JAMES SACCO

Bioinformatics Scientist
Department of Gene and Cell Therapy
ASC Therapeutics
Milpitas, CA 95035

Self-motivated **bioinformatics programming professional** with **six years** of experience in support of **rare disease and cancer therapies**. Areas of expertise include **transcriptomics**, **CRISPR gene editing**, development of robust, reproducible **data workflows** for scientific user interfaces.

An analytical thinker and quick learner, with a broad understanding of **statistical and machine learning methods**. Published author of research studies in the integration of genotype and phenotype data to **model disease progression and epidemiology**.

SELECTED EXPERIENCE

Present | 2021

Bioinformatics Scientist

ASC Therapeutics (Contract)

Milpitas, CA

- Quantify and characterize **CRISPR gene modifications**, with both custom and open-source tools for genomic data analysis.
- Query, retrieve, and integrate data from public genomics databases, to enhance on-target gene editing with sequence alignment protocols.
- Examine concordance and sensitivity of five open-source CRISPR off-target detection methods (**CALITAS**, **CHANGE-Seq**, **CRISPResso**, **GUIDE-Seq**, and **Cas-OFFinder**).
- Streamline high-performance analyses for bench scientists by constructing **custom data pipelines**, **command line tools**, and **user interfaces**.
- Generate reproducible reports for publication-quality manuscripts.

2020 | 2019

Biomarker Data Analyst II

Genentech (Contract)

South San Francisco, CA

- Established department-first **machine learning pipeline** to study effects of biomarker operations on quality of cancer immunotherapy assays.
- Upgraded **data operations** for six data streams, including flow cytometry, biomarker operations, and oncology data warehouses.
- Monitored and solved sample data quality issues for two cancer immunotherapy trial arms.

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

CONTACT INFO

jsacco001@gmail.com
jsacco1.github.io
github.com
linkedin.com
+1 (786)-473-9993

For more information, please contact me via email

KEY SKILLS

Expert programming in **Python**, **R**, and **Linux shell scripting**

Extensive experience with **statistical, NGS data analyses** and public genomic repositories

Proficiency with *in silico* **CRISPR gene editing** modules

Machine learning and applied statistics

Data engineering & pipeline development (AWS, Airflow, PySpark)

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

2019 | 2018

Data Curator

Genentech (Contract)

South San Francisco, CA

- Collaborated with AI engineers to create an ETL pipeline for **multi-modal survival prediction and patient stratification**, by using Python ML and R to integrate **gene panel** and **RNA-Seq data**.
- Designed ETL data pipeline for integration of **real-world EHRs** into **deep learning** module, using Python, SQL, and Apache Spark.
- Organized and managed close coordination of Artificial Intelligence, data management, and DevOps functions, to bring **clinical machine learning** from prototype to production.

2018 | 2017

Data Curator, Bioinformatics Analyst

BioMarin Pharmaceutical (Contract)

San Rafael, CA

- Established an integrated genetic and curated literature workflow to predict incidence and prevalence rates of **over sixty rare, genetic disorders,** with **statistical and data visualization R packages (ggplot2)** and **MATLAB**.
- Upgraded and tested **RNA-Seq data pipeline (STAR2, SAMtools, Bioconductor)** to discern disease contribution of **rare variants** to neural disease.
- Spearheaded development of three relational databases derived from **Hail population genomics platform**, in collaboration with software engineer.
- Co-authored manuscript submitted to peer-reviewed publication, on **prediction of disease severity** in metachromatic leukodystrophy.
- Developed **allele-specific association models and GWAS analysis software** for two later published manuscripts.

2017 | 2016

Research Associate

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

Miami, FL

- Mentored research associates in developing bioinformatic skills, specifically in **multiple sequence alignment** and **standard molecular biological protocols** (restriction enzyme digestion, gel electrophoresis, spectrophotometry).
- Identified over forty putative **conserved vertebrate phosphodiesterase proteins**, by using multiple sequence alignment and genomic evidence.

PROGRAMMING SKILLS

- Python ★★★★★
- R ★★★★
- Linux ★★★★
- Git ★★★★★
- GitHub ★★★★★
- Markdown★★★★★
- SQL ★★★
- Dash ★★★
- C++ ★★★
- MATLAB ★★★

EDUCATION

Rowan University, Graduate School of Biomedical Sciences

Master of Biomedical Science (Distinction) — Stratford, NJ

"Hippocampal long-term potentiation in neurodegenerative disorders"

Related coursework:

- Biochemistry
- Molecular Biology
- Cell Biology
- Microbiology
- Pharmacology

Florida International University, The Honors College

BSc, Biological Sciences — Miami, FL

PROFESSIONAL DEVELOPMENT

2022 — **Bulk and Single Cell Transcriptomics Training Program**
Pine Biotech, Tauber Bioinformatics Research Center — Online

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

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**

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

Spanish: Bilingual Proficiency

CONFERENCE PAPERS

2020

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

American Society of Human Genetics Virtual Meeting. [Poster] **[Forthcoming on bioRxiv.org]**

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 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.

EXTRA-CURRICULAR

Avid supporter of the San Francisco Symphony and San Francisco Conservatory of Music

Ultimate frisbee at *Big Gay Frisbee.* San Francisco, CA. 2018 – Present.

Salsa dance at *In Lak'ech Dance Academy.* Oakland, CA. 2019.

Rugby at *Seattle Quake RFC*. Seattle, WA. 2016 – 2017.

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

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: 2022-02-16.The most recent version of this resume is available here.