**JOHN J. CAPERELLA Jr., MS**

13 Crotona Avenue, Harrison, NY 10528 • 914-484-0257 • [jcaperella@gmail.com](mailto:jcaperella@gmail.com)

<https://github.com/jcaperella29>

**Bioinformatician & Data Analyst**

**MS in Bioinformatics with 2+ years computational research experience**

**EDUCATION**

**Brandeis University,** Waltham, MA May 2024

MS, Master of Science in Bioinformatics, GPA: 4.0

**Pace University**, **Dyson College of Arts and Sciences,** Pleasantville, NY

Bachelor of Science in Biochemistry, cum laude; Mathematics Minor May 2019

**Honors/Awards:** Biochemistry Excellence Award, Induction into the Pace University Dyson Society of Fellows, Dean’s Scholarship, Dean’s List, Multiple Outside Presentation Awards

**SKILLS**

**Coding languages /deployment/ data processing tools:** Proficient in R (statistical packages, machine learning, Bioconductor, Shiny); Python (machine learning, Biopython); Bash; Nextflow, Docker, CWL, Google Cloud Platform. Knowledge of SQL and web development.

**Genomics/Genetics software and methods:** Experience with analytical tools including genome-wide association studies with PLINK and R, burden tests, meta-analysis with METAL, command line FASTQ processing tools, variant calling, ATAC peak analysis and De novo genome assembly.

**Transcriptomics:** Some experience in the development and usage of scripts for processing microarray, RNA -seq and ScRNA-seq data, preparing counts matrices as well as performing meta-analysis on RNA-seq studies in R as well as weighted correlation network analysis.

**Machine learning and dimension reduction:** Development and usage of Linear Regression, Logistic Regression, Large Language models, Random Forest and Neural Network models in R and Python. Usage of hierarchical and k-means clustering and dimension reduction methods such as principal components analysis (PCA) and Uniform Manifold Approximation and Projection (UMAP).

**Structural Biology:** Experience with tools such as PyMOL, AutoDock and AlphaFold.

**Databases:** Experience finding relevant data from databases such as Genebass, Ensembl, UCSC, NCBI, Protein Data Bank, Open Targets Genetics, UniProt, Gene Ontology.

**RESEARCH & PROJECTS**

**“RNA\_SEQ\_Processing App” –** An all-in-one R Shiny web app for interactive RNA-seq analysis — no coding required. It includes differential expression, power analysis, Random Forest, PCA/UMAP, and pathway enrichment, a clean, tabbed UI with built-in visualizations and downloadable results. It is fully Docker, Singularity, and Apptainer compatible — perfect for HPC or local use.

Designed for bioinformatics teams and core facilities. Built for real-world datasets, reproducibility, and ease of use.

The GitHub repository for it can be found at <https://github.com/jcaperella29/RNA_SEQ_APP>.

**“JCAP GWAS Shiny App”** – Interactive R Shiny platform for SNP QC, GWAS, gene mapping, pathway enrichment, and machine learning-based prediction. Accepts VCF/phenotype/covariate files, supports data visualization (PCA, UMAP, Q-Q, Manhattan plots), gene annotation, and power analysis. Fully containerized (Singularity/HPC-ready). It can be found at <https://github.com/jcaperella29/GWAS_SHINY_APP>.

**“Bioinformatics Tools Hub”-**

Front-End Web Development + Scientific Tool Integration | HTML, CSS, JS | Designed and deployed a custom biotech-themed web portal to showcase RNA-Seq, ATAC-Seq, miRNA-Seq, GWAS and other sequence analysis applications. Implemented a responsive, multi-tab interface using HTML, CSS, and JavaScript for clear navigation of bioinformatics tools. Integrated animated DNA-inspired visuals and deployed using GitHub Pages for public access and open science collaboration. Enhanced accessibility for non-programming researchers by centralizing complex data workflows into a clean, web-based user interface. It can be found at <https://jcaperella29.github.io/JC_BIOINFORMATICS_HUB/>.

**RELEVANT EXPERIENCE**

**Independent Software Engineer**  2024–Present

Designed, built, and deployed a range of tools including Flask and Shiny web applications, Python command-line interfaces, and APIs. Undertaking projects focused on providing user-friendly solutions to real-world problems in both biology and business contexts. Selected work is available on my GitHub @ <https://github.com/jcaperella29>.

**Pace University**, Pleasantville, NY

***Computational Research Assistant- Department of Biology*** 2021 - 2024

* Automated zebrafish hair cell counting with Python GUI and reporting
* Used ML and AlphaFold to predict Cxcr4b receptor partners.
* Provided computational support for the lab.
* Introduced the team to text mining tools and their use in literature review.
* Supported wet lab work and zebrafish system upkeep.

**Pace University**, Pleasantville, NY

***Research Assistant- Department of Biology***  2020 -2021

* Provided lab support via zebrafish husbandry and dechorionation; RNA extraction; RT-PCR; agarose gel electrophoresis; EdU labeling and Click chemistry; confocal microscopy, and sequence analysis.

**Pace University**, New York, NY

***Biology and Chemistry Tutor*** 2019 -2020

* Provided support to undergraduates for General Biology, General Chemistry I and II, and Organic Chemistry; Clarified course materials to enhance students’ understanding; Conducted pre-exams reviews; Reviewed students’ exam results and provided reinforcement of course materials.