

# Cancer-associated Fibroblast Spatial Heterogeneity and *EMILIN1* Expression in Cancer Stroma Modulate TGF- $\beta$ Activity and CD8 $^{+}$ T-Cell Infiltration in Breast Cancer

Chikako Honda<sup>1</sup>, Sasagu Kurozumi<sup>1,2</sup>, Takaaki Fujii<sup>1</sup>, Didier Pourquier<sup>3,4,5</sup>, Lakhdar Khellaf<sup>3,4,5</sup>, Florence Boissiere<sup>3</sup>, Jun Horiguchi<sup>2</sup>, Tetsunari Oyama<sup>6</sup>, Ken Shirabe<sup>1</sup>, Jacques Colinge<sup>3,5,7\*</sup>, Takehiko Yokobori<sup>8\*</sup>, Andrei Turtoi<sup>3,4,5,8\*</sup>

<sup>1</sup>Department of General Surgical Science, Gunma University Graduate School of Medicine, Maebashi, Gunma, Japan

<sup>2</sup>Department of Breast Surgery, International University of Health and Welfare, Chiba, Japan

<sup>3</sup>Institut régional du Cancer de Montpellier (ICM)-Val d'Aurelle, Montpellier, France

<sup>4</sup>Tumor Microenvironment and Resistance to Treatment Lab, INSERM U1194, Montpellier, France  
<sup>5</sup>Université de Montpellier, Montpellier, France

<sup>6</sup>Department of Pathology and Diagnostics, Gunma University Graduate School of Medicine, Maebashi, Gunma, Japan

<sup>7</sup>Cancer Bioinformatics and Systems Biology Team, INSERM U1194, Montpellier, France

<sup>8</sup>Division of Integrated Oncology Research, Gunma University, Initiative for Advanced Research (GIAR), Maebashi, Gunma, Japan

**Running title:** CAF-Mediated T-Cell Exclusion in Breast Cancer.

\* Co-corresponding authors/ contributed equally.

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

## ABSTRACT

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 frequently precludes conclusions on function. Therefore, 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. 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 IFN $\gamma$ -iCAFs tended to cluster together, while Wound-myCAFs, TGF $\beta$ -myCAFs, and ECM-myCAFs formed another group that overlapped with elevated TGF- $\beta$  signaling. Differential gene expression analysis of areas with CD8 $^+$  T-cell infiltration/exclusion within the TGF- $\beta$  signaling-rich zones identified elastin microfibrillar interface protein 1 (*EMILIN1*) as a top modulated gene. *EMILIN1*, a TGF- $\beta$  inhibitor, was upregulated in IFN $\gamma$ -iCAFs directly modulating TGF $\beta$  immunosuppressive function. Histological analysis of 74 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. In conclusion, our data show that correlating TGF- $\beta$  signaling to a CAF subpopulation is not enough because proteins with TGF- $\beta$ -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.

## INTRODUCTION

Breast cancer (BC) is the second most frequent cause of cancer death in women worldwide<sup>1</sup>. Molecular and histological classifications of BC have significantly improved its clinical management. Today, histopathological assessment of needle biopsies for morphological type and histological grade determination are complemented by the assessment of estrogen receptor (ER), progesterone receptor (PgR), human epidermal growth factor receptor type 2 (HER2) status, and Ki-67 proliferative index<sup>2</sup>. Therefore, BC are classified into three types: (a) hormone receptor (ER and/or PgR)-positive and HER2-negative, (b) HER2-positive, and (c) hormone receptor- and HER2-negative (triple negative breast cancer, TNBC)<sup>2</sup>. Based on this classification, drug treatment regimens for early invasive BC include endocrine therapy (a) and chemotherapy and anti-HER2 agents (b), alone or in combination<sup>3</sup>. No specific treatment is available for TNBC (c), and this represents a tremendous clinical challenge<sup>3,4</sup>. Indeed, no druggable vulnerability has been identified in TNBC cells to date, precluding their direct targeting<sup>4,5</sup>.

A complementary approach to cancer cell targeting is to target the environment in which they reside<sup>6</sup> and that is called tumor microenvironment (TME). The TME consists of a cellular part (stroma) and a supportive extracellular matrix (ECM) with specific physical and chemical properties. The stroma is primarily composed of endothelial, immune and fibrotic components. All three have attracted considerable attention for novel drug development in solid tumors. Indeed, in some cancers, targeting endothelial and immune cells is more effective than killing cancer cells (e.g. melanoma<sup>7,8</sup> and hepatocellular<sup>9</sup> carcinoma). This demonstrates the potential of TME-directed therapies, possibly in combination with molecules against cancer cells. Despite these encouraging results, BC (like many other solid tumors) has not really benefited from TME targeting yet. Clinical trials produced rather mitigated results. For example, in BC, endothelial cells have been mainly targeted with anti-angiogenic drugs (e.g. anti-vascular endothelial growth factor (VEGF) antibodies), alone or in combination with chemotherapy<sup>10,11</sup>. Unfortunately, the survival benefit for patients with BC was minimal, and several potential resistance mechanisms were described<sup>12</sup>. The immune component of BC has been mainly targeted using novel monoclonal antibodies against immune checkpoint proteins (e.g. programmed cell-death protein 1 (PD-1) and its ligand PD-L1) with the aim of restoring the anti-tumor immunity. The immune checkpoint inhibitors (ICI) pembrolizumab (anti-PD-1) and atezolizumab (anti-PD-L1) have been clinically tested in patients with metastatic TNBC with heterogeneous results<sup>13,14</sup>. A good clinical response is achieved in a small subpopulation of patients, and no clear biomarker exists to predict which patients will respond. Patients with BC characterized by high mutational burden or with immunologically inflamed tumors (with high proportion of CD8<sup>+</sup> T cells in the tumor center) might respond to ICIs<sup>15,16</sup>. Indeed, ICIs that activate cytotoxic T cells against tumors are now considered an important therapeutic tool<sup>17,18</sup>. Lastly, clinical trials on cancer-associated fibroblasts (CAFs) and the targeting of the tumor fibrotic component have not brought any conclusive results in BC, leaving this area unexplored. Despite the plethora of experimental data on CAF significance in BC progression<sup>19</sup>, very few targetable

molecules have been identified in CAFs. Additionally, recent data on CAF subpopulations in BC<sup>20</sup> raise the question of whether specific subtypes should be targeted.

However, we believe that successful TME targeting should be process-oriented and not cell-oriented. Indeed, immune exclusion and angiogenesis are promoted and regulated by the concerted action of several stromal cell types, and therefore disrupting a specific cell population is unlikely to abolish such processes 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 engage to support tumor-promoting programs.

Recent advances in spatial OMICS technologies have been a true game changer for characterizing the TME and tumor heterogeneity<sup>21</sup>. They have allowed, for the first time, to link the spatial occurrence of different TME cell subtypes and of cancer cells with enhanced proliferative or therapy-resistance features<sup>22</sup>. In the present study, we investigated the spatial distribution of CAF subpopulations in BC and their relationship with infiltrating cytotoxic T cells.

## MATERIALS AND METHODS

### Patient Material

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 2020-2021 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) primers to capture the mRNA. After 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  $\mu$ m-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: Read1- 28 bp and Read2 - 100bp.

### 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<sup>23</sup> 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<sup>23</sup>, 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<sup>24</sup> 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<sup>23</sup> and for CAFs<sup>25</sup>. 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<sup>26</sup> to these gene signatures due to the bulk nature of Visium spatial data. A first scoring of cycling cancer cells was obtained using BisqueRNA scores for the Cycling population<sup>23</sup>. 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)<sup>27</sup> spatial transcriptomic features. TGF- $\beta$  signaling was scored with the BulkSignalR scoreSignatures function to generate dendrograms to relate this process to CAF subpopulations. TGF- $\beta$  signaling genes were obtained from

MSSigDB (c5.bp.v7.0.symbols.gmt.txt,

GO\_TRANSFORMING\_GROWTH\_FACTOR\_BETA\_RECECTOR\_SIGNALING\_PATHWAY). Spatial co-localization between cell populations or between TGF- $\beta$  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- $\beta$  signaling tumor areas was performed with edgeR (PMID: 22287627) 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  $\mu$ m-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-EMILIN-1 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, EMILIN-1 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.

EMILIN-1 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 EMILIN-1-stained 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 EMILIN-1 score (0-12). BC samples with higher EMILIN-1 score in the margin than central area were defined as a high EMILIN-1 group, and the others as low EMILIN-1 group. The total number of CD8<sup>+</sup> cells was counted in the 200x view fields where the EMILIN-1 score was evaluated.

### Immunofluorescence analysis

Multicolor immunofluorescence staining was performed in tissue sections of BC surgically resected from five patients to detect EMILIN-1, CD8, and TGFBI expression and from seven patients to detect EMILIN-1, 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, EMILIN-1 staining (anti-EMILIN-1 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 (**Figure 5a**). In the other seven BC samples, EMILIN-1 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 (**Figure 5c**). All sections were lightly counterstained with hematoxylin and examined under an All-in-One BZ-X710 fluorescence microscope (KEYENCE Corporation, Osaka, Japan).

## 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 D; 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])<sup>23</sup>. 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<sup>23</sup>, and the second one characterized >18,000 CAFs from 8 BC samples<sup>25</sup>. To avoid redundancy, we removed the CAF populations from the first study and kept only the CAF populations from the second study. The resulting atlas is featured in **Figure 1a**, and representative cell-specific markers are in **Figure 1b**. Then, we used this BC cell atlas to annotate spatial transcriptomic data for our ten BC samples and to project each cell population on histological sections (**Figures S1-S10**). The annotation relevance was verified relative to the presence of the typical histological structures observed in the H-E stained histological sections (e.g. vasculature, fibrosis, *in situ* versus invasive cancer). An example is shown in **Figure 1c**. This annotation process highlighted that individual cell types compartmentalized differently in different BC samples, suggesting 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 1d & S11**). 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 (**Figure 1d**). TGF- $\beta$  signaling was particularly interesting due to its capacity to suppress tumor immune response<sup>28</sup>. The T-cell and macrophage activation processes showed a distinct compartmentalization, but with a spatial pattern opposite to that of TGF- $\beta$  signaling. However, some tumor regions were rich in both TGF- $\beta$  signaling and T-cell activation. This finding was particularly intriguing and required additional analyses.

**BC areas with low proliferation potential are characterized by high macrophage and CD8<sup>+</sup> T-cell infiltration.** The spatial transcriptomic data projection on histological sections (**Figures S1-S10**) 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 S12**, S+G2M score plots). Spatial deconvolution of the immune cell infiltrate indicated that CD8<sup>+</sup> 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). As CD8<sup>+</sup> T cells have a crucial role in tumor growth inhibition, we next determined whether locoregional differences in TGF- $\beta$  signaling were correlated with the differential presence of CD8<sup>+</sup> T cells. Indeed, TGF- $\beta$  is a well-known master regulator of normal and pathologic inflammation. We did not find any significant correlation between TGF- $\beta$  signaling and CD8<sup>+</sup> T-cell abundance (data not shown), suggesting a more complex relationship between immune exclusion and TGF- $\beta$  signaling. CAFs are major TGF- $\beta$  producers in tumors and also regulate TGF- $\beta$  activity through the secretion of modulatory proteins<sup>29</sup>. Therefore, we hypothesized that CAF populations and their tissue distribution might explain the link between TGF- $\beta$  activity and CD8<sup>+</sup> T-cell exclusion.

**ECM-myCAFs, wound-myCAFs and TGF $\beta$ -myCAFs are in regions with high TGF- $\beta$  signaling.** Recent single-cell studies<sup>25</sup> determined that in BC, there are several major CAF subpopulations: ECM-myCAF, TGF $\beta$ -myCAF, wound-myCAF, IFN $\alpha\beta$ -myCAF, acto-myCAF, IFN $\gamma$ -iCAF, detox-iCAF, and IL-iCAF. 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 S13**. 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 S1-S10**). Overall, their distribution profiles could be classified in two main patterns (**Figure 3a-b**) that included ECM-myCAF, wound-myCAF and TGF $\beta$ -myCAF (first pattern) and detox-iCAF, IL-iCAF and IFN $\gamma$ -iCAF (second

pattern). Conversely, IFN $\alpha\beta$ -myCAFs frequently grouped separately. We also found that high TGF- $\beta$  signaling was associated with the first pattern (ECM-myCAFs, wound-myCAFs and TGF $\beta$ -myCAFs). Having established a link between these three CAF subpopulations and TGF- $\beta$  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- $\beta$  signaling, although this is in contradiction with TGF- $\beta$  immune suppressor role (**Figure 4a**). We spatially scored TGF- $\beta$  signaling in each BC sample using a gene set (Materials and Methods), and defined a tumor-specific high TGF- $\beta$  area that corresponded to the top TGF- $\beta$  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- $\beta$  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- $\beta$  signaling, but in a different fashion. *COL3A1* is produced by fibroblasts in response to TGF- $\beta$  activation<sup>30</sup>, whereas *EMILIN1* is an inhibitor of TGF- $\beta$  signaling<sup>31</sup>. We were particularly interested in *EMILIN1* because its expression may modulate TGF- $\beta$  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 S14**. Targeted analysis of the single-cell RNA-seq dataset reported by Wu et al.<sup>23</sup> 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.<sup>25</sup> showed that *EMILIN1* was expressed by most CAF subpopulations, except IL-iCAFs. The strongest expression was observed in IFN $\gamma$ -iCAFs, followed by ECM-myCAFs, IFN $\alpha\beta$ -myCAFs and TGF $\beta$ -myCAFs (**Figure 4d**). Interestingly, wound-myCAFs, which showed the strongest expression of TGF- $\beta$  signature genes (**Figure 4e**), displayed low *EMILIN1* expression.

**Spatial modulation of EMILIN-1 expression coincides with CD8 $^+$  T-cell infiltration and is predictive of patient survival.** To support the hypothesis that EMILIN1 expression is locally inhibiting TGF- $\beta$  signaling, we monitored EMILIN-1 and TGFBI spatial expression by immunofluorescence analysis in 5 patients with BC. We selected TGFBI because this protein is a known TGF- $\beta$  activity reporter in cancer<sup>32</sup> and its expression is inversely correlated with CD8 $^+$  T-cell tumor infiltration<sup>33</sup>. TGFBI $^{\text{high}}$  and EMILIN-1 $^{\text{high}}$  CAFs constituted two distinct cell populations (**Figure 5a**). CD8 $^+$  T cells were predominantly found in the EMILIN-rich areas, while they were excluded from regions with high TGFBI expression. As EMILIN-1 expression is limited to CAFs and EMILIN-1 functions as TGF- $\beta$  activity suppressor<sup>34</sup>, we examined its expression by immunohistochemistry in 75 patients with BC and its relationship with CD8 $^+$  T-cell infiltration. We found that EMILIN-1 was clearly overexpressed in BC areas rich in infiltrating CD8 $^+$  T cells (**Figure 5b-c**). Moreover, EMILIN-1-rich areas had a significant proportion of Ki-67-negative cancer cells (**Figure 5c-e**), whereas many CD8 $^+$  T cells expressed Ki-67 (**Figure 5c**, yellow arrowheads). Lastly,

survival analysis showed that high EMILIN-1 expression in BC was associated with increased survival (**Figure 5f**). The results of the multivariate analysis of EMILIN-1 expression in this patient cohort are in **Table S3**.

## DISCUSSION

ICIs have been 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<sup>+</sup> T cells. Although the underlying mechanisms of their infiltration are poorly understood, CD8<sup>+</sup> T cells are one of the most relevant effector cell types recruited by ICIs<sup>35,36</sup>. To shed additional light on CD8<sup>+</sup> 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, opposite, spatial correlation between proliferating cancer cells and CD8<sup>+</sup> T cells in BC. It has also been shown that the tissue localization of CD8<sup>+</sup> T cells is important for the patient outcome. CD8<sup>+</sup> T cell presence in the tumor margin correlates with better clinical outcome<sup>37,38</sup>. However, we do not know which stromal parameters influence CD8<sup>+</sup> T-cell composition, infiltration extent, activation, or exhaustion in BC. This limits our ability to turn immunologically cold tumors into hot tumors and in consequence subject them to effective ICI treatment. TGF- $\beta$  signaling in the tumor stroma mediates this immunomodulatory process. TGF- $\beta$  is a potent immune suppressor with direct effects on the proliferation, differentiation and survival of various immune cell sub-populations<sup>39,40</sup>. Experiments in mice suggest that TGF- $\beta$  restricts CD8<sup>+</sup> T-cell trafficking into tumors by suppressing CXCR3 expression<sup>41</sup>. These and other findings motivated the design of clinical trials to assess TGF- $\beta$  blockade in combination with ICIs. However, the results were rather contrasted and surprisingly modest, in sharp contrast to the clear importance of TGF- $\beta$  in tumor immunity<sup>42</sup>. The reason for this failure remains unclear, and might be related to the actual TGF- $\beta$  activity, which is difficult to measure *in situ*. Indeed, TGF- $\beta$  activity is modulated by factors secreted from the TME<sup>29</sup>. Therefore, TGF- $\beta$  expression level may not be necessarily directly related to its activity level. To try to shed light into TGF- $\beta$  activity, we spatially correlated the relationship between CD8<sup>+</sup> T-cell tumor infiltration, individual cancer cell populations, and TGF- $\beta$  activity in BC. CAFs are a TME cell type with specific features: they are major TGF- $\beta$  producers in the tumor and among the largest modulators of its activity by expressing soluble matrix proteins that can efficiently inhibit this cytokine<sup>43,44</sup>. Therefore, it is not surprising that recent studies highlighted CAFs and some CAF subpopulations as key modulators of T-cell exclusion<sup>45</sup>. 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- $\beta$  signaling. This is in line with previous studies in pancreatic cancer where such spatial heterogeneity was first reported<sup>45,46</sup>. A finer spatial distinction between 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- $\beta$  signaling and different degrees of

CD8<sup>+</sup> T-cell infiltration. This allowed us to propose a molecular explanation of the overlaps between TGF- $\beta$ -driven myCAF<sup>s</sup> and areas of CD8<sup>+</sup> T-cell infiltration, despite the known immunosuppressing effect of TGF- $\beta$ . This analysis also highlighted a novel immune modulator protein, EMILIN-1, that was previously reported as a TGF- $\beta$  inhibitor. We found that EMILIN-1 promotes CD8<sup>+</sup> 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- $\beta$ . 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 EMILIN-1 as an important determinant of CAF anti-tumor program. Future studies should elucidate how EMILIN-1 expression is modulated and how it could be upregulated in CAFs.

## 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 Starting 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.

## REFERENCES

1. Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2021). Cancer statistics, 2021. *Ca Cancer J Clin*, 71(1), 7-33.
2. Li, J., Chen, Z., Su, K., and Zeng, J. Clinicopathological classification and traditional prognostic indicators of breast cancer.
3. Moo, T.-A., Sanford, R., Dang, C., and Morrow, M. (2018). Overview of Breast Cancer Therapy. *PET Clinics* 13, 339–354. [10.1016/j.cpet.2018.02.006](https://doi.org/10.1016/j.cpet.2018.02.006).
4. Asleh, K., Riaz, N., and Nielsen, T.O. (2022). Heterogeneity of triple negative breast cancer: Current advances in subtyping and treatment implications. *J Exp Clin Cancer Res* 41, 265. [10.1186/s13046-022-02476-1](https://doi.org/10.1186/s13046-022-02476-1).
5. Deepak, K. G. K., Vempati, R., Nagaraju, G. P., Dasari, V. R., S, N., Rao, D. N., & Malla, R. R. (2020). Tumor microenvironment: Challenges and opportunities in targeting metastasis of triple negative breast cancer. *Pharmacological research*, 153, 104683. [10.1016/j.phrs.2020.104683](https://doi.org/10.1016/j.phrs.2020.104683)

6. Bejarano, L., Jordāo, M.J.C., and Joyce, J.A. (2021). Therapeutic Targeting of the Tumor Microenvironment. *Cancer Discovery* 11, 933–959. [10.1158/2159-8290.CD-20-1808](https://doi.org/10.1158/2159-8290.CD-20-1808).
7. Atkins, M.B., Lee, S.J., Chmielowski, B., Tarhini, A.A., Cohen, G.I., Truong, T.-G., Moon, H.H., Davar, D., O'Rourke, M., Stephenson, J.J., et al. (2023). Combination Dabrafenib and Trametinib Versus Combination Nivolumab and Ipilimumab for Patients With Advanced *BRAF* -Mutant Melanoma: The DREAMseq Trial—ECOG-ACRIN EA6134. *JCO* 41, 186–197. [10.1200/JCO.22.01763](https://doi.org/10.1200/JCO.22.01763).
8. Tarhini, A.A., Toor, K., Chan, K., McDermott, D.F., Mohr, P., Larkin, J., Hodi, F.S., Lee, C.-H., Rizzo, J.I., Johnson, H., et al. (2021). 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* 6, 100050. [10.1016/j.esmoop.2021.100050](https://doi.org/10.1016/j.esmoop.2021.100050).
9. Finn, R.S., Qin, S., Ikeda, M., Galle, P.R., Ducreux, M., Kim, T.-Y., Kudo, M., Breder, V., Merle, P., Kaseb, A.O., et al. (2020). Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *N Engl J Med* 382, 1894–1905. [10.1056/NEJMoa1915745](https://doi.org/10.1056/NEJMoa1915745).
10. Miller, K.D., Chap, L.I., Holmes, F.A., Cobleigh, M.A., Marcom, P.K., Fehrenbacher, L., Dickler, M., Overmoyer, B.A., Reimann, J.D., Sing, A.P., et al. (2005). Randomized Phase III Trial of Capecitabine Compared With Bevacizumab Plus Capecitabine in Patients With Previously Treated Metastatic Breast Cancer. *JCO* 23, 792–799. [10.1200/JCO.2005.05.098](https://doi.org/10.1200/JCO.2005.05.098).
11. Cameron, D., Brown, J., Dent, R., Jackisch, C., Mackey, J., Pivot, X., Steger, G. G., Suter, T. M., Toi, M., Parmar, M., Laeufle, R., Im, Y. H., Romieu, G., Harvey, V., Lipatov, O., Pienkowski, T., Cottu, P., Chan, A., Im, S. A., Hall, P. S., ... Bell, R. (2013). Adjuvant bevacizumab-containing therapy in triple-negative breast cancer (BEATRICE): primary results of a randomised, phase 3 trial. *The Lancet. Oncology*, 14(10), 933–942. [10.1016/S1470-2045\(13\)70335-8](https://doi.org/10.1016/S1470-2045(13)70335-8)
12. Ayoub, N.M., Jaradat, S.K., Al-Shami, K.M., and Alkhalifa, A.E. (2022). Targeting Angiogenesis in Breast Cancer: Current Evidence and Future Perspectives of Novel Anti-Angiogenic Approaches. *Front. Pharmacol.* 13, 838133. [10.3389/fphar.2022.838133](https://doi.org/10.3389/fphar.2022.838133).
13. Thomas, R., Al-Khadairi, G., and Decock, J. (2021). Immune Checkpoint Inhibitors in Triple Negative Breast Cancer Treatment: Promising Future Prospects. *Front. Oncol.* 10, 600573. [10.3389/fonc.2020.600573](https://doi.org/10.3389/fonc.2020.600573).
14. Swoboda, A., and Nanda, R. (2018). Immune Checkpoint Blockade for Breast Cancer. In *Optimizing Breast Cancer Management Cancer Treatment and Research.*, W. J. Gradishar, ed. (Springer International Publishing), pp. 155–165. [10.1007/978-3-319-70197-4\\_10](https://doi.org/10.1007/978-3-319-70197-4_10).
15. Narang, P., Chen, M., Sharma, A.A., Anderson, K.S., and Wilson, M.A. (2019). The neoepitope landscape of breast cancer: implications for immunotherapy. *BMC Cancer* 19, 200. [10.1186/s12885-019-5402-1](https://doi.org/10.1186/s12885-019-5402-1).
16. Miyan, M., Schmidt-Mende, J., Kiessling, R., Poschke, I., and De Boniface, J. (2016). Differential tumor infiltration by T-cells characterizes intrinsic molecular subtypes in breast cancer. *J Transl Med* 14, 227. [10.1186/s12967-016-0983-9](https://doi.org/10.1186/s12967-016-0983-9).

17. Schmid P., Cortes J., Dent R., Pusztai L., McArthur H., Kümmel S., Bergh J., Denkert C., Park Y.H., Hui R. et al. (2022). Event-free Survival with Pembrolizumab in Early Triple-Negative Breast Cancer. *N Engl J Med.* 386 (6): 556-567. 10.1056/NEJMoa2112651.
18. Cortes J., Rugo H.S., Cescon D.W., Im S.A., Yusof M.M., Gallardo C., Lipatov O., Barrios C.H., Perez-Garcia J., Iwata H. et al. (2022). Pembrolizumab plus Chemotherapy in Advanced Triple-Negative Breast Cancer. *N Engl J Med.* 387 (3): 217-226. 10.1056/NEJMoa2202809.
19. Mhaidly, R., and Mechta-Grigoriou, F. (2021). Role of cancer-associated fibroblast subpopulations in immune infiltration, as a new means of treatment in cancer. *Immunological Reviews* 302, 259–272. [10.1111/imr.12978](https://doi.org/10.1111/imr.12978).
20. Costa, A., Kieffer, Y., Scholer-Dahirel, A., Pelon, F., Bourachot, B., Cardon, M., Sirven, P., Magagna, I., Fuhrmann, L., Bernard, C., et al. (2018). Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* 33, 463-479.e10. [10.1016/j.ccr.2018.01.011](https://doi.org/10.1016/j.ccr.2018.01.011).
21. Van Oekelen, O., & Laganà, A. (2022). Multi-Omics Profiling of the Tumor Microenvironment. *Advances in experimental medicine and biology*, 1361, 283–326. 10.1007/978-3-030-91836-1\_16
22. Binnewies, M., Roberts, E.W., Kersten, K., Chan, V., Fearon, D.F., Merad, M., Coussens, L.M., Gabrilovich, D.I., Ostrand-Rosenberg, S., Hedrick, C.C., et al. (2018). Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med* 24, 541–550. [10.1038/s41591-018-0014-x](https://doi.org/10.1038/s41591-018-0014-x).
23. Wu, S. Z., Al-Eryani, G., Roden, D. L., Junankar, S., Harvey, K., Andersson, A., Thennavan, A., Wang, C., Torpy, J. R., Bartonicek, N., Wang, T., Larsson, L., Kaczorowski, D., Weisenfeld, N. I., Uytingco, C. R., Chew, J. G., Bent, Z. W., Chan, C. L., Gnanasambandapillai, V., Dutertre, C. A., ... Swarbrick, A. (2021). A single-cell and spatially resolved atlas of human breast cancers. *Nature genetics*, 53(9), 1334–1347. 10.1038/s41588-021-00911-1
24. Villemain, J.-P., Bassaganyas, L., Pourquier, D., Boissière, F., Cabello-Aguilar, S., Crapez, E., Tanos, R., Cornillot, E., Turtoi, A., and Colinge, J. (2023). Inferring ligand-receptor cellular networks from bulk and spatial transcriptomic datasets with BulkSignalR. *Nucleic Acids Research* 51, 4726–4744. 10.1093/nar/gkad352.
25. Kieffer, Y., Hocine, H.R., Gentric, G., Pelon, F., Bernard, C., Bourachot, B., Lameiras, S., Albergante, L., Bonneau, C., Guyard, A., et al. (2020). Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discovery* 10, 1330–1351. 10.1158/2159-8290.CD-19-1384.
26. Jew B., Alvarez M., Rahmani E., Miao Z., Ko A., Garske K.M., Sul J.H., Pietiläinen K.H., Pajukanta P., Halperin E. (2020). Accurate estimation of cell composition in bulk expression through robust integration of single-cell information. *Nat Commun.* 11 (1): 1971. 10.1038/s41467-020-15816-6.
27. Korotkevich G., Sukhov V., Budin N., Shpak B., Artyomov M.N., Sergushichev A. (2021). Fast gene set enrichment analysis. *BioRxiv*. doi: <https://doi.org/10.1101/060012>

28. Tie Y., Tang F., Peng D., Zhang Y., Shi H. (2022). TGF-beta signal transduction: biology, function and therapy for diseases. *Mol Biomed.* 3 (1): 45. 10.1186/s43556-022-00109-9.
29. Costanza B, Umelo IA, Bellier J, Castronovo V, Turtoi A. (2017). Stromal Modulators of TGF- $\beta$  in Cancer. *J Clin Med.* 6 (1): 7. 10.3390/jcm6010007.
30. Verrecchia, F., Chu, M.-L., and Mauviel, A. (2001). Identification of Novel TGF- $\beta$ /Smad Gene Targets in Dermal Fibroblasts using a Combined cDNA Microarray/Promoter Transactivation Approach. *Journal of Biological Chemistry* 276, 17058–17062. 10.1074/jbc.M100754200.
31. Zacchigna, L., Vecchione, C., Notte, A., Cordenonsi, M., Dupont, S., Maretto, S., Cifelli, G., Ferrari, A., Maffei, A., Fabbro, C., et al. (2006). Emilin1 Links TGF- $\beta$  Maturation to Blood Pressure Homeostasis. *Cell* 124, 929–942. 10.1016/j.cell.2005.12.035.
32. Chiavarina B., Costanza B., Ronca R., Blomme A., Rezzola S., Chiodelli P., Giguely A., Belthier G., Doumont G., Van Simaeys G., et al. (2021). Metastatic colorectal cancer cells maintain the TGF $\beta$  program and use TGFBI to fuel angiogenesis. *Theranostics.* 11 (4): 1626-1640. 10.7150/thno.51507.
33. Nakazawa N., Yokobori T., Kaira K., Turtoi A., Baatar S., Gombodorj N., Handa T., Tsukagoshi M., Ubukata Y., Kimura A., et al. (2020). 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.* 27 (3): 933-942. 10.1245/s10434-019-07878-8.
34. Zacchigna L., Vecchione C., Notte A., Cordenonsi M., Dupont S., Maretto S., Cifelli G., Ferrari A., Maffei A., Fabbro C., et al. (2006). Emilin1 links TGF-beta maturation to blood pressure homeostasis. *Cell.* 124 (5): 929-42. 10.1016/j.cell.2005.12.035.
35. Li F., Li C., Cai X., Xie Z., Zhou L., Cheng B., Zhong R., Xiong S., Li J., Chen Z., et al. (2021). The association between CD8+ tumor-infiltrating lymphocytes and the clinical outcome of cancer immunotherapy: A systematic review and meta-analysis. *EClinicalMedicine.* 41: 101134. 10.1016/j.eclinm.2021.101134.
36. Kist de Ruijter L., van de Donk P.P., Hooiveld-Noeken J.S., Giesen D., Elias S.G., Lub-de Hooge M.N., Oosting S.F., Jalving M., Timens W., Brouwers A.H., et al. (2022). Whole-body CD8+ T cell visualization before and during cancer immunotherapy: a phase 1/2 trial. *Nat Med.* 28 (12): 2601-2610. 10.1038/s41591-022-02084-8.
37. Loi S., Sirtaine N., Piette F., Salgado R., Viale G., Van Eenoo F., Rouas G., Francis P., Crown J.P., Hitre E., et al. (2013). 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. *Clin Oncol.* 31(7):860-7.
38. Mahmoud S.M., Paish E.C., Powe D.G., Macmillan R.D., Grainge M.J., Lee A.H., Ellis I.O., Green A.R. (2011). Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer. *J Clin Oncol.* 29(15):1949-55.

39. Li M.O., Flavell R.A. (2008). TGF-beta: a master of all T cell trades. *Cell.* 134 (3): 392-404. 10.1016/j.cell.2008.07.025.
40. Li M.O., Wan Y.Y., Sanjabi S., Robertson A.K., Flavell R.A. (2006). Transforming growth factor-beta regulation of immune responses. *Annu Rev Immunol.* 24: 99-146. 10.1146/annurev.immunol.24.021605.090737.
41. Gunderson A.J., Yamazaki T., McCarty K., Fox N., Phillips M., Alice A., Blair T., Whiteford M., O'Brien D., Ahmad R., et al. (2020). TGF $\beta$  suppresses CD8 $^{+}$  T cell expression of CXCR3 and tumor trafficking. *Nat Commun.* 11 (1): 1749. 10.1038/s41467-020-15404-8.
42. Metropulos A.E., Munshi H.G., Principe D.R. (2022). The difficulty in translating the preclinical success of combined TGF $\beta$  and immune checkpoint inhibition to clinical trial. *EBioMedicine.* 86: 104380. 10.1016/j.ebiom.2022.104380.
43. Maris P., Blomme A., Palacios A.P., Costanza B., Bellahcène A., Bianchi E., Goflot S., Drion P., Trombino G.E., Di Valentin E., et al. (2015). Asporin Is a Fibroblast-Derived TGF- $\beta$ 1 Inhibitor and a Tumor Suppressor Associated with Good Prognosis in Breast Cancer. *PLoS Med.* 12 (9): e1001871. 10.1371/journal.pmed.1001871.
44. Xie C., Mondal D.K., Ulas M., Neill T., Iozzo R.V. (2022). Oncosuppressive roles of decorin through regulation of multiple receptors and diverse signaling pathways. *Am J Physiol Cell Physiol.* 322 (3): C554-C566. 10.1152/ajpcell.00016.2022.
45. Maia, A., Schöllhorn, A., Schuhmacher, J., & Gouttefangeas, C. (2023). CAF-immune cell crosstalk and its impact in immunotherapy. *Seminars in immunopathology,* 45(2), 203–214. 10.1007/s00281-022-00977-x
46. Öhlund, D., Handly-Santana, A., Biffi, G., Elyada, E., Almeida, A. S., Ponz-Sarvise, M., Corbo, V., Oni, T. E., Hearn, S. A., Lee, E. J., et al. (2017). Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *The Journal of experimental medicine,* 214(3), 579–596. 10.1084/jem.20162024

## FIGURE LEGENDS

**Figure 1:** Breast cancer atlas for cellular and functional annotation of spatial single-cell RNA-seq data. **(a)** UMAP plot showing the breast cancer atlas based on two previously published single-cell RNA-seq datasets.<sup>21,23</sup> **(b)** Verification of the cellular annotation of the newly generated breast cancer atlas using several cell-specific genes. **(c)** Deconvolution of two cell populations in the spatial single-cell RNA-seq data using the breast cancer cell atlas (left panels) and confirmation of the deconvolution by histology analysis (right panel). Different colors denote regions with certain predominant cell types, such as: cancer-associated fibroblasts (CAF), in-situ cancer cells (CC-is), invasive cancer cells (CC-inv) and immune cells (IC). **(d)** Spatial distribution of selected GO processes in BC samples (two representative samples are shown: C1 and 1160920F; other samples are displayed in **Figure S1**). The following GO processes are displayed: ECM Structural Organization, Wound Healing, TGF- $\beta$  Receptor Signaling, Regulation of Immune System, Macrophage Activation, and T-Cell Activation.

**Figure 2:** Spatial analysis of proliferating cancer cells and immune infiltrate in breast cancer samples. **(a)** Histological annotation of two representative breast cancer samples (other samples are shown in **Figure S12**), 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 the correlation analysis for these four populations. **(b)** Correlation analysis for the four selected cell populations in the other eight breast cancer samples.

**Figure 3:** Spatial relationship of TGF- $\beta$  signaling and CAF subpopulations in breast cancer. **(a)** Spatial distribution of genes implicated in TGF- $\beta$  signaling (top) and spatial distribution of different CAF subpopulations in two breast cancer samples. The dendrogram (bottom, right) shows the spatial co-occurrence between CAF subpopulations and TGF- $\beta$  signaling. **(b)** Dendrograms showing the co-occurrence of different CAF populations and TGF- $\beta$  signaling in the other eight breast cancer samples.

**Figure 4:** Differential gene expression analysis of areas with high in TGF- $\beta$  signaling and with/without CD8 $^{+}$  T-cell exclusion. **(a)** Spatial distribution of areas with high *versus* low TGF- $\beta$  signaling and

presence/absence of CD8<sup>+</sup> T cells in two breast cancer samples. (b) Differential gene expression analysis showing the frequency of top-modulated genes overexpressed in the areas where CD8<sup>+</sup> T cells are present despite high TGF- $\beta$  signaling. (c) *EMILIN1* expression in the indicated cell subpopulation (from the breast cancer atlas in **Figure 1a**). (d) *EMILIN1* expression in the indicated CAF subpopulations. (e) Upregulation of TGF- $\beta$  signature genes in the indicated CAF subpopulations. The patient-wise statistical analysis of *EMILIN1* overexpression in CD8<sup>+</sup> cells with high TGF- $\beta$  signaling regions is provided in **Figure S14**.

**Figure 5:** EMILIN-1 is a good prognostic marker in breast cancer. (a) Multiplexed immunofluorescence analysis displaying the localization of CD8<sup>+</sup> T cells (red) and EMILIN-1 expression (blue) in CAFs in a representative breast cancer sample (N=5). Expression of TGFBI, a TGF- $\beta$  signaling activity marker, was in green. (b) Multiplexed immunohistochemistry analysis showing examples of EMILIN-1 (red) and CD8 (brown) co-staining in breast cancer samples (N=75, all subtypes; for multivariate analysis see **Table S2**). (c) Low- and high-power views of the multiplexed immunofluorescence analysis displaying the localization of EMILIN-1, Ki-67 and CD8 in representative breast cancer samples (N=7). (d-e) Violin plots of CD8<sup>+</sup> cell counts and Ki-67 positivity in areas of high *versus* low EMILIN-1 expression in breast cancer samples (N=75). (f) Survival analysis of patients with breast cancer (N=75) in function of EMILIN-1 expression level (high *versus* low) in the tumor. Maybe you should remind what you used as cut-off for high/low levels.

## SUPPLEMENTARY TABLES AND FIGURES

**Table S1:** Clinical and pathological information on the patients whose breast cancer samples (A1, B1, C1 and D1) were used for the spatial analysis.

**Table S2:** Clinical and pathological information on the patients whose breast cancer samples (N=75) were used for the validation analysis.

**Table S3:** Multivariate analysis of EMILIN-1 expression and clinical-pathological variables in the cohort of 75 patients with breast cancer.

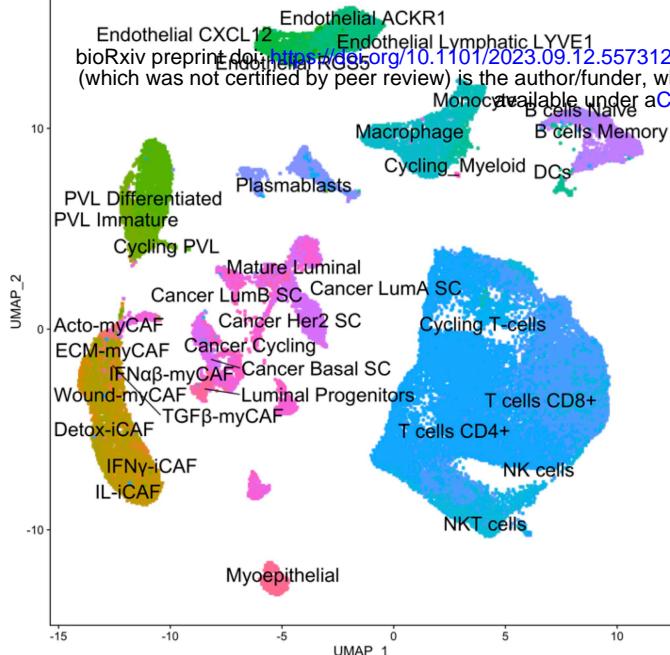
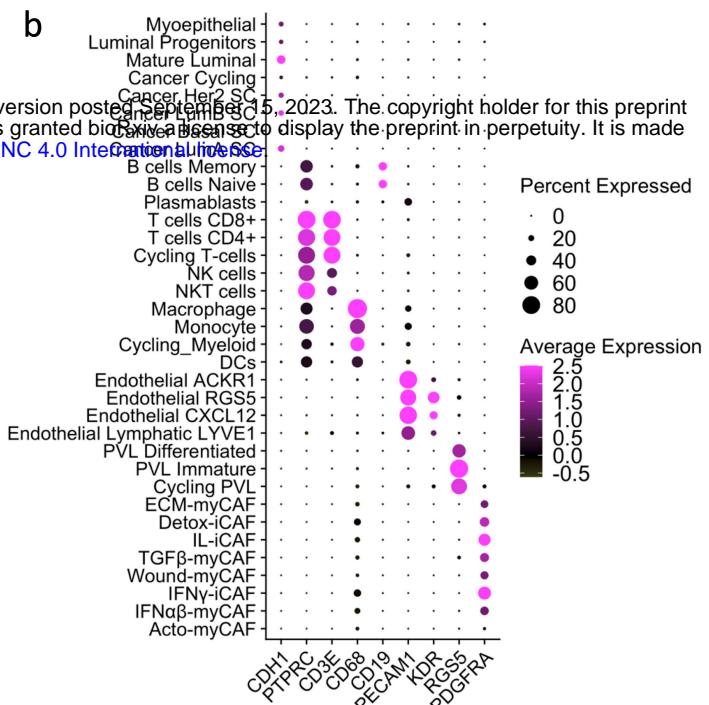
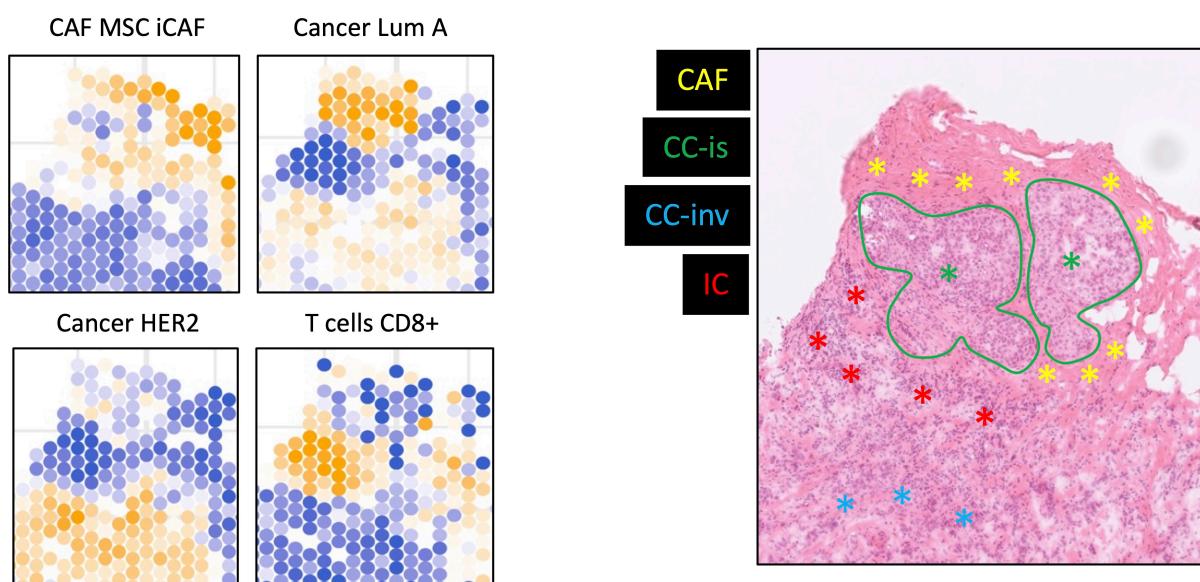
**Figure S1-S10:** Spatial distribution of different cellular populations in breast cancer samples. You need to specify in each figure which breast cancer sample you are showing.

**Figure S11:** Spatial distribution of selected GO processes in the indicated ten breast cancer samples. Shown are the following GO processes: ECM Structural Organization, Wound Healing, TGF- $\beta$  Receptor Signaling, Regulation of Immune System, Macrophage Activation, and T-Cell Activation. In some samples, some GO processes were not significantly enriched and are not shown.

**Figure S12:** Histological annotation of breast cancer samples (Reference tissue) and estimation of highly proliferative regions (S+G2M), actively cycling cancer cells and two immune populations (macrophages and CD8<sup>+</sup> T cells) in the indicated ten breast cancer samples.

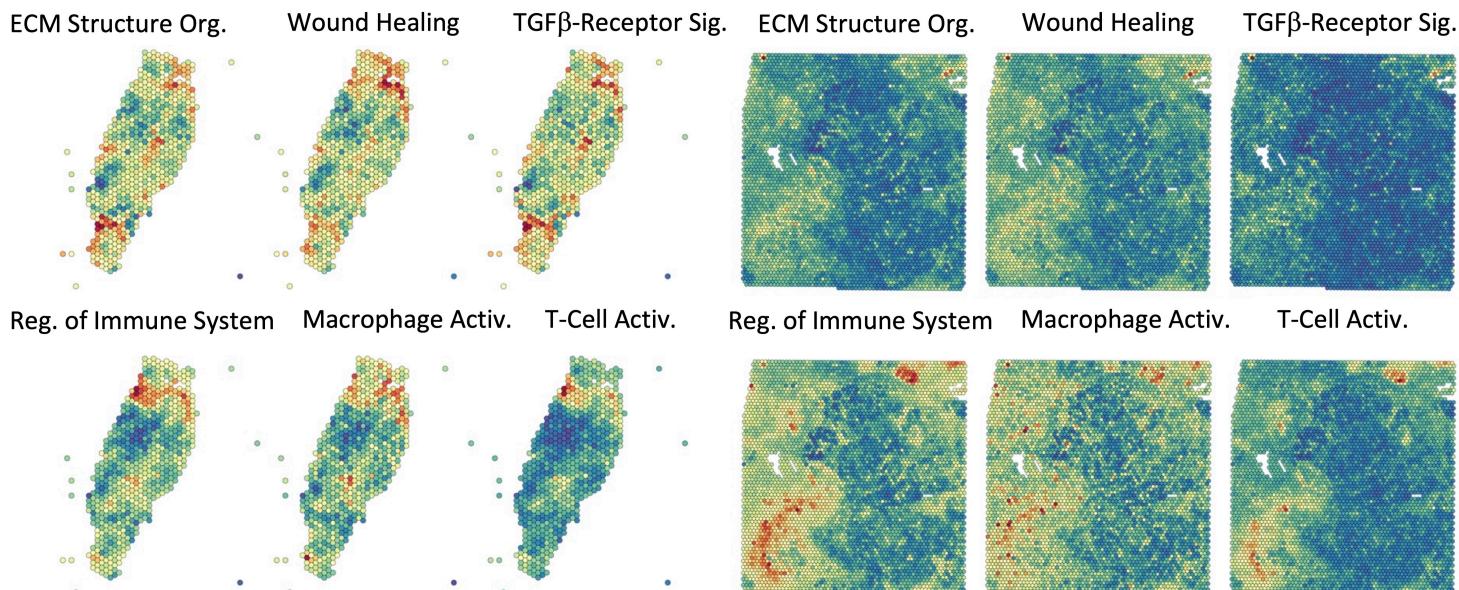
**Figure S13:** Gene ontology analysis of the top 5 significantly enriched biological processes in the indicated CAF subpopulations. No significant enrichment was found for Acto-myCAFs.

**Figure S14:** Box plots showing the patient-wise statistical analysis of *EMILIN1* expression in CD8<sup>+</sup> cells/TGF $\beta$ + regions in the indicated ten breast cancer samples.

**a****b****c****d**

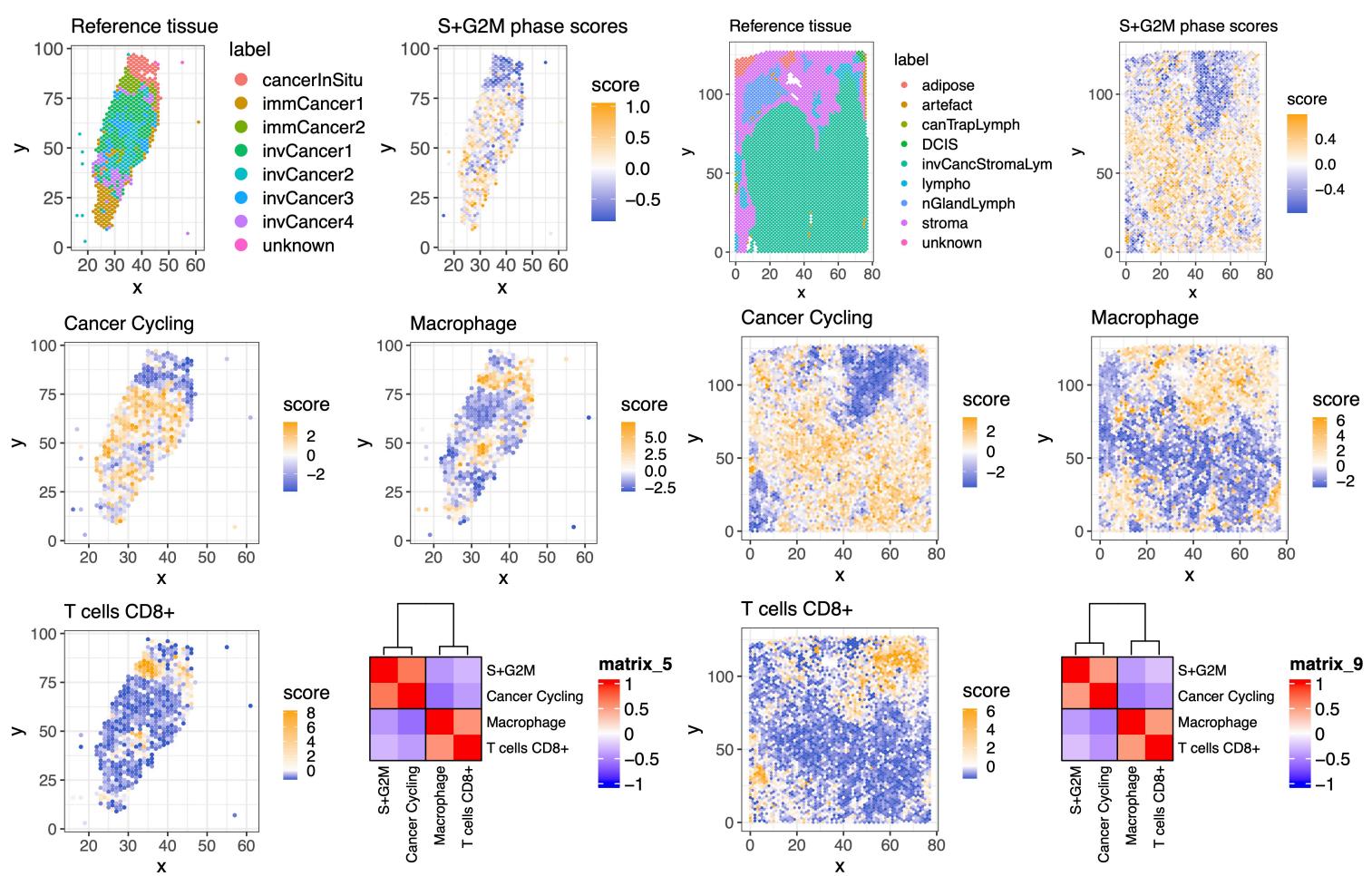
C1

1160920F

**Figure 1**

1160920F

a



b

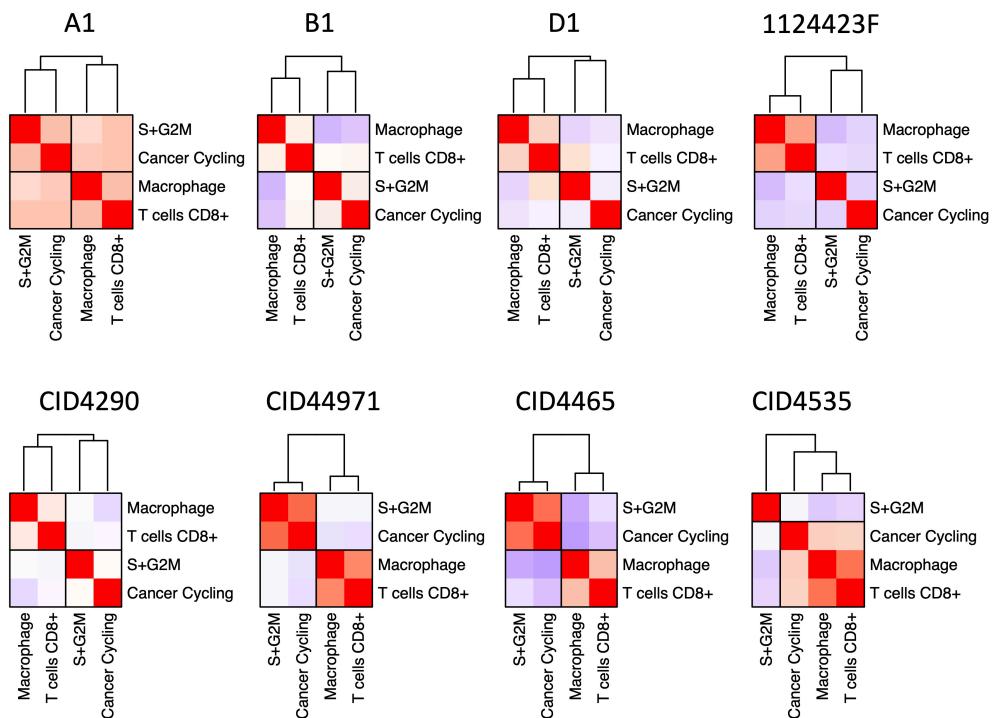
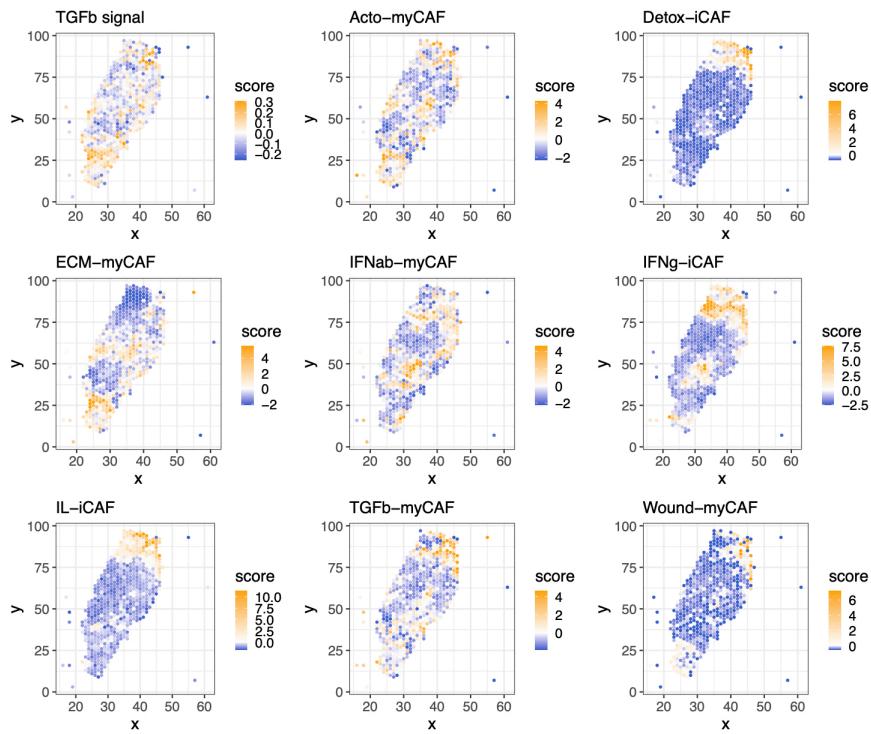


Figure 2

a



1160920F

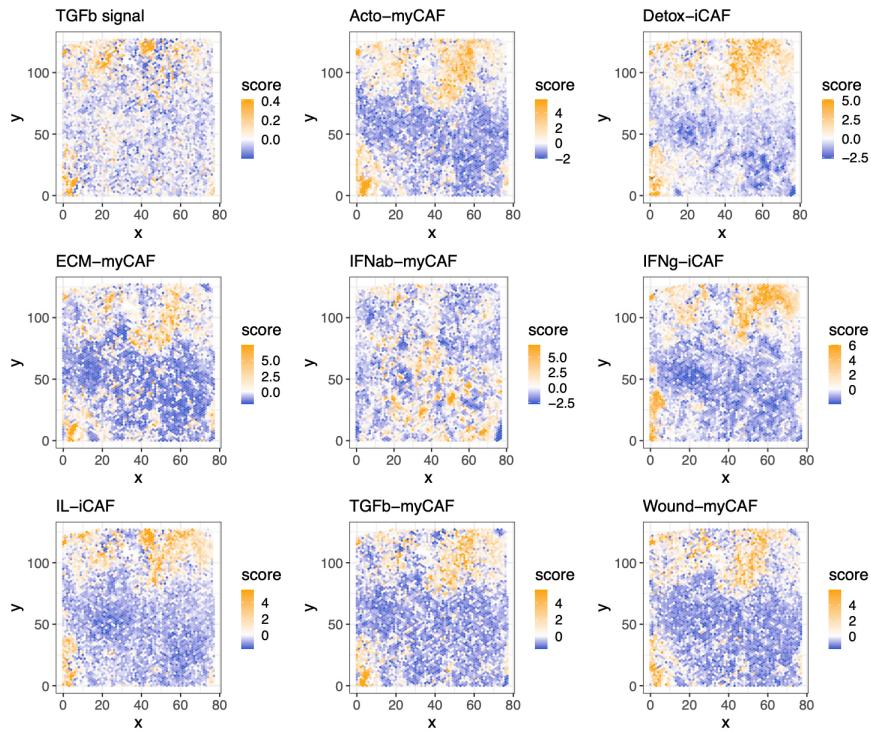
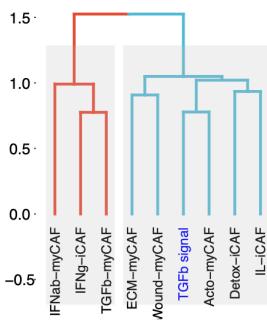
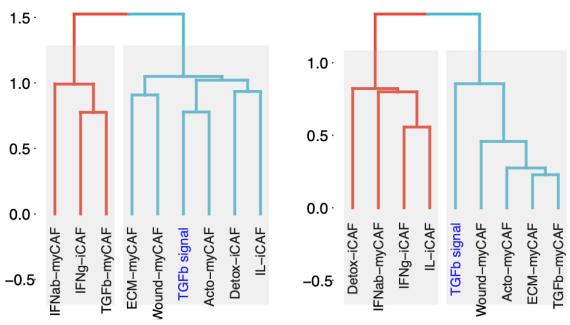


Figure 3

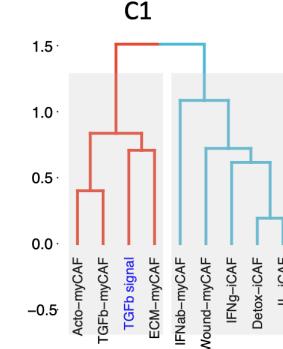
A1



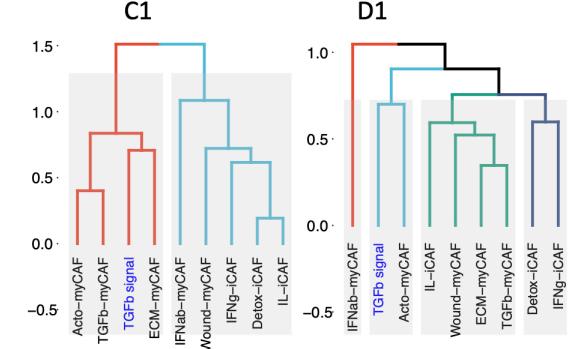
B1



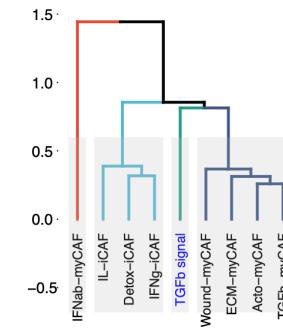
C1



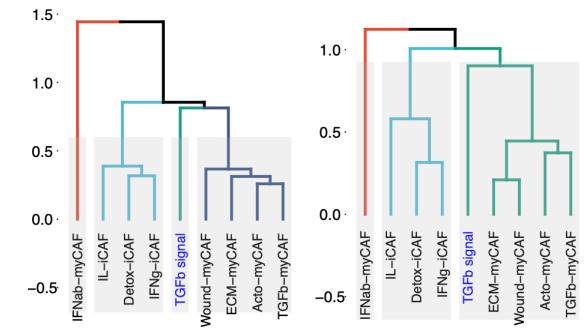
D1



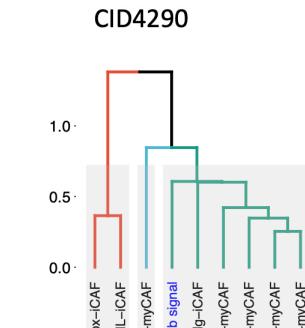
1160920F



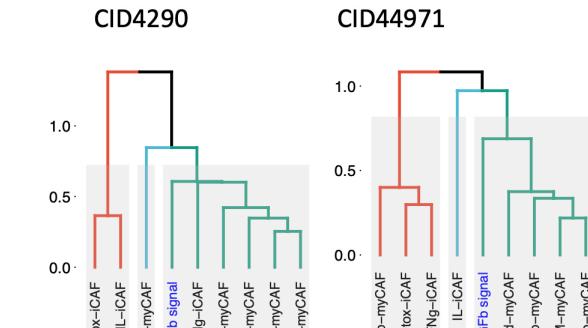
1124423F



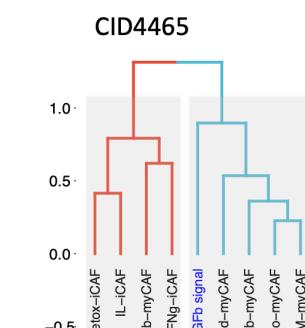
CID4290



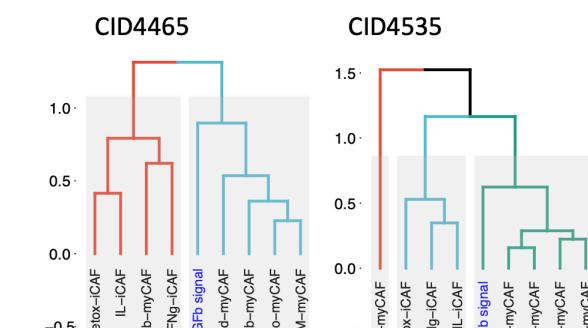
CID44971



CID4465

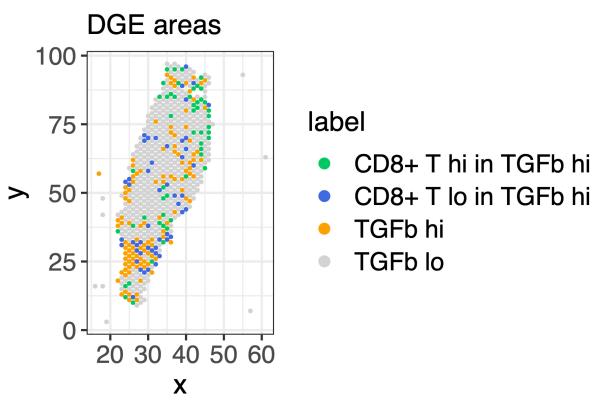


CID4535

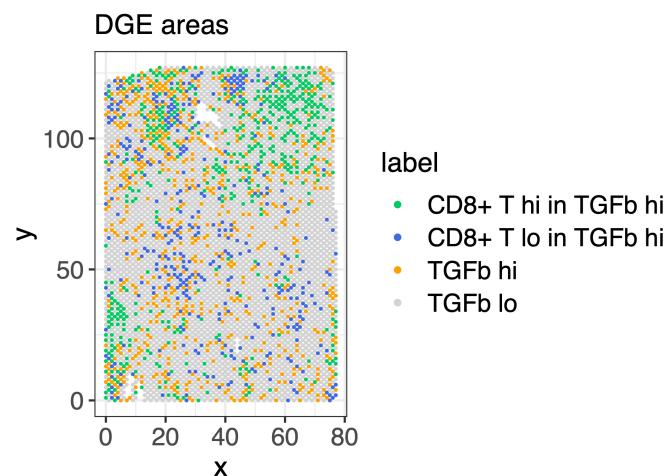


a

C1

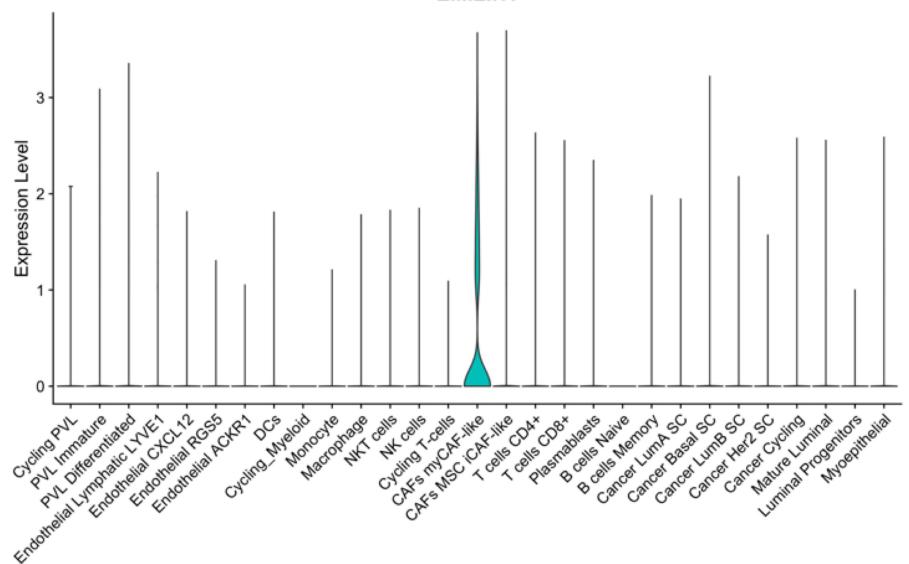


1160920F



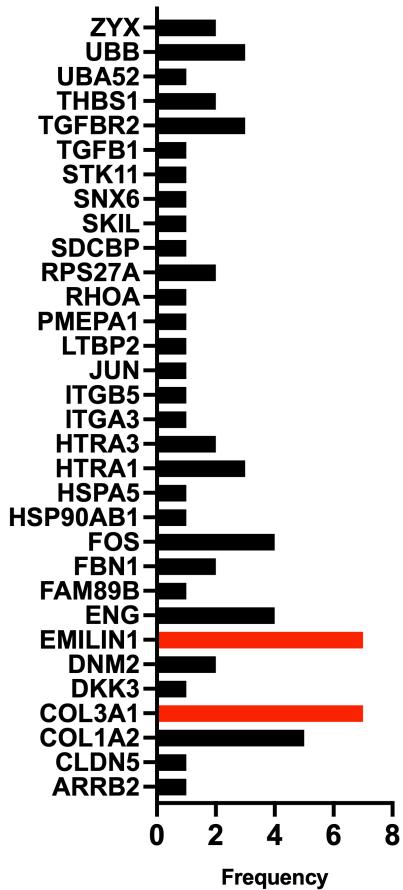
c

EMILIN1



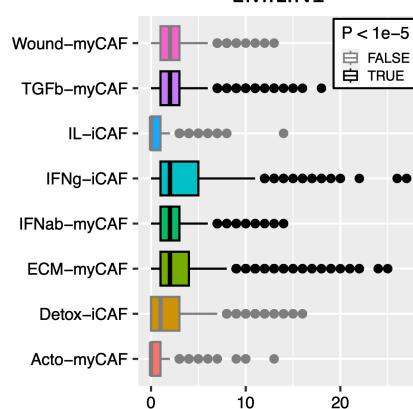
b

### Top Modulated Genes



d

EMILIN1



e

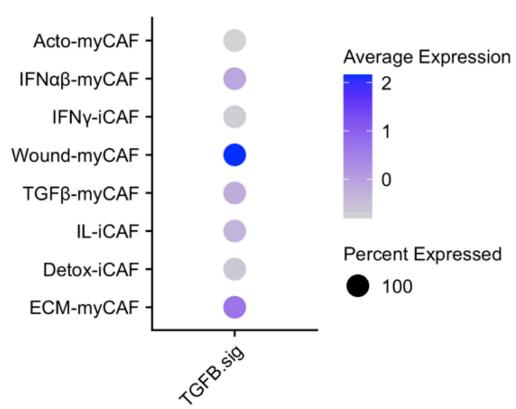
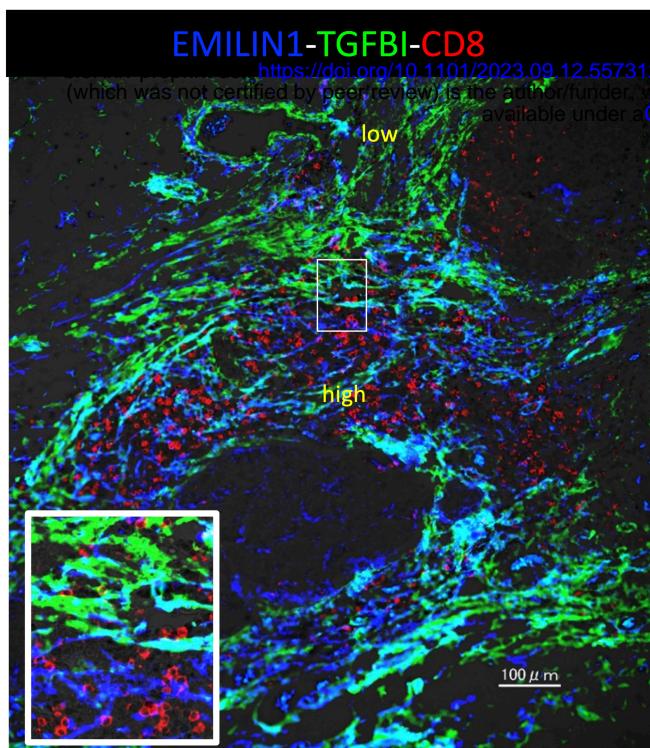
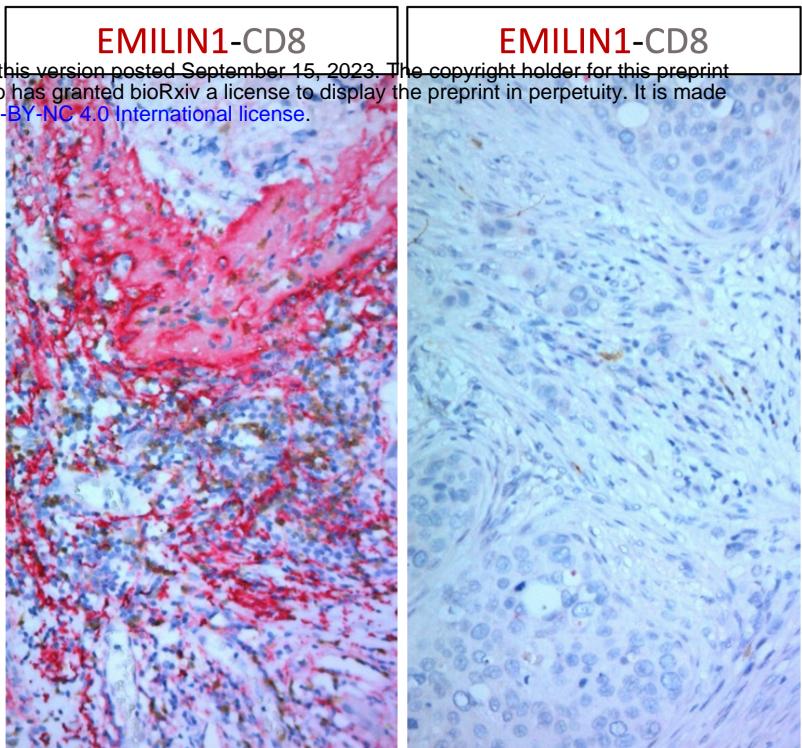


Figure 4

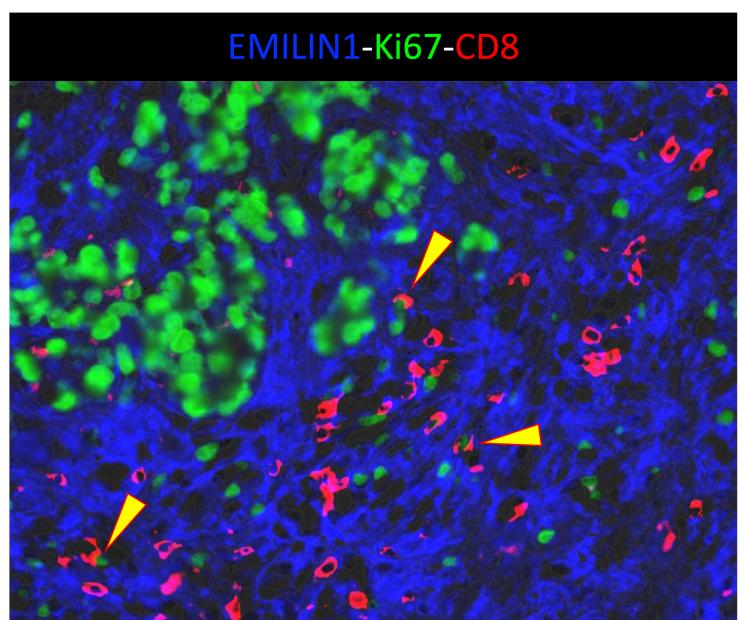
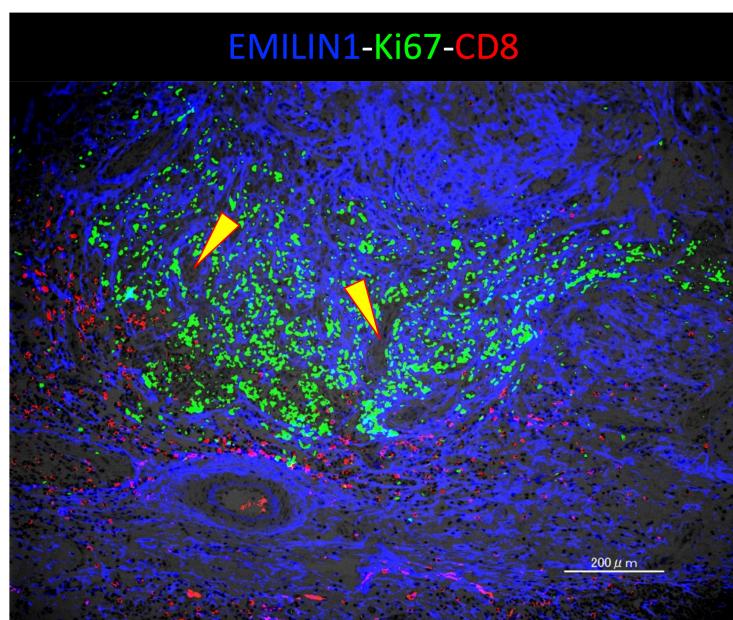
a



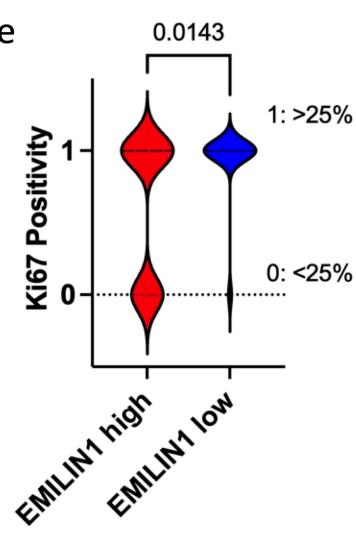
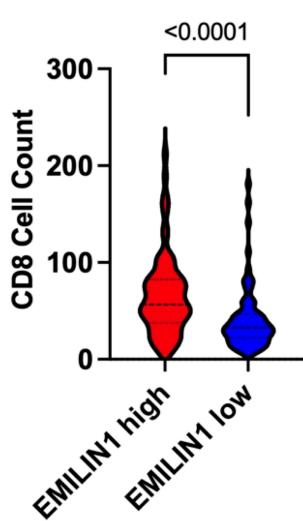
b



c



d



f

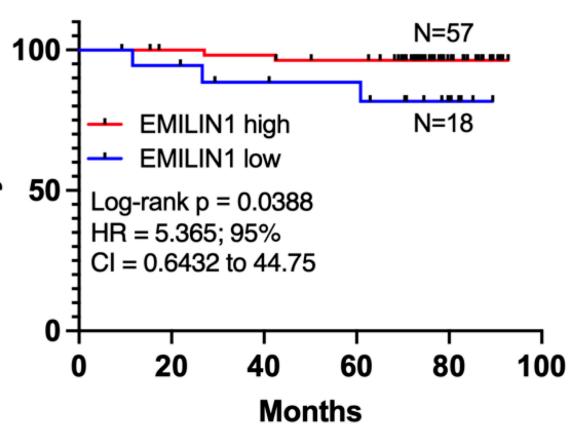


Figure 5