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.

Matrix stiffness-related molecular subtypes and prognostic signature in breast cancer

pu shen
Peking Union Medical College Hospital
yunru chen
China-Japan Friendship Hospital
meng yang
mengyang7012@163.com

China-Japan Friendship Hospital

Research Article

Keywords: breast cancer, matrix stiffness, cancer-associated fibroblasts, tumor microenvironment, machine learning, breast cancer

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11595411/v1

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

Additional Declarations: No competing interests reported.

Matrix stiffness-related molecular subtypes and prognostic signature in breast cancer

Shen Pu[1, 2], Chen Yunru[3] , Yang Meng[2]

Affiliations

[1] Peking Union Medical College, Beijing, China. 17838537820@163.com

[2] Department of General Surgery, China-Japan Friendship Hospital, Beijing, China. mengyang7012@163.com

[3] Department of Surgery of Traditional Cinese Medicine, China-Japan Friendship Hospital, Beijing, China, chenyunru2013@126.com

Corresponding Author: Yang Meng

Abstract

Background: Matrix stiffness, a hallmark of the tumor microenvironment, plays a critical role in breast cancer progression by regulating cancer-associated fibroblasts activation and immune remodeling. However, large-scale clinical evidence linking matrix stiffness-related molecular features to CAF heterogeneity, prognosis, and therapeutic response remains scarce.

Methods: Matrix stiffness-related genes were compiled from MSigDB, literature mining, and a pan-cancer signature. Differentially expressed genes between tumor and normal tissues were intersected with Matrix stiffness-related genes, and univariate Cox regression was used to identify prognostic genes. Consensus clustering was performed to define molecular subtypes. Single-cell RNA-seq (GSE176078) was used to identify stiffness-associated CAF subtypes. Bulk transcriptomic (TCGA-BRCA, GSE96058), and machine learning analyses were integrated to construct a prognostic signature. Immune infiltration and drug response were evaluated.

Results: A total of 1,695 MSRGs were compiled, finally yielding 40 prognostic genes. Two molecular subtypes were identified: Cluster 1 exhibited enriched ECM remodeling and mechanotransduction pathways, whereas Cluster 2 showed an

inflamed, stroma-rich immune microenvironment with elevated immune checkpoint expression and superior survival. Single-cell analysis revealed that CAFs had the highest matrix stiffness scores, with myCAFs as the dominant contributor, driving ECM-related crosstalk with epithelial and endothelial cells. Among 101 machine learning models, the Elastic Net ($\alpha = 0.5$) signature was optimal, stratifying patients into high- and low-risk groups with distinct survival outcomes ($P < 0.0001$). The high-risk group was characterized by activated proliferation and metabolic pathways, an immunosuppressive microenvironment with reduced immune infiltration and downregulated immune checkpoints, and broad-spectrum multi-drug resistance.

Conclusions: This study establishes a matrix stiffness-related prognostic signature and, through single-cell analysis, localizes matrix stiffness to myCAFs. The signature stratifies patients into distinct risk groups with different prognosis, immune microenvironments, and drug responses, providing a rationale for targeting myCAF-mediated matrix stiffness and related pathways in high-risk patients.

Keywords: breast cancer; matrix stiffness; cancer-associated fibroblasts; tumor microenvironment; machine learning; breast cancer

Introduction

Breast cancer, accounting for over 25% of all new cancer cases, has surpassed lung cancer as the most prevalent malignancy among women worldwide and remains the leading cause of cancer-related mortality(1). Its complex etiology, diverse molecular subtypes, and environmental factors contribute to profound heterogeneity, posing considerable challenges for precision medicine(2). Despite remarkable breakthroughs in immunotherapy, the concurrent emergence of drug resistance and treatment-related toxicities underscores the urgent need for strategies that can stratify patients based on their therapeutic responses(3). Intriguingly, growing evidence indicates that, beyond intrinsic genomic and transcriptomic alterations in tumor cells, the tumor micro-environment (TME) also exert a critical regulatory influence on tumor initiation and progression(4).

The extracellular matrix (ECM) constitutes one of the most prominent components of the TME. During breast cancer progression, extensive deposition and cross-linking of collagen lead to a marked increase in local matrix stiffness(5). This mechanical cue is sensed by various cell types within the TME and transduced into intracellular biochemical signals, thereby modulating cellular phenotypes such as proliferation, invasion, apoptosis, and differentiation(6).

Multiple core molecules are involved in the mechanical signal transduction pathways associated with matrix stiffness: the mechanosensitive ion channel PIEZO1 directly perceives extracellular mechanical tension(7); the RHOA–ROCK1/2 axis governs cytoskeletal contraction; and the Hippo–YAP/TAZ cascade serves as the most critical downstream effector pathway(8), wherein nuclear translocation of YAP1 and WWTR1 (TAZ) subsequently activates transcription of target genes including CTGF and CYR61, which mediate ECM remodeling and malignant transformation(9).

Cancer-associated fibroblasts (CAFs) are the most pivotal effector cells responsible for matrix remodeling in the TME. Under stimulation by tumor-derived soluble factors and mechanical signals from the stiffened matrix, normal fibroblasts undergo activation to become CAFs(10). Currently, CAFs are classified into three main subtypes: myofibroblastic CAFs (myCAFs), inflammatory CAFs (iCAFs), and antigen-presenting CAFs (apCAFs); myCAFs highly express ECM and contraction-related genes, such as COL1A1, COL1A2, and ACTA2, and dominate collagen secretion and matrix stiffening, whereas iCAFs abundantly secrete inflammatory chemokines and cytokines, and apCAFs exhibit molecular machinery for antigen presentation(11–13). These three subtypes possess distinct functional profiles and collectively contribute to the construction of the TME and the regulation of tumor progression.

Although in vitro studies using three-dimensional collagen hydrogels have demonstrated that matrix stiffness can modulate the activation status and subtype phenotypes of CAFs, large-scale clinical cohort evidence remains scarce regarding

matrix stiffness-related molecular signatures and CAF heterogeneity in patient tumor tissues(14,15). Matrix stiffness not only governs the functional phenotypes of tumor cells and CAFs but also serves as a critical driver of tumor immune microenvironment (TIME) remodeling(16). Nevertheless, a comprehensive dissection of the interplay between matrix stiffness-related transcriptomic features and the immunological landscape remains lacking in clinical breast cancer cohorts.

To address these gaps, our study aimed to construct a matrix stiffness-related prognostic signature (MSRS) and explore its relationship with CAFs heterogeneity, clinical prognosis, and immune features in breast cancer. By integrating bulk and single-cell transcriptomic data, we identified matrix stiffness-related prognostic genes and CAF subsets, and further evaluated the prognostic and immunological value of the MSRS, thereby providing clinical evidence linking matrix stiffness-related molecular features with CAF heterogeneity, prognosis, and the immune microenvironment in breast cancer.

Materials and methods

Data Collection and Processing

The overall design and analytical workflow of this study are depicted in Figure 1. The STAR-Counts transcriptome quantification data were downloaded from the GDC TCGA BRCA via the UCSC Xena platform (https://xenabrowser.net/datapages/?cohort=TCGA%20Breast%20Cancer%20(BRCA)). We obtained the unstranded raw count matrix, TPM expression matrix, and accompanying clinical information files. Ensembl IDs in the gene expression matrix were converted to gene symbols via probe annotation information. For external validation, the GSE96058 cohort was retrieved from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE96058), which includes 3273 breast cancer samples with available survival time and outcomes. Additionally, the GSE176078 dataset was also obtained from GEO (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE176078), containing 26 primary breast cancer samples profiled by the 10× Genomics platform.

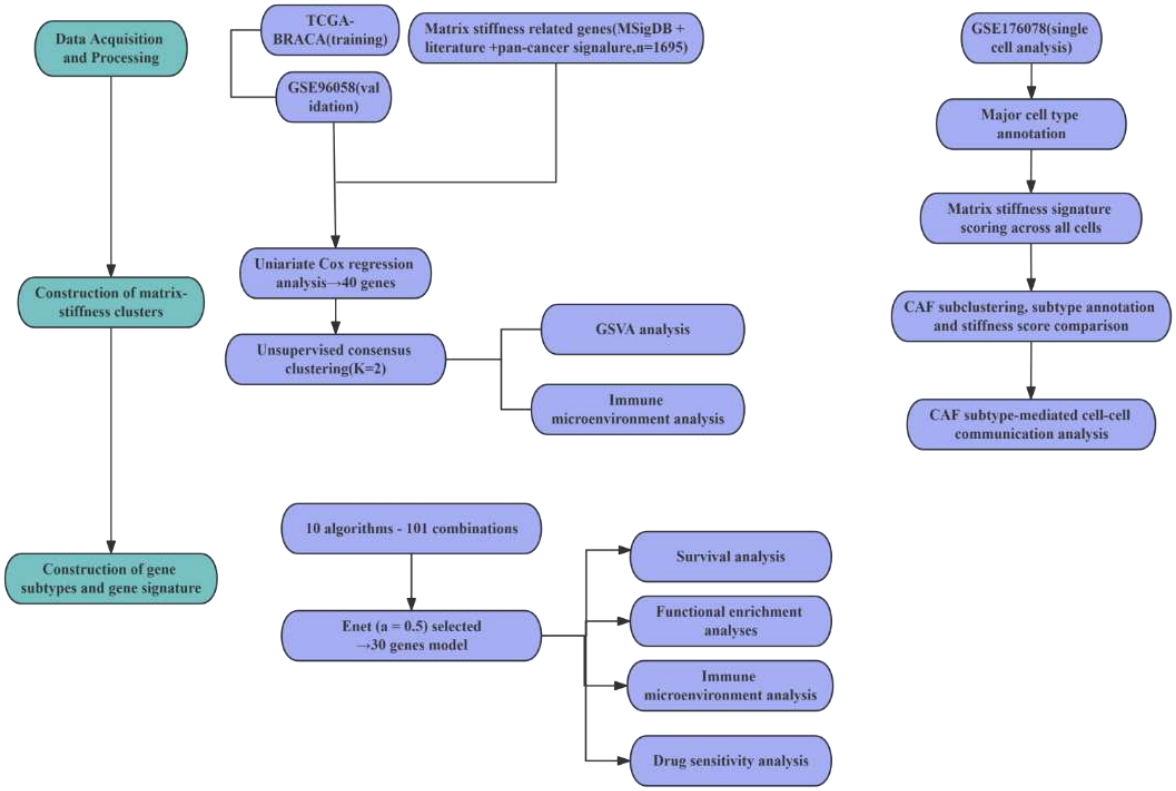


Figure 1 Workflow of the present research.

Identification of prognostic Matrix Stiffness-Related Gene Set

The matrix stiffness-related gene set was compiled by integrating the MSigDB database and previously published datasets (https://www.gsea-msigdb.org/gsea/msigdb/human/collections.jsp)(17). Matrix stiffness-related pathways were derived from MSigDB, including NABA_CORE_MATRISOME, NABA_ECM_REGULATORS, GOBP_FOCAL_ADHESION_ASSEMBLY, and GOBP_HIPPO_SIGNALING(18). Literature mining supplemented classic mechanotransduction genes, including YAP1, WWTR1 (TAZ), CCN2/CTGF, CCN1/CYR61, RHOA, ROCK1, ROCK2, and PIEZO1(19–22). In addition, eight core genes from a pan-cancer matrix stiffness signature were included: ODC1, SDC1, CREB3L1, LINC00842, CGB8, PI3, ATOH8, and ADD3(23). After removing duplicates, the final matrix stiffness-related gene set was used for downstream intersection analysis. The complete gene list is provided in Supplementary Table 1.

Differentially expressed genes (DEGs) between breast cancer and adjacent normal tissues were identified using the limma R package(24), with thresholds of adjusted P-value < 0.01, FDR < 0.05, and |log2FC| > 1. The DEGs were then intersected with the matrix stiffness-related gene set (MSRGs) to obtain overlapping genes. Univariate Cox regression analysis was subsequently performed on these overlapping genes to identify genes significantly associated with prognosis, with $P < 0.01$ considered statistically significant.

Unsupervised consensus clustering and Gene Set Variation Analysis of Matrix Stiffness-Related Subtypes

Unsupervised consensus clustering was performed on the matrix stiffness-related genes (MSRGs) to classify BRCA patients into molecular subtypes using the ConsensusClusterPlus R package(25). Clustering was repeated 500 times with a pItem value of 0.8 to assess cluster stability. Kaplan–Meier survival analysis and the log-rank test were used to compare overall survival between the resulting clusters.

Gene Set Variation Analysis (GSVA) was performed on the TPM expression matrix using the GSVA R package with the c5.go.bp.v2023.2.Hs.symbols.gmt gene set(26). Per-sample enrichment scores were calculated with a Gaussian kernel and compared between Cluster 1 and Cluster 2 using Wilcoxon rank-sum tests, with $P < 0.05$ defined as statistically significant. A heatmap was constructed to visualize differential pathway activity.

Immune infiltration and tumor microenvironment evaluation

To characterize the matrix stiffness-associated TME, we systematically profiled the immune landscape. The ESTIMATE gene signature was applied to compute immune and stromal scores for individual samples(27). The relative fractions of immune cell subsets were quantified using the online CIBERSORTx platform (https://cibersortx.stanford.edu/) with the LM22 immune signature matrix. Wilcoxon rank-sum tests were utilized to compare enrichment scores across molecular subtypes. Furthermore, inter-subtype comparisons of expression levels for 32 immune

checkpoint genes were performed using Wilcoxon rank-sum tests(28). Gene Ontology (GO) biological process enrichment analysis was conducted on matrix stiffness-related genes(29), and significantly enriched terms were visualized in a bubble diagram to display gene counts and false discovery rates (FDR).

Single-cell RNA sequencing analysis

Single-cell RNA-seq analysis was performed using the Seurat R package(30). The GSE176078 dataset (26 primary breast cancer samples, 10× Genomics) was imported with raw count matrices, gene/barcode annotations, and metadata. After barcode harmonization, cells with 200–6000 genes and <10% mitochondrial reads were retained, followed by SCTransform normalization, Harmony integration, clustering at resolution 0.5, and UMAP visualization(31). Major cell types were assigned from dataset-provided metadata and validated with canonical marker genes using DotPlot. Matrix stiffness module scores were calculated per cell using AddModuleScore(32), and their distribution across cell types was visualized by violin plots and UMAP feature plots.

CAF subtype analysis and cell–cell communication

CAFs were isolated from the full single-cell dataset via combined keyword and marker-gene filtering. The CAF subset was visualized with UMAP. Three CAF subtypes (myCAF, iCAF, and apCAF) were annotated based on the average expression of published subtype-specific gene signatures(33). Matrix stiffness module scores were calculated using AddModuleScore; score distributions were visualized with violin and UMAP plots, and inter-subtype differences were assessed using Wilcoxon rank-sum tests. Cell–cell communication was inferred with CellChat(34). CAF subtype annotations were transferred back to the full Seurat object, and ligand–receptor interactions were predicted using the human reference database CellChatDB.human. Intercellular signaling originating from CAF subtypes toward tumor epithelial and immune cells was visualized using interaction heatmaps, circular network diagrams, and ligand–receptor bubble plots.

Construction and validation of prognostic risk model

We integrated 10 machine learning algorithms and systematically evaluated 101 algorithm combinations derived from them, including single algorithms and combined feature selection-modeling pipelines(35). The algorithms comprised random survival forest (RSF), elastic net (Enet), Lasso, Ridge, stepwise Cox, CoxBoost, partial least squares regression for Cox (plsRcox), supervised principal components (SuperPC), generalized boosted regression modeling (GBM), and survival support vector machine (survival-SVM). To build a robust breast cancer prognostic signature, these 101 machine learning-based survival models were systematically compared. The TCGA cohort was used as the training dataset, and the independent GSE96058 cohort served as the external validation set. All models were evaluated and ranked by concordance index (C-index) in both cohorts.

The optimal Elastic Net (Enet) model (alpha = 0.5) was selected to construct the final prognostic risk signature based on superior predictive performance(36). Gene coefficients obtained from 10-fold cross-validation were used to calculate individualized risk scores for each patient. Patients were stratified into high- and low-risk groups according to the median risk score. Kaplan–Meier survival analysis was performed using the survival R package to assess survival differences between the two groups(37). Time-dependent ROC curves at 1-, 3-, and 5-year survival were plotted using the timeROC R package to evaluate predictive accuracy(38).

Characteristics of the Tumor Microenvironment in the High- and Low-Risk Groups

GSEA and GSVA were performed using the KEGG gene set (c5.go.bp.v2023.2.Hs.symbols.gmt) obtained from the MSigDB database. GSEA was applied to compare differential biological pathways between high- and low-risk subgroups, while GSVA was used to quantify pathway activity variations across risk groups with dedicated R packages.

TME stromal and immune infiltration levels were evaluated via the ESTIMATE

algorithm. The relative fractions of 22 immune cell subsets pre-computed by CIBERSORTx were used to compare immune cell infiltration abundance between high-risk and low-risk groups.. In addition, the expression profiles of 32 immune checkpoint genes were extracted and compared between different risk groups.

Prediction and correlation analysis of drug sensitivity

To obtain anti-tumor drug profiles, we accessed the Genomics of Drug Sensitivity in Cancer (GDSC) dataset from the Open Science Framework (OSF) repository (https://osf.io/c6tfx/)(39). Drug sensitivity was predicted using the R package oncoPredict(40). The half-maximum inhibitory concentration (IC50) for each drug was calculated for drug sensitivity comparisons. Differences in predicted IC50 between high- and low-risk groups were assessed using the Wilcoxon rank-sum test, and Spearman's correlation analysis was performed to determine the correlation between drug sensitivity and risk scores.

Statistical analysis

All statistical analyses and graphing were performed using R software (version 4.4.2) along with its associated packages. Statistical significance was denoted as *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, and ns for not significant.

Results

Identification of matrix stiffness-related molecular subtypes in BRCA

To identify matrix stiffness-related molecular subtypes in breast cancer, we first compiled a comprehensive matrix stiffness-related gene set by merging and deduplicating relevant gene sets, yielding 1,695 genes (Figure 2A). Differentially expressed genes (DEGs) between tumor and paired adjacent normal tissues were then identified and visualized in a volcano plot (Figure 2B). Intersection of these DEGs with the matrix stiffness-related gene set yielded 510 overlapping genes. Univariate Cox regression analysis further pinpointed 40 genes with significant prognostic relevance to overall survival (P < 0.01) (Figure 2C).

Based on these 40 prognosis-associated genes, unsupervised consensus clustering was

implemented. The optimal number of clusters was determined to be k = 2 according to the consensus matrix and cumulative distribution function (CDF) curves (Figure 2D, 2E). Patients were stratified into two molecular subtypes, namely Cluster 1 and Cluster 2. Kaplan–Meier survival analysis revealed that Cluster 2 exhibited markedly superior overall survival relative to Cluster 1 (p = 0.0078; Figure 2F), demonstrating that molecular subtypes defined by matrix stiffness-related genes are tightly linked to clinical prognosis.

GSVA was performed to evaluate the activity of matrix stiffness-associated biological pathways across the two subtypes. The heatmap illustrated distinct pathway enrichment patterns between Cluster 1 and Cluster 2(Figure 2G). Pathways pertaining to extracellular matrix remodeling, cell adhesion, integrin activation, cytoskeleton organization and leukocyte-endothelial adhesion were markedly enriched in Cluster 1, whereas these mechanotransduction-related pathways exhibited suppressed activity in Cluster 2, implying divergent matrix stiffness signaling and TME remodeling status between the two molecular subtypes.

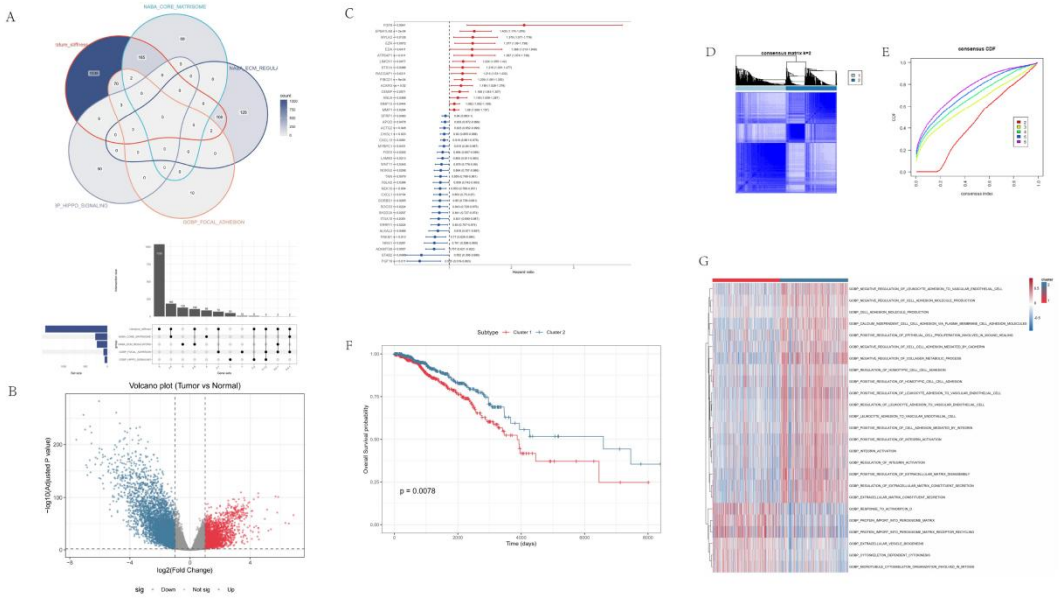


Figure 2. Identification of matrix stiffness-related molecular subtypes in BRCA. (A) UpSet plot showing the overlap of matrix stiffness-related gene sets. (B) Volcano plot of DEGs between tumor and normal tissues. (C) Forest plot of univariate Cox

regression analysis for prognostic matrix stiffness-related genes. (D, E) Consensus matrix for k = 2. (F) Consensus CDF curves for k = 2–6. (F) Kaplan–Meier survival curves of the two subtypes. (G) Heatmap of GSVA pathway activity between Cluster 1 and Cluster 2.

Landscape of the tumor immune microenvironment across matrix stiffness-defined subtypes

To dissect the immunological features of matrix stiffness-associated molecular subtypes, we comprehensively profiled the TME landscape. ESTIMATE analysis revealed that Cluster 2 exhibited markedly higher immune and stromal scores compared with Cluster 1 (Figure 3A; $P < 0.0001$). Consistent with this, ssGSEA revealed uniformly elevated enrichment scores in Cluster 2 for major immune cell populations (including dendritic cells, $CD4^+$ and $CD8^+$ T cells, B cells, plasma cells, NK cells, and monocytes) and immune functional modules (such as IFN-γ response, cytotoxicity, T cell exhaustion, immune checkpoint signature, and angiogenesis; Figure 3B). Further quantification of immune cell subsets confirmed that the majority of immune cell types were significantly enriched in Cluster 2 relative to Cluster 1 (Figure 3C; $P < 0.0001$). In line with the heightened immune infiltration, the majority of the 32 immune checkpoint genes, including PD-L1, PD-1, CTLA-4, and LAG-3, were significantly upregulated in Cluster 2 versus Cluster 1 (Figure 2D; $P < 0.0001$). Collectively, these findings demonstrate that the two subtypes harbor distinct TME phenotypes: Cluster 2 features a stroma-rich, inflamed immune microenvironment with elevated checkpoint expression, whereas Cluster 1 displays a relatively immune-excluded phenotype with suppressed ECM remodeling and immune signaling.

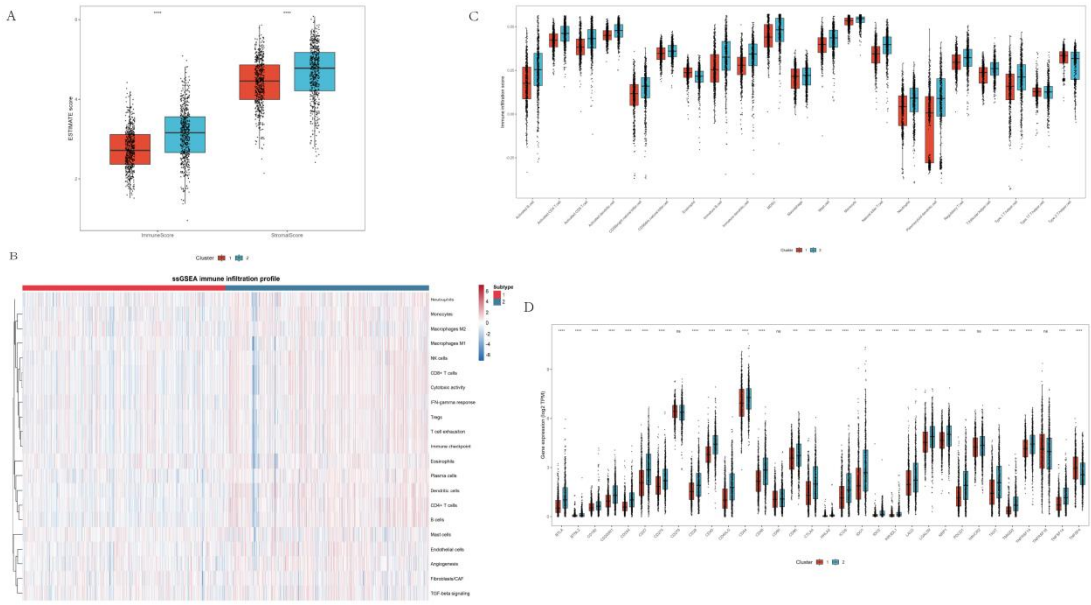


Figure 3. Characterization of tumor microenvironment and immune infiltration between two matrix stiffness-associated subtypes. (A) ESTIMATE analysis comparing immune score and stromal score between Cluster 1 and Cluster 2. (B) Heatmap showing ssGSEA scores of immune cell subsets and immune functional modules in two subtypes. (C) Boxplots illustrating the infiltration levels of 22 immune cell types (CIBERSORT) between Cluster 1 and Cluster 2. (D) Boxplots showing expression differences of 32 immune checkpoint genes between subtypes. Wilcoxon rank-sum test.

Single-cell analysis

Unsupervised UMAP dimensionality reduction was performed on the single-cell transcriptomic dataset of breast cancer, and nine major cell types were annotated using the predefined metadata from GSE176078, including cancer epithelial cells, normal epithelial cells, CAFs, endothelial cells, T cells, B cells, myeloid cells, plasmablasts, and PVL cells (Figure 4A). Marker gene expression profiles confirmed the identity of each cell population (Figure 4C). Matrix stiffness module scores were calculated for each cell and projected onto the UMAP embedding, revealing a cell type–dependent distribution (Figure 4B). Violin plots further showed that CAFs exhibited the highest stiffness scores among all cell types, followed by PVL and

endothelial cells (Figure 4D). These results indicate that matrix stiffness–related transcriptional programs are predominantly enriched in CAFs and selected stromal/immune compartments.

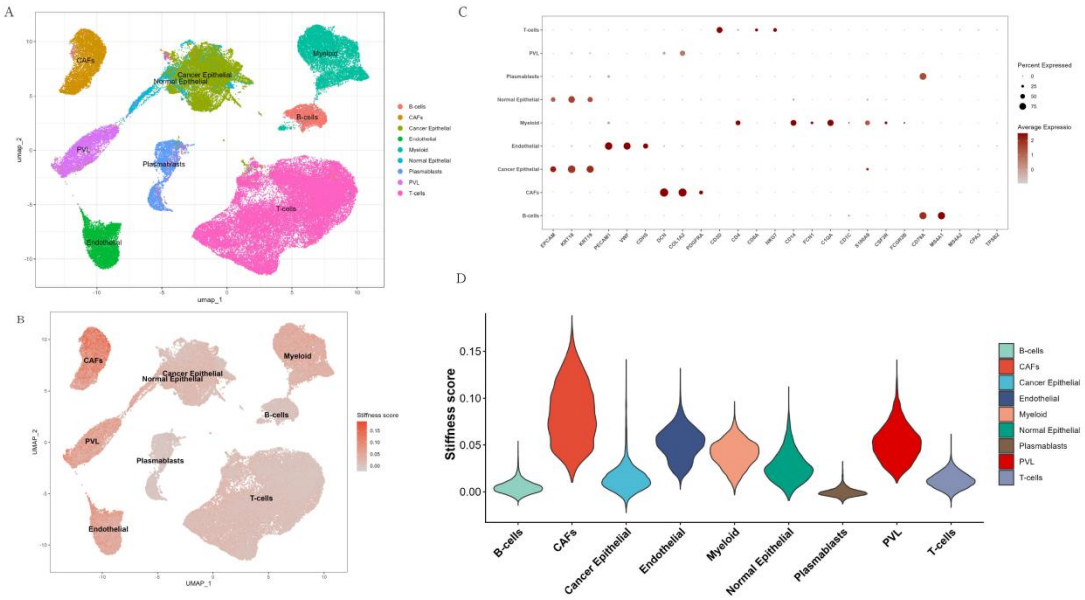


Figure 4. Single-cell landscape and matrix stiffness distribution in the breast cancer microenvironment. (A) UMAP visualization showing the nine major cell types. (B) UMAP feature plot showing the distribution of matrix stiffness module scores across all single cells. (C) Dot plot validating the expression of canonical marker genes across the identified cell populations. (D) Violin plot comparing matrix stiffness scores among distinct cell types.

To resolve CAF heterogeneity, we performed subclustering and identified three distinct subtypes: apCAF, iCAF, and myCAF (Figure 5A). Matrix stiffness scores were then projected onto the CAF subpopulations, revealing a subtype-dependent distribution (Figure 5B). Quantitative comparison demonstrated that myCAF exhibited the highest stiffness score, followed by iCAF, whereas apCAF displayed the lowest score among the three subsets (Figure 5C). These findings identify myCAF as the predominant CAF subtype contributing to matrix stiffness in the breast cancer stroma.

To systematically characterize intercellular crosstalk within the tumor microenvironment, we performed ligand–receptor interaction prediction using CellChat. Global communication network analysis revealed that CAF subsets, particularly myCAFs, served as predominant signaling sources and actively communicated with cancer epithelial and endothelial cells (Figure 5D). The interaction quantification heatmap further confirmed that myCAFs exhibited the broadest and most robust crosstalk with both cancer epithelial and endothelial compartments (Figure 5E). Focused analysis of CAF subtype-derived signaling demonstrated distinct communication patterns among apCAFs, iCAFs, and myCAFs. Notably, myCAF-mediated signaling was predominantly mediated by ECM-associated ligand–receptor pairs, including collagens, laminins, fibronectin, and thrombospondins (Figure 5F). Collectively, these findings indicate that myCAFs act as critical stromal regulators, driving ECM remodeling and mechanotransduction signaling through extensive cell–cell communication in the breast cancer microenvironment.

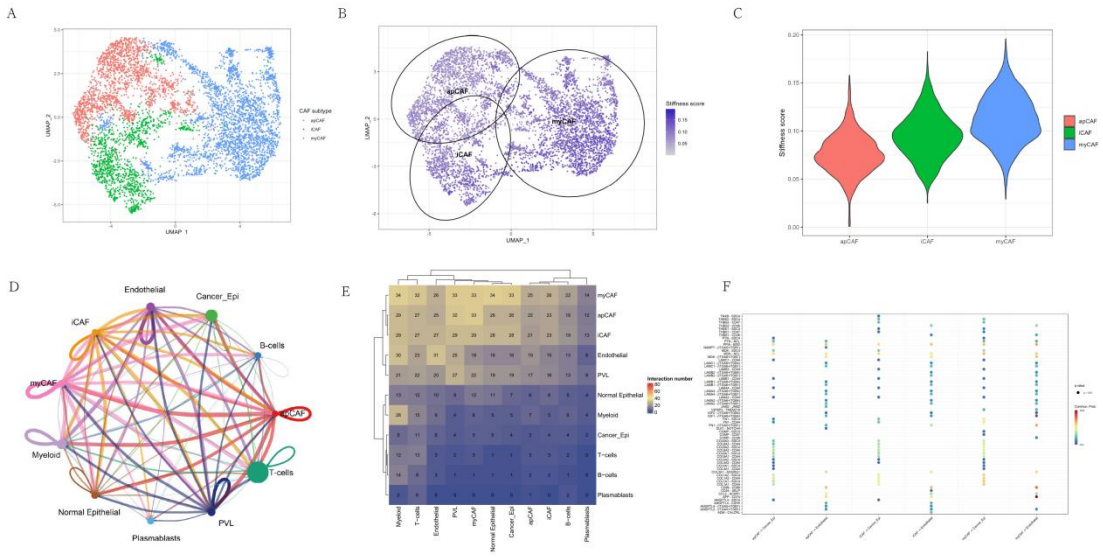


Figure 5. Identification of CAF subtypes and their matrix stiffness and cell–cell communication features. (A) UMAP of CAF subclusters. (B) UMAP feature plot of matrix stiffness scores across CAF subtypes. Ellipses indicate major distribution

regions. (C) Violin plot comparing stiffness scores among CAF subtypes. (D) CellChat circular network of cell–cell communication among all cell populations. (E) Heatmap of interaction counts between sender and receiver cell groups. (F) Bubble plot of ligand–receptor interactions from CAF subtypes to cancer epithelial and endothelial cells.

Construction and validation of a machine-learning-based prognostic risk model

To identify a robust prognostic signature for breast cancer, we systematically compared 101 machine learning-based survival models, including standalone algorithms and combined feature selection–modeling pipelines. Model performance was quantified by the C-index in the TCGA training cohort and the independent GSE96058 external validation cohort (Figure 6A). Across all models, random survival forest (RSF)-integrated models achieved the highest C-index values in the TCGA training set, with several reaching >0.98. However, these models exhibited substantial performance decay in the GSE96058 cohort, with C-index dropping to approximately 0.62, indicating poor generalizability. In contrast, Elastic Net (Enet) models demonstrated stable and consistent predictive capacity across both cohorts. Among them, the Enet signature with $\alpha = 0.5$ yielded a TCGA C-index of 0.729 and a GSE96058 C-index of 0.651, representing the optimal balance between predictive power and cross-cohort robustness. We therefore selected this Enet ($\alpha = 0.5$) model to construct the final prognostic signature. Using this algorithm with 30 matrix stiffness-related genes, patients were stratified into high- and low-risk groups.

Kaplan–Meier survival analysis demonstrated that the high-risk group had significantly worse overall survival than the low-risk group in both the TCGA training cohort ($P < 0.0001$; Figure 6B) and the GSE96058 validation cohort ($P < 0.0001$; Figure 6E). Time-dependent ROC curves were generated to assess the predictive accuracy of the Enet ($\alpha = 0.5$) signature for 1-, 3-, and 5-year overall survival. In the TCGA training cohort, the signature achieved AUC values of 0.736 (1-year), 0.756 (3-year), and 0.725 (5-year) (Figure 6C). In the independent GSE96058 validation

cohort, the corresponding AUCs were 0.694 (1-year), 0.653 (3-year), and 0.659 (5-year) (Figure 6F). The risk score distribution and survival status of patients further confirmed this stratification, with high-risk patients exhibiting higher mortality rates in both cohorts (Figure 6D, 6G). Collectively, these results demonstrated that the Enet-derived risk signature maintained reliable prognostic discrimination in both training and external validation cohorts.

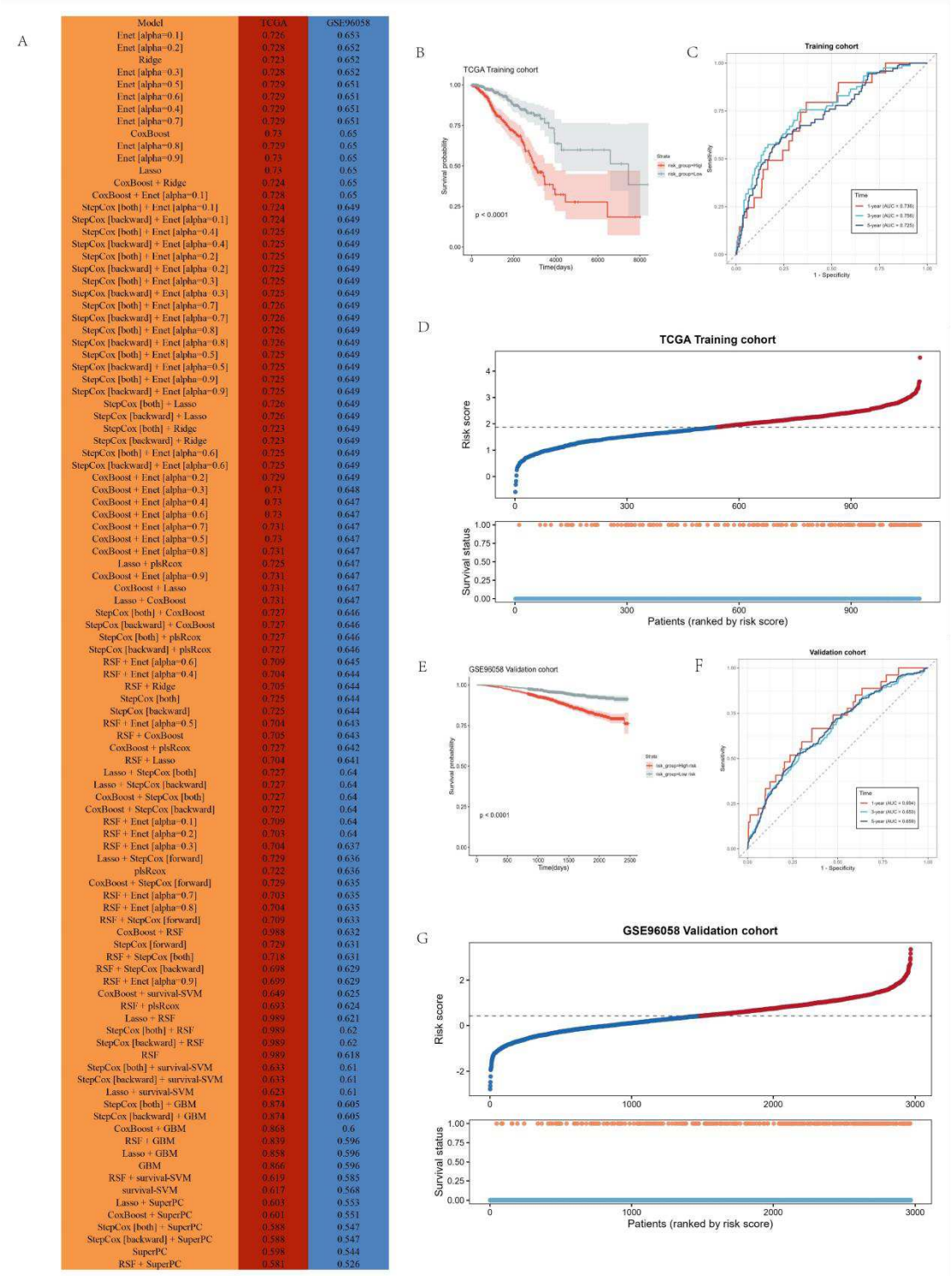


Figure 6. Construction and validation of the machine-learning-based prognostic risk model. (A) C-index comparison of 101 machine learning models in the TCGA training and GSE96058 validation cohorts. (B) Kaplan–Meier survival curves for high- and low-risk groups in the TCGA training cohort. (C) Time-dependent ROC curves at 1, 3, and 5 years in the TCGA training cohort. (D) Risk score distribution

and survival status of patients ranked by risk score in the TCGA training cohort. (E) Kaplan–Meier survival curves for high- and low-risk groups in the GSE96058 validation cohort. (F) Time-dependent ROC curves at 1, 3, and 5 years in the GSE96058 validation cohort. (G) Risk score distribution and survival status of patients ranked by risk score in the GSE96058 validation cohort.

Functional enrichment analysis of prognostic risk subgroups

GSEA based on GO-BP was performed between high- and low-risk subgroups. The high-risk group was enriched in extracellular matrix organization and structural remodeling pathways, including extracellular matrix organization, extracellular structure organization, and external encapsulating structure organization (Figure 7A). The low-risk group was enriched in adaptive humoral immune responses, such as B-cell–mediated immunity, immunoglobulin-mediated immune response, lymphocyte-mediated immunity, and humoral immune response (Figure 7B). Overlaid GSEA curves confirmed positive NES for matrix remodeling in high-risk samples and negative NES for immune signatures in low-risk samples (Figure 7C). KEGG-GSEA showed that the high-risk group was enriched in proliferation and metabolic reprogramming pathways, including cell cycle, protein processing in endoplasmic reticulum, oxidative phosphorylation, nucleocytoplasmic transport, and carbon metabolism (Figure 7E). The low-risk group was enriched in immune-associated pathways, such as chemokine signaling, cytokine–cytokine receptor interaction, Th1 and Th2 cell differentiation, hematopoietic cell lineage, and cell adhesion molecule–mediated interactions (Figure 7F). Overlaid curves showed positive NES for proliferation–metabolism gene sets in high-risk samples and negative NES for immune signaling in low-risk samples (Figure 7G). Collectively, these results link the prognostic signature to stromal remodeling and anti-tumor adaptive immunity, and distinguish patients with distinct proliferative–metabolic and immune statuses.

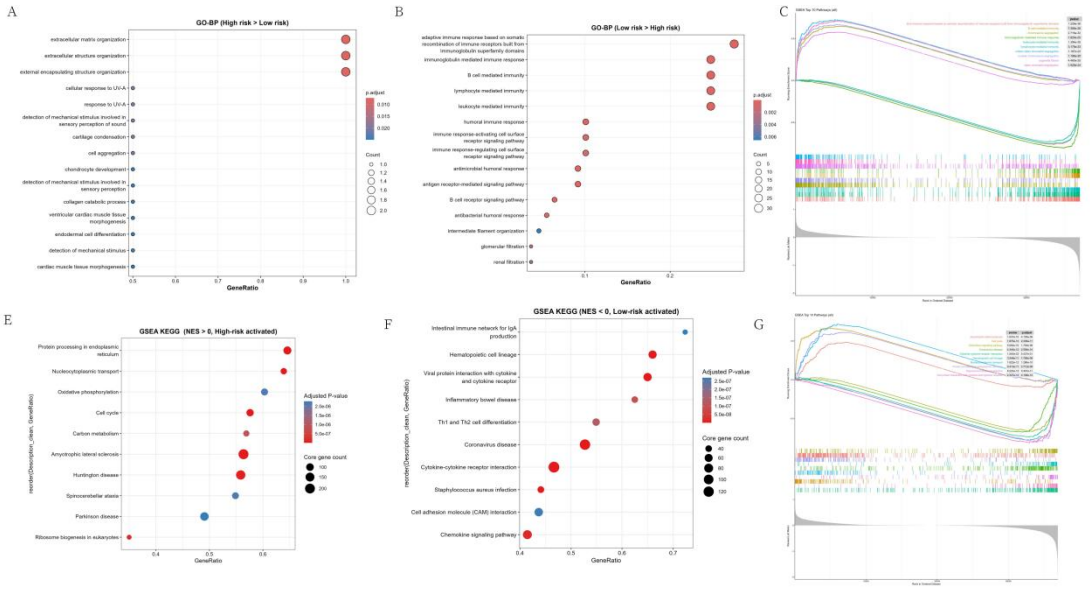


Figure 7. Functional enrichment analysis of prognostic risk subgroups. (A) GO-BP enrichment analysis in the high-risk group (positive NES). (B) GO-BP enrichment analysis in the low-risk group (negative NES). (C) GSEA curves of representative GO-BP pathways. (E) KEGG enrichment analysis in the high-risk group (positive NES). (F) KEGG enrichment analysis in the low-risk group (negative NES). (G) GSEA curves of representative KEGG pathways.

Immunological features of high- and low-risk breast cancer

The immune microenvironment plays a crucial role in tumorigenesis and response to immunotherapy. We therefore investigated the TME landscape of breast cancer patients stratified into high- and low-risk groups. The overall immune infiltration score was significantly lower in the high-risk group than in the low-risk group (Figure 8A, $P < 0.0001$). Consistent with this, CIBERSORTx analysis of 22 immune cell subsets revealed that the high-risk group had significantly higher proportions of M0 macrophages and neutrophils, and lower proportions of plasma cells, $CD8^+$ T cells, resting $CD4^+$ memory T cells, M1 and M2 macrophages, resting dendritic cells, and eosinophils (Figure 8B, all $P < 0.05$). The heatmap further visualized the distinct immune infiltration patterns between the two groups (Figure 8C). Furthermore, the expression of 32 immune checkpoint genes was compared between risk groups. The

majority of these genes, including PD-L1 (CD274), PD-1 (PDCD1), CTLA-4, LAG-3, and HAVCR2, were significantly upregulated in the low-risk group (Figure 8D, $P < 0.05$), suggesting a more active immune microenvironment in low-risk patients. In summary, high-risk breast cancer exhibits an immune-excluded phenotype characterized by reduced immune infiltration and downregulated immune checkpoint expression, which may contribute to poor prognosis and immunotherapy resistance.

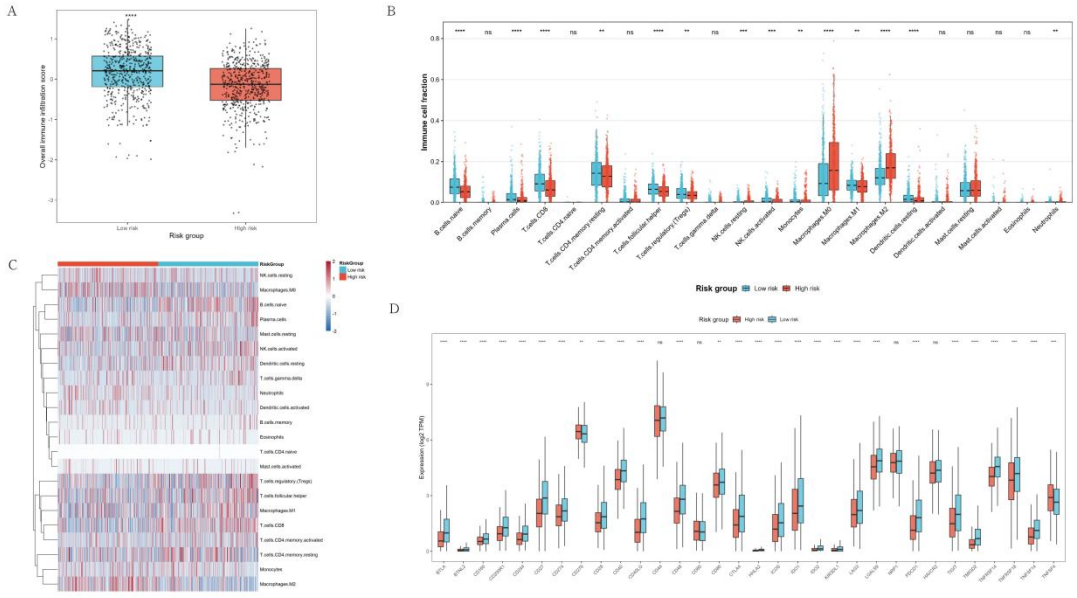


Figure 8. Immune microenvironment and immune checkpoint expression between high- and low-risk groups. (A) Overall immune infiltration score in low- and high-risk groups. (B) Relative fractions of 22 immune cell subsets (CIBERSORTx) between risk groups. (C) Heatmap of immune cell infiltration profiles. Top bar: risk group (red, high risk; cyan, low risk). (D) Expression of 32 immune checkpoint genes between risk groups.

High-risk breast cancer exhibits a broad-spectrum multi-drug resistance phenotype

Drug sensitivity was predicted across 198 drugs. Differential analysis identified 178 drugs with significantly altered IC50 between high- and low-risk groups (Wilcoxon test, $p < 0.05$), of which 164 (92.1%) showed higher IC50 in the high-risk group, indicating a multi-drug resistance phenotype, while only 14 (7.9%) showed the opposite trend (Figure 9A). The top resistant agents included DNA-damaging

chemotherapeutics (Topotecan, Mitoxantrone, Teniposide), anti-apoptotic inhibitors (Obatoclax, Sabutoclax), and kinase inhibitors targeting IGF1R, JAK1, and MRE11 (Figure 9B). Spearman correlation analysis further revealed positive correlations between risk score and IC50 for the top 20 drugs (rho = 0.309–0.394, all p < 2.22e-16), with the strongest associations for JAK1_8709, IGF1R_3801, and Irinotecan (Figure 9C, 9D). These findings demonstrate that high-risk breast cancer exhibits intrinsic multi-drug resistance, potentially linked to activated DNA replication and chromosome segregation pathways, warranting further experimental validation.

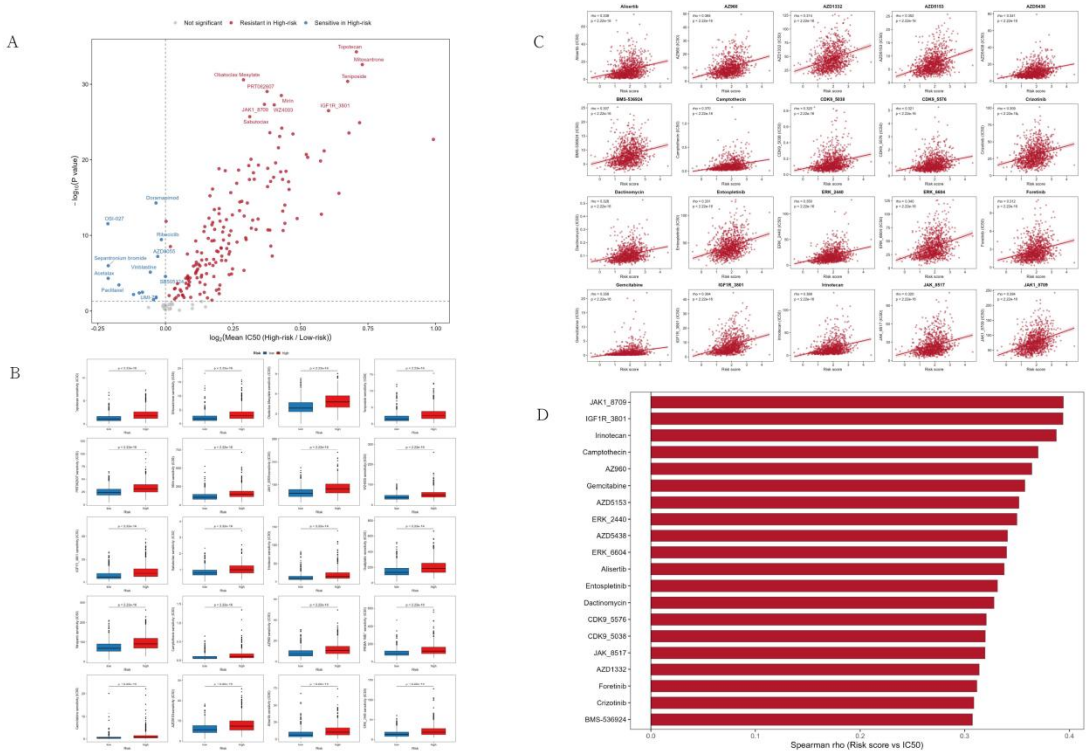


Figure 9. Drug sensitivity analysis between high- and low-risk groups. (A) Volcano plot of differentially sensitive drugs. (B) Boxplots of predicted IC50 for representative resistant drugs in high- and low-risk groups. (C) Scatter plots showing the correlation between risk score and predicted IC50 for the top 20 resistant drugs. (D) Bar chart of Spearman's rho for the top 20 drugs positively correlated with risk score.

Discussion

Breast cancer is characterized by pronounced molecular and microenvironmental

heterogeneity, which substantially compromises individualized treatment efficacy and contributes to immune resistance, recurrence, and metastasis(2). Extracellular matrix stiffness, a distinctive biomechanical property of the TME, has emerged as a central regulator of malignant progression, immune dysregulation, and therapeutic resistance in breast cancer mechanobiology(41). Although accumulating evidence has demonstrated that matrix fibrosis and aberrant stiffening can drive breast cancer progression, most previous studies have remained at the level of phenotypic observation, lacking systematic integration of biomechanical properties, immune landscape, and therapeutic responsiveness (42). To date, few robust prognostic signatures have been established to quantify matrix stiffness-associated heterogeneity, and the dominant cellular source driving mechanical remodeling, particularly regarding CAF heterogeneity in this context, as well as its clinical implications, remains poorly defined(43,44).

To address these gaps, we integrated bulk and single-cell transcriptomic data, multiple machine learning algorithms, and in silico drug sensitivity analysis to systematically dissect the molecular heterogeneity, immune microenvironment remodeling, and therapeutic response associated with matrix stiffness in breast cancer. Our main findings include: a 30-gene matrix stiffness-related prognostic signature with robust performance in both training and external validation cohorts; single-cell analysis that precisely localized matrix stiffness to the myCAF subtype and revealed myCAF-mediated ECM communication networks; and a high-risk phenotype characterized by an immunosuppressive microenvironment and broad-spectrum multidrug resistance, in contrast to the immune-activated and metabolically stable phenotype of low-risk patients.

Reliable prognostic tools are crucial for clinical decision-making in cancer. To overcome the limitations of single-algorithm modeling, we systematically compared 101 machine learning combinations to select optimal matrix stiffness-related feature genes and optimize model fitting. The final Elastic Net (α=0.5) model demonstrated favorable survival stratification in both the TCGA training cohort and the GSE96058 external validation cohort (C-index = 0.729 and 0.651, respectively). Risk scores were

significantly positively correlated with poor prognosis, indicating that aberrant matrix stiffness remodeling is tightly coupled with malignant progression in breast cancer. Moreover, the model could distinguish patient responses to chemotherapy, targeted therapy, and immunotherapy, providing a potential basis for individualized stratified treatment.

Single-cell transcriptomic analysis further revealed the cellular origin of matrix stiffness. We found that stiffness signatures were not broadly expressed across all stromal cells but were specifically enriched in myCAFs and malignant tumor cells, suggesting that myCAFs may be the core effector cells mediating matrix stiffening, mechanical signal transduction, and microenvironmental remodeling. This observation is consistent with previous studies showing that myCAFs are key mediators of ECM remodeling and mechanotransduction, which sustain a protumorigenic CAF state through feed-forward mechanosignaling loops(45,46). Cell–cell communication analysis showed that myCAFs extensively communicated with tumor epithelial and endothelial cells through aberrant ligand–receptor interactions, potentially suppressing anti-tumor immune cell function and constructing an immunosuppressive microenvironment. While CAF heterogeneity has been increasingly recognized(47,48), the differential contribution of CAF subtypes to matrix stiffness in clinical breast cancer has remained largely unexplored. Here, we quantitatively localized matrix stiffness to myCAFs and linked this to an immunosuppressive microenvironment, representing an important mechanistic innovation.

Matrix stiffness-mediated immune microenvironment imbalance may serve as a core intermediate mechanism underlying poor prognosis in high-risk patients(49). In this study, the high-stiffness, high-risk group exhibited significant enrichment of M0, M1, and M2 macrophages and neutrophils, whereas the low-risk group showed prominent infiltration of anti-tumor immune cells, including $CD8^+$ T cells, NK cells, dendritic cells, and resting mast cells, along with a higher overall immune infiltration score. These findings are consistent with previous studies: M2 macrophages promote angiogenesis, invasion, metastasis, and immune evasion through secretion of

immunosuppressive cytokines(50). Matrix stiffness can drive macrophage polarization toward the M2 phenotype through mechanotransduction cascades such as Piezo1-YAP signalling signaling(51). Notably, recent evidence indicates that M0 macrophages are not merely non-functional precursors; their abnormal enrichment is closely associated with high-grade tumors and poor prognosis, providing a novel explanation for the greater malignancy observed in the high-risk group(52). Conversely, the enriched $CD8^+$ T cells, NK cells, mature dendritic cells, and resting mast cells in the low-risk group enhance tumor immune surveillance and serve as protective factors for favorable prognosis. Immune checkpoint expression differed markedly between risk groups. The majority of the 32 immune checkpoint genes were significantly upregulated in the low-risk group, consistent with its higher immune infiltration and activated immune microenvironment(53). In contrast, the widespread downregulation of immune checkpoints in the high-risk group aligns with its immune-excluded phenotype, suggesting that checkpoint-mediated immune evasion is not the dominant resistance mechanism(54). These findings indicate that high-risk patients may derive limited benefit from immune checkpoint inhibitor monotherapy, whereas strategies targeting matrix stiffness remodeling or other immunosuppressive pathways may be more rational(55).

GSEA revealed distinct pathway activation patterns between the two risk groups. The high-risk group was characterized by enrichment of EMT, cell cycle, and matrix remodeling pathways, which are known to promote tumor proliferation, invasion, and resistance to chemotherapy and targeted therapy(56–58), consistent with the aggressive phenotype and poor prognosis observed in these patients. In contrast, the low-risk group was enriched in immune activation and metabolic homeostasis pathways, suggesting a more favorable anti-tumor immune status(58). These pathway-level differences link matrix stiffness remodeling to altered tumor signaling, providing a molecular basis for the prognostic disparities and microenvironmental dysregulation in breast cancer.

Drug sensitivity analysis revealed a broad-spectrum multi-drug resistance phenotype in the high-risk group. The top resistant agents included DNA-damaging

chemotherapeutics and anti-apoptotic inhibitors, consistent with enhanced DNA repair and apoptotic evasion in aggressive breast cancer(59). Kinase inhibitors targeting IGF1R and JAK1 also showed elevated IC50, supporting the role of alternative survival signaling in resistance(60,61). These findings suggest that combination strategies targeting DNA repair, anti-apoptotic pathways, or kinase signaling may be explored to overcome intrinsic resistance in high-risk patients.

To our knowledge, this is the first study to construct a matrix stiffness-related prognostic signature and localize matrix stiffness to myCAFs in breast cancer. Unlike previous studies focusing on conventional genomic or metabolic markers, we translated a biomechanical property of the tumor microenvironment into a gene expression-based prognostic model. By integrating ten machine learning algorithms, we improved model robustness and generalizability. Furthermore, by combining bulk transcriptomics for clinical prognosis with single-cell profiling for mechanistic exploration, we identified CAF subtypes associated with matrix stiffness and localized myCAFs as the predominant cellular source, thereby linking clinical risk stratification with mechanistic insights into matrix stiffness-associated heterogeneity.

However, several limitations remain. First, all analyses relied on retrospective public datasets, requiring further validation in clinical samples and experimental models. Second, the biological functions and regulatory mechanisms of the core matrix stiffness-related genes, as well as their involvement in CAF–immune cell crosstalk, need further investigation. Third, the clinical predictive value of our model requires validation in prospective multicenter cohorts. Despite these limitations, our findings highlight potential prognostic markers and therapeutic targets, particularly myCAF-mediated matrix stiffness, for future breast cancer research.

Conclusion

In summary, this study developed a matrix stiffness-related prognostic signature and, by integrating bulk and single-cell transcriptomic analyses, identified myCAFs as the predominant cellular source of matrix stiffness in breast cancer. The signature effectively stratified patients into distinct risk groups differing in prognosis, immune microenvironment, and drug response. High-risk patients exhibited an

immunosuppressive microenvironment and broad-spectrum multidrug resistance, whereas low-risk patients displayed activated anti-tumor immunity and metabolic homeostasis. These findings establish a link between matrix stiffness remodeling and CAF heterogeneity, immune dysregulation, and therapeutic resistance, offering a rationale for targeting myCAF-mediated matrix stiffness in high-risk breast cancer. Prospective cohort studies and experimental validation are warranted to translate these insights into clinical practice.

Declaration

1. Ethical Approval and Participant Consent

Not applicable

2. Consent for Publication

Not applicable

3. Data Availability Statement

The datasets generated and/or analyzed during the current study are available in the following public repositories:

- TCGA-BRCA: The STAR-Counts transcriptome quantification data and clinical information were downloaded from the UCSC Xena platform (https://xenabrowser.net/datapages/?cohort=TCGA%20Breast%20Cancer%20(BRCA)).
- GEO: The external validation datasets were obtained from the Gene Expression Omnibus (GEO) database. The GSE96058 cohort is available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE96058, and the GSE176078 dataset (single-cell RNA-seq) is available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE176078.
- MSigDB: Matrix stiffness-related pathways and gene sets, including c5.go.bp.v2023.2.Hs.symbols.gmt, were derived from the MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/human/collections.jsp).
- CIBERSORTx: The relative fractions of immune cell subsets were quantified using the CIBERSORTx web platform (https://cibersortx.stanford.edu/) with the LM22 immune signature matrix.

- GDSC & OSF: The anti-tumor drug profiles and drug sensitivity data were obtained from the GDSC dataset, which is openly available in the OSF repository(https://osf.io/c6tfx/).

4. Competing Interests

All authors have confirmed that there are no competing interests.

5. Funding

No.

6. Authors' Contributions

All authors have made substantial contributions to the conception, design, data analysis, manuscript writing, or revision. The specific contributions are as follows:

Shen Pu: concept of theme, illustration design. statistical analysis and manuscript writing.

Chen Yunru: statistical analysis and data visualization.

Yang Meng: Corresponding Author.

Reference

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA A Cancer J Clinicians. 2021 May;71(3):209–49. doi:10.3322/caac.21660

2. Barzaman K, Karami J, Zarei Z, Hosseinzadeh A, Kazemi MH, Moradi-Kalbolandi S, et al. Breast cancer: Biology, biomarkers, and treatments. International Immunopharmacology. 2020 Jul;84:106535. doi:10.1016/j.intimp.2020.106535

3. Wu GY, Xiao MZ, Hao WC, Yang ZS, Liu XR, Xu DS, et al. Drug resistance in breast cancer: Mechanisms and strategies for management. Drug Resistance Updates. 2025 Nov;83:101288. doi:10.1016/j.drup.2025.101288

4. Kundu M, Butti R, Panda VK, Malhotra D, Das S, Mitra T, et al. Modulation of the tumor microenvironment and mechanism of immunotherapy-based drug resistance in breast cancer. Mol Cancer. 2024 May 7;23(1):92. doi:10.1186/s12943-024-01990-4

5. Herrera-Quintana L, Vázquez-Lorente H, Plaza-Diaz J. Breast Cancer: Extracellular Matrix and Microbiome Interactions. Int J Mol Sci. 2024 Jun 30;25(13):7226. doi:10.3390/ijms25137226 PubMed PMID: 39000333; PubMed Central PMCID: PMC11242809.

6. Pickup MW, Mouw JK, Weaver VM. The extracellular matrix modulates the hallmarks of cancer. EMBO Rep. 2014 Dec 1;15(12):1243–53. doi:10.15252/embr.201439246

7. Tang H, Zeng R, He E, Zhang I, Ding C, Zhang A. Piezo-Type Mechanosensitive Ion Channel Component 1 (Piezo1): A Promising Therapeutic Target and Its Modulators. J Med Chem. 2022 May 12;65(9):6441–53. doi:10.1021/acs.jmedchem.2c00085 PubMed PMID: 35466678.

8. Kiang KM, Ahad L, Zhong X, Lu QR. Biomolecular condensates: hubs of Hippo-YAP/TAZ signaling in cancer. Trends Cell Biol. 2024 Jul;34(7):566–77. doi:10.1016/j.tcb.2024.04.009 PubMed PMID: 38806345.

9. Yu H, Qu T, Yang J, Dai Q. Serotonin acts through YAP to promote cell proliferation: mechanism and implication in colorectal cancer progression. Cell Commun Signal. 2023 Apr 12;21(1):75. doi:10.1186/s12964-023-01096-2

10. Jia H, Chen X, Zhang L, Chen M. Cancer associated fibroblasts in cancer development and therapy. J Hematol Oncol. 2025 Mar 28;18(1):36. doi:10.1186/s13045-025-01688-0

11. Thorlacius-Ussing J, Jensen C, Nissen NI, Cox TR, Kalluri R, Karsdal M, et al. The collagen landscape in cancer: profiling collagens in tumors and in circulation reveals novel markers of cancer-associated fibroblast subtypes. The Journal of Pathology. 2024 Jan;262(1):22–36. doi:10.1002/path.6207

12. Ye J, Baer JM, Faget DV, Morikis VA, Ren Q, Melam A, et al. Senescent CAFs Mediate Immunosuppression and Drive Breast Cancer Progression. Cancer Discovery. 2024 Jul 1;14(7):1302–23. doi:10.1158/2159-8290.CD-23-0426

13. Song J, Wei R, Liu C, Zhao Z, Liu X, Wang Y, et al. Antigen-presenting cancer associated fibroblasts enhance antitumor immunity and predict immunotherapy response. Nat Commun. 2025 Mar 4;16(1):2175. doi:10.1038/s41467-025-57465-7

14. Gao H, Tian Q, Zhu L, Feng J, Zhou Y, Yang J. 3D Extracellular Matrix Regulates the

Activity of T Cells and Cancer Associated Fibroblasts in Breast Cancer. Front Oncol. 2021 Dec 9;11:764204. doi:10.3389/fonc.2021.764204

15. Cao H, Cheng HS, Wang JK, Tan NS, Tay CY. A 3D physio-mimetic interpenetrating network-based platform to decode the pro and anti-tumorigenic properties of cancer-associated fibroblasts. Acta Biomater. 2021 Sep 15;132:448–60. doi:10.1016/j.actbio.2021.03.037 PubMed PMID: 33766799.

16. Zhang M, Zhang B. Extracellular matrix stiffness: mechanisms in tumor progression and therapeutic potential in cancer. Exp Hematol Oncol. 2025 Apr 10;14(1):54. doi:10.1186/s40164-025-00647-2

17. Li H, Zhang J, Shi Y, Wang H, Yang R, Wu S, et al. Identification of matrix stiffness-related molecular subtypes in HCC via integrating multi-omics analysis and machine learning algorithms. J Transl Med. 2025 Jul 1;23(1):716. doi:10.1186/s12967-025-06733-7 PubMed PMID: 40598539; PubMed Central PMCID: PMC12329936.

18. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011 Jun 15;27(12):1739–40. doi:10.1093/bioinformatics/btr260

19. Zhu B, Li F, Yu J, Liang Z, Ke X, Wang Y, et al. PIEZO1 mediates matrix stiffness-induced tumor progression in kidney renal clear cell carcinoma by activating the Ca2+/Calpain/YAP pathway. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2025 Jan;1872(1):119871. doi:10.1016/j.bbamcr.2024.119871

20. Peng Y, Chen Z, Chen Y, Li S, Jiang Y, Yang H, et al. ROCK isoforms differentially modulate cancer cell motility by mechanosensing the substrate stiffness. Acta Biomaterialia. 2019 Apr;88:86–101. doi:10.1016/j.actbio.2019.02.015

21. Reid SE, Kay EJ, Neilson LJ, Henze A, Serneels J, McGhee EJ, et al. Tumor matrix stiffness promotes metastatic cancer cell interaction with the endothelium. EMBO J. 2017 Aug 15;36(16):2373–89. doi:10.15252/embj.201694912

22. Dupont S, Morsut L, Aragona M, Enzo E, Giulitti S, Cordenonsi M, et al. Role of YAP/TAZ in mechanotransduction. Nature. 2011 Jun;474(7350):179–83. doi:10.1038/nature10137

23. Tang G, Liu X, Li Y, Ta Y, Cho M, Li H, et al. Pan-cancer analysis reveals molecular signatures for predicting matrix stiffness in solid tumors. Intl Journal of Cancer. 2026 Feb

15;158(4):1093–105. doi:10.1002/ijc.70175

24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 2015 Apr 20;43(7):e47–e47. doi:10.1093/nar/gkv007

25. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010 Jun 15;26(12):1572–3. doi:10.1093/bioinformatics/btq170

26. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics. 2013 Dec;14(1):7. doi:10.1186/1471-2105-14-7

27. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013 Oct 11;4(1):2612. doi:10.1038/ncomms3612

28. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. Immunity. 2018 Apr;48(4):812-830.e14. doi:10.1016/j.immuni.2018.03.023

29. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000 May;25(1):25–9. doi:10.1038/75556

30. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 2024 Feb;42(2):293–304. doi:10.1038/s41587-023-01767-y

31. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol. 2019 Jan;37(1):38–44. doi:10.1038/nbt.4314

32. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016 Apr 8;352(6282):189–96. doi:10.1126/science.aad0501

33. Elyada E, Bolisetty M, Laise P, Flynn WF, Courtois ET, Burkhart RA, et al. Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. Cancer Discovery. 2019 Aug 1;9(8):1102–23. doi:10.1158/2159-8290.CD-19-0094

34. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and

analysis of cell-cell communication using CellChat. Nat Commun. 2021 Feb 17;12(1):1088. doi:10.1038/s41467-021-21246-9

35. Liu Z, Liu L, Weng S, Guo C, Dang Q, Xu H, et al. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. Nat Commun. 2022 Feb 10;13(1):816. doi:10.1038/s41467-022-28421-6

36. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Soft. 2010;33(1). doi:10.18637/jss.v033.i01

37. Therneau TM. survival: Survival Analysis [Internet]. 2001 [cited 2026 Oct 4]. p. 3.8-12. Available from: https://CRAN.R-project.org/package=survival doi:10.32614/CRAN.package.survival

38. Blanche P, Dartigues J, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Statistics in Medicine. 2013 Dec 30;32(30):5381–97. doi:10.1002/sim.5958

39. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Research. 2012 Nov 22;41(D1):D955–61. doi:10.1093/nar/gks1111

40. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting *in vivo* or cancer patient drug response and biomarkers from cell line screening data. Briefings in Bioinformatics. 2021 Nov 5;22(6):bbab260. doi:10.1093/bib/bbab260

41. Parra JPRLDL, Abuná RPF, Garcia MHH, Barbalho SM, Miglino MA. Extracellular Matrix Remodeling as a Mechanobiological Driver of Breast Cancer Aggressiveness: Comparative Oncology, Multi-Omics, and Artificial Intelligence Perspectives. Biology. 2026 Jul 16;15(14):1164. doi:10.3390/biology15141164

42. Piersma B, Hayward MK, Weaver VM. Fibrosis and cancer: A strained relationship. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2020 Apr;1873(2):188356. doi:10.1016/j.bbcan.2020.188356

43. Knipper K, Lyu S, Quaas A, Bruns C, Schmidt T. Cancer-Associated Fibroblast Heterogeneity and Its Influence on the Extracellular Matrix and the Tumor Microenvironment. IJMS. 2023 Aug 30;24(17):13482. doi:10.3390/ijms241713482

44. Shi W, Wang H, Guan Y, Zhou Y, Cui J, Wang X, et al. Novel insights into triple-negative

breast cancer heterogeneity, prognosis, and treatment response based on matrix stiffness: a combined single-Cell and transcriptome analysis. Front Oncol. 2026 May 7;16:1821887. doi:10.3389/fonc.2026.1821887

45. Virumbrales-Muñoz M, Fertal SA, Lugo-Cintrón KM, Loken JR, Numan-Vazquez Y, Burkel BM, et al. Laminin – myCAF interplay promotes α6-integrin-mediated breast cancer cell invasion. Lab Chip. 2026;26(19):5216–31. doi:10.1039/D5LC00711A

46. Kim HS, Nam JS. The multifaceted role of YAP in the tumor microenvironment and its therapeutic implications in cancer. Exp Mol Med. 2025 Oct 1;57(10):2201–13. doi:10.1038/s12276-025-01551-9

47. Liu L, Wan W, Chang Y, Ao L, Xu Y, Xu X. Crosstalk between heterogeneous cancer-associated fibroblast subpopulations and the immune system in breast cancer: key players and promising therapeutic targets. J Exp Clin Cancer Res. 2025 Sep 1;44(1):263. doi:10.1186/s13046-025-03527-z

48. Do KV, Tran AV, Pham AD, Mac TT, Pham TL, Do HN. Molecular remodeling of cancer-associated fibroblasts in breast cancer patients receiving anti–PD-1 immunotherapy. Front Oncol. 2026 Feb 24;16:1754311. doi:10.3389/fonc.2026.1754311

49. Guo S, Yuan F, Qin Y, Ji X, Wu X, Zheng H, et al. Fibrotic mechanobiological microenvironments drive mechano-immune exclusion and immunotherapy resistance in breast cancer: mechanisms, stratification and therapeutic opportunities. International Immunopharmacology. 2026 Nov;189:117456. doi:10.1016/j.intimp.2026.117456

50. Nasir I, McGuinness C, Poh AR, Ernst M, Darcy PK, Britt KL. Tumor macrophage functional heterogeneity can inform the development of novel cancer therapies. Trends in Immunology. 2023 Dec;44(12):971–85. doi:10.1016/j.it.2023.10.007

51. Mei F, Guo Y, Wang Y, Zhou Y, Heng BC, Xie M, et al. Matrix stiffness regulates macrophage polarisation via the PIEZO1-YAP signalling axis. Cell Proliferation. 2024 Aug;57(8):e13640. doi:10.1111/cpr.13640

52. Zhang Y, Zou J, Chen R. An M0 macrophage-related prognostic model for hepatocellular carcinoma. BMC Cancer. 2022 Dec;22(1):791. doi:10.1186/s12885-022-09872-y

53. Gu Y, Feng Z, Xu X, Jin L. Identification of a novel immune-related gene signature by single-cell and bulk sequencing for the prediction of the immune landscape and prognosis of

breast cancer. Cancer Cell Int. 2024 Dec 3;24(1):393. doi:10.1186/s12935-024-03589-7

54. Wu B, Thant W, Bitman E, Liu T, Liu J, Paschalis EI, et al. TROP2/claudin program mediates immune exclusion to impede checkpoint blockade in breast cancer. J Immunother Cancer. 2026 Apr 3;14(4):e012265. doi:10.1136/jitc-2025-012265 PubMed PMID: 41932810; PubMed Central PMCID: PMC13052784.

55. Shen H, Zheng Q, Wang Z, Zheng D, Gong Z, Wang H, et al. Breaking the heterogeneity barrier: a robust prognostic signature for survival stratification and immune profiling in triple-negative breast cancer. Front Immunol. 2025;16:1611917. doi:10.3389/fimmu.2025.1611917 PubMed PMID: 41098738; PubMed Central PMCID: PMC12518079.

56. Błaszczak E, Miziak P, Odrzywolski A, Baran M, Gumbarewicz E, Stepulak A. Triple-Negative Breast Cancer Progression and Drug Resistance in the Context of Epithelial-Mesenchymal Transition. Cancers (Basel). 2025 Jan 12;17(2):228. doi:10.3390/cancers17020228 PubMed PMID: 39858010; PubMed Central PMCID: PMC11764116.

57. Hawes ML, Moody MA, McCauley CR, Huddleston AG, Solanky M, Khosravi DH, et al. Oncogenic effects of ECM remodeling in obesity and breast cancer. Oncogene. 2025 Sep 29;44(37):3409–21. doi:10.1038/s41388-025-03521-x

58. Ren D, Xue D, Hou Z, Xu H, Li X. ST6GAL1 Promotes Epithelial-to-Mesenchymal Transition in Breast Cancer via the HIF-HK2 Signaling Pathway. BCTT. 2025 Dec;Volume 17:1199–211. doi:10.2147/BCTT.S555609

59. Bhavsar PP, Kalita B, Taunk K, Rapole S. Decoding the key hallmarks of chemoresistance: A proteomic tale from breast cancer research. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2025 Oct;1880(5):189404. doi:10.1016/j.bbcan.2025.189404

60. Guven DC, Ahmed J, Stephen B, Naing A. IGF-1R inhibitors in cancer: A review of available evidence and future outlook. Critical Reviews in Oncology/Hematology. 2025 Oct;214:104809. doi:10.1016/j.critrevonc.2025.104809

61. Vargas C, Aguirre-Ducler A, Cereceda K, Quijada S, Escobar-Gómez N, Castillo RL, et al. CD44 Marks Dormant Tumor Cells After HER2 Inhibition in Breast Cancer Cells. IJMS. 2025 May 20;26(10):4907. doi:10.3390/ijms26104907

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarytable1.xlsx