

Research Paper



2024; 14(5): 1873-1885. doi: 10.7150/thno.90627

Cancer-associated fibroblast spatial heterogeneity and *EMILIN1* expression in the tumor microenvironment modulate TGF- β activity and CD8⁺ T-cell infiltration in breast cancer

Chikako Kanno Honda¹, Sasagu Kurozumi^{1,2}, Takaaki Fujii¹, Didier Pourquier^{3,4,5}, Lakhdar Khellaf^{3,4,5}, Florence Boissiere³, Jun Horiguchi², Tetsunari Oyama⁶, Ken Shirabe¹, Jacques Colinge^{3,5,7}, Takehiko Yokobori⁸, Andrei Turtoi^{3,4,5,8}

- 1. Department of General Surgical Science, Gunma University Graduate School of Medicine, Maebashi, Gunma, Japan.
- 2. Department of Breast Surgery, International University of Health and Welfare, Chiba, Japan.
- 3. Institut régional du Cancer de Montpellier (ICM)-Val d'Aurelle, Montpellier, France.
- 4. Tumor Microenvironment and Resistance to Treatment Lab, INSERM U1194, Montpellier, France.
- 5. Université de Montpellier, Montpellier, France.
- 6. Department of Pathology and Diagnostics, Gunma University Graduate School of Medicine, Maebashi, Gunma, Japan.
- 7. Cancer Bioinformatics and Systems Biology Team, INSERM U1194, Montpellier, France.
- 8. Division of Integrated Oncology Research, Gunma University, Initiative for Advanced Research (GIAR), Maebashi, Gunma, Japan.

Corresponding author: Andrei Turtoi, PhD; Email: andrei.turtoi@inserm.fr; Tel.: +33(0)411.28.31.88. Takehiko Yokobori, MD, PhD; Email: bori45@gunma-u.ac.jp; Tel.: +81(0)70.7594.6065. Jacques Colinge, PhD; Email: jacques.colinge@umontpellier.fr; Tel.: +33(0)411.28.31.18.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). See http://ivyspring.com/terms for full terms and conditions.

Received: 2023.09.29; Accepted: 2024.02.13; Published: 2024.02.24

Abstract

Rationale: The tumor microenvironment (TME) and its multifaceted interactions with cancer cells are major targets for cancer treatment. Single-cell technologies have brought major insights into the TME, but the resulting complexity often precludes conclusions on function.

Methods: We combined single-cell RNA sequencing and spatial transcriptomic data to explore the relationship between different cancer-associated fibroblast (CAF) populations and immune cell exclusion in breast tumors. The significance of the findings was then evaluated in a cohort of tumors (N=75) from breast cancer patients using immunohistochemistry analysis.

Results: Our data show for the first time the degree of spatial organization of different CAF populations in breast cancer. We found that IL-iCAFs, Detox-iCAFs, and IFNy-iCAFs tended to cluster together, while Wound-myCAFs, TGF β -myCAFs, and ECM-myCAFs formed another group that overlapped with elevated TGF- β signaling. Differential gene expression analysis of areas with CD8⁺ T-cell infiltration/exclusion within the TGF- β signaling-rich zones identified elastin microfibrillar interface protein 1 (*EMILIN1*) as a top modulated gene. *EMILIN1*, a TGF- β inhibitor, was upregulated in IFNy-iCAFs directly modulating TGF β immunosuppressive function. Histological analysis of 75 breast cancer samples confirmed that high EMILIN1 expression in the tumor margins was related to high CD8⁺ T-cell infiltration, consistent with our spatial gene expression analysis. High EMILIN1 expression was also associated with better prognosis of patients with breast cancer, underscoring its functional significance for the recruitment of cytotoxic T cells into the tumor area.

Conclusion: Our data show that correlating TGF- β signaling to a CAF subpopulation is not enough because proteins with TGF- β -modulating activity originating from other CAF subpopulations can alter its activity. Therefore, therapeutic targeting should remain focused on biological processes rather than on specific CAF subtypes.

Keywords: spatial transcriptomics, cancer invasion, tumor immunity, CAF subpopulations, patient outcome

Introduction

Breast cancer (BC) is the second most frequent cause of cancer death in women worldwide [1]. Molecular and histological classifications of BC have significantly improved its clinical management [2, 3]. Yet, a subtype of BC called triple negative breast cancer (TNBC), named after the absence of hormone receptors and HER2 receptor protein-tyrosine kinase, remains a clinical challenge [3, 4]. Reminiscent of other difficult to cure cancers, TNBC is a prime example where alternative approaches to those aiming to target the cancer cells are urgently needed [4]. A complementary approach to cancer cell targeting is to target the environment in which they reside [5] and that is called tumor microenvironment (TME). Major components of the TME are non-parenchymal resident cells (e.g. endothelial cells, fibroblasts, adipocytes), immune cells and the supportive extracellular matrix (ECM) with specific physical and chemical properties [6]. TME has for decades attracted a considerable attention for novel drug development in solid tumors. Indeed today, in some cancers, modulating angiogenesis (leading to a reduced or normalized blood supply) or activating the cytotoxic immune responses is more effective than killing cancer cells (e.g. melanoma [7, 8] and hepatocellular [9] carcinoma), and as such represents the first-line treatment. This demonstrates the potential of TME-directed therapies. Despite these encouraging results, BC (like many other solid tumors) has not really benefited from TME targeting yet. Clinical trials with anti-angiogenic therapy or immune-checkpoint inhibitors (ICIs) produced rather unimpressive results [10-14]. However, a very good clinical response is achieved in a small subpopulation of patients [15, 16], which raises hope that the concept of TME targeting has significant potential that needs to be unlocked.

Of the different TME cells, cancer-associated fibroblasts (CAF) have been known as critical for BC progression and recognized as important modulators of tumoral immune function [17]. Recent data on CAF subpopulations in BC raised the question whether specific CAF subtypes should be targeted to allow for better anti-tumor immune response [18, 19]. Answering this question requires a fairly detailed understanding of the interaction between the different cells in the TME, including CAF and immune cells. A step towards this objective has been facilitated by advances in spatial OMICS technologies, which have been a true game changer for characterizing the TME and possible intercellular interactions [20, 21]. They have allowed, for the first time, to link the spatial occurrence of different TME cell subtypes and of cancer cells with enhanced proliferative, immunosuppressive or therapy-resistance features [22]. In the present study, we investigated the spatial distribution of CAF subpopulations in BC and their relationship with infiltrating cytotoxic T cells. To do so we reconstituted a comprehensive single-cell BC atlas, where multiple subpopulations of CAF were identified. We then projected single-cell data on spatially resolved BC transcriptomes, specifically analyzing spatial proximity between infiltrating cytotoxic T cells and CAF subpopulations. Our findings globally showed that successful TME targeting should be process-oriented and not cell-oriented. Indeed, immune exclusion is promoted and regulated by the concerted action of several stromal cell types, and therefore disrupting a specific cell population is unlikely to abolish such a process entirely. Conversely, it might be more relevant to target the potentially trans-cellular molecular network at the heart of a crucial tumor process. However, knowledge on this area is still limited. Particularly, it is crucial to understand how different stromal cells molecularly support engage to or oppose tumor-promoting programs. Our study offers an example of such a program, where CAF-derived elastin microfibrillar interface protein 1 (EMILIN1) counteracts CAF-mediated immunosuppressive function of TGF-β.

Materials and Methods

Patients

Four patients with invasive ductal BC who underwent surgical resection at Gunma University Hospital (Gunma, Japan) in 2020-2021 were enrolled for the Visium Spatial Gene Expression experiments (clinical data are in Table S1). For immunohistochemical staining (validation study), 75 patients with invasive BC who underwent breast-conserving surgery or modified total mastectomy at Gunma University Hospital (Gunma, Japan) in 2014-2015 were enrolled (Table S2). Men with BC were not included in the study. None of the patients received neoadjuvant treatment. Their median age was 60 years (range, 35-82 years). Pathological tumor size, nodal status and lymphovascular invasion were determined using the pathological records. The present study was approved by the Gunma University Hospital Institutional Review Board (reference no. HS2021-071) and was conducted according to the tenets of the Declaration of Helsinki. All patients gave their consent via the opt-out system.

Tissue optimization

Tissue optimization was performed following the 10x Genomics Visium Spatial Tissue Optimization

Reagents Kits User Guide (CG000238, 10x Genomics) to optimize the permeabilization time for the subsequent gene expression profiling. BC tissue cryosections (10 µm-thick) were placed on a Visium Spatial Tissue Optimization Slide (10x Genomics). Different permeabilization times were tested with different tissue sections on the slide with poly(dT) capture the mRNA. After primers to the permeabilization and the mRNA capture steps, reverse transcription followed by addition of fluorescently labeled oligonucleotides to the cDNA allowed detecting the resulting cDNAs as fluorescence signals. Hematoxylin and eosin (H-E) staining and the fluorescence signals were imaged with a BZ-X800 microscope (Keyence). The optimal permeabilization time was the incubation time that gave the strongest fluorescence signal.

Gene expression analysis library preparation

Spatial gene expression analysis was done with the Visium Spatial Gene Expression Reagent Kit (10x Genomics) following the manufacturer's user guide (CG000239, 10x Genomics). BC tissue cryosections (10 µm-thick) were placed on a Visium Spatial Gene Expression Slide (10x Genomics). Images of H-E-stained sections were taken with a BZ-X800 microscope (Keyence). After tissue permeabilization for the optimal time (see above), mRNA capture with the poly(dT) probes in the slide and reverse transcription resulted in the construction of the full-length cDNA. After second strand synthesis and denaturation, cDNAs were amplified in a Veriti 96-Well Thermal Cycler (Thermo Fisher Scientific) and quantified with a LabChip GX Touch HT Nucleic Acid Analyzer (PerkinElmer) to ensure that sufficient cDNA amounts were generated for the library construction. Enzymatic fragmentation and size selection with the SPRIselect reagent (Beckman Coulter) were used to optimize the cDNA fragment size for sequencing. Then, sample index PCR allowed preparing sequence-ready libraries. The final library quantification was done with LabChip GX.

Sequencing

The MGIEasy Universal Library Conversion Kit (MGI Tech) was used to convert the libraries to DNBSEQ-compatible libraries. Sequencing was done by DNBSEQ-G400 (MGI Tech) with a DNBSEQ-G400RS High-throughput Sequencing Set (App-A FCL PE100) following the manufacturer's instructions. The resulting read lengths were as follows: Read 1-28 bp and Read 2–100 bp. Raw data were deposited at GEO with the accession code GSE243022.

Bioinformatics

The raw fastq files were processed with the SpaceRanger software 1.0 (10x Genomics) using the human genome reference set GRCh38-3.0.0 and default parameters. Data obtained from our four BC samples were complemented with published data [23] retrieved from GEO (reference GSE176078). For our four samples, tissue areas were defined using the Seurat clustering default algorithm (functions FindNeighbors and FindClusters). The cluster number was adjusted to the maximum value where distinct, cluster-specific gene expression patterns were detected with the Seurat differential search tool (function FindAllMarkers). These tissue areas were named by referring to the original areas defined by a pathologist. For the publicly available datasets [22], the original area definitions were used. Of note, these tumor areas played no role in the analysis, and they were defined only for reference and descriptive purposes.

The gene spatial expression analysis mainly relied on our library BulkSignalR [24] and project-specific R scripts. Count matrices were filtered for non-expressed genes by imposing a minimum read count of 1 in at least 1% of the Visium spots. Subsequently, normalization was achieved by total count. Cell population-specific gene signatures were retrieved from sequence data for BC general cell populations [23] and for CAFs [25]. In all cases, the top 20 genes reported for each population were used. The spatial abundance of each cell population was estimated by applying BisqueRNA [26] to these gene signatures due to the bulk nature of Visium spatial data. Copy number variation analysis using InferCNV [27] was performed to validate the annotation of cancerous epithelial cell populations, and distinguish those cells from the cells of the TME. A first scoring of cycling cancer cells was obtained using BisqueRNA scores for the Cycling population [23]. An alternative score was provided by scoring a gene signature available from Seurat (cc.genes.updated.2019\$s.genes and cc.genes.updated.2019\$g2m.genes). In this case, scoring was done using the BulkSignalR function scoreSignatures. The various plots reporting the localization of cell types or cycling cells were generated using BulkSignalR standard functions.

Several biological processes were scored using fast Gene Set Enrichment Analysis (fGSEA) [28] spatial transcriptomic features. TGF- β signaling was scored with the BulkSignalR scoreSignatures function to generate dendrograms to relate this process to CAF subpopulations. TGF- β signaling genes were obtained from MSSigDB (c5.bp.v7.0.symbols.gmt.txt, GO_TRANSFORMING_GROWTH_FACTOR_BETA_ RECEPTOR_SIGNALING_PATHWAY). Spatial co-localization between cell populations or between TGF- β and cell populations was determined using a Pearson correlation-based distance matrix (distance = 1 – correlation). Correlations within one sample were computed over the whole set of spots.

The differential gene expression analysis to compare CD8⁺ T cell-rich *versus* -poor areas with the top TGF- β signaling tumor areas was performed with edgeR [29] and the following parameters: maximum false discovery rate of 5%, minimum fold-change of 1.5, and normalized read count >1 in at least 25% of spots. For one sample (tumor 114223F), this last threshold was decreased to 20%.

Immunohistochemistry

Paraffin-embedded BC specimens (n=75; 10 luminal A, 10 luminal B, 20 luminal HER2, 15 HER2, and 20 TNBC) were cut into 4 µm-thick sections and mounted on glass slides. All sections were incubated at 60 °C for 60 min, deparaffinized in xylene, rehydrated, and incubated with fresh 0.3% hydrogen peroxide in 100% methanol at room temperature for 30 min to block endogenous peroxidase activity. After rehydration through a graded series of ethanol solutions, antigen retrieval was performed using an Immunosaver (Nishin EM, Tokyo, Japan) at 98 °C-100 °C for 30 min. Sections were passively cooled to room temperature and then incubated in Protein Block Serum-Free Reagent (Agilent (Dako), Santa Clara, CA, USA) for 30 min. This was followed by incubation with an anti-EMILIN1 rabbit polyclonal antibody (x400, HPA002822; Sigma Aldrich, Saint Louis, MO, USA) in Dako REAL Antibody Diluent at 4 °C for 24 h. According to the manufacturer's instructions, EMILIN1 staining was visualized as a red color using the Histofine Simple Stain AP (Multi) Kit (Nichirei, Tokyo, Japan) and the FastRed II reagent (Nichirei, Tokyo, Japan). Then, sections were boiled in a microwave oven for 10 min to inactivate the antibodies and enzyme activity. Next, they were incubated with an anti-CD8 rabbit polyclonal antibody (x500, ab4055; Abcam, Cambridge, UK) in Dako REAL Antibody Diluent at 4 °C for 24 h. CD8 staining was visualized as a brown color using the Histofine Simple Stain MAX-PO (Multi) Kit (Nichirei, Tokyo, Japan) and DAB substrate. Sections were lightly counterstained with hematoxylin and mounted. Negative controls were incubated without the primary antibody, and no staining was detected.

EMILIN1 expression was evaluated as staining intensity and staining ratio in 200x view fields from two tumor margin areas and one center area. Staining intensity was evaluated as 0 (none), 1 (weak), 2 (moderate), and 3 (strong). The ratio of EMILIN1stained area to the whole field of view was evaluated as 0 (none), 1 (1 %-25 %), 2 (26 %-50 %), 3 (51 %-75 %), 4 (\geq 76 %). The staining intensity and ratio were multiplied to obtain the EMILIN1 score (0-12). The total number of CD8⁺ cells was counted in the 200x view fields where the EMILIN1 score was evaluated. Breast cancer samples with higher EMILIN1 score in the margin than central area (score ratio of margin to center > 1) were defined as a high EMILIN1 group, and the others (score ratio \leq 1) as low EMILIN1 group.

Immunofluorescence analysis

Multicolor immunofluorescence staining was performed in tissue sections of BC surgically resected from five patients to detect EMILIN1, CD8, and TGFBI expression and from seven patients to detect EMILIN1, CD8, and Ki-67 using the Akoya Biosciences Opal Kit following the manufacturer's instructions. All patients were selected from the validation group of BC samples. In the first five samples, EMILIN1 staining (anti-EMILIN1 rabbit polyclonal antibody: x400, HPA002822, Sigma) was visualized using the Opal 480 Fluorophore; CD8 staining (anti-CD8 rabbit polyclonal antibody: x500, ab4055, Abcam) with the Opal 570 Fluorophore; and TGFBI staining (anti-TGFBI rabbit polyclonal antibody: x400, 10188-1-AP; Proteintech, Rosemont, IL, USA) with the Opal 520 Fluorophore. In the other seven BC samples, EMILIN1 staining was visualized using the Opal 480 Fluorophore, CD8 staining with the Opal 570 Fluorophore, as above, and Ki-67 staining (anti-Ki67 rabbit monoclonal antibody: x500, #9027; Cell Signaling Technology, Danvers, MA, USA) with the Opal 520 Fluorophore. All sections were lightly counterstained with hematoxylin and examined under an All-in-One BZ-X710 fluorescence microscope (KEYENCE Corporation, Osaka, Japan).

Statistical analysis

Immunohistochemical data were subjected were appropriate to statistical analysis. Mann-Whitney U test and χ^2 test were used to identify statistically significant differences between the EMILIN1 high and low groups. The Kaplan-Meier graphs were generated for the overall survival and statistical significance was determined by using the log-rank test. Univariate and multivariate survival analyses were performed using the Cox proportional hazards model. A p-value of < 0.05 was considered to indicate statistical significance. The statistical analyses were performed using JMP software (SAS Institute, Cary, NC, USA).

Results

The integration of single-cell RNA-seq and spatial transcriptomic data unveils functional heterogeneity across BC samples. We generated spatial transcriptomic data from four untreated invasive BC samples (A1, B1, C1, and D1; see Table S1 for the tumor classification). We also included published data on six BC samples (1160920F [TNBC], 1142243F [TNBC], CID4290 [ER+], CID4535 [ER+], CID4465 [TNBC], CID44971 [TNBC]) [23]. To gain a deeper insight into the cellular composition of each BC sample, we decided to assemble a single-cell RNA-seq BC atlas. To this end, we merged data from two recently published studies. The first one characterized all cell populations in 26 BC samples [23], and the second one characterized > 18,000 CAFs from 8 BC samples [25]. The resulting atlas is featured in Figure 1A, and representative cell-specific markers are in Figure 1B. Original cell annotations were kept from the respective studies, while their accuracy was additionally verified by inferring copy number variations (Figure S1A). While the original dataset of Wu et al. [23] contained a distinct CAF population (annotated as myCAF like and iCAF like), these cells perfectly superimposed (Figure 1C) the more detailed classification of Kieffer et al. [25]. Globally, all CAF could be divided in two major groups, the myCAF and iCAF, while they all belonged mainly to s1 type and none were of s4 type (as defined by Costa et al. [18]) (Figure 1D). We were unable to verify if any of the CAF were of s2 and s3 subtype as, to the best of our knowledge, no gene signatures of these subtypes are publicly available. However, these CAF subtypes have been reported in non-tumoral breast tissues [18] and as such were assumed as absent in the present BC atlas. Gene expression analysis of all CAF subtypes in the dataset (Figure 1D and Figure S1B) has clearly shown that CAF annotation from Wu et al. [23] could be replaced with the one from Kieffer et al. [25]. Indeed, the myCAF population as defined by Wu et al. had intermediate gene expression signature of all myCAF subpopulations as proposed by Kieffer et al. (Figure 1D). This was also observable for iCAF. Finally, the CAF nomenclature of Kieffer et al. could be seemingly transposed to Wu et al. dataset (Figure S2), where all the 8 CAF subpopulations were readily found. Consequently, we choose to further use the more nuanced CAF definition by Kieffer et al. in all subsequent analyses.

Following the establishment of the BC cell atlas and the clarification of CAF subpopulations, we next used it to annotate our spatial transcriptomic data (**Figures S3-S12**). Annotation relevance was checked relative to the presence of the typical histological structures observed in the H-E-stained histological sections (data not shown). Following the annotation process, visual inspection highlighted that individual cell subtypes (within a cellular population) compartmentalized differently in different BC samples. Knowing that the cell subtypes are characterized by distinct gene expression signatures (Table S3) we expected that this spatial heterogeneity would also imply а significant functional heterogeneity. To assess this, we performed a spatial gene ontology (GO) analysis using fGSEA. All BC specimens (n=10) showed a significant modulation of > 300 biological processes (data not shown) that could be grouped in two to three spatial patterns *per* sample (Figures 1E & S13). Among the spatially modulated GO processes, those relating to ECM remodeling and immunity regulation were particularly relevant for studying the CAF-immune cell interactions. Both processes showed a clear tissue compartmentalization, suggesting that such interactions may be enriched in specific BC tissue regions. TGF-β signaling was especially interesting due to its ability to suppress tumor immune response [30]. T-cell and macrophage activation processes displayed slightly different compartmentalization, but a spatial pattern opposite to that of TGF- β signaling in the majority of tumors. However, some tumor regions were rich in both TGF-β signaling and T-cell activation. This finding was particularly intriguing and motivated additional analyses.

BC areas with low proliferation potential are characterized by high macrophage and CD8⁺ T-cell infiltration. The spatial transcriptomic data projection on histological sections (Figures S3-S12) showed that BC samples were composed of distinct and heterogeneously distributed cancer cell subtypes. However, we observed a high degree of spatial consistency among cycling cancer cells and regions characterized by high-proliferative potential (Figure S14, S+G2M score plots). Spatial deconvolution of the immune cell infiltrate indicated that CD8+ T cells and macrophages were abundant in regions with low proliferative capacity (Figure 2A-B). The correlation analysis confirmed this pattern in 9/10 BC spatial datasets (the only exception was CID4535). When this correlation analysis was extended to include all cancer cells found in the BC atlas, the inversed correlation trend between CD8+ T cells and macrophages on one side and proliferative cancer cells on the other was confirmed (Figure S15). Since CD8+ T cells have a crucial role in tumor growth inhibition, we next determined whether locoregional differences in TGF- β signaling were correlated with the differential presence of CD8⁺ T cells. Indeed, TGF-β is a well-known master regulator of normal and pathologic inflammation. We did not find any significant correlation between TGF-B signaling and CD8⁺ T-cell abundance (data not shown), suggesting a more complex relationship between immune exclusion and TGF- β signaling. CAFs are major TGF- β

producers in tumors and also regulate TGF- β activity through the secretion of modulatory proteins [31]. Therefore, we hypothesized that CAF populations

and their tissue distribution might explain the link between TGF- β activity and CD8⁺T-cell exclusion.



Figure 1: BC atlas for cellular and functional annotation of spatial single-cell RNA-seq data. (A) UMAP plot showing the BC atlas based on two previously published single-cell RNA-seq datasets [23, 25]. (B) Validation of the cellular annotation using several cell-specific genes. (C) Enlarged UMAP plot of CAF subpopulations from the panel (A). (D) (left

and center) Scoring of myCAF, iCAF, CAF s1 and CAF s4 signatures in the CAF subpopulations from the panel (**C**); highest score is denoted in red, lowest score in dark blue. (*right*) Dot plot of marker genes delineating individual CAF subpopulations; same color-code. Genes used for scoring are outlined in **Table S3**. (**E**) Spatial distribution of selected GO processes in BC samples (two representative samples are shown: C1 and 1160920F; other samples are displayed in **Figure S13**). The following GO processes are displayed: ECM Structural Organization, Wound Healing, TGF-β Receptor Signaling, Regulation of Immune System, Macrophage Activation, and T-Cell Activation.



Figure 2: Spatial analysis of proliferating cancer cells and immune infiltrate in BC samples. (A) Histological annotation of two representative BC samples (other samples are shown in Figure S14), and estimation of highly proliferative regions (S+G2M phases) (*higher panels*); actively cycling cancer cells and two immune populations (macrophages and CD8⁺ T cells) (*middle and lower panels*). The heat map shows spatial correlation between these four populations. (B) Correlation analysis for the four selected cell populations in the other eight BC samples.

ECM-myCAFs, wound-myCAFs and TGFβmyCAFs are in regions with high TGF-β signaling. Recent single-cell studies [25] determined that in BC, there are several major CAF subpopulations: ECMmyCAFs, TGFβ-myCAFs, wound-myCAFs, IFNαβmyCAFs, acto-myCAFs, IFNy-iCAFs, detox-iCAFs, and IL-iCAFs. These CAF subpopulations are characterized by distinct gene expression profiles that suggest their involvement in specific cancer-relevant biological pathways. An overview of the GO enrichment analysis in each CAF subpopulations is provided in Figure S16. No significant acto-myCAF enrichment was observed. This was the smallest CAF population, at the periphery of the CAF cluster in Figure 1a. Next, we used single-cell RNA-seq data to spatially map CAF subpopulations in the BC samples. CAF subpopulations showed a rather compartmentalized distribution pattern (Figure 3A and Figures S3-S12). Overall, their distribution profiles could be classified in two main patterns (Figure 3A-B) that included ECM-myCAFs, wound-myCAFs and TGFβ-myCAFs (first pattern) and detox-iCAFs, IFNγ-iCAFs IL-iCAFs and (second pattern). Conversely, IFNaβ-myCAFs frequently grouped separately. We also found that high TGF- β signaling was associated with the first pattern (ECM-myCAFs, wound-myCAFs and TGF β -myCAFs). Having established a link between these three CAF subpopulations and TGF- β signaling, we wanted to understand its potential effect on the spatial localization of CD8⁺ T cells. Specifically, we asked why CD8⁺ T cells could accumulate in some BC areas that were rich in TGF- β signaling, although this is in contradiction with TGF-ß immune suppressor role (**Figure 4A**). We spatially scored TGF-β signaling in each BC sample using a gene set (Materials and Methods), and defined a tumor-specific high TGF- β area that corresponded to the top TGF- β signaling scores. Independently, we defined CD8+ T cell-rich and -poor areas in each tumor sample (i.e. the locations with top and bottom quarter CD8⁺ T-cell abundance scores, respectively). By intersecting these areas, we compared gene expression in CD8+ T cell-rich versus -poor areas within high TGF-B signaling locations (Figure 4A). Differential gene analysis in each BC sample identified the top modulated genes and their frequency (Figure 4B). Two genes emerged as significantly modulated in most BC samples: EMILIN1 and COL3A1. Both are related to TGF- β signaling, but in a different fashion. COL3A1 is produced by fibroblasts in response to TGF- β activation [32], whereas *EMILIN1* is an inhibitor of TGF- β signaling [33]. We were particularly interested in EMILIN1 because its

expression may modulate TGF-β activity and thus explain the selective CD8+ T-cell infiltration. A detailed comparison of EMILIN1 expression in CD8+T cell-low *versus* -high areas for each patient is provided in the Figure S17. Targeted analysis of the single-cell RNA-seq dataset reported by Wu et al. [23] revealed that EMILIN1 was a bona fide CAF gene, and was expressed only by myCAFs (Figure 4C). A more detailed analysis using the dataset reported by Kieffer et al. [25] showed that EMILIN1 was expressed by most CAF subpopulations, except IL-iCAFs. The strongest expression was observed in IFNy-iCAFs, followed by ECM-myCAFs, IFNaB-myCAFs and TGFβ-myCAFs (Figure 4D). Interestingly, wound-myCAFs, which showed the strongest expression of TGF- β signature genes (Figure 4E), displayed low EMILIN1 expression.

Spatial modulation of EMILIN1 expression coincides with CD8⁺ T-cell infiltration and is predictive of patient survival. To support the hypothesis that EMILIN1 expression is locally inhibiting TGF- β signaling, we monitored EMILIN1 and TGFBI spatial expression by immunofluorescence analysis in 5 patients with BC. We selected TGFBI because this protein is a known TGF-B activity reporter in cancer [34] and its expression is inversely correlated with CD8⁺ T-cell tumor infiltration [35]. TGFBIhigh and EMILIN1high CAFs constituted two distinct cell populations (Figure 5A and Figure S18A). CD8+ T cells were predominantly found in the EMILIN-rich areas, while they were excluded from regions with high TGFBI expression. As EMILIN1 expression is limited to CAFs and EMILIN1 functions as TGF- β activity suppressor [33], we examined its expression by immunohistochemistry in 75 patients with BC and its relationship with CD8⁺ T-cell infiltration. We found that EMILIN1 was clearly overexpressed in BC areas rich in infiltrating CD8+ T cells (Figure 5B-C). Moreover, breast cancer cases defined as EMILIN1-high had a significantly lower proportion of Ki-67-positive cancer cells (Figure 5D and Table S4). Indeed, 94 % of EMILIN1-low cases had Ki-67-positivity above 20 % threshold. This percentage sunk in EMILIN1-high cases to 62 %. Additional analysis in restricted number of EMILIN1-high cases showed that many CD8⁺ T cells infiltrating in the tumor were Ki-67 positive (Figure S18B, right panel, yellow arrowheads). In support of this observation, survival analysis showed that high EMILIN1 expression in BC was associated with increased survival (Figure 5E). Further correlation analysis between patient age, T factor, N factor, stage, hormonal status, HER2 status or subtype did not significant modulation between show any EMILIN1-high and -low groups (Table S4).



Figure 3: Spatial relationship between TGF- β signaling and CAF subpopulations in BC. (**A**) Spatial distribution of genes implicated in TGF- β signaling (top) and spatial distribution of different CAF subpopulations in two BC samples. The dendrogram (bottom, right) shows the spatial co-occurrence between CAF subpopulations and TGF- β signaling. (**B**) Dendrograms showing the co-occurrence of different CAF populations and TGF- β signaling in the other eight BC samples. (**A**-**B**) Labels TGFb, IFNab and IFNg refer to TGF β , IFN $\alpha\beta$ and IFN γ respectively.



Figure 4: Differential gene expression analysis of areas with high TGF- β signaling and with/without CD8⁺ T-cell exclusion. (**A**) Spatial distribution of areas with high versus low TGF- β signaling and presence/absence of CD8⁺ T cells in two BC samples (patient samples C1 and 1160920F are shown as examples; remaining patients, data not shown). (**B**) Differential gene expression analysis performed in all 8 BC samples; displayed are the number of patients in which each of the top-modulated genes was found as significantly overexpressed (in the areas where CD8⁺ T cells are present despite high TGF- β signaling). Highlighted in red are *EMILIN1* and *COL3A1*. (**C**) *EMILIN1* expression in the indicated cell subpopulation (from the BC atlas in **Figure 1A**). (**D**) *EMILIN1* expression in the indicated CAF subpopulations. (**E**) Upregulation of TGF- β signalurg genes in the indicated CAF subpopulations. The patient-wise statistical analysis of *EMILIN1* overexpression in CD8⁺ cells with high TGF- β signaling regions is provided in **Figure S17**. (**A**, **D**) Labels TGFb, IFNab and IFNg refer to TGF β , IFNa β and IFN γ respectively.



Figure 5: EMILINI is a good prognostic marker in BC. (A) Multiplexed immunofluorescence analysis displaying the localization of CD8⁺ T cells (red) and EMILINI expression (blue) in CAFs in a representative BC sample (N=5). Expression of TGFBI, a TGF- β signaling activity marker, was in green. Cell nuclei were examined by phase-contrast microscopy (data shown in Figure S18). (B) Multiplexed immunohistochemistry analysis showing examples of EMILINI (red) and CD8 (brown) co-staining in breast cancer samples (N=75, all subtypes; see also Table S2). (C) Violin plots of CD8⁺ cell counts in areas of high versus low EMILINI expression in BC samples (N=75). (D) Ki-67 positivity (as evaluated by retrospective analysis; see Table S4) in EMILINI-high versus EMILINI-low cases (N=70); (C-D) p-values were calculated using Mann-Whitney U test. (E) Survival analysis of patients with BC (N=75) in function of EMILINI expression level (high versus low). (D-E) The cut-off value of I was used to assign patients to EMILINI how group. The cut-off value was calculated as ratio of EMILINI score in the margin and the score in the central area. Clinical and pathological information regarding the patient cohort are displayed in the Table S2 and S4. Details on scoring methodology are provided in the Materials and Methods section.

Discussion

ICIs represent a major breakthrough for the systemic treatment of some tumor types and patient subpopulations. The success of immunotherapy is influenced by the tumor immunological status and the infiltration of cytotoxic CD8⁺ T cells. Although the underlying mechanisms of their infiltration are poorly understood, CD8⁺ T cells are one of the most relevant effector cell types recruited by ICIs [36, 37]. To shed additional light on CD8⁺ T-cell infiltration, the present

study used comprehensive single-cell RNA-seq BC datasets to project the different cell populations spatially in BC tissue sections. In accordance with the literature, we found a clear, opposed, spatial localisation between proliferating cancer cells and CD8⁺ T cells in BC. It has also been shown that the tissue localization of CD8⁺ T cells is important for the patient outcome, and CD8⁺ T cell presence in the tumor margin correlates with better clinical outcome [38, 39]. However, we do not know which stromal parameters influence CD8⁺ T-cell composition,

infiltration extent, localization, activation, or exhaustion in BC. This limits our ability to turn immunologically cold tumors into hot tumors and subject them to effective ICI treatment. TGF-B signaling in the tumor stroma mediates this immunomodulatory process. TGF- β is a potent immune suppressor with direct effects on the proliferation, differentiation and survival of various immune cell sub-populations [40, 41]. Experiments in mice suggest that TGF-β restricts CD8⁺ T-cell trafficking into tumors by suppressing CXCR3 expression [42]. These and other findings motivated the design of clinical trials to assess TGF-β blockade in combination with ICIs. However, the results were rather contrasted and surprisingly modest, in sharp contrast to the clear importance of TGF- β in tumor immunity [43]. The reason for this failure remains unclear, and might be related to the actual TGF- β activity, which is difficult to measure in situ. Indeed, TGF- β activity is modulated by factors secreted from the TME [31]. Therefore, TGF- β expression level may not reflect its actual activity.

To try to better understand TGF-β activity, we spatially correlated the relationship between CD8+ T-cell tumor infiltration, individual cancer cell populations, and TGF-B activity in BC. CAFs are a TME cell type with specific features: they are major TGF- β producers in the tumor and among the largest modulators of its activity by expressing soluble matrix proteins that can efficiently inhibit this cytokine [44, 45]. Therefore, it is not surprising that recent studies highlighted CAFs and some CAF subpopulations as key modulators of T-cell exclusion [46]. In the present study, spatial differences were mainly observed between the broad myCAF and iCAF subtypes, and myCAF were frequently spatially associated with higher TGF- β signaling. This is in agreement with previous studies in pancreatic cancer where such spatial heterogeneity was first reported [46, 47]. Previous spatial transcriptomic analysis of breast cancer [23] showed that CD8+/CD4+ T-cell tumor infiltration was more frequently associated with iCAF populations than myCAF. Despite the effort to further spatially deconvolute these major CAF populations, a distinction between finer spatial all CAF subpopulations was not accessible, possibly due to limitations of the current spatial and single-cell transcriptomic data depth. Therefore, we performed differential analysis not based on individual CAF subpopulations but on regions with high TGF- β signaling and different degrees of CD8+ T-cell infiltration. This allowed us to propose a molecular explanation of the overlaps between TGF-β-driven myCAFs and areas of CD8+ T-cell infiltration, despite

the known immunosuppressing effect of TGF- β . This analysis also highlighted a novel immune modulator protein, EMILIN1, that was previously reported as a TGF- β inhibitor. We found that EMILIN1 promotes CD8⁺ T-cell infiltration and is associated with better outcome in patients with BC.

The results highlight the fact that different CAF populations cannot be simply categorized as immune-promoting or -suppressing cells and that their status is finely tuned by the expression of modulator genes. Such modulators can mitigate the activity of key cytokines, such as TGF-B. Moreover, this finding suggests that CAF subpopulations can be largely regarded as cell programing states, probably with few exceptions. Such exceptions may occur in organs/tissues where different CAF sources are possible because of the intrinsic presence of different fibroblast-like cells (e.g. stellate cells in liver). Viewing CAF heterogeneity as cell states rather than actual subpopulations implies that harnessing CAFs for therapy would require their re-programing rather than the elimination of a specific CAF subpopulation. In this regard the current study highlights EMILIN1 as an important determinant of CAF anti-tumor program. Future studies should elucidate how EMILIN1 expression is modulated and how it could be upregulated in CAFs.

Supplementary Material

Supplementary figures and tables. https://www.thno.org/v14p1873s1.pdf

Acknowledgements

The authors want to thank Ms. Mariko Nakamura, Kao Abe, Kei Masuda, Kumiko Sudo, Yukiko Suto, Kyoko Miyoshi, Miyoko Suzuki, Chiho Noguchi, Saori Suto, Sayaka Okada, and Harumi Kanai for experimental assistance. This work was supported by grants from Gunma University. A.T. is supported by a LabEx MabImprove Grant. J.C. was supported by a Fondation ARC grant PJA 20141201975. No funding body had any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests

The authors have declared that no competing interest exists.

References

- Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. CA Cancer J Clin. 2021; 71: 7-33.
- Li J, Chen Z, Su K, Zeng J. Clinicopathological classification and traditional prognostic indicators of breast cancer. Int J Clin Exp Pathol. 2015; 8: 8500-5.
- 3. Moo TA, Sanford R, Dang C, Morrow M. Overview of breast cancer therapy. PET Clinic. 2018; 13: 339–354.

- Asleh K, Riaz N, Nielsen TO. Heterogeneity of triple negative breast cancer: Current advances in subtyping and treatment implications. J Exp Clin Cancer Res. 2022; 41: 265.
- Bejarano L, Jordão MJC, Joyce JA. Therapeutic targeting of the tumor microenvironment. Cancer Discov. 2021; 11: 933–959.
- Anderson NM, Simon MC. The tumor microenvironment. Curr Biol. 2020; 30: R921-R925.
- Atkins MB, Lee SJ, Chmielowski B, Tarhini AA, Cohen GI, Truong TG, et al. Combination dabrafenib and trametinib versus combination nivolumab and ipilimumab for patients with advanced BRAF -mutant melanoma: The DREAMseq Trial – ECOG-ACRIN EA6134. J Clin Oncol. 2023; 41:186–197.
- Tarhini AA, Toor K, Chan K, McDermott DF, Mohr P, Larkin J, et al. A matching-adjusted indirect comparison of combination nivolumab plus ipilimumab with BRAF plus MEK inhibitors for the treatment of BRAF-mutant advanced melanoma. ESMO Open. 2021; 6: 100050.
- FinnRS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim TY, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. N Engl J Med. 2020; 382: 1894–1905.
- Miller KD, Chap LI, Holmes FA, Cobleigh MA, Marcom PK, Fehrenbacher L, et al. Randomized phase III trial of capecitabine compared with bevacizumab plus capecitabine in patients with previously treated metastatic breast cancer. J Clin Oncol. 2005; 23: 792–799.
- Cameron D, Brown J, Dent R, Jackisch C, Mackey J, Pivot X, et al. Adjuvant bevacizumab-containing therapy in triple-negative breast cancer (BEATRICE): primary results of a randomised, phase 3 trial. Lancet Oncol. 2013; 14: 933–942.
- Ayoub NM, Jaradat SK, Al-Shami KM, Alkhalifa AE. Targeting angiogenesis in breast cancer: current evidence and future perspectives of novel anti-angiogenic approaches. Front Pharmacol. 2022; 13: 838133.
- Thomas R, Al-Khadairi G, Decock J. Immune checkpoint inhibitors in triple negative breast cancer treatment: promising future prospects. Front Oncol. 2021; 10: 600573.
- Swoboda A, Nanda R. Immune checkpoint blockade for breast cancer. Cancer Treat Res. 2018; 173: 155–165.
- Schmid P, Cortes J, Dent R, Pusztai L, McArthur H, Kümmel S, et al. Event-free survival with pembrolizumab in early triple-negative breast cancer. N Engl J Med. 2022; 386: 556-567.
- Cortes J, Rugo HS, Cescon DW, Im SA, Yusof MM, Gallardo C, et al. Pembrolizumab plus chemotherapy in advanced triple-negative breast cancer. N Engl J Med. 2022; 387: 217-226.
- Mhaidly R, Mechta-Grigoriou F. Role of cancer-associated fibroblast subpopulations in immune infiltration, as a new means of treatment in cancer. Immunol Rev. 2021; 302: 259–272.
- 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; 33: 463-479. e10.
- 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; 14: 4294.
- Van Oekelen O, Laganà A. Multi-omics profiling of the tumor microenvironment. Adv Exp Med Biol. 2022; 1361: 283–326.
- Guo W, Zhou B, Yang Z, Liu X, Huai Q, Guo L, et al. Integrating microarray-based spatial transcriptomics and single-cell RNA-sequencing reveals tissue architecture in esophageal squamous cell carcinoma. EBioMedicine. 2022; 84: 104281.
- 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; 24: 541–550.
- 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; 53: 1334–1347.
- Villemin JP, Bassaganyas L, Pourquier D, Boissière F, Cabello-Aguilar S, Crapez E, et al. Inferring ligand-receptor cellular networks from bulk and spatial transcriptomic datasets with BulkSignalR. Nucleic Acids Res. 2023; 51: 4726–4744.
- 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; 10: 1330–1351.
- Jew B, Alvarez M, Rahmani E, Miao Z, Ko A, Garske KM, et al. Accurate estimation of cell composition in bulk expression through robust integration of single-cell information. Nat Commun. 2020; 11: 1971.
- Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014; 344: 1396-401.
- Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. BioRxiv. doi: https://doi.org/10.1101/060012
- McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 2012; 40: 4288-97.
- Tie Y, Tang F, Peng D, Zhang Y, Shi H. TGF-beta signal transduction: biology, function and therapy for diseases. Mol Biomed. 2022; 3: 45.
- 31. Costanza B, Umelo IA, Bellier J, Castronovo V, Turtoi A. Stromal modulators of TGF- β in cancer. J Clin Med. 2017; 6: 7.

- Verrecchia F, Chu ML, Mauviel A. Identification of novel TGF-β/Smad gene targets in dermal fibroblasts using a combined cDNA microarray/promoter transactivation approach. J Biol Chem. 2001; 276: 17058-17062.
- Zacchigna L, Vecchione C, Notte A, Cordenonsi M, Dupont S, Maretto S, et al. Emilin1 Links TGF-β maturation to blood pressure homeostasis. Cell. 2006; 124: 929–942.
- Chiavarina B, Costanza B, Ronca R, Blomme A, Rezzola S, Chiodelli P, et al. Metastatic colorectal cancer cells maintain the TGFβ program and use TGFBI to fuel angiogenesis. Theranostics. 2021; 11: 1626-1640.
- 35. Nakazawa N, Yokobori T, Kaira K, Turtoi A, Baatar S, Gombodorj N, et al. High stromal TGFBI in lung cancer and intratumoral CD8-positive T Cells were associated with poor prognosis and therapeutic resistance to immune checkpoint inhibitors. Ann Surg Oncol. 2020; 27: 933-942.
- Li F, Li C, Cai X, Xie Z, Zhou L, Cheng B, et al. The association between CD8+ tumor-infiltrating lymphocytes and the clinical outcome of cancer immunotherapy: A systematic review and meta-analysis. EClinicalMedicine. 2021; 41: 101134.
- Kist de Ruijter L, van de Donk PP, Hooiveld-Noeken JS, Giesen D, Elias SG, Lub-de Hooge MN, et al. Whole-body CD8+ T cell visualization before and during cancer immunotherapy: a phase 1/2 trial. Nat Med. 2022; 28: 2601-2610.
- Loi S, Sirtaine N, Piette F, Salgado R, Viale G, Van Eenoo F, et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. J Clin Oncol. 2013; 31: 860-867.
- Mahmoud SM, Paish EC, Powe DG, Macmillan RD, Grainge MJ, Lee AH, et al. Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer. J Clin Oncol. 2011; 29: 1949-55.
- Li MO, Flavell RA. TGF-beta: a master of all T cell trades. Cell. 2008; 134 : 392-404.
- Li MO, Wan YY, Sanjabi S, Robertson AK, Flavell RA. Transforming growth factor-beta regulation of immune responses. Annu Rev Immunol. 2006; 24: 99-146.
- Gunderson AJ, Yamazaki T, McCarty K, Fox N, Phillips M, Alice A, et al. TGFβ suppresses CD8+ T cell expression of CXCR3 and tumor trafficking. Nat Commun. 2020; 11: 1749.
- Metropulos AE, Munshi HG, Principe DR. The difficulty in translating the preclinical success of combined TGFβ and immune checkpoint inhibition to clinical trial. EBioMedicine. 2022; 86: 104380.
- Maris P, Blomme A, Palacios AP, Costanza B, Bellahcène A, Bianchi E, et al. Asporin is a fibroblast-derived TGF-β1 inhibitor and a tumor suppressor associated with good prognosis in breast cancer. PLoS Med. 2015; 12: e1001871.
- Xie C, Mondal DK, Ulas M, Neill T, Iozzo RV. Oncosuppressive roles of decorin through regulation of multiple receptors and diverse signaling pathways. Am J Physiol Cell Physiol. 2022; 322: C554-C566.
- Maia A, Schöllhorn A, Schuhmacher J, Gouttefangeas C. CAF-immune cell crossstalk and its impact in immunotherapy. Semin Immunopathol. 2023; 45: 203–214.
- 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; 214: 579–596.