Tumor Intrinsic Subtype Is Reflected in Cancer-Adjacent Tissue

Patricia Casbas-Hernandez1, Xuezheng Sun2, Erick Roman-Perez2, Monica D’Arcy2, Rupninder Sandhu3, Asahi Hishida3, Kirk K. McNaughton4, Xiaohong R. Yang5, Liza Makowski3,6, Mark E. Sherman5, Jonine D. Figueroa5, and Melissa A. Troester1,2,3

Abstract

Introduction: Overall survival of early-stage breast cancer patients is similar for those who undergo breast-conserving therapy (BCT) and mastectomy; however, 10% to 15% of women undergoing BCT suffer ipsilateral breast tumor recurrence. The risk of recurrence may vary with breast cancer subtype. Understanding the gene expression of the cancer-adjacent tissue and the stromal response to specific tumor subtypes is important for developing clinical strategies to reduce recurrence risk.

Methods: We utilized two independent datasets to study gene expression in cancer-adjacent tissue from invasive breast cancer patients. Complementary in vitro cocultures were used to study cell–cell communication between fibroblasts and specific breast cancer subtypes.

Results: Our results suggest that intrinsic tumor subtypes are reflected in histologically normal cancer-adjacent tissue. Gene expression of cancer-adjacent tissues shows that triple-negative (Claudin-low or basal-like) tumors exhibit increased expression of genes involved in inflammation and immune response. Although such changes could reflect distinct immune populations present in the microenvironment, altered immune response gene expression was also observed in cocultures in the absence of immune cell infiltrates, emphasizing that these inflammatory mediators are secreted by breast-specific cells. In addition, although triple-negative breast cancers are associated with upregulated immune response genes, luminal breast cancers are more commonly associated with estrogen-response pathways in adjacent tissues.

Conclusions: Specific characteristics of breast cancers are reflected in the surrounding histologically normal tissue. This commonality between tumor and cancer-adjacent tissue may underlie second primaries and local recurrences.

Impact: Biomarkers derived from cancer-adjacent tissue may be helpful in defining personalized surgical strategies or in predicting recurrence risk. Cancer Epidemiol Biomarkers Prev; 24(2); 406–14. ©2014 AACR.

Introduction

Breast-conserving therapy (BCT) with lumpectomy and radiotherapy and mastectomy is equally effective in treating early-stage breast cancer. However, approximately 10% to 15% of women undergoing BCT suffer ipsilateral breast tumor recurrence (1–3), often with metastasis (4, 5). Younger age has been associated with higher recurrence (6), but tumor characteristics may also be responsible because aggressive breast cancers tend to be diagnosed in younger women (7) and have higher local recurrence rates (8, 9).

Breast stromal microenvironments (including fibroblasts, endothelial cells, and immune cells), change during carcinogenesis. Cancer-associated fibroblasts may play a critical role in maintaining chronic inflammation around breast cancers (10), and may also have regulatory effects independent of immune cells (11–13). Stromal microenvironments also vary by breast cancer subtype, and may influence progression (14–16). Recent studies have examined benign, cancer-adjacent tissue and found substantial interindividual variation. These studies show two distinct subtypes of cancer-adjacent tissues with distinct survival patterns (17), and show that “molecular histology” of epithelium in cancer-adjacent tissues surrounding estrogen receptor (ER)-negative tumors differ from those of ER-positive cancers (18). ER-positive tumors are associated with high expression of ER mRNA in cancer-adjacent tissue (19). Thus, understanding the microenvironment surrounding breast cancer subtypes is important for recurrence and in targeting surgical strategies.

We hypothesized that genomic features of histologically normal, cancer-adjacent tissue differ by intrinsic subtype. On the basis of previous findings from cell culture and mouse models showing upregulation of key chemokines and growth factors in fibroblast interactions with basal-like breast cancers (14, 20, 21), it is important to characterize the microenvironment response to basal-like breast cancer in human tissue. We also sought to validate previous reports of differences in estrogen responsiveness of ER-positive tumor-adjacent tissue. We therefore investigated...
gene expression profiles of cancer-adjacent tissue using data from two independent sources: the National Cancer Institute's Polish Breast Cancer Study and The Cancer Genome Atlas (TCGA) Project. We then used insights from these in vivo studies to further interrogate subtype-specific gene expression in vitro.

Materials and Methods

Polish Women's Breast Cancer Study, the TCGA study, and the Normal Breast Study (UNC)

This study included 139 women from the Polish Women's Breast Cancer Study (PWBCS) with snap-frozen extratumoral breast and tumor tissues (Supplementary Table S1). The PWBCS is a population-based case-control study conducted in Poland (Warsaw and Lodz) during 2000–2003 (22). PWBCS cases were women ages 20 to 74 years with pathologically confirmed in situ or invasive breast carcinoma. Tissues from invasive tumors and non-neoplastic cancer-adjacent breast tissue were collected at the time of breast surgery. Histologically normal, cancer-adjacent tissues were <2 cm from the tumor margin. On the basis of in vitro evidence of their distinctive microenvironments, basal-like and luminal tumors were oversampled for these analyses. Patient data were collected from medical records and in-person interviews as described previously. All participants provided written informed consent under a protocol approved by the National Cancer Institute (Rockville, MD) and Polish Institutional Review Boards (IRB).

An additional 60 snap-frozen cancer-adjacent samples collected and analyzed by the TCGA were used as a validation dataset (Supplementary Table S1). These samples were all histologically normal, cancer-adjacent (<2 cm from tumor margin) to invasive breast carcinoma and tumor subtype was classified for these samples and reported previously (23).

A subset of 36 cancer-adjacent samples from "The Normal Breast Study" (NBS) was used for this study to evaluate distance from tumor margin. NBS is a hospital-based cross-sectional study conducted in University of North Carolina (UNC) Hospitals (Chapel Hill, NC) from 2009 to 2014. All patients had a newly diagnosed invasive breast carcinoma. Fresh tissues were collected at the time of breast surgery and snap frozen in liquid nitrogen. Tumor adjacent breast tissues used in this study were classified as peritumoral (<2 cm from the tumor margin) and remote (>2 cm from the tumor margin). Information on clinicopathological, demographic, and anthropometric factors was collected from medical records and in-person interviews. Detailed data on tumor subtype were unavailable for these patients so they could not be used for primary analyses of subtype-specific microenvironments. All of the participants provided written informed consent under a protocol approved by the IRB.

Tumor expression analysis: molecular classification of the tumor using PAM50

Tumor samples from PWBCS and TCGA were used to determine molecular subtype of the cancer. RNA was isolated using previously published methods. For PWBCS data, Illumina Ref-8 Beadchip Version2 microarray platform was used and data normalization was performed using Lumi in R. For TCGA data, custom Agilent arrays or RNA sequencing were performed as described in (ref. 23). To classify tumors, genes were median centered and samples were standardized to zero mean and unit variance. The PAM50 predictor was performed (24) to categorize the tumors into five subtypes (luminal A, luminal B, Her2-enriched, basal-like, and normal-like). The Claudin-low predictor was applied as in (ref. 25).

Cancer-adjacent expression analysis

For PWBCS, two-color 4 × 44K Agilent whole-genome arrays were performed on a frozen section of cancer-adjacent tissue, with sections on either side used for imaging of cellular composition. Tissue for microarrays was homogenized using a MagnaLyser homogenizer (Roche), and RNA was isolated and quality was checked as described in Troester and colleagues (15). Microarrays were performed as previously described (26). Briefly, Cy3-labeled reference was produced from total RNA from Stratagene Universal Human Reference (spiked 1:1,000 with MCF-7 RNA and 1:1,000 with ME16C RNA to increase expression of breast cancer genes) following amplification with Agilent low RNA input amplification kit. Patient samples were labeled with Cy5. Data were Lowess normalized, and probes with a signal <10 dpm in either channel were excluded as missing. Probes with more than 20% missing data across samples were excluded. In data preprocessing, we (i) eliminated probes without corresponding ENTREZ ID, (ii) collapsed duplicate probes by averaging, (iii) imputed missing data using k-nearest neighbors (KNN) method with k = 10, and (iv) median centered genes. Microarray data are publicly available through the Gene Expression Omnibus (GEO; GSE49175 (26), GSE50393). TCGA data and methods are available at the TCGA Data Portal (https://tcga-data.nci.nih.gov/).

Supervised analysis of cancer-adjacent tissue

Using the PWBCS as a training set, four-class significance analysis of microarrays (SAM; ref. 27) was used to identify differentially expressed genes associated with breast cancer subtypes (luminal A, luminal B, Her2, and Basal-like plus Claudin-low; ref. 27). Significance was defined as FDR <0.1%. Tumors classified as "normal-like" may result from extensive normal or stromal content in the tumor (28), so we excluded normal-like tumors. The genes identified as differentially expressed in these four groups from the SAM analysis are henceforth referred to as the "in vivo triple-negative microenvironment signature" (Fig. 1; full list of genes can be found in Supplementary Table S2). Because there was a common pattern for both triple-negative tumor subtypes versus other subtypes, we collapsed tumor subtypes to conduct a two-class comparison (basal-like/Claudin low vs. HER2/luminal). This two-class gene list was used only for gene ontology; the more parsimonious (fewer genes) four-class list was used for all subsequent classification. Gene ontology analysis was done using Ingenuity Pathway Analysis with Benjamin–Hochberg multiple testing correction to identify significant functions and pathways (P < 0.05). Pathways and functions with less than two genes were excluded (Supplementary Tables S3 and S4).

To evaluate the association of gene expression for each patient with the defined biologic signature, Pearson correlation coefficients were obtained as described in (refs. 17, 26, 29). Briefly, for a given gene signature [i.e., EREs (30) and the "in vivo triple-negative microenvironment signature"], "1" was assigned to upregulated and "-1" to downregulated genes. Pearson correlation coefficients were calculated by comparing this standard vector to the measured, median-centered gene expression level for each patient. Patients were classified as positive if the Pearson correlation coefficient was >0, and negative if the coefficient was <0. These classes were further evaluated for their association with other tumor characteristics. The estrogen response signature
(ERE5) signature was analyzed in both datasets and the "in vivo triple-negative microenvironment signature" was identified in the PWBCS and tested in the TCGA dataset. Because all analyses were conducted with median-centered data, expression is relative to other subtypes. That is, high expression of the triple-negative signature in tissue adjacent to basal-like and Claudin-low tumors implies lower expression of these inflammatory genes among tissue adjacent to luminal tumors. Conversely, high expression of the ERE5 in tissue adjacent to luminal tumors implies relatively lower estrogen responsiveness among tissue adjacent to triple-negative tumors.

Composition analysis of cancer-adjacent tissues from PWBCS

Frozen sections were obtained from the same frozen block used for the microarray analysis and were taken immediately adjacent to the piece used for RNA extraction. One hundred and twenty-seven samples from the PWBCS and 47 samples from TCGA had hematoxylin and eosin (H&E)-stained sections of sufficient quality to be analyzed for tissue composition. Composition analysis was performed as explained in (ref. 26). Briefly, 20 μm (PWBCS) or 5 μm (TCGA) frozen sections were H&E stained. A training set of slides was scanned using Aperio ScanScope CS V11.0.2.725 and images were manually annotated for composition of adipose, epithelium, and nonfatty stroma. This training set was used to create a Genie (Aperio Technologies) algorithm with high accuracy in segmenting adipose tissue, epithelium, nonfatty stroma, and glass on each slide. Agreement between manual and digital assessment exceeded 98%, and thus the data from Genie digital segmentation was used in analyses. Composition data from both datasets (mean values) were used to determine a cutoff point for dichotomized analyses of composition; an epithelium cutoff point of 10% (mean values PWBCS = 9.8% and TCGA = 10.2%) and a stroma cutoff point of 20% (mean values PWBCS = 26.8% and TCGA = 14.4%) were selected.

Cell lines and coculture conditions

Cell lines were purchased and maintained as in ref. (14); MCF7, SKBR3, MDA-MB-231, HCC1937, and MDA-MB-468.
were purchased from ATCC and passaged for less than 6 months before being used for experiments. Sum159 were purchased from Asterand and passaged for less than 6 months before being used for experiments, authentication was provided at time of purchase. ZR75, T47D, SUM149, ME16C, and SUM102 were described in Troester and colleagues (31). Direct cocultures, wherein cells are seeded in a single well and in physical contact (not separated by Transwell membranes), were performed as previously described (14). Briefly, cancer cells lines (MCF7, ZR75, T47D, SKBR3, MDA-MB-231, SUM159, SUM149, HCC1937, ME16C, SUM102, and MDA-MB-468) and immortalized reduction mammary fibroblasts (RMF; ref. 32) were plated on plastic in direct contact, potentially interacting both through secreted factors and through cell-cell contact. RMFs were tested for viability in all cancer cell media, and direct cocultures were maintained in the appropriate cancer cell media (e.g., MCF-7 in RPMI). The following RMF:cancer cell ratios were plated for most direct cocultures: 0:1, 1:4, 1:2, 1:1, 2:1, 1:0. Cocultures and monocultures (for comparison) were maintained for 48 hours before RNA isolation.

RNA and expression microarrays: cell lines
Monocultures and cocultures were harvested by scraping in RNA lysis buffer. Total RNA was isolated using the RNeasy mini kit (Qiagen) as previously described in (Camp and colleagues; ref. 14). Microarrays were performed according to Agilent protocol using 2-color Agilent 4×44K (Agilent G4112F) human arrays and 244K (Agilent G4502A) custom human arrays. Only probes present on both platforms were utilized. Samples were run in batches together with appropriate monoculture controls to minimize the effect of batch. Agilent’s Quick Amp labeling kit was used to synthesize Cy3-labeled reference from Stratagene Universal Human Reference (as described above) and Cy5-labeled RNA from cocultured or monocultured cell lines (as previously described in ref. 14). Data are available through the GEO (GSE26411).

Coculture data normalization and analysis
Data from 122 microarrays (representing monocultures and direct cocultures from 12 different cell lines described above) were included. Only genes where >70% of microarrays had signal >10 dpi in both channels were included. Data were Lowess normalized and missing data were imputed using k-nearest neighbors’ imputation. For the direct coculture analyses, we excluded genes that did not have at least 2-fold deviation from the mean in at least one sample, and the method of Bueb and colleagues (33, 34) was used to normalize cocultures to appropriate monocultures as described in (ref. 14). Briefly, the Bueb method is an established expression deconvolution approach for direct cocultures of two different cell lines that estimates the percentage of fibroblasts and cancer cells in each coculture, and normalizes the data for composition differences before estimating the effect of epithelial-stromal interaction on gene expression. The Bueb interaction coefficient ’I’ was calculated as the ratio of observed to expected gene expression for each gene and an ‘I-matrix’ representing the epithelial-stromal interaction coefficients for each gene in each coculture was generated. This I-matrix was analyzed using multiclass SAM (27) in R 1.14, comparing basal-like and Claudin-low cocultures (i.e., direct cocultures of MBA-MB-231, SUM159, SUM149, HCC1937, SUM102, and MDA-MB-468 with RMFs) to Her2-enriched and luminal cocultures (direct cocultures of MCF7, ZR75, T47D, SKBR3, ME16C with RMFs). Heatmaps were generated using Cluster 3.0 and Java treeview was used to visualize data.

Evaluating correlation between breast cancer subtype and in vitro triple-negative gene signature
The "in vitro triple-negative signature" (Supplementary Table S2) was defined by all genes significantly upregulated in basal-like and Claudin-low cocultures after performing a multiclass SAM analysis. Because, all of the genes defining this "in vitro triple-negative signature" were upregulated, the average expression levels across all genes were used to score tumors and normal tissue for the triple-negative signature. Expression values were median-centered by gene before summing. Boxplots were used to compare the triple-negative score across intrinsic subtypes. We further evaluated the difference in mean expression by subtype using ANOVA.

Statistical analysis
R version 1.14 was used to generate box plots, evaluate signature correlations, and to perform χ² significance tests. SAS 9.2 (32) was used for logistic regression to estimate ORs for expression of estrogen response and the triple-negative stromal response signatures by subtype. Fisher exact test was used to test associations with clinical features and for subtype-stratified tissue composition analysis.

Results
Tumor intrinsic subtype is reflected in cancer-adjacent tissue
We used samples from the PWBCS (population characteristics in Supplementary Table S1) to identify subtype-associated changes in cancer-adjacent tissue. Triple-negative cancer-adjacent tissues had a unique stromal response, with Fig. 1 showing a heatmap of 126 genes (Supplementary Table S2) whose expression differed significantly between tissues adjacent to basal-like, Claudin-low, luminal B, and luminal A tumors. Gene ontology analysis (using Ingenuity Pathway Analysis) revealed that genes upregulated in tissues adjacent to triple-negative tumors are involved in functions and pathways such as activation of leukocytes, proliferation of mononuclear leukocytes, cell movement of leukocytes, IFN signaling, hepatic fibrosis, T-helper cell differentiation, or antigen presentation pathway (full list in Supplementary Tables S3 and S4).

Our results suggest that the cancer-adjacent tissue shares biology of the tumors themselves. Four transcripts (NAT1, FOXA1, MLPH, ESR1) used in the PAM50 subtyping (24) have differential expression in the tissue adjacent to luminal breast cancers. Having observed high expression of estrogen receptor 1 (ESR1) adjacent to luminal breast cancers, and in light of previous reports suggesting similarities between ER-positive tumors and their adjacent tissue (19), we utilized a published ERes (30) to characterize the estrogen response of each cancer-adjacent tissue (positive or negative for the ERes) in both populations. Table 1 shows that in both populations there was a significant association between the expression of this signature and breast cancer subtype, with the majority of luminal A cancer-adjacent tissue being positive for ERes (PWBCS: 62.26% and TCGA: 68.00%) and the vast majority of the more aggressive Claudin-low cancer-adjacent tissue being negative (PWBCS: 92.31% and TCGA: 100.00%). Moreover, there was a significant association between
expression of the *in vivo* triple-negative signature by subtype (*P* value: 0.0033 in PWBCS, training set and *P* value: 0.0005 in TCGA, test set, respectively). Claudin-low cancer-adjacent tissues had the highest percentage of *in vivo* triple-negative signature-positive tumors (84.65% in PWBCS and 100% in TCGA) and luminal A cancer-adjacent tissues had the lowest percentage of positive tumors (only 34.00% in PWBCS and 28.00% in TCGA; Supplementary Table S1).

Cancer-adjacent expression is not associated with tissue composition

One source of variation in cancer-adjacent tissue is the heterogeneous composition, with some patients having more or less stroma and epithelium than others. To evaluate the role of tissue composition in expression of the *in vivo* triple-negative microenvironment signature, we quantified the proportion of epithelium, nonfatty stroma, and fatty/adipose tissue in each sample. Histologic sections adjacent to the portion of tissue used for RNA extraction were used for H&E staining and analysis. Table 2 dichotomizes the samples according to composition. There were no statistically significant associations between tumor subtype (from which the adjacent tissue came) and epithelial content (Fisher exact *P* value = 0.213 and *P* value = 0.177) or nonfatty stromal content (Fisher exact *P* value = 0.235 and *P* value = 0.104) in either dataset of cancer-adjacent tissue, underscoring that the triple-negative gene expression profile is not due to differences in composition by subtype. However, the contrary is true for the EReS signature, it was significantly associated with nonfatty stromal content (*P* value < 0.001 in PWBCS and *P* value = 0.049 in TCGA) in both datasets, suggesting a role for the stroma in estrogen response of normal breast tissue.

Given that tumor margins may vary from tumor to tumor, we addressed the EReS signature in an additional dataset from UNC-stroma in estrogen response of normal breast tissue.

Cancer-adjacent biology can be recapitulated *in vitro*

We next identified a triple-negative signature in coculture (*in vitro*) and evaluated whether this *in vitro* signature also accurately identifies triple-negative adjacent samples. Twelve breast cancer subtype models were cocultured with immortalized RMFs and we identified a unique set of genes upregulated in triple-negative cocultures (Fig. 2A, gray bar “*in vitro* triple-negative signature”; the full list of genes can be found in Supplementary Table S2). To address whether this *in vitro* gene signature is associated with tissue adjacent to triple-negative tumors, two approaches were used: (i) comparison of pathways and biologic functions obtained *in vitro* versus in tissue and (ii) evaluation of the *in vitro* signatures in breast tumors and the cancer adjacent tissue.

Using the first approach, genes identified through the *in vitro* cocultures (Supplementary Tables S5 and S6) and *in vivo* (Supplementary Tables S3 and S4) were in similar pathways. Statistically significant biologic functions and pathways in common were activation of cells, proliferation of mononuclear leukocytes, cell movement of leukocytes, inflammatory response, hepatic fibrosis, and role of cytokines in mediating communication between immune cells and IL6 signaling. Using the second approach, basal-like and Claudin-low tumors had high expression of genes identified in cocultures of triple-negative cancer cell lines, both within the tumor (Fig. 2B, one-way ANOVA by subtype, *P* = 8.44e−15) and in the cancer-adjacent tissue (of the PWBCS; Fig. 2C). The association with cancer-adjacent tissue was weaker and not significant (*P* = 0.196); however, the expression from cancer-adjacent tissues qualitatively mirrors the expression patterns from tumors.

### Table 1. Expression of the estrogen response gene signature (EReS) and the "in vivo triple-negative microenvironment signature" according to tumor subtype in PWBCS and TCGA

<table>
<thead>
<tr>
<th></th>
<th>PWBCS, N (%)</th>
<th>TCGA, N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basal-like</td>
<td>Claudin-low</td>
</tr>
<tr>
<td><strong>EReS (+)</strong></td>
<td>7 (41.2)</td>
<td>1 (7.7)</td>
</tr>
<tr>
<td><strong>EReS (−)</strong></td>
<td>10 (58.3)</td>
<td>12 (92.3)</td>
</tr>
</tbody>
</table>

**In vivo TN sig (+)**

<table>
<thead>
<tr>
<th></th>
<th>PWBCS, N (%)</th>
<th>TCGA, N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basal-like</td>
<td>Claudin-low</td>
</tr>
<tr>
<td><strong>In vivo TN sig (+)</strong></td>
<td>12 (70.6)</td>
<td>11 (84.6)</td>
</tr>
</tbody>
</table>

**In vivo TN sig (−)**

<table>
<thead>
<tr>
<th></th>
<th>PWBCS, N (%)</th>
<th>TCGA, N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basal-like</td>
<td>Claudin-low</td>
</tr>
<tr>
<td><strong>In vivo TN sig (−)</strong></td>
<td>5 (29.4)</td>
<td>2 (15.4)</td>
</tr>
</tbody>
</table>

### Table 2. Cancer-adjacent tissue composition according to tumor subtype and estrogen response gene signature (EReS) in PWBCS and TCGA

<table>
<thead>
<tr>
<th></th>
<th>PWBCS</th>
<th>TCGA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basal-like</td>
<td>Claudin-low</td>
</tr>
<tr>
<td>% Epithelium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10%</td>
<td>15 (76.5)</td>
<td>4 (36.3)</td>
</tr>
<tr>
<td>≥10%</td>
<td>4 (23.5)</td>
<td>7 (63.6)</td>
</tr>
<tr>
<td>% Stromal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20%</td>
<td>9 (52.9)</td>
<td>8 (72.7)</td>
</tr>
<tr>
<td>≥20%</td>
<td>8 (47.1)</td>
<td>5 (27.3)</td>
</tr>
<tr>
<td>% Epithelium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10%</td>
<td>6 (75.0)</td>
<td>8 (88.9)</td>
</tr>
<tr>
<td>≥10%</td>
<td>2 (25.0)</td>
<td>1 (11.1)</td>
</tr>
<tr>
<td>% Stromal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20%</td>
<td>7 (87.5)</td>
<td>9 (100.0)</td>
</tr>
<tr>
<td>≥20%</td>
<td>1 (12.5)</td>
<td>—</td>
</tr>
</tbody>
</table>
Discussion

Studies of breast cancer microenvironment by subtype (19, 35) have important implications for local recurrence. Locoregional recurrence may be higher among basal-like breast cancers (36), and we hypothesized that these cancers may induce a permissive microenvironment for local recurrence. The results presented in this manuscript suggest that cancer-adjacent tissue of basal-like and Claudin-low breast cancers differs substantially from that of luminal cancers and that these differences are strongly dependent upon fibroblast interactions and/or stromal composition. That is, the microenvironment may be primed for inflammation (37) or estrogen response by the extratumoral stroma.

Previous studies have isolated intratumoral and extratumoral stroma and evaluated gene expression in association with outcome. For example, Finak and colleagues reported that stroma
adjacent to basal-like breast cancers had high levels of immune response genes and that these genes predicted progression (16). However, microdissection cannot perfectly exclude immune infiltrates, so it was difficult to identify the cell type responsible for the upregulation of immune mediators. Our use of cocultures allowed us to identify fibroblasts as key contributors to cytokine/chemokine expression. Even in the absence of immune cells, fibroblasts and epithelial cells produce molecules that affect immune cell recruitment and activation and that directly regulate epithelial cell differentiation (e.g., IL6 alters epithelial cell phenotypes; ref. 38). Thus, although previous studies established that triple-negative breast cancers are associated with a proinflammatory milieu (25, 37, 39, 40), the current findings suggest that this reaction may initiate in epithelial–fibroblast interactions and occurs in both tumors and surrounding histologically normal tissue.

Our results also addressed unique features of luminal microenvironements in vivo. Cancer-adjacent gene expression for luminal tumor subtypes is markedly different from that of basal-like and Claudin-low tumors, with luminal-adjacent tissues expressing high levels of luminal breast tumor markers (24). In other words, the cancer-adjacent tissue around the tumor reflects or may even predict the biology of the tumor that arises. If ER-positive tumors are more likely to occur in pervasively estrogen-responsive benign tissues (41, 42), extra-tumoral signatures could be candidate biomarkers for predicting subtype-specific risk. Others have hypothesized that host factors or widespread field effects cause first and second primaries to have similar phenotypes, with most second primaries having the same phenotype as the first primary (43, 44). Thus, the idea of an estrogen-responsive "field effect" is supported by different types of human data. The viability of using estrogen response in normal tissue to predict risk or recurrence risk depends upon whether these signatures occur early or late in carcinogenesis. Because our data were collected at a single time point after tumor onset, the genomic differences we observed could represent (i) patterns of predisposition that preexist tumor formation, or (ii) a reaction to the tumor. We hypothesize that the triple-negative signatures are a response to tumor, whereas the EReS reflects susceptibility. Future longitudinal studies are needed to test these hypotheses in additional populations and should consider other epidemiologic variables (such as body mass index, age, or estrogen exposure).

Cancer-adjacent tissue is composed of a high percentage of stroma (both fatty and nonfatty) and gene expression analyses of these tissues are enriched for stromal pathways. Because of this, other studies have approached the study of adjacent tissue by microdissecting and studying individual cellular components (16, 45, 46). However, cellular heterogeneity is important, and inclusion of all cell types enabled us to evaluate how stromal composition relates to pathway expression. Our coculture results confirm that immune response may initiate with stromal–epithelial interactions. Our tissue studies show concordance between EReS and stromal content, suggesting a role for stroma in modulating estrogen activity. Previous literature also suggests biologic plausibility for this association (47), but the mechanisms of estrogen action in stroma remain to be fully elucidated. However, regardless of the ultimate mechanisms, both of the signatures evaluated were originally defined in cell lines, underscoring our findings and those of others that show that cell line-derived signatures have accuracy in elucidating in vivo biology (14, 33, 48, 49).

Strengths of this analysis include use of two distinct sample sets: one as a training dataset (PWBCS), and the other as an independent validation set (TCGA). In addition, tissue composition was considered as a potential modifier. It is also a strength that we have validated the signatures in a controlled, experimental coculture system. Future work should focus on investigating the role of distance from tumor more thoroughly, preferably in larger study populations.

In conclusion, we found distinct biologic characteristics of cancer-adjacent tissue, depending upon tumor intrinsic subtype. This commonality between tumor and surrounding tissue may underlie second primaries and provides plausible explanations for local recurrence. These results also suggest that tissue biomarkers derived from cancer-adjacent tissue may help in predicting risk and in defining appropriate, personalized surgical strategies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Concept and design: X. Sun, R. Sandhu, L. Makowski, M.E. Sherman, J.D. Figueroa, M.A. Troester


Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P. Casbas-Hernandez, X. Sun, R. Sandhu, X. R. Yang, J.D. Figueroa, M.A. Troester

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P. Casbas-Hernandez, X. Sun, E. Roman-Perez, M. D’Arcy, R. Sandhu, A. Hishida, M.E. Sherman, J.D. Figueroa

Writing, review, and/or revision of the manuscript: P. Casbas-Hernandez, M. D’Arcy, R. Sandhu, A. Hishida, K.K. McNaughton, L. Makowski, M.E. Sherman, J.D. Figueroa, M.A. Troester

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): R. Sandhu, M.A. Troester

Study supervision: M.A. Troester

Acknowledgments

The authors thank Dr. Keith Amos for all his support during the initial stages of this manuscript. The authors gratefully acknowledge Monte Garcia-Closas, Louise Brinton, and Jolanta Lissowska for contributions to the Polish study; Paul Melzer, Sean Davis, and Sarah Anzick for profiling data of tumor tissues; and the UNC Translational Pathology Laboratory for scoring and analyzing histology.

Grant Support

X. R. Yang, M.E. Sherman, and J.D. Figueroa received NCI intramural funds. M.A. Troester received funds from the Avon Foundation, the University Cancer Research Fund, the National Cancer Institute (RO1CA-138255), the NCI/National Institutes of Environmental Health Sciences (NIEHS) Breast Cancer and the Environment Research Program (BCERP; Z01-ES-019472), and a SPORE Grant (P50-CA-058223). P. Casbas-Hernandez was supported by a Howard Hughes Medical Institute Fellowship. The Polish Breast Cancer Study was supported by the National Cancer Institute Intramural funds.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 13, 2014; revised November 14, 2014; accepted November 19, 2014; published OnlineFirst December 2, 2014.
References

Cancer Epidemiology, Biomarkers & Prevention

Tumor Intrinsic Subtype Is Reflected in Cancer-Adjacent Tissue
Patricia Casbas-Hernandez, Xuezheng Sun, Erick Roman-Perez, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-14-0934

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2014/12/03/1055-9965.EPI-14-0934.DC1

Cited articles
This article cites 49 articles, 19 of which you can access for free at:
http://cebp.aacrjournals.org/content/24/2/406.full.html#ref-list-1

Citing articles
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
/content/24/2/406.full.html#related-urls

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.