Feng Chen

713-366-1375 | Boston, MA (Open to Relocation) | fchen3357@gmail.com | GitHub | LinkedIn | Portfolio

SUMMARY

Bioinformatics graduate with experience in NGS data analysis (RNA-seq, ATAC-seq, ChIP-seq), reproducible workflow development, and large-scale data processing across Linux/HPC environments. Skilled in Python, R, and Nextflow for scalable bioinformatics analysis and workflow automation.

EDUCATION

Northeastern University, Boston, MA | *Master of Science in Bioinformatics* | GPA: 3.75 December 2025

University of Houston, Houston, TX | *Bachelor of Engineering in Biotechnology* May 2024

TECHNICAL PROJECTS

Multi-Assay Bulk Sequencing Pipeline (Nextflow DSL2) | GitHub October 2025 – January 2026

- Developed modular Nextflow (DSL2) pipelines for RNA-seq, ATAC-seq, and ChIP-seq data processing, including testing and validation using publicly available datasets across local and HPC (SLURM) environments
- Built end-to-end workflows for data processing, including FASTQ quality control and pipeline validation across multi-sample sequencing datasets
- Automated multi-sample NGS data processing, improving reproducibility and reducing manual steps
- Designed scalable and modular pipeline workflows to support efficient analysis of large sequencing datasets

ATAC-seq and RNA-seq Analysis of Triple-Negative Breast Cancer | GitHub January 2026 – March 2026

- Performed RNA-seq differential expression analysis using DESeq2, including normalization and statistical testing, and integrated results with ATAC-seq data to identify regulatory patterns in TNBC
- Identified 1,435 differentially accessible regions (FDR < 0.05), highlighting increased chromatin accessibility in tumor samples
- Integrated multi-omics datasets to identify relationships between chromatin accessibility and gene expression
- Conducted enrichment analysis (GO, KEGG, GSEA) to identify hormone signaling and immune pathways
- Generated and customized data visualizations (e.g., PCA, volcano plots) to support data interpretation and communication of biological findings

Breast Cancer Classification using k-Nearest-Neighbors | GitHub September 2025 – October 2025

- Built and evaluated a machine learning (k-NN) model to classify breast cancer samples (benign vs. malignant)
- Performed data preprocessing and model validation using cross-validation and ROC, achieving >95% accuracy

EXPERIENCE

University of Houston, Houston, TX | *Lab Technician* December 2023 – August 2024

- Performed laboratory experiments and supported data collection workflows in a research environment
- Maintained detailed experimental records to support reproducibility and data integrity
- Collaborated with cross-functional research teams to support experimental design and data collection workflows

GenviewDx, Houston, TX | *Medical Records Clerk* January 2024 – March 2024

- Managed and updated patient records in the electronic health record (EHR) system, ensuring accurate data entry and compliance with HIPAA regulations
- Organized and digitized 500+ patient files, reducing retrieval time by 30% and improving access for medical staff

TECHNICAL SKILLS

Programming: Python, R, SQL, Bash

Systems: Linux, HPC (SLURM), UNIX command line

NGS & Pipelines: Nextflow (DSL2), RNA-seq, ATAC-seq, ChIP-seq, FASTQ processing, alignment, quality control

Tools: FastQC, STAR, Bowtie2, SAMtools, featureCounts, MultiQC

Analysis: DESeq2, DiffBind, clusterProfiler, fgsea, ggplot2