Research Square

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

Spatial localization of RNAs within the breast tumor microenvironment may affect progression and recurrence of DCIS

Kyra Lee
University of Vermont
Nicole A Bouffard
University of Vermont
Mark F Evans
University of Vermont
Agnes Balla
University of Vermont
Donald L Weaver
University of Vermont
Brian L Sprague
University of Vermont
Janet L Stein
University of Vermont
Gary S Stein
University of Vermont
Jonathan AR Gordon
University of Vermont
Prachi N Ghule
prachi.ghule@med.uvm.edu
University of Vermont

Research Article

Keywords: Tumor microenvironment (TME), breast cancer, DCIS, spatial proximity analysis, TMA

Posted Date: February 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8787311/v1

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

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Breast Cancer Research on October 16th, 2026. See the published version at https://doi.org/10.1186/s13058-026-02402-9.

Abstract

Background

Ductal carcinoma in situ (DCIS) is a cancerous growth of breast duct cells that may remain indolent or progress to invasive ductal carcinoma (IDC). As screening rates increase, the prevalence of DCIS diagnosis has risen with many women undergoing aggressive treatment with surgery, radiation, and endocrine therapy when diagnosed with DCIS. It is critical to identify factors that predict progression to invasive disease for improved outcomes. Studies involving the breast tumor microenvironment (TME) provide potential understanding of crosstalk between tumor epithelium and stroma to identify factors that influence progression of DCIS lesions.

Methods

To identify such biomarkers of breast cancer progression, we examined the TME using breast tissue microarrays (TMAs) representing matched benign (normal adjacent), DCIS, and IDC samples. We examined 139 cores which provided data from 47 unique patients. We characterized the expression and spatial distribution patterns of regulatory RNAs (mRNAs and long-noncoding RNAs) including Runt-related transcription factor 1 (RUNX1), Runt-related transcription factor 2 (RUNX2), Mitotically activated long non-coding RNA (MANCR), Cluster of Differentiation 90 (CD90), C-X-C motif chemokine 12 (CXCL12), C-X-C chemokine receptor type 6 (CXCR6), and tumor protein 63 (TP63), using RNAScope fluorescence in situ hybridization (RNA-FISH) and a panel of stromal marker proteins (i.e., Cluster of Differentiation 3 (CD3), Cluster of Differentiation 68 (CD68), Cluster of Differentiation 34 (CD34), Alpha Smooth Muscle Actin (aSMA) and TP63) using multiplex immunofluorescence (IF).

Results

We identified several temporal and spatial expression signatures of RNAs and/or proteins throughout breast cancer progression, both in the epithelial and stromal compartment of benign, DCIS or IDC lesions. Spatial proximity analysis to assess location of markers away from epithelial boundary or within the myoepithelial layer of these lesions identified significant association with clinical parameters including tumor recurrence status.

Conclusions

The finding that decreased myoepithelial continuity maybe protective against recurrence could help better patient stratification for immunotherapy. This work emphasizes the importance of spatial localization of markers in the breast cancer tumor microenvironment and their importance for clinical outcome.

Background

Female breast cancer is the leading cause of cancer death among women younger than 50 years as of 2025 and incidence rates continue to increase at an estimated rate of 1.6% yearly [1]. Additionally, about 59,080 new cases of ductal carcinoma in situ (DCIS), which is a precursor lesion to invasive breast cancer, are expected in the United States this year [2]. The increased incidence is at least in part due to improved breast cancer screening which allows DCIS to be diagnosed earlier. Less than half of DCIS cases managed with excisional biopsy (12–50%) progress to invasive ductal carcinoma (IDC) [3–6]. This indicates that many DCIS cases will not progress to IDC and therefore may not require aggressive treatment. The lack of reliable predictive markers that distinguish between indolence and progression means that many DCIS cases will unnecessarily be treated aggressively with surgery (i.e. complete or partial mastectomy), radiation therapy, and endocrine therapy. This causes morbidity for patients undergoing these treatments when many DCIS lesions may never progress to malignant disease. One study found that aggressive treatments for DCIS were found to be associated with a lower quality of life score for breast cancer and DCIS patients [7]. An example of excessive patient harm is the increased rate of contralateral prophylactic mastectomy (CPM) from 3.9% in 2002 to 12.7% in 2012 [8] despite the absence of a guidelines recommending the surgery in the majority of women diagnosed with unilateral breast cancer. While some institutions have recently shown a decrease in CPM rates with increased physician education and patient counseling [9] there remains a requirement for identification of markers that predict progression to reduce surgical morbidity.

Conversely, aggressive treatment of DCIS can be justified in cases with elevated risk of recurrence of DCIS or IDC after breast conserving surgery. Several methods of predicting breast tumor progression and recurrence risk have been suggested. One of the more successful methods is stratification of DCIS into molecular subtypes based on expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Antigen Kiel 67 (Ki67). These breast tumor categories are defined as luminal A (ER/PR+, HER2-, low Ki67), luminal B (ER+, PR+/-, high Ki67), HER2+, triple negative (ER-/PR-/HER2-), and basal-like (triple negative with high expression of basal epithelial markers). These epithelial tumor molecular subtypes correlate with some recurrence risk; however, these predictors alone are insufficient to adequately satisfy risk of DCIS progression [10, 11]. In addition to epithelial based tumor cells, many other cells (e.g. T cells, myeloid cells, B cells, plasma blasts, endothelial cells and mesenchymal cells (fibroblasts and perivascular-like cells)) are major components of the tumor microenvironment (TME) and provide extensive variability in individual tumors or lesions. The abundance and molecular composition of these diverse cells have been linked to roles in tumor progression and recurrence [12–14]. Adding further complication to this environment, each distinct cell type found in the TME has unique transcriptome and proteome that may convey specialized functions to small populations or individual cells [15] and may be further influenced by spatial relationship (proximity) to malignant or transformed epithelial cells.

Several studies have identified microenvironmental signatures associated with DCIS progression and recurrence of breast cancer that involve specific immune cell population signatures [16] based on cell

populations such as hematopoietic-derived immune cells (e.g. T-cells, B-cells) or myeloid-derived immune cells (e.g. mast cells, neutrophils and macrophages) [17]. In general, T-cells can be identified by the expression of CD3, however, the level of expression can vary depending on the T-cell subset and its activation state [18]. Similarly, markers such as CD68 identify myeloid-derived cells that include monocytes and macrophages although evidence suggests that CD68 RNA may be detected in some fibroblasts, endothelial and tumor cells [19]. Although these cells are minimally represented in the TME, mesenchymal or mesenchymal-derived cells have been implicated in supporting many facets of tumor progression including tumor cell growth, dormancy, migration, invasion, metastasis, and drug resistance [20]. Mesenchymal stromal cells (MSCs) express a wide range of context dependent cell surface antigens; however, CD90 (Thy-1) is a reliable marker to identify MSCs or MSC-derived fibroblasts [21]. As MSCs are multipotent stem-like cells, they have many diverse roles. Perhaps due to the complexity of these functions there are many distinct populations of tumor-resident MSCs that are characterized mainly by unique cell surface markers or RNA expression patterns [22, 23]. Within tumor-resident MSC or MSC-derived cells, the expression levels of specific transcriptional regulators related to epithelial to mesenchymal transition (EMT) (e.g. RUNX1) or metastatic potential (e.g. RUNX2), cytokine signaling mediators (e.g. CXCL12, CXCR6) or specific long non-coding RNAs (lncRNAs) (e.g. HHLA3, TP53TG1, MIR22HG) have been proposed to be potential prognostic biomarkers or associated with functions promoting oncogenic progression in a variety of human tumors. [24–28].

In this study we addressed the compelling requirement to identify how the spatial distribution of both phenotypic protein and RNA transcripts within the tumor microenvironment relate to DCIS and IDC progression, and recurrence. Our study aimed to identify the importance of spatial distribution of cells and their association with clinical parameters within tumor microenvironment. We identified cells expressing CXC-cytokine signaling, cancer associated fibroblasts (aSMA), myoepithelial cells (TP63 and aSMA), immune infiltrates (CD68, CD3, CD34), stromal cells expressing RUNX1/RUNX2, and MANCR expressing cells. The lncRNA MANCR is upregulated in metastatic breast cancer cells and found to have a role in stabilizing the genome of cancer cells [29]. In gastric cancer, upregulation of MANCR is associated with poor survival [30]. Additionally, higher expression of MANCR is associated with aggressive clinical parameters in breast cancer [31]. The RUNX family of transcription factors are known to play an important role in breast cancer development. RUNX1 is typically downregulated in breast cancer compared to normal mammary tissue, and RUNX1 mutations were identified as driver mutations in breast cancer, which suggests loss of RUNX1 promotes breast cancer progression [32–34]. RUNX2 has the opposite effect, functioning as an oncogene while RUNX1 functions as a tumor suppressor [35]. RUNX2 was also found to be upregulated in a mouse xenograft model of tumor progression and played a role in promoting the development and metastasis of breast cancer by regulating the proportion of breast cancer stem cells (BCSCs) [36]. In mantle cell lymphoma (MCL), MANCR and RUNX2 were both upregulated and MANCR overexpression was found to promote RUNX2 expression in MCL cells [37]. The potential interaction between MANCR and the RUNX family of transcription factors is not well understood in breast cancer. We assessed expression and spatial distribution of MANCR, RUNX1 and RUNX2 expressing cells with clinical outcome to evaluate role of these genes as biomarkers.

We developed three specific RNA in situ hybridization panels to identify key components of the TME and defined spatial relationship between these features in breast cancer patient samples. The three RNA panels consisted of probes for transcripts associated with CXC-cytokine signaling pathways and mesenchymal markers (e.g. CD90, CXCL12, and CXCR6); lncRNA and basal-like myoepithelial cells (e.g. MANCR, TP63); and transcriptional regulation (e.g. RUNX1 and RUNX2). RNA expression was evaluated in conjunction with multiplex IF staining for cell surface markers (e.g. CD3, CD68, CD34, aSMA, TP63). This multi-modal strategy was used to evaluate benign (normal adjacent) ducts, DCIS, invasive ductal carcinoma (IDC) and associated stroma areas from 47 individual patients using a custom generated tissue microarray (TMA) containing multiple cores per individual sample. In total 139 individual cores were evaluated for marker expression and spatial interactions, and we identified several significantly associated expression patterns and spatial relationship signatures that correspond to DCIS subtypes, tumor progression, recurrence and other clinical parameters.

Methods

Table 1
Clinical parameters of breast cancer cases (N = 47 patients)

Clinical Parameter	Number of Patients	% of patients
Grade		
Low	2	4.26%
Intermediate	28	59.57%
High	12	25.53%
Data Not Available	5	10.64%
HER2 Status		
Positive	19	40.43%
Negative	25	53.19%
Data Not Available	3	6.38%
Hormone Receptor Status		
Positive	40	85.11%
Negative	4	8.51%
Data Not Available	3	6.38%
Differentiation Status		
Well	16	34.04%
Moderate	18	38.30%
Poor	12	25.53%
Data Not Available	1	2.13%
Lymphovascular Invasion (LVI)		
Present	14	29.79%
Absent	32	68.09%
Data Not Available	1	2.13%
Invasive Tumor Classification		
Invasive Ductal	35	74.47%
Invasive Ductal with Tubular	1	2.13%
Invasive Ductal with Lobular	5	10.64%
Other	5	10.64%

Clinical Parameter	Number of Patients	% of patients
Grade		
Data Not Available	1	2.13%
DCIS Type		
Cribriform Only	5	10.64%
Solid Only	11	23.40%
Mixed	30	63.83%

Breast tissue Microarray Generation

To enhance the efficiency of histopathological simultaneous analysis of multiple patient-derived tissue specimens under uniform experimental conditions, five tissue microarrays (TMAs) were created using breast cancer cases available through the Vermont Breast Cancer Surveillance System (VBCSS)[38]. Patients were selected as described by Evans et. al. [39] and included cases with DCIS and invasive breast cancer diagnosed. Histological sections from selected patient cases were subjected to pathology assessment and regions of benign (normal adjacent), DCIS, and IDC were identified and 1.5 mm (circumference) cores were punched from paraffin-embedded (formalin-fixed) tissue samples. The extracted cores were embedded into a single paraffin block in a 10 x 3 grid arrangement. The generated TMA represented ten patients, with three core punches with pathologist-assessed regions (i.e. normal/benign, DCIS, IDC/IBC) from each patient. 5µm serial sections of the TMAs were mounted on glass slides, deparaffinized via graded ethanol series, and first section in each of the series was stained with hematoxylin and eosin (H&E) (**Supplementary Fig. 1**). H&E stained slides were imaged on the Leica-Aperio Versa whole slide imager, digital scans were then annotated in Aperio ImageScope (Leica Biosystems) and reassessed by pathologist (DLW or AB) to identify and distinguish benign, DCIS, and invasive areas for each core. This study was approved by the University of Vermont Committee on Human Research in the Medical Sciences (IRB protocol 15–629), with a waiver of consent for the use of existing clinical specimens and data in this study. Clinical parameters are summarized in Table 1.

RNAScope Assay: RNA FISH was performed on serially sectioned TMA slides using ACDbio's RNAScope Multiplex Fluorescent Assay kit (Cat. 322800, ACDBio, Biotechne, Newark, CA, USA) on the Leica Biosystems' BOND RXm Research Advanced Staining System as per manufacturer's protocols. Briefly, pre-baked slides were dewaxed on the BOND system and pretreated with ER2 for 15 minutes at 95°C. They were then treated with RNAscope® 2.5 LS Protease III from ACD for 15 minutes and RNAscope® 2.5 LS Hydrogen Peroxide for 10 minutes, both at 44°C. TMAs were hybridized with the RNAscope® 2.5 LS Multiplex Positive Control Probe- Hs, RNAscope® 2.5 LS Multiplex Negative Control Probe (Cat. 321838) or a target probe cocktail for 2 hours at 42°C. RNAscope® 2.5 LS Multiplex Positive Control Probe- Hs (Cat. 321808) targets POLR2A (C1 channel), PPIB (C2 channel), UBC (C3 channel), and HPRT-1 (C4 channel). The RNAscope® 2.5 LS Multiplex Negative Control Probe (Cat. 321838) targets DapB

(*Bacillus subtilis* strain) in all four channels. There were three probe cocktails, the first consisted of C1-MANCR (Cat. 411088-C1), C2-RUNX2 (Cat. 440078-C2), and C3-RUNX1 (Cat. 419908-C3). The second consisted of C1-MANCR (Cat. 411088-C1), C2 – TP63 (Cat. 601898-C2), and C3 – CXCL12 (Cat. 422998-C3). The third pool consisted of C1- CD90 (Cat. 430618), C3 – CXCL12 (Cat. 422998-C3), and C4-CXCR6 (Cat. 468468-C4). After the 2-hour probe hybridization, target signal was amplified through three 30-minute incubations at 42°C with RNAscope® LS Multiplex AMP 1, RNAscope® LS Multiplex AMP 2, and RNAscope® LS Multiplex AMP 3. Fluorescent signal was developed for each channel using RNAscope® LS Multiplex HRP and Opal™ fluorophores from PerkinElmer diluted 1:1000 in TSA Buffer. Finally, DAPI was used to stain the nuclei. The probe-opal dye combinations are described in **Supplementary Table 1**.

Multiplex Immunohistochemistry (IHC) Staining

The tissue microarrays were additionally stained with a custom antibody panel (**Supplementary Table 2**) using an Akoya Opal 7-Color Automation IHC kit (NEL821001KT; Akoya Biosciences, Marlborough MA) on a Leica BOND RXm autostaining system (Leica Biosystems, Buffalo Grove, IL) software version 6. Immunostaining protocol steps were established and developed according to the Perkin Elmer user manual titled, “OPAL 4-Color and 7-Color Automation IHC Kits for Leica Biosystems BOND RX System Software version 4.0”.

Image Capture

Stained TMAs were imaged on a Nikon A1R-ER point scanning confocal microscope. A JOB was written using Nikon Elements software to take an overview scan of the whole slide upon which regions of interest could be drawn to be imaged at higher magnification. An overview scan was acquired using the 4x Plan Apo λ (NA 0.20) objective at 256 x 256-pixel resolution with 405 nm (DAPI) excitation. Regions of interest were defined on the low-resolution scan to be imaged at 40x (CFI Plan Apochromat Lambda 40x; NA 0.95). The final images were captured using the high-resolution galvanometer scanner and spectral detector equipped with 405 nm, 445 nm, 488 nm, 514 nm, 561 nm, and 640 nm laser lines and acquired in 10nm bandwidth passes [40].

Spectral Unmixing

A spectral library for the Opal™ fluorophores was created by performing Single Plex RNAscope® with the RNAscope® 2.5 LS Multiplex Positive Control Probe-Hs on formalin fixed paraffin embedded (FFPE) cells pellets made of MDA-MB-231 cells [41]. C2-PPIB was developed with each Opal™ fluorophore (520, 540, 570, 620, 650, and 690). A DAPI library slide and a background autofluorescence library slide were also created. The spectral images from the TMA JOB were then unmixed using the Nikon Elements Spectral Unmixing module. A similar spectral library was also created for protein marker panel using aSMA stained breast sections with each opal dye as described by Taatjes et al [31].

Image Analysis

The unmixed spectral images were imported into HALO (v3.5.3577.140) (Indica Labs, Albuquerque, New Mexico USA) and annotated according to the pathologist's annotations from the H&E sections. The annotations made on the fluorescent images reflected either benign, DCIS, or invasive breast cancer regions and their respective stroma types as well as adipose tissue and blood vessels. Any folded or out of focus regions were annotated and excluded from final analysis. The FISH module (v3.1.3) from Indica Labs HALO was used to segment and analyze cells for the RNAScope panels and cells were classified as low expressors if they had less than 9 detectable copies of a given RNA per cell, and high expressors if they had more than 9 detectable copies of a given RNA per cell. The Highplex FL module (v3.2.1) was used to segment and analyze cells for the multiplex antibody panel. The number of cells positive for each type of RNA or protein was measured both on a whole core basis, and within each distinct annotation type. The area in which the number of positive cells was measured was referred to as an "analysis region". The area of each "analysis region" was also measured. Expression data was compared between analysis regions for each marker to determine how the epithelial tissue and the surrounding microenvironment changed in progression from benign, to DCIS, to IDC. The infiltration analysis algorithm from the spatial analysis module (v3.5) was used to generate spatial plots and measure stromal infiltration of the RNA or protein marker signals Away from epithelial boundary. The nearest neighbor algorithm from the spatial analysis module that measures the average distance and number of unique neighbors between any two cells or objects was used to measure the average distance between adjacent TP63 RNA positive cells or TP63/aSMA protein co-positive cells to measure myoepithelial continuity within different epithelial regions.

Statistical Analysis

The object data generated by the FISH analysis, Highplex FL analysis and spatial analysis in HALO (Indica Labs.) was exported and statistical analysis was performed using R (version 4.3.0). First, cell size outliers were removed, then single positive and co-expression phenotypes were calculated on a cell-by-cell basis. The percent of positive cells for each marker or combination were calculated by dividing the number of cells expressing a given marker or combo by the total number of cells. The spatial analysis data was also tidied using custom R scripts and the distance measurements were averaged by patient and type of epithelium (Invasive or DCIS or Benign/Normal). Both the percentage of positive cells and spatial data were tested for association with clinical parameters using the Kruskal-Wallis test. The clinical parameters tested included the type of DCIS, differentiation status, ER status, PR status, HER2 status, nuclear grade, invasive tumor type, and presence of lymphovascular invasion (LVI) (Table 1). The resulting associations were plotted as dot matrices using the ggplot2 package. Associations identified as significant using the Kruskal-Wallis test were also graphed as boxplots using the ggpubr package and further tested for significance using the Wilcoxcon rank sum test. Significant p values were automatically assigned using the stat_compare_means function from ggpubr, with p </= 0.05 being considered significant. All p values reported in the text and figures are adjusted p values. The myoepithelial continuity score was calculated according to the methods described by Borowsky et al [42]. The overall experimental schematic is shown in Fig. 1.

RESULTS

Breast tumors are comprised of epithelial-derived tumor cells, together with a diverse and complex network of cells including immune cells, fibroblasts, and an extracellular matrix residing in the TME. Each cell has a distinct transcriptome (i.e. RNA), proteome and spatially integrated relationship to other cells. To define relational marker signatures between cells, we created a combination of three RNAScope® panels and a multiplex IHC panel to examine serial sections of breast cancer progression TMAs that included areas of DCIS and IDC. Representative individual channel signal images for each of the RNAs or proteins analyzed within benign, DCIS or invasive cores that include the three RNAscope panels (MANCR/RUNX1/RUNX2, CD90/CXCL12/CXCR6, and MANCR/TP63/CXCL12) as well as multiplex IHC protein panel are shown in **Supplementary Figs. 2–5**. The mean positive percent and standard deviations for each RNA marker in each analysis region is reported in **Supplementary Table 3.** The mean positive percentage and standard deviations for each protein marker are reported in **Supplementary Table 4**.

RUNX transcripts exhibit differential expression across breast tumor progression and stroma

RUNX1 and RUNX2 both showed differential expression patterns across both the epithelium and stromal compartments of patient samples. In addition, as lesion type progressed from benign to DCIS to invasive cancer there was loss of RUNX1/RUNX2 expression in the epithelium of invasive lesions (Fig. 2A-H). Conversely there was a measurable increase in RUNX1/RUNX2 abundance in stroma of invasive lesions. The average percentage of RUNX1 positive cells in benign lesions was 60.4% +/-27.6%, compared to 57.3 +/- 29.2% in DCIS ($p = 0.68$) and 51.6+/-24.9% in invasive lesions ($p = 0.1$) compared to benign epithelium). In stroma, the percentage of RUNX1 positive cells was lowest in benign stroma with an average of 18.0+/-14.6% and increased to an average of 24.3+/- 20.2% in DCIS stroma ($p = 0.29$), with the highest average of 35.3 +/- 16.7% in invasive stroma ($p < 0.0001$ compared to benign stroma). Although overall expression levels of RUNX2 were lower compared to RUNX1, RUNX2 showed similar patterns by lesion type, with an average of 15.3 +/- 15.0% in benign epithelium, an average of 13.9 +/- 13.3% in DCIS epithelium ($p = 0.09$), and an average of 12.1+/- 10.6% in invasive epithelium ($p = 0.8$ from benign). RUNX2 was lowest overall in benign stroma with an average of 9.3 +/- 8.7% cells, increasing to 14.5 +/- 11.2% in DCIS stroma ($p = 0.44$) and further increasing to 17.2 +/- 11.2% in invasive stroma ($p = 0.01$ as compared to benign stroma).

As the lesion type progressed towards IDC, they lost TP63 RNA expression, going from a high of 36.2 +/- 22.0% in benign epithelium to an average of 14.2+/-11.1% in DCIS ($p < 0.0001$) to 4.9 +/- 3.7% in invasive lesions ($p < 0.0001$ as compared to benign). TP63 protein expression showed similar loss of expression from benign to invasive epithelium (Fig. 2I-P). The long-noncoding RNA, MANCR, showed no significant changes in expression pattern during breast tumor progression, with similar expression levels across both the epithelium and stroma of all lesion types (Fig. 2Q-T**)**. There was no significant difference in

staining pattern between the MANCR RNAScope stains performed on serial sections (**Supplementary Fig. 6)**.

Mesenchymal stromal cell marker CD90 is elevated in IDC stroma

MSCs and MSC-derived fibroblast express a variety of cell surface markers however, CD90 (Thy-1) is a reliable marker of mesenchymal cells normally found in adipose tissue such as breast [43]. We measured CD90 RNA transcript levels in patient samples. The number of CD90 positive cells were significantly increased across IDC stroma regions (24.9 +/- 16.6%) compared to benign (12.6+/- 10.1%) ($p = 0.002$) and DCIS stromal regions (14.4 +/- 15.4%) ($p = 0.01$) (Fig. 3A-D). This overall 10.4% net increase observed in the number of CD90 positive cells in stroma regions adjacent to IDC (versus DCIS) strongly suggests that MSC or MSC-derived populations are elevated in response to progression from DCIS to IDC. In single cell studies of tumor-derived CD90 + MSCs, several genes associated with CXC chemokine signaling pathway were found to be upregulated in specific, rare populations of patient-derived mesenchymal stromal cells [23]. To evaluate if these rare MSC were observable by FISH/IHC in patient samples, TMA sections were labelled with CXCR6 and CXCL12 probes. Although the overall level of CXCR6 signal was low across all samples, there was a consistently detectable number of CXCR6 positive in stroma regions. Stroma regions adjacent to IDC had the highest number of CXCR6 positive cells (2.0 +/- 2.6%) compared to DCIS stroma (1.1+/- 1.5%) ($p = 0.2$) or benign stroma (1.7 +/- 2.3%) ($p = 0.4$) **(**Fig. 3E-H). In contrast to CXCR6 which encodes a membrane bound G-protein receptor, the CXCL12 gene encodes a secreted chemokine: stromal cell-derived factor 1 (SDF-1) that binds to the CXCR4 receptor to promote cell migration and immune response. Although our previous studies identified the concomitant upregulation of CXCR6 and CXCL12 in patient-derived MSCs [23], we observed that the number of CXCL12 positive cells were significantly decreased ($p < 0.0001$ for both staining iterations) in stroma associated with IDC (36.6 +/- 17.2%) compared to benign stroma (62.9 +/- 18.1%) (Fig. 3I-L). Reproducibility and robustness of the CXCL12 staining patterns were assessed by staining serial sections and showed no significantly different staining across samples (**Supplementary Fig. 7)**. This finding suggests that although CXCL12 expression is present in all – benign-, DCIS- or IDC-associated stroma, it is reduced as a function of tumor progression and may be independent of the gain of CXCR6-mediated expression and/or signaling within MSCs.

T-lymphocyte-related CD3 expression in DCIS. Several studies have demonstrated a link between intratumoral and stroma lymphocyte infiltration as a prognostic delineator of indolent DCIS, DCIS associated with IBC/IDC and invasive carcinoma [44]. To identify T-lymphocytes, TMAs were subjected to IHC with the pan T-cell marker: CD3. Although not significantly different, stroma associated with DCIS had an average increase in CD3 positive cells (9.0 +/- 5.6%) compared to benign stroma (5.1 +/- 4.2% cells) ($p = 0.2$) or IDC-associated stroma ($p = 0.2$) (6.3+/- 5.6%) (Fig. 4A-D). Interestingly, only in DCIS stroma, significantly higher CD3 positive cells were present compared to epithelial regions (Fig. 4A). To gauge potential macrophage involvement/infiltration TMAs were labelled with CD68, a cell surface glycoprotein normally associated with macrophages. Although there was detectable CD68 signal in all

matched tissue (benign lesion/stroma, DCIS lesion/stroma, IDC lesion/stroma), there was no significant variation in CD68 positive macrophages in patient samples (Fig. 4E-H). Immune invasion in tumors is frequently correlated with an increase in the number of cancer-associated fibroblasts (CAFs) in the tumor stroma. Studies have shown that invasive lobular carcinoma (ILC) progression is accompanied by CD34 positive cells surrounding the normal breast tissue differentiating or transforming to aSMA positive CAFs as the lesions become more malignant [45] and may potentially occur in DCIS to IDC progression. To identify CAFs in our patient cohort, aSMA and CD34 were labelled by IHC **(**Fig. 4I-P). In all samples there was a large number of aSMA positive cells which included benign lesion (50.8 +/- 20.5%), DCIS lesion (65.9+/- 26.1%) and IDC lesion (71.6 +/- 25.1%) areas due to aSMA being expressed in myoepithelial cells **(**Fig. 4M**).** Our pathological assessment of stroma regions excludes epithelium and therefore aSMA signal in stomal regions was not contributed from myoepithelial, epithelial or cells related to blood vessels (e.g. endothelial, smooth muscle, pericyte cells) as these were masked from the analysis. In stroma there was a significant difference (p = 0.04) in the number of aSMA positive cells in benign stroma (32.9+/- 20.5%) compared to IDC stroma (64.3+/-22.8%). It was also not significantly (p = 0.1) increased in DCIS stroma (54.2+/- 26.1%). In contrast to aSMA, CD34 average positive cell percents were highest in benign stroma (23.7 +/- 16.6%) compared to DCIS (15.3 +/- 14.3%) stroma (p = 0.6) or IDC stroma (6.7 +/- 13.7%)(p = 0.04). This represented a significant (p = 0.04) decrease in CD34 expression in benign stroma compared to IDC stroma.

Integrative spatial analysis of marker expression and clinical parameters

Using our panel of RNA and protein markers we observed clear differences in the cellular makeup of stroma of benign, DCIS and IDC lesions. However, when comparing these features to identified clinical pathological grading and demographic information (Table 1**)** from each individual patient there was no significant correlation between the percent positive cell values for individual markers and clinical parameters. Spatial proximity analysis was performed for cells expressing individual RNA or protein markers within 20um or 40um of each type of epithelium using infiltration analysis. There was very little difference between clinical associations observed between 20um or 40um findings (data not shown). We chose to assess 20um distance for further analysis as it represented approximately one cell height. The associations are summarized in a dot matrix plot indicating positive association (red) or negative association (blue) with pathological or clinical features **(**Fig. 5**)**. Spatial proximity patterns showed several associations with clinical parameters that include hormone status, DCIS type and disease aggressiveness, as well as presence of recurrence. Clinical associations were found with all types of lesions (benign, DCIS and invasive) and it was observed that the location of cellular RNA expression, defined by the number of copies of RNA present within 20um of the epithelium, played a role in association with clinical factors.

CXCR6 or MANCR spatial proximity to epithelium is significantly associated with DCIS subtypes

DCIS architectural subtypes have been associated with disease recurrence, although it is still not completely evident as most DCIS lesions exhibit mixed architectural patterns. However, there is data to suggest the solid and micropapillary DCIS subtypes are more often associated with recurrence than the cribriform subtype [46]. In this study, we tested spatial expression patterns of the protein and RNA markers within the different subtypes (Table 1) to determine if correlation between spatial proximity to types of epithelium and DCIS subtypes contributes to more aggressive forms of breast cancer.

Patients with a mixed DCIS subtype had a significantly lower percentage of CXCR6 expressing cells within 20um of benign lesions with an average percent of 24.2 +/- 22.0% compared to averages of 75.0 +/- 35.4% in cribriform lesions and 82.1 +/- 16.9% ($p = 0.01$) in solid lesions (Fig. 6A). Although non-significantly associated, an opposite trend was observed near invasive lesions with mixed lesion having an average of 47.5 +/- 13.6% CXCR6 positive cells, cribriform lesions having an average 21.9 +/- 20.4% CXCR6 positive cells, and solid lesions having an average of 40.0 +/- 6.3% CXCR6 positive cells (data not shown). Although the patterns of CXCR6 near benign or invasive regions varied in patients with mixed DCIS lesions, they were consistently higher in patients with solid lesions compared to cribriform lesions. Other studies have shown increased CXCR6 expression in both invasive breast cancer cell lines [47] and breast cancer tissues [48] suggesting a role for CXCR6 in increasing cell migration, invasion, and metastasis. Our findings reveal presence of CXCR6 expressing cells residing closer to more aggressive DCIS lesions, potentially suggesting a role in promoting invasion.

We observed that spatial proximity of MANCR positive cells from the DCIS epithelial boundary is significantly associated with solid vs mixed DCIS subtypes. MANCR expression was highest near solid DCIS lesions at an average of 87.9 +/- 18.5% MANCR positive cells within 20um of DCIS lesions. Expression near cribriform lesions was lower at an average of 75.0+/- 35.4% and lowest near mixed lesions at an average of 41.3 +/- 16.3% ($p = 0.01$) (Fig. 6B). The association between MANCR and solid DCIS lesions suggests that MANCR expressing cells may play a direct role in recurrence or be involved in re-sculpting the microenvironment to promote recurrence. MANCR has been reported to be involved in more aggressive breast cancers. We (and others) have reported that MANCR is upregulated in triple negative breast cancer cells and contributes to stabilizing the cancer cell genome [29].

Spatial proximity of cells expressing CD90 or CXCL12 is associated with hormone receptor status

Spatial proximity of CD90 high positive cells (9 + RNA copies per cell) to invasive epithelium was significantly associated with positive hormone receptor (HR) status ($p = 0.02$). HR negative specimens had an average of 27.0 +/- 19.0% CD90 high positive cells within 20um of invasive epithelium as compared to HR positive specimens which had an average of 58.0 +/- 19.0% (Fig. 7A-E). CD90 is commonly expressed on mesenchymal stromal cells (MSCs) and is upregulated in invasive breast cancer cell lines [49]. Studies in equine and murine endometrium have found that sex hormones may

play a role in regulating MSC populations, with increased estradiol and progesterone increasing MSC proliferation [50, 51] supporting our observed correlation of CD90 + cells and HR status.

A similar spatial expression pattern was observed with CXCL12 high positive cells and HR status showing significantly higher spatial proximity to benign epithelium in HR positive specimens (p = 0.039). HR negative specimens had an average of 25.5 +/- 11.0% CXCL12 high positive cells within 20um of benign lesions compared to an average of 60.1 +/- 19.1% in HR positive specimens (Fig. 7F-J). CXCL12 overexpression is similarly associated with metastasis and tumor growth and has been shown to be induced by estradiol in lung cancer[52] [53]. Our results suggest that sex hormone induced overexpression of CXCL12 near benign epithelium may be an early indicator of breast cancer invasion.

Figure 7. **CD90 or CXCL12 spatial distribution associate with hormone receptor status(A-J).** Boxplot showing association between hormone receptor status and percent of CD90 high positive cells in stroma of invasive lesions **(A)**. Example images of CD90 RNA expression in HR negative case **(B, C)** vs HR positive case **(D, E)**. Boxplot showing association between hormone receptor and percent of CXCL12 RNA high positive cells in stroma of benign lesions **(F)**. Example images of CXCL12 RNA expression in HR negative case **(G, H)** vs HR positive case **(I, J).** Adjusted P < 0.05 is labelled as significant. *: p < = 0.05, **: p < = 0.01, ***: p < = 0.001, ****: p < = 0.0001.

Spatial proximity of MANCR or RUNX1 expressing cells is associated with recurrence status

Presence of higher spatial proximity of MANCR expressing cells to invasive epithelial boundary, with an average of 50.7 +/- 21.2%, was associated with no recurrence, whereas a lower percentage of MANCR expressing cells proximal to invasive lesion with an average of 37.7 +/- 7.4% was significantly associated with recurrence (P value = 0.04) (Fig. 8A-E). In addition, non-recurring cases had higher percentage of MANCR expressing cells closer to DCIS epithelium with an average of 48.2 +/- 19.2% positive cells than recurring cases with an average of 30.0 +/- 16.6% positive cells (P value = 0.04) (Fig. 8F-J). We have observed in our other studies that MANCR is expressed in MSCs (data not shown). It would be important to identify and phenotype these MANCR or RUNX1 expressing cells present in the tumor stroma near specific types of lesions in future studies. TGFb regulated Runx1 expression has been shown to be necessary for MSC proliferation and myofibroblast differentiation in prostate cancer [54].

Spatial localization of RUNX1/RUNX2 co-expressing cells in stroma is associated with recurrence

Previous studies show a reciprocal relationship between RUNX1 and RUNX2 in breast cancer cell lines, with RUNX1 functioning to repress the breast cancer stem cell phenotype and suppressing tumor growth in vivo [55] and RUNX2 functioning to promote metastasis [56, 57]. These studies were primarily carried out in epithelial breast tumor cell lines; in our sample set we observed that there is a measurable number

of RUNX1/RUNX2 co-expressing cells in the TME surrounding the epithelia that are significantly different between benign, DCIS and invasive epithelium (**Supplementary Fig. 8**). We tested spatial proximity of cells that co-expressed either high RUNX1/low RUNX2 or low RUNX1/high RUNX2 near each type of epithelium within TME. The clinical association analysis identified several spatial proximity relationships from various epithelia boundaries that correlated with clinical metrics that include differentiation status, grade, and recurrence status (Fig. 5). As recurrence was highly relevant to this study, these metrics were examined further. We observed a higher percentage high RUNX1/low RUNX2 co-positive cells in stroma of DCIS lesions was associated with non-recurrence **(**Fig. 9A-E**)**. Non-recurrent tumors had an average of 0.4 +/- 0.3% high RUNX1/low RUNX2 cells near DCIS epithelium whereas recurrent tumors had an average of 0.1 +/- 0.02% high RUNX1/low RUNX2 cells (P value = 0.02). A higher percentage of low RUNX1/high RUNX2 positive cells near benign lesions was associated with non-recurrence (Fig. 9F-J). Non-recurrent samples had an average of 0.5 +/- 0.4% low RUNX1/high RUNX2 positive cells near benign regions whereas recurrent samples had an average of 0.03 +/- 0.05% low RUNX1/high RUNX2 cells (P value = 0.02).

Higher myoepithelial layer continuity is associated with recurrence status

Previous studies have shown that myoepithelial layer continuity in DCIS epithelium could be predictive of progression, with a more continuous myoepithelial layer associated with progressive disease state [58, 59]. All the cases examined in this study were progressors. The myoepithelial continuity of all types of lesions for each patient was measured and tested as a potential predictor of recurrence. Cells co-expressing aSMA and TP63 within individual epithelial regions were designated as myoepithelial cells and the distance between adjacent myoepithelial cells was measured by nearest neighbor analysis. The continuity score was calculated by determining the maximum distance between myoepithelial cells, adding one and subtracting the other measured values from the maximum. The myoepithelial continuity score was compared across benign, DCIS and invasive epithelium. The myoepithelium of benign lesions was most continuous with a score of 145.5 +/- 13.5 and there was progressive loss of continuity in DCIS, 101.6 +/- 44.1, and invasive lesions, 94.2 +/- 43.7 (P values < 0.0001) (Fig. 10A). The continuity scores per patient were then tested against the recurrence status. Recurring cases were found to have a more continuous myoepithelial layer with an average score of 135.8+/- 21.9 than non-recurring cases with an average score of 113.9 +/- 45.1 (P value = 0.05) (Fig. 10B-D).

Discussion

In this study, we aimed to identify RNA and protein signatures within the breast tumor microenvironment that are associated with progression and recurrence of breast cancer by examining breast tumors representing regions with benign (normal adjacent) ducts, DCIS, and IDC areas from individual patients. We focused on protein and RNA markers reflecting several cell types and regulatory processes of interest, including cytokine signaling (CXCL12, CXCR6), MSCs (CD90), myoepithelial cells (TP63, aSMA),

T cells (CD3), macrophages (CD68), and CAFs (aSMA, CD34). We selected RNAs that exhibit cancer-associated expression patterns in breast cancer cell lines (MANCR, RUNX1, RUNX2).

We compared the expression patterns of each marker in breast tumor and in breast TME across the different types of epithelia and stroma to identify changes in expression as breast cancer lesion progresses from benign to DCIS to IDC. We found increased expression of RUNX1, RUNX2, CD90, CXCR6, and aSMA in the stroma of invasive lesions compared to normal adjacent ducts while CXCL12 and CD34 expression levels decreased. Together, these results indicate increase in mesenchymal stromal cells, CAFs, and perhaps chemokine signaling.

TP63, RUNX1, and RUNX2 RNA levels decreased in epithelium as lesion types progressed. The loss of TP63 RNA may represent a decrease in myoepithelial cells, which was further investigated using TP63/aSMA protein co-expressing cells to reinforce the observation with TP63 RNA expression. Other studies have reported similar finding with a loss of myoepithelial cells during breast cancer progression [58]. RUNX1 is known to play a tumor suppressive role by stabilizing the mammary epithelial cell phenotype and preventing an epithelial to mesenchymal transition. There is decreased RUNX1 expression in tumorigenic and metastatic breast cancer cells [55]. Our finding that the RUNX1 expression levels in epithelium were lower in areas of IDC is consistent with its tumor suppressor activity. Increase expression of RUNX2 in metastatic or triple negative breast cancer (TNBC) epithelium is correlated with increased invasion, metastasis and poor outcomes [60]. In our cohort of mostly HR-positive breast tumor specimens we did not observe significant difference in expression of RUNX2 between benign, DCIS or IDC epithelium. Interestingly we did observe significant difference for expression of RUNX2 between benign and invasive stroma. RUNX2 overexpression has been shown to enhance proliferation of specific subtypes of T-cells [61], suggesting a dual role for RUNX2 in a context dependent manner that needs to be explored further.

The increased levels of CD3 in DCIS stroma specifically suggest an increased immune response with heightened T cell activity in DCIS. This is consistent with reports that T cells are upregulated during the transition from DCIS to IDC [62]. As these cases all included IDC, increased CD3 expression near DCIS would be expected. Tumor infiltrating lymphocytes (TILs) are prognostic in IDC, and in DCIS their spatial arrangement has been shown to correlate with Oncotype DX (ODX) Breast DCIS Score which measures local recurrence risk [63]. TILs touching the ducts were associated with a higher score and worse prognosis. In our spatial analysis, we did not observe correlation with spatial proximity of CD3 expression near and recurrence status. The predictive prognostic value of TILs is specific to breast cancer subtypes with higher correlation associated with HER + or TNBC tumors. Higher TILs can predict improved treatment response and survival outcomes in early TNBC tumors. Presence of TILs in HR positive breast tumors has lower mean counts and can be heterogenous and is not prognostic for disease free survival[64].

Although the expression patterns of RNAs and proteins studies exhibited differential expression between specific types of epithelia and stroma, the average expression patterns were not found to associate

significantly with any clinical parameters. Therefore, we performed spatial proximity analysis to explore whether the spatial location of cells expressing these RNAs and proteins would influence the disease phenotype or clinical outcomes of breast cancer. We anticipated that spatial distribution would provide further insights into how the tumor microenvironment affects tumor progression and recurrence because the identity of cells in proximity to the epithelial boundary could indicate how the local environment is influencing the tumors. We chose a tumor stroma boundary of 20um away from the epithelium which is approximately one-two cell heights to evaluate the edge effect [65]. The spatial analysis provided several correlations with clinical parameters that can be further examined.

Increased MANCR expression near DCIS epithelium significantly correlated with presence of solid DCIS lesions compared to cribriform and mixed lesion, suggesting that focal MANCR expression may be related to tumor progression. Cell line studies have shown increased expression of MANCR in aggressive breast cancer cell lines[29] and the association of MANCR with progressive, solid DCIS lesions observed in this study, indicates that MANCR may be an indicator if progression towards invasive cancer through metastasis.

Hormone receptor status is an important prognostic marker for identifying various breast cancer phenotypes and their potential for progression to aggressive cancer [11]. Our spatial analysis of the stromal compartment revealed a higher percentage CD90/ Thy1 expressing cells were associated with positive HR status when expressed near IDC epithelium and a higher percentage of CXCL12 expressing cells were present near the benign ducts of HR positive tumors. Studies in other cancer models have shown a link between estradiol and progesterone increasing MSC proliferation [50, 51] and another study showed a link between CXCL12 overexpression and estradiol [52, 53]. These studies and our observations suggest a link between sex hormones, chemokine signaling, and MSC proliferation which can be an important avenue for further exploration. It is critical to identify ways MSCs are regulated in breast cancer because the presence of MSCs in the microenvironment increase the metastatic potential of breast cancer cells [66]. MSCs are regulated through chemokine signaling, which is the basis for focusing on CXCL12 and CXCR6 to examine the MSC role in the tumor microenvironment along with CD90 [67]. We observed that expression of CD90 RNA investigated, using RNA in situ hybridization, was relatively stable compared to CD90 protein detected using antibody-based methods. We tested several commercially available CD90 antibodies and found they lacked reliability. Especially in FFPE samples, CD90 protein expression was unreliable and not well preserved in comparison to RNA expression. The correlation between MSCs, chemokine signaling and changing dynamics based on HR status could be important to identify novel treatment targets.

Larger populations of RUNX1 expressing cells and cells expressing high RUNX1/low RUNX2 near the DCIS epithelial boundary were found to be associated with non-recurrence. RUNX transcription factors play a context dependent role in many cancers, including breast cancer, where they can act as both tumor promoters and tumor suppressors. [68, 69]. RUNX factors are also known to play a role in proliferation and maintenance of immune cells (e.g., CD4 + T cells, CD8 + T cells) as well as proliferation of CAFs and MSCs [61, 68, 70, 71]. Our findings of predictive spatial relationship between RUNX1 /

RUNX2 expressing cells in the stromal cell population near DCIS or benign epithelium could be further assessed in an expanded cohort of stratified samples that include aggressive and non-aggressive tumors.

MANCR positive cells were elevated near non-recurring invasive lesions. This is an unexpected result as MANCR is generally elevated in more aggressive breast cancer cell lines, particularly in TNBC epithelial cells. In our analysis of tumor stroma, we observed that MANCR was consistently expressed at low levels in the stomal cell population surrounding the epithelium.

Another factor associated with recurrence status by spatial proximity analysis was a higher myoepithelial layer continuity, as measured by the distance between aSMA/TP63 co-expressing cells within ductal epithelium. Our data corroborated a reported finding that reduced myoepithelial continuity observed as breast tumors progressed from benign to DCIS to IDC can predict the progression of DCIS to invasive cancer [42, 58]. Although counterintuitive, the observation that fewer cells bordering an epithelial lesion could be protective against recurrence, it has been reported that low myoepithelial continuity corresponds with adaptive immune response, leukocyte proliferation, tumor necrosis factor (TNF) superfamily cytokine production, and major histocompatibility (MHC) Class II antigen assembly, processing, and presentation[42]. A lack of myoepithelial continuity may allow immune cells to penetrate a lesion and facilitate an immune response.

Potential future directions include studying the mechanism by which the myoepithelial layer becomes discontinuous to identify potential drug targets for preventing recurrence as well as to perform an expanded analysis with markers for several different types of immune cells to determine which types of immune cells are present near a discontinuous myoepithelial layer versus a continuous myoepithelial layer. A larger panel of markers would be particularly useful in phenotyping cells that are expressing MANCR, CXCR6, and RUNX1 in tumor stroma. Highplex imaging-based proteomics or spatial transcriptomics studies to better identify and phenotype cells expressing the markers used in this study can further define the importance of spatial distribution of some of the markers within TMEs.

Conclusions

Overall, our study highlights the importance of studying the spatial localization of cells within the tumor epithelium, tumor microenvironment, and tumor stroma boundary. Global expression levels of RNAs and proteins in the entire lesion may not identify localized stromal dynamics acting on the tumor cells whereas a dedicated spatial analysis can. This work suggests several parameters that may broaden understanding by further mechanistic studies with potential implications for drug development. The finding that decreased myoepithelial continuity is protective against recurrence could help to identify patients who are better candidates for immunotherapy. This work emphasizes the importance of spatial analysis in imaging data and a basis for further exploration of the breast cancer tumor microenvironment.

Abbreviations

DCIS
Ductal Carcinoma in Situ
IDC
Invasive Ductal Carcinoma
TME
Tumor Microenvironment
TMAs
Tissue Microarrays
RUNX1
Runt–related transcription factor 1
RUNX2
Runt–related transcription factor 2
MANCR
Mitotically activated long non–coding RNA
CD90
Cluster of Differentiation 90
CXCL12
C–X–C motif chemokine 12
CXCR6
C–X–C chemokine receptor type 6
TP63
Tumor Protein 63
aSMA
Alpha Smooth Muscle Actin
CD34
Cluster of Differentiation 34
CD68
Cluster of Differentiation 68
CD3
Cluster of Differentiation 3
FISH
Fluorescence In Situ Hybridization
IF
Immuno Fluorescence
CPM
Contralateral Prophylactic Mastectomy
ER
Estrogen Receptor

PR
Progesterone Receptor
HER2
Human Epidermal Growth Factor Receptor 2
Ki67
Antigen Kiel 67
MSCs
Mesenchymal Stromal Cells
EMT
Epithelial to Mesenchymal Transition
lncRNAs
Long Non Coding RNAs
BCSCs
Breast Cancer Stem Cells
MCL
Mantle Cell Lymphoma
IHC
Immunohistochemistry
VBCSS
Vermont Breast Cancer Surveillance System
FFPE
Formalin Fixed Paraffin Embedded
LVI
Lymphovascular Invasion
CAFs
Cancer Associated Fibroblasts
ILC
Invasive Lobular Carcinoma
HR
Hormone Receptor
TNBC
Triple Negative Breast Cancer
TILs
Tumor Infiltrating Lymphocytes
TNF
Tumor Necrosis Factor
MHC
Major Histocompatibility Complex

Declarations

Ethics approval and consent to participate:

This study was approved by the University of Vermont Committee on Human Research in the Medical Sciences (IRB protocol 15-629), with a waiver of consent for the use of existing clinical specimens and data in this study.

Availability of data and materials:

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests:

The authors declare that they have no competing interests.

Funding:

This work is supported by P01CA240685 and the Arthur Jason Perelman Professorship and the Charlotte Perelman Cancer Fund. NIH grant U54 GM115516 for the Northern New England Clinical Translational Research (NNE-CTR) and Technology Development Initiative award through NNE-CTR.

Authors contributions:

KL and PNG conceptualized and designed the work; KL, PNG, and NAB acquired and analyzed the imaging data; MFE, DLW, AB and BLS facilitated with acquisition of breast TMAs, annotations of the TMAs with appropriate histological properties/features and compilation clinical data. All authors contributed to interpretation of data and manuscript writing.

Acknowledgements:

Imaging and/or Cytometry work was performed at Microscopy Imaging and Cytometry at the University of Vermont (RRID# SCR_018821). The Leica-Aperio VERSA whole slide imager was purchased using the College of Medicine Shared Instrumentation Award (Douglas Taatjes). We would like to thank Douglas Taatjes, director of MIC at UVM for invaluable support in imaging studies.

References

1. Sherman, R.L., et al., *Annual Report to the Nation on the Status of Cancer, featuring state-level statistics after the onset of the COVID-19 pandemic.* Cancer, 2025. **131**(9): p. e35833.
2. Siegel, R.L., et al., *Cancer statistics, 2025.* CA: A Cancer Journal for Clinicians, 2025. **75**(1): p. 10-45.
3. Collins, L.C., et al., *Outcome of patients with ductal carcinoma in situ untreated after diagnostic biopsy.* Cancer, 2005. **103**(9): p. 1778-1784.

4. Page, D.L., et al., *Continued local recurrence of carcinoma 15-25 years after a diagnosis of low grade ductal carcinoma in situ of the breast treated only by biopsy.* Cancer, 1995. **76**(7): p. 1197-200.
5. Page, D.L., et al., *Intraductal carcinoma of the breast: follow-up after biopsy only.* Cancer, 1982. **49**(4): p. 751-8.
6. Sanders, M.E., et al., *The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up.* Cancer, 2005. **103**(12): p. 2481-2484.
7. Dominici, L., et al., *Association of Local Therapy With Quality-of-Life Outcomes in Young Women With Breast Cancer.* JAMA Surg, 2021. **156**(10): p. e213758.
8. Wong, S.M., et al., *Growing Use of Contralateral Prophylactic Mastectomy Despite no Improvement in Long-term Survival for Invasive Breast Cancer.* Annals of Surgery, 2017. **265**(3): p. 581-589.
9. Kapur, H., et al., *Decreasing contralateral prophylactic mastectomy rates in average-risk women with unilateral breast cancer.* The American Journal of Surgery, 2021. **221**(6): p. 1172-1176.
10. Pawloski, K.R., et al., *Patterns of invasive recurrence among patients originally treated for ductal carcinoma in situ by breast-conserving surgery versus mastectomy.* Breast Cancer Res Treat, 2021. **186**(3): p. 617-624.
11. Williams, K.E., et al., *Molecular phenotypes of DCIS predict overall and invasive recurrence†.* Annals of Oncology, 2015. **26**(5): p. 1019-1025.
12. Vargas, A.C., et al., *Gene expression profiling of tumour epithelial and stromal compartments during breast cancer progression.* Breast Cancer Res Treat, 2012. **135**(1): p. 153-65.
13. Ma, X.J., et al., *Gene expression profiling of the tumor microenvironment during breast cancer progression.* Breast Cancer Res, 2009. **11**(1): p. R7.
14. Allinen, M., et al., *Molecular characterization of the tumor microenvironment in breast cancer.* Cancer Cell, 2004. **6**(1): p. 17-32.
15. Wu, S.Z., et al., *A single-cell and spatially resolved atlas of human breast cancers.* Nat Genet, 2021. **53**(9): p. 1334-1347.
16. Wang, J., et al., *Progression from ductal carcinoma in situ to invasive breast cancer: molecular features and clinical significance.* Signal Transduction and Targeted Therapy, 2024. **9**(1): p. 83.
17. Kim, M., et al., *Immune microenvironment in ductal carcinoma in situ: a comparison with invasive carcinoma of the breast.* Breast Cancer Research, 2020. **22**(1): p. 32.
18. El Hentati, F.Z., et al., *Variability of CD3 membrane expression and T cell activation capacity.* Cytometry B Clin Cytom, 2010. **78**(2): p. 105-14.
19. Gottfried, E., et al., *Expression of CD68 in non-myeloid cell types.* Scand J Immunol, 2008. **67**(5): p. 453-63.
20. Cammarota, F. and M.O. Laukkanen, *Mesenchymal Stem/Stromal Cells in Stromal Evolution and Cancer Progression.* Stem Cells Int, 2016. **2016**: p. 4824573.

21. Zeng, F., et al., *Role and mechanism of CD90(+) fibroblasts in inflammatory diseases and malignant tumors.* Mol Med, 2023. **29**(1): p. 20.
22. Svensson, A., et al., *Identification of two distinct mesenchymal stromal cell populations in human malignant glioma.* J Neurooncol, 2017. **131**(2): p. 245-254.
23. Gordon, J.A.R., et al., *Identification of molecularly unique tumor-associated mesenchymal stromal cells in breast cancer patients.* PLOS ONE, 2023. **18**(3): p. e0282473.
24. Tuo, Z., et al., *RUNX1 is a promising prognostic biomarker and related to immune infiltrates of cancer-associated fibroblasts in human cancers.* BMC Cancer, 2022. **22**(1): p. 523.
25. Zhang, D., et al., *Cancer-associated fibroblasts promote tumor progression by lncRNA-mediated RUNX2/GDF10 signaling in oral squamous cell carcinoma.* Mol Oncol, 2022. **16**(3): p. 780-794.
26. Wietecha, M.S., et al., *Phase-specific signatures of wound fibroblasts and matrix patterns define cancer-associated fibroblast subtypes.* Matrix Biol, 2023. **119**: p. 19-56.
27. Jung, Y., et al., *Recruitment of mesenchymal stem cells into prostate tumours promotes metastasis.* Nat Commun, 2013. **4**: p. 1795.
28. Li, Z., et al., *Cancer-associated fibroblasts-derived lncRNA signature as a putative biomarker in breast cancer.* Front Oncol, 2022. **12**: p. 1028664.
29. Tracy, K.M., et al., *Mitotically-Associated lncRNA (MANCR) Affects Genomic Stability and Cell Division in Aggressive Breast Cancer.* Mol Cancer Res, 2018. **16**(4): p. 587-598.
30. Yao, L., et al., *Upregulation of MANCR predicts poor survival in patients with gastric cancer.* Oncology letters, 2019. **18**(6): p. 6801-6806.
31. Tahmouresi, F., et al., *Upregulation of the long noncoding RNAs DSCAM-AS1 and MANCR is a potential diagnostic marker for breast carcinoma.* Biotechnol Appl Biochem, 2021. **68**(6): p. 1250-1256.
32. Kas, S.M., et al., *Insertional mutagenesis identifies drivers of a novel oncogenic pathway in invasive lobular breast carcinoma.* Nat Genet, 2017. **49**(8): p. 1219-1230.
33. Pereira, B., et al., *The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes.* Nat Commun, 2016. **7**: p. 11479.
34. Browne, G., et al., *Runx1 is associated with breast cancer progression in MMTV-PyMT transgenic mice and its depletion in vitro inhibits migration and invasion.* J Cell Physiol, 2015. **230**(10): p. 2522-32.
35. Chuang, L.S., K. Ito, and Y. Ito, *Roles of RUNX in Solid Tumors.* Adv Exp Med Biol, 2017. **962**: p. 299-320.
36. Fritz, A.J., et al., *RUNX1 and RUNX2 transcription factors function in opposing roles to regulate breast cancer stem cells.* Journal of Cellular Physiology, 2020. **235**(10): p. 7261-7272.
37. Wen, S., et al., *Downregulation of MANCR inhibits cancer cell proliferation in mantle cell lymphoma possibly by interacting with RUNX2.* Acta Biochim Biophys Sin (Shanghai), 2019. **51**(11): p. 1142-1147.
38. Sprague, B.L., et al., *Long-term Mammography Screening Trends and Predictors of Return to Screening after the COVID-19 Pandemic: Results from a Statewide Registry.* Radiol Imaging Cancer, 2024. **6**(3):

p. e230161.

39. Evans, M.F., et al., *Microarray and RNA in situ hybridization assay for recurrence risk markers of breast carcinoma and ductal carcinoma in situ: Evidence supporting the use of diverse pathways panels.* J Cell Biochem, 2020. **121**(2): p. 1736-1746.
40. Taatjes, D.J., et al., *The Shared Core Resource as a Partner in Innovative Scientific Research: Illustration from an Academic Microscopy Imaging Center.* J Biomol Tech, 2022. **33**(1).
41. Lee, K., et al., *Novel method of paraffin embedding cultured cells and organoids using silicone molds.* Histochemistry and Cell Biology, 2025. **163**(1): p. 37.
42. Borowsky, A., et al., *Tumor microenvironmental determinants of high-risk DCIS progression.* Res Sq, 2024.
43. Lin, C.S., et al., *Commonly used mesenchymal stem cell markers and tracking labels: Limitations and challenges.* Histol Histopathol, 2013. **28**(9): p. 1109-16.
44. Toss, M.S., et al., *The prognostic significance of immune microenvironment in breast ductal carcinoma in situ.* Br J Cancer, 2020. **122**(10): p. 1496-1506.
45. Díaz-Flores, L., et al., *CD34+ Stromal Cells/Telocytes as a Source of Cancer-Associated Fibroblasts (CAFs) in Invasive Lobular Carcinoma of the Breast.* Int J Mol Sci, 2021. **22**(7).
46. Thiago Acrux, D.A., Débora Geudêncio, Caroline Rocha, , *Ductal carcinoma in situ of the breast: correlation of architectural, cytological, IHC findings and recurrence analysis.* Jornal Brasilieiro de Patologia e Medicina Laboratorial, 2020. **56**.
47. Cheng, H. and N.Y. Chen, *[Expression of Chemokine receptor CXCR6 and its significance in breast cancer cell lines].* Sichuan Da Xue Xue Bao Yi Xue Ban, 2014. **45**(3): p. 405-9.
48. Xiao, G., et al., *CXCL16/CXCR6 chemokine signaling mediates breast cancer progression by pERK1/2-dependent mechanisms.* Oncotarget, 2015. **6**(16): p. 14165-78.
49. Lobba, A.R.M., et al., *Differential expression of CD90 and CD14 stem cell markers in malignant breast cancer cell lines.* Cytometry Part A, 2012. **81A**(12): p. 1084-1091.
50. Rink, B.E., et al., *Reproductive stage and sex steroid hormone levels influence the expression of mesenchymal stromal cell (MSC) markers in the equine endometrium.* Theriogenology, 2018. **116**: p. 34-40.
51. Gaidamaka, A.O., et al., *Novel insight into CD90 mesenchymal stem cell marker in mouse endometrial stroma during embryo implantation.* Reproduction, 2025. **169**(3): p. e240375.
52. Rodriguez-Lara, V., et al., *Estrogen Receptor Beta and CXCR4/CXCL12 Expression: Differences by Sex and Hormonal Status in Lung Adenocarcinoma.* Archives of Medical Research, 2014. **45**(2): p. 158-169.
53. Rodriguez-Lara, V., G.-S. Ignacio, and M.A. Cerbón Cervantes, *Estrogen induces CXCR4 overexpression and CXCR4/CXL12 pathway activation in lung adenocarcinoma cells in vitro.* Endocrine Research, 2017. **42**(3): p. 219-231.

54. Kim, W., et al., *RUNX1 is essential for mesenchymal stem cell proliferation and myofibroblast differentiation.* Proc Natl Acad Sci U S A, 2014. **111**(46): p. 16389-94.
55. Hong, D., et al., *Runx1 stabilizes the mammary epithelial cell phenotype and prevents epithelial to mesenchymal transition.* Oncotarget, 2017. **8**(11): p. 17610-17627.
56. Binay, S. and E. Kaptan, *Transcription factor Runx2 changes the expression of some matricellular proteins in metastatic breast cancer cells.* Molecular Biology Reports, 2022. **49**(7): p. 6433-6441.
57. Zhang, P., et al., *Runx2 is required for activity of CD44(+)/CD24(-/low) breast cancer stem cell in breast cancer development.* Am J Transl Res, 2020. **12**(5): p. 2305-2318.
58. Risom, T., et al., *Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma.* Cell, 2022. **185**(2): p. 299-310.e18.
59. Borowsky, A., et al., *Tumor microenvironmental determinants of high-risk DCIS progression.* Research Square, 2024: p. rs.3.rs-4126092.
60. McDonald, L., et al., *RUNX2 correlates with subtype-specific breast cancer in a human tissue microarray, and ectopic expression of Runx2 perturbs differentiation in the mouse mammary gland.* Disease Models & Mechanisms, 2014. **7**(5): p. 525-534.
61. DeGolier, K.R., et al., *Antigen experience history directs distinct functional states of CD8(+) CAR T cells during the antileukemia response.* Nat Immunol, 2025. **26**(1): p. 68-81.
62. Lv, S., et al., *Functional CD3+CD8+PD1− T Cell Accumulation and PD-L1 Expression Increases During Tumor Invasion in DCIS of the Breast.* Clinical Breast Cancer, 2019. **19**(5): p. e617-e623.
63. Komforti, M., et al., *Tumour-infiltrating lymphocytes in ductal carcinoma in situ (DCIS)-assessment with three different methodologies and correlation with Oncotype DX DCIS Score.* Histopathology, 2020. **77**(5): p. 749-759.
64. Valenza, C., et al., *Tumor Infiltrating Lymphocytes across Breast Cancer Subtypes: Current Issues for Biomarker Assessment.* Cancers (Basel), 2023. **15**(3).
65. Feng, Y., et al., *Spatially organized tumor-stroma boundary determines the efficacy of immunotherapy in colorectal cancer patients.* Nature Communications, 2024. **15**(1): p. 10259.
66. Karnoub, A.E., et al., *Mesenchymal stem cells within tumour stroma promote breast cancer metastasis.* Nature, 2007. **449**(7162): p. 557-563.
67. Liu, S., et al., *Breast Cancer Stem Cells Are Regulated by Mesenchymal Stem Cells through Cytokine Networks.* Cancer Research, 2011. **71**(2): p. 614-624.
68. Gao, L. and F. Zhou, *Comprehensive Analysis of RUNX and TGF-β Mediated Regulation of Immune Cell Infiltration in Breast Cancer.* Frontiers in Cell and Developmental Biology, 2021. **Volume 9 - 2021**.
69. Hong, D., et al., *Suppression of Breast Cancer Stem Cells and Tumor Growth by the RUNX1 Transcription Factor.* Molecular Cancer Research, 2018. **16**(12): p. 1952-1964.
70. Halperin, C., et al., *Global DNA Methylation Analysis of Cancer-Associated Fibroblasts Reveals Extensive Epigenetic Rewiring Linked with RUNX1 Upregulation in Breast Cancer Stroma.* Cancer Research, 2022. **82**(22): p. 4139-4152.

71. Wu, Y., et al., *Identification of cancer-associated fibroblast subtypes and prognostic model development in breast cancer: role of the RUNX1/SDC1 axis in promoting invasion and metastasis.* Cell Biology and Toxicology, 2025. **41**(1): p. 21.

Figures

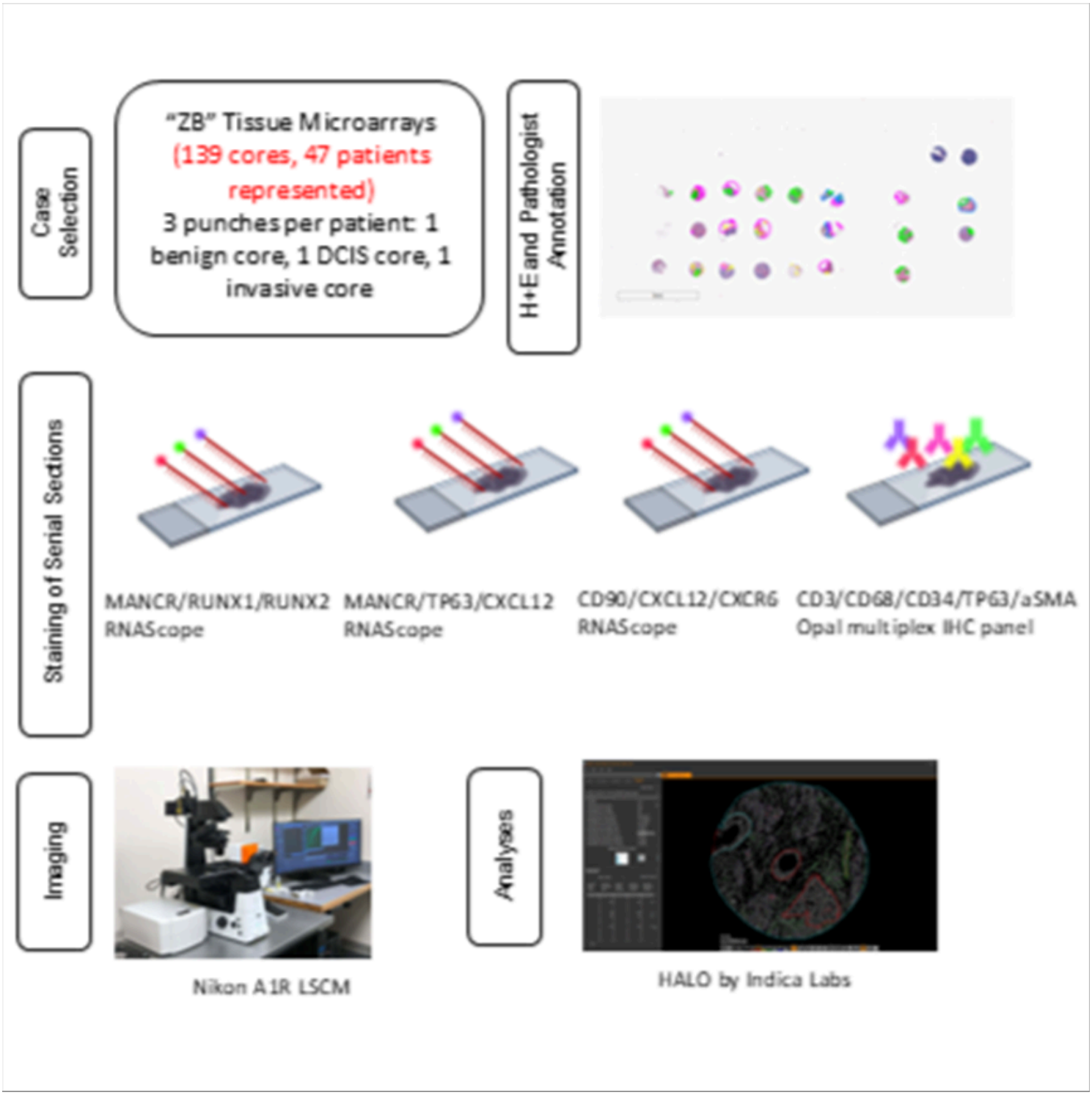


Figure 1

Overview of tissue selection and staining panels. 5 tissue microarrays with invasive, DCIS, and benign lesions from 47 patients were serially stained with the multiplex immunofluorescence panel and 3 RNA FISH panels. Stained slides were imaged on the Nikon A1RHD laser scanning confocal microscope. All images were analyzed in HALO by Indica Labs. Created with BioRender.com

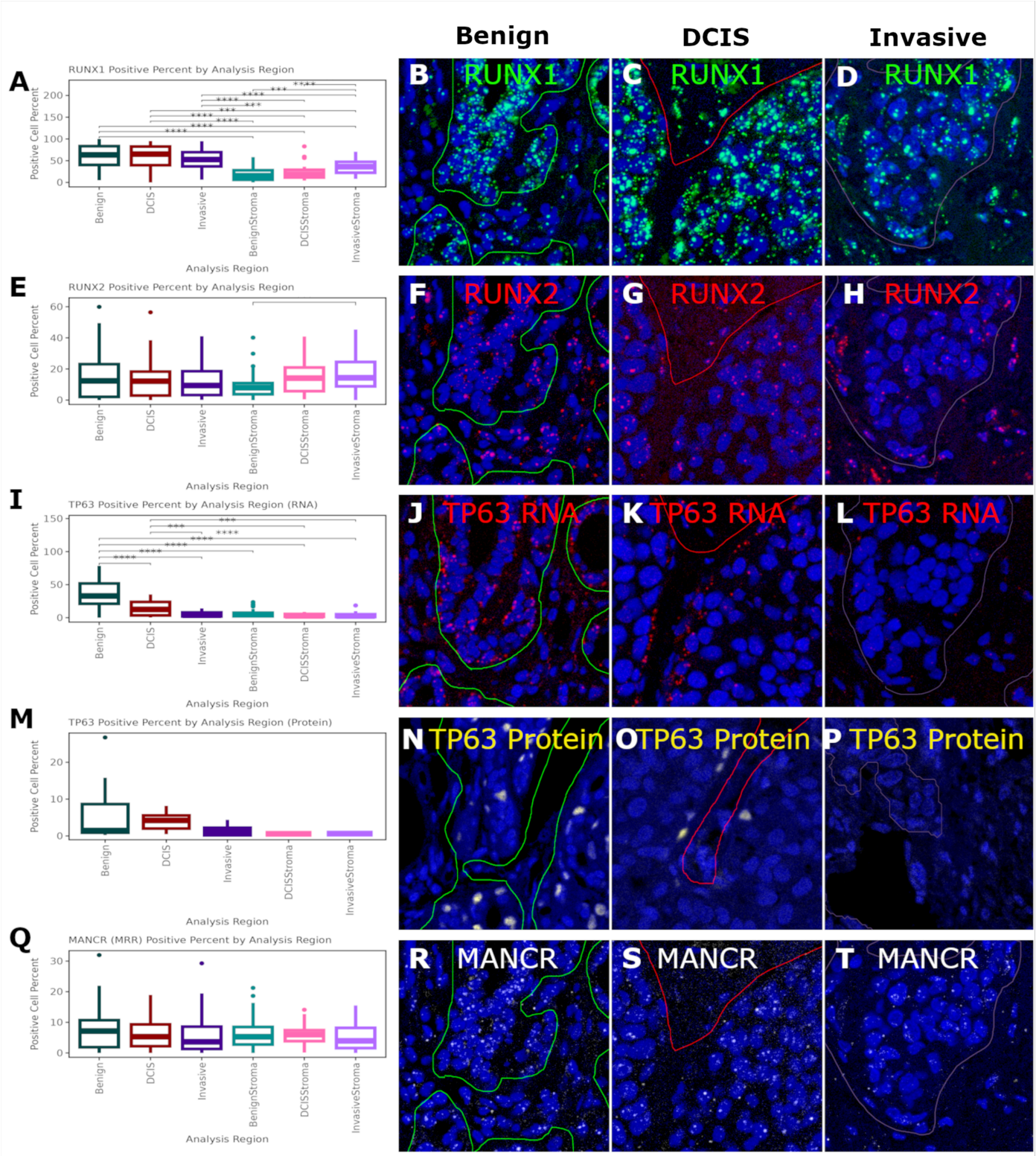


Figure 2

Distribution of tumor markers in breast cancer progression lesions (A-T). Boxplots showing distribution of individual RNAs/protein across epithelium and stroma. Representative micrographs show expression of individual RNAs within either benign, DCIS or invasive regions. RUNX1 RNA (A-D). RUNX2 RNA (E-H). TP63 RNA (I-L). TP63 protein (M-P). MANCR RNA (Q-T). Adjusted P<0.05 is labelled as significant. *: p<= 0.05, **: p<= 0.01, ***: p<= 0.001, ****: p<= 0.0001

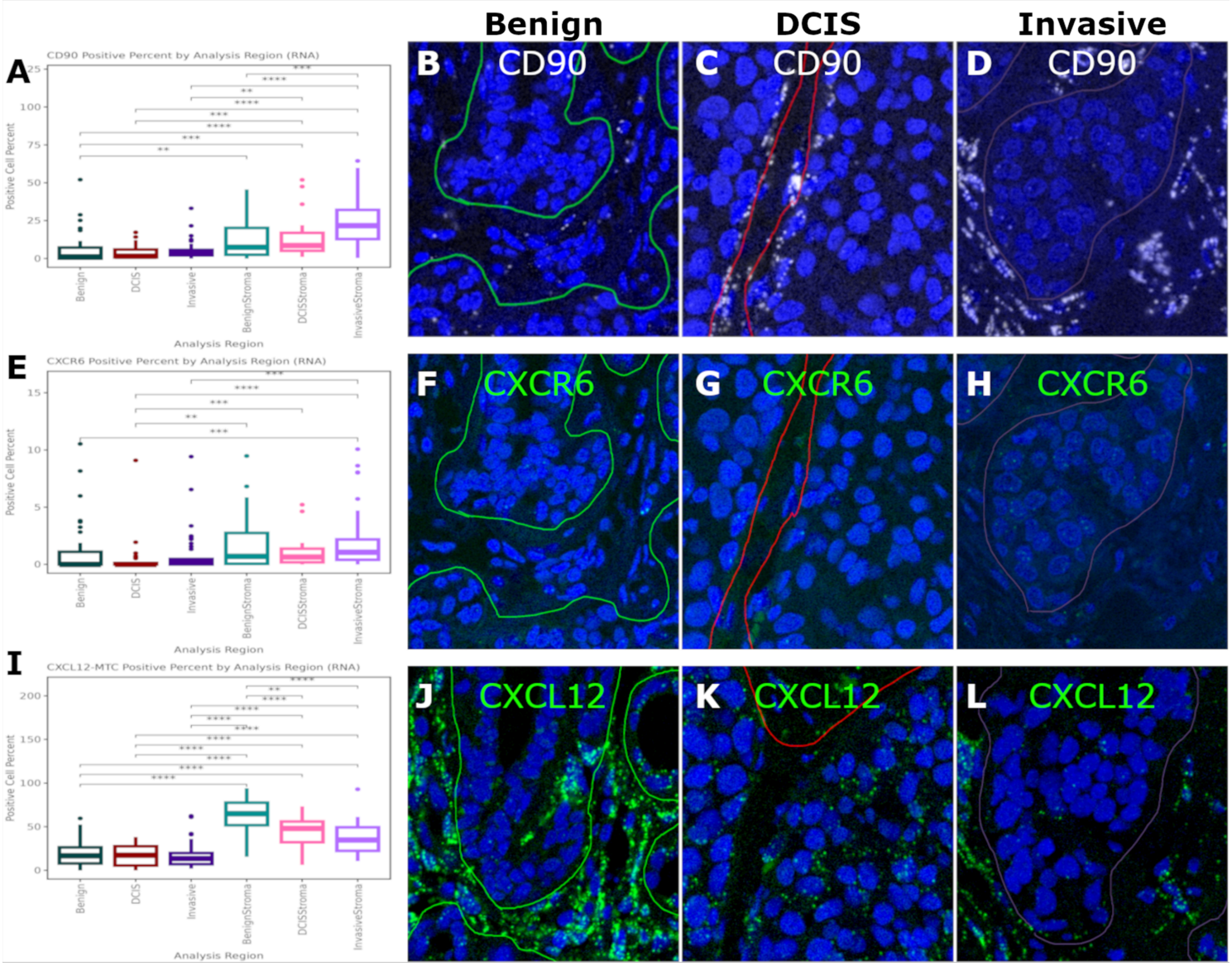


Figure 3

Distribution of stromal markers in breast cancer progression lesions (A-L). Boxplots showing distribution of individual RNAs across epithelium and stroma. Representative micrographs show expression of individual RNAs within either benign, DCIS or invasive regions. CD90 RNA (A-D). CXCR6 RNA (E-H). CXCL12 RNA (I-L). Adjusted P<0.05 is labelled as significant. *: p<= 0.05, **: p<= 0.01, ***: p<= 0.001, ****: p<= 0.0001.

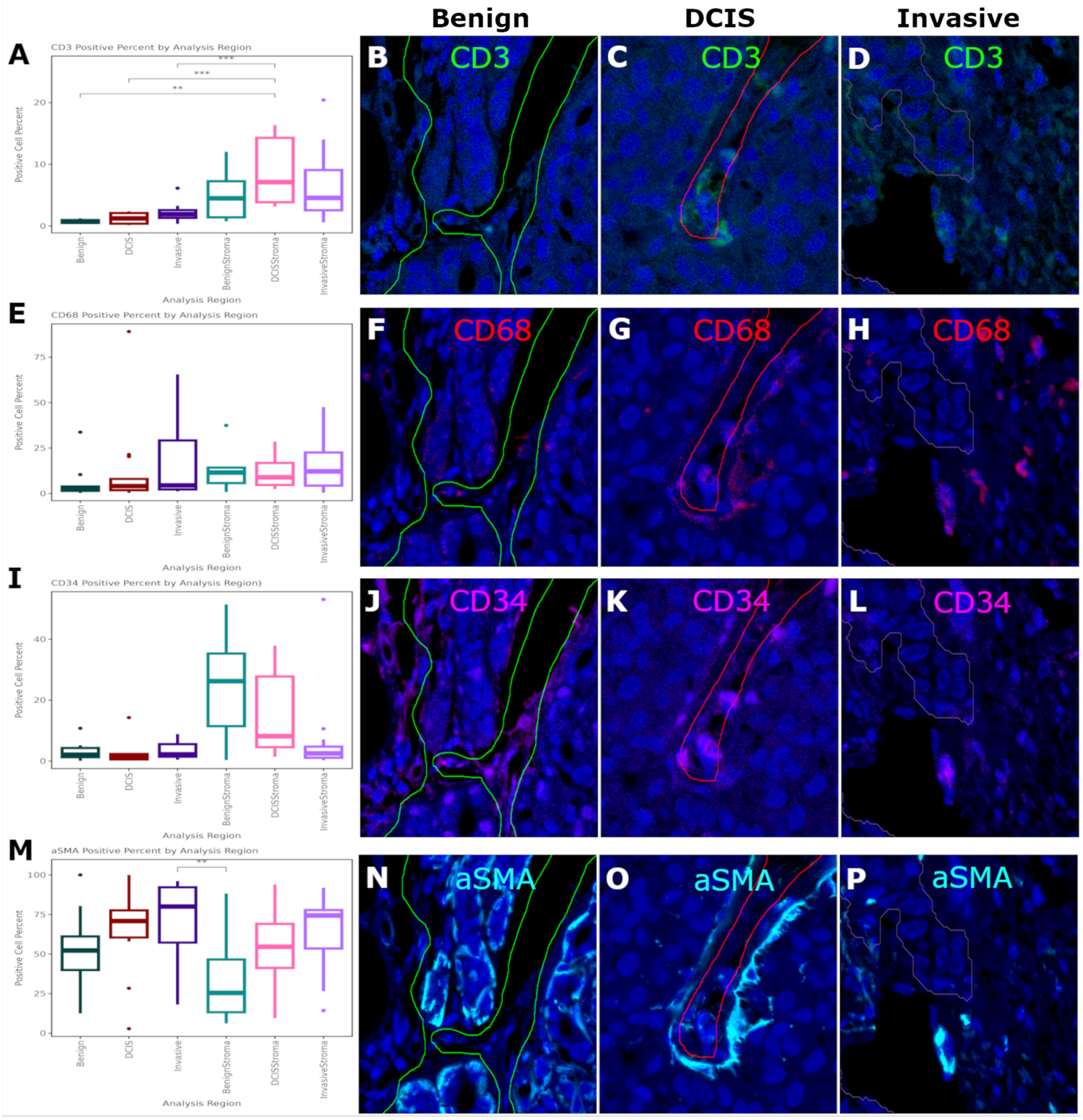


Figure 4

Distribution of TME marker protein in breast cancer progression lesions (A-P). Boxplots showing distribution of individual proteins across epithelium and stroma. Representative micrographs show expression of individual proteins within either benign, DCIS or invasive regions. CD3 protein (A-D). CD68 protein (E-H). CD34 protein (I-L). aSMA protein (M-P). Adjusted P<0.05 is labelled as significant. *: p<= 0.05, **: p<= 0.01, ***: p<= 0.001, ****: p<= 0.0001.

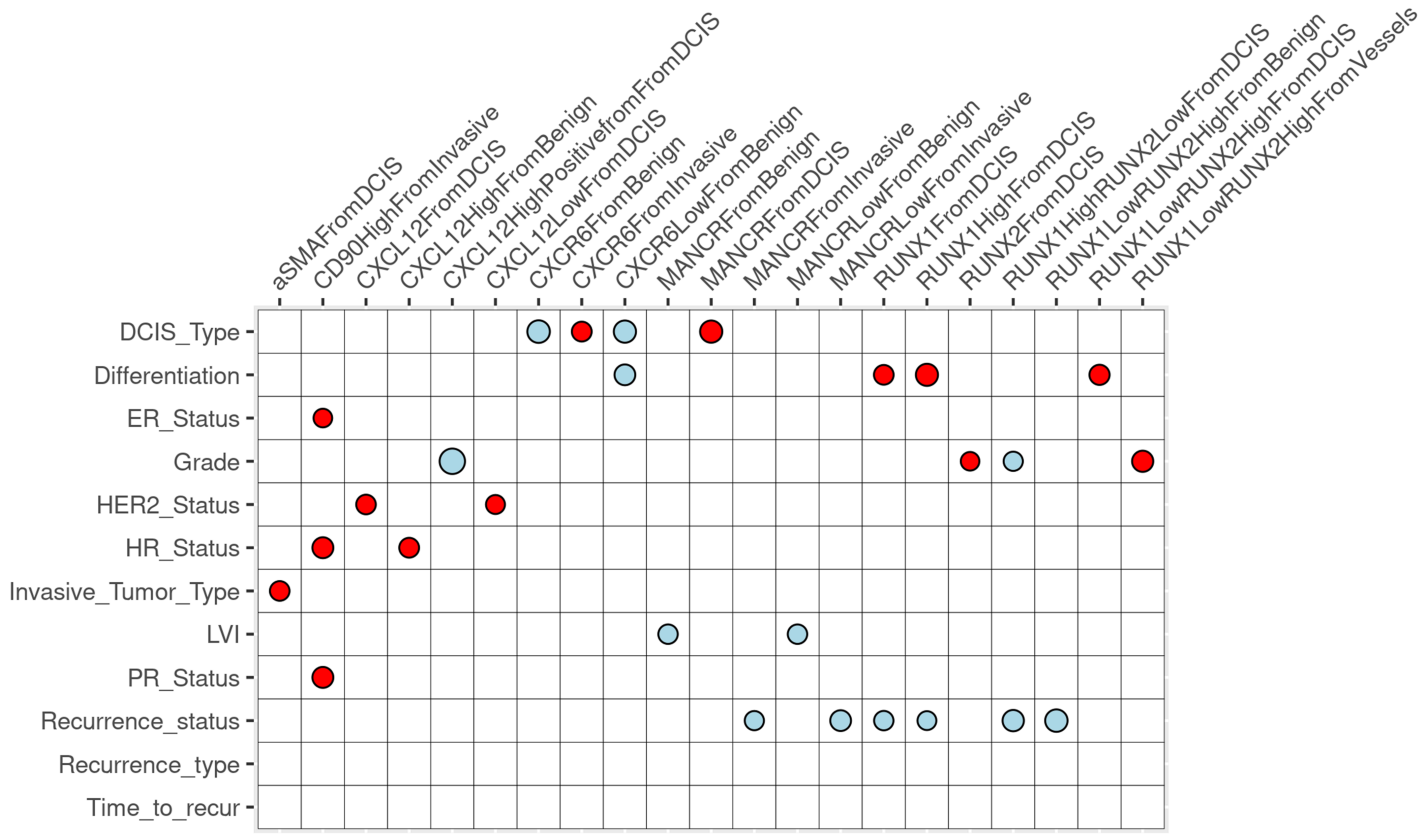


Figure 5

Integrative spatial analysis of marker expression and clinical parameters. Dot matrix of average positive cell percent within 20um from benign, DCIS, and invasive lesions showing association with clinical parameters using Kruskal-Wallis test. Color of dot indicates direction of association (red – positive association; blue – negative). Size of dot is proportional to significance (larger dot → smaller p-value).

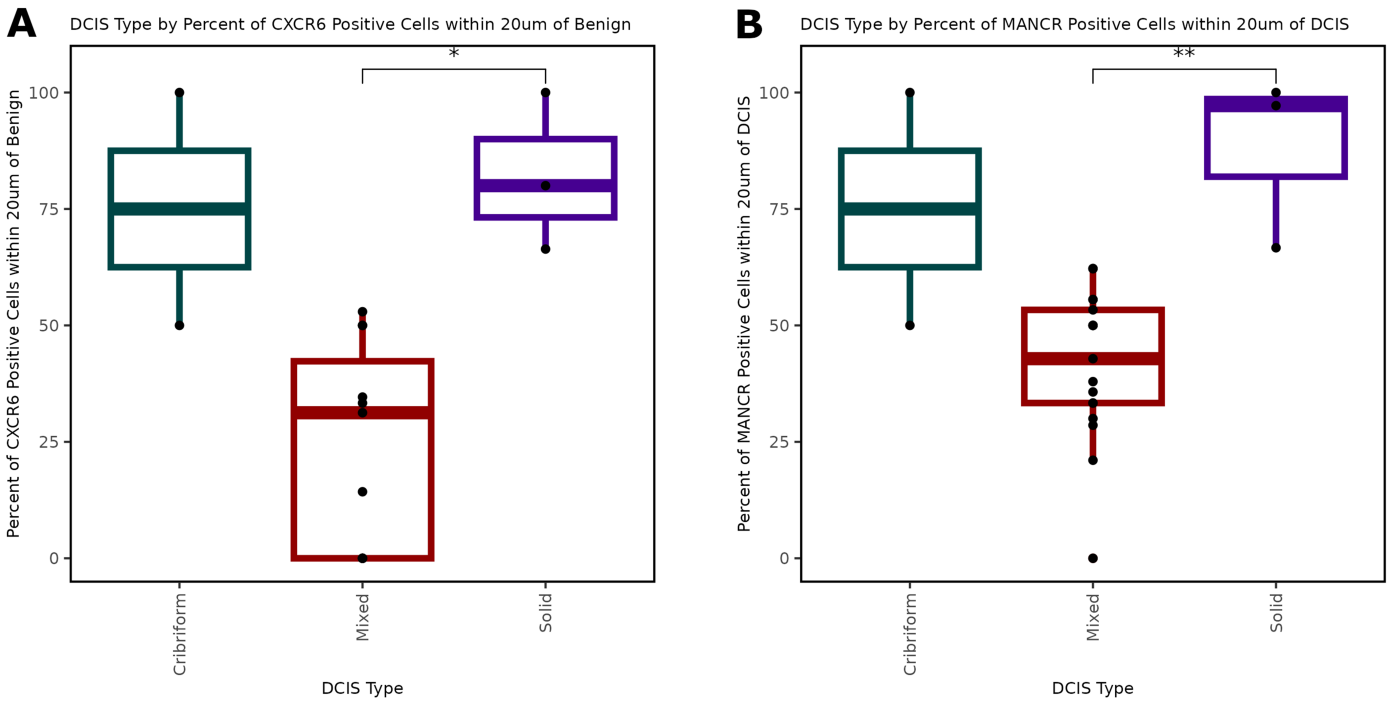


Figure 6

CXCR6 and MANCR spatial distribution associate with DCIS subtypes(A-B). Boxplot showing spatial proximity of CXCR6 positive cells from benign epithelial boundary is significantly associated with solid vs mixed DCIS subtype (A). Boxplot showing spatial proximity of MANCR positive cells from DCIS epithelial boundary is significantly associated with solid vs mixed DCIS subtype. (B). Adjusted P<0.05 is labelled as significant. *: p<= 0.05, **: p<= 0.01, ***: p<= 0.001, ****: p<= 0.0001.

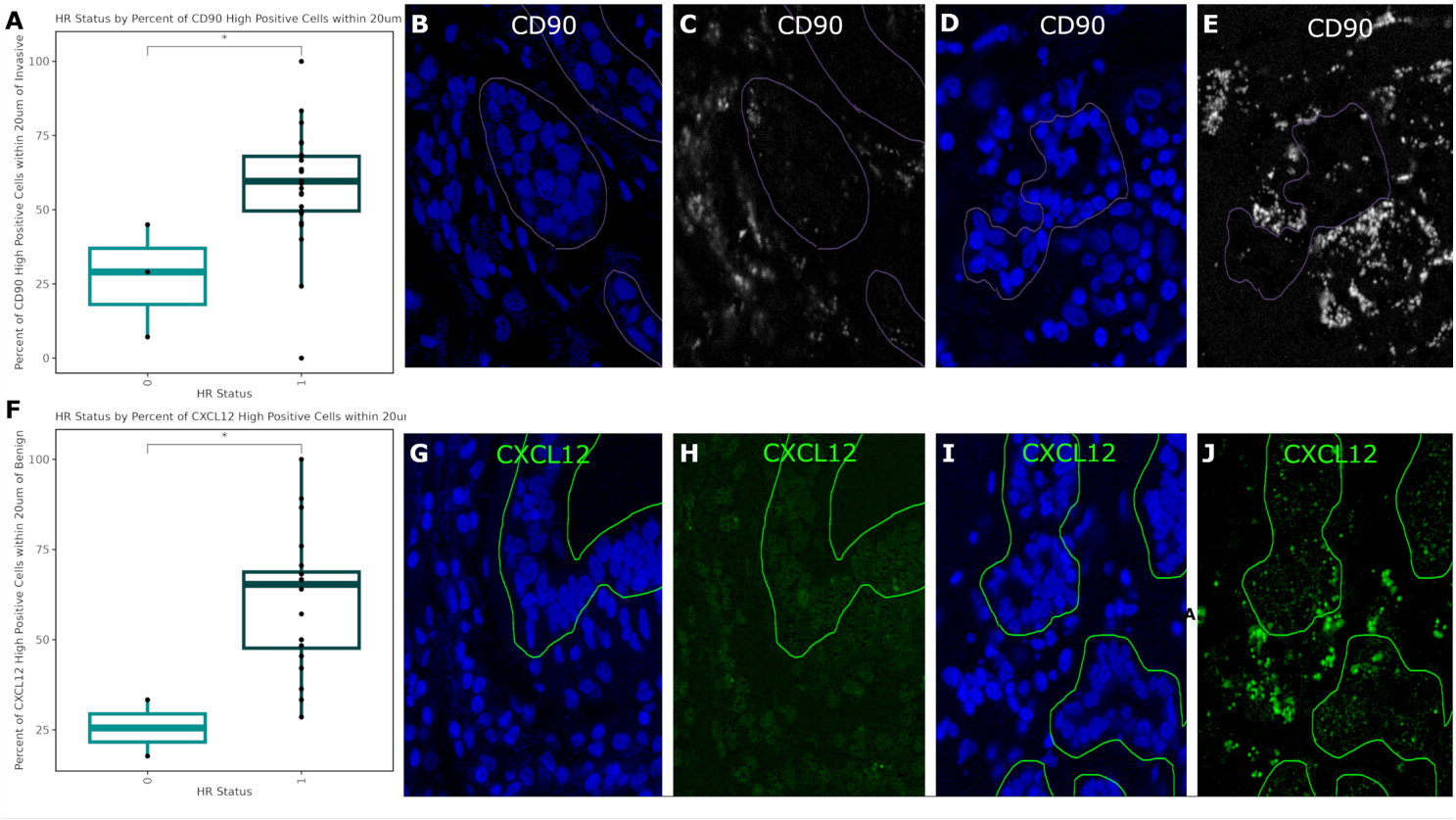


Figure 7

CD90 or CXCL12 spatial distribution associate with hormone receptor status(A-J). Boxplot showing association between hormone receptor status and percent of CD90 high positive cells in stroma of invasive lesions (A). Example images of CD90 RNA expression in HR negative case (B,C) vs HR positive case (D,E). Boxplot showing association between hormone receptor and percent of CXCL12 RNA high positive cells in stroma of benign lesions (F). Example images of CXCL12 RNA expression in HR negative case (G,H) vs HR positive case (I,J). Adjusted P<0.05 is labelled as significant. *: p<= 0.05, **: p<= 0.01, ***: p<= 0.001, ****: p<= 0.0001.

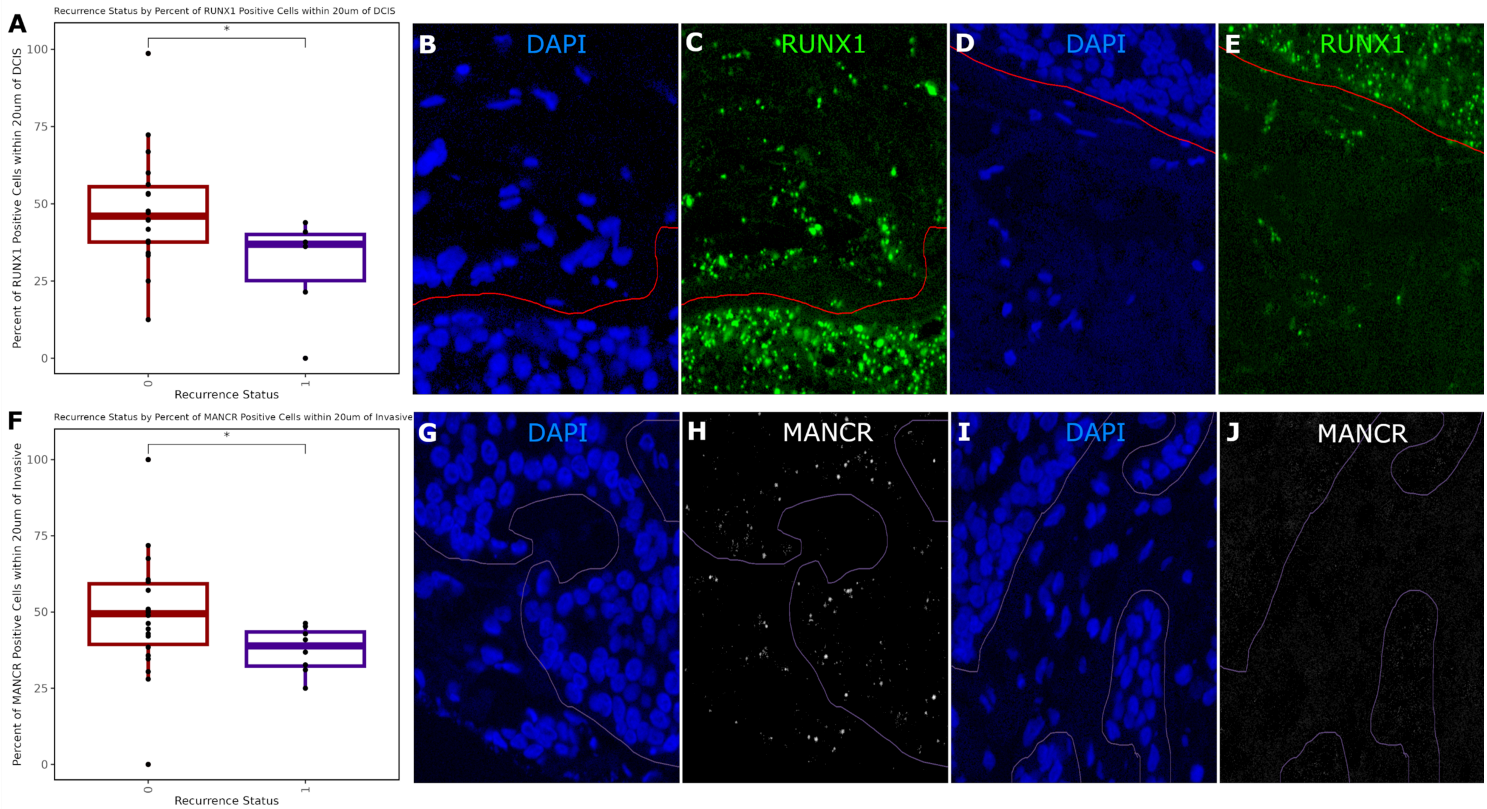


Figure 8

RNAs associated with recurrence of breast cancer (A-J). Boxplot showing association between recurrence risk and percent of MANCR positive cells in stroma of invasive lesion(A). Example images of MANCR RNA expression in non-recurring case (B,C) vs recurring case (D,E). Boxplot showing association between recurrence risk and percent of RUNX1 RNA positive cells in stroma of DCIS lesions (F). Example images of RUNX1 RNA expression in non-recurring case (G,H) vs recurring case (I,J). Adjusted P<0.05 is labelled as significant. *: p<= 0.05, **: p<= 0.01, ***: p<= 0.001, ****: p<= 0.0001.

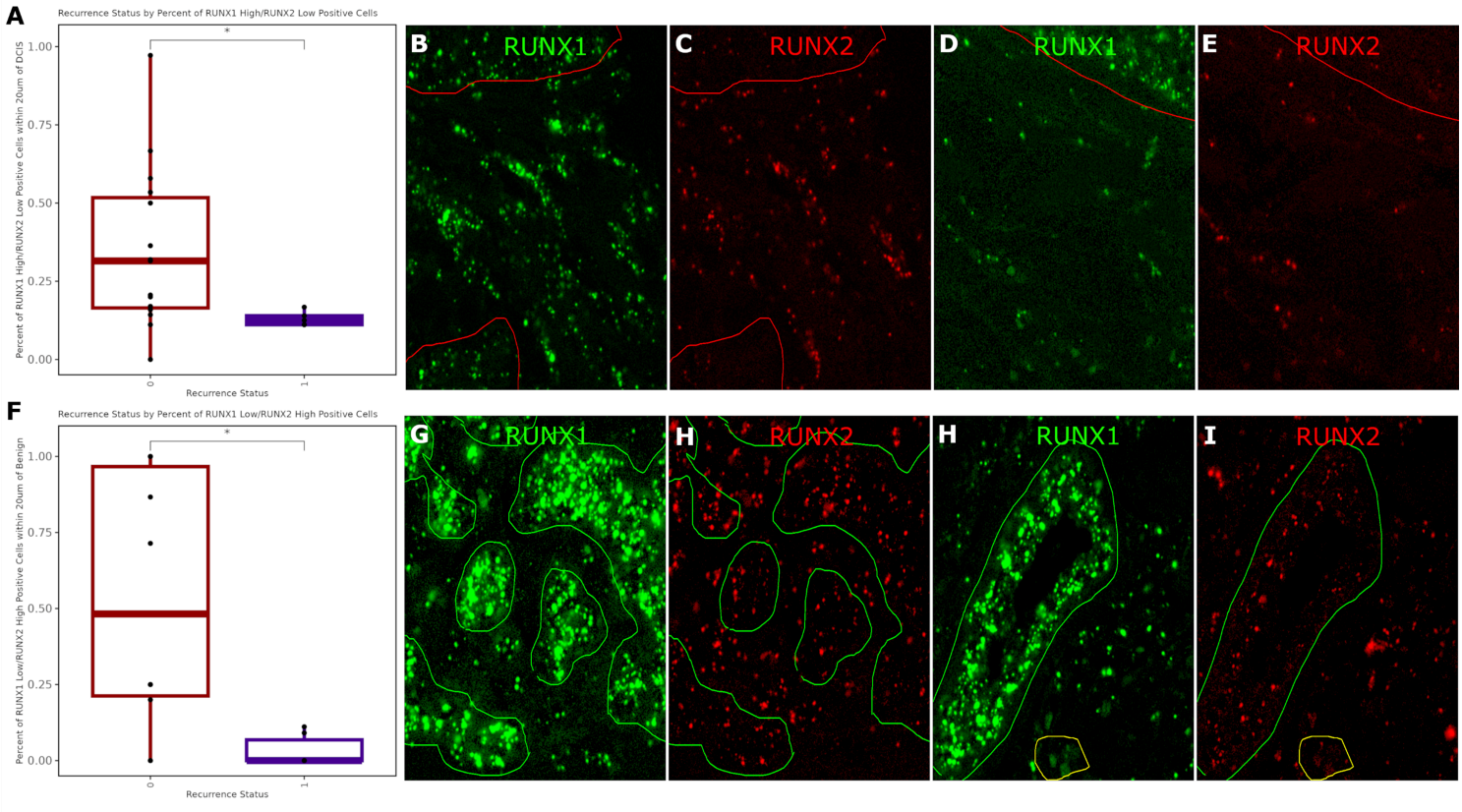


Figure 9

Spatial distribution of RUNX1/RUNX2 co-expressing cells is associated with recurrence status

(A-J) Boxplot showing association between recurrence risk and percent of high RUNX1 positive/low RUNX2 positive cells in stroma of DCIS lesions (A). Example images of high RUNX1 positive/low RUNX2 positive cells in non-recurring case (B,C) vs recurring case (D,E). Boxplot showing association between recurrence risk and percent of low RUNX1 positive/high RUNX2 positive cells in stroma of benign lesions (F). Example images of high RUNX1 positive/low RUNX2 positive cells in non-recurring case (G,H) vs recurring case (I,J). Adjusted P<0.05 is labelled as significant. *: p<= 0.05, **: p<= 0.01, ***: p<= 0.001, ****: p<= 0.0001.

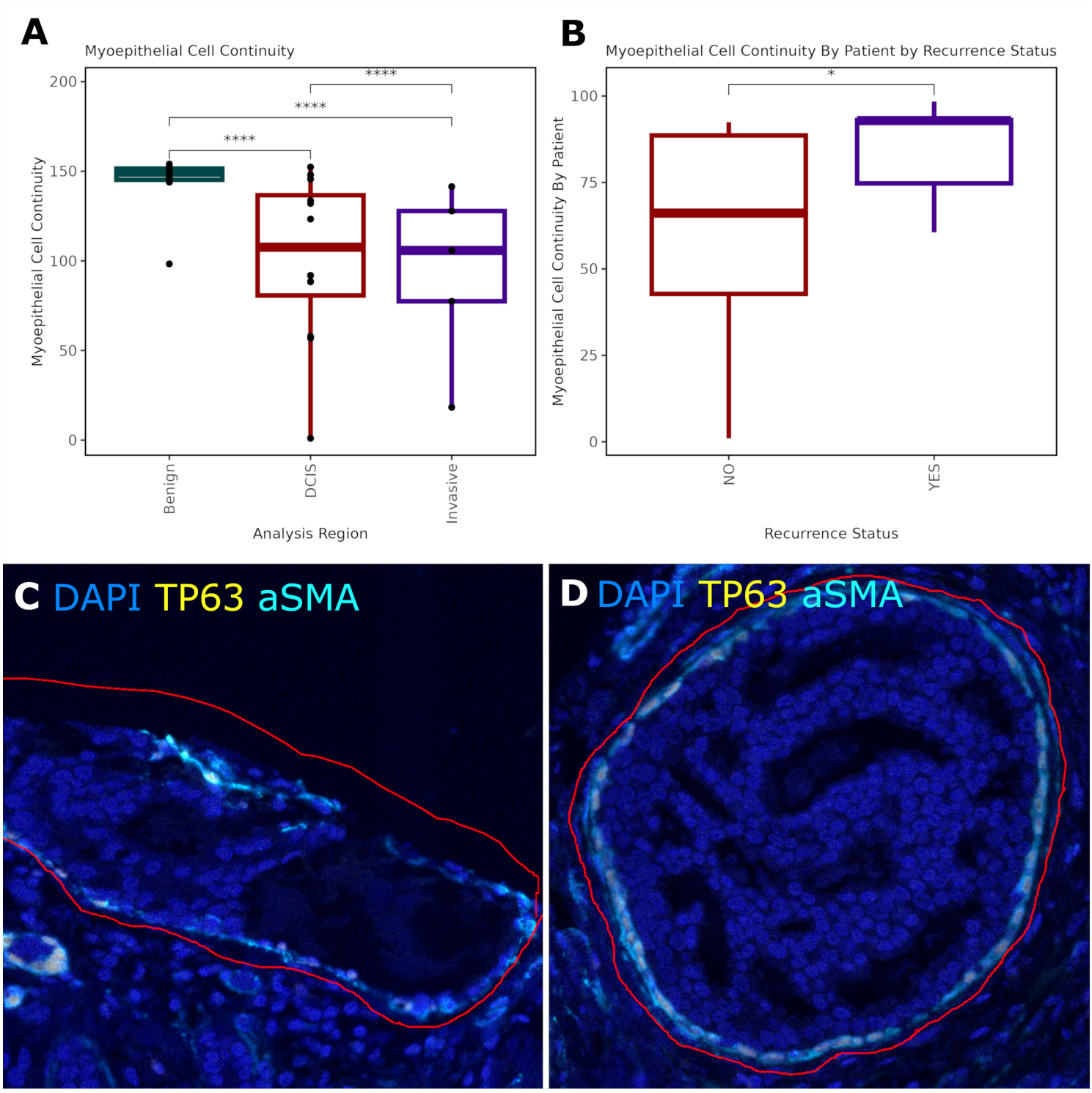


Figure 10

Higher myoepithelial continuity is associated with higher risk of recurrence (A-D). Boxplot showing association between myoepithelial continuity and lesion type(A). Boxplot of myoepithelial continuity association with recurrence risk (B). Example image of myoepithelial continuity in a DCIS region (red annotation) of a non-recurring case (C) vs a DCIS region in a recurring case (red annotation) (D). Myoepithelial cells marked using aSMA (cyan) and TP63 (yellow). Adjusted P<0.05 is labelled as significant. *: p<= 0.05, **: p<= 0.01, ***: p<= 0.001, ****: p<= 0.0001.

Supplementary Files

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

- SupplementaryFigure1HEsNew.tiff
- SupplementaryFigure2MRRZ1BFIndividualChannelsSeparateDAPIWithReferenceFrames.tif
- SupplementaryFigure3MTCZ1BFIndividualChannelsSeparateDAPIWithReferenceFrames.tif
- SupplementaryFigure4CCCZ1BFIndividualChannelsSeparateDAPIWithReferenceFrames.tif
- SupplementaryFigure5WPZ1BFIndividualChannelswithReferenceFrameswithDAPIInset.tif
- SupplementaryFigure6MANCRRepeatComparison.tiff
- SupplementaryFigure7CXCL12RepeatComparison.tiff
- SupplementaryFigure8R1R2Stroma.tiff
- SupplementaryTablesFinal.docx