Sofia Salazar-Magaña

1400 E 55th Pl, Chicago IL, USA 60637 | sogusama@gmail.com | GitHub | ORCID

Education

Bachelor's degree in Genomic Sciences 2020 – 2024
National Autonomous University of Mexico (UNAM), Querétaro, México.
Grade: 9.89/10 (3.96 GPA)

Research Experience

Research laboratory data analyst October 2024 – Present
University of Chicago, Chicago, Department of Medicine, Section of Genetic Medicine, IL, USA.
PI: Hae Kyung Im, PhD
- Applied a novel lab-developed method leveraging deep learning-based DNA-to-epigenome predictions to investigate TF-binding alterations associated with breast cancer risk.
- Developing a Transcriptome-Wide Association Study (TWAS)-like framework by training deep-learning models that predict cancer-associated alternative polyadenylation (APA) events.
- Developing approach to optimize DNA sequence-based genomic models for accurate predictions on individual genomes.

Undergraduate research assistant July 2023 – October 2024
University of Chicago, Remotely.
PI: Hae Kyung Im, PhD
- Compared the performance of Borzoi and Enformer deep learning models for DNA sequence-based prediction of epigenomic features using LCL expression measurements from the 1000 Genomes Project.
- Applied epigenomic prediction software to reconstruct and analyze Neanderthal epigenomes.
- Provided technical support for lab-developed software enabling TWAS analyses.

Undergraduate research assistant January 2022 – October 2024
National Autonomous University of Mexico (UNAM), International Laboratory for Human Genome Research (LIIGH), Querétaro, México.
PI: Alejandra Medina-Rivera, PhD
- Analyzed RNA-seq data from the Mexican Lupus Registry (LupusMX) to identify key differentially expressed genes and assembled gene regulatory networks, contributing novel insights into the disease etiology within an understudied population.
- Assembled a bioinformatic pipeline for RNA-seq data analysis, enabling lab-wide workflows.

Contributions

Zhou, Y., Adeluwa, T., Zhu, L., **Salazar-Magaña, S.**, Sumner, S., Kim, H., Gona, S., Nyasimi, F., Kulkarni, R., Powell, J. E., Madduri, R., Liu, B., Chen, M., & Im, H. K. (2025). scPrediXcan integrates deep learning methods and single-cell data into a cell-type-specific transcriptome-wide association study framework. *Cell Genomics*, *5*(5), 100875. https://doi.org/10.1016/j.xgen.2025.100875

Hernández-Ledesma, A. L., **Salazar-Magaña, S**., Coss-Navarrete, E. L., Nuñez-Reza, K. J., Tinajero Nieto, L., Peña-Ayala, A., Torres-Valdez, E., Felix Rodríguez, G., Frontana-Vázquez, G., Rosetti, F., Fernández-Valverde, S. L., Gutierrez Arcelus, M., Alpizar-Rodriguez, D., & Medina-Rivera, A. (2024). AB1113 allele specific gene expression during tolerance induction on dendritic cells from Mexican women with systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, *83*, 1887. https://doi.org/10.1136/annrheumdis-2024-eular.2067 **(Conference abstract)**

Adeluwa, T., **Salazar-Magaña, S**., Sumner, S., Gona, S., Nyasimi, F., Carbonetto, P., Baca, S., Freedman, M., Gusev, A., Liu, B., Madduri, R., Stephens, M., Gao, G., Amariuta, T., & Im, H. K. (2025). TFXcan reveals transcription factor programs driving GWAS loci. **(Under revision)**

Salazar-Magaña, S., Gao, G., Adeluwa, T., Gona, S., Nyasimi, F., Madduri, R., Epstein, M., Marks, J., Pharoah, P., Schildkraut, J. M., & Im, H. K. **(Manuscript in preparation).** Linking transcription factor binding variation to overall and subtype-specific breast cancer risk variants using a deep learning-based approach.

Herández-Ledesma, A. L., **Salazar-Magaña, S**., Coss-Navarrete, E., Ramírez-Espinosa, D., Rosetti, F., Fernandez-Valverde, S. L., Gutierrez-Arcelus, M., Alpízar-Rodríguez, D., & Medina-Rivera, A. **(Manuscript in preparation).** Analysis of gene expression profiles during tolerance induction in dendritic cells from women with Systemic Lupus Erythematosus.

Ogonowski, N. S., Cao, F., Flores-Ocampo, V., **Salazar-Magaña, S**., Seviiri, M., Song, L., Kugelman, J., MacGregor, S., Cuellar-Partida, G., Gharahkhani, P., Im, H. K., Harding, I. H., Yang, J., Nayler, S., Kumar, K. R., Martin, N. G., Diaz-Torres, S., & Rentería, M. E. **(Manuscript in preparation)**. GWAS Meta-Analysis of Essential Tremor Uncovers 47 Novel Genetic Loci and Biological Pathways Linked to Brain Morphometry of the Ventral Diencephalon and Cerebellum.

Conference presentations

Deep learning predicts TF binding variation, illuminating breast and ovarian cancer biology. (Poster). *American Society of Human Genetics (ASHG) Annual Conference,* 2024.

Studying the Systemic Lupus Erythematosus expression signature and contextualizing it in a Mexican cohort. (Talk). *Academic Days at LIIGH Institutional Symposium,* 2024.

Lupus peripheral blood meta-analysis with Recount3 & Recount3-like processed RNA-seq data. (Poster). *Academic Days at LIIGH Institutional Symposium,* 2023.

Lupus peripheral blood meta-analysis with Recount3 & Recount3-like processed RNA-seq data. (Poster). *ISCB-LA SoIBio BioNetMX International Conference on Bioinformatics,* 2022.

Teaching Experience

TA at Quantitative Genomics Training Boot Camps, *Columbia University.*	Summer '25 & '24
TA at Deep Learning in Genomics course, *University of Chicago.*	Spring '25
Lecturer at bioinformatics institutional courses, *International Laboratory for Human Genome Research (LIIGH).*	2024 & 2023

Certifications & Relevant courses
- **Intro to AI-driven Science on Supercomputers,** *Argonne National Laboratory.* Credential ID: a1ff9ac49d6d6754ce667fae36f8ad83f8fe93a5
- **Google's Data Analytics Professional Certificate** Credential ID: KEBDX9FHLDSG
- **Workshop: Analysis of biological and biomedical transcriptomic data using Artificial Intelligence (AI) and Machine Learning (ML) methods with R,** *ISCB-LA SoIBio BioNetMX International Conference on Bioinformatics, 2022.*
- **Workshop: Introduction to the logical modeling of cellular networks, with applications to cancer and immunology**, *International Laboratory for Human Genome Research (LIIGH).*

Skills
- Programming (R, Python, Bash, C++, Git, PHP, SQL, Perl).
- Bioinformatics (bulk and single-cell RNA-seq workflow, genome alignment, epigenomic data analysis e.g ChIP-seq, ATAC-seq).
- Deep learning and hardware acceleration: Pytorch, Tensorflow, CUDA, Intel XPU, MPS.
- Statistics, TWAS.
- High Performance Computing (SGE, SLURM).

Languages
- Spanish (Native)
- English (Proficient – TOEFL iBT 118)
- French (Basic)