Characterization and clinical evaluation of CD10+ stroma cells in the breast cancer microenvironment

Christine Desmedt¹, Samira Majjaj¹, Naima Kheddoumi¹, Sandeep K. Singhal¹, Benjamin Haibe-Kains¹², Frank El Ouriaghli¹, Carole Chaboteaux¹, Stefan Michiels¹, Françoise Lallemant¹, Fabrice Journe³, Hughes Duvillier⁴, Sherene Loi¹, John Quackenbush², Sophie Dekoninck⁵, Cédric Blanpain⁵, Laurence Lagneaux⁶, Nawal Houhou⁷, Mauro Delorenzi⁷, Denis Larsimont⁸, Martine Piccart⁹, Christos Sotiriou¹⁹

1. Breast Cancer Translational Research Laboratory JC Heuson, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium;
2. Computational Biology and Functional Genomics Laboratory, Center for Cancer Computational Biology, Dana-Farber Cancer Institute, Harvard School of Public Health;
3. Laboratory of Oncology and Experimental Surgery, Institut J Bordet, Centre des Tumeurs de l’Université Libre de Bruxelles, Brussels, Belgium;
4. Department of flow cytometry, Institut Jules Bordet; Université Libre de Bruxelles, Brussels, Belgium;
5. Interdisciplinary Research Institute (IRIBHM), Université Libre de Bruxelles (ULB), Brussels, Belgium
6. Department of Experimental Hematology, Institut Jules Bordet, Brussels, Belgium;
7. Département de formation et recherche, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland;
8. Department of Pathology, Institut Jules Bordet, Brussels, Belgium;
9. Medical oncology, Institut Jules Bordet, Brussels, Belgium;

Corresponding author:
Christine Desmedt
Institut Jules Bordet
Breast Cancer Translational Research Laboratory JC Heuson
125 Bld de Waterloo
1000 Brussels – Belgium
Tel: +32 2 541 31 07/ Fax: +32 2 538 08 58
christine.desmedt@bordet.be

Running title: Characterization and clinical relevance of CD10+ stroma cells in breast cancer

Keywords: Breast cancer, stroma, CD10, fibroblasts, mesenchymal stem cells, gene expression, prognosis, molecular subtypes.
Financial support: This project was supported by the MEDIC Foundation, Les Amis de l’Institut Bordet, the “Fonds National de la Recherche Scientifique” and the Brussels Region (Impulse Programme 2007).
STATEMENT OF CLINICAL RELEVANCE

In this study, we aimed at getting further insight on the molecular characteristics differentiating tumor-associated stroma from normal stroma, and to investigate the potential clinical relevance of the breast stroma. We specifically concentrated on CD10+ cells since the most striking changes in gene expression were reported during breast cancer progression in this particular cell population from the breast cancer microenvironment (Allinen et al. Cancer Cell 2004). We highlighted that this stroma contribution, defined by the CD10+ stroma signature, has a clinical relevance, particularly in patients with HER2+ tumors, where it is associated with worse prognosis and potentially to non-response to chemotherapy. We believe these results may help to better identify those patients for which the cancer stroma is, at least partly, responsible of their worse prognosis, and for which the stroma should thus be specifically targeted.
ABSTRACT

Purpose: There is growing evidence that interaction between stromal and tumor cells is pivotal in breast cancer progression and response to therapy. Since the pioneer work of Allinen et al. suggested that during breast cancer progression striking changes occur in CD10+ stromal cells, we aimed to better characterize this cell population and its clinical relevance.

Experimental design: We developed a CD10+ stroma gene expression signature (using HG U133 Plus 2.0) based on the comparison of CD10 cells isolated from tumoral (n=28) and normal (n=3) breast tissue. We further characterized the CD10+ cells by co-culture experiments of representative breast cancer cell lines with the different CD10+ stromal cell types (fibroblasts, myoepithelial and mesenchymal stem cells (MSC)).

We then evaluated its clinical relevance in terms of in situ to invasive progression, invasive breast cancer prognosis and prediction of efficacy of chemotherapy using publicly available datasets.

Results: This 12-gene CD10+ stroma signature includes among others genes involved in matrix remodeling (MMP11, MMP13, COL10A1) and genes related to osteoblast differentiation (periostin). The co-culture experiments demonstrated all three CD10+ cell types contribute to the CD10+ stroma signature, although MSCs have the highest CD10+ stroma signature score.

Of interest, this signature demonstrated an important role in differentiating in situ from invasive breast cancer, in prognosis of the HER2+ subpopulation of breast cancer only, and potentially in non-response to chemotherapy for those patients.
Conclusions: Our results highlight the importance of CD10+ cells in breast cancer prognosis and efficacy of chemotherapy, particularly within the HER2+ breast cancer disease.
INTRODUCTION

Tumor epithelial cells are surrounded by the tumor microenvironment which is composed by the extracellular matrix (ECM) and various other cell types such as endothelial cells, (myo)fibroblasts and leukocytes. There is growing evidence that interaction of these other cell types with tumor cells is pivotal in breast cancer progression and response to therapy.

Several studies have provided insight on the molecular characteristics differentiating tumor-associated stroma from normal stroma (1-5). Allinen et al. were the first to perform systematic profiling of different stromal cell types isolated through cell-type specific cell surface markers and magnetic beads (1). Their work suggested that during cancer progression, striking changes in gene expression occurred in almost every cell type with the most dramatic and consistent changes (other than in the malignant epithelial cells) detected in cells which were characterized by the surface marker CD10, that encompasses myoepithelial cells and myofibroblasts. Later, two research groups have conducted exploratory global gene expression analysis of the tumor microenvironment using laser capture microdissected tumor and normal breast samples (2,3). Both publications reported important expression changes in genes related to the extracellular matrix in the cancer stroma compared to the normal stroma. Additionally, the study by Ma et al. also compared the epithelium and stromal compartment of ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC) lesions: they observed no significant differences at the transcriptome level between the malignant epithelium of DCIS and IDC, whereas they found a significant number of genes differentiating DCIS and IDC lesions in the stromal compartment (2).
Very recently, Bauer et al. analyzed gene expression profiles from paired cancer associated and normal fibroblasts from 6 breast tumor specimen isolated through a method similar to the one developed by Allinen et al. (1) followed by short-term cell culture (4). They identified a list of 31 genes differentiating cancer associated from normal fibroblasts. Based on these genes and consistent with previous works, they suggest that the cancer associated fibroblasts contribute to cancer growth and progression by enhancing ECM production, promoting stromal-epithelial paracrine signaling and altering steroid hormone metabolism.

However the analysis of the clinical relevance of stroma-related molecular signatures has been relatively limited to date. Finak and colleagues identified a stromal signature (stroma-derived prognostic predictor, further referred to as SDPP) by comparing gene expression profiles of tumor stroma and matched normal stroma isolated by laser capture microdissection (LCM) from 53 primary breast tumors (5). SDPP predicted outcome in several published whole tumor derived gene expression datasets, and especially in the HER2+ breast cancer subgroup. In parallel, using a prototype-based clustering approach, we developed a stroma-metagene (further referred to as PLAU) as a list of genes specifically correlated to PLAU, which was associated with prognosis only in the HER2+ breast cancer subgroup (6). Using a similar approach, Farmer et al. defined their stromal metagene (further referred to as DCN) as a list of genes correlated to decorin; DCN was able to predict chemoresistance in patients treated with preoperative chemotherapy (7).

In this paper, we aimed to gain further knowledge about the breast cancer stroma cellular compartment which according to Allinen et al. (1) undergoes the most extensive
changes when compared to normal breast tissue. We therefore compared CD10+ cells issued from tumor and normal breast tissue based on an approach similar to the one of Allinen et al. but differing from that used by Bauer et al. (4) in the sense that we did not submit the isolated cells to short-term culture. We further aimed to characterize the cell-type specific contribution of the stroma by performing co-culture experiments and to investigate the three-fold clinical relevance of our findings, with regard to the DCIS/invasive transition, breast cancer prognosis and prediction of the efficacy of preoperative chemotherapy. From the results, it appears that CD10+ cells may be particularly clinically relevant in HER2+ breast cancer.
MATERIAL AND METHODS

Patients

Patients with invasive ductal breast carcinoma were recruited for the study at the Bordet Institute. Breast tissues were obtained at the time of the surgery. Since Finak et al. demonstrated that stroma surrounding histologically normal ducts from tumor specimen and stroma isolated from reduction mammoplasties were highly similar at the gene expression level (8); we used histologically normal tissue distant from at least 2 cm from the tumor as normal breast tissue.

The study was approved by the ethics committee of the Bordet Institute. All patients included in this study signed the consent form.

Isolation of CD10+ cells

To isolate CD10+ cells from breast tissues, we established a procedure similar to the one described by Allinen et al. (1) and schematically illustrated in Supplementary Figure 1. Briefly, carcinoma and normal samples were mechanically dissociated using a scalpel, after confirmation of the pathologist (DL) on the tumor and normal nature of the sample. Fragments were incubated in 12-well culture dish with a mixture of collagenase, type 4 (LS004189, Worthington Biochemical Corporation, Lakewood, NJ USA) in x-vivo 20™ Serum-free Medium (BE04-448Q, Cambrex, Walkersville, MD USA) in a 37°C incubator with 5% CO2 with constant agitation for 1-2 hours, depending of the size of the sample. Following dissociation, the digestion product were filtered through a BD Falcon™ Cell Strainers (352340, BD Biosciences, San Jose, CA USA) using piston syringe and washed with x-vivo medium. The Epcam+ cells were first isolated using the
Human Epcam selection kit, followed by a depletion of the CD45+ cells using the Human CD45 depletion kit from the same company. The CD10+ cells were further isolated from the unicellular suspension using human CD10+ selection kit. All selection kits were from STEMCELL technologies and were used according to the manufacturer’s instructions. The purity of the population was checked by flow cytometry (Antibodies CD326(Epcam) FITC from Milteny Biotec, APC anti human CD45 from Pharmingen and PE Mouse anti-human CD10 from Pharmingen).

**Cell cultures**

The 7 breast cancer cell lines MCF-7, T47D, LY2 (ER+/HER2-), MDA-MB-231, (ER-/HER2-), SKBR3, MDA-MB-361, BT474 (HER2+) and the CD10+ cell lines HS57BST (myoepithelial cells) and CRL2564 (fibroblasts) were obtained from the ATCC. They were cultured according to the media indicated by ATCC, except the HS57BST which we did not stimulate with EGF (Supplementary Table 1). Mesenchymal stem cells (MSCs) were isolated from bone marrow from a single female healthy donor and cultured as reported previously (9).

For the co-culture experiments (each possible combination of a breast cancer cell line with a specific type of CD10+ cells), the CD10+ were first plated in petri dishes until sub-confluence. At that time the epithelial cells were added on top of the CD10+ cells. After 4 days of co-culture, the CD10+ and epithelial cells were isolated using flow cytometry. The CD10+ cells used as control were not co-cultured with breast cancer cells. All co-culture experiments were done in duplicate. FACS sorting was then performed with a 3 lasers FACS Aria I flow cytometer (Becton Dickinson) at high pressure (70 psi)
using the same anti-CD10 and anti-Epcam antibodies as above. Compensation adjustments were performed with co-cultures stained with single color. Double staining on each cell culture was used as a positive control for the gating strategy. Cells were sorted in 1.5ml collection tubes containing 350µl of lysis buffer (Absolutely RNA microprep kit, Agilent).

**Gene expression data**

We used the Human Genome U133-2.0 plus GeneChips to generate the gene expression data. More information about the experiments can be found in Supplementary methods.

Gene expression data were normalized by using the quantile normalization method of affy package version 1.28.0 (10) of the Bioconductor software project (http://www.bioconductor.org/). Subsequently data were log2 transformed. Microarray data are deposited in the Gene Expression Omnibus database under accession number GSE27120.

**The CD10+ signature**

To select the genes for the CD10+ signature, we carried out a class comparison of the tumor-associated CD10+ cells (n=28) with CD10+ cells from normal breast tissues (n=3) using the 10% most variant probesets across the entire series (n=5467 probesets). The t-test was used to compare the mean values of probes across groups. To overcome the problem due to multiple testing, false discovery rate (FDR) (11) approach was used to select the top probesets.
We first selected those genes with an absolute fold change (FC) \( \geq 1.5 \) and false discovery rate (FDR) < 0.05 for describing the genes differentially expressed and an absolute FC\( >4 \) and FDR\( <0.01 \) to generate the CD10+ signature respectively.

Using those genes, the CD10+ signature was computed for each sample as

\[
s = \frac{\sum w_ix_i}{\sum |w_i|}
\]

where \( x_i \) is the expression of a gene included in the set of genes of interest and \( w_i \) is either +1 or -1 depending on the increased or decreased expression in tumor-associated CD10 cells vs CD10 cells from normal breast tissue respectively. When different probe-sets represented the same gene, only the most variant one was included in the signature. Also, when a probe-set did not match a Entrez GeneID, it was not included in the signature.

The F-test was used to compare the variances of genes of the CD10+ signature between tumor-associated and normal CD10+ cells.

The list of genes in the CD10+ stromal signature and the corresponding annotations have been submitted to GeneSigDB (12).

**Analysis of gene expression data**

For the cell culture data, the Kruskal-Wallis test was applied to identify the differently expressed probes 1) between cell lines and 2) between the subtypes of each cell line. The Wilcoxon test was used to compute the differences between the different cell lines. It was not applied to the analyses within the different cell lines due to the too small number of samples for these comparisons. To overcome the problem due to multiple testing, false discovery rate (FDR) (11) approach was used for the different
CD10+ stroma signature genes, except again for comparisons within the different cell lines due to the too small number of samples. The FC was calculated by the average difference between those groups.

To assess whether stromal genes can also be expressed in breast cancer epithelial cells, we examined the data from the invasive breast cancer cell lines from Neve et al. (13). We first classified the cell lines according to the main 3 molecular subgroups of breast cancer: ER-/HER-, HER2+ and ER+/HER2- based on the ER and HER2 mRNA expression levels using the SCMGENE classifier as implanted in the genefu R package.

To evaluate the implication of the stromal genes in the transition from DCIS to invasive breast cancer, we used publicly available gene expression data from matched DCIS/invasive cancer (14).

To investigate the prognostic value of the CD10+ stroma signature, gene expression datasets were retrieved from public databases or authors’ websites. We used normalized data (log2 intensity in single-channel platforms or log2 ratio in dual-channel platforms) as published by the original studies. Hybridization probes were mapped to Entrez GeneID as described (15), using RefSeq and Entrez database version 2007.01.21. When multiple probes were mapped to the same GeneID, the one with the highest variance in a particular dataset was selected. Seven public breast cancer microarray datasets were used (16-24, Supplementary Table 2). Since we wanted to evaluate the pure prognostic value of the signature, we only considered patients who did not receive any systemic anti-cancer treatment. Distant metastasis free survival (DMFS) was used as survival endpoint. However, when DMFS was not available, we used relapse free
survival (RFS). We censored the survival data at 10 years in order to have comparable follow-up across the different studies as described (6,16,25). The hazard ratio (HR) was estimated for each dataset separately and combined with inverse variance-weighted method (26). The Kaplan-Meier product-limit estimator was used to display time to event curves. The tertiles for the signature scores were computed within each breast cancer molecular subtypes separately, in order to estimate the signature's prognostic value for each subtype.

To investigate the prognostic value of the CD10+ stroma signature when treating patients with anthracycline-based chemotherapy, we used gene expression data treated in the TOP trial, NCT00162812 (27). In this trial, patients with estrogen receptor (ER)-negative tumors were treated with anthracycline (epirubicin) monotherapy, allowing investigating anthracycline-specific predictors. The area under the curve (AUC) and its confidence interval was estimated through the concordance index (28), and its one-sided p-value was determined by the Wilcoxon Rank Sum test.

All the analysis were performed using R software version 2.9.2 (www.r-project.org)
RESULTS

Development of the CD10+ stroma signature

Using an approach similar to that of Allinen et al. (1), we were able to isolate tumor-associated CD10+ cells (n=28) as well as CD10+ samples from normal breast tissues (n=3). The characteristics of the tumors are illustrated in Table 1.

A comparison between gene expression profiles of tumor-associated CD10+ cells and CD10+ cells isolated from normal breast tissue (FDR<0.05 and |FC|>1.5, Supplementary Table 3) revealed 242 probesets that were significantly differentially expressed including genes involved in matrix remodeling such as matrix metalloproteinases (MMP3, MMP11, MMP13), fibronectin, collagens (COL1A2, COL8A1, COL10A1 and COL11A1). This list also included genes related to skeletal development such as periostin, matrilin-3, CHRDL1, a BMP-4 antagonist and the integrin-binding sialoprotein (see Supplementary Table 4). We computed the gene overlap between this list and the various published signatures (see Supplementary Table 5). All signatures showed some gene overlap except SDPP. COL10A1, the gene which in our dataset was the most differentially expressed between tumor and normal CD10+ cells was present in all signatures, except PLAU and SDPP.

To develop a signature we selected from the above reported comparison the most differentially expressed genes, with |FC| ≥ 4 and FDR < 0.01 (Supplementary Table 6). Twelve genes were selected based on these criteria. The reason we developed a signature based on a small number of genes was to facilitate its potential future translation into clinic by evaluating a small number of genes using an RT-PCR assay for instance. Figure 1 illustrates the expression of these genes according to the molecular subtype of the
tumors from which the CD10+ cells were isolated. This figure shows that the vast majority of CD10+ cells isolated from tumors show a clear differential expression pattern, with PCOLCE2, GPX3 and ADH1B being over-expressed in normal-associated CD10+ cells compared to the tumor-associated CD10+ cells and the remaining genes being over-expressed in the tumor-associated CD10+ cells. For some tumors, the expression pattern of CD10+ cells does not seem to be very different from the normal CD10+ cells, as it is clearly the case for sample IJB_CD10_27, suggesting that in some tumors these cells might still exhibit the same expression pattern as in normal breast tissue.

Since Bauer et al. (4) surprisingly reported a higher variability of global gene expression in normal fibroblasts compared to fibroblasts present in the tumor; we sought to compare the variance of the genes from our signature between the tumor-associated CD10+ cells and the CD10+ cells from normal breast tissues. We found that in contrast to Bauer et al.’s findings, the variances for the genes from our signature are nominally higher for the tumor-associated CD10+ cells than for the normal CD10+ cells, although the p-value generated by the F-test was only significant for half of the genes (Supplementary Table 7). This confirms that the variability of these specific genes is higher in tumor-associated CD10+ compared to normal CD10+ cells.

**Characterization of the CD10+ stroma signature**

In order to investigate whether our CD10+ stroma signature genes were exclusively expressed in stromal cells, we examined their expression levels in the breast cancer cell line dataset made available by Neve et al. (13). We observed that for half of
the genes included in the signature (ADH1B, COL10A1, GRP, HSD17B6, MATN3 and MMP13) the expression level was below or around the 25th percentile of the expression values of the array and did not show any variation across the different cell lines, suggesting that high expression values of those genes are most probably not derived from the epithelial compartment. However, some exceptions were observed. For example, GPX3 expression levels were superior to the median of the arrays for almost half of the cell lines and we observed significant variation across the cell lines. The expression levels of MMP11 were above the 25th percentile, however no strong variation was observed across the cell lines. PCOLCE2 showed low expression in all cell lines except in the cell lines characterized by the absence of ER and HER2 expression. POSTN showed low expression in all cell lines except in HCC38, HS578T and BT549, cell lines also characterized by the negativity of the hormone receptors. Detailed results are available as Supplementary Figure 2.

Several studies have revealed that some CD10+ stroma cells can be derived from bone marrow-derived progenitor cells, such as mesenchymal stem cells (MSCs), which can be mobilized into the circulation and incorporated into the tumor microenvironment. In order to analyze the contribution of the different CD10+ cell types to the CD10+ signature, we performed co-culture analyses of representative breast cancer cell lines (see methods) with different stroma cell types expressing CD10+ ([myo]fibroblast and myoepithelial cell lines, as well as MSCs obtained from one healthy donor). We then generated gene expression data of the different CD10+ stromal components which were and were not (controls) co-cultured with the different breast cancer cell lines. We first
compared the CD10+ stroma signature scores according to the different CD10+ cell types which were or were not co-cultured with the different breast cancer cell lines (first panel of Figure 2). We observed significantly different signature scores in the three different cell types (p-values: fibroblasts/MSC=0.002, MSC/myo-epithelial=8.10^{-9}, fibroblast/myo-epithelial=3.10^{-8}), with the highest scores observed in the MSCs followed by the fibroblasts. However, within each CD10+ cell type, we also observed differences in the CD10+ stroma signature scores between the different co-cultured and the control cells (Figure 2).

Second, we compared the expression levels of the different CD10+ stroma signature genes according to the different CD10+ cell types. We observed significant changes in expression levels of 11/12 signature genes, with the myoepithelial cells differing most from the fibroblasts and MSCs (Supplementary Table 8). For 7/12 genes, expression levels were significantly different between the MSCs and the fibroblasts. Figure 2 and Supplementary Figure 3 illustrate the expression values of the CD10+ stroma signature genes for the control and different co-culture conditions for each CD10+ cell type. Globally, COL10A1 and MMP13 clearly have higher expression values in MSCs than in the two other cell types, with the exception for COL10A1 of MSCs co-cultured with ER-HER2- cell lines. On the contrary, GJB2 and HSD17B6 display higher expression values in myoepithelial cells compared to the other cell types, whereas high expression levels of ADH1B and POSTN can clearly be attributed to the fibroblasts and MSCs.

Third, we analyzed the differences in gene expression of the 12 genes from the signature within each CD10+ subtype, to see whether the expression of those genes
varied between the controls, the co-cultures with HER2+, the co-cultures with the ER-/HER2- and the co-cultures with the ER+/HER2- breast cancer cell lines. There were significant differences observed for some genes (Figure 2, Supplementary Figure 3 and Supplementary Tables 9, 10 and 11 for the results regarding the fibroblasts, MSCs and myoepithelial cells respectively). Fibroblasts co-cultured with ER-/HER2- breast cancer cells expressed higher levels of MMP13 compared to the other fibroblasts. In MSCs, we observed lower levels of HSD17B6 in the co-cultures compared to the control MSCs. Also, higher levels of POSTN were observed in MSCs co-cultured with HER2+ and ER+/HER2- breast cancer cells compared to the remaining MSCs, with the highest levels observed when co-cultured with the HER2+ cells. In the myoepithelial cells, we observed higher levels of GJB2 and MATN3 in the co-cultured cells compared to the control myoepithelial cells. Also, we observed a slight increase in HSD17B6 expression when co-cultured with ER+ breast cancer cells compared to the remaining myoepithelial cells.

Of note, no significant differences in cell proliferation were noticed in the co-culture experiments, except that the ER+/HER2- breast cancer cell lines were more proliferative after co-culture with fibroblasts than were the other breast cancer cell lines co-cultured with fibroblasts (Supplementary Figure 4).

Clinical relevance of the CD10+ stroma signature

Since it has repeatedly been suggested that the progression of DCIS to invasive breast cancer could be strongly dependent upon changes in the stroma (reviewed in 29), we aimed at investigating whether our CD10+ stroma signature was able to differentiate
DCIS samples from their associated invasive counterparts. To this end, we used publicly available gene expression data from DCIS and their associated invasive counterparts isolated by laser capture microdissection (LCM) (14).

Interestingly, as illustrated in Figure 3, our CD10+ stroma signature was able to differentiate DCIS from invasive cancer. This figure also illustrates the different expression levels of the signature genes in DCIS and IDC.

Recently, three studies have investigated the clinical relevance of global gene expression signatures describing the stromal compartment in invasive breast cancer (5-7) using gene expression data generated from whole breast tumor tissue comprising both tumor epithelium and stroma. To assess whether our CD10+ stromal signature could predict outcome from such data, we analyzed available microarray data generated from whole tumors of 1,140 patients who did not receive any systemic treatment (16-24). Elevated values of our CD10+ stromal signature were associated with poor clinical outcome (HR=2.17 (CI 1.23-3.85), p=0.008) in the HER2+ subpopulation only (Supplementary Table 12). The Kaplan-Meier curves are illustrated in Figure 4. In agreement with previous reports, SDPP, PLAU and DCN also carried prognostic information in this same HER2+ population (Supplementary Table 12). The CD10+ stroma signature was only correlated with PLAU and DCN but not with SDPP (Supplementary Table 13) suggesting that these signatures represent different biological signals.

In order to test whether the clinical relevance of the CD10+ stroma signature was specific for the HER2+ subgroup, we carried out interaction tests. The CD10+ signature was
significantly different in the HER2+ subgroup compared to the ER+/HER2- subgroup (p=0.028) and showed a trend towards significance compared to the ER-/HER2- subgroup (p=0.051).

We further wanted to evaluate whether our stroma signature was correlated with response/resistance to chemotherapy or whether it was just associated with the natural history of the disease. Collagens, which are part of the transport barrier imposed by the ECM and also part of our signature (COL10A1 and COL11A1), have been demonstrated to be, at least partly, responsible for decreased penetration of therapeutic agents in the tumor (30). We therefore investigated the association of our CD10+ signature with the pathological complete response (pCR) status in the well-defined prospective TOP trial (27). Patients in this dataset had estrogen-receptor negative tumors and were treated with pre-operative epirubicin only. Similarly to the above results, we observed that the CD10+ stroma signature was only significantly associated with lack of pCR in patients with HER2+ tumors (Figure 5). Consistent with the fact that our signature was correlated with PLAU and DCN, these two signatures also were associated with pCR in this subgroup of patients, but not SDPP (Supplementary Table 14). However, these results should be interpreted with caution given the small number of patients in this HER2+ subgroup (n=30).
DISCUSSION

In this paper, we compared gene expression profiles from tumor-associated CD10+ cells with CD10+ cells from normal breast tissue. Our hypothesis was that by selecting only CD10+ cells we investigated a more uniform stromal compartment at the cellular level compared to previous publications which used LCM to dissect the stroma fraction and which thus considered the stroma as a uniform entity (2,3,5). Of note, the recent publication of Bauer et al. (4) used a similar approach as the one we used, except that they subjected the isolated cells to short term culture, with the potential disadvantage that it could mask some of the originally existing differences in gene expression, and that they did not investigate their potential clinical relevance.

As in previous descriptive studies, we found many genes associated with the remodeling of the extracellular matrix, such as matrix metalloproteases, collagens and fibronectin. Interestingly, we also observed a certain number of genes linked with skeletal development such as periostin (POSTN). It has been shown that POSTN is undetectable in normal breast tissues, whereas elevated levels are observed in breast cancer tumor samples. It has also been demonstrated that serum POSTN levels are elevated in breast cancer patients with bone metastases from breast cancer (reviewed in 31). Another recent report showed, using a well-defined animal model, that bone metastases from breast cancer are associated with a marked overexpression of this stroma-derived POSTN (32).

From the comparison between tumor and normal CD10+ cells, we selected those genes with the highest FC and lowest corrected p-value to develop the CD10+ stroma signature. To assess whether the genes included in this signature could also be expressed
by breast cancer epithelial cells, we examined the cell line data from Neve et al. (13). We observed that half of the genes (ADH1B, COL10A1, GRP, HSD17B6, MATN3 and MMP13) showed a homogeneous low expression in the different breast cancer cell lines, suggesting that these genes are not expressed in cultures of breast cancer cells. Since CD10+ cells encompass different cell types, (myo)fibroblasts, myoepithelial cells and MSCs, we performed cell line co-culture experiments to understand the contribution of each cell type. Although the cell culture results should be taken with caution since these co-culture experiments are just an artificial representation of these CD10+/breast cancer epithelial cells interaction, three main messages could be derived from the cell culture analyses. First, by comparing the CD10+ stroma signature scores between the three CD10+ cell types, we observed that the highest signature scores were observed in MSCs followed by the fibroblasts. We further also observed changes in the signature score within each cell type following co-culture with breast cancer cells. Second, by comparing the expression levels of the different CD10+ stroma signature genes between the three CD10+ cell types, it clearly appears that all three cell types contribute to the CD10+ stroma signature. Indeed, some genes are more highly expressed both in fibroblasts and MSCs (POSTN and ADH1B), whereas other genes are more expressed in myoepithelial cells only (such as GJB2, PCOLCE2 and HSD17B6) or MSCs only (MMP13 and COL10A1). Third, by investigating the expression levels of the different CD10+ stroma signature genes within each CD10+ cell type, we observed that the expression of some genes clearly differed in some co-culture conditions compared to controls: MMP13 for the fibroblasts, HSD17B6 and POSTN for the MSCs, and GJB2, MATN3 and HSD17B6 for the myoepithelial cells.
We further aimed to investigate the clinical relevance of our CD10+ stroma signature. We first demonstrated that our signature and the majority of its genes were able to discriminate DCIS from associated invasive samples. This finding confirms previous findings that suggested that progression from in situ to invasive breast cancer may not be due to the changes in the epithelial cells but dependent upon the tumor microenvironment (33-35).

Using gene expression data generated from whole tumor tissue, we further demonstrated that our CD10+ stromal signature carried prognostic value in the HER2+ breast cancer population of untreated breast cancer patients only. We observed similar results for the previously developed stroma signatures, the PLAU module and SDPP (6,7). A moderate but significant correlation was observed between our CD10+ stroma signature and the PLAU module, although these were generated using a completely different approach (CD10+cells isolation versus in silico development using whole tumor tissue gene expression data). On the contrary, no correlation was observed between our signature and SDPP although both were developed from the comparison of tumor versus normal stroma. One possible reason for this is that the authors from SDPP isolated the stroma using LCM and therefore most probably considered additional cell types. Why would this stroma signature be particularly relevant in HER2+ breast cancer? In these HER2+ cancers, the activity of HER2 can be due to its transactivation. Cabioglu et al recently suggested a novel mechanism of HER2 transactivation through CXCL12/SDF1 stimulation of the CXC4R4 chemokine receptor, and this only in breast cancer cells that express high levels of HER2 (36). In our cell line data, we observed that MSCs and
especially fibroblasts co-cultured with HER2+ breast cancer cells express very high levels of CXCL12 (data not shown). Alternatively, POSTN could be playing a role in breast cancer progression of HER2+ breast cancer since increased expression has been correlated with a worse prognosis (32) and since we observed the highest levels of POSTN in our cell culture experiments in MSCs after co-culture with HER2+ breast cancer cells. However, the exact tumor-epithelium interactions explaining the specific clinical relevance of the stroma in HER2+ breast cancer still need to be investigated and will be the focus of future research.

We recognize the limitation of our study, i.e. the under-representation of the different molecular subgroups of breast cancer in the tumor population used for CD10+ isolation. However, to our knowledge, this is the first study which isolates a specific cellular stromal compartment, in contrast to the majority of the previous studies which used LCM, analysed extensively its clinical value and further characterized it by co-culture experiments.

To conclude, we believe that the comprehensive characterization of this well-defined stroma compartment reinforces the importance of tumor epithelial-stroma cell interactions in tumorigenesis. We highlighted that this stroma contribution, defined by the CD10+ stroma signature, seems particularly important in HER2+ breast tumors. We further demonstrated an important role of this signature in differentiating in situ and invasive cancers, in prognosis of HER2+ patients and potentially in non-response to chemotherapy for those patients. Altogether, we believe these results may help to
better identify those patients for which the cancer stroma is, at least partly, responsible of their worse prognosis, and for which the stroma should thus be specifically targeted.
REFERENCES


FIGURES

**Figure 1:** Heatmap representing the expression of the genes included in the CD10+ stroma signature in the normal and tumor-associated CD10+ cells (red represents high expression values and green low expression values). The molecular subtypes of the tumor samples are illustrated for each tumor sample. ER (estrogen receptor) and HER2 were both assessed by IHC. ER+/HER2- (low) were ER+/HER2- tumors of grade 1 or 2 and ER+/HER2- (high) were ER+/HER2- tumors of grade 3.

**Figure 2:** Boxplots representing the CD10+ stroma signature and the CD10+ signature genes in the different CD10+ cell types (co-cultures with different subtypes of breast cancer cell lines and controls are represented separately.

**Figure 3:** Boxplots representing the CD10+ stroma signature score and the individual gene expression values between the invasive ductal carcinomas (IDC) and DCIS from Schuetz et al. originally profiled on the Affymetrix HG U133Plus2.0. * refers to p-values< 0.01 and ** to p-values <0.001.

**Figure 4:** Survival curves for untreated patients stratified by the CD10+ stroma signature: all patients (A), ER-/HER2- (B), HER2+ (C), and ER+/HER2- (D).
Figure 5: ROC analysis of the ability of the CD10+ stroma signature and the A-Score to discriminate patients with pCR from patients with residual disease in the TOP trial: all ER- patients (A), ER-/HER2- (B) and ER-/HER2+ (C).

TABLES

Table 1: Characteristics of the tumors used to develop the CD10+ stroma signature.
### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Nr of patients</th>
<th>Percentage of patients (n=28)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 50</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>&gt; 50</td>
<td>21</td>
<td>75</td>
</tr>
<tr>
<td><strong>Tumor size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 2 cm</td>
<td>11</td>
<td>39</td>
</tr>
<tr>
<td>&gt; 2 cm</td>
<td>16</td>
<td>57</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><strong>Nodal status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>14</td>
<td>50</td>
</tr>
<tr>
<td>N1-3</td>
<td>14</td>
<td>50</td>
</tr>
<tr>
<td><strong>Histological grade</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>G2</td>
<td>11</td>
<td>39</td>
</tr>
<tr>
<td>G3</td>
<td>14</td>
<td>50</td>
</tr>
<tr>
<td><strong>ER status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Positive</td>
<td>24</td>
<td>86</td>
</tr>
<tr>
<td><strong>PgR status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>12</td>
<td>43</td>
</tr>
<tr>
<td>Positive</td>
<td>16</td>
<td>57</td>
</tr>
<tr>
<td><strong>HER2 status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative (IHC 1+ or 2+ or not amplified by FISH)</td>
<td>26</td>
<td>93</td>
</tr>
<tr>
<td>Positive (3+ or amplified)</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>
FIGURE 1

Author Manuscript Published OnlineFirst on January 10, 2012; DOI: 10.1158/1078-0432.CCR-11-0383

Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.
FIGURE 2

Fibroblast Myoepithelial MSC co-culture with ER−/HER2− cell lines co-culture with HER2+ cell lines co-culture with ER+/HER2− cell lines controls

CD10+ stroma signature

COL10A1

MMP13

GJB2

HSD17B6

POSTN

Log2 gene expression
Gene expression (log2)

CD10+ stroma sig**.

COL10A1**

GJB2*

MMP11**

MMP13*

HSD17B6*

POSTN*

PPAPDC1A**

ADH1B

GPX3

PCOLCE2

GRP

MATN3

IDC

DCIS

FIGURE 3
FIGURE 4

A)

All samples

HR=1.2, 95%CI[0.91,1.5], p=2.2E−01

No. At Risk
Low 355 336 302 281 272 260 236 209 173 139 102
Intermediate 369 354 331 300 276 248 223 206 167 138 118
High 386 370 339 311 284 263 235 201 177 153 130

B)

ER-/HER2- samples

HR=1.1, 95%CI[0.64,2], p=6.5E−01

No. At Risk
Low 68 63 55 49 47 45 45 40 30 28 15
Intermediate 75 68 59 55 54 51 45 43 33 22 18
High 81 76 64 59 56 51 49 45 35 28 21

C)

HER2+ samples

HR=2.5, 95%CI[1.3,5], p=8.7E−03

No. At Risk
Low 52 50 46 43 40 39 35 31 25 21 16
Intermediate 52 50 46 43 38 34 32 26 23 20 18
High 54 49 41 33 29 27 23 20 19 16 13

D)

ER+/HER2- samples

HR=1.1, 95%CI[0.82,1.6], p=4.3E−01

No. At Risk
Low 238 232 213 200 193 181 160 144 129 107 84
Intermediate 239 220 197 180 163 145 129 107 88 77
High 250 242 219 199 185 164 141 122 104 88
FIGURE 5

A
All ER- samples

B
All ER-/HER2- samples

C
All ER-/HER2+ samples

AUC = 0.53 (95% CI: 0.40-0.66)
p = 0.35

AUC = 0.48 (95% CI: 0.30-0.67)
p = 0.43

AUC = 0.73 (95% CI: 0.59-0.87)
p = 0.046
Clinical Cancer Research

Characterization and clinical evaluation of CD10+ stroma cells in the breast cancer microenvironment

Christine Desmedt, Samira Majjaj, Naima Kheddoumi, et al.

Clin Cancer Res  Published OnlineFirst January 10, 2012.

Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-11-0383

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2012/01/10/1078-0432.CCR-11-0383.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.