

Bioinformatics Profile

Bioinformatician with two MSc degrees, with the second in Bioinformatics (both with **distinction**), over four years of experience in R and proficiency in Python. Known for developing [innovative bioinformatics pipelines](#) and [interactive tools](#) (e.g., [Shiny apps](#)) by combining **strong programming skills** with **creativity** and **problem-solving**. While my expertise to date has mainly focused on **transcriptomics** (bulk and **single-cell RNA-Seq**), I am confident in applying these techniques to other omics data types. Experienced in extracting **meaningful biological insights** from complex datasets, including **pathway interpretation**, functional annotation, and **molecular profiling**. I also have evolving experience with **AI/ML/DL**, including neural networks, and am keen to expand my programming and analytical expertise to develop innovative AI-driven solutions. Committed to translational research in **cancer** and other **disease** areas, and able to **adapt quickly**, **communicate clearly**, meeting **deadlines**, and balancing **multiple projects** while acting as a **reliable point of contact** for resolving **technical challenges** and **work effectively under pressure** in **multidisciplinary teams**. Created and maintain a [personal portfolio website](#) showcasing my tools, code, and research outputs.

Areas of Expertise

- Programming (R / Python)
- Biological data interpretation
- Functional enrichment & pathway
- ML & AI (evolving expertise)
- Transcriptomics (bulk/ single cell)
- Bioinformatics pipeline development
- Shiny Web applications
- Web design with HTML/CSS
- Cancer and disease research
- Data Analysis & Visualisation
- Data Integration
- Linux & GitHub

Career Experience

[Shiny Web Application](#) | *Independent Bioinformatics Project*

01/2025 – Present

- Designed & developed a Shiny app (~2,000 lines R) to support biomarker discovery from single-cell RNA-seq data.
- Enables cell-type-specific exploration by integrating metadata, streamlining hypothesis generation for drug discovery.
- Supports 10X input, metadata, EDA, QC, normalization; identifies HVGs; performs GSEA/ORA for GO terms and KEGG.
- Tested on kidney cancer data (ICB-treated); designed for users without programming or bioinformatics expertise.
- Presented at Norwich Single-Cell & Spatial Symposium 2025 ([abstract published](#)).

[Biologically Informed HVG Selection Pipeline](#) | *Independent Bioinformatics Project*

06/2024 – 12/2024

Developed an end-to-end R-based bioinformatics pipeline to generate a biologically refined HVG list from scRNA-seq data for training deep learning models. Identifies an optimal HVG threshold and intersects it with GO/pathway-informed gene sets to produce interpretable, context-relevant features.

- **Step 1:** Filtering input 10X-format scRNA-seq data by cell type to generate multiple HVG sets of varying sizes.
 - **Step 2:** Evaluating biological relevance of each HVG set by calculating overlap with immune-related gene sets derived from GO terms and pathways (via biomaRt and msigdb), identifying the point at which biological signal plateaus.
 - **Step 3:** Producing a final, biologically curated HVG list by intersecting the selected HVG set with reference genes.
- This project was motivated by the need to select biologically meaningful genes for model training and outcome prediction.

Collaborative Research Project (R-based) | Teesside University

10/2023 – 06/2024

Kidney Cancer scRNA-seq Analysis with Prof. Claudio Angione's AI Research Group

- Collaborated with AI researchers to extract biological meaning from scRNA-seq data & support model development.
- Developed a full Seurat pipeline: cell filtering, HVG selection, clustering, annotation, DEA, and functional analysis
- Conducted end-to-end scRNA-seq analysis to extract complex biological insights from ICB-treated kidney cancer data
- Interpreted AI model outputs for biological relevance and alignment with known pathways and immune mechanisms
- Supported a multidisciplinary team with clear analyses and visualizations to communicate complex biological insights
- Engaged in weekly team meetings, providing analysis updates and contributing to interdisciplinary discussions

[RNA-Seq Analysis of AD: DESeq2 Differential Gene Expression](#) | *Independent Bioinformatics Project*

02/2023 – 09/2023

- Analysed GSE157194 (166 Atopic Dermatitis samples) using R (DESeq2): lesional vs. non-lesional skin.
- Conducted exploratory data analysis including PCA, boxplots, histograms, heatmaps, and other visualizations.
- Performed differential expression & functional enrichment analysis (PPI, ORA & GSEA for GO terms and KEGG pathways).
- Published visuals and methods on my personal website (full project update in progress).

[Website Design & Development \(My Personal Portfolio Website\)](#) | *Self-Directed Project*

07/2022 – 01/2023

- Independently designed and developed a personal portfolio website from scratch using HTML and CSS coding.
- Created a multi-page, interactive platform to showcase my bioinformatics projects and communicate complex results.
- Demonstrated visual & written communication skills through well-structured documentation & intuitive structure.

- Applied creativity and problem-solving to design custom layouts and engaging interactive components.

Multi-Phase Transcriptomic Analysis of Breast Cancer | Dissertation Research Project | Teesside University 01/2022 – 06/2022

- Developed a 3-step pipeline using custom R functions with GEOquery, limma, GeneMeta, fgsea/msigdb, ggplot2/plotly.
- Implemented grade-based comparisons using limma and meta-analysis to identify aggressiveness-related DEGs.
- **Phase 1:** DEG analysis on one dataset comparing tumor aggressiveness via grade vs. subtype; grade prioritized.
- **Phase 2:** Integrated 4 datasets using grade-based DEG analysis via limma package and custom R functions.
- **Phase 3:** Conducted meta-analysis with REM model to integrate DEGs, followed by pathway and functional enrichment.

Meta-Analysis Study Selection & Classification | Internship Project | Teesside University 09/2021 – 12/2021

- Queried PubMed with defined keywords; screened & shortlisted breast cancer transcriptomic studies for meta-analysis. Selected 4 eligible pre-treatment datasets & based on literature, identified grade & subtype as aggressiveness indicators.
- Defined classification groups (e.g. Grade 1 vs. 3, Lum A/Normal-like vs. Basal-like/Lum B/HER2) for downstream analysis.
- Performed initial DEG and functional analyses using G02R and NetworkAnalyst platforms to assess group-level differences.
- Established classification approach and prepared datasets for integration & DEG analysis in dissertation research project.

Research Assistant, Pasteur Institute of Iran, Tehran, Iran

05/2018–09/2020

- Supported whole exome sequencing (WES) and disease projects with literature reviews and weekly research updates.
- Assisted in preparing reports, posters, presentations, manuscript drafts for the team, and delivering workshops.
- Co-authored a publication on NF1-related mutation analysis and contributed to VSD and cancer reviews.

National Institute of Genetic Engineering and Biotechnology (NIGEB), Iran (Training & Transition)

08/2016–05/2018

- Completed multiple workshops in NGS, multi-omics including transcriptomics, genomics, and epigenomics, etc.
- Built core skills through self-study and literature review during transition to human genomics.

Education

MSc Bioinformatics (with advanced practice) (Distinction 73/100)

Teesside University, Middlesbrough, UK

MSc Agricultural engineering (GPA: 89.15/100)

SRB University, Tehran, Iran

BSc Plant Products Engineering (GPA: 76.35/100)

Karaj University, Karaj, Iran

Bioinformatics Training

Intensive Course in Drug Discovery, Bioprocessing & Immunological Techniques (11 days) | Real-time PCR (16 hours) | Bioinformatics & NGS Methodologies, Bio-Linux, SNP Discovery & Analysis, RNA-Seq Data Analysis, Predict New Transcript/Isoforms, Clinical Applications & Diagnostics (24 hours) | Gene Cloning (RNA Extraction, cDNA Synthesis, Primer Designing, PCR, DNA Extraction, Cloning, Gene Transformation, Electrophoresis, Protein Expression & Purification, SDS-PAGE, Western blot) (32 hours) | Next-generation Sequencing Data Analysis: RNA Seq and Small RNA Seq (2 days) | Meta-analysis of High-throughput Sequencing data (2 days) | Computational Systems Biology: Network Analysis, Gene Ontology and Promoter analysis (2days) | Next-generation Sequencing Data Analysis: Metagenomics (2 days) | Epigenomic Data Analysis (ChIP-seq & Genome-wide Methylation) (2 days) | Data Mining and Machine Learning in Bioinformatics (2 days) | NGS Analysis and Clinical Report Writing (1 day) | Introduction to Genomic Technologies (Johns Hopkins University; Coursera)

Programming & Analytical Tools

- **R:** dplyr, ggplot2, Seurat, clusterProfiler, limma, tidyr, readr, GEOquery, plotly, data.table, ComplexHeatmap, sva, shiny, Matrix, msigdb, org.Hs.eg.db, biomaRt, fgsea, caret, tidyverse, enrichplot, RColorBrewer, ggpubr, shinyjs, circlize, ggvenn, wordcloud, reshape2, vsn, GeneMeta, rgl
- **Python:** tensorflow, scikit-learn, shap, pandas, numpy, matplotlib, seaborn

Publications

1. Mohajeri, O. (M.R.) (2025). **An Interactive Shiny App for Biomarker Identification: Cell-Type-Specific HVGs & Functional Analysis in Single-cell RNA-Sequencing**. Norwich Single-Cell & Spatial Symposium Abstracts, p. 28. Earlham Institute.
2. E.S. Rahmani, H. Azarpara, M. F. Abazari, **M.R. Mohajeri**, M. Nasimi, R. Ghorbani, A. Azizpour & H. Rahimi. 2020. **Novel Mutation C.7348C>T in NF1 Gene Identified by Whole-Exome Sequencing in Patient with Overlapping Clinical Symptoms of Neurofibromatosis Type 1 and Bannayan–Riley–Ruvalcaba Syndrome**, Cytol Genet. vol. 54, no. 4, pp. 353–362.
3. **Mohajeri, M.R.**, Ghamari Zare, A., Naderi Shahab M. A., Kalateh Jari S., 2014. Seed Germination of Liliun ledebourii (Baker) Boiss after Cryopreservation. Journal of Rangeland Science, 4 (4): 279- 286.