. Cardiovasc Res. 2024 Sep 4:cvae153. doi: 10.1093/cvr/cvae153. Epub ahead of print. PMID: 39229899
- **Armstrong, S. S.,** Nettersheim, F. S., Durant, C., Blanco-Dominguez, R., Roy, P., Orecchioni, M., Suryawanshi, V., & Ley, K. (2022). Titration of 124 antibodies using CITE-Seq on human PBMCs. Scientific Reports, 12 (1), Article 1
- **Armstrong, S. S.,** Chen, D. G., Kumar, S., Heath, J. R., Feinstein, M. J., Greenland, J. R., Calabrese, D. R., Lanier, L. L., Ley, K., & Shemesh, A. (2024). CITE-seq analysis reveals human cytomegalovirus and diabetes-associated adaptive NK cell alterations in cardiovascular disease. bioRxiv, 2024.03.22.581997
- **Armstrong Suthahar, S. S.,** ROY, P., Makings, J., & Ley, K. (2023). Identification of apolipoprotein B-reactive CDR3 motifs allow tracking of atherosclerosisrelated memory CD4+ T cells in multiple donors. Frontiers in Immunology, 15 , 1302031
- **Armstrong Suthahar, S. S.** (2021). A comparative transcriptomic analysis of differentiating neural crest cells and neuroectoderm or ectoderm cells at early development
- Nettersheim, F. S., Brunel, S., Sinkovits, R. S., **Armstrong, S. S**., Roy, P., Billitti, M., Kobiyama, K., Alimadadi, A., Bombin, S., Lu, L., Zoccheddu, M., Oliaeimotlagh, M., Benedict, C. A., Sette, A., & Ley, K. (2024). PD-1 and CD73 on naive CD4+ T cells synergistically limit responses to self. *Nature Immunology*, 1–11. https://doi.org/10.1038/s41590-024-02021-6
- Freuchet, **Armstrong, S. S.,** A., Roy, P., Oliaeimotlagh, M., Kumar, S., Orecchioni, M., Ali, A. J., Khan, A., Makings, J., Lyu, Q.,Winkels, H., Wang, E., Durant, C., Ghosheh, Y., Gulati, R., Nettersheim, F., & Ley, K. (2023). Identification of human exTreg cells as CD16+CD56+ cytotoxic CD4+ T cells. Nature Immunology, 24 (10), 1748–1761
- Nettersheim, F. S., Brunel, S., Sinkovits, R. S**., Armstrong, S. S.,** Roy, P., Billitti, M., Kobiyama, K., Alimadadi, A., Bombin, S., Lu, L., Zoccheddu, M., Oliaeimotlagh, M., Benedict, C. A., Sette, A., & Ley, K. (2024). PD-1 and CD73 on naive CD4+ T cells synergistically limit responses to self. *Nature Immunology*, 1–11
- Athreya, S., Melehy, A., **Armstrong Suthahar, S. S.,** Ivezić, V., Radhachandran, A., Sant, V., ... & Speier, W. (2024). Reducing Overtreatment of Indeterminate Thyroid Nodules Using a Multimodal Deep Learning Model. arXiv preprint arXiv:2409.19171
- Vallejo, J., Saigusa, R., Gulati, R., **Armstrong Suthahar, S. S.,** Suryawanshi, V., Alimadadi, A., Durant, C. P., Ghosheh, Y., Roy, P., Ehinger, E., Pattarabanjird, T., Hanna, D. B., Landay, A. L., Tracy, R. P., Lazar, J. M., Mack, W. J., Weber, K. M., Adimora, A. A., Hodis, H. N., ... Ley, K. (2022). Combined protein and transcript single-cell RNA sequencing in human peripheral blood mononuclear cells. BMC Biology, 20 (1), 193
- Saigusa, R., Roy, P., Freuchet, A., Gulati, R., Ghosheh, Y., **Armstrong Suthahar, S. S.,** Durant, C. P., Hanna, D. B., Kiosses, W. B., Orecchioni, M., Wen, L., Wu, R., Kuniholm, M. H., Landay, A. L., Anastos, K., Tien, P. C., Gange, S. J., Kassaye, S., Vallejo, J., Ley, K. (2022). Single cell transcriptomics and TCR reconstruction reveal CD4 T cell response to MHC-II-restricted APOB epitope in human cardiovascular disease. Nature Cardiovascular Research , Article 5

ABSTRACTS AND PRESENTATIONS

- *Multiomic Analysis of RNA Sequencing, Reduced Representation Bisulfite Sequencing, and Metabolomics Data Using Coupled Matrix Factorizations,* Society of Toxicology Conference, January 2025 – abstract & poster presentation
- *Wasserstein Cycle-consistent Generative Adversarial Network for RNA -seq and scRNA -seq Translation,* UCLA Bioengineering Conference, February 2024 – abstract, and poster presentation
- *scRNA -Seq with CITE -Seq reveals differential surface marker and gene expression in natural killer cells from subjects with and without coronary artery disease*, American Association of Immunology Conference, February 2023 – abstract
- *Enhancing thyroid cancer diagnosis of indeterminate nodules through multi-modal analysis of ultrasound images*, UCLA Jonsson Comprehensive Cancer Center Annual All-Center Symposium, March 2024 – abstract and poster

EDUCATION

University of California, Los Angeles (UCLA) | Ph.D. Bioengineering	GPA: 4.00/4.00
University of California, Los Angeles (UCLA) | M.S. Bioengineering (Biomedical Data Sciences)	GPA: 3.97/4.00
University of California, Riverside (UCR) | B.S. Biology, Honors (Biochemistry and Bioinformatics)	GPA: 3.50/4.00

SKILLS

Programming: Proficient in Python, R, Bash/Shell, Go, SQL, JavaScript, C/C++, CSS/HTML, and LaTex

Software: MS Word, PowerPoint, Excel, Pymol, Linux, Blender, AutoCAD, IBM SPSS, SAS/STATA, Docker, Git

Machine Learning / Deep Learning: TensorFlow, Keras, PyTorch, Scikit-learn, CUDA, TensorLy, numpy, GLIPH (TCR analysis)

Bioinformatics Tools: STAR aligner, Bowtie2, Cellranger, GATK4, SALMON, DESeq2, Seurat, Scanpy, NextFlow, PyMol, BioPython, Bioconductor, PSIPRED, DeepFold, edgeR, dplyr, tidyr, ggplot2, CibersortX, Cytoscape, matplotlib, seaborn, WGCNA, MiXCR, MethylKit, ChIPseeker, DSS, GSEA, GenomicRanges, Monocle3, VEP, Immunarch, VDJtools, CellChat