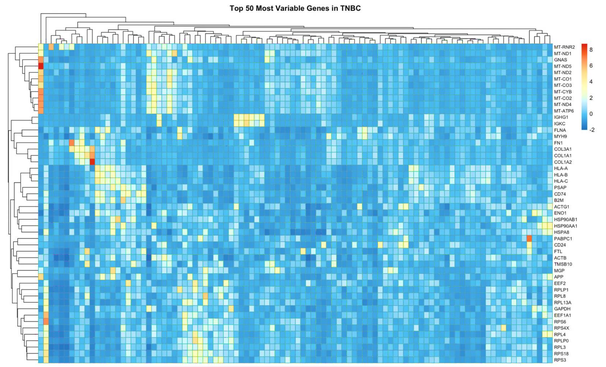
Identifying GABRA3 and MKI67 as Key Proliferation-Associated Markers in Triple-Negative Breast Cancer Using TCGA RNA-seq Data

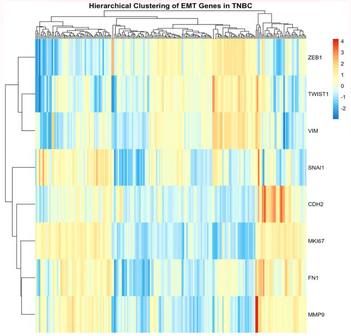
Omnia Abdelrahman, MPH[1] · Yasmeen Ahmed, PharmD, MS[2] · Lukas Weber, PhD[1]
[1]Boston University School of Public Health [2]Saint Joseph's University

Introduction

- Triple-negative breast cancer (TNBC) is an aggressive subtype with no targeted therapies.
- Recent studies suggest GABA receptors may promote tumor growth and invasion.
- Wet-lab data showed reduced EMT and invasion after GABA receptor knockdown, but the specific subunit and downstream targets remain unclear.
- This study uses TCGA RNA-seq data to identify key players in this pathway.



Unsupervised Clustering of the top 500 most variable genes in TNBC revealed distinct transcriptional patterns. Early emergence of proliferation-related genes like FN1 and MKI67 suggested a potential regulatory role for upstream pathways.



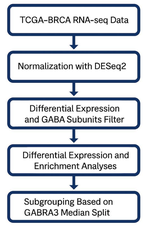
EMT Marker Heatmap reveals molecular heterogeneity in TNBC, suggesting that specific upstream regulators (ex. GABA receptor subunits) may drive EMT in certain tumors.

Methodology

Data source: TCGA-BRCA RNA-seq (HTSeq counts)

Tools: TCGAbiolinks, DESeq2, biomaRt, fgsea, pheatmap, Naive Bayes classifier

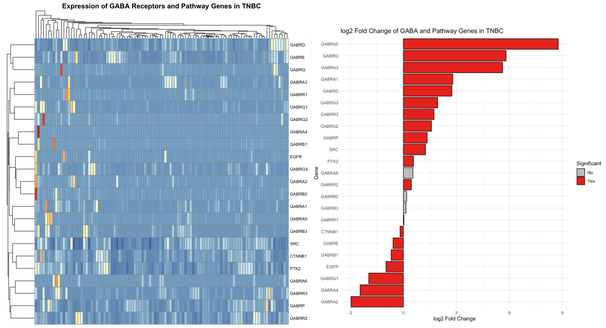
Key steps:
- TNBC samples were identified from the TCGA-BRCA dataset as those falling below the 20th percentile of ESR1, PGR, and ERBB2 expression, reflecting triple-negative receptor status.
- Normalization with DESeq2
- EMT markers and GABA subunit filtering
- Differential expression and enrichment analyses
- Subgrouping based on GABRA3 median split
- All analyses were performed using R (version 4.2) and Bioconductor packages including TCGAbiolinks, DESeq2, biomaRt, and fgsea.



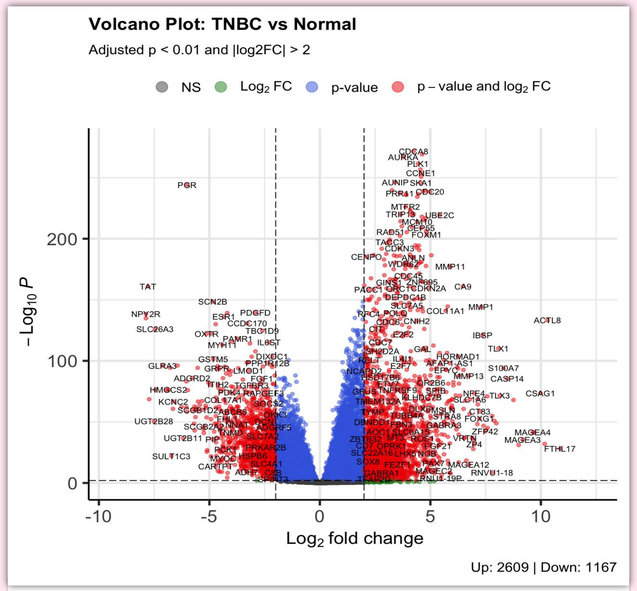
Analysis Pipeline: TCGA RNA-seq data were processed through normalization, gene filtering, differential expression, and subgrouping based on GABRA3 expression.

Principal Component Analysis of the top 500 most variable genes shows clear separation between TNBC and normal samples.

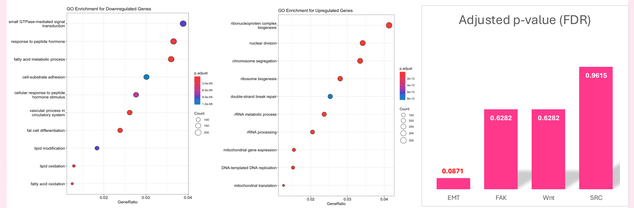
Results



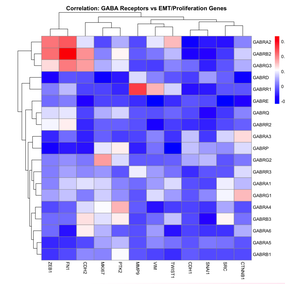
GABA Receptors and Pathway Gene Expression in TNBC
GABRA3, GABRA5, and GABRQ were significantly upregulated in TNBC. Co-expression with genes such as SRC and PTK2 suggests a role in tumor proliferation and migration.
(Left) Heatmap of GABA receptors and key pathway genes across TNBC samples.
(Right) Log2 fold changes for each gene, grouped by statistical significance.

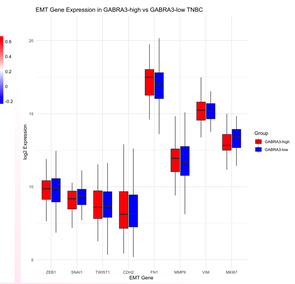


GABRA3 and MKI67 Among Top Upregulated Genes in TNBC"
Genome-wide differential expression analysis revealed significant upregulation of GABRA3, MKI67, and EMT-related genes in TNBC.



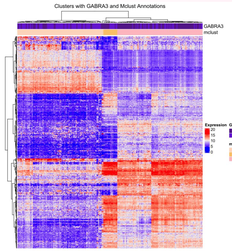
Functional Enrichment Analysis Reveals EMT Activation in TNBC.
Gene Ontology (GO) enrichment of upregulated genes in TNBC highlighted processes related to cell cycle and DNA repair, while downregulated genes were enriched in lipid metabolism and adhesion pathways (left, center). EMT showed moderate enrichment (NES = 1.53, FDR = 0.0871) among curated gene sets, as visualized by adjusted p-values across selected pathways (right).

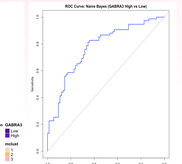




Correlation Heatmap between GABA Subunits and EMT/Proliferation GenesGABRA2 correlated with ZEB1 and FN1; GABRR1 with MKI67; GABRD with MMP9.

Boxplots show a non-significant trend toward higher EMT gene expression in GABRA3-high samples. MKI67 showed the strongest difference (p = 0.08), suggesting a possible link to proliferation.





GABRA3 Defines a Distinct and Predictable Molecular Subgroup in TNBC

A **Naive Bayes classifier** trained on the top 20 differentially expressed genes achieved a balanced accuracy of 77.8%, with an AUC of 0.772. The strong alignment between unsupervised clustering and supervised classification supports the role of GABRA3 as a biologically informative and computationally identifiable feature in TNBC stratification.

Conclusion

Main Finding
GABRA3 and MKI67 were among the most significantly upregulated genes in TNBC.
Biological Relevance
GABRA3-high tumors showed higher expression of EMT and proliferation markers, suggesting its role in driving aggressive tumor behavior.
Functional Validation
EMT pathway enrichment and correlation analysis supported GABRA3's involvement in mesenchymal activation.
Predictive Power
ROC analysis confirmed that GABRA3 expression defines a distinct molecular subgroup within TNBC (AUC ≈ 0.81).

Acknowledgements

This project was completed as part of **BS831 A1: Genomics Data Mining and Statistics (Spring 2025)** at Boston University School of Public Health. I would like to express my sincere gratitude to **Prof. Lukas Weber** for his continuous mentorship, insightful feedback, and encouragement throughout this project. lmweber@bu.edu
I am also deeply thankful to **Dr. Yasmeen Ahmed** for her thoughtful questions and generous support during the development of this work, her perspective helped me refine my analytical focus and strengthen the final product. yasmeen.ahmed@sju.edu