

# JAMES SACCO, M.SC.

Senior Bioinformatician | Machine Learning Enthusiast



## EXPERIENCE

2024  
|  
2022

### Bioinformatics Programmer III



Laboratory for Genomics Research, University of California, San Francisco

San Francisco, CA

- **Conceptualized and designed** a CRISPR-Cas12a library of guide RNAs, with genomic, transcriptomic, and epigenomic data to ensure comprehensive coverage.
- **Investigated drivers** of CRISPR efficacy, using ML, resulting in a 15% increase in targeting accuracy.
- **Led computational analyses** of multiple CRISPR screens in rare diseases, identifying 400 gene hits.
- Co-authored manuscript on key regulators of human stem cell differentiation.
- **Collaborated with wet-lab scientists** to support 5 projects, accelerating project timelines by up to 20%.

2022  
|  
2021

### Bioinformatics Scientist



ASC Therapeutics

Milpitas, CA

- **Conducted comparative analyses** of off-target detection methods in CRISPR screens using single-cell data, employing statistical models to evaluate accuracy.
- Developed a **custom sequence alignment** pipeline for gene editing, improving alignment speed by 80%.
- Automated workflows for sequence alignment, NGS data analysis, and off-target detection, reducing processing time by 60% and enhancing collaboration with experimental teams.
- Implemented company's **first version control** processes, standardizing collaborative software development and data analysis.

View this CV online with links at  
<https://jsacco1.github.io/cv/>

## CONTACT INFO

✉ [jsacco001@gmail.com](mailto:jsacco001@gmail.com)

🔗 [jsacco1.github.io](https://jsacco1.github.io)

🐙 [github.com](https://github.com/jsacco1)

in [linkedin.com](https://www.linkedin.com/in/jsacco1)

📞 +1 (786)-473-9993

Work authorization: US citizen

For more information, please  
contact me via email.

## KEY SKILLS

**Machine learning** (scikit-learn, Keras)

**Single cell multi-omics, experimental design, and robust and reproducible analyses**

Bioinformatics databases and workflows (Snakemake, Tableau, Docker, Singularity)

Cloud data analysis (PySpark, Hail), **AWS deep learning stack** (EC2, SageMaker, S3)

High scientific rigor & willingness to teach and learn about new computational methods and biology

2020  
|  
2019

## Biomarker Data Analyst II



Genentech (Contract)

📍 South San Francisco, CA

- **Optimized clinical trials** by streamlining biomarker analysis using ML models, including regression, random forest, and Xgboost, as well as explainable AI.
- **Improved sample quality** by 10% across 49 clinical trial arms with department's first ML pipeline to study the effects of biomarker operations on the quality of cancer immunotherapy assays.
- **Upgraded data operations** for 6 data streams, including flow cytometry, biomarker operations, and oncology data warehouses, resulting in more efficient data sharing.

2019  
|  
2018

## Data Curator



Genentech (Contract)

📍 South San Francisco, CA

- Integrated RNA assays, enhancing **AI/ML module performance** efficacy by at least 10%.
- **Enabled AI/ML team** in patient stratification with robust data pipelines for static and time series data points.
- Coordinated data governance and integration between AI/ML, data management, and DevOps functions, to bring **clinical ML** from prototype to production.

2018  
|  
2017

## Data Curator, Bioinformatics Analyst



BioMarin Pharmaceutical (Contract)

📍 San Rafael, CA

- **Published author** on ARSA gene research for metachromatic leukodystrophy, adding to rare disease drug development.
- **Streamlined data analysis** and visualization using R, Python, and MATLAB, improving data analysis speed by 70%.
- **Developed workflows** for pathogenic variant identification across 60 rare disorders, resulting in more accurate prediction of disease prevalence and incidence.
- **Prototyped relational databases** with Hail for efficient genomic data storage, improving data management and accessibility.

2017  
|  
2016

## Research Associate



Pearse Lab: The Miami Project to Cure Paralysis, The Miller School of Medicine, University of Miami

📍 Miami, FL

- Identified over forty putative **conserved vertebrate phosphodiesterase proteins**, using multiple sequence alignment and genomic evidence.

## PROGRAMMING SKILLS

🐍 Python

📊 R

🐧 Linux

🔑 Git

📦 C++

📐 MATLAB

Active member of Biostars



**Biostars**  
BIOINFORMATICS EXPLAINED

I have participated in bioinformatics and software workshops and *Meetup.com* groups, in both Seattle, WA and San Francisco, CA, including:

*SF/Bay AI Developers Group*

*San Jose Data Science and AI/ML Meetup*

*Hacker Dojo*

*Seattle Sequencing*

## LANGUAGES



**English:** Native Proficiency

**Spanish:** Native Proficiency

**Dutch:** Intermediate Proficiency



## EDUCATION

- Rowan University, Graduate School of Biomedical Sciences  
Master of Biomedical Science  Stratford, NJ
- Florida International University, The Honors College  
BSc, Biological Sciences  Miami, FL



## PROFESSIONAL DEVELOPMENT

- 2022 ● Python for Data Science: Fundamentals Part I Course  
Dataquest.io  Online
- 2021 ● Introduction to Genomic Technologies  
Johns Hopkins University  Online
- 2021 ● Statistics for Genomic Data Science  
Johns Hopkins University  Online
- 2016 ● Software Carpentry for R and UNIX  
University of Miami  Miami, FL





## PUBLICATIONS


- 2023 ● Predicting disease severity in metachromatic leukodystrophy using protein activity and a patient phenotype matrix.  
*Genome Biology*. 2023 Jul 24:172.  
Trinidad M., Hong X., Froelich S., Daiker J., **Sacco J.**, Phuc Nguyen H., Campagna M., Suhr D., Suhr T., LeBowitz J.H., Gelb M.H., Clark W.T.
- 2016 ● Regulating Axonal Responses to Injury: The Intersection between Signaling Pathways Involved in Axon Myelination and The Inhibition of Axon Regeneration.  
*Frontiers in Molecular Neuroscience*. 2016 Jun 8;9:33.  
Rao S.N., Pearse D.D. **Manuscript Editor**



## CONFERENCES

- 2023 ● Single Cell Analyses  
 Cold Spring Harbor Laboratory Single Cell Analyses Meeting.  
 Cold Spring Harbor, New York

2020 • Prediction of disease severity in metachromatic leukodystrophy using measures of protein activity and a novel phenotype matrix.

 American Society of Human Genetics *American Society of Human Genetics Virtual Meeting*. [Poster]

📍 Online

Trinidad M., Hong X., **Sacco J.**, Nguyen H.P., Clark W.T., Froelich S., LeBowitz J.H., Gelb M.H.

2018 • Association of HGMD and gnomAD variants of unknown significance with prediction of disease incidence and prevalence.

BIOMARIN

*BioMarin Internal Research and Development Asilomar Conference*. [Poster]

📍 Pacific Grove, CA

**Sacco J.**, Clark W.T., Yu K., Wu K., LeBowitz J.H.



## RESEARCH PROJECTS

2019 • Characterization of glycan substrates accumulating in GM1 gangliosidosis

*Molecular Genetics and Metabolism Reports*. 2019. 21, p.100524.

Lawrence R., et al.

2018 • Utilizing ExAC to assess the hidden contribution of variants of unknown significance to Sanfilippo Type B incidence

*PLoS One*. 2018. 13(7):e0200008.

Clark W.T., et al.

2017 • Phosphodiesterase Inhibitors as a Therapeutic Approach to Neuroprotection and Repair

*International Journal of Molecular Sciences*. 2017. 18(4): 696.

Knott E.P., et al.



## ORAL COMMUNICATIONS

2022 • Data Visualization of RNA-Seq Differential Expression Analysis with DESeq2

**hhmi**

**Doudna Lab**: Howard Hughes Medical Institute, Department of Molecular and Cell Biology, University of California, Berkeley

📍 Berkeley, CA

Topics: Introduction to RNA-Seq, Standard Workflows with DESeq2, Reference-Based Assembly, Exploratory Data Analysis, Statistical Analysis, Data Visualization, Gene Set Enrichment, Variant Annotation

• Presented publication-quality RNA-Seq data summaries and representations to post-doctoral associates

## HOBBIES

Stonewall Sports kickball. San Francisco, CA. 2023.

Ultimate frisbee player. San Francisco, CA. 2018 – Present.

Rugby at *Seattle Quake RFC*. Seattle, WA. 2017.

Literature

Classical Music

Language Learning

2022

- Changes in chromatin accessibility are not concordant with transcriptional changes for single-factor perturbations



Laboratory for Genomics Research & Department

Biochemistry & Biophysics, University of California, San Francisco

📍 San Francisco, CA

- Presented now-published preprint in contrast with established research at internal Journal Club to inform future experimental design in epigenetic editing.



## TEACHING

2009

- Peer-Led Team Learning

Florida International University

📍 Miami, FL

General Biology II facilitator for two concurrent PLTL courses

Made with 🍷 using R [pagedown](#).

Last update: 2024-06-20. The most recent version of this resume is [available here](#).