Journal of Translational Medicine

https://doi.org/10.1186/s12967-026-08446-x

Article in Press

Single-cell profiling reveals a novel CAF subpopulation linking stromal heterogeneity to immune suppression in breast cancer subtypes

Received: 22 February 2026

Accepted: 5 June 2026

Published online: 25 June 2026

Cite this article as: Vigiano Benedetti V., Conte F., Santoni D. *et al.* Single-cell profiling reveals a novel CAF subpopulation linking stromal heterogeneity to immune suppression in breast cancer subtypes. *J Transl Med* (2026). https://doi.org/10.1186/s12967-026-08446-x

Virginia Vigiano Benedetti, Federica Conte, Daniele Santoni, Tancredi Massimo Pentimalli, Roberto Bei, Loredana Cifaldi, Giovanni Barillari, Francesco Spallotta, Chiara Cencioni & Ombretta Melaiu

Single-cell profiling reveals a novel CAF subpopulation linking stromal heterogeneity to immune suppression in breast cancer subtypes

Authors: Virginia Vigiano Benedetti[1,2]*, Federica Conte[1]*, Daniele Santoni[1], Tancredi Massimo Pentimalli[3], Roberto Bei[4], Loredana Cifaldi[4], Giovanni Barillari[4], Francesco Spallotta[5,6], Chiara Cencioni[1,§], Ombretta Melaiu[4,§]

[1]Institute of System Analysis and Informatics "Antonio Ruberti", National Research Council (IASI-CNR), 00185, Rome, Italy.

[2] Department of Translational Medicine, Catholic University of the Sacred Heart, 00168 Rome, Italy.

[3]Laboratory for Systems Biology of Regulatory Elements, Berlin Institute for Medical Systems Biology (BIMSB), Max-Delbrück-Centrum for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany; Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin.

[4]Department of Clinical Sciences and Translational Medicine, University of Rome Tor Vergata, Rome, 00133, Rome, Italy.

[5]Department of Biology and Biotechnologies 'Charles Darwin', Sapienza University of Rome, Laboratory affiliated to Istituto Pasteur Italia-Fondazione Cenci Bolognetti, 00185, Rome, Italy.

[6]Istituto Pasteur Italia-Fondazione Cenci Bolognetti, Sapienza University, 00185, Rome, Italy.

* Correspondence: Ombretta.Melaiu@uniroma2.it; chiara.cencioni@cnr.it

* These Authors equally contributed to the work as first authors

§ These Authors equally contributed to the work as last authors

Abstract

Background: The tumor microenvironment critically influences breast cancer (BC) progression, immune surveillance, and therapeutic response. Cancer-associated fibroblasts (CAFs), a heterogeneous stromal population, are key regulators of these processes, yet their subtype-specific contributions in BC remain insufficiently defined.

Methods: We integrated three single-cell RNA sequencing datasets from 29 BC patients to characterize stromal populations. Bulk RNA-seq data from The Cancer Genome Atlas (TCGA) were analyzed to assess correlations between CAF subsets and immune infiltration. Gene signatures were derived to identify subtype-specific CAF–immune interactions, prognostic markers, and potential predictors of chemotherapy response.

Results: Three conserved stromal populations (iCAFs, myCAFs, and pericytes) were identified, along with a previously unrecognized subset, the

cluster 3 (CL3) CAF-like cells, referred as metabolic stressed CAF (msCAF). msCAF cells displayed transcriptional programs associated with antigen presentation, stress response, glycolysis, and extracellular matrix remodeling. Their abundance was inversely correlated with T-cell infiltration and function, in a subtype-specific manner: triple negative breast cancer (TNBC) was enriched for msCAFs in immune-infiltrated but functionally constrained microenvironments, whereas Luminal A tumors exhibited weaker immune infiltration with heterogeneous CAF–immune associations. msCAFs were characterized by a conserved gene signature (HLA-A, HLA-C, IL32, EMP3) and subtype-specific genes related to T-cell exhaustion. Several genes demonstrated prognostic relevance with distinct patterns in Luminal A (IER3, TIMP1, TBX3, SEC61G) and TNBC (ADM, C4orf3, LDHA) tumors, as well as shared biomarkers (FN1, LOXL2, P4HA1). Multiple msCAF genes also predicted chemotherapy response, suggesting utility as treatment stratification biomarkers.

Conclusion: msCAFs represent a clinically relevant CAF subset that drives immune suppression, impacts subtype-specific prognosis, and influences therapy response in BC. These findings highlight msCAFs as promising targets for enhancing immunotherapy and personalizing treatment strategies.

Keywords: Breast Cancer, cancer-associated fibroblasts, T cells, immunomodulation, single-cell RNA-seq.

Introduction

The tumor microenvironment (TME) is a highly dynamic ecosystem supporting tumor development, progression and resistance to therapy. It is composed of malignant and non-malignant cells, including fibroblasts, endothelial cells and immune cell populations, signaling molecules, and extracellular matrix (ECM) (1). Cancer-associated fibroblasts (CAFs), a heterogenous population of activated fibroblasts residing in the TME, exert a pro-tumorigenic role orchestrating ECM deposition, secreting growth factors and cytokines, inducing angiogenesis, and modulating anti-tumoral

responses (2). Recently, they attracted scientific interest as critical regulators of tumor biology and potential targets for cancer therapy (2).

Breast cancer (BC), one of the malignancies most enriched in CAFs (3), provides an ideal model for investigating the influence of stromal cells on tumor behaviour. BC is a significant global health burden, representing the most frequently diagnosed cancer and the second leading cause of cancer-related death among women worldwide (4). Its progression and therapeutic outcomes strictly depend on cellular composition and dynamic interactions within the TME (5). Despite recent advances in early detection and treatment, including screening campaigns, chemotherapy protocols, endocrine treatment, HER2-targeted therapies, and precision surgery, drug resistance, recurrence, and metastasis formation continue to limit long-lasting clinical benefit (6). BC is broadly classified into four molecular subtypes based on gene expression profiles and hormone receptor status. Luminal A, luminal B, HER2-enriched, and triple-negative BC (TNBC)/basal-like are primary molecular subclasses listed from less to more aggressive subtype, respectively (7). Substantial BC inter- and intra-tumoral heterogeneity drives the variability in patient outcomes, with CAFs conditioning the structural and molecular landscape of TME (3).

CAF origins in BC are complex and multifactorial, involving contributions from several precursor cell types through diverse molecular pathways (8–10). Unambiguous identification of human breast CAFs is still an intricate and complex matter, due to their ample heterogeneity and plasticity. Recently, compelling evidence showed that CAFs belong to a non-homogenous population, subcategorized into four distinct CAF subsets according to the expression of key markers derived from flow cytometry analysis (11). Typical CAF markers analysed are: α-smooth muscle actin (α-SMA), fibroblast activation protein (FAP), fibroblast-specific protein 1 (FSP1/S100A4), platelet-derived growth factor receptor (PDGFR) α (11). Of note, these different identified CAFs (from S1 to S4 subset) showed not only a specific flow cytometric pattern, but also peculiar functional features (11,12). Indeed,

S1-CAF subset exerts immunosuppressive functions, by recruiting and entrapping T lymphocytes, sustaining their differentiation into regulatory T-cells (Treg) (12,13). Instead, S4-CAF subset, characterized by high positivity to αSMA paralleled to low positivity to FAP, supports invasiveness and metastatic properties of BC cells (11). Moreover, S1-S4 CAF subpopulations characterize the four BC subtypes for their abundance and spatial distribution, suggesting how stromal heterogeneity might affect cancer progression, anti-tumoral response and therapy outcomes. In this light, phenotypic and functional heterogeneity of CAFs might be exploited for prognostic and therapeutic purposes aimed to target the TME.

The interplay between CAFs and the tumor immune microenvironment (TIME) (14) leads to an immunologically “cold” microenvironment, hindering an effective anti-tumor immune response. This bidirectional communication between CAFs and immune cells, including T lymphocytes, macrophages, and myeloid-derived suppressor cells (MDSCs), arises from continuous soluble mediator release and direct cell-to-cell contact (11,15,16), and might result in T cell exhaustion induction with progressive loss of effector function and sustained expression of inhibitory checkpoint molecules (i.e. PD-1 and CTLA-4) (17). Breast CAFs hamper T cell function by secretion of immunosuppressive cytokines, ECM remodeling, and modulation of chemokine gradients impairing tumor T cell infiltration. Indeed, CAF-rich BC often exhibits reduced numbers of cytotoxic $CD8^+$ T cells (18,19), altered PD-L1 expression, and resistance to immune checkpoint blockade (ICB) therapies (20). Moreover, increased ECM deposition is associated with CAF-mediated T-cell exclusion, further consolidating their role in sustaining BC immune escape (21). In addition to their effects on T cells, CAFs interact with other immune components of the TME, including natural killer (NK) cells and dendritic cells (DCs), further contributing immune suppression and tumor progression (22). Despite all these CAF features, their clinical and prognostic relevance in BC needs further in-depth analysis. Traditional approaches based on single markers are limited for their inability to resolve the full

spectrum of CAF subtypes. High-resolution analytical methods, able to dissect complex stromal populations at the single-cell level, were therefore increasingly applied (23–26). Among these, single-cell RNA sequencing (scRNA-seq) revolutionized the analysis of transcriptomic landscapes, allowing detailed characterization of individual CAF subpopulations, their plasticity, and dynamic interactions within the TIME, thus paving the way for the identification of new therapeutic targets and prognostic BC biomarkers (23,26).

The present study fulfills the knowledge gap regarding the heterogeneity and immunomodulatory functions of breast CAF subtypes, with a particular focus on Luminal A and TNBC across different levels of BC aggressiveness. By integrating multiple scRNA-seq datasets, the work addresses the limited understanding of CAF heterogeneity and their immunomodulatory functions, especially focusing on their impact on T-cell activity. Indeed, the present analysis revealed four major stromal populations in these BC patients, each differentially modulating immune cell function, with a specific impact on T lymphocytes. The identification of a previously unrecognized immunosuppressive CAF cluster fulfills an important gap in understanding potential prognostic markers and therapeutic targets for precision medicine approaches in BC.

Materials and Methods

2.1 scRNA-seq data

Data retrieval

Publicly available single-cell RNA sequencing (scRNA-seq) datasets were obtained from three BC studies: EMBO Molecular Medicine (Pal et al., (27)), Nature Genetics (Wu et al., (28)), and Nature Medicine (Bassez et al., (29)). Processed expression matrices and metadata were retrieved as provided by the original authors. Analyses were restricted to CAF populations.

Data preprocessing

All analyses were performed in R (v4.4.3) using Seurat (v5.1.0). Gene expression was normalized using SCTransform with regularized negative

binomial regression; where specified, the glmGamPoi backend and SCT v2 flavor were used. Mitochondrial gene content was quantified per cell using genes matching the ^MT- pattern.

For the EMBO dataset (27), cells were normalized and subjected to principal component analysis (PCA), UMAP embedding, graph-based neighbor detection, and Louvain clustering. Sample identifiers were extracted from cell barcodes and used for patient-level integration. Cells were split by sample and integrated using SCT-based integration with selected variable features and anchor identification, followed by re-computation of dimensionality reduction and clustering on the integrated assay. Marker genes were identified after SCT marker preparation. CAF clusters were defined based on enrichment of canonical fibroblast markers (*COL1A1*, *FN1*, *VIM*). Clusters 15 and 22 were retained for downstream analyses. As previously reported for stromal populations, partial overlap with perivascular or pericyte-like transcriptional programs cannot be excluded.

For the Nature Genetics dataset (28), a Seurat object was generated using the published expression matrix and metadata. Cells annotated as CAFs by the original authors' major cell type classification were directly subset without further reclustering.

For the Nature Medicine dataset (29), only treatment-naïve triple-negative breast cancer samples collected at baseline were retained. Cells annotated as fibroblasts with valid expansion status were subset, normalized using SCTransform, and subjected to PCA, UMAP, neighbor graph construction, and clustering.

Data integration

CAF populations from all studies were merged and integrated at the patient level. Patient identifiers, original annotations, and study-of-origin labels were added as metadata. Cells were split by patient and normalized independently using SCTransform. Patients with fewer than 100 cells were excluded to ensure robust anchor detection. Integration features were selected, SCT-based anchors were identified, and data were integrated into a shared latent

space. The integrated CAF dataset was subjected to PCA, UMAP embedding, graph-based neighbor detection, and unsupervised clustering.

Identification of CAF markers

Cluster-specific marker genes were identified from the patient-integrated CAF dataset using differential expression analysis. Only positively enriched genes were retained. Marker detection was performed using permissive thresholds, including a minimum expression fraction of 10% and no log-fold change cutoff, to enable comprehensive marker discovery.

Spatial transcriptomics validation of CL3 CAFs

To spatially validate the CL3 CAF population in an independent cohort, we analyzed a published spatial transcriptomics dataset comprising 92 TNBC patients (30), in which polyadenylated transcripts were profiled across 100-µm spatial bins. A CL3 CAF gene signature was defined using significantly enriched cluster-specific marker genes identified from the integrated scRNA-seq dataset (adjusted p-value < 0.05, average log2 fold-change > 1, pct.1 > 0.6, pct.2 < 0.4).

Signature enrichment was quantified in each spatial bin using the AUCell R package. Briefly, genes were ranked according to their expression within each spatial bin, and an area under the curve (AUC) score was computed based on the recovery of CL3 signature genes among the top-ranked genes. A maximum ranking threshold corresponding to the top 5% of expressed genes was used for AUC calculation. AUCell scores were z-score normalized across samples for downstream visualization and comparison.

Spatial bins with a CL3 AUCell score > 0.15 were classified as CL3-positive. The spatial distribution of CL3-positive bins was subsequently visualized using the corresponding spatial coordinates across all patient sections.

2.2 Bulk RNA-seq data

Data retrival

The Cancer Genome Atlas (TCGA) repository (https://www.cancer.gov/tcga) was exploited to obtain bulk RNA-seq data referring to patients affected by BC. We considered data of tissue samples of patients affected by two specific

BC subtypes defined based on the well-established molecular PAM50 classification (31): Basal-like/TNBC subtype (98 samples) and Luminal A subtype (229 samples). High-throughput RNA-seq data correspond to level 3 data (i.e., normalized expression data) given in terms of FPKM (i.e., fragments per kilobase of exon per million fragments mapped). Additional details of this dataset can be found in the ref. (32). After the analysis of TCGA data, the expression patterns of selected genes across Luminal A and TNBC samples were validated in three independent Gene Expression Omnibus (GEO) datasets, namely GSE45827, GSE65194, and GSE43358.

Data analysis

Bulk RNA-seq data were analysed by using different computational and statistical approaches, including clustering analysis, differential expression analysis and correlation-based analysis. Heatmaps of gene expression levels across BC samples were generated with the **"**pheatmap**"** R function, normalizing and scaling by gene (z-scores) and applying hierarchical clustering based on correlation distance and complete linkage method.

Box-and-whisker plots of gene expression levels of selected genes across normal, Luminal A, and TNBC samples were generated with the "ggboxplot" R function. Differences between groups were assessed using pairwise statistical comparisons, and significance levels were indicated as follows: p-value > 0.05 (ns), p-value < 0.05 (*), p < 0.01 (**), p < 0.001 (***), and p < 0.0001 (****).

Specifically, comparisons of means between two groups were evaluated using Student's t-test or Mann–Whitney U test, depending on the variable distributions and sample size. Comparisons of means across different groups were evaluated using ANOVA test or Kruskal–Wallis's test, depending on the variable distributions and sample size. Correlation matrix plots were generated with the "corrplot" R function. Correlations between different gene signatures were assessed by Pearson or Spearman correlation coefficients, depending on the variable distributions and sample size. False Discovery Rate (FDR) method was used as correction method in multiple

comparisons. In all statistical analyses a p-value<0.05 was considered as statistically significant. All these analyses were performed by using R statistical software (version 4.1.3). Biological Process gene ontology analysis on genes describing CL03-CAF like was performed using DAVID (https://david.ncifcrf.gov/).

2.3 Immune-related signatures

Eight immune signatures derived from Bindea et al. (33) were used for correlation analysis with stromal signatures. Each immune signature comprised a variable number of genes and was classified as representing either adaptive or innate immune cell populations. Adaptive immune cell signatures included T cells, type 1 helper (Th1) T cells, and $CD8^+$ T cells, whereas innate immune cell signatures included natural killer (NK) cells, NK CD56^dim cells, DCs, and immature DCs (iDC). To capture the functional status of immune cells, the cytotoxic cell signature described by Bindea et al. (33) was also included. In addition, an exhaustion gene signature was developed specifically for this study and consisted of 12 genes selected based on markers most frequently used in literature to characterize immune cell exhaustion (34,35). The exhaustion signature included *PDCD1*, *CTLA4*, *HAVCR2*, *LAG3*, *CD160*, *CD244*, *TIGIT*, *ENTPD1*, *BTLA*, *CD38*, *CD101* and *EOMES*.

2.4 Correlation distribution analysis

To evaluate correlations between CAF Clusters (CC) and Immune System Cells (ISC), we developed a computational framework to compute pairwise Pearson correlations between the gene signatures of each CC–ISC pair and to compare the number of significant correlations with the background expected values. The analysis was performed separately for Luminal A and TNBC subtypes. Specifically, for any given CC–ISC pair, we calculated the correlation value and the corresponding p-value for each gene pair c_i - m_j (where c_i is any gene in cluster CC, i=1,2,…,CC_n, and m_j is any gene in cluster ISC, j=1,2,…, ISC_n), and then counted the number of significantly correlated pairs (p-value < 0.05). We then extracted gene expression values for all genes

in Luminal A and TNBC samples separately and retained only those genes showing a non-constant profile (stdev > 0.5).

For each CC–ISC pair, we randomly selected the same number of genes and computed correlations between random gene pairs, counting the number of significantly correlated pairs (p-value < 0.05). This procedure was repeated 10,000 times, and the mean and standard deviation of the resulting counts were calculated. Finally, a z-score was computed for each CC–ISC pair for both positive and negative correlations.

2.5 Expression and survival analysis

The tissue-wide expression of *GADPH*, *HLA-A*, *HLA-C*, *IL-32*, *LDHA*, *SEC61G*, *TPI1*, *BNIP3L*, *C4orf3*, *TIMP3*, *IER3*, *TBX3*, *TIMP1*, *ADM*, *FN1*, *LOXL2*, *P4HA1* genes was analyzed in BC specimens according to the Luminal A, Luminal B, HER2+, and TNBC subtype stratifications, using the “subtype filter” option, and compared with normal tissues. Kaplan–Meier survival plots for *IER3*, *TBX3*, *TIMP1*, *SEC61G*, *ADM*, *C4orf3*, *LDHA*, *FN1*, *LOXL2* and *P4HA1* genes were generated using the “Survival analysis” tool (kmplot.com/analysis/; (36)), based on OS data and median expression levels, retrieving the Hazard Ratio (HR) with a 95% confidence interval. The present tool integrates transcriptomic and clinical data from multiple public databases, including METABRIC (37,38), The Cancer Genome Atlas (TCGA) and IMPACT (39).

Moreover, to assess the independent prognostic contribution of the CAF signature beyond standard clinical variables, multivariate Cox proportional hazards analyses were performed on the TCGA-BRCA dataset, for which complete clinical annotation was available. Composite CAF scores were computed as the median of z-score normalized expression values of the corresponding gene sets and included as continuous variables in Cox models alongside age at diagnosis and tumor stage as standard clinical covariates. Statistical power was evaluated according to the criterion of at least 10 events per variable (40). Model performance was assessed by the Likelihood Ratio Test, the Akaike Information Criterion (AIC), and Harrell's concordance

index (C-index) (41). All analyses were performed using the survival R package.

2.6 Cell-type deconvolution and CL3 CAF (msCAFs) bulk score

To assess the stromal origin of the CL3 CAF-like (msCAF) transcriptomic signal in bulk RNA-seq data, cell-type deconvolution was performed using MuSiC (Multi-Subject Single Cell deconvolution) method (42). The integrated scRNA-seq dataset used as the single-cell reference, with the four identified stromal populations (iCAF, myCAF, Pericytes, and CL3 (msCAFs)) as cell-type labels. Deconvolution was applied to TCGA bulk RNA-seq profiles of Luminal A and TNBC samples. The CL3 CAF (msCAF) bulk score was computed for each sample as the median of z-score standardized log2-transformed expression values of the CL3 (msCAF)-derived genes showing progressive expression changes across normal tissue, Luminal A, and TNBC (Supplementary Figure S3). For genes showing progressive downregulation (BNIP3L, C4orf3, TIMP3), the z-score sign was inverted prior to computing the median, so that higher scores consistently reflect greater CL3 CAF (msCAF) transcriptional activity. Spearman correlation between the bulk score and MuSiC-estimated CL3 CAF (msCAF) proportions was computed separately for Luminal A and TNBC samples.

2.7 Response to chemotherapy evaluation

ROC curves of *ADM*, *C4ORF3*, *FN1*, *IER3*, *LDHA*, *LOXL2*, *P4HA1*, *SEC61G*, *TBX3* and *TIMP1* genes were generated taking advantage from the Receiver Operating Characteristic (ROC) plotter platform (ROCplot.com; (43)). The pathological Complete Response (pCR) was evaluated in chemotherapy treated BC patients stratified into Luminal A (n = 1272) and TNBC (n = 658) subgroups according to the St. Gallen criteria (44). An AUC value greater than 0.5 was considered statistically significant.

Results

Single-cell transcriptomic integration uncovers heterogeneous CAF and pericyte populations in breast cancer

To comprehensively characterize CAF heterogeneity in BC, we performed an integrated scRNA-seq analysis by combining datasets from three independent studies (EMBO Molecular Medicine (27); Nature Genetics (28); and Nature Medicine (29)).

To address potential study-specific biases and assess the robustness of fibroblast clustering across datasets, we generated a combined dataset including all cell populations and performed dataset integration after down-sampling to 10,000 randomly selected cells per study. Prior to integration, UMAP visualization of the complete dataset coloured by study revealed strong batch effects, with cells primarily segregating according to their dataset of origin (**Fig. 1A-B, left panels**). CAF-focused analysis before integration similarly demonstrated pronounced study-specific clustering patterns (**Fig. 1C**), underscoring the necessity for batch correction and data harmonization.

As detailed in the Methods section (2.1 Data Integration), CAFs were split by patient and normalized using SCTransform before integration. Batch correction and dataset integration were subsequently performed using Seurat FindIntegrationAnchors and IntegrateData. The resulting merged and integrated expression matrix was then used for PCA, clustering, and UMAP embedding. Following integration, cells from all studies showed efficient overlap in the low-dimensional space, confirming successful correction of study-driven batch effects while preserving biologically meaningful transcriptional variability (**Fig. 1B, right panel**). After this stringent quality control and filtering, the integrated dataset comprised 10,498 fibroblasts derived from 29 BC patients. Unsupervised analysis identified four conserved stromal cell populations that were consistently shared across patients and studies (**Fig. 1D**), indicating the presence of stable and reproducible fibroblast-related states in human BC. These included three CAF subtypes and one non-CAF stromal population: cluster 0 resembles inflammatory CAFs (iCAFs); cluster 1 the myofibroblast-like CAFs (myCAF-like); cluster 2 (CL2) corresponds to pericytes; and cluster 3 (CL3) represented a distinct CAF

subset characterized by elevated expression of TGFBI and MMP2 (**Supplementary Table S1**). Consistently, independent analysis of each dataset confirmed the presence of comparable fibroblast populations across studies (**Supplementary Fig. S1A**), while the relative contribution of each study to the integrated clusters demonstrated balanced representation among datasets following integration (**Supplementary Fig. S1B**). This integrative approach confirmed the robustness and reproducibility of fibroblast clustering across independent datasets and provided a unified tool to compare transcriptional profiles among distinct fibroblast subtypes.

Exploiting this harmonized dataset, specific gene expression programs were searched to distinguish pericytes (cluster 2) from proper CAF subtypes (i.e. cluster 1, 4 and 3) (**Supplementary Table S1**). Specifically, cluster 2 displayed a gene signature typical of pericytes or perivascular fibroblasts, including *IGFBP7*, *SPARCL1*, *CAV1*, *COL4A1*, *COL4A2*, *NOTCH3*, and *JAG1*, suggesting vascular support, ECM assembly, and mechanosensory regulation functions. myCAF-like population exhibited strong expression of ECM and contractile genes, including *ACTA2*, *TAGLN*, and *COL1A1*, consistent with their role in orchestrating TME structure by ECM deposition (45). Conversely, iCAF-like population showed elevated expression of inflammatory mediators including *IL6*, *CXCL12*, and *LIF*, reflecting a secretory phenotype associated with paracrine cytokine release and immune modulation. CL3 CAF-like subtype showed intrinsic features, expressing a distinct set of stress-response, metabolic, and ECM remodeling genes, including *GAPDH*, *PGK1*, *ENO1*, *LGALS1*, *FTH1*, *MT2A*, *TPI1*, *BGN*, *FN1*, *ANGPTL4*, *TIMP1*, *PLOD2*, *LOXL2*, and *SERPINE1*. Notably, this cluster also expressed immune-modulatory (i.e. *HLA-A*, *HLA-C*) and hypoxia-responsive genes (i.e. *ADM*, *NDRG1*, *DDIT4*, and *BNIP3)*, as well as apoptosis and stress-associated regulators (*BNIP3L*, *DDIT4*, *PHLDA1*). This distinct signature suggests a stress-responsive and metabolically active fibroblast subtype, potentially contributing to ECM remodeling, hypoxic adaptation, and modulation of local immune responses. The metabolic gene enrichment

suggests a high glycolytic activity, possibly supporting energetic needs for ECM deposition and paracrine signaling (**Fig. 1E**). Indeed, pathway enrichment analysis related to full gene set of the CL3 CAF-like cluster (**Fig. 1E**) revealed a prominent metabolic and stress-response signature, with top enriched pathways encompassing glycolysis and canonical glycolysis, reactive oxygen species metabolism, and mitochondrial catabolic processes. These were accompanied by immune-related categories, including defence responses to viruses, immune activation, and antigen processing and presentation. These findings point to a fibroblast population characterized by a metabolically activated, stress-conditioned state that might modulate the TIME through coordinated metabolic and antigen-related signaling programs. Based on this distinctive transcriptional programme, combining glycolytic reprogramming, hypoxic adaptation, ECM remodelling, and atypical MHC class I antigen presentation, this population is hereafter referred to as metabolic-stressed CAFs (msCAFs).

To validate the physical existence of msCAFs in an independent cohort and investigate their spatial distribution within BC tissues, we analyzed a published spatial transcriptomic dataset comprising 92 TNBC patients (30). Using AUCell scoring based on the msCAF transcriptional signature, we

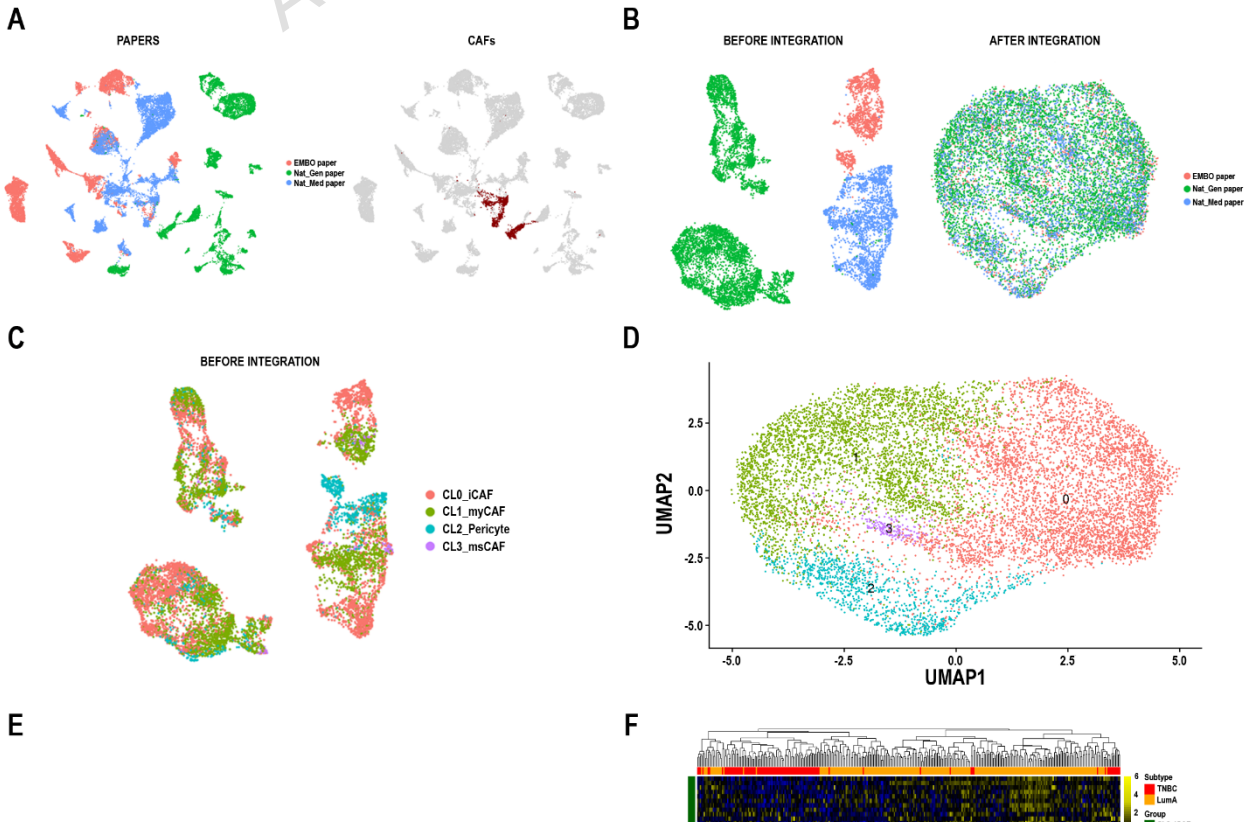

Figure 1. Single-cell transcriptomic analysis identifies conserved and novel CAF subtypes in BC

A. UMAP representation of the complete scRNA-seq dataset generated from three independent BC studies (EMBO Molecular Medicine, Nature Genetics, and Nature Medicine), shown before integration and coloured according to dataset origin, illustrating strong study-specific batch effects across the combined dataset. CAF populations identified in each study are highlighted. **B.** Left, UMAP representation of CAF populations from the three independent studies before integration, demonstrating pronounced dataset-driven segregation. Right, UMAP representation after SCTransform normalization and Seurat-based integration, showing efficient overlap of cells from the three studies following batch correction. **C.** UMAP representation of the integrated fibroblast compartment comprising 10,498 fibroblasts derived from 29 BC patients. Four conserved stromal populations shared across datasets were identified, including inflammatory CAFs (iCAFs; cluster 0), myofibroblast-like CAFs (myCAF-like; cluster 1), pericytes (cluster 2), and a distinct metabolic-stress CAF population (msCAFs; cluster 3). **D.** Integrated UMAP embedding of fibroblast populations coloured according to cluster annotation, highlighting the spatial distribution of iCAFs, myCAF-like CAFs, pericytes, and msCAFs following integration. **E.** Gene Ontology term enrichment analysis of genes highly expressed in the msCAF cluster, performed using DAVID (https://david.ncifcrf.gov/). The top 13 significantly enriched pathways are shown. **F, G.** Heatmaps displaying (D) CAF subtype–specific gene signatures (iCAFs, myCAF-like, cluster 2, and msCAF) and (E) immune cell–associated transcriptional programs (T cells, T helper cells, CD8 T cells, cytotoxic cells, NK cells, NK CD56^dim cells, DCs, iDCs, and exhausted T cells) across BC tumors stratified by molecular subtype (i.e Luminal A and TNBC). In C-E, Analyses were performed using RNA-seq data from The Cancer Genome Atlas (TCGA), focusing on patients with Luminal A BC (n = 229) and those with TNBC (n = 98).

Distinct stromal–immune landscapes define luminal A and triple negative breast cancer

As growing evidence indicates a complex interplay between CAFs and immune cells, supporting tumor progression and therapy response (3), expression levels of the identified CAF signatures and immune-related genes (**Supplementary Table S2**), derived from the Bindea Immune Signature Panel (33), were analysed in two out of four widely recognized molecular BC subtypes: Luminal A and TNBC BC. Luminal A breast tumors are typically hormone receptor-positive and less aggressive, while TNBCs are triple-negative, highly proliferative, and more aggressive (46). To this end, The Cancer Genome Atlas (TCGA) repository was queried to obtain bulk RNA-seq

transcriptomic profiles from BC patients. The analysis was performed focusing on these two extremes of BC pathogenesis.

Unsupervised hierarchical clustering heatmaps display CAF subtype–specific gene signatures (**Fig. 1F**) and immune cell–associated transcriptional programs (**Fig. 1G**), each stratified according to BC subtype. In the CAF-centred heatmap, tumors clustered according to their enrichment for distinct CAF states, revealing subtype-specific stromal profiles (**Fig. 1F**).

TNBCs (red) were particularly enriched in msCAFs, suggesting that these aggressive tumors might promote or recruit a more activated, glycolytic CAF phenotype that supports tumor progression through enhanced ECM remodeling, stress adaptation, and immune modulation. In contrast, Luminal A tumors (orange) exhibited increased expression of gene programs associated with iCAFs, and to a greater extent, myCAFs and pericytes. These populations are typically linked to structural maintenance and ECM organization rather than inflammatory signaling, consistent with a more quiescent stromal landscape oriented toward tissue integrity and regulated ECM turnover.

In the immune heatmap (**Fig. 1G**), a parallel pattern emerged. TNBC displayed elevated expression of genes associated with $CD8^+$ cytotoxic T cells, NK cells, and exhausted T cells, consistent with a more inflamed and immunologically active TIME. Luminal A breast tumors, in contrast, exhibited lower overall immune gene expression.

Together, these findings suggest a coordinated relationship between stromal and immune cell programs in BC, with TNBCs harbouring a pro-inflammatory, immune-enriched microenvironment shaped by iCAFs and msCAFs, while Luminal A breast tumors are associated with a myCAF- and pericyte-dominant stroma and reduced immune infiltration.

Correlation analysis reveals divergent CAF–immune associations across breast cancer subtypes

To elucidate the extent of interaction between specific CAF states and immune cell composition within the breast TME, we computed Z-score–based

correlations between each identified CAF cluster and above-mentioned immune cell gene signatures in Luminal A BC and TNBC samples (**Table 1**). Indeed, in TNBC tumors, iCAFs showed the strongest positive correlations with total T cells (Z = 16.2), cytotoxic immune cells (Z = 8.6), exhausted T cells (Z = 14.7), DCs (Z = 9.6), and iDCs (Z = 10.0), suggesting a potential role in sustaining a highly immune-infiltrated microenvironment. Similarly, myCAFs were positively correlated not only with T cells (Z = 7.6) and T cell exhaustion signature (Z = 7.7), but also with iDCs (Z = 4.6) and $CD56^+$ NK cells (Z = 5.3), suggesting a possible crosstalk with innate immune components. In Luminal A breast tumors, iCAFs remained positively correlated with several immune cell populations (e.g., T cells, DCs, and cytotoxic cells), but the overall magnitude of correlations was lower than in TNBCs. In contrast, pericytes exhibited stronger positive correlations with immune cell signatures compared with their TNBC counterparts. myCAFs show less uniform behaviour in Luminal A BC subtypes, showing positive correlation with DCs, NK cells, and signatures of cytotoxicity and immune exhaustion, but negative correlation with T $CD4^+$ cells and $CD8^+$ T cells (**Table 1**).

Another notable finding was that msCAFs showed extensive negative correlations with immune cell signatures in both TNBC and Luminal A breast tumors (**Table 1**). In TNBC subtype, this was evident for $CD8^+$ T cells and their associated cytotoxicity and exhaustion signatures. In Luminal A BC subtype, negative correlations were especially observed with T $CD4^+$ cells, followed by cytotoxic T cells, and NK cells. Except for the correlation with T $CD4^+$ cells, the magnitude of these negative associations was greater in TNBC subtype (**Table 1**).

Overall, although positive correlations, among iCAFs, myCAFs and different immune cell populations, might suggest a relatively immune-permissive microenvironment in TNBC tumors, the marked enrichment of msCAFs identified in the latter (**Fig. 1**) suggests that this type of CAFs rather favours the establishment of a predominantly immune-excluded TME. This implies

that, despite the presence of pro-inflammatory CAF populations, the substantial abundance of immune-suppressive msCAFs likely constrains immune cell infiltration, generating a more compartmentalized and less accessible immune landscape.

		Luminal A				TNBC			
		CL0_ iCAF	CL1_ myCAF	CL2_ Pericytes	CL3_ msCAF	CL0_ iCAF	CL1_ myCAF	CL2_ Pericytes	CL3_ CAF Like
POSITIVE CORRELATION	T cells	13,797	4,556	8,529	1,872	16,222	7,666	3,799	1,586
	CD4+ T cells	0,354	-2,822	-2,644	-2,942	4,321	-0,166	-1,359	-1,896
	CD8+ T cells	4,411	0,328	8,529	-1,762	4,871	2,044	1,762	-1,541
	T_exhaustion	8,864	4,811	-2,644	1,935	14,650	7,748	2,346	1,784
	T_cytotoxic	9,763	3,677	7,623	1,596	8,663	5,412	1,936	1,359
	NK_cells	3,081	-0,025	2,971	-1,473	2,478	2,329	2,686	-0,694
	NK_CD56	2,166	2,846	2,330	0,921	3,002	5,265	-0,640	0,831
	DC	8,298	2,462	5,974	2,288	9,604	3,738	3,673	0,812
	iDC	6,017	3,571	5,123	0,595	10,020	4,596	4,302	1,048
NEGATIVE CORRELATION	T cells	-3,425	-2,165	-2,794	0,009	-2,577	-2,620	-0,481	7,247
	CD4+ T cells	-1,045	4,231	1,661	4,386	-1,020	1,245	1,641	2,533
	CD8+ T cells	-0,675	4,857	0,052	4,834	-2,577	-2,620	-0,481	3,294
	T_exhaustion	-3,144	-2,349	-2,689	-0,620	-1,020	1,245	1,641	5,619
	T_cytotoxic	-2,069	-1,769	-1,561	0,086	-1,541	-1,089	-0,807	3,575
	NK_cells	-0,326	3,392	2,591	4,422	-1,279	-1,590	-2,248	1,434
	NK_CD56	-2,435	-2,793	-2,933	-2,240	-1,155	-2,295	-1,088	-0,936
	DC	-2,532	-2,025	-2,743	-0,877	-1,411	-1,313	-1,034	1,857
	iDC	-1,387	-0,760	-0,449	0,406	-2,536	-2,860	-1,796	0,556

Table 1. Correlation analysis of CAF signatures and immune cell signatures.
Strongly correlated: POSITIVE: Z > 3.46; NEGATIVE: Z > 3.46; p < 0.0003
Correlated: POSITIVE: 2.33 < Z < 3.46; NEGATIVE: 2.33 < Z < 3.46; p < 0.01

msCAF subtype harbours genes modulating T cell infiltration and function

Given the distinctive negative immune correlations observed between the msCAF subtype and immune cells, particularly T-cell populations (**Table 1**), we focused on specific genes within this cluster that were strongly associated with T cell–related transcriptional programs in both TNBC and Luminal A BC subtypes. The msCAF population is a novel interesting fibroblast population,

which needs further characterization. The identification of genes with a predominant role in immune regulation could provide valuable insights to improve BC patient clinical management.

To this end, we performed gene-gene correlation analyses between each gene belonging to msCAF signature and the ones of the immune signatures (33), focusing on total T cells, $CD4^+$ T cells, $CD8^+$ T cells, and cytotoxicity and exhaustion signatures. Only statistically significant correlations ($p < 0.05$) were considered. Stromal genes were included if they: 1) showed positive or negative correlations with at least three genes from a given T-cell signature; 2) displayed exclusively positive or exclusively negative correlations with the individual genes within each T-cell signature; and 3) exhibited more than 60% positive versus negative correlations (or vice versa). This approach identified a subset of msCAF derived stromal genes consistently associated with multiple T-cell programmes in both Luminal A and TNBC tumours, revealing both subtype-specific stromal wiring and a conserved core programme associated with T-cell infiltration (**Fig. 2 and Supplementary Fig. S2**). Across the overall T-cell signature (**Fig. 2A)**, Luminal A breast tumors displayed a broader set of positively correlated stromal genes than TNBC. A conserved group of five genes (*EMP3, HLA-A, HLA-C, IFI6,* and *IL32*) was positively correlated in both subtypes. Furthermore, specific CAF genes (*BGN, C4orf3, DDIT4, MT2A, SCG5,* and *SERPINE1*) resulted exclusively positively correlated in Luminal A breast tumors. In contrast, negative correlations showed an opposite pattern: Luminal A BC displayed only one subtype-specific negatively associated gene (*LDHA*), whereas TNBC harboured five negatively correlated stromal genes (*ADM, ANGPTL4, DDIT4, GAPDH,* and *TPI1*). The shared negatively correlated set (*BNIP3*, *IER3*, *P4HA1*, and *TBX3*) implicated hypoxia adaptation and ECM remodelling as conserved CAF mechanisms limiting global T-cell accumulation (**Fig. 2A**).

For what concerns $CD4^+$ T-cell signatures (**Supplementary Fig. S2A)**, positive correlations were limited across both subtypes. Luminal A BC exhibited two positively associated stromal genes (*BNIP3L* and *C4orf3*),

whereas TNBC showed three ones (*EMP3, HLA-A,* and *IL32*). *HLA-C* was the only positive gene commonly correlated to both subtypes, underscoring a narrow but conserved antigen-presentation axis. Negative correlations were substantially more frequent in Luminal A BC, which showed ten uniquely negatively associated stromal genes, in comparison to the four ones related to TNBC. Three genes (*GAPDH, LGALS1,* and *TIMP1*) were shared, pointing to glycolytic and galectin-driven stromal pathways that inversely track with $CD4^+$ T-cell activity. Luminal A BC also displayed a large group of subtype-specific negatively correlated genes (*AEBP1, BGN, EMP3, FN1, LOXL2, PLOD2, SEC61G, TIMP3,* and *TPI1*), whereas TNBC specificity was limited to *DDIT4, IER3, P4HA1,* and *BNIP3* (**Supplementary Fig. S2A)**. Correlation patterns with $CD8^+$ T cells (**Supplementary Fig. S2B)** showed marked subtype specificity. Luminal A breast tumors displayed four unique positive correlations (*C4orf3, DDIT4, SCG5,* and *SERPINE1*), while TNBC two ones (*EMP3* and *TIMP1*). Three stromal genes (*HLA-A, HLA-C,* and *IL32*) were positively correlated in both subtypes. Conversely, negative correlations seemed to be typical of Luminal A BC with six unique negatively correlated stromal genes associated with metabolism and ECM modulation (*ENO1, LDHA, LGALS1, LOXL2, SEC61G,* and *TGFB1*), whereas TNBC had three unique negative correlations (*BNIP3, GAPDH,* and *TPI1*). Three stress-response genes (*DDIT4* (47)*, IER3* (48)*,* and *NDRG1* (49)) were negatively associated in both subtypes, forming a core CAF stress-survival module antagonising $CD8^+$ T-cell activity (**Supplementary Fig. S2B**).

To determine whether the subtype-specific stromal associations with T-cell abundance also extended to the functional competence of infiltrating lymphocytes, we examined correlations between msCAF genes and transcriptional programmes reflecting cytotoxic activity (**Fig. 2B**) and T-cell exhaustion (**Fig. 2C**). For the cytotoxic T cell programme, Luminal A breast tumours showed a larger panel of positively correlated stromal genes (*BGN, C4orf3, DDIT4, MT2A, SCG5 and SERPINE1*), than TNBC (*FTH1* and *IFI6*). Four genes, *EMP3, HLA-A, HLA-C* and *IL32* were shared across subtypes.

Negative correlations showed similar subtype biases: Luminal A BC and TNBC displayed two unique negatively correlated genes associated with glycolysis and cellular stress (*LDHA* and *IER3* vs *ADM* and *GAPDH*, respectively), while *BNIP3*, *P4HA1*, and *TBX3* formed a shared negative set highlighting conserved hypoxia- and ECM-regulated suppressive mechanisms (**Fig. 2B**). Interestingly, stromal associations with the T-cell exhaustion programme reflected even more divergencies. Luminal A BC exhibited five subtype-specific stromal genes that were positively correlated with the exhaustion-related program (*C4orf3, DDIT4, MT2A, SCG5,* and *SERPINE1*). In contrast, no subtype-specific positive correlations were detected in TNBC, indicating an apparent loss of stromal–exhaustion coupling with this subtype. In addition, a shared cluster of positively associated genes (*EMP3, HLA-A, HLA-C, IFI6,* and *IL32*) was identified in both subtypes, mirroring correlation patterns observed in other T-cell–related transcriptional programs. Negative correlations were balanced between subtypes: three genes resulted subtype-specific whereas three genes were shared in both subtypes (*ADM*, *ANGPTL4*, and *GAPDH*), highlighting conserved angiogenic, anti-inflammatory, and glycolytic stromal pathways that might broadly suppress T-cell activity (**Fig. 2C**).

Collectively, these analyses reveal that CAF-mediated T-cell regulatory programmes contain a conserved stromal core mainly involving antigen presentation, IL32 signaling, stress-response pathways, hypoxia adaptation, and ECM remodeling. Superimposed on this shared architecture, each BC subtype exhibits unique stromal programmes: Luminal A BC maintains broad, coherent, and predominantly positive stromal–immune associations, consistent with a more permissive stromal–immune interface, whereas TNBC shows more restricted positive interactions and more pronounced negative correlations driven by glycolytic and stress-response pathways, reflecting stronger CAF-mediated constraints on T-cell infiltration and effector function.

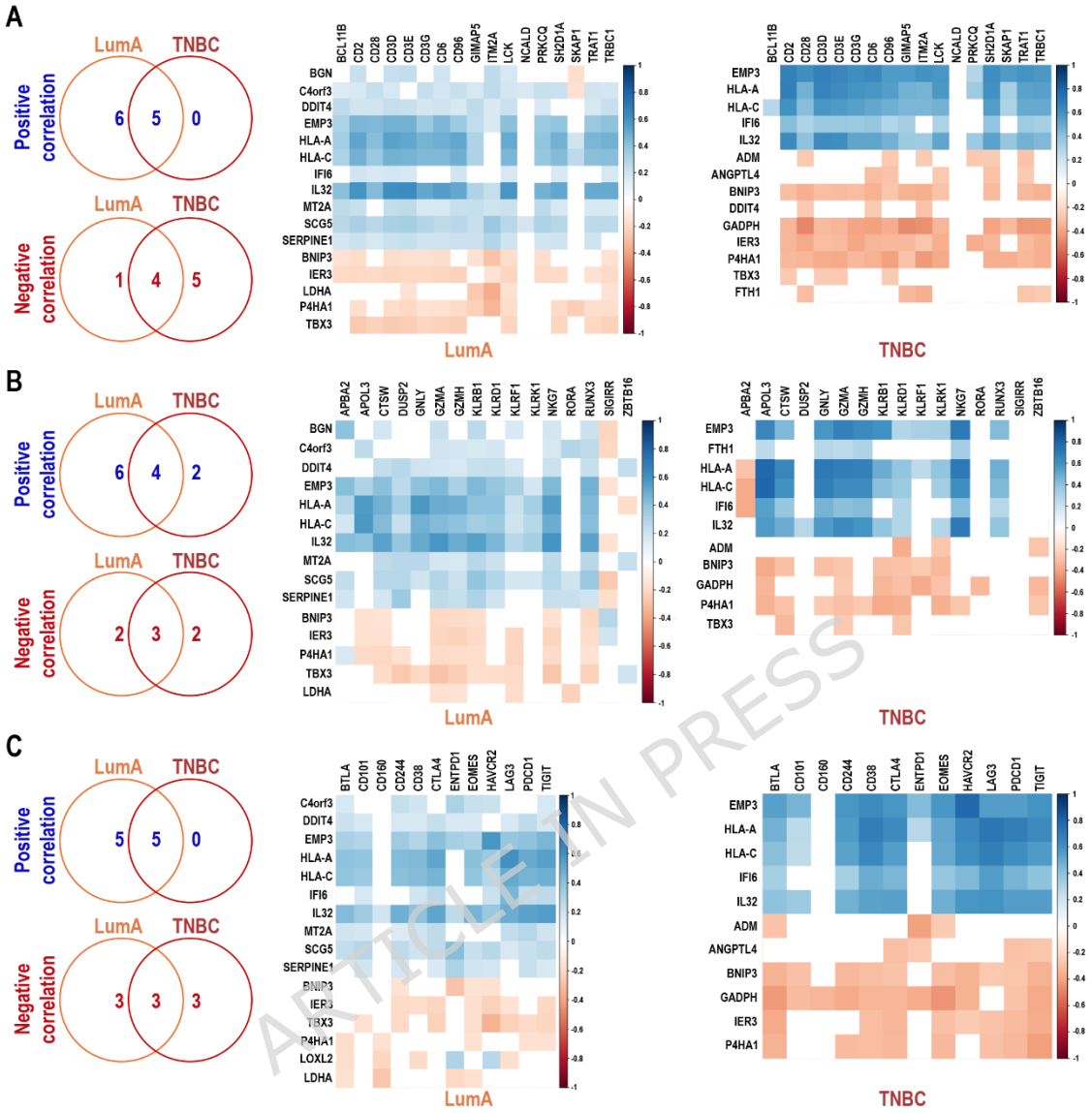

Figure 2. Correlation of msCAF stromal genes with T-cell–associated transcriptional programs across BC subtypes. A-C. Left panels show Venn diagram–based summaries of msCAF stromal genes exhibiting significant positive (POS) or negative (NEG) correlations with (A) overall T-cell signatures, (B) cytotoxic T-cell transcriptional programs, and (C) T-cell exhaustion signatures in Luminal A and TNBC tumors. Numbers in the center indicate genes shared between subtypes, while lateral numbers represent subtype-specific genes. Right panels display heatmaps of correlation patterns between selected msCAF stromal genes and individual genes within each T-cell signature, shown separately for Luminal A and TNBC. In all heatmaps, blue and red denote statistically significant positive and negative correlations ($p < 0.05$), respectively. Stromal genes were selected based on consistent correlation directionality across multiple genes within each immune signature. In A-C, analyses were performed using RNA-seq data from TCGA, focusing on patients with Luminal A BC (n = 229) and those with TNBC (n = 98).

Distinct msCAF-like transcriptional programs define Breast Cancer subtypes and clinical outcome

To further assess the clinical relevance of msCAF genes, positively and negatively correlated with T-cell–associated programs in both BC subtypes, we examined their transcriptional expression in normal breast tissues and in primary Luminal A tumors and TNBC exploiting the TCGA dataset.

Comparative analysis of Luminal A BC and normal breast tissue showed significant upregulation of 18 genes in tumors, including *FN1, BGN, IFI6, IER3, AEBP1, TIMP1, SERPINE1, GAPDH, P4HA1, TPI1, SEC61G, HLA-C, IL32, LOXL2, LDHA, TBX3, HLA-A,* and *FTH1* (**Supplementary Table S3**). Conversely, 12 genes (*EMP3, ENO1, BNIP3, LGALS1, TIMP3, C4orf3, BNIP3L, SCG5, MT2A, NDRG1, ADM,* and *ANGPTL4*) were significantly downregulated, indicating marked stromal transcriptional remodeling in Luminal A BC involving ECM organization, metabolic pathways, and stress-response mechanisms (**Supplementary Table S3**).

Instead, comparative analysis of TNBC and normal breast tissue identified a distinct and pronounced transcriptional signature. 19 genes were significantly upregulated in TNBC, as reflected by increased median expression values and highly significant p-values (**Supplementary Table S4**). The most prominently overexpressed genes included *FN1, BGN, IFI6,*

IL32, GAPDH, TPI1, SEC61G, DDIT4, ENO1, HLA-A, PLOD2, LDHA, P4HA1, HLA-C, TIMP1, NDRG1, TGFBI, SERPINE1, and *LOXL2*. This expression profile highlights a strong activation of pathways related to ECM remodeling, hypoxia signaling, and glycolytic metabolism in TNBC. In contrast, 8 genes (*MT2A, C4orf3, SCG5, ADM, BNIP3L, TIMP3, TBX3,* and *ANGPTL4*) were significantly downregulated in TNBC in comparison to normal tissue, indicating selective repression of specific stromal components within this aggressive BC subtype (**Supplementary Table S4**).

Moreover, comparative analysis of Luminal A BC and TNBC revealed significant subtype-specific differences, with TNBC generally exhibiting higher expression of genes associated with tumor aggressiveness. Sixteen genes (*NDRG1, IL32, ADM, ENO1, GAPDH, TPI1, PLOD2, DDIT4, HLA-A, SEC61G, ANGPTL4, LDHA, MT2A, TGFBI, HLA-C,* and *BNIP3*) were significantly upregulated in TNBC compared with Luminal A tumors, whereas 10 genes (*C4orf3, BGN, TIMP1, SERPINE1, FN1, BNIP3L, TIMP3, AEBP1, IER3,* and *TBX3*) showed higher expression in Luminal A tumors (**Supplementary Table S5**).

Notably, 10 genes of all these analysed ones exhibited a gradual upregulation pattern along with increased tumor aggressiveness: 7 genes (*GAPDH, HLA-A, HLA-C, IL32, LDHA, SEC61G,* and *TPI1*) showed a stepwise increase in their expression from normal tissue to Luminal A tumors and TNBC (**Fig. 3A**), whereas 3 genes (*BNIP3L, C4orf3,* and *TIMP3*) a progressive decrease across the same sequence (**Fig. 3B**). To validate that the msCAF transcriptomic signal detected in bulk RNA-seq data reflects stromal cell content rather than ubiquitous gene expression across tumor and immune compartments, we performed cell-type deconvolution of TCGA bulk samples exploiting MuSiC (42), with our integrated scRNA-seq dataset as reference. The msCAF bulk score (see Materials and Methods) showed a significant positive correlation with MuSiC-estimated msCAF proportions in both TNBC (Spearman R = 0.43, $p = 1\times10^{-5}$) and Luminal A tumors (R = 0.42, $p = 2.2\times10^{-11}$; **Supplementary Figure S3A**). Furthermore, MuSiC-estimated msCAF proportions were

significantly higher in TNBC compared to Luminal A tumors (Student's t-test, $p = 6.4 \times 10^{-12}$; **Supplementary Figure S3B**), consistent with the scRNA-seq findings. These results support the stromal origin of the msCAF bulk transcriptomic signal and strengthen the clinical relevance of the msCAF gene expression changes observed in TCGA.

Thereafter, all significantly differentially expressed genes were assessed for their association with the overall survival (OS). In Luminal A BC, survival analysis identified ten genes whose expression levels were significantly associated with OS ($p<0.05$). Most of these genes (*FN1, IER3, AEBP1, TIMP1, P4HA1, SEC61G, LOXL2,* and *TBX3*) displayed an adverse prognostic pattern, with higher expression levels correlating with poorer patient survival. In contrast, increased expression of *BGN* and *IL32*, as well as reduced expression of *C4orf3* and *ADM*, were associated with favourable OS. The remaining genes that were significantly dysregulated in Luminal A tumors did not show a statistically significant association with OS, indicating that not all expression changes translate into prognostic relevance (**Supplementary Table S3**).

In TNBC, 12 genes were significantly associated with OS. Low expression levels of *FN1, LDHA, P4HA1,* and *LOXL2*, as well as high expression of *C4orf3* and *ADM*, were associated with improved survival outcomes ($p \leq 0.05$). Conversely, increased expression of *BGN, IFI6, IL32, HLA-A,* and *HLA-C* was associated with a favourable OS, whereas reduced expression of *SCG5* was associated with poorer survival. As observed in Luminal A tumors, several genes exhibiting strong transcriptional deregulation in TNBC did not show a significant association with OS, underscoring that differential expression alone is insufficient to predict clinical outcome (**Supplementary Table S4**).

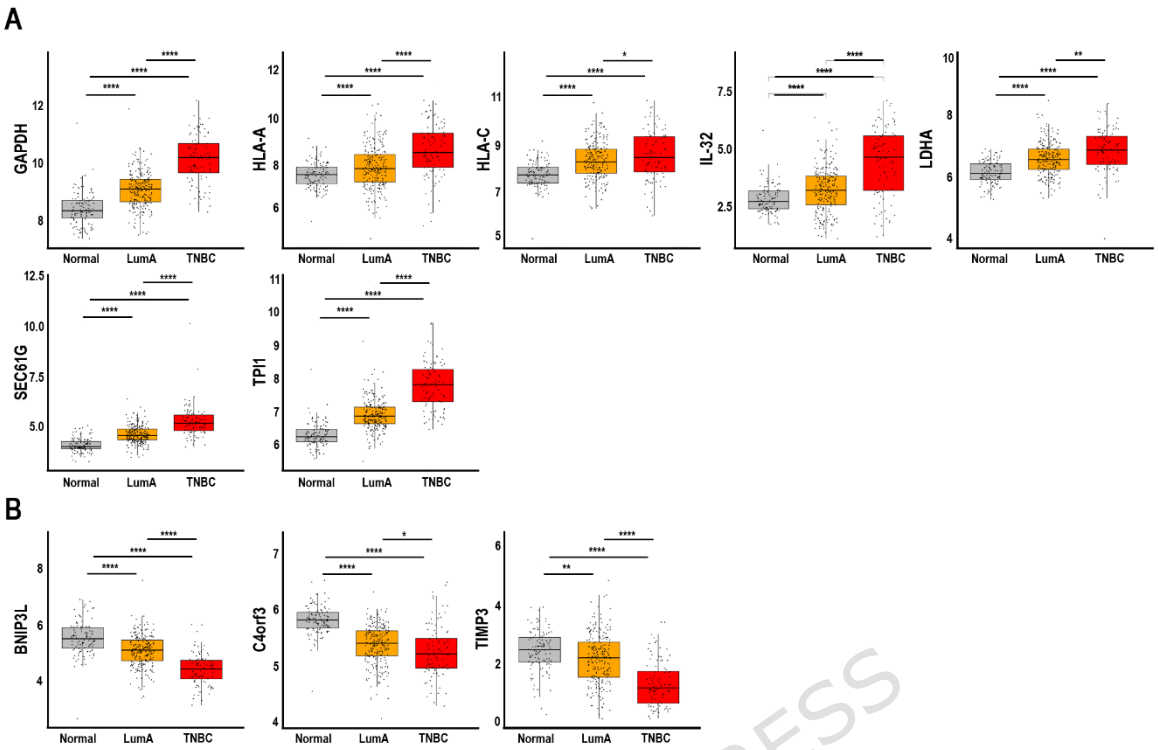

Figure 3. Differential expression of msCAF-like genes across normal breast tissue and Luminal A and TNBC subtypes. A, B. Box plots showing msCAF genes associated with T-cell–related programs that exhibit (A) progressively increased expression and (B) progressively decreased expression from normal breast tissue (n=113) to Luminal A tumors (n=229) and then to TNBC (n=98). Boxes represent the interquartile range, center lines indicate the median, and whiskers denote the data range. RNA-seq data were obtained from The Cancer Genome Atlas (TCGA). Statistical significance between groups was assessed using Student's t-test, with significance level indicated by asterisks (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$).

By integrating subtype-specific expression differences with OS associations that were concordant with the observed expression patterns, this analysis identified distinct groups of prognostically relevant genes. A set of genes emerged as specific of Luminal A subtype, including *IER3, TIMP1, TBX3*, and *SEC61G,* whose differential expression was associated with OS exclusively in Luminal A tumors (**Fig. 4A**). Conversely, a TNBC-specific prognostic signature was identified comprising *LDHA, C4orf3,* and *ADM*. Notably, *LDHA* and *C4orf3* were also differentially expressed in Luminal A tumors. However, their expression levels did not retain prognostic significance in this subtype, indicating subtype-dependent differences in their clinical relevance (**Fig. 4B**). Finally, a subset of genes, *FN1, LOXL2*, and *P4HA1,* displayed consistent deregulation and strong prognostic impact in both Luminal A BC and TNBC subtypes, suggesting shared stromal determinants of clinical outcome independent of the tumor subtype (**Fig. 4C**). Next, to evaluate whether the identified prognostic genes provide independent prognostic information beyond established clinical variables, multivariate Cox proportional hazards analyses were performed on the TCGA-BRCA dataset, the only cohort for which complete clinical annotation was available. Analyses were first conducted separately for Luminal A and TNBC subgroups using subtype-specific composite CAF scores (see Materials and Methods). However, none of the subtype-specific scores reached statistical significance in the multivariate setting, likely due to the limited number of events per subgroup (fewer than 10), which resulted in insufficient statistical power (40). The

independent prognostic contribution of the CAF signature at the subtype level awaits validation in larger cohorts with extended follow-up and complete clinical annotation. To overcome this limitation, we focused on the three genes showing consistent prognostic relevance across both Luminal A and TNBC subtypes (*FN1*, *LOXL2*, and *P4HA1*), which we defined as the shared CAF prognostic signature, and performed the analysis on the full TCGA-BRCA tumor cohort, ensuring adequate statistical power with more than 10 events per variable. The shared CAF score emerged as a significant independent prognostic factor in the multivariate model (HR = 1.53, 95% CI: 1.071–2.185, p-value = 0.019), alongside age (HR = 1.05, 95% CI: 1.024–1.074, p-value < 0.001) and tumor stage (HR = 3.15, 95% CI: 1.761–5.644, p-value < 0.001). The addition of the CAF score significantly improved model fit over the clinical-variables-only baseline (Likelihood Ratio Test p-value = 0.019, ΔAIC = 3.49) and increased the C-index from 0.730 to 0.748, demonstrating added discriminative value. These results confirm that the shared CAF prognostic signature provides independent prognostic value beyond standard clinical variables.

To further investigate the clinical relevance of the identified 10 genes with the strongest differential expression and prognostic value, we queried three independent BC datasets (GSE45827, GSE65194, GSE43358) (**Supplementary Fig.S4**). The available information allowed us to focus exclusively on differential expression between Luminal A BC and TNBC, to perform a validation, since data related to OS are not registered yet. Interestingly, most genes exhibited the same significant trends observed in the TCGA dataset. Specifically, *ADM*, *LDHA*, and *SEC61G* were significantly upregulated in TNBC, whereas *C4orf3*, *FN1*, *IER3*, *TBX3*, and *TIMP1* were upregulated in Luminal A BC. Notably, *FN1* was significantly upregulated in Luminal A tumor in two out of three datasets (GSE45827, GSE65194), while *P4HA1* showed conflicting results in these three cohorts. The analysis of cohorts of BC patients not included in the TCGA further supports the correlations analysed in the present manuscript.

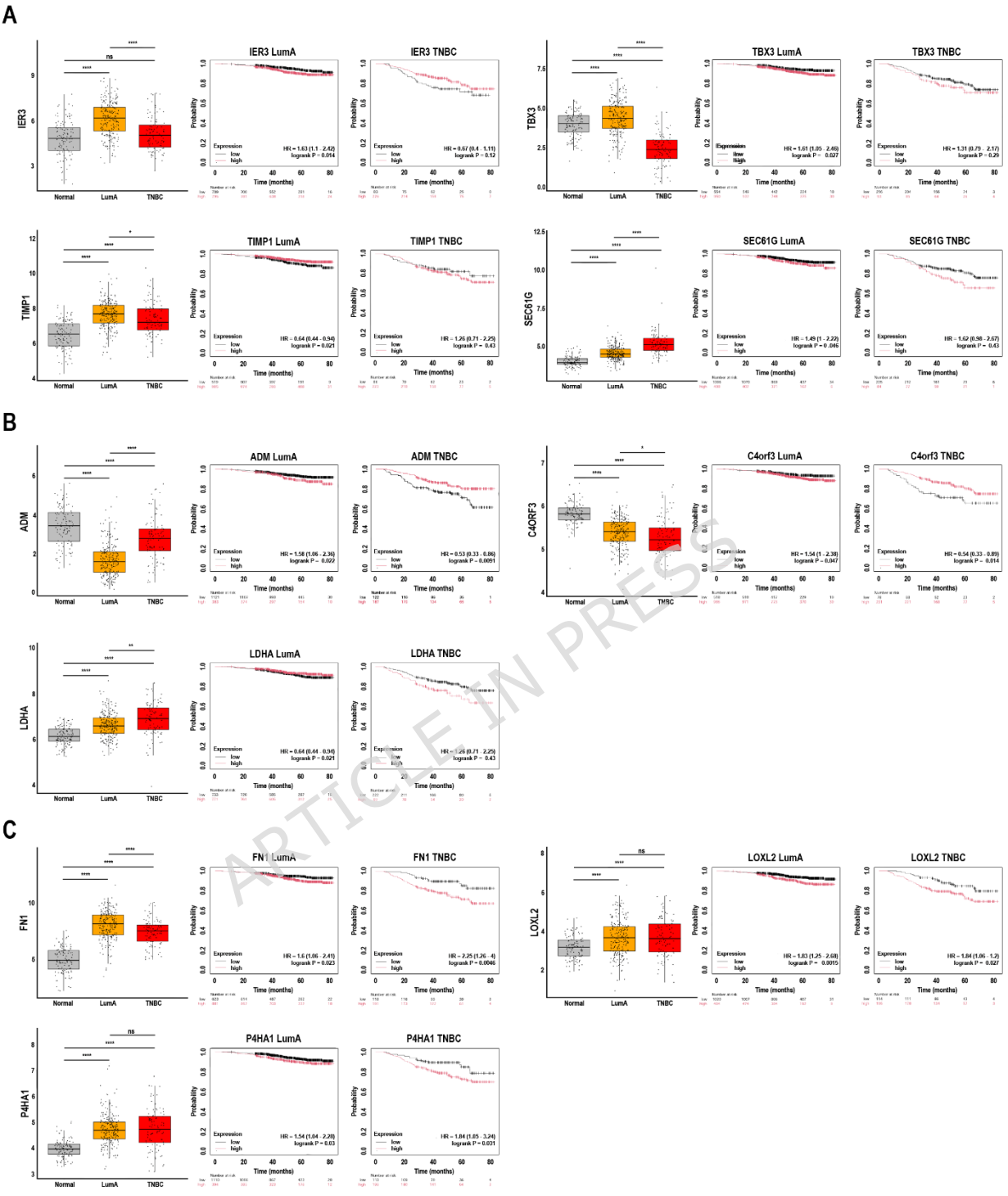

Figure 4. Subtype-specific and shared prognostic genes in Luminal A breast cancer and TNBC.
A-C. BC subtype–specific prognostic genes identified by integrating subtype-restricted expression differences with OS associations that mirror the direction of gene deregulation. For each panel, left, box plots depict relative expression levels of the indicated genes in normal breast tissue (n=113), Luminal A BC tumors (n=229), and TNBC tumors (n=98; boxes represent the interquartile range, center lines indicate the median, and whiskers denote the data range); right, Kaplan-Meier OS analyses stratified by high versus low expression reveal significant prognostic associations for the corresponding indicated genes in (A) Luminal A BC patients, (B) TNBC patients, and (C) both Luminal A BC and TNBC patients. Expression differences were evaluated using Student's t-test, with significance denoted by asterisks (*P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001). RNA-seq data were obtained from TCGA. Survival analyses were performed using the Kaplan–Meier Plotter (https://kmplot.com/analysis/index.php?p=service&cancer=breast_rnaseq_gse96058), selecting the PAM50 subtype option for Luminal A BC (n =1504) and TNBC (n = 309). Patients were dichotomized based on median gene expression, and hazard ratios (HRs) and exact log-rank *p* values are displayed within each plot.

Subtype-specific msCAF-like signatures predict chemotherapy response in Luminal A and TNBC breast cancer

We next compared the expression of genes showing statistically significant differences between normal tissue and tumors and prognostic relevance in TCGA BC patients, using the Receiver Operating Characteristic (ROC) Plotter platform (ROCplot.com; (43)) (**Fig.5 and Supplementary Fig. S5**). Analyses were performed in Luminal A (n = 1272) and TNBC (n = 658) patients treated with chemotherapy, stratified according to treatment response. Clear and subtype-specific expression patterns emerged, supporting a role for these stromal signals in modulating therapeutic efficacy. In Luminal A tumors, *IER3* and *TBX3* were significantly upregulated in non-responders, consistent with their reported association with immunosuppressive CAF status and impaired T-cell function. In contrast, *ADM*, *FN1*, *LOXL2*, and *TIMP1* were significantly enriched in responders, suggesting potential links to treatment-permissive immune microenvironments or stress-adaptive programs that enhance tumor vulnerability. Other candidates, including *C4orf3*, *LDHA*, *P4HA1*, and

SEC61G, did not reach statistical significance in this subtype (**Fig.5 and Supplementary Fig. S3**).

ROC analysis revealed that predictive performance was overall maintained across the msCAF gene set, with area-under-the-curve (AUC) values ranging from ~0.50 to 0.71. Four genes showed poor predictive performance (AUC < 0.58, p value: ns), whereas six genes demonstrated moderate predictive capacity (AUC ≈ 0.57–0.71, $p < 0.05$). Genes reaching statistically significant predictive accuracy included *ADM* (AUC = 0.57), *FN1* (AUC = 0.64), *LOXL2* (AUC = 0.59), *TIMP1* (AUC = 0.61), *TBX3* (AUC = 0.71), and *IER3* (AUC = 0.58). The remaining genes (*C4orf3*, *LDHA*, *P4HA1*, and *SEC61G*) displayed weaker yet directionally consistent discriminatory power, with AUC values ranging from 0.53 to 0.56 (**Fig.5 and Supplementary Fig. S5**).

In TNBC, a distinct transcriptional program associated with chemotherapy response was observed. Responders showed significantly higher expression of *FN1*, *SEC61G*, and *TBX3*, whereas *C4orf3* and *IER3* were significantly downregulated compared with non-responders. In contrast, *ADM*, *LDHA*, *LOXL2*, *TIMP1*, and *P4HA1* did not exhibit statistically significant differential expression patterns in responder and non-responder groups in this subtype. ROC analyses indicated that five genes displayed poor predictive performance (AUC < 0.50, p value: ns), while five genes achieved moderate and statistically significant predictive accuracy (AUC ≥ 0.57, $p < 0.05$). Among these, *FN1* (AUC = 0.56), *SEC61G* (AUC = 0.59), *TBX3* (AUC = 0.69), *C4ORF3* (AUC = 0.63), and *IER3* (AUC = 0.55) provided fair discrimination according to chemotherapy response in TNBC, with significantly improved AUC values relative to genes lacking expression differences (*ADM*, *LDHA*, *P4HA1*, *TIMP1*, and *LOXL2*) (**Fig.5 and Supplementary Fig. S5**).

Of note, *FN1*, *TBX3*, and *IER3* demonstrated predictive value in both Luminal A BC and TNBC. While *FN1* and *IER3* were consistently upregulated and downregulated, respectively, in responders irrespective of subtype, *TBX3* exhibited a clear subtype-dependent behaviour, being downregulated in

responders belonging to Luminal A tumors but upregulated in responders belonging to TNBC.

The reproducible signal observed across multiple CL3 CAF-associated markers highlights the collective contribution of this stromal transcriptional program to therapeutic response. Together, these findings support a model in which immunomodulatory, stress-adaptive, and ECM–remodeling CAF-dependent pathways actively shape anti-tumor immunity and treatment efficacy, reinforcing the clinical relevance of targeting stromal–immune interactions to enhance therapeutic response in BC.

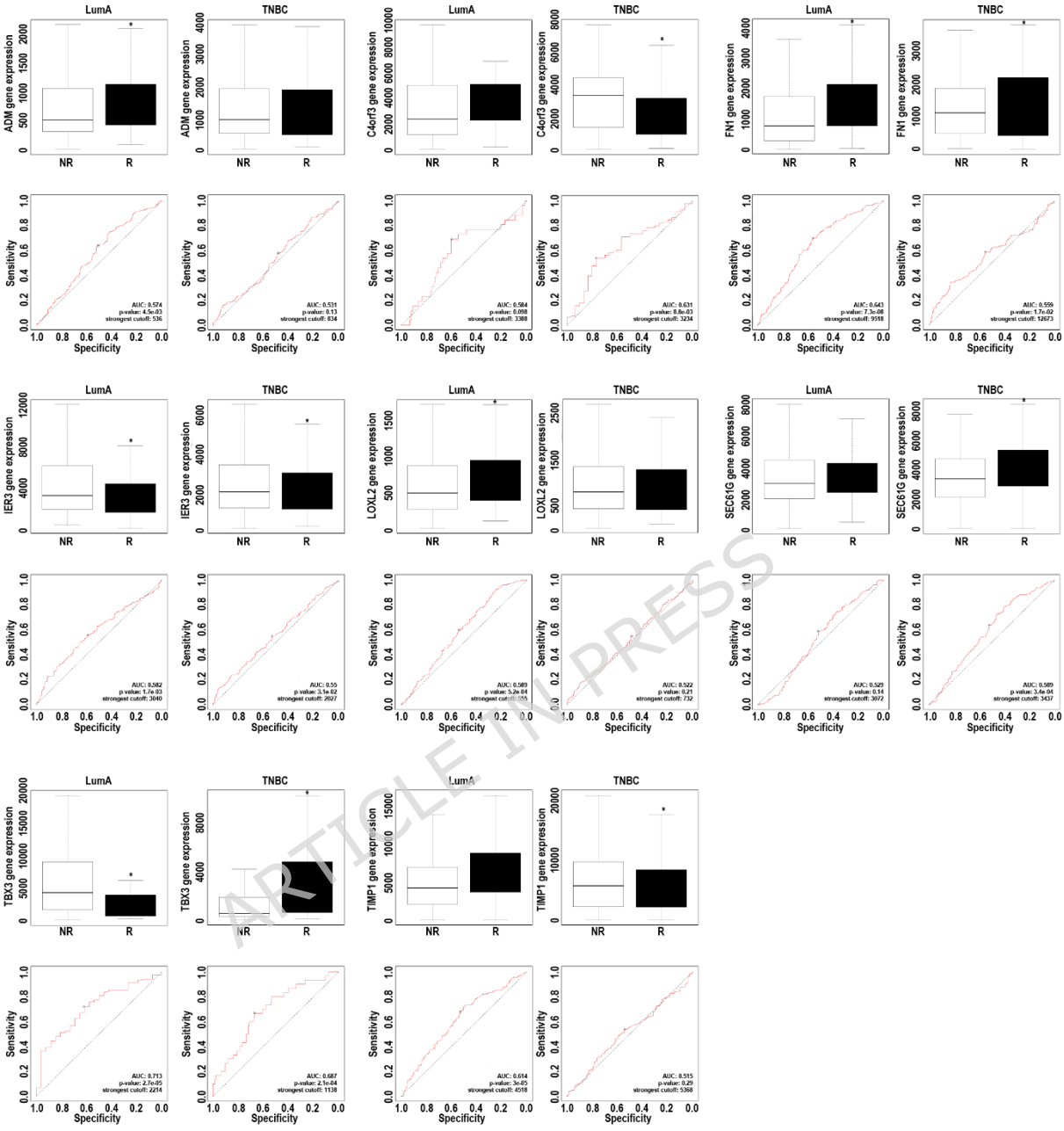

Figure 5. Association of msCAF–related gene expression with chemotherapy response in Luminal A breast cancer and TNBC.
Expression of selected CL3 CAF–associated genes with prognostic relevance was evaluated in chemotherapy-treated patients with Luminal A BC (n=1.272) and TNBC (n=658) using the ROC Plotter platform (https://rocplot.com/). BC patients were stratified according to pathological complete response (pCR). For each indicated gene, top panels show box plots of expression differences between responders and non-responders, while bottom panels display receiver operating characteristic (ROC) curves assessing predictive performance in Luminal A and TNBC tumors. Box plots depict the interquartile range with median values indicated by center lines and whiskers representing the data range. Statistical significance of expression differences is denoted by asterisks (*P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001). For ROC analyses, the area under the curve (AUC) and exact *p* values are reported within each plot.

Discussion

Tumor immune evasion is a process involving both malignant cells and surrounding stroma cells whose interaction establishes an immunosuppressive TME, ultimately impairing effective immune surveillance (50). While tumor cell intrinsic mechanisms of immune escape have been extensively characterized, the immunoregulatory roles of CAFs have been increasingly appreciated but still barely characterized, despite growing interest in their therapeutic targeting (51).

In this study, we provide an integrated single-cell and bulk transcriptomic analysis of CAF heterogeneity in BC, revealing how distinct stromal programs can differentially shape TIME, clinical outcome, and response to chemotherapy according to the molecular BC subtype.

Integration of multiple scRNA-seq datasets (27–29), together with correlation analyses between CAF and immune signatures stratified by clinical parameters, allowed us to refine a model of stroma–immune crosstalk in BC. Essentially, we identified a previously unrecognized CAF subpopulation with immunomodulatory and prognostic relevance, highlighting the functional diversity of stroma compartment beyond canonical classifications.

Our integrated scRNA-seq analysis validates the prominent presence of well-described CAF subtypes, containing iCAFs and myCAFs, and pericytes in independent BC datasets. These stroma populations were consistently

detected in both Luminal A BC and TNBC, albeit at variable abundance, reinforcing the perception that CAF identity is conserved but contextually regulated.

These findings agree with previous works describing the presence of different myCAF subpopulations enriched in collagen-rich BC invasive regions, contributing to ECM deposition and tumor invasion (52,53). Further, scRNA-seq studies have already described iCAFs in BC, classifying them according to inflammatory cytokine expression or immune-modulatory functions, including regulation of leukocyte trafficking *via* chemokines including CXCL12 (54). Recently, Costa et al. identified four distinct subclasses of CAFs differentially associated with BC molecular subtypes, with CAF-S1 and CAF-S4 preferentially enriched in TNBC. Specifically, CAF-S1 promotes immunosuppression through expansion and functional enhancement of regulatory T cells (Tregs) (11).

Recent spatially resolved studies further support the existence of specialised immunoregulatory stromal niches in human BC as suggested also by our findings. Croizer and colleagues identified ten spatially organised FAP^+ CAF niches and demonstrated that tumour cells can drive the transition of detoxification-associated iCAFs toward ECM-producing immunosuppressive myCAFs through DPP4- and YAP-dependent signalling (12). These ECM-myCAFs subsequently polarise $TREM2^+$ macrophages together with regulatory T and NK cells, thereby consolidating immunosuppressive microenvironments. This spatial immune-exclusion programme closely mirrors our observation that msCAFs are the only CAF population consistently inversely correlated with $CD8^+$ T-cell infiltration in both Luminal A BC and TNBC. Similarly, Kieffer and colleagues identified eight functionally distinct CAF-S1 clusters associated with immunotherapy resistance, including ECM-producing and wound-healing CAF states capable of promoting T-cell dysfunction (26). Our data extend this model by showing that msCAF programme integrates ECM remodelling, glycolytic activation, and hypoxia-associated stress pathways into a coordinated stromal state

strongly associated with impaired cytotoxic T-cell activity and T-cell exhaustion, particularly in TNBC.

Beyond these established CAF subtypes, msCAFs emerged as a transcriptionally distinct stromal population characterised by coordinated activation of metabolic, hypoxic, ECM-remodelling, and immune-associated programmes, distinguishing them from previously described CAF states (**Supplementary Table S6**). Systematic comparison with established CAF classification frameworks consistently supported the biological uniqueness of this stromal population. The canonical iCAF/myCAF dichotomy established by Öhlund et al. (55) defines iCAFs by inflammatory cytokine secretion and myCAFs by contractile, αSMA-high matrix-producing programmes; notably, msCAFs exhibit neither of these defining features. Instead, msCAFs are characterised by enrichment of glycolytic enzymes, including GAPDH, PGK1, ENO1, TPI1, and LDHA, together with hypoxia-responsive genes such as ADM, NDRG1, DDIT4, and BNIP3, consistent with a metabolically stressed and hypoxia-adapted stromal identity distinctive from classical CAF categories. Importantly, this coordinated multi-module transcriptional programme was reproducibly identified across three independent scRNA-seq cohorts and lacked expression of canonical EMT regulators, including ZEB1, TWIST1, and SNAIL, arguing strongly against the interpretation of msCAFs as a transient epithelial-to-mesenchymal intermediate state or technical integration artefact.

Comparison with antigen-presenting CAFs (apCAFs) further highlights the distinct biological identity of msCAFs. Canonical apCAFs, originally described by Elyada et al. in pancreatic cancer (56) and subsequently identified in BC (26), are characterised by expression of MHC class II molecules, including CD74, HLA-DRA, HLA-DQ, and HLA-DP, and have been linked to regulatory T-cell expansion and immunosuppressive $CD4^+$ T-cell interactions. In contrast, msCAFs selectively express MHC class I genes, including HLA-A and HLA-C, while lacking CD74 and canonical MHC-II components, indicating a non-canonical antigen-presentation programme

fundamentally distinct from conventional apCAFs. Consistently, msCAFs are not associated with Treg enrichment but instead display a strong and reproducible inverse correlation with $CD8^+$ cytotoxic T-cell infiltration across independent BC cohorts, supporting a specialised role in immune exclusion rather than classical apCAF-mediated immune modulation.

Within BC-specific CAF classification frameworks, Cords and others (2,52) identified tumour-like CAFs (tCAFs) characterised by activation of hypoxia-related pathways and association with adverse clinical outcome. Although msCAFs partially overlap with tCAFs through enrichment of hypoxia-responsive programmes, they additionally exhibit pronounced glycolytic reprogramming, atypical MHC class I antigen-presentation features, and persistent negative associations with $CD8^+$ cytotoxic T-cell infiltration, features not specifically attributed to tCAFs in previous classification systems. These observations therefore support the interpretation of msCAFs as a biologically more specialised stromal state extending beyond a conventional hypoxic CAF phenotype.

The subtype-specific stromal distribution observed in our study, with msCAFs preferentially enriched in TNBC and myCAFs together with pericytes predominating in Luminal A tumours, is consistent with the classification framework proposed by Cords and colleagues (52), who analysed more than 16,000 stromal cells from 14 human BC samples and demonstrated that matrix-remodelling and stress-response CAF states are selectively enriched in more aggressive tumour contexts. Specifically, Wu and colleagues identified an energy-metabolism CAF subtype (emCAF) in TNBC characterised by activation of glycolytic pathways, partially overlapping with the glycolytic component of the msCAF programme. However, emCAFs were not associated with coordinated hypoxia-response signalling, ECM remodelling, or immune-exclusion features, including reduced $CD8^+$ T-cell infiltration, distinguishing them from msCAFs and indicating that the msCAF state represents a broader and more functionally integrated stromal adaptation than metabolic reprogramming alone.

Comparison with other metabolically active CAF populations further emphasises the distinctiveness of msCAFs. Metabolic CAFs (meCAFs) (57) and glycolytic CAFs (glyCAFs) (58) share enhanced glycolytic pathway activity with msCAFs; however, neither population was reported to simultaneously co-express the hypoxia-stress programme, atypical antigen-presentation features, and ECM-remodelling components that collectively define the msCAF state. Importantly, whereas meCAFs have been associated with improved responsiveness to immunotherapy in pancreatic cancer, msCAFs consistently exhibit strong inverse correlations with $CD8^+$ cytotoxic T-cell infiltration across independent human BC cohorts. Together, these findings indicate that metabolic reprogramming alone is insufficient to determine immunological function and support the interpretation that the integrated metabolic, hypoxic, ECM-remodelling, and immune-associated architecture of msCAFs constitutes a biologically and functionally distinct stromal identity.

The transcriptional architecture of msCAFs also aligns with emerging mechanistic concepts in CAF biology linking stromal metabolic adaptation to immune suppression. Recent cross-cancer studies identified metabolic reprogramming, ECM remodelling, and stress adaptation as major drivers of stromal immunosuppression (59). The msCAF signature strongly recapitulates each of these processes. Elevated expression of glycolytic enzymes suggests increased lactate production, a mechanism known to suppress T-cell proliferation and cytotoxicity (60,61). Concurrent upregulation of ECM-remodelling genes, including *FN1*, *LOXL2*, *PLOD2*, *SERPINE1*, *P4HA1*, and *TGFBI*, suggests the formation of dense stromal barriers capable of restricting immune-cell infiltration (59). In parallel, activation of hypoxia-response genes, including *NDRG1*, *DDIT4*, *BNIP3*, *BNIP3L*, *ADM*, and *ANGPTL4*, indicates adaptation to metabolically stressed TME that might further support immune evasion and tumour progression. Notably, the convergence of metabolic reprogramming, hypoxic adaptation, ECM remodelling, and immune-associated signalling within a single

fibroblast population has not previously been described as a discrete CAF entity in BC.

Several additional observations further argue against msCAFs representing either a transient epithelial-to-mesenchymal transition intermediate state or a technical integration artefact. msCAFs consistently express canonical stromal markers, including COL1A1, FN1, VIM, and LGALS1, while lacking epithelial markers and canonical EMT master regulators such as ZEB1, TWIST1, and SNAIL (62). Moreover, the transcriptional stability of the msCAF signature across three independent scRNA-seq cohorts comprising 29 BC patients supports its interpretation as a reproducible and biologically stable stromal population rather than a transient intermediate state.

Collectively, these findings support the interpretation of msCAFs as a biologically distinct stromal population in human BC, rather than a conventional iCAF, myCAF, apCAF, tCAF, or EMT-like transitional state. Their defining combination of glycolytic reprogramming, hypoxic adaptation, ECM remodelling, atypical MHC class I antigen presentation, and broad association with $CD8^+$ T-cell exclusion suggests a specialised role in metabolic adaptation, immune evasion, and stromal-mediated tumour progression within the BC microenvironment.

The coexistence of antigen presentation genes with metabolic and ECM-associated immunosuppressive programs is particularly noteworthy. Although expression of HLA molecules might imply immune activation (63), accumulating evidence suggests that antigen presentation by non-professional antigen-presenting cells (APCs) can instead induce T-cell dysfunction or exhaustion (64). While professional APCs, including DCs, B cells, and macrophages, deliver the full set of signals required for naïve T-cell activation, atypical APCs, including stromal cells, often lack sufficient co-stimulatory capacity (65). This incomplete signaling might lead to tolerance rather than immunity.

Recently, apCAFs have been described in different tumor types although their functional roles remain controversial (11,26,56,66,67). While apCAFs in lung

cancer can activate effector T cells, studies in colorectal cancer demonstrate that CAF-mediated antigen cross-presentation can directly suppress $CD8^+$ T-cell cytotoxicity and induce exhaustion markers, including TIM3, LAG3, and CD39 (68). Consistent with this immunosuppressive paradigm, our data indicate that msCAFs associated with depletion and functional impairment of multiple T-cell subsets, suggesting an immune-inhibitory role in BC.

By integrating CAF signatures with immune transcriptional programs in TCGA BC cohorts of Luminal A BC and TNBC patients, we revealed subtype-specific differences in stroma–immune crosstalk. TNBC tumors displayed a globally immune-enriched phenotype, particularly enriched in NK cell, cytotoxic and exhausted T-cell signatures, consistent with their recognized immunogenicity (69). However, this immune infiltration co-occurred with high levels of msCAF programs, which showed persistent negative correlations with T-cell abundance and function. This apparent paradox suggests that TNBC tumors might be functionally “immune-excluded”, characterized by CAF-dependent metabolic and structural barriers limiting effective antitumor immunity despite immune cell presence. In contrast, Luminal A tumors exhibited lower overall immune infiltration, included in a landscape enriched in myCAFs and pericytes and reduced iCAF contribution. In these tumors, immune activity appeared more diffusely regulated and less constrained by overt immunosuppressive CAF programs.

In both subtypes, iCAFs and myCAFs generally correlated positively with immune system, whereas msCAFs emerged as the only CAF population negatively associated with immune cells, particularly T-cell programs. Notably, although the msCAF signature exhibits an overall inverse association with T-cell signatures, single gene-level correlation analyses revealed positive associations between specific T-cell–related genes and individual CL3-CAF genes. This nuanced pattern suggests that distinct CAF-derived signals might selectively sustain or modulate discrete T-cell functions, despite a broadly immunosuppressive environment. Moreover, another important consideration is that these gene–gene correlations were

derived from bulk RNA-seq data from TCGA. Unlike single-cell RNA-seq, bulk RNA-seq captures the average expression of each gene across all TME populations rather than within defined cellular subsets. Therefore, it is plausible that some of these genes might also be overexpressed at the tumor level, rather than being restricted to specific stromal or immune compartments. Nevertheless, bulk transcriptomic profiling remains the most widely used approach in clinical and translational settings, including diagnostic applications and for the identification of prognostic and therapy-response biomarkers.

In this context, a central finding of our study is the identification of a conserved stromal gene core within msCAFs that is consistently linked to T-cell regulation in BC subtypes. Genes involved in antigen presentation (*HLA-A*, *HLA-C* (70)), inflammatory signaling (*IL32* (71)), stress response (*IFI6* (72)), and immune modulation (*EMP3* (73)) emerged as key nodes connecting CAFs to immune cell homing and dysfunction.

These genes have well-documented but context-dependent roles in cancer biology. IL32, for example, enhances BC invasion through integrin β3–p38 MAPK signaling, while exerting dual immunological effects by promoting IFNγ expression in $CD8^+$ T cells yet inducing Foxp3 expression in $CD4^+$ T cells (74). *IFI6*, an estrogen-responsive and anti-apoptotic gene (75,76), has been linked to therapy resistance and immunosuppressive microenvironments across multiple cancers, including BC (77,78). *EMP3* has been found to correlate with the extent of CAF infiltration (79) and has been implicated in limiting T-cell infiltration through macrophage-mediated chemokine suppression, reinforcing immune exclusion (80). Together, these observations support a model in which msCAFs orchestrate a multifaceted immunosuppressive niche through antigen presentation coupled to metabolic stress and ECM remodeling.

Beyond this shared stromal gene core, our analyses uncovered pronounced subtype-specific msCAF–immune wiring. Luminal A tumors exhibited broader positive stromal–immune correlations compared with TNBC, with enrichment

of msCAF–derived genes linked to T-cell exhaustion and cytotoxic programs, suggesting a state of sustained but functionally constrained immune engagement shaped by chronic CAF-derived signaling. In contrast, TNBC displayed a more polarized stromal architecture, characterized by fewer positive msCAF–immune associations and stronger msCAF-negative correlations driven by glycolytic and hypoxia-associated CAF programs, indicative of a more stringent and metabolically reinforced role in CAF-mediated immune suppression.

These differences translated into distinct prognostic landscapes. While several msCAF associated genes, including *FN1*, *LOXL2*, and *P4HA1*, were consistently associated with poor survival in analysed BC subtypes, other genes displayed context-dependent prognostic value. For instance, *IER3*, *TIMP1*, *TBX3*, and *SEC61G* were prognostically relevant in Luminal A tumors, whereas *LDHA*, *ADM*, and *C4orf3* were more strongly associated with clinical outcome in TNBC. These findings emphasize that the clinical impact of stromal programs cannot be inferred from expression alone but depends on tumor subtype, immune context, and metabolic state.

The prognostic relevance of these genes is supported by prior studies. *FN1* is significantly upregulated in BC compared with normal breast (81), with high *FN1* expression strongly associated with poor clinical outcomes and predictive of patient survival (82). *LOXL2*, a member of the amine oxidase family, promotes collagen and elastin cross-linking within the tumor ECM (83), thereby facilitating cell migration and metastatic dissemination (84), and is implicated in BC progression and metastasis (85). Likewise, P4HA family members are overexpressed in BC, including early-stage disease, and elevated expression correlates with adverse survival outcomes (86), consistent with their role in hypoxia-driven ECM remodeling (87).

At the prognostic level, our observation that *FN1*, *LOXL2*, and *P4HA1* are consistently associated with poor OS across BC subtypes is supported by independent multi-cohort scRNA-seq studies from Chen and colleagues (88), which identified $FN1^{+}$ CAF subpopulations as major drivers of metastasis and

macrophage polarisation. Notably, FN1 silencing in primary CAFs significantly impaired tumour cell migratory capacity, directly supporting a functional role for these stromal programmes in tumour progression. The convergence between our bulk transcriptomic prognostic analyses and independent single-cell functional studies therefore reinforces the biological and clinical relevance of the msCAF programme beyond the original discovery cohorts.

In line with these observations, several additional msCAF-associated genes displayed subtype-specific prognostic associations that further emphasise the context-dependent clinical impact of stromal programmes in BC. *IER3*, *TIMP1*, *TBX3*, and *SEC61G*, while previously linked to prognosis across mixed BC cohorts (48,89–91), emerged here as preferentially relevant in Luminal A tumors, highlighting the importance of molecular context. In contrast, *LDHA* is markedly upregulated in TNBC and HER2-positive tumors (92), where high expression is associated with worse overall, recurrence-free, and distant metastasis–free survival. Functional analyses have linked LDHA to enhanced glycolytic activity and metabolic reprogramming, and its expression correlates with multiple immune cell populations, implicating it in metabolic–immune crosstalk within TME (92).

Notably, *ADM* exhibited a subtype-specific prognostic association in our analysis in contrast with previous reports (93). While we observed higher *ADM* expression to be associated with poor outcome in TNBC, Li-Li Liu and colleagues reported reduced *ADM* expression in TNBC samples and cell lines, with low *ADM* levels correlating with increased recurrence, metastasis, and poor prognosis (93). These discrepancies suggest that ADM might exert context-dependent or stage-dependent functions, potentially influenced by stromal composition, metabolic stress, or immune context, underscoring the complexity of CAF-associated prognostic markers.

Importantly, msCAF transcriptional signatures were associated with chemotherapy response in both Luminal A BC and TNBC, supporting a role of stroma composition in shaping treatment sensitivity. In TNBC, response

was linked to elevated expression of *ADM*, *FN1*, *LOXL2*, and *TIMP1*, consistent with a stress-responsive and dynamically remodeling CAF program. These genes are characteristic of wound-healing and hypoxia-associated pathways (94–97) and might reflect increased ECM plasticity, which could attenuate cell adhesion–mediated drug resistance, transiently enhance drug delivery, and sensitize highly proliferative tumor cells to cytotoxic stress. ADM- and TIMP1-related signaling might further contribute to vascular adaptation (94) and immune–stromal interactions (98) that potentiate chemotherapy-induced tumor cell death.

By contrast, Luminal A BC responders displayed higher expression of *FN1*, *SEC61G*, and *TBX3*, indicating a distinct CAF-associated context potentially related to development transcriptional programs (99), and epithelial–stromal crosstalk (100) in hormone receptor–positive disease. Notably, *TBX3* exhibited opposing associations across CAF subtypes, underscoring the context-dependent functions of stromal regulators. Together, these findings suggest that CAF-mediated modulation of chemotherapy response is highly BC subtype specific and driven by distinct microenvironment-dependent mechanisms rather than uniform stromal effects.

The subtype-dependent predictive value of msCAF-associated signatures for chemotherapy response is further supported by the recently proposed DeCAF framework developed by Peng and colleagues, which demonstrated that CAF subtype composition critically influences treatment responsiveness across multiple solid tumour types (101). Although the DeCAF model was primarily developed in the context of immune checkpoint blockade rather than cytotoxic chemotherapy, its central concept, namely, that tumour-promoting CAF states are associated with inferior therapeutic outcomes whereas tumour-restraining CAF states correlate with improved response, is highly consistent with our findings. Our data demonstrate that msCAF abundance and expression of specific msCAF-associated genes can discriminate chemotherapy responders from non-responders in both Luminal A and TNBC cohorts, highlighting the clinical relevance of stromal composition in shaping

treatment sensitivity. Collectively, the convergence of evidence across bulk transcriptomic analyses, single-cell studies, spatial profiling approaches, and functional CAF classification frameworks establishes stromal heterogeneity as a critical and clinically actionable determinant of immune landscape, prognosis, and therapeutic response in BC.

Conclusions

Collectively, our findings demonstrate that CAF heterogeneity plays a central role in shaping immune surveillance, clinical outcome, and therapeutic response in BC. The identification of msCAFs as an immunosuppressive stromal population with antigen-presenting features provides new insight into mechanisms of immune dysfunction beyond tumor cell intrinsic pathways. These findings support the view that the stromal compartment actively contributes to treatment resistance and immune evasion, and suggest that targeting specific CAF subpopulations, particularly those associated with immune exclusion, could have beneficial effect on responsiveness to both chemotherapy and immunotherapy in BC. Among limitations, the present study lacks a direct functional validation of the immunosuppressive activity of the msCAF subset. Overall, conclusions are primarily based on transcriptomic and correlative analyses, and no *in vitro* or *in vivo* assays were performed to directly demonstrate suppression of T-cell function or modulation of immune cell activity by msCAFs. To reinforce the robustness of our observations, we sought external validation of our scRNA-seq results using an independent available scRNA-seq dataset from Kieffer, Y and colleagues (26). The analysis of this independent cohort revealed consistent transcriptional features associated with the CAF population identified in our study, further supporting the reproducibility and biological relevance of the msCAF-associated signature. Nevertheless, future studies should include functional approaches, including co-culture systems with purified msCAFs and immune cells, as well as in vivo models, to validate the mechanistic role of this CAF subset in shaping an immunosuppressive TME.

Acknowledgements

N/A

Author contributions

Virginia Vigiano Benedetti: Investigation, Methodology, Data curation, Writing—review and editing. Federica Conte: Investigation, Methodology, Data curation, Writing—review and editing. Daniele Santoni: Methodology, Data curation, Writing—review and editing. Tancredi Massimo Pentimalli: Methodology, Data curation, Writing—review and editing. Roberto Bei: Writing—review and editing. Loredana Cifaldi: Writing—review and editing. Giovanni Barillari: Writing—review and editing. Francesco Spallotta: Investigation, Writing—review and editing. Chiara Cencioni: Conceptualization, Funding acquisition, Investigation, Writing—original draft. Ombretta Melaiu: Conceptualization, Funding acquisition, Investigation, Writing—original draft.

Funding

This research was supported by PRIN-MIUR 2022 grant n. 2022ZFFALH to P.I.: O.M and co-PI: C.C; and by AIRC, Associazione Italiana per la Ricerca sul Cancro (AIRC), My First AIRC “Giorgio e Adriana Squinzi” MFAG number 23099 to F.S.

Data availability

Data will be made available on request

Conflict of interest

The authors declare no competing interests.

Ethics declarations

Ethics approval and consent to participate

Not applicable.

References

1. Xiao Y, Yu D. Tumor microenvironment as a therapeutic target in cancer. Pharmacol Ther. 2021 May;221:107753.
2. Liu Y, Sinjab A, Min J, Han G, Paradiso F, Zhang Y, et al. Conserved spatial subtypes and cellular neighborhoods of cancer-associated fibroblasts revealed by single-cell spatial multi-omics. Cancer Cell.

2025 May 12;43(5):905-924.e6.

3. 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.
4. Kim J, Harper A, McCormack V, Sung H, Houssami N, Morgan E, et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. Nat Med. 2025 Apr;31(4):1154–62.
5. Harris MA, Savas P, Virassamy B, O'Malley MMR, Kay J, Mueller SN, et al. Towards targeting the breast cancer immune microenvironment. Nat Rev Cancer. 2024 Aug;24(8):554–77.
6. Xiong X, Zheng L-W, Ding Y, Chen Y-F, Cai Y-W, Wang L-P, et al. Breast cancer: pathogenesis and treatments. Signal Transduct Target Ther. 2025 Feb 19;10(1):49.
7. Orrantia-Borunda E, Anchondo-Nuñez P, Acuña-Aguilar LE, Gómez-Valles FO, Ramírez-Valdespino CA. Subtypes of Breast Cancer. 2022.
8. Kojima Y, Acar A, Eaton EN, Mellody KT, Scheel C, Ben-Porath I, et al. Autocrine TGF-beta and stromal cell-derived factor-1 (SDF-1) signaling drives the evolution of tumor-promoting mammary stromal myofibroblasts. Proc Natl Acad Sci U S A. 2010 Nov 16;107(46):20009–14.
9. Nair N, Calle AS, Zahra MH, Prieto-Vila M, Oo AKK, Hurley L, et al. A cancer stem cell model as the point of origin of cancer-associated fibroblasts in tumor microenvironment. Sci Rep. 2017 Jul 28;7(1):6838.
10. Zeisberg EM, Potenta S, Xie L, Zeisberg M, Kalluri R. Discovery of endothelial to mesenchymal transition as a source for carcinoma-associated fibroblasts. Cancer Res. 2007 Nov 1;67(21):10123–8.
11. Costa A, Kieffer Y, Scholer-Dahirel A, Pelon F, Bourachot B, Cardon M, et al. Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. Cancer Cell. 2018 Mar 12;33(3):463-479.e10.

12. Croizer H, Mhaidly R, Kieffer Y, Gentric G, Djerroudi L, Leclere R, et al. Deciphering the spatial landscape and plasticity of immunosuppressive fibroblasts in breast cancer. Nat Commun. 2024 Apr 1;15(1):2806.
13. Magagna I, Gourdin N, Kieffer Y, Licaj M, Mhaidly R, Andre P, et al. CD73-Mediated Immunosuppression Is Linked to a Specific Fibroblast Population That Paves the Way for New Therapy in Breast Cancer. Cancers (Basel). 2021 Nov 23;13(23).
14. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. Nat Med. 2018 May;24(5):541–50.
15. Ziani L, Chouaib S, Thiery J. Alteration of the Antitumor Immune Response by Cancer-Associated Fibroblasts. Front Immunol. 2018;9:414.
16. Altieri A, Buechler MB. Breast cancer-associated fibroblasts maintain self-renewing immunosuppressive macrophages. Nat Rev Immunol. 2024 Jul;24(7):459.
17. Lakins MA, Ghorani E, Munir H, Martins CP, Shields JD. Cancer-associated fibroblasts induce antigen-specific deletion of CD8 + T Cells to protect tumour cells. Nat Commun. 2018 Mar 5;9(1):948.
18. Kim HS, Noh Y-K, Min K-W, Kim D-H, Kwon MJ, Pyo JS, et al. Cancer-Associated Fibroblasts Together with a Decline in CD8+ T Cells Predict a Worse Prognosis for Breast Cancer Patients. Ann Surg Oncol. 2024 Mar;31(3):2114–26.
19. Kim HS, Min K-W. ASO Author Reflections: Cancer-Associated Fibroblasts with Decreased CD8+ T Cells are Associated with a Poor Prognosis in the Treatment of Breast Cancer Patients. Ann Surg Oncol. 2024 Mar;31(3):2127–8.
20. Jenkins L, Jungwirth U, Avgustinova A, Iravani M, Mills A, Haider S, et al. Cancer-Associated Fibroblasts Suppress CD8+ T-cell Infiltration and Confer Resistance to Immune-Checkpoint Blockade. Cancer Res. 2022 Aug 16;82(16):2904–17.

21. Zheng S, Zou Y, Tang Y, Yang A, Liang J-Y, Wu L, et al. Landscape of cancer-associated fibroblasts identifies the secreted biglycan as a protumor and immunosuppressive factor in triple-negative breast cancer. Oncoimmunology. 2022;11(1):2020984.
22. Ielpo S, Barberini F, Dabbagh Moghaddam F, Pesce S, Cencioni C, Spallotta F, et al. Crosstalk and communication of cancer-associated fibroblasts with natural killer and dendritic cells: New frontiers and unveiled opportunities for cancer immunotherapy. Cancer Treat Rev. 2024 Dec;131:102843.
23. Yang J, Xin B, Wang X, Wan Y. Cancer-associated fibroblasts in breast cancer in the single-cell era: Opportunities and challenges. Biochim Biophys acta Rev cancer. 2025 Apr;1880(2):189291.
24. Zhang X, Ren B, Liu B, Wang R, Li S, Zhao Y, et al. Single-cell RNA sequencing and spatial transcriptomics reveal the heterogeneity and intercellular communication of cancer-associated fibroblasts in gastric cancer. J Transl Med. 2025 Mar 18;23(1):344.
25. Liu Y, Dong G, Yu J, Liang P. Integration of single-cell and spatial transcriptomics reveals fibroblast subtypes in hepatocellular carcinoma: spatial distribution, differentiation trajectories, and therapeutic potential. J Transl Med. 2025 Feb 18;23(1):198.
26. Kieffer Y, Hocine HR, Gentric G, Pelon F, Bernard C, Bourachot B, et al. Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. Cancer Discov. 2020 Sep;10(9):1330–51.
27. Pal B, Chen Y, Vaillant F, Capaldo BD, Joyce R, Song X, et al. A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. EMBO J. 2021 Jun 1;40(11):e107333.
28. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. Nat Genet. 2021 Sep;53(9):1334–47.
29. Bassez A, Vos H, Van Dyck L, Floris G, Arijs I, Desmedt C, et al. A

single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. Nat Med. 2021 May;27(5):820–32.

30. Wang X, Venet D, Lifrange F, Larsimont D, Rediti M, Stenbeck L, et al. Spatial transcriptomics reveals substantial heterogeneity in triple-negative breast cancer with potential clinical implications. Nat Commun. 2024 Nov 26;15(1):10232.
31. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol. 2009 Mar 10;27(8):1160–7.
32. Grimaldi AM, Conte F, Pane K, Fiscon G, Mirabelli P, Baselice S, et al. The New Paradigm of Network Medicine to Analyze Breast Cancer Phenotypes. Int J Mol Sci. 2020 Sep 12;21(18).
33. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013 Oct 17;39(4):782–95.
34. Woroniecka K, Chongsathidkiet P, Rhodin K, Kemeny H, Dechant C, Farber SH, et al. T-Cell Exhaustion Signatures Vary with Tumor Type and Are Severe in Glioblastoma. Clin Cancer Res. 2018 Sep 1;24(17):4175–86.
35. Belk JA, Daniel B, Satpathy AT. Epigenetic regulation of T cell exhaustion. Nat Immunol. 2022 Jun;23(6):848–60.
36. Posta M, Győrffy B. Pathway-level mutational signatures predict breast cancer outcomes and reveal therapeutic targets. Br J Pharmacol. 2025 Dec;182(23):5734–47.
37. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012 Apr 18;486(7403):346–52.
38. Pereira B, Chin S-F, Rueda OM, Vollan H-KM, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun.

2016 May 10;7:11479.

39. Hintzsche J, Kim J, Yadav V, Amato C, Robinson SE, Seelenfreund E, et al. IMPACT: a whole-exome sequencing analysis pipeline for integrating molecular profiles with actionable therapeutics in clinical samples. J Am Med Inform Assoc. 2016 Jul;23(4):721–30.
40. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol. 1995 Dec;48(12):1503–10.
41. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996 Feb 28;15(4):361–87.
42. Wang X, Park J, Susztak K, Zhang NR, Li M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. Nat Commun. 2019 Jan 22;10(1):380.
43. Fekete JT, Győrffy B. ROCplot.org: Validating predictive biomarkers of chemotherapy/hormonal therapy/anti-HER2 therapy using transcriptomic data of 3,104 breast cancer patients. Int J cancer. 2019 Dec 1;145(11):3140–51.
44. Mihály Z, Győrffy B. Improving Pathological Assessment of Breast Cancer by Employing Array-Based Transcriptome Analysis. Microarrays (Basel, Switzerland). 2013 Aug 29;2(3):228–42.
45. Lv K, He T. Cancer-associated fibroblasts: heterogeneity, tumorigenicity and therapeutic targets. Mol Biomed. 2024 Dec 16;5(1):70.
46. Nolan E, Lindeman GJ, Visvader JE. Deciphering breast cancer: from biology to the clinic. Cell. 2023 Apr 13;186(8):1708–28.
47. Tirado-Hurtado I, Fajardo W, Pinto JA. DNA Damage Inducible Transcript 4 Gene: The Switch of the Metabolism as Potential Target in Cancer. Front Oncol. 2018;8:106.
48. Ye J, Zhang Y, Cai Z, Jiang M, Li B, Chen G, et al. Increased expression of immediate early response gene 3 protein promotes aggressive

progression and predicts poor prognosis in human bladder cancer. BMC Urol. 2018 Sep 24;18(1):82.

49. Merlot AM, Porter GM, Sahni S, Lim EG, Peres P, Richardson DR. The metastasis suppressor, NDRG1, differentially modulates the endoplasmic reticulum stress response. Biochim Biophys acta Mol basis Dis. 2019 Sep 1;1865(9):2094–110.
50. Gupta I, Hussein O, Sastry KS, Bougarn S, Gopinath N, Chin-Smith E, et al. Deciphering the complexities of cancer cell immune evasion: Mechanisms and therapeutic implications. Adv Cancer Biol - Metastasis. 2023 Oct;8:100107.
51. Huang Z, Chen J, Zhu T, Li J, Ng HY, Zhou Y, et al. Cancer-associated fibroblasts in the tumor microenvironment: heterogeneity, crosstalk mechanisms, and therapeutic implications. Mol Cancer. 2025 Dec 18;25(1):19.
52. Cords L, Tietscher S, Anzeneder T, Langwieder C, Rees M, de Souza N, et al. Cancer-associated fibroblast classification in single-cell and spatial proteomics data. Nat Commun. 2023 Jul 18;14(1):4294.
53. Wu X, Lu W, Zhang W, Zhang D, Mei H, Zhang M, et al. Integrated analysis of single-cell RNA-seq and bulk RNA-seq unravels the heterogeneity of cancer-associated fibroblasts in TNBC. Aging (Albany NY). 2023 Nov 13;15(21):12674–97.
54. Friedman G, Levi-Galibov O, David E, Bornstein C, Giladi A, Dadiani M, et al. Cancer-associated fibroblast compositions change with breast cancer progression linking the ratio of S100A4+ and PDPN+ CAFs to clinical outcome. Nat cancer. 2020 Jul;1(7):692–708.
55. Öhlund D, Handly-Santana A, Biffi G, Elyada E, Almeida AS, Ponz-Sarvise M, et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. J Exp Med. 2017 Mar 6;214(3):579–96.
56. 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 Discov. 2019 Aug;9(8):1102–23.

57. Wang Y, Liang Y, Xu H, Zhang X, Mao T, Cui J, et al. Single-cell analysis of pancreatic ductal adenocarcinoma identifies a novel fibroblast subtype associated with poor prognosis but better immunotherapy response. Cell Discov. 2021 May 25;7(1):36.
58. Broz MT, Ko EY, Ishaya K, Xiao J, De Simone M, Hoi XP, et al. Metabolic targeting of cancer associated fibroblasts overcomes T-cell exclusion and chemoresistance in soft-tissue sarcomas. Nat Commun. 2024 Mar 20;15(1):2498.
59. Fan C, Zhu W, Chen Y, Zhu W, Ding J. Cancer-Associated Fibroblasts: Origin, Classification, Tumorigenicity, and Targeting for Cancer Therapy. MedComm. 2025 Nov;6(11):e70415.
60. Comito G, Iscaro A, Bacci M, Morandi A, Ippolito L, Parri M, et al. Lactate modulates CD4+ T-cell polarization and induces an immunosuppressive environment, which sustains prostate carcinoma progression via TLR8/miR21 axis. Oncogene. 2019 May;38(19):3681–95.
61. Ielpo S, Barberini F, Gaiba A, Baronti C, Greppi M, Obino V, et al. Lactate-mediated NK cell dysfunction as a prognostic marker and therapeutic target in breast cancer. Cell death Discov. 2026 Mar 27;12(1).
62. Migault M, Sapkota S, Bracken CP. Transcriptional and post-transcriptional control of epithelial-mesenchymal plasticity: why so many regulators? Cell Mol Life Sci. 2022 Mar 12;79(3):182.
63. da Silva GBRF, Silva TGA, Duarte RA, Neto NL, Carrara HHA, Donadi EA, et al. Expression of the Classical and Nonclassical HLA Molecules in Breast Cancer. Int J Breast Cancer. 2013;2013:250435.
64. Combes F, Meyer E, Sanders NN. Immune cells as tumor drug delivery vehicles. J Control Release. 2020 Nov 10;327:70–87.
65. Kambayashi T, Laufer TM. Atypical MHC class II-expressing antigen-

presenting cells: can anything replace a dendritic cell? Nat Rev Immunol. 2014 Nov;14(11):719–30.

66. Kerdidani D, Aerakis E, Verrou K-M, Angelidis I, Douka K, Maniou M-A, et al. Lung tumor MHCII immunity depends on in situ antigen presentation by fibroblasts. J Exp Med. 2022 Feb 7;219(2).
67. Zhang M, Yang H, Wan L, Wang Z, Wang H, Ge C, et al. Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. J Hepatol. 2020 Nov;73(5):1118–30.
68. Harryvan TJ, Visser M, de Bruin L, Plug L, Griffioen L, Mulder A, et al. Enhanced antigen cross-presentation in human colorectal cancer-associated fibroblasts through upregulation of the lysosomal protease cathepsin S. J Immunother cancer. 2022 Mar;10(3).
69. Loizides S, Constantinidou A. Triple negative breast cancer: Immunogenicity, tumor microenvironment, and immunotherapy. Front Genet. 2022;13:1095839.
70. Karnaukhov V, Paes W, Woodhouse IB, Partridge T, Nicastri A, Brackenridge S, et al. HLA variants have different preferences to present proteins with specific molecular functions which are complemented in frequent haplotypes. Front Immunol. 2022;13:1067463.
71. Meng D, Dong H, Wang C, Zang R, Wang J. Role of interleukin-32 in cancer progression (Review). Oncol Lett. 2024 Feb;27(2):54.
72. Jia H, Mo W, Hong M, Jiang S, Zhang Y-Y, He D, et al. Interferon-α inducible protein 6 (IFI6) confers protection against ionizing radiation in skin cells. J Dermatol Sci. 2020 Nov;100(2):139–47.
73. Chen Q, Jin J, Huang X, Wu F, Huang H, Zhan R. EMP3 mediates glioblastoma-associated macrophage infiltration to drive T cell exclusion. J Exp Clin Cancer Res. 2021 May 8;40(1):160.
74. Han L, Chen S, Chen Z, Zhou B, Zheng Y, Shen L. Interleukin 32 Promotes Foxp3+ Treg Cell Development and CD8+ T Cell Function in Human Esophageal Squamous Cell Carcinoma Microenvironment.

Front cell Dev Biol. 2021;9:704853.

75. Mueller C, Haymond A, Davis JB, Williams A, Espina V. Protein biomarkers for subtyping breast cancer and implications for future research. Expert Rev Proteomics. 2018 Feb 3;15(2):131–52.
76. Cocco S, Piezzo M, Calabrese A, Cianniello D, Caputo R, Lauro V Di, et al. Biomarkers in Triple-Negative Breast Cancer: State-of-the-Art and Future Perspectives. Int J Mol Sci. 2020 Jun 27;21(13).
77. Barzaman K, Karami J, Zarei Z, Hosseinzadeh A, Kazemi MH, Moradi-Kalbolandi S, et al. Breast cancer: Biology, biomarkers, and treatments. Int Immunopharmacol. 2020 Jul;84:106535.
78. Bertoli G, Cava C, Castiglioni I. MicroRNAs: New Biomarkers for Diagnosis, Prognosis, Therapy Prediction and Therapeutic Tools for Breast Cancer. Theranostics. 2015;5(10):1122–43.
79. Zeng W-J, Yang Y-L, Liu Z-Z, Wen Z-P, Chen Y-H, Hu X-L, et al. Integrative Analysis of DNA Methylation and Gene Expression Identify a Three-Gene Signature for Predicting Prognosis in Lower-Grade Gliomas. Cell Physiol Biochem. 2018;47(1):428–39.
80. Dong J, Wang F, Gao X, Zhao H, Zhang J, Wang N, et al. Integrated analysis of genome-wide DNA methylation and cancer-associated fibroblasts identified prognostic biomarkers and immune checkpoint blockade in lower grade gliomas. Front Oncol. 2022;12:977251.
81. Pomella S, Bei R, Melaiu O, Barillari G. Fibronectin Is a Likely Therapeutic Target Shared by Oral and Breast Carcinomas. Int J Mol Sci. 2026 Jan 23;27(3):1148.
82. Zhang X-X, Luo J-H, Wu L-Q. FN1 overexpression is correlated with unfavorable prognosis and immune infiltrates in breast cancer. Front Genet. 2022;13:913659.
83. Ye C, Jiang S, Zeng T, He S, Cao J, Xiao J. The role of LOXL2 in tumor progression, immune response and cellular senescence: a comprehensive analysis. Discov Oncol. 2024 Jun 26;15(1):245.
84. Ferreira S, Saraiva N, Rijo P, Fernandes AS. LOXL2 Inhibitors and

Breast Cancer Progression. Antioxidants (Basel, Switzerland). 2021 Feb 19;10(2).

85. Ramos S, Ferreira S, Fernandes AS, Saraiva N. Lysyl Oxidases Expression and Breast Cancer Progression: A Bioinformatic Analysis. Front Pharmacol. 2022;13:883998.
86. Li M, Wang Q, Zheng Q, Wu L, Zhao B, Wu Y. Prognostic and diagnostic roles of prolyl 4-hydroxylase subunit α members in breast cancer. Biomark Med. 2021 Sep;15(13):1085–95.
87. Aggarwal V, Sahoo S, Donnenberg VS, Chakraborty P, Jolly MK, Sant S. P4HA2: A link between tumor-intrinsic hypoxia, partial EMT and collective migration. Adv cancer Biol - metastasis. 2022 Oct;5.
88. Chen B, Tang H, Zheng X, Xie F, Yu P, Lyu Y, et al. Spatial and functional dissection of cancer-associated fibroblasts-mediated immune modulation in H. pylori-associated gastric cancer. Mol Cancer. 2025 Nov 6;24(1):282.
89. Cheng G, Fan X, Hao M, Wang J, Zhou X, Sun X. Higher levels of TIMP-1 expression are associated with a poor prognosis in triple-negative breast cancer. Mol Cancer. 2016 Apr 30;15(1):30.
90. Aliwaini S, Lubbad AM, Shourfa A, Hamada HAA, Ayesh B, Abu Tayem HEM, et al. Overexpression of TBX3 transcription factor as a potential diagnostic marker for breast cancer. Mol Clin Oncol. 2019 Jan;10(1):105–12.
91. Ma J, He Z, Zhang H, Zhang W, Gao S, Ni X. SEC61G promotes breast cancer development and metastasis via modulating glycolysis and is transcriptionally regulated by E2F1. Cell Death Dis. 2021 May 27;12(6):550.
92. Xia JQ, Wang YL, Li XL, Wang JN, Li FC, Jia SY, et al. Lactate dehydrogenase a is a crucial biomarker that affects the prognosis, chemotherapy effect, and immune infiltration of breast cancer. BMC Womens Health. 2025 Oct 17;25(1):497.
93. Liu L-L, Chen S-L, Huang Y-H, Yang X, Wang C-H, He J-H, et al.

Adrenomedullin inhibits tumor metastasis and is associated with good prognosis in triple-negative breast cancer patients. Am J Transl Res. 2020;12(3):773–86.

94. Zhang Y, Xu Y, Ma J, Pang X, Dong M. Adrenomedullin promotes angiogenesis in epithelial ovarian cancer through upregulating hypoxia-inducible factor-1α and vascular endothelial growth factor. Sci Rep. 2017 Jan 16;7:40524.
95. Xu H, Chai H, Chen M, Zhu R, Jiang S, Liu X, et al. Single-cell RNA sequencing identifies a subtype of FN1 + tumor-associated macrophages associated with glioma recurrence and as a biomarker for immunotherapy. Biomark Res. 2024 Oct 7;12(1):114.
96. Fan Z, Zheng W, Li H, Wu W, Liu X, Sun Z, et al. LOXL2 upregulates hypoxia-inducible factor-1α signaling through Snail-FBP1 axis in hepatocellular carcinoma cells. Oncol Rep. 2020 May;43(5):1641–9.
97. Manresa MC, Godson C, Taylor CT. Hypoxia-sensitive pathways in inflammation-driven fibrosis. Am J Physiol Regul Integr Comp Physiol. 2014 Dec 15;307(12):R1369-80.
98. Li Q, Wei K, Zhang X, Lv Y, Li M, Zhou C, et al. TIMP1 shapes an immunosuppressive microenvironment by regulating anoikis to promote the progression of clear cell renal cell carcinoma. Aging (Albany NY). 2023 Sep 15;15(17):8908–29.
99. Khan SF, Damerell V, Omar R, Du Toit M, Khan M, Maranyane HM, et al. The roles and regulation of TBX3 in development and disease. Gene. 2020 Feb 5;726:144223.
100. Hashemi M, Arani HZ, Orouei S, Fallah S, Ghorbani A, Khaledabadi M, et al. EMT mechanism in breast cancer metastasis and drug resistance: Revisiting molecular interactions and biological functions. Biomed Pharmacother. 2022 Nov;155:113774.
101. Peng XL, McCabe IC, Kharitonova E V, Xu Y, Zhao RT, Luan C, et al. DeCAF defines clinical fibroblast subtypes and multidimensional tumor-stroma crosstalk shaping prognosis and immunotherapy

response. Cell reports Med. 2026 Feb 17;7(2):102611.