

MENNA ARAFAT

Mansoura, Egypt | +20 1094107444

Portfolio: <https://menna-arafat.github.io/research/>

GitHub: <https://github.com/Menna-Arafat>

Email: mennaarafat.md@gmail.com



EDUCATION

Bachelor of Medicine and Surgery Degree (M.B.B. Ch),
Faculty of Medicine, Mansoura university, GPA [81.3%, 3.5/4].

2014 Sep –
2021 Feb

WORK EXPERIENCE

Bioinformatician & Statistician at El-Beltagy Therapeutics Company

2024 May –
2025 Apr

- Modeled temporal exosome proteomics from gingival mesenchymal stem cell differentiation; revealed time-aligned regulatory programs, key TF–pathway interactions, and the cell-type contributions that together explain enhanced regenerative responses under different treatment conditions.
- Contributed to grant applications, study design, and data-driven research proposals.

Bioinformatics Research Assistant at 57357 Children Cancer Hospital

2023 May –
2024 Apr

- Applied End-to-end metabolomic analysis of colorectal cancer and used weighted network modeling to identify metabolite modules associated with tumor phenotype, published in Amino Acid (Springer Nature).
- Modeled metabolomic and proteomic profiles to define subtype-specific network modules in rhabdomyosarcoma, published in Cells.

Remote Collaborations

Current

- Contributed to projects across genomics, transcriptomics, epigenomics, proteomics, and metabolomics, using both bulk and single-cell datasets.
- Developing frameworks using multi-omics integration and machine learning that characterize the molecular architecture of disease and identify therapeutic vulnerabilities.
- Implemented advanced analysis for pathway crosstalk inference, drug target prioritization, drug-pathway sensitivity analysis, and drug synergy prediction.

PUBLICATIONS

1. [Metabolomic Analysis of Gut Metabolites in Colorectal Cancer Patients \(CRC\)](#)

Performed preprocessing of colorectal cancer (CRC) metabolomic data, including PQN normalization and differential analysis, and applied weighted network analysis (WGCNA) to identify co-expressed metabolite modules associated with disease status. Used PCA and PLS-DA to prioritize discriminative features underlying group separation, highlighting biomarkers and network-level metabolic dysregulation.

2. [Integrative Multi-Omics Profiling of Rhabdomyosarcoma Subtypes Reveals Distinct Molecular Pathways and Biomarker Signatures](#)

Conducted the systems-level analysis that mapped subtype-specific molecular features in rhabdomyosarcoma. I applied weighted network analysis (WGCNA) to define co-expressed protein modules and used eigengene-based Wilcoxon testing to identify modules that significantly distinguish between clinical subtypes. Integrated hub proteins with DEPs into PPI networks and centrality analysis to reveal key molecular drivers.

3. Integrative Transcriptomic Profiling of Plasma Extracellular Vesicles Reveals miRNA Signatures for Breast Cancer Subtype Classification

Developed a framework that integrates and harmonizes extracellular vesicle miRNA datasets to identify reproducible subtype-relevant signatures in breast cancer, combining meta-analysis with MINT sPLS-DA for feature selection and classification. Constructed miRNA–target interaction networks and performed pathway enrichment analysis to characterize the regulatory processes distinguishing each subtype.

4. CrosstalkX: Mutual Information–Based Pathway Interaction Inference

Developed CrosstalkX, an R package that uses mutual information to detect nonlinear interactions among signaling pathways and prioritize core regulatory components. The method identifies pathway–pathway and pathway–TF interactions using biologically informed constraints and permutation testing for robust inference. It allows to quantify pathway–pathway and pathway–TF dependencies in complex omics data.

5. A Multi-Omics and Mutual Information Framework to Decode Therapeutic Vulnerabilities in TNBC

Developed a framework for identifying therapeutic vulnerabilities in triple-negative breast cancer assembling six omics layers from TCGA. By combining multi-omics factor analysis with mutual information–based pathway crosstalk modeling, I prioritized key genes, transcription factors, and coordinated pathway activity associated with aggressive tumor behavior. I then linked these regulatory and pathway features to candidate compounds using CMAP, Drug Repurposing Hub, and DepMap essentiality data. Finally, I applied Macau to relate pathway activity to drug sensitivity profiles, enabling the identification of drug–target associations tailored to TNBC-specific molecular states.

6. Enhanced Bone Regeneration in Rat Gingival Mesenchymal Stem Cells: A Comparative Study of Inductive Media, Polycaprolactone Nano Scaffold, and Low-Level Laser Therapy

Analyzed temporal proteomic data in rat gingival mesenchymal stem cells using MEFISTO for time-resolved multi-omics patterns, decoupleR to infer key transcriptional regulators, pathway crosstalk analysis to identify regulatory interactions. Using CIBERSORT, I linked these molecular patterns to specific cell populations, revealing coordinated TF–pathway interactions and cell-type contributions underlying enhanced regenerative responses under different treatment conditions.

7. Proteomic Profiling of Pancreatic Cancer Patients: Time Series Study Before and After Whipple Surgery

Analyzed longitudinal proteomic profiles from pancreatic cancer patients undergoing Whipple surgery by identifying recovery-associated co-expression modules with WGCNA and modeling their temporal

trajectories using generalized additive models. Integrating hub proteins into PPI networks and applying clustering and enrichment analyses revealed coordinated molecular programs linked to stress responses, tissue repair, and long-term glycemic stability.

SKILLS

Advanced R (WGCNA, PCA/PLS-DA, GAMs), **Python**, **Bash** scripting, and Linux for high-throughput bioinformatics workflows; **multi-omics** analysis and **integration** (metabolomics, proteomics, transcriptomics, genomics, methylome) using MOFA2 and MINT-sPLS-DA; time-series modeling (GAM, GLMM), single-sample and **single-cell** analyses (CIBERSORT, niche detection, CCC, lineage analysis); **network analysis** including weighted network analysis, centrality measures, modularity clustering, mutual information-based crosstalk, and functional enrichment (clusterProfiler, STRING, KEGG); downstream integration with **public databases** (TCGA, cBioPortal, CMAP, LINCS L1000, DepMap); advanced statistical modeling (including regression, confounder detection, and effect modification); **machine learning** for supervised and unsupervised analyses; reproducible reporting with R Markdown and bookdown.

COURSES

- Deep Learning Architecture for Single-Cell Genomics Data– Abdelrahman Mahmoud.
- Integrative Systems Biology – Mohamed Maysara (Nile University).
- CS224W: Machine Learning with Graphs – Jure Leskovec.
- Applied Computational Genomics – Aaron Quinlan.
- Advanced Bioconductor – Kasper Hansen.
- Statistics for Genomics – Rafael Irizarry.
- Statistics and R Programming – Mike Marin.
- Systems Biology and Systems Medicine – Uri Alon.
- Design and Interpretation of Clinical Trials – Coursera.
- Writing in the Sciences – Kristin Sainani (Coursera).

VOLUNTEERING

Coordinator of Foundation Year Conference- Pathways after graduation held in Mansura University founded by European council for continuing medical Education.

Campaign assistant for raising awareness of covid19 and administering vaccines to patients in MUH.

LANGUAGES

Arabic (Mother Language) & English (IELTS Score: 7)

REFERENCES

- **Nouran Al-Shehaby**, Junior Scientist, Basic Research Unit, Children's Cancer Hospital 57357, Egypt, +201278629538, nouran.alshehaby@57357.org
- **Mustafa Nematallah**, Professor of Molecular Genetics and Biochemistry, Mansoura Faculty of Medicine, Egypt, +201020751759, mneamatallah@mans.edu.eg

PORTFOLIO

<https://mennaa-arafat.github.io/research/>