Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

GPNMB-Mediated Immune Modulation in Triple-Negative Breast Cancer: Transcriptomic Analysis of GPNMB Knockdown in TAM Co-Cultures Reveals Therapeutic Targeting Opportunities

Ranit Ghosh
Jadavpur University
Arghya Adhikary
arghyaa.lifesc@jadavpuruniversity.in
Jadavpur University

Research Article

Keywords: Triple-negative breast cancer (TNBC), Glycoprotein non-metastatic melanoma protein B (GP-NMB), tumor-associated macrophages (TAMs), transcriptomics, RNA-seq, GSEA, immunotherapy

Posted Date: March 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8976503/v3

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

GPNMB-Mediated Immune Modulation in Triple-Negative Breast Cancer: Transcriptomic Analysis of GPNMB Knockdown in TAM Co-Cultures Reveals Therapeutic Targeting Opportunities

Ranit Ghosh, Arghya Adhikary*

Department of Life Sciences & Biotechnology,
Jadavpur University, Kolkata, India
ranitg.lsbt.pg@jadavpuruniversity.in

March 3, 2026

Keywords: Triple-negative breast cancer (TNBC), Glycoprotein non-metastatic melanoma protein B (GPNMB), tumor-associated macrophages (TAMs), transcriptomics, RNA-seq, GSEA, immunotherapy

Abstract

Triple Negative Breast Cancer (TNBC) can be regarded as one of the most aggressive and therapeutically challenging forms of cancer. It is responsible for 10-15% of all breast cancer cases worldwide along with significantly low success rate in prognosis when compared to hormone receptor positive (ER, PR, HER2) tumors. The immunosuppressive tumor microenvironment (TME) characterized by abundant tumor associated macrophages (TAMs) with M2-like pro tumoral phenotypes, plays a critical role in TNBC progression and also in therapy failure. Glycoprotein non-metastatic melanoma protein B(GPNMB) has emerged as one of the key immunomodulatory factors with the ability to polarize macrophages toward immunosuppressive states and facilitating immune evasion. This analytical study presents a comprehensive transcriptomic analysis of GPNMB knockdown (siGPNMB1 and siGPNMB2) in co-cultured TNBC cancer cells with THP1 derived macrophages when compared to control conditions. Using RNAseq differential gene expression (DGE) analysis with DESeq2 and pathway enrichment analysis through Gene Set Enrichment Analysis (GSEA), we were able to identify 232 genes which show significant differential expression. GPNMB knockdown resulted in substantial downregulation of pro-tumoral inflammatory programs including IL6-JAK-STAT3 signalling, TGFβ signalling, and epithelial mesenchymal transition (EMT), while upregulating oxidative phosphorylation and MYC target pathways. Our findings demonstrate that GPNMB functions as a critical orchestrator of macrophage mediated immune suppression in TNBC, making it an attractive therapeutic target for re-educating the TME and enhancing anti tumor immunity.

1. Introduction

1.1 Triple-Negative Breast Cancer: Clinical Significance and Therapeutic Challenges

By definition, TNBC lacks the three receptors ER (Estrogen Receptor), PR (Progesterone Receptor) and HER 2(Human Epidermal Growth Factor Receptor) [1]. This Intrinsic immunophenotypic classification causes serious clinical implications as described later. Due to this triple negative condition, TNBC tumors are inherently resistant to endocrine therapies or HER2 targeted therapeutic agents [1, 2, 3]. As a result, chemotherapy remains the primary mode of treatment although often turning out to be aggressive [4, 5]. Even with modern innovations like neoadjuvant chemotherapy, surgery and adjuvant therapies, the rate of survival is greatly compromised in TNBC patients compared to non-TNBC subtypes of breast cancer, with mortality rate being 40% within the first 5 years of diagnosis and 75% within 3 months of recurrence [6, 7]. While the 5 year survival rate for luminal cancer was found to be 77.6%- 82.8%. [8] This aggressive nature is compounded by the early onset of TNBC, being more prevalent in young women (<40) and distant migration, often, leading as far as the brain. [6]

The poor prognosis observed characteristically in TNBC is also a consequence of the immunosuppressive milieu, otherwise known as the Tumor Micro Environment or TME. Despite being one of the tumors where heavy immune cell infiltration is seen, the paradox of TNBC microenvironment is in that the majority of the infiltrating immune cells are pro-tumoral M2 polarized macrophages [9, 10, 11] with other immunosuppressive T cell populations, instead of the anti-tumoral effector cells. A critical translational aspect of this lies in that, therapeutic intervention driving re-education of this TME into anti-tumoral immunophenotype could possibly turn out beneficial for TNBC prognosis with enhanced immunotherapy responses [12, 13].

1.2 The Tumor Microenvironment and Tumor-Associated Macrophages

The tumor microenvironment (TME) is a complex but often, precisely coordinated population of immune cells, stromal components etc. The main components forming TME include Tumor-Associated Macrophages or TAMs, Tumor-infiltrating lymphocytes or TILs, Cancer-associated fibroblasts or CAFs, endothelial cells and extracellular matrix components or ECM [14, 15, 16]. Of this, TAMs are one of the major immune infiltrates, counting for 20-70% of the total immune infiltrate population in solid tumors, depending upon type of tumor and anatomical location. [17] The highly plastic nature of TAM confers upon them the ability to differentiate into either anti-tumor subtype of inflammatory M1 macrophages via classical activation or into the entirely antagonistic pro-tumor immunosuppressive M2 phenotype via alternate activation in response to environmental cues.

In case of TNBC, the majority of the TAMs morph into the pro-tumoral M2 subtype, with the characteristic expression of anti-inflammatory cytokines (IL-10, TGFβ ETC), immune-suppressive markers (PD-L1, IL-6 etc) and tissue remodelling enzymes [18, 19]. They serve particular roles in promoting the tumor progression in tumor microenvironment via following mechanisms –

1. Secretion of pro-tumoral cytokines that promote cancer cell proliferation and survival [20, 21].
2. Remodelling the extracellular matrix (ECM) by tissue remodelling enzymes to further facilitate invasion and metastasis, including processes like EMT. [12, 18]
3. Promoting angiogenesis and neovascularization [22, 12].
4. Suppressing anti-tumor T cell response, for example blocking CD4+ and CD8+ T cells via PD1-PDL1 axis. [23, 24, 25, 26, 18]

1.3 GPNMB: An Emerging Immunomodulatory Axis in Cancer

Glycoprotein non-metastatic melanoma protein B or in short GPNMB is a class I transmembrane glycoprotein first characterized in melanoma cells. [27] It is also known as Hematopoietic Growth Factor Inducible neurokinin-1 type (HGFIN) or Osteoactivin (OA, rat ortholog) or dendritic cell- heparin integrin ligand (DC-HIL, mouse ortholog) [28]. [29, 30] GPNMB has two isoforms with 572 aa and 560 aa lengths [31, 28]. It has a single pass transmembrane domain, an extracellular N-terminal part with several domains and a cytosolic C-terminal domain containing hemITAM motif [31]. Beyond the historical association with melanoma cell lines, it has also been associated in immunomodulatory mechanisms involved not only in cancer but also in auto-immune diseases [31]. GPNMB expression has been observed to be upregulated in M2 macrophages, dendritic cells (DC) and monocytes in several types of cancer, which includes breast cancer, melanoma, pancreatic cancer etc. [32]

The multifaceted immunological role of GPNMB is highly dependent on context. In this case, i.e., in the tumor microenvironment, GPNMB engagement on myeloid cells promotes anti-inflammatory signalling cascades via its interaction with receptors like Integrin αVβ3 and T cell immunoreceptor with Ig and ITIM domains (TIGIT), amplifying immunosuppression. GPNMB expressing macrophages suppress pro-inflammatory signalling mechanisms and upregulates production of IL-10 and TGFβ [28]. It has also been observed that GPNMB directly inhibits dendritic cell maturation and T cell activation, thereby impairing the anti-tumor adaptive immunity development. [32]

1.4 Rationale for GPNMB-Targeted Therapy in TNBC

GPNMB, with its well-proven role in macrophage mediated immunosuppressive mechanisms as well as the correlation to poor prognosis in various cancer subtypes, arises as one of the most promising and rational therapeutic targets for TNBC. The selective inhibition of GPNMB and subsequent reprogramming of M2 TAMs to more pro-inflammatory M1 phenotypes is a possible strategy to enhance anti-tumor responses and sensitizing TNBC tumors via both conventional mechanisms and immunotherapies. [31, 32, 33]

1.5 Study Objectives

The objective of this study is to characterize the transcriptomic consequences of GPNMB knockdown (kd) within a physiologically relevant co-culture system. The primary goals in this study include –

1. Identifying the differential gene expression patterns and systemic pathways dysregulated in GPNMB kd cases.
2. Identifying novel biomarkers and therapeutic targets that are assoiciated with GPNMB knockdown.
3. Defining potential GPNMB antagonistic mechanisms with respect to enhancing anti-tumor immunity.
4. Illuminate the role of GPNMB in pro-tumoral inflammatory and immunosuppressive mechanisms.

2. Methods

2.1 Data Source and Study Design

2.1.1 Public Dataset Acquisition

In this study, the publicly accessible bulk RNAseq dataset with accession id GSE270890 was chosen from Gene Expression Omnibus (GEO) database. [33]This dataset contained the transcriptomic data of THP-1 derived macrophage cells that were co-cultured with TNBC spheroids. [33]Four sets of conditions were taken with three biological replicates for each condition in this dataset.

Dataset Specifications:

- **Clinical trial number:** not applicable.
- **GEO Accession ID**: GSE270890
- **Sample Cohort**: N = 12 biological replicates
- **Experimental Conditions** (n=3 replicates per condition):
    1. **THP-1 alone** (baseline): THP-1 macrophages cultured in isolation (denoted "only" in raw data)
    2. **Co-culture Control**: THP-1 macrophages, co-cultured with control siRNA transfected Hs578T spheroids (denoted as "C")
    3. **GPNMB Knockdown - siRNA 1**: THP-1 macrophages, co-cultured with GPNMB-knockdown siRNA sequence 1 Hs578T spheroids (denoted as "Si1")
    4. **GPNMB Knockdown - siRNA 2**: THP-1 macrophages co-cultured with GPNMB-knockdown siRNA sequence 2 Hs578T spheroids (denoted as "Si2")

2.1.2 Experimental Context and Source

The original RNAseq experimental design consisted of three-dimensional (3D) co-culture systems, where TNBC cells (Hs578T) were formed into spheroids. The TNBC spheroids were then co-cultured with THP-1 derived macrophages (macrophage identity induced by PMA treatment) for 3 days. Two independent siRNA sequences were utilized to knockdown GPNMB expression in two sets of culture, respectively. Use of two different siRNA further strengthens the knockdown efficiency and lessens possibility of off-target silencing. This approach enabled enquiry of macrophage transcriptional responses to cancer cell-derived GPNMB relevant to that in a physiologically relevant context. [33]

2.2 Data Preprocessing and Quality Control

2.2.1 Count Matrix Assembly from GEO Supplementary Files

Raw gene expression counts were obtained from 12 GEO supplementary files in the format of count matrices where each row represents a gene and each column represents a biological sample. Count matrices contained feature-level quantification with gene identifiers in the "Name" column and quantitative expression counts in the "Total.counts" column, representing the total number of sequencing reads assigned to each gene per sample.

2.2.2 Gene Filtering and Quality Thresholding

The count matrix can contain many lowly expressed genes and many such artefacts, which, if not optimised, will increase noise and signal to noise ratio would be compromised. Therefore, standard methods were employed on the count matrix for quality checking purpose before differential expression analysis to remove biologically irrelevant genes:

1. **Minimum expression threshold**: Genes with mean read count <10 across all 12 samples were excluded from analysis. This conservative threshold ensures that identified significant differences represent biologically meaningful expression changes rather than noise from lowly expressed genes.

2. **Minimum sample representation**: Genes were retained only if they demonstrated 10 read counts in at least 3 samples (25% of cohort), ensuring that genes were genuinely expressed in multiple biological replicates rather than sporadic expression in isolated outlier samples. This step further filtered the biologically and statistically relevant genes.

3. **Gene annotation**: Only genes with assigned gene names/symbols were retained. Gene entries lacking annotation, that is N/A were excluded to maintain mapping to curated functional databases for pathway analysis. Also, duplicated gene names were fixed to contain single entry for each gene. Keeping these blank genes, duplicate genes will significantly cause error in downstream processing of data.

Following these filtering criteria, the count matrix was reduced from ~60,000 initial features to approximately ~ 15,000 mRNA-coding genes meeting expression criteria. This filtered count matrix with dimensions ~15000x12 served as input for all downstream differential expression and enrichment analyses.

2.3 Differential Gene Expression Analysis Using DESeq2

2.3.1 DESeq2 Statistical Framework

Differential gene expression analysis was performed using the **R/Bioconductor package DESeq2 v1.28+** [34] [35], the gold-standard methodology for RNAseq count data statistical analysis. DESeq2 implements a robust statistical framework that is specifically designed to handle the distributional properties of RNAseq count data, which violates assumptions of normality underlying parametric statistical methods like simple t-tests. The distribution, being skewed, instead of assumptions of bell- shaped, requires the special treatment of DEseq2, which is mostly non-parametric in nature. [36]

Normalization Strategy: **The median-of-ratios** Normalization procedure was implemented to account for library size variation across samples. This approach calculates a size factor for each sample reflecting sequencing depth:

Size_Factor_i = median (count_i,g/ geometric_mean_counts_i)

For example, size factor for sample i is the median of the ratio of count of a gene (counti,g) in i sample with geometric mean of the gene's counts across all samples.

This normalization is robust to highly expressed genes and to genes that are differentially expressed between samples, providing superior normalization compared to methods that assume the majority of genes are not differentially expressed. Taking the median of ratios, rather than the mean, down-weights outlier genes dominance.

Variance Modelling: Gene-level overdispersion parameters were estimated to account for biological and technical variation exceeding Poisson expectations. DESeq2 employs an empirical Bayes approach combining observed gene-specific dispersions with global dispersion trends, generating regularized estimates that balance gene-level accuracy with global stability.

2.3.2 Experimental Design Model Specification

Two complementary statistical models were selected for differential expression testing:

Model 1 - Primary Comparison (GPNMB Status Grouping):

- Design formula: ~GPNMB_status
- GPNMB_status factor levels (ordered):
    1. **Baseline** (reference): THP-1 macrophages cultured alone (condition "only")
    2. **CoCulture_GPNMB_High**: THP-1 co-cultured with control TNBC cancer cells (condition "C")
    3. **CoCulture_GPNMB_KD**: THP-1 co-cultured with GPNMB-knockdown cancer cells (conditions "Si1" and "Si2" combined)

This model aids the testing of the primary biological hypothesis, that GPNMB knockdown in cancer cells alters macrophage transcriptional profile differently than GPNMB-high control conditions.

Model 2 - Secondary Comparison (Individual Condition Analysis):

- Design formula: ~condition
- Enables detailed comparison of individual siRNA effects (Si1 vs. Si2) to identify siRNA-specific or off-target effects, thereby increasing resolution

2.3.3 Hypothesis Testing and Statistical Significance

For each gene, differential expression was tested using the Wald Test, which compares the observed log-fold change against the null hypothesis of no expression difference. The Wald statistic follows an approximately normal distribution under the null hypothesis, enabling calculation of two-tailed p-values.

Multiple Testing Correction: Given the high-dimensional nature of simultaneous hypothesis testing across ~18,000 genes, the Benjamini-Hochberg method was applied to control the False Discovery Rate (FDR) at $\alpha = 0.05$. This ensures that among all genes called significant, an expected 5% (or fewer) are false positives. This approach provides superior statistical power compared to stringent family-wise error rate controls for genomics applications.

2.3.4 Significance Thresholds for DEG Identification

Genes were classified to be significantly differentially expressed between the given parameters if they simultaneously meet all of the following criteria:

1. **Log2 Fold Change (|logFC| 1.0)**: Expression change of 2-fold in either direction. Here, log2FC > 1.0 translates to upregulation, increased expression while log2FC < -1.0 translates to downregulation.
   - Rationale: Changes <2-fold often fail to translate to meaningful biological consequences despite statistical significance
   - Provides practical balance between statistical stringency and biological relevance

   FDR-Adjusted p-value (padj < 0.05): Benjamini-Hochberg corrected significance threshold. The Adj.P.val gives greater significance for this kind of large expression data instead of raw P-value.
2. • Rationale: Controls false discovery rate at 5% level
3. **Minimum Expression Level (*baseMean* >= 5)**: Average read count across all 12 samples
   - Rationale: Ensures robust quantification with adequate sequencing depth
   - Excludes noisy, marginally-expressed genes prone to technical artefact

2.3.5 Primary Contrast: GPNMB Knockdown vs. GPNMB-High

The primary statistical contrast extracted was:

CoCulture_GPNMB_KD vs. CoCulture_GPNMB_High

This contrast directly tests the effect of GPNMB knockdown on macrophage (THP-1) transcriptional programs With contrast to control GPNMB-high for co-culture conditions, isolating the GPNMB-specific effect. Genes reaching significance in this contrast represent direct or indirect transcriptional targets of tumor-expressed GPNMB.

2.3.6 Validation Through Dual siRNA Approach

To ensure identified genes actually represent robust, sequence-independent GPNMB-regulated transcription, supplementary analyses examined:

- Individual siRNA effects (Si1 vs. Co-culture Control, Si2 vs. Co-culture Control)
- Genes dysregulated by both Si1 AND Si2 conditions represent high-confidence GPNMB targets
- Genes dysregulated by only one siRNA may indicate off-target or siRNA-specific effects

2.4 Data Transformation and Quality Assessment

2.4.1 Variance Stabilizing Transformation (VST)

To facilitate unsupervised clustering and visualization, count data underwent **variance stabilizing transformation (VST)** using the DESeq2 vst() function. VST applies a regularized logarithmic transformation and thereby stabilizing variance across the range of expression levels:

$vst(count) = \log(q + (q^2 + \sigma^2))$

Where q represents quasi-counts estimated from the fitted statistical model and σ^2 represents prior variance. This transformation achieves:

- Variance stabilization (similar variance for low and high-expression genes)
- Noise reduction in low-expression genes
- Preservation of biological signal

2.4.2 Principal Component Analysis (PCA)

Comprehensive quality assessment also included Principal Component Analysis (PCA) on VST-transformed count data, performing singular value decomposition on the top 500 most variable genes to identify sample relationships and assess batch effects.

Interpretation:

- **Expected Results**: Samples from identical experimental conditions should cluster tightly, indicating reproducibility. Distinct experimental conditions should separate along principal components, indicating robust biological responses.
- **Quality Indicators**: Tight clustering within conditions validates technical quality and biological replicability. Unexpected clustering or outlier samples indicate potential quality issues.

2.4.3 Heatmap Visualization

Hierarchical clustering heatmaps displayed the top 40 most significantly dysregulated genes (ranked by adjusted p-value). Heatmaps employed:

- **Distance metric**: Euclidean distance on VST-transformed counts
- **Linkage method**: Ward.D2 hierarchical agglomeration
- **Row normalization**: Z-score scaling to emphasize relative expression patterns
- **Color scale**: Red-Blue diverging palette (red = high expression, blue = low expression, white = mean)

2.5 Pathway Enrichment Analysis

2.5.1 Gene Set Enrichment Analysis (GSEA)

Functional pathway analysis employed Gene Set Enrichment Analysis (GSEA) v4.1.x, a rank based statistical approach identifying groups of functionally related genes showing co-ordinated expression changes. GSEA does not require predefined significance thresholds, instead testing whether gene sets show statistically significant bias toward the top (upregulated) or bottom (downregulated) of the complete ranked gene list.

GSEA Parameters:

- **Gene set collection**: MSigDB Hallmark Gene Set Collection [37, 38] (h.all.v2024.1.Hs.symbols.gmt)
  - Contains 50 non-overlapping gene sets representing well-defined biological processes (immune response, signaling pathways, development, metabolism, cell cycle)

  Ranking metric: Wald test statistic from DESeq2 (combines effect size and statistical significance)
- **Permutation strategy**: 1,000 random permutations of phenotypic labels to generate empirical null distribution
- **Gene set size filters**: Minimum 15 genes, maximum 500 genes per set
- **FDR significance threshold**: q-value 0.05 (5% false discovery rate)

Statistical Outputs:

- **Normalized Enrichment Score (NES)**: Quantifies direction and magnitude of gene set enrichment
  - Positive NES: Gene set predominantly upregulated in GPNMB_KD condition
  - Negative NES: Gene set predominantly downregulated in GPNMB_KD condition

- |NES| > 1.5 with FDR < 0.05 indicates robust, biologically meaningful enrichment

p-value: Empirical p-value derived from permutation distribution

- **FDR q-value**: Benjamini-Hochberg adjusted p-value
- **Leading edge genes**: Core subset of genes driving enrichment signal

2.5.2 GSEA Output Processing and Visualization

Significantly enriched pathways (padj < 0.05) were extracted and ranked by NES magnitude to identify most prominent biological alterations. Results were visualized using:

Dotplots: Simultaneous display of multiple enrichment metrics

- X-axis: GeneRatio (proportion of dysregulated genes in pathway)
- Y-axis: Pathway name
- Point size: Number of core enrichment genes
- Point color: FDR-adjusted p-value (red = most significant)

Enrichment Plots: For top dysregulated pathways, GSEA enrichment plots display:

- Ranked gene list with up/down-regulated genes marked
- Running enrichment score showing cumulative bias toward pathway genes
- Normalized enrichment score (NES) and FDR q-value

2.6 Analysis of DEG Magnitude and Direction

2.6.1 Directional Gene Classification

Significantly dysregulated genes were classified directionally:

- **Downregulated genes** (in GPNMB_KD vs. GPNMB_High): logFC < -1.0
  - Genes suppressed upon GPNMB knockdown in cancer cells
  - Represent pro-tumoral macrophage programs dependent on tumor-derived GPNMB

Upregulated genes (in GPNMB_KD vs. GPNMB_High): logFC > 1.0

- Genes induced upon GPNMB knockdown in cancer cells
- Represent anti-tumoral or metabolic shifts upon GPNMB loss

2.6.2 Effect Size Distribution

Log2 fold-change values quantify magnitude of expression changes independent of statistical significance. Distribution analysis reveals:

- Median and range of effect sizes
- Identification of genes with exceptionally large fold-changes (potential key regulators)
- Separation of modest (|logFC| 1-2) vs. dramatic (|logFC| >2) expression changes

2.7 Computational Implementation and Data Analysis

2.7.1 R Statistical Environment and Package Versions

All analyses were conducted in the R statistical environment (v4.2.x) with Bioconductor packages (v3.16+). Key packages and versions:

- **DESeq2** (v1.28+): Differential expression analysis, normalization, variance stabilization [35], specialized for RNAseq data analysis.

- **fgsea** (v1.14+): Fast GSEA implementation for rapid pathway analysis [39]
- **tidyverse** (v1.3+): Data manipulation (dplyr, tidyr, stringr) [40]
- **ggplot2** (v3.3+): Publication-quality visualization [41]
- **pheatmap** (v1.0+): Advanced heatmap construction with dendrograms [42]
- **clusterProfiler** (v4.0+): Functional enrichment analysis [43]
- **msigdbr** [38] (v7.4+): MSigDB gene set retrieval
- **enrichplot** (v1.10+): Enrichment result visualization [44]
- **RColorBrewer** (v1.1+): Color palette management [45]
- **ggrepel** (v0.9+): Text label repulsion in plots [46]
- **viridis** (v0.6+): Colorblind-safe color scales [47]

2.7.2 Workflow Implementation

The complete analysis pipeline was implemented as a single R script encompassing:

1. Package installation and loading
2. GEO data download and extraction
3. Count matrix assembly from supplementary files
4. Sample metadata creation with condition assignments
5. Gene filtering based on expression thresholds
6. DESeq Dataset object creation
7. DESeq2 differential expression analysis
8. Results extraction and filtering
9. PCA generation for quality assessment
10. Volcano plot visualization
11. Heatmap generation for top DEGs
12. GSEA pathway enrichment analysis
13. Results export to CSV and visualization files

2.7.3 Output File Generation

The analysis pipeline generated multiple output files:

- **CSV files**: Count data with DESeq2 statistics, significant DEGs, GSEA results.
- **PDF/PNG figures**: PCA plots, volcano plots, heatmaps, GSEA dot plots and enrichment curves.

2.8 Reproducibility and Data Availability

2.8.1 Raw Data Deposition and Access

Raw RNAseq data analysed in this study are publicly available in the NCBI Gene Expression Omnibus (GEO) database under series accession number GSE270890. [33] Count matrices and sample metadata are directly downloadable from GEO, ensuring complete reproducibility and independent verification by research community.

2.8.2 Analysis Code Availability

Complete R code documenting all analyses from GEO data download through final visualization have been made available in supplementary materials as well as GitHub profile. Scripts include extensive inline comments explaining analytical decisions, parameter specifications and function usage, facilitating reproducibility and adaptation to similar datasets.

2.9 Statistical Rigor and Interpretation Framework

2.9.1 Multiple Comparison Correction Strategy

All statistical tests incorporated Benjamini-Hochberg FDR adjustment at $\alpha = 0.05$, controlling expected proportion of false positives among rejected hypotheses. This stringent threshold minimizes false discovery rate while maintaining adequate statistical power for the sample size and effect size distribution observed.

2.9.2 Effect Size Reporting Alongside P-values

In addition to adjusted p-values, all findings are accompanied by log2 fold-change values and standard errors, enabling assessment of effect magnitude independent of statistical significance. This practice prevents conflation of statistical significance with biological importance.

2.9.3 Interpretation Framework for DEG Results

Identified differentially expressed genes represent macrophage transcriptional responses to cancer cell-derived GPNMB status, but causality cannot be definitively established from transcriptomic data alone. Functional validation through complementary approaches (e.g., in vitro manipulation, pathway inhibition) would be required to establish mechanistic links.

3. Results

3.1 Sample Quality Assessment and Quality Control

Principal component analysis of the 12 samples revealed excellent quality metrics and clear biologically relevant segregation between the mentioned experimental conditions (Figure 1). PC1 (with 46.7% of variance) distinctly separates the control co-culture samples (C1 - C3) from both GPNMB knockdown conditions, indicating highly robust gene expression changes or differential expression caused due to GPNMB knockdown. While PC2 (with 27.9% of variance) further discriminated between siGPNMB1 and siGPNMB2 conditions, suggesting potentially distinct off-target or partial knockdown effects. Notably, THP1 macrophages cultured alone (only1 - only3, red) clustered separately from co-culture conditions, implying the critical role of cancer cell-macrophage interactions on the transcriptomic landscape of the macrophages. Within each condition, biological replicates clustered tightly, confirming high experimental reproducibility and minimal batch effects. It is also important to mention that the si2_1 was found isolated from the clusters and this outlier nature could be due to technical artefacts.

3.2 Differential Gene Expression Analysis

The RNAseq analysis was able to identify 232 significantly differentially expressed genes (DEGs) between the GPNMB knockdown and control co-culture samples ($|\log2FC| \geq 1.0$, p.adj < 0.05). Among these 155 genes were found to be downregulated upon GPNMB knockdown, while 77 genes were upregulated in comparison to control co-cultures. The magnitude of expression change was ranged widely, with the most dramatically downregulated genes including FLOT1 (log2FC= -1.17; baseMean= 772), CEMIP (log2FC= -4.64; baseMean= 115), INAVA (log2FC= -1.77, baseMean= 75), IGFBP4 (log2FC= -1.63, baseMean= 375) and LSP1 (log2FC= -1.01, baseMean= 1269).

These highly downregulated genes are predominantly involved in immune suppression mechanisms and other pro-tumoral functions. FLOT1 (flotillin-1) is a lipid-raft associated membrane protein involved in cellular signaling and endocytosis. It has been associated with cancer cell survival, exosome formation and metastasis. CEMIP (cell migration inducing hyaluronic acid binding protein), another gene that is very strongly downregulated in GPNMB knockdown, has its role in cancer cell migration and invasion through hyaluronate mediated signaling. IGFBP4 (insulin-like growth factor binding protein 4) modulates insulin like growth factor signaling. This pathway is critical for cancer cell proliferation and survival.

The genes that were found to be upregulated in GPNMB knockdown samples include SOCS2 (log2FC= 1.46, baseMean= 197), EPX (log2FC= 1.16, baseMean= 195), PRG2 (log2FC= 1.11, baseMean= 399), NKD2 (log2FC=

1.09, baseMean= 119) and NOG (log2FC= 1.50, baseMean= 54). These genes are related to the negative regulation of inflammatory signaling pathways. This further implies the alternate hypothesis that GPNMB knockdown significantly shifts the cellular state towards reduced levels of immune activation and increased levels of feedback suppression of inflammatory pathways.

3.3 Functional Enrichment and Pathway Analysis

Gene Set Enrichment Analysis or GSEA revealed profound dysregulation of pro-tumoral and immunosuppressive pathways upon GPNMB knockdown. The pathways to be most significantly downregulated include:

IL6-JAK-STAT3 Signaling (NES= -2.33, q= 4.6e-09): IL6-JAK-STAT3 represents a canonical immunosuppressive axis. Elevated IL-6 and STAT3 activation drives macrophage polarization toward pro-tumoral M2 phenotype and inhibits anti-tumor T cell differentiation. GPNMB knockdown dramatically suppressed the core genes involved in this pathway including IL6, IL6ST (IL-6 signal transducer), STAT3, JAK2 and multiple downstream effectors of the same. This finding suggests that GPNMB functions to stabilize IL-6 autocrine/paracrine signaling in TAMs, maintaining immunosuppressive macrophage states.

TGF-β Signaling (NES= -2.37, q= 4.5e-09): Transforming growth factor-β (TGF-β) is a pleiotropic cytokine with widely established immunosuppressive functions - driving regulatory T cell differentiation, inhibiting effector T cell responses and promoting epithelial to mesenchymal transition in cancer cells. The core TGF-β pathway genes that were downregulated upon GPNMB knockdown included TGFB1 (TGF-β ligand itself), ACVR1B and BMPR2 (BMP/TGF-β serine/threonine kinase receptor) and SMAD1/6/7 (SMAD family signal transduction molecules). This suggests GPNMB maintains a pro-fibrotic, immunosuppressive state through sustained TGF-β signaling in the TME.

Inflammatory Response (NES = -2.82, q= 1e-10): Despite nomenclature suggesting activation, GPNMB knockdown suppressed multiple pro-inflammatory mediators including TNF, IL1B, IL18, CCL2, CCL20, CXCL6, CXCL8 (IL-8) and complement cascade components (C3, C5a). The root of this apparent paradox lies in the complexity of "inflammatory" responses in the tumor context — these mediators drive tumor-promoting inflammation rather than anti-tumor immunity.

Epithelial to Mesenchymal Transition (NES= -2.75, q= 1e-10): EMT is one of the most critical processes that enable the cancer cells to acquire migratory and invasive properties, facilitating in metastasis. GPNMB knockdown suppressed EMT hallmark genes like SNAI2 (Slug), WNT5A, FZD receptors and multiple metalloproteases, suggesting GPNMB promotes metastatic capacity through EMT induction.

TNF α Signaling via NFκB (NES= -2.51, q= 1e-10): TNF-α mediated NFκB signaling pathway drives pro-tumoral inflammation. GPNMB knockdown resulted in the suppression of core NFκB targets including CCND1 (cyclin D1), BCL2, IL10, IL6 and multiple chemokines, indicating GPNMB sustains proliferative and pro-inflammatory NFκB signaling.

In contrast, GPNMB knockdown also upregulated several pathways associated with oxidative metabolism and cell cycle regulation:

MYC Targets V1 and V2 (NES= 2.39 and 2.50, q= 1e-10): MYC is the master transcription factor driving cell proliferation, ribosomal biogenesis and biosynthesis. Upregulation of MYC target pathways upon GPNMB depletion suggests a metabolic shift toward anabolic biosynthetic processes, potentially indicating the restoration of metabolic fitness following the removal of immunosuppressive signaling.

Oxidative Phosphorylation (NES= 2.13, q= 1.1e-09): This pathway includes components of ATP synthase, electron transport chain proteins and also mitochondrial biogenesis factors. This indicates enhanced oxidative metabolism. Upregulation of oxidative phosphorylation further suggests that GPNMB knockdown promotes a metabolic shift away from the lactate producing Warburg phenotype characteristic of immunosuppressive myeloid cells.

3.4 GPNMB Regulated Immune and Inflammatory Networks

Detailed analysis of significantly downregulated genes revealed networks of immunosuppressive factors that were directly or indirectly regulated by GPNMB. The key downregulated genes include:

Macrophage Derived Cytokines and Chemokines: IL6 (interleukin-6), IL10, IL1B, TNF, CXCL8 (IL-8), CCL2, CCL20 are some of the classical pro-tumoral secreted factors produced by M2 TAMs. Their well-coordinated downregulation upon GPNMB knockdown implies the role of GPNMB as a master regulator of TAM directed cytokine production.

Complement and Coagulation Cascade: C3, C5a (complement component 5a), CFB (complement factor B), CFH (complement factor H), tissue factor F3 and multiple serpins were also significantly downregulated. These factors promote thrombosis, angiogenesis and immune cell infiltration into tumors, facilitating the progression of metastasis.

Adhesion and Migration Molecules: CCL2, ICAM1, PECAM1 and multiple integrins coordinate immune cell recruitment and effusion into tumors. Their suppression upon GPNMB knockdown suggests impaired immune cell trafficking and tumor infiltration.

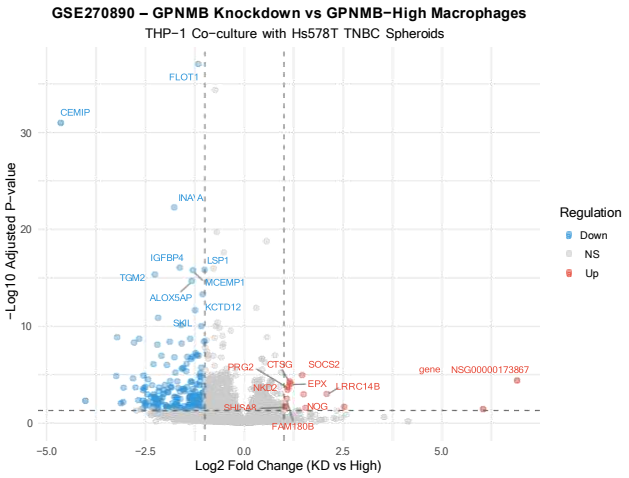
Metabolic Enzymes and Nutrient Transporters: SLC2A3 (GLUT3, glucose transporter), SLC6A6 (taurine transporter), LDHA (lactate dehydrogenase A) and multiple amino acid transporters were downregulated, which indicates metabolic reprogramming away from Warburg-like glycolysis and toward more oxidative metabolism.

3.5 Core Pathway Enrichment in GPNMB Downregulation

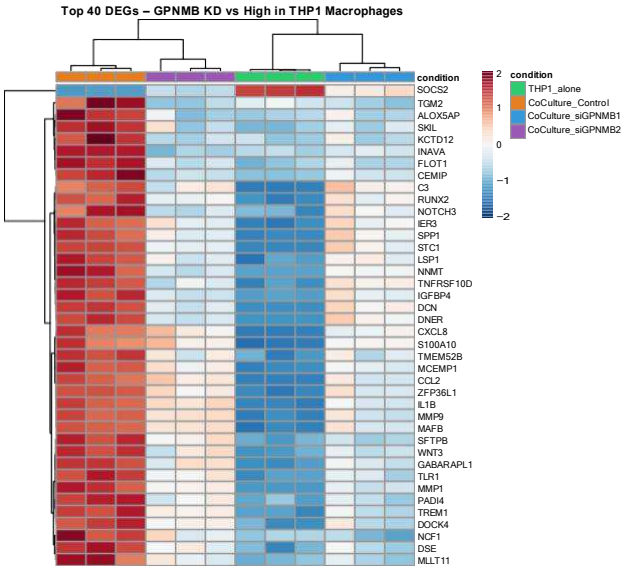
Among the 232 significant DEGs, 89 genes were identified as "core enrichment" genes and were present in multiple significantly dysregulated Hallmark pathways. Which represents nodes of high functional connectivity. The most frequently affected core enrichment genes in downregulated pathways include:

- **IL6:** Core to IL6-JAK-STAT3 signaling, TNF-α signaling, inflammatory response.
- **STAT3:** JAK-STAT cascade convergence point.
- **TNF:** Multiple pathway participation including TNF-α signaling, inflammatory response.
- **CCL2/CCL20:** Chemokine axis relevant to immune infiltration.
- **C3/Complement factors:** Coagulation and complement cascades.
- **TGFB1/TGF-β pathway components:** Multi-pathway integration point.
- **Multiple collagens and ECM components:** EMT and cancer cell invasion.

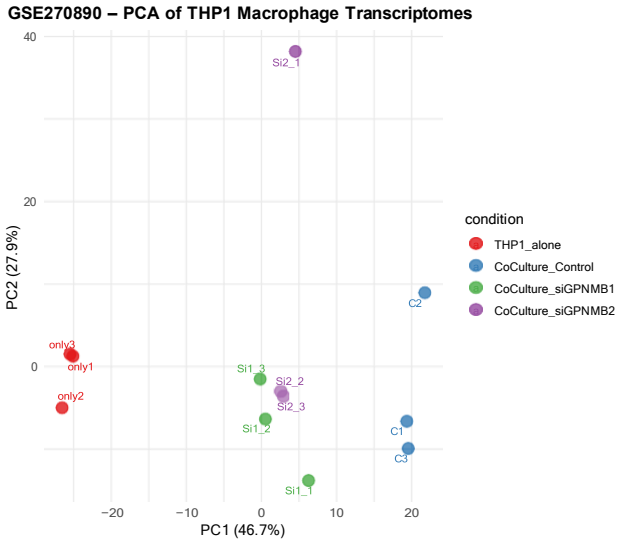
This network analysis emphasizes that GPNMB functions not in isolation but as a master orchestrator of interconnected immune suppression and pro-tumoral networks.



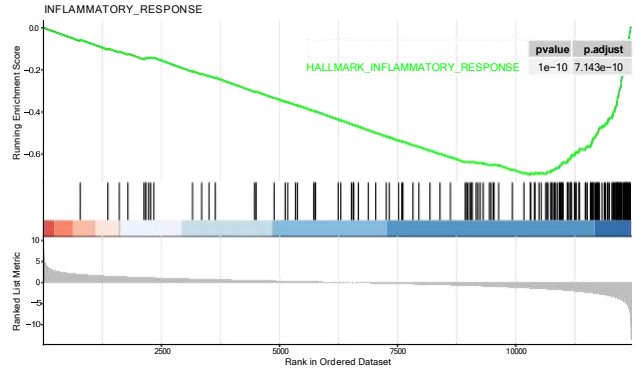
(a) Volcano plot shows all upregulated and downregulated genes in GPNMB knockdown vs. high



(b) Heatmap of top 40 differentially expressed genes in GPNMB knockdown vs. high



(c) PCA plot of differentially expressed genes in GPNMB knockdown vs. high



(d) GSEA enrichment plots for selected MSigDB Hallmark pathways significantly dysregulated in GPNMB knockdown versus GPNMB-high co-culture conditions in THP-1 macrophages from GSE270890

Figure 1: Result figures from DEG Analysis and GSEA

4. Discussion

4.1 GPNMB as a Central Hub in TAM-Mediated Immunosuppression

The present transcriptomic analysis reveals GPNMB as a critical orchestrator of macrophage-mediated immunosuppression in the TNBC microenvironment. The coordinate downregulation of IL6-JAK-STAT3, TGF-β, TNF-α signaling, and inflammatory response pathways upon GPNMB knockdown indicates these networks are not independently regulated but rather integrated under GPNMB control. This functional architecture suggests GPNMB acts as a master regulator integrating multiple immunosuppressive signals into a coordinated program.

4.2 GPNMB-IL6-STAT3 Axis as Primary Mediator

The dramatic downregulation of IL6-JAK-STAT3 signaling (NES = -2.33) represents the most striking pathway dysregulation upon GPNMB knockdown. IL-6 is a pleiotropic cytokine with complex roles in tumor immunity. In the myeloid compartment, IL-6 drives STAT3 phosphorylation and activation. This, in turn, promotes a self-amplifying feedback loop of IL-6 production and M2 polarization. In the T cell compartment, IL-6 drives Th17 differentiation while also simultaneously inhibiting Foxp3+ regulatory T cell development, thereby promoting a pro-inflammatory but immunosuppressive Th17-skewed response.

The evidence substantially suggests that GPNMB directly or indirectly promotes IL-6 transcription and/or autocrine IL-6 signaling. This could occur through multiple mechanisms like : (1) direct transcriptional upregulation of IL6 promoter through GPNMB-mediated signaling, (2) enhanced stability of IL6 mRNA through post-transcriptional mechanisms, (3) paracrine amplification where GPNMB-expressing macrophages stimulate neighboring myeloid cells or stromal cells to produce IL-6 or (4) GPNMB-mediated suppression of negative regulators of IL-6 signaling, such as SOCS3 (interestingly, SOCS3 was not significantly regulated in this analysis, suggesting SOCS3-independent mechanisms).

4.3 TGF-β Signaling and Pro-Fibrotic Immunity

The profound downregulation of TGF-β signaling upon GPNMB knockdown (NES = -2.37) unveils a role for GPNMB in maintaining a pro-fibrotic, immunosuppressive microenvironment. TGF-β drives differentiation and maintenance of Foxp3+ regulatory T cells, a cell population crucial for tumor immune evasion. Additionally, TGF-β suppresses effector T cell responses while promoting EMT in cancer cells and myofibroblast differentiation in cancer-associated fibroblasts. The TAM compartment itself produces substantial quantities of TGF-β, which suggests GPNMB sustains a feedforward loop of TGF-β autocrine signaling.

Notably, while TGFB1 mRNA levels decreased, many TGF-β target genes involved in EMT, fibrosis, and immune suppression showed even more dramatic reductions, suggesting GPNMB may regulate both TGF-β production and signaling cascade amplification.

4.4 GPNMB, Metabolic Reprogramming, and Immune Function

A surprising but significant finding was the upregulation of oxidative phosphorylation (OXPHOS, NES = 2.13) and MYC targets (NES = 2.39-2.50) upon GPNMB knockdown. This metabolic reprogramming illuminates a mechanistic link between immunosuppression and metabolism. Pro-tumoral M2 macrophages characteristically rely on glycolytic metabolism (Warburg phenotype) with associated production of lactate, which directly suppresses anti-tumoral T cell function.

The upregulation of OXPHOS and MYC targets suggests GPNMB knockdown recalibrates TAMs toward a more oxidative, biosynthetically active metabolism. This metabolic shift may have dual consequences: (1) reduced lactate production, improving local pH conditions for effector T cell function and (2) increased anabolic metabolism supporting production of molecules required for pro-inflammatory M1 macrophage functions such as nitric oxide synthase and reactive oxygen species generation.

4.5 Clinical Implications and Therapeutic Opportunities

The present findings support GPNMB antagonism as a rational and practical therapeutic strategy for TNBC through multiple complementary mechanisms:

Macrophage Re-Education: GPNMB targeting would not only suppress multiple pro-tumoral TAM programs (IL-6, TGF-β, EMT support, immune cell suppression) but also would potentially promote metabolic shifts towards a more immunostimulatory macrophage status. This represents a fundamentally more potent approach than targeting individual cytokines, utilizing the interconnected nature of GPNMB-regulated networks.

Synergy with Immunotherapy: Suppression of IL-6-STAT3 signaling, TGF-β, and T cell suppression would enhance susceptibility to checkpoint inhibitor therapy. While TNBC has demonstrated higher immunogenicity than other breast cancer subtypes, the effectiveness of anti-PD-1/PD-L1 approaches has been, at best, modest, suggesting

the TME provides multiple layers of suppression. GPNMB targeting could address some of these layers, enhancing the efficacy of immunotherapy.

Combination Strategies: The downregulation of EMT and metastasis-associated pathways upon GPNMB knockdown suggests potential synergy with non-invasive therapies, anti-angiogenic approaches, or conventional chemotherapy. Furthermore, combination with small molecule JAK inhibitors, TGF-β pathway inhibitors, or IL-6 antagonists could provide additive benefits.

4.6 Limitations and Future Directions

Several limitations warrant discussion. First, this analysis employed a co-culture model combining cancer cells with differentiated monocytes rather than primary patient-derived TAMs. While THP1-macrophage co-cultures provide excellent tractability and reproducibility, they may not fully recapitulate the complexity of TME resident macrophages that have undergone months of in situ conditioning. Second, the analysis represents a snapshot of steady-state transcriptomics; longitudinal studies or single-cell sequencing approaches could reveal temporal dynamics and cell-state-specific GPNMB functions. Third, while transcriptomics identifies dysregulated pathways, validation at protein and functional levels (e.g., T cell co-culture assays, macrophage polarization flow cytometry, cytokine production assays) would strengthen mechanistic conclusions.

Future studies should incorporate:

1. **Single-Cell RNA-Seq:** SCRNAseq would enable identification of GPNMB-regulated programs in specific myeloid subsets and identify rare cell populations disproportionately affected by GPNMB manipulation.
2. **Functional Validation:** Macrophage polarization assays, M1/M2 marker expression flow cytometry, and T cell co-culture experiments measuring T cell proliferation and activation would validate transcriptomic predictions.
3. **Primary Sample Analysis:** Studies employing primary macrophages from TNBC patient tumors would provide clinical relevance and potentially identify patient-specific GPNMB-regulated networks.
4. **Longitudinal Tracking:** Time-course RNA-seq could reveal primary immediate-early responses to GPNMB disruption versus secondary adaptive changes.
5. **GPNMB Protein Studies:** Immunohistochemistry and immunofluorescence on clinical TNBC samples would determine GPNMB expression patterns in situ and correlate with immune infiltration, EMT markers, and patient outcomes.

5. Conclusions

This comprehensive transcriptomic analysis demonstrates that GPNMB functions as a critical orchestrator of pro-tumoral immunity in TNBC through coordinate regulation of IL6-JAK-STAT3, TGF-β, TNF-α, inflammatory response, and epithelial-mesenchymal transition pathways. GPNMB-expressing macrophages represent a key component of the immunosuppressive TNBC microenvironment, with therapeutic antagonism of GPNMB offering an attractive strategy to reprogram TAMs, enhance anti-tumor immunity, and potentially improve outcomes in this challenging cancer subtype.

The identification of GPNMB as a convergence point for multiple immunosuppressive networks suggests that GPNMB-targeted therapies could provide superior efficacy compared to single-target approaches through simultaneous disruption of interconnected suppressive programs. Combined with conventional and immunotherapeutic approaches, GPNMB antagonism represents a promising avenue for improving TNBC patient outcomes and deserves further pre-clinical investigation and clinical evaluation.

6. Declarations

1. **Ethical Approval and Consent to Participate** Not applicable. This study used only publicly available, de-identified data (GEO accession GSE270890) and did not involve any new experiments on human participants or animals.

2. **Consent for Publication** Not applicable.

3. **Funding** The authors received no specific funding for this work.

References

[1] William D. Foulkes, Ian E. Smith, and Jorge S. Reis-Filho. *Triple-Negative Breast Cancer*. Nov. 2010. doi: 10.1056/nejmra1001389. url: https://doi.org/10.1056/nejmra1001389.

[2] Rebecca Dent et al. "Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence". In: *Clinical Cancer Research* 13 (15 Aug. 2007), pp. 4429–4434. issn: 1078-0432, 1557-3265. doi: 10.1158/1078-0432.ccr-06-3045. url: http://dx.doi.org/10.1158/1078-0432.ccr-06-3045.

[3] Paola Zagami and Lisa A. Carey. *Triple negative breast cancer: Pitfalls and progress*. Aug. 2022. doi: 10.1038/s41523-022-00468-0. url: https://doi.org/10.1038/s41523-022-00468-0.

[4] Shinjit Mani, Abhishek Hande, and Sergei Boichuk. "Triple-Negative Breast Cancer: the Current Aspects of Pathogenesis and Therapies". In: *BioNanoScience* 12 (4 May 2022), pp. 1404–1435. issn: 2191-1630, 2191-1649. doi: 10.1007/s12668-022-00991-1. url: https://doi.org/10.1007/s12668-022-00991-1.

[5] Yun Li et al. *Recent advances in therapeutic strategies for triple-negative breast cancer*. Aug. 2022. doi: 10.1186/s13045-022-01341-0. url: https://doi.org/10.1186/s13045-022-01341-0.

[6] Maria Eleni Hatzipanagiotou et al. "Does timing of neoadjuvant chemotherapy influence the prognosis in patients with early triple negative breast cancer?" In: *Journal of Cancer Research and Clinical Oncology* 149 (13 July 2023), pp. 11941–11950. issn: 0171-5216, 1432-1335. doi: 10.1007/s00432-023-05060-y. url: https://doi.org/10.1007/s00432-023-05060-y.

[7] Li Yin et al. "Triple-negative breast cancer molecular subtyping and treatment progress". In: *Breast Cancer Research* 22 (1 June 2020), pp. 61–61. issn: 1465-5411, 1465-542X. doi: 10.1186/s13058-020-01296-5. url: https://doi.org/10.1186/s13058-020-01296-5.

[8] Chagkrit Ditsatham et al. "Breast Cancer Subtypes in Northern Thailand and Barriers to satisfactory survival outcomes". In: *BMC Cancer* 22 (1 Nov. 2022). issn: 1471-2407. doi: 10.1186/s12885-022-10196-0. url: https://doi.org/10.1186/s12885-022-10196-0.

[9] Hong Jin, Xinyue Meng, and Jianwei Feng. *Mechanisms of tumor-associated macrophages in breast cancer and treatment strategy*. Feb. 2025. doi: 10.3389/fimmu.2025.1560393. url: https://doi.org/10.3389/fimmu.2025.1560393.

[10] Yang Liu et al. *Advances in immunotherapy for triple-negative breast cancer*. Sept. 2023. doi: 10.1186/s12943-023-01850-7. url: https://doi.org/10.1186/s12943-023-01850-7.

[11] Tian-Jian Yu and Gen-Hong Di. "Role of tumor microenvironment in triple-negative breast cancer and its prognostic significance". In: *Chinese Journal of Cancer Research* 29 (3 Jan. 2017), pp. 237–252. issn: 1000-9604, 1993-0631. doi: 10.21147/j.issn.1000-9604.2017.03.10. url: https://doi.org/10.21147/j.issn.1000-9604.2017.03.10.

[12] Daniela F. Quail and Johanna A. Joyce. *Microenvironmental regulation of tumor progression and metastasis*. Nov. 2013. doi: 10.1038/nm.3394. url: https://doi.org/10.1038/nm.3394.

[13] Zijie Guo et al. *Tumor microenvironment and immunotherapy for triple-negative breast cancer*. Dec. 2024. doi: 10.1186/s40364-024-00714-6. url: https://doi.org/10.1186/s40364-024-00714-6.

[14] Nicole Anderson and M. Celeste Simon. "The tumor microenvironment". In: *Current Biology* 30 (16 Aug. 2020). issn: 0960-9822, 1879-0445. doi: 10.1016/j.cub.2020.06.081. url: https://doi.org/10.1016/j.cub.2020.06.081.

[15] Maonan Wang et al. *Role of tumor microenvironment in tumorigenesis*. Jan. 2017. doi: 10.7150/jca.17648. url: https://doi.org/10.7150/jca.17648.

[16] Yiqi Fan and Shuai He. *The Characteristics of Tumor Microenvironment in Triple Negative Breast Cancer*. Jan. 2022. url: https://doaj.org/article/a87f2752ab1b4cf6b75219e13bc35ee5.

[17] Anghesom Ghebremedhin et al. *Tumor-Associated Macrophages as Major Immunosuppressive Cells in the Tumor Microenvironment*. Oct. 2024. doi: 10.3390/cancers16193410. url: https://doi.org/10.3390/cancers16193410.

[18] Maija Hollmén et al. "Characterization of macrophage - cancer cell crosstalk in estrogen receptor positive and triple-negative breast cancer". In: *Scientific Reports* 5 (1 Mar. 2015). issn: 2045-2322. doi: 10.1038/srep09188. url: https://doi.org/10.1038/srep09188.

[19] Kristine Cate S. et al. "Triple-negative breast cancer influences a mixed M1/M2 macrophage phenotype associated with tumor aggressiveness". In: *PLoS ONE* 17 (8 Aug. 2022). issn: 1932-6203. doi: 10 . 1371 / journal.pone.0273044. url: https://doi.org/10.1371/journal.pone.0273044.

[20] Huakan Zhao et al. "Inflammation and tumor progression: signaling pathways and targeted intervention". In: *Signal Transduction and Targeted Therapy* 6 (1 July 2021), pp. 263–263. issn: 2059-3635, 2095-9907. doi: 10.1038/s41392-021-00658-5. url: https://doi.org/10.1038/s41392-021-00658-5.

[21] Glauben Landskron et al. *Chronic Inflammation and Cytokines in the Tumor Microenvironment*. Jan. 2014. doi: 10.1155/2014/149185. url: https://doi.org/10.1155/2014/149185.

[22] Douglas Hanahan and Lisa M. Coussens. *Accessories to the Crime: Functions of Cells Recruited to the Tumor Microenvironment*. Mar. 2012. doi: 10.1016/j.ccr.2012.02.022. url: https://doi.org/10.1016/j.ccr.2012.02.022.

[23] Asmaa El-Kenawi, Kay Hänggi, and Brian Ruffell. *The Immune Microenvironment and Cancer Metastasis*. Sept. 2019. doi: 10.1101/cshperspect.a037424. url: https://doi.org/10.1101/cshperspect.a037424.

[24] Shalini Kashipathi Sureshbabu et al. "Immunosuppressive microenvironment in oral cancer: implications for cancer immunotherapy". In: *Exploration of Immunology* (Aug. 2021). issn: 2768-6655. doi: 10.37349/ei.2021.00013. url: https://doi.org/10.37349/ei.2021.00013.

[25] Frank Riedel et al. "Expression of IL-18 in patients with head and neck squamous cell carcinoma". In: *International Journal of Molecular Medicine* 13 (2 Feb. 2004), pp. 267–72. issn: 1107-3756, 1791-244X. doi: 10.3892/ijmm.13.2.267. url: https://doi.org/10.3892/ijmm.13.2.267.

[26] WanJun Chen et al. "Conversion of Peripheral CD4+CD25 Naive T Cells to CD4+CD25+ Regulatory T Cells by TGF-β Induction of Transcription Factor Foxp3". In: *The Journal of Experimental Medicine* 198 (12 Dec. 2003), pp. 1875–1886. issn: 0022-1007, 1540-9538. doi: 10 . 1084 / jem . 20030152. url: https://doi.org/10.1084/jem.20030152.

[27] Manisha Taya and Stephen R. Hammes. *Glycoprotein Non-Metastatic Melanoma Protein B (GPNMB) and Cancer: A Novel Potential Therapeutic Target*. Oct. 2017. doi: 10.1016/j.steroids.2017.10.013. url: https://doi.org/10.1016/j.steroids.2017.10.013.

[28] Hengfang Zhang et al. "GPNMB plays an active role in the M1/M2 balance". In: *Tissue and Cell* 74 (Nov. 2021), pp. 101683–101683. issn: 0040-8166, 1532-3072. doi: 10.1016/j.tice.2021.101683. url: https://doi.org/10.1016/j.tice.2021.101683.

[29] April A. N. Rose et al. "Osteoactivin Promotes Breast Cancer Metastasis to Bone". In: *Molecular Cancer Research* 5 (10 Oct. 2007), pp. 1001–1014. issn: 1541-7786, 1557-3125. doi: 10.1158/1541-7786.mcr-07-0119. url: https://doi.org/10.1158/1541-7786.mcr-07-0119.

[30] Oneida A. Arosarena et al. "Osteoactivin regulates head and neck squamous cell carcinoma invasion by modulating matrix metalloproteases". In: *Journal of Cellular Physiology* 233 (1 Mar. 2017), pp. 409–421. issn: 0021-9541, 1097-4652. doi: 10.1002/jcp.25900. url: https://doi.org/10.1002/jcp.25900.

[31] Pei-Suen Tsou and Amr H. Sawalha. "Glycoprotein nonmetastatic melanoma protein B: A key mediator and an emerging therapeutic target in autoimmune diseases". In: *The FASEB Journal* 34 (7 May 2020), pp. 8810–8823. issn: 0892-6638, 1530-6860. doi: 10.1096/fj.202000651. url: https://doi.org/10.1096/fj.202000651.

[32] Zhenyu Wu et al. “The Role of Glycoprotein Non-Metastatic Melanoma Protein B (GPNMB) in Tumors and Other Diseases”. In: *MedComm – Oncology* 4 (3 Aug. 2025). issn: 2769-6448. doi: 10.1002/mog2.70033. url: https://doi.org/10.1002/mog2.70033.

[33] Ting-Lan Cao et al. “Tumor-expressed GPNMB orchestrates Siglec-9 + TAM polarization and EMT to promote metastasis in triple-negative breast cancer”. In: *Proceedings of the National Academy of Sciences* 122 (36 Sept. 2025). issn: 0027-8424, 1091-6490. doi: 10.1073/pnas.2503081122. url: https://doi.org/10.1073/pnas.2503081122.

[34] Nick Owens, Elena De Domenico, and Michael J. Gilchrist. “An RNA-Seq Protocol for Differential Expression Analysis”. In: *Cold Spring Harbor Protocols* 2019 (6 Apr. 2019). issn: 1559-6095, 1940-3402. doi: 10.1101/pdb.prot098368. url: https://doi.org/10.1101/pdb.prot098368.

[35] Michael I. Love, Wolfgang Huber, and Simon Anders. “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2”. In: *Genome biology* 15 (12 Dec. 2014), pp. 550–550. issn: 1474-7596, 1474-760X. doi: 10.1186/s13059-014-0550-8. url: https://doi.org/10.1186/s13059-014-0550-8.

[36] Simon Anders et al. “Count-based differential expression analysis of RNA sequencing data using R and Bioconductor”. In: *Nature Protocols* 8 (9 Aug. 2013), pp. 1765–1786. issn: 1750-2799, 1754-2189. doi: 10.1038/nprot.2013.099. url: https://doi.org/10.1038/nprot.2013.099.

[37] Aravind Subramanian et al. “Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles”. In: *Proceedings of the National Academy of Sciences* 102 (43 Sept. 2005), pp. 15545–15550. issn: 0027-8424, 1091-6490. doi: 10.1073/pnas.0506580102. url: https://doi.org/10.1073/pnas.0506580102.

[38] Igor Dolgalev. “msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format”. In: (Apr. 2018). doi: 10.32614/cran.package.msigdbr. url: https://doi.org/10.32614/cran.package.msigdbr.

[39] Gennady Korotkevich et al. “Fast gene set enrichment analysis”. In: *bioRxiv (Cold Spring Harbor Laboratory)* (June 2016). doi: 10.1101/060012. url: https://doi.org/10.1101/060012.

[40] Hadley Wickham et al. “Welcome to the Tidyverse”. In: *The Journal of Open Source Software* 4 (43 Nov. 2019), pp. 1686–1686. issn: 2475-9066. doi: 10.21105/joss.01686. url: https://doi.org/10.21105/joss.01686.

[41] Hadley Wickham. *ggplot2*. Springer International Publishing, Jan. 2016. doi: 10.1007/978-3-319-24277-4. url: https://doi.org/10.1007/978-3-319-24277-4.

[42] Raivo Kolde. “pheatmap: Pretty Heatmaps”. In: (Oct. 2010). doi: 10.32614/cran.package.pheatmap. url: https://doi.org/10.32614/cran.package.pheatmap.

[43] Guangchuang Yu et al. “clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters”. In: *OMICS A Journal of Integrative Biology* 16 (5 Mar. 2012), pp. 284–287. issn: 1536-2310, 1557-8100. doi: 10.1089/omi.2011.0118. url: https://doi.org/10.1089/omi.2011.0118.

[44] Zhi Zhao, Manuela Zucknick, and Tero Aittokallio. “EnrichIntersect: an R package for custom set enrichment analysis and interactive visualization of intersecting sets”. In: *Bioinformatics Advances* 2 (1 Jan. 2022). issn: 2635-0041. doi: 10.1093/bioadv/vbac073. url: https://doi.org/10.1093/bioadv/vbac073.

[45] Erich Neuwirth. “RColorBrewer: ColorBrewer Palettes”. In: (Oct. 2002). doi: 10.32614/cran.package.rcolorbrewer. url: https://doi.org/10.32614/cran.package.rcolorbrewer.

[46] Kamil Slowikowski. “ggrepel: Automatically Position Non-Overlapping Text Labels with ’ggplot2’”. In: (Jan. 2016). doi: 10.32614/cran.package.ggrepel. url: https://doi.org/10.32614/cran.package.ggrepel.

[47] Simon Garnier et al. “sjmgarnier/viridis: CRAN release v0.6.3”. In: (May 2023). doi: 10.5281/ZENODO.4679423. url: https://zenodo.org/record/4679423.