

Shan A. Sabri

BIOINFORMATICS SCIENTIST · COMPUTATIONAL BIOLOGIST

☎ (630) 201-1911 | ✉ ShanASabri@gmail.com | 🏠 ShanSabri.github.io | 📄 ShanSabri | 📺 ShanSabri | 📖 Publications

Summary

Computational biologist with expertise in performing end-to-end data analysis, visualization, and interpretation of data from various NGS-based platforms (ChIP-/DNase-/ATAC-/RNA-seq, single cell RNA-/ATAC-seq, HiC). Actively developing innovative, sophisticated computational methods and tools for data analysis. Strong interpersonal, written and oral communication skills as evidenced by multiple collaborations, fellowships, and presentation awards. A complete vitae can be found [here](#).

Education

University of California, Los Angeles

PH.D. IN BIOINFORMATICS, EMPHASIS MACHINE LEARNING

Los Angeles, CA

June 2020

Johns Hopkins University

M.Sc. IN BIOINFORMATICS

Baltimore, MD

Dec. 2013

Loyola University Chicago

B.Sc. IN BIOINFORMATICS AND MATHEMATICS, MINOR BIOSTATISTICS

Chicago, IL

May 2012

Experience

Ernst & Plath Labs

GRADUATE STUDENT RESEARCH

Los Angeles, CA

Aug. 2014 - Present

- Applying genomic (ATAC-seq, ChIP-seq, bulk RNA and single cell-seq) and computational methods (trajectory inference, gene regulatory network reconstruction, imputation, ChromHMM modeling) to diverse genomic data, to decipher the changes that occur during cell-fate transition using somatic to induced-pluripotent stem cell conversion as a platform
- Developing a computational method that utilizes a stacked ensemble of regression models to predict single cell chromatin landscape using an atlas of population-level data, alleviating the need for additional assays
- Writing R packages ([JubRI](#) and [PubMedScrapeR](#)) to perform literature-based NLP topic modeling of genes using latent Dirichlet allocation (LDA)
- Responsible for the management of 50Tb+ data warehouse hosted on HPC and Cloud services
- Deploying data processing and analysis pipelines for single cell RNA-seq being used as a standard in our labs and among collaborators

Genentech

COMPUTATIONAL BIOLOGY RESEARCH INTERN

South San Francisco, CA

June 2018 - Sept. 2018

- Performed integrated computational analyses from various genomic data platforms (RNA-seq, ATAC-seq, and scRNA-seq) to inform drug discovery pipelines of the metastatic process of colorectal cancer
- Wrote production grade software packages and Shiny applications in R to pinpoint shifts in population numbers due to targeted depletion of cancer stem cells

Sample Projects

SoCal Data Open	Investigated racial bias of ridesharing services on communities of New York City; placed 3rd overall (report)
deconR	R package to deconvolve cell type profiles from bulk assays using an ensemble of regressions (github)
Movie-Recommender	A model-based movie recommendation system utilizing PySpark's Alternating Least Squares (ALS) (github)
modified-Levenshtein	An efficient dynamic programming algorithm of the Sequence-Levenshtein distance written in Cython (github)
SSFold	An algorithm for predicting the optimal RNA secondary structure written in Python (github)
oriFinder	A method to pinpoint the origin of replication for bacterial genomes (github)

Skills

Programming	R, Python, BASH; working knowledge of SQL, C++ & Apache Spark and Hadoop frameworks; experience developing R Shiny dashboards and packages
Bioinformatics	RNA-/ChIP-/DNase-/ATAC-seq, single-cell RNA-/ATAC-seq, Whole genome sequencing & promoter-capture HiC processing & analysis techniques; variant calling, batch correction, gene regulatory networks & data visualization
Machine Learning	Classification & regression modeling, hyperparameter tuning & optimization, feature selection; H2O, Keras, PyTorch, Scikit-learn & TensorFlow frameworks
Applications	☒, Git/Subversion, AWS/Cloud computing platforms; Linux/Unix; common spreadsheet & presentation software

Publications

(SELECTED OF 12)

Identification of Conserved Mechanisms Underlying the Gene Expression Changes During Reprogramming Processes

J. Langerman*, **S. Sabri***, C. Chronis, J. Ernst, K. Plath

Nature; in revision (Apr. 2020)

Developmental Trajectory of Human Skeletal Muscle Progenitor and Stem Cells across Development and from Pluripotent Stem Cells

H. Xi, J. Langerman, **S. Sabri**, P. Chien, C. S. Young, S. Younesi, M. Hicks, K. Gonzalez, W. Fujiwara, J. Marzi, S. Liebscher, M. Spencer, B. V. Handel, D. Evseenko, K. Schenke-Layland, K. Plath, A. D. Pyle

Cell Stem Cell; recently accepted (Apr. 2020)

A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation

D. Polioudakis, L. Torre-Ubieta, J. Langerman, A. Elkins, X. Shi, J. Stein, C. Vuong, S. Nichterwitz, M. Gevorgian, C. Opland, D. Lu, W. Connell, E. Ruzzo, J. Lowe, T. Hadzic, F. Hinz, **S. Sabri**, W. Lowry, M. Gerstein, K. Plath, D. Geschwind

Neuron 103.5 (Sept. 2019) pp. 785–801

Reduced MEK Inhibition Preserves Genomic Stability in Naïve Human ES Cells

B. D. Stefano, M. Ueda, **S. Sabri**, J. Brumbaugh, A. J. Huebner, A. Sahakyan, K. Clement, K. J. Clowers, A. R. Erickson, K. Shioda, S. P. Gygi, H. Gu, T. Shioda, A. Meissner, Y. Takashima, K. Plath, K. Hochedlinger

Nature Methods 15.9 (Aug. 2019) pp. 732–740

Human Naïve Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation

A. Sahakyan, R. Kim, C. Chronis, **S. Sabri**, G. Bonora, T. W. Theunissen, E. Kuoy, J. Langerman, A. T. Clark, R. Jaenisch, K. Plath

Cell Stem Cell 20.1 (Jan. 2017) pp. 87–101

Cooperative Binding of Transcription Factors Orchestrates Reprogramming

C. Chronis*, P. Fiziev*, B. Papp, S. Butz, G. Bonora, **S. Sabri**, J. Ernst, K. Plath

Cell 168.3 (Jan. 2017) 442–459.e20

Honors & Awards

PROFESSIONAL RECOGNITION

2019	Finalist , 1st Annual AI LA Life Summit	Pasadena, CA
2019	Recipient , Outstanding Mentor Award, Bruins-In-Genomics (B.I.G.) Summer Research Program	Los Angeles, CA
2019	Recipient , Travel Award, Deep learning for Science at Lawrence Berkeley National Laboratory	Berkeley, CA
2019	Finalist , Citadel & Correlation One Datathon	Berkeley, CA
2018	Finalist , Citadel & Correlation One Datathon	Pasadena, CA
2017	Recipient , Molecular Biology Institute (MBI) Outstanding Poster Presentation Award	Los Angeles, CA
2017	2nd Place , Broad Stem Cell Research Center (BSCRC) Elevator Pitch (source)	Los Angeles, CA
2017	3rd Place , Citadel & Correlation One Datathon (\$2,500 cash prize)	Pasadena, CA
2016	Recipient , Best Research Pitch Award at Cornell University's ACLS International Summer School	New York, NY

ACADEMIC FELLOWSHIPS

2019	Recipient , UCLA Dissertation Year Fellowship (\$37,272/year)	Los Angeles, CA
2016-2018	Recipient , Broad Stem Cell Research Center (BSCRC) Fellowship (\$40,000/year)	Los Angeles, CA
2017	Recipient , Philip J. Whitcome Fellowship (awarded but declined; \$29,376/year)	Los Angeles, CA
2013	Recipient , Noblis Bioinformatics Scholarship (\$9,750/year)	Baltimore, MD
2008-2012	Recipient , Trustee Academic Scholarship (\$10,000/year)	Chicago, IL

Teaching & Mentorship

Spg. 2020	Research Mentor , Computer Science PhD graduate student Jingyuan Fu	Los Angeles, CA
Sum. 2019	Research Mentor , Undergraduate students Jeremy Wang (Brown U.) & Rebecca Castillo (U. of New Mexico)	Los Angeles, CA
Win. 2017	Research Mentor , Bioinformatics PhD graduate student Harry Yang	Los Angeles, CA
Fall 2015	Research Mentor , Bioinformatics PhD graduate student Kiku Koyano	Los Angeles, CA
Fall 2015	Teaching Assistant , Intro to Bioinformatics and Genomics; student evaluation rating of 8.94/10	Los Angeles, CA
Fall 2014	Teaching Assistant , Practical Computer Concepts for Bioinformatics; student evaluation rating of 11.21/12	Baltimore, MD
Fall 2013	Teaching Assistant , Computational Genomics; student evaluation rating of 10.60/12	Baltimore, MD