



Published in final edited form as:

*Clin Cancer Res.* 2013 September 15; 19(18): 4972–4982. doi:10.1158/1078-0432.CCR-13-0029.

## RELATIONSHIP OF MAMMOGRAPHIC DENSITY AND GENE EXPRESSION: ANALYSIS OF NORMAL BREAST TISSUE SURROUNDING BREAST CANCER

Xuezheng Sun<sup>#1</sup>, Gretchen L. Gierach<sup>#2</sup>, Rupninder Sandhu<sup>3</sup>, Tyisha Williams<sup>4</sup>, Bentley R. Midkiff<sup>3</sup>, Jolanta Lissowska<sup>5</sup>, Ewa Wesolowska<sup>5</sup>, Norman F. Boyd<sup>6</sup>, Nicole B. Johnson<sup>7</sup>, Jonine D. Figueroa<sup>2</sup>, Mark E. Sherman<sup>2</sup>, and Melissa A. Troester<sup>1,3,†</sup>

<sup>1</sup>Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

<sup>2</sup>Hormonal and Reproductive Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA <sup>3</sup>Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA <sup>4</sup>Department of Biology, Trinity University, San Antonio, TX, USA <sup>5</sup> Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland <sup>6</sup>Campbell Family Institute for Breast Cancer Research, Ontario Cancer Institute, Toronto, ON, Canada <sup>7</sup>Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

# These authors contributed equally to this work.

### Abstract

**Purpose**—Previous studies of breast tissue gene expression have demonstrated that the extratumoral microenvironment has substantial variability across individuals, some of which can be attributed to epidemiologic factors. To evaluate how mammographic density (MD) and breast tissue composition relate to extratumoral microenvironment gene expression, we used data on 121 breast cancer patients from the population-based Polish Women's Breast Cancer Study.

**Design**—Breast cancer cases were classified based on a previously reported, biologically-defined extratumoral gene expression signature with two subtypes: an Active subtype, which is associated with high expression of genes related to fibrosis and wound response, and an Inactive subtype, which has high expression of cellular adhesion genes. MD of the contralateral breast was assessed using pre-treatment mammograms and a quantitative, reliable computer-assisted thresholding method. Breast tissue composition was evaluated based on digital image analysis of tissue sections.

**Results**—The Inactive extratumoral subtype was associated with significantly higher percentage mammographic density (PD) and dense area (DA) in univariate analysis (PD:  $p=0.001$ ; DA:  $p=0.049$ ) and in multivariable analyses adjusted for age and body mass index (PD:  $p=0.004$ ; DA:  $p=0.049$ ). Inactive/higher MD tissue was characterized by a significantly higher percentage of

<sup>†</sup>Corresponding author Correspondence to: Melissa A. Troester, Ph.D. Department of Epidemiology, CB 7435 University of North Carolina at Chapel Hill Chapel Hill, NC 27599 (919) 966-7408 troester@unc.edu.

Conflict of interest statement: None declared

stroma and a significantly lower percentage of adipose tissue, with no significant change in epithelial content. Analysis of published gene expression signatures suggested that Inactive/higher MD tissue expressed increased estrogen response and decreased TGF- $\beta$  signaling.

**Conclusions**—By linking novel molecular phenotypes with MD, our results indicate that MD reflects broad transcriptional changes, including changes in both epithelia- and stroma-derived signaling.

### Keywords

mammographic density; breast microenvironment; gene expression

## Background

Molecular profiling of gene expression of breast cancers has demonstrated that tumors are remarkably heterogeneous, which has profound influences on etiological and clinical research (1, 2). More recently, molecular analyses of the microenvironment have demonstrated similar heterogeneity (3-7), but the epidemiological, clinical and pathological correlates of this variation are not well studied. Specifically, tandem analyses of breast cancers and the surrounding microenvironment may reveal important stromal-epithelial interactions. In fact, previous work suggests that stromal changes may precede tumor invasiveness and may reflect tumor characteristics (8-13). Findings such as these have led to speculation that the microenvironment may be dominant over tumor biology early in progression, when invasive cancers are still forming (8-13). While the importance of tumor microenvironment is increasingly established in the cancer biology literature (8, 14), the epidemiologic factors that affect the microenvironment remain poorly understood.

We recently reported an extratumoral signature, so-called Active signature, classifying extratumoral stromal microenvironments into two primary gene expression phenotypes (Active and Inactive) based on unsupervised clustering on 72 normal tissue samples adjacent to invasive breast cancer or ductal carcinoma in situ (Active, n=27; Inactive, n=45) (6). The Active subtype had high expression of genes involved in activation of fibrosis, cellular movement, increased TWIST expression and positive expression of TGF- $\beta$  signatures. The Inactive subtype expressed higher levels of cell adhesion and cell-cell contact genes. Compared with the Inactive subtype, ER+ and hormone-treated patients with the Active subtype had poorer overall survival, suggesting possible prognostic value. However, the Active subtype appeared to be independent of breast cancer subtype and standard clinicopathological parameters, such as tumor size and grade (6). These findings raise the hypothesis that extratumoral subtypes may be host factors rather than tumor-dependent factors.

We hypothesized that host factors influence the microenvironment that exists prior to tumor development and that these changes may be etiologically relevant. Thus, to further evaluate whether the Active/Inactive signature is related to tumor factors, host factors, or both, we performed gene expression profiling on extratumoral non-neoplastic breast tissues from 121 breast cancer patients of the population-based Polish Women's Breast Cancer Study (PWBCS) (15). In particular, based on differential expression of fibrosis, wound response,

and cell adhesion genes in the two subtypes, we hypothesized that the Inactive microenvironment would be associated with high mammographic density. We further hypothesized that because mammographic density is a radiologic reflection of variations in breast tissue composition, high mammographic density/inactive microenvironment would be associated with high non-fatty stroma and epithelial content. PWBCS participants are well-characterized with respect to established breast cancer risk factors, including mammographic density, providing us with the opportunity to link the Active/Inactive molecular phenotype with mammographic density.

## Materials and Methods

### Study population

The study population included 121 women from the PWBCS with available snap frozen extratumoral breast tissues and mammographic density. The PWBCS is a population-based case-control study conducted in two major cities in Poland (Warsaw and Łódź) during 2000-2003 (15). PWBCS cases were women aged 20-74 years with newly-diagnosed, pathologically-confirmed *in situ* or invasive breast carcinoma identified through a rapid identification system organized at five participating hospitals and via cancer registries. Fresh tissues from invasive tumors, non-neoplastic adjacent breast tissue and mammary fat tissue were collected at the time of breast surgery and snap frozen in liquid nitrogen. Tumor adjacent breast tissues used in this study were <2cm from the tumor margin. Based on *in vitro* evidence of their distinctive microenvironments (16), basal-like and luminal tumors were oversampled in this study. Information on clinicopathological, demographic, and anthropometric factors was collected from medical records and in-person interviews as described previously (15). All of the participants provided written informed consent under a protocol approved by the U.S. National Cancer Institute and local (Polish) institutional review boards.

### Mammographic density measurement

Pre-treatment mammograms of the unaffected breast were sent to the Ontario Cancer Institute in Toronto, Canada where they were digitized using a Lumisys® 85 laser film scanner. Patient identifiers were permanently deleted from the electronic images. Craniocaudal views of digitized films were used to assess mammographic density with Cumulus®, an interactive computer-assisted thresholding program developed at the University of Toronto (17). One expert reader (NFB) measured absolute dense area (cm<sup>2</sup>) and total breast area (cm<sup>2</sup>) using the methods as described previously (17); percentage mammographic density was calculated by dividing the dense breast area by the total breast area and multiplying by 100. A repeat set of 49 images was assessed for reliability. The intra-class correlation coefficients for percent mammographic density, dense area, and total breast area were 0.95, 0.93, and 0.99, respectively, documenting excellent reproducibility.

### Breast tissue composition measurement

Frozen non-neoplastic breast specimens of approximately 100 mg were cut over dry ice and then used to cut frozen sections. Sections were collected at both ends of the specimen and then constructed into 20 µm slides. The central portion was used for RNA extraction. After

H&E staining, the slides were scanned into high-resolution digital images using the Aperio Scan-Scope XT Slide Scanner (Aperio Technologies, Vista, CA, USA) in the UNC Translational Pathology Laboratory. After excluding slides with poor resolution or having folded tissues, slides from 118 women (97.5%) were subjected to breast tissue composition analysis. To train the composition estimator in Aperio's Genie software, 15 representative digital slides were selected and manually annotated for epithelial area, stromal area, and total area (mm<sup>2</sup>) using Aperio ImageScope software. These digital area-based, quantitative estimates were used to train Aperio's Genie Classifier to partition epithelium, adipose tissue, non-fatty stroma, and glass into percentages. Examples of annotated digital images are presented in Supplementary Figure 1. The regular H&E counterparts of these 15 digital slides were also evaluated by eye by a pathologist who provided semi-quantitative estimates of the percentage of adipose tissue (10% bin width), epithelium (1% bin width), and non-fatty stroma (10% bin width). To assess the performance of Genie classifiers, we compared the results of three methods (by Genie, by pathologist digital slide-based, and by pathologist regular H&E slide-based). The trained classifier was positively and strongly correlated with manually scored area based on the digital images, for all three tissue compartments. The trained Genie classifier was strongly correlated with pathologist review based on regular H&E slides for stroma and adipose (Supplementary Table 1, Pearson correlation coefficient ranged 0.95-0.96), while relatively lower for epithelium (Pearson correlation coefficient=0.68). Compared with digital assessment, visual assessment (by human eye on regular H&E slides) of small percentage differences is weaker, such as epithelial tissue, which is sparse (<10%) in benign breast. Thus the digital image analysis data were used in analyses, and the trained Genie Classifier was then applied to the remaining slides.

### RNA isolation and microarrays

All microarrays on non-neoplastic breast tissue were performed at the University of North Carolina at Chapel Hill. The central section of fresh frozen non-neoplastic tissue (as described above) was homogenized using a MagnaLyser homogenizer (Roche), and RNA was isolated by Qiazol extraction followed by purification on an RNeasy column as described in Troester et al. (5). RNA quality and quantity were analyzed on an Agilent 2100 Bioanalyzer and a ND-1000 Nanodrop spectrophotometer, respectively, before running two-color 4X44K Agilent whole genome arrays. 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. The same protocol was applied to total RNA from breast tissues, with all patient samples labeled with Cy5. Data were lowess-normalized, and probes that had a signal of <10 dpi in either channel were excluded as missing. Probes that had more than 20% missing data across all samples were excluded from further analysis. In expression data preprocessing, we 1) eliminated the probes without corresponding ENTREZ ID, 2) collapsed the duplicate probes by averaging, 3) imputed missing data using k-nearest neighbors (KNN) method with k=10, and 4) median-centered each gene. Microarray data for 121 specimens used in this analysis are publicly available through the Gene Expression Omnibus (GSE49175).

## Statistical analysis

Samples were classified as having Active or Inactive extratumoral subtype using unsupervised hierarchical clustering (average linkage) on the genes in the published Active/Inactive extratumoral signature (6). 324 of 3518 genes in the original Active signature were excluded in the clustering analysis as they did not pass filtering criteria described above. Consensus clustering was used to evaluate whether two clusters resulted in optimal segregation of the sample classes (18). The Active or Inactive extratumoral subtype obtained by clustering was validated using the Creighton correlation method (19), described in Supplementary Figure 2. Briefly, a standard vector corresponding to all genes in the Active/Inactive signature was constructed, with 1 assigned to up-regulated genes and -1 assigned to down-regulated genes. A Pearson correlation coefficient was calculated for this standard vector vs. the vector of median centered gene expression for each patient. The identities of the 3194 genes and their corresponding standard vector for Creighton correlation in the present study were shown in Supplementary Table 2. Patients were classified as Active if the Pearson correlation coefficient was greater than zero, and Inactive if the coefficient was less than zero.

The distributions of breast cancer risk factors by the Active/Inactive subtype were assessed, including age at diagnosis (continuously and categorically as <50 y and ≥50 y), BMI (continuously and categorically as <30 kg/m<sup>2</sup> and ≥30 kg/m<sup>2</sup>), age at menarche (<12 y and ≥12 y), parity (nulliparous and parous), age at 1<sup>st</sup> full term birth in parous women (<25 y and ≥25 y), use of oral menopausal hormone therapy (MHT, never and ever), menopausal status (premenopausal and postmenopausal), family history of breast cancer in the first degree relative (yes and no), previous benign breast disease history (defined as if a subject ever had any biopsy or partial removal procedure, but no cancer detected; yes and no), percentage mammographic density (continuously and categorically as <25% and ≥25%), and dense area (continuously and categorically as <median and ≥median; median=36.52 cm<sup>2</sup>). Moreover, the distributions of clinicopathologic characteristics by the Active/Inactive subtype were evaluated, including tumor size (<2 cm and ≥2 cm), histological type (ductal and others), differentiation stage (well/moderate and poor), number of positive axillary lymph nodes (0 and ≥1), ER status and intrinsic subtype. To identify potential confounding factors, the distributions of clinicopathologic characteristics and breast cancer risk factors by percent mammographic density and dense area were also examined. The corresponding associations were evaluated using Chi-square tests for categorical variables and Student's t-tests for continuous variables. Moreover, to assess the magnitude of association between risk factor exposure (e.g. obese vs. nonobese) and extratumoral microenvironment subtype, we computed odds ratios (OR) for the Inactive subtype (vs. Active, referent) and corresponding 95% confidence intervals (CI). Analysis of covariance (ANCOVA) was used to examine the relationship between Active/Inactive subtype and mammographic density, adjusting for potential confounders, age and BMI, continuously.

To explore the biological meaning of microenvironment subtypes, the gene expression profile in each sample was compared to five previously reported gene expression signatures of related phenotypes, including young age (20), obesity (21), desmoid-type fibrosis (DTF) (22), TGF-β (6) and estrogen (E2) response (23). To ensure that each of these signatures was

independent, overlap between gene sets was assessed and found to be small (Supplementary Table 3). The median centered expression profile of each patient was evaluated by calculating Pearson correlation coefficients, using the method of Creighton et al. described above (19). More information regarding the gene ID and direction of each signature is listed in Supplementary Table 4. The associations of existing signatures with the extratumoral subtype (Active and Inactive), percentage mammographic density (<25% and ≥25%), and breast tissue composition (percentage adipose, epithelium and stroma) were assessed by Chi-square tests and Student's t-test, respectively. Non-parametric exact methods were used when expected cell count was less than 5. Logistic regression and generalized linear regression were used to further evaluate the associations after adjusting for age and/or BMI. Probability values of less than 0.05 were considered statistically significant. All analyses were performed using R (version 2.9.2).

## Results

### Associations between Active/Inactive subtype and breast cancer risk factors

Unsupervised hierarchical clustering (average linkage) on the 3,000 genes from the published Active/Inactive signature was used to classify extratumoral breast tissues from 121 cases into two homogeneous groups. Consensus clustering documented that two groups provided an optimal segregation of the data (Supplementary Figure 3). Using the Creighton correlation method to categorize the microenvironment subtype as described in Methods also confirmed that the direction of association for Active and Inactive groups by clustering was concordant with the previous publication of the signature, showing high agreement of the two classification methods with Kappa statistic of 0.92 ( $p < 0.0001$ ). Approximately 49% of patients ( $n=59$ ) were Active in the current study, compared to 38% of patients ( $n=27$ ) in Roman-Perez et al. (6).

Consistent with our previous findings (6), we observed no association between extratumoral subtype and ER status and other tumor characteristics (Supplemental Table 5). Table 1 shows the distribution of participant characteristics, overall and categorized according to extratumoral microenvironment subtype. Of all the breast cancer risk factors evaluated, mammographic density measures showed the strongest differences by Active/Inactive subtype. Compared with the Active group, mean percentage density was 9% higher and mean dense area was 8.8 cm<sup>2</sup> more in the Inactive group (percentage density:  $p=0.0013$ ; dense area:  $p=0.049$ ). Substantial (but statistically insignificant) associations were noted for age [age <50y vs. ≥50y: OR = 2.00, 95% CI=(0.88, 4.57)] and BMI [BMI <30kg/m<sup>2</sup> vs. ≥30kg/m<sup>2</sup>: OR = 1.76, 95% CI=(0.79, 3.93)], but no other breast cancer risk factors showed substantial associations with Active/Inactive subtype.

We also evaluated potential confounders of the association between mammographic density and extratumoral microenvironment subtype. Both elevated percentage density and dense area were significantly associated with younger age, lower BMI, ever use of oral MHT and premenopausal status (Supplementary Table 6). Of these factors, only age and BMI showed substantial associations with extratumoral subtype. Thus, we performed multivariable analysis to study mean differences in mammographic density by extratumoral microenvironment subtype, controlling for age and BMI (Table 2). In multivariable models,

percent density and dense area remained higher among the Inactive group (percentage density:  $p=0.0038$ ; dense area:  $p=0.049$ ).

### Biological relevance of Active/Inactive subtype

To better understand the molecular characteristics of extratumoral microenvironment subtype and mammographic density-associated breast cancer risk, associations were evaluated between Active/Inactive subtype, mammographic density, and biologically-relevant, published gene expression signatures. As shown in Table 3, both the Inactive subtype and higher percentage density (defined as  $\geq 25\%$ ) were significantly positively associated with a young-like age signature and an increased estrogen response expression signature (E2), while they were significantly inversely associated with an obesity signature and a TGF- $\beta$  signature. After adjusting for patient age and/or BMI, most of these associations remained statistically significant, but the association of high percentage density with the young-like age signature and low TGF- $\beta$  signature was no longer significant (Supplementary Table 7). The previously reported signature derived from desmoid-type fibroid tumors and indicative of fibroblast response (DTF) was only associated with lower percentage density, but not after adjusting for BMI and age.

### Associations between subtype/mammographic density and breast tissue composition

Given variation in stroma-derived signatures, we expected that stromal composition may vary by extratumoral microenvironment subtype. Indeed, as shown in Figure 1, there were substantial differences in breast tissue composition by Active/Inactive subtype. These differences paralleled the differences in tissue composition by percentage mammographic density and dense area. The percentage of adipose tissue was significantly lower and the percentage of non-fatty stromal tissue was significantly higher among Inactive/mammographically dense patients (Supplementary Table 8). No differences were detected in the percentage of epithelial tissue by extratumoral microenvironment subtype or measures of mammographic density.

Finally, because Active/Inactive subtype and mammographic density appeared to be tracking cellular composition, the association between published gene expression signatures and cellular composition was also considered (Table 3). Again mirroring microenvironment subtype and mammographic density, the percentage of non-fatty stroma was associated with young-like gene expression, decreased expression of obesity-associated genes, decreased expression of a TGF- $\beta$  signature, and increased expression of estrogen response signatures. Epithelial composition was modestly associated with gene expression signatures for age and obesity (higher proportion in young-like and non-obese groups) after adjusting for BMI and age respectively (Supplementary Table 7).

## Discussion

By linking gene expression patterns of non-neoplastic breast tissue from breast cancer patients to mammographic density measures of their unaffected breast, we observed that distinct extratumoral microenvironments were independently associated with mammographic density. In particular, higher percentage density and dense area were related

to the Inactive subtype. Histologically we confirmed these findings and found that the Inactive subtype and higher mammographic density were also strongly and positively associated with the proportion of stromal composition. Further analysis of published gene expression signatures implicated estrogen response and TGF- $\beta$  signaling pathways, providing clues as to the biology underlying mammographic density, a strong and consistent risk factor for breast cancer.

Mammographic density is believed to reflect the combined effects of cell proliferation and genetic damage to proliferating epithelial cells (24). This hypothesis is supported by strong associations between mammographic density and exogenous estrogen exposure (25, 26), and in the current study, our data provide further molecular support for the role of estrogen in density. We observed strong estrogen-response gene expression among those with dense tissue and among those with the Inactive extratumoral microenvironment subtype. The strength of the molecular relationships is striking in light of recent findings showing relatively weak molecular evidence of estrogen response in dense breast tissue. Haakensen found a signature of 24 genes differentially expressed between high and low density breasts, including the gene coding for estrogen receptor (ESR1) and three uridine 5'-diphosphoglucuronosyltransferase (UGT) family genes (27) with postulated roles in protecting epithelial cells from genetic damage by local estrogen metabolites (27, 28). However, in a later study, Haakensen et al. observed that while serum estradiol was associated with mammographic density, none of the gene expression changes related to serum estradiol was significantly associated with mammographic density (29). It is possible that the lack of significant associations with estrogen response in previous studies is attributable to differences in the patient populations or due to the emphasis on fewer genes.

Interestingly, we also observed a strong association between estrogen response and stromal content (higher non-fatty stroma/lower breast adiposity) in non-neoplastic breast tissue, while the association between estrogen response and epithelial proportion was not significant. This null result for epithelium may be attributed to lower variation in epithelial percentage across patients (interquartile range for percent composition was 4.4-13.2%) and random measurement error in our estimation of epithelial percentage, but our molecular data and the published literature underscore the importance of stroma in mammographic density (30-32). Few gene expression studies have emphasized stromal biology in mammographic density, but one study by Yang et al. used canonical pathway analysis of MD-associated gene expression to identify decreased TGF- $\beta$  signaling in dense breast tissue (33). Our findings confirmed this association by showing the strong association between TGF- $\beta$  signature and adipose and stromal composition. TGF- $\beta$  is an important morphogen in normal mammary tissue, with inhibitory roles on proliferation of both ER-positive and ER-negative cells (34). Thus, the decreased TGF- $\beta$  signaling in dense tissue may be upstream of differences in epithelial proliferation or may reflect altered stromal-epithelial interaction during oncogenesis (35-37).

Previous models to interpret the association between mammographic density and breast cancer risk have focused on mitogens and mutagens, where luminal epithelial cells are central in interpreting the MD-breast cancer association. Although a breast tumor originates from epithelial cells, the extended model suggests a more dominant role of stromal



microenvironment and morphogenesis in tumorigenesis (38), which supports the important role of interactions between stroma and epithelial cells in malignancy transformation proposed in the previous studies (39). Consequently, studies focusing on the association between mammography and breast tissue composition have helped to improve our understanding of the MD-breast cancer association (30, 40-42). Given that the molecular findings of this study support and extend previously proposed mechanisms, we suggest a revised conceptual model for the MD-breast cancer association (Figure 2) (24). Indeed, other recent data also support the fundamental importance of microenvironment in density (43), and further suggest a critical interplay between mechanics, morphogenesis and malignancy (44-46). Mammographic density may be a powerful biological marker in its ability to comprehensively summarize the variation in mitogenesis, mutagenesis, and morphogenesis of breast.

The observation that a strong component of genomic variation in breast tissue is associated with a strong risk factor for breast cancer suggests that further investigation of normal tissue gene expression could yield novel insights about the biology of breast cancer risk and mammographic-density associated risk in particular. In fact, the normal tissue gene expression seems to reflect host factors and risk factor exposure more strongly than tumor characteristics; none of the tumor characteristics evaluated (ER status, grade, breast tumor subtype, size, etc.) showed significant associations with the Active/Inactive subtype. While there may be genes whose extratumoral expression is affected by tumor subtype, the set of genes that determine the major variation (and Active/Inactive subtype) are not tumor associated. Previous epidemiologic studies evaluating whether mammographic density is related to breast cancer characteristics are conflicting, with some suggesting high density is associated with more aggressive disease, and some speculating that observed associations between MD and tumor biology may be explained by delayed diagnosis among high density cases (i.e. masking bias) (47-50). Our current and previous research on Active/Inactive subtypes suggests no link with tumor characteristics, supporting the latter hypothesis, that mammographic density does not have distinct influences on etiology of subtypes.

The strong association between mammographic density and gene expression in the current study may be a result of looking at broad transcriptional changes rather than individual gene-level correlates of density. While a limited number (27, 33) of studies previously addressed genomic signatures of mammographic density, these studies identified very few density-associated genes (<100 genes), limiting their value in defining the biology of mammographic density. Weak associations with individual gene expression changes in previous studies may reflect low statistical power to detect a small differentially expressed signature (51, 52), but it is striking that such a strong breast cancer risk factor should produce such weak genomic expression, when other factors such as age and BMI have broad effects on normal tissue gene expression (20, 21). Another explanation for the strength of our pattern-focused analysis may be that mammographic density summarizes several biologically complex aspects of breast physiological and pathological conditions and itself may not represent a singular biological state. If this latter explanation were true, analysis starting from mammographic density would be poorly correlated with individual gene expression values, as gene-by-gene supervised analysis works best in homogeneously defined exposures. In contrast, Active/Inactive subtype summarizes the effect of various

factors influencing breast microenvironment (i.e. strong associations with age, obesity, TFG- $\beta$ , and E2 signatures) and therefore better captures genomic variation. Future research should test other biological themes for their association with mammographic density and Active/Inactive microenvironment.

Our results should be interpreted with consideration of several limitations. The non-neoplastic breast tissues used in the present study were taken from breast cancer patients. The influence of tumor on extratumoral microenvironment has been reported previously (5, 53) and we cannot exclude the possibility that the tissue of women with disease are systematically different from those without disease. However, as discussed above, the evidence that Active/Inactive microenvironment is not associated with tumor characteristics has led us to hypothesize that Active/Inactive subtype is a host-dependent factor. Caution is necessary when extrapolating these findings to the normal breast microenvironment of non-diseased women. Future research on normal breast tissues from healthy women is needed and will clarify the role of the Inactive microenvironment in disease development. Another key limitation is sample size. We cannot rule out the possibility that negative results for associations between extratumoral environment subtype and parity, menopause and MHT were a result of limited power. Besides inherited factors (54), other factors also play a role in mammographic density and would be expected to show associations with Active/Inactive subtype (55, 56). Indeed, age and BMI, the two strongest identified endogenous factors for mammographic density, showed substantial associations with Active/Inactive microenvironment (estimated ORs greater than 1.5). These associations were stronger when genomic surrogates for age and BMI, developed based on normal breast and therefore capturing underlying biological influence on local organs, were examined. We also note that we used mammograms of the unaffected breast to estimate density, avoiding the potential interference of tumor. While this may have introduced some measurement error, mammographic density is highly correlated within a woman and density in the unaffected breast has been found to be comparable to that in the affected breast (49, 56), so this is not expected to bias our results.

## Conclusion

In summary, we evaluated extratumoral microenvironment subtype for associations with mammographic density and breast tissue composition. Based on strong associations between the genomic subtypes and mammographic density, it is likely that Active/Inactive subtype is also associated with and reflective of breast cancer risk. Further research to better understand the molecular characteristics of mammographic density and microenvironment subtypes could identify pathways that are targetable in preventing mammographic-density associated risk.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors gratefully acknowledge Lineberger Comprehensive Cancer Center's Translational Pathology Laboratory for support of the image analysis work.

### Grant Support

This project was supported by grants from the National Cancer Institute (U01-ES019472 and R01-CA138255), a Breast SPORE (P50CA058223) Career Development Award to M.A.T., and a grant from the Avon Foundation. The authors gratefully acknowledge Lineberger Comprehensive Cancer Center's Translational Pathology Laboratory for support of the image analysis work. This work was supported (in part) by the Intramural Research Program of the National Institutes of Health, National Cancer Institute.

## References

1. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature*. 2000; 406(6797):747–52. [PubMed: 10963602]
2. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*. 2001; 98(19):10869–74. [PubMed: 11553815]
3. Cichon MA, Degnim AC, Visscher DW, Radisky DC. Microenvironmental influences that drive progression from benign breast disease to invasive breast cancer. *J Mammary Gland Biol Neoplasia*. 2010; 15(4):389–97. [PubMed: 21161341]
4. Bissell MJ, Radisky DC, Rizki A, Weaver VM, Petersen OW. The organizing principle: Microenvironmental influences in the normal and malignant breast. *Differentiation*. 2002; 70(9-10): 537–46. [PubMed: 12492495]
5. Troester MA, Lee MH, Carter M, Fan C, Cowan DW, Perez ER, et al. Activation of host wound responses in breast cancer microenvironment. *Clin Cancer Res*. 2009; 15(22):7020–8. [PubMed: 19887484]
6. Roman-Perez E, Casbas-Hernandez P, Pirone JR, Rein J, Carey LA, Lubet RA, et al. Gene expression in extratumoral microenvironment predicts clinical outcome in breast cancer patients. *Breast Cancer Res*. 2012; 14(2):R51. [PubMed: 22429463]
7. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med*. 2008; 14(5):518–27. [PubMed: 18438415]
8. Polyak K, Kalluri R. The role of the microenvironment in mammary gland development and cancer. *Cold Spring Harb Perspect Biol*. 2010; 2(11):a003244. [PubMed: 20591988]
9. Fukino K, Shen L, Matsumoto S, Morrison CD, Mutter GL, Eng C. Combined total genome loss of heterozygosity scan of breast cancer stroma and epithelium reveals multiplicity of stromal targets. *Cancer Res*. 2004; 64(20):7231–6. [PubMed: 15492239]
10. Weber F, Shen L, Fukino K, Patocs A, Mutter GL, Caldes T, et al. Total-genome analysis of BRCA1/2-related invasive carcinomas of the breast identifies tumor stroma as potential landscaper for neoplastic initiation. *Am J Hum Genet*. 2006; 78(6):961–72. [PubMed: 16685647]
11. Fukino K, Shen L, Patocs A, Mutter GL, Eng C. Genomic instability within tumor stroma and clinicopathological characteristics of sporadic primary invasive breast carcinoma. *JAMA*. 2007; 297(19):2103–11. [PubMed: 17507346]
12. Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res*. 2009; 11(1):R7. [PubMed: 19187537]
13. Schedin P, Elias A. Multistep tumorigenesis and the microenvironment. *Breast Cancer Res*. 2004; 6(2):93–101. [PubMed: 14979914]
14. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med*. 2011; 17(3):320–9. [PubMed: 21383745]
15. Garcia-Closas M, Brinton LA, Lissowska J, Chatterjee N, Peplonska B, Anderson WF, et al. Established breast cancer risk factors by clinically important tumour characteristics. *Br J Cancer*. 2006; 95(1):123–9. [PubMed: 16755295]

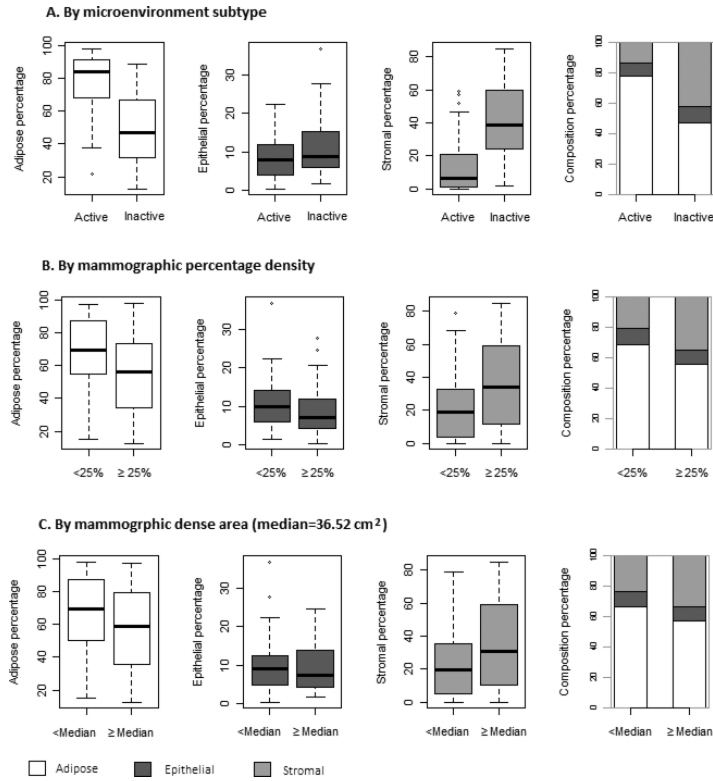
16. Camp JT, Elloumi F, Roman-Perez E, Rein J, Stewart DA, Harrell JC, et al. Interactions with fibroblasts are distinct in basal-like and luminal breast cancers. *Mol Cancer Res*. 2011; 9(1):3–13. [PubMed: 21131600]
17. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. *Phys Med Biol*. 1994; 39(10):1629–38. [PubMed: 15551535]
18. Wilkerson MD, Hayes DN. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010; 26(12):1572–3. [PubMed: 20427518]
19. Creighton CJ, Casa A, Lazard Z, Huang S, Tsimelzon A, Hilsenbeck SG, et al. Insulin-like growth factor-I activates gene transcription programs strongly associated with poor breast cancer prognosis. *J Clin Oncol*. 2008; 26(25):4078–85. [PubMed: 18757322]
20. Pirone JR, D'arcy M, Stewart DA, Hines WC, Johnson M, Gould MN, et al. Age-associated gene expression in normal breast tissue mirrors qualitative age-at-incidence patterns for breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2012
21. Sun X, Casbas-Hernandez P, Bigelow C, Makowski L, Joseph Jerry D, Smith Schneider S, et al. Normal breast tissue of obese women is enriched for macrophage markers and macrophage-associated gene expression. *Breast Cancer Res Treat*. 2012; 131(3):1003–12. [PubMed: 22002519]
22. Beck AH, Espinosa I, Gilks CB, van de Rijn M, West RB. The fibromatosis signature defines a robust stromal response in breast carcinoma. *Lab Invest*. 2008; 88(6):591–601. [PubMed: 18414401]
23. Oh DS, Troester MA, Usary J, Hu Z, He X, Fan C, et al. Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers. *J Clin Oncol*. 2006; 24(11):1656–64. [PubMed: 16505416]
24. Martin LJ, Boyd NF. Mammographic density: potential mechanisms of breast cancer risk associated with mammographic density: Hypotheses based on epidemiological evidence. *Breast Cancer Res*. 2008; 10(1):201. [PubMed: 18226174]
25. Chlebowski RT, Hendrix SL, Langer RD, Stefanick ML, Gass M, Lane D, et al. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: The women's health initiative randomized trial. *JAMA*. 2003; 289(24):3243–53. [PubMed: 12824205]
26. Kerlikowske K, Cook AJ, Buist DS, Cummings SR, Vachon C, Vacek P, et al. Breast cancer risk by breast density, menopause, and postmenopausal hormone therapy use. *J Clin Oncol*. 2010; 28(24):3830–7. [PubMed: 20644098]
27. Haakensen VD, Biong M, Lingjaerde OC, Holmen MM, Frantzen JO, Chen Y, et al. Expression levels of uridine 5'-diphospho-glucuronosyltransferase genes in breast tissue from healthy women are associated with mammographic density. *Breast Cancer Res*. 2010; 12(4):R65. [PubMed: 20799965]
28. Guillemette C, Belanger A, Lepine J. Metabolic inactivation of estrogens in breast tissue by UDP-glucuronosyltransferase enzymes: An overview. *Breast Cancer Res*. 2004; 6(6):246–54. [PubMed: 15535854]
29. Haakensen VD, Bjoro T, Luders T, Riis M, Bukholm IK, Kristensen VN, et al. Serum estradiol levels associated with specific gene expression patterns in normal breast tissue and in breast carcinomas. *BMC Cancer*. 2011; 11:332. [PubMed: 21812955]
30. Li T, Sun L, Miller N, Nicklee T, Woo J, Hulse-Smith L, et al. The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2005; 14(2):343–9. [PubMed: 15734956]
31. Yamaguchi Y. Microenvironmental regulation of estrogen signals in breast cancer. *Breast Cancer*. 2007; 14(2):175–81. [PubMed: 17485903]
32. Miki Y, Suzuki T, Sasano H. Controversies of aromatase localization in human breast cancer--stromal versus parenchymal cells. *J Steroid Biochem Mol Biol*. 2007; 106(1-5):97–101. [PubMed: 17624762]
33. Yang WT, Lewis MT, Hess K, Wong H, Tsimelzon A, Karadag N, et al. Decreased TGFbeta signaling and increased COX2 expression in high risk women with increased mammographic breast density. *Breast Cancer Res Treat*. 2010; 119(2):305–14. [PubMed: 19241157]

34. Moses H, Barcellos-Hoff MH. TGF-beta biology in mammary development and breast cancer. *Cold Spring Harb Perspect Biol.* 2011; 3(1):a003277. [PubMed: 20810549]
35. Kuperwasser C, Chavarria T, Wu M, Magrane G, Gray JW, Carey L, et al. Reconstruction of functionally normal and malignant human breast tissues in mice. *Proc Natl Acad Sci U S A.* 2004; 101(14):4966–71. [PubMed: 15051869]
36. Bhowmick NA, Chytil A, Plieth D, Gorska AE, Dumont N, Shappell S, et al. TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. *Science.* 2004; 303(5659): 848–51. [PubMed: 14764882]
37. Daly AC, Randall RA, Hill CS. Transforming growth factor beta-induced Smad1/5 phosphorylation in epithelial cells is mediated by novel receptor complexes and is essential for anchorage-independent growth. *Mol Cell Biol.* 2008; 28(22):6889–902. [PubMed: 18794361]
38. Potter JD. Morphogens, morphostats, microarchitecture and malignancy. *Nat Rev Cancer.* 2007; 7(6):464–74. [PubMed: 17522715]
39. Maller O, Martinson H, Schedin P. Extracellular matrix composition reveals complex and dynamic stromal-epithelial interactions in the mammary gland. *J Mammary Gland Biol Neoplasia.* 2010; 15(3):301–18. [PubMed: 20811805]
40. Warren R, Lakhani SR. Can the stroma provide the clue to the cellular basis for mammographic density? *Breast Cancer Res.* 2003; 5(5):225–7. [PubMed: 12927028]
41. Alowami S, Troup S, Al-Haddad S, Kirkpatrick I, Watson PH. Mammographic density is related to stroma and stromal proteoglycan expression. *Breast Cancer Res.* 2003; 5(5):R129–35. [PubMed: 12927043]
42. Lin SJ, Cawson J, Hill P, Haviv I, Jenkins M, Hopper JL, et al. Image-guided sampling reveals increased stroma and lower glandular complexity in mammographically dense breast tissue. *Breast Cancer Res Treat.* 2011; 128(2):505–16. [PubMed: 21258862]
43. DeFilippis RA, Chang H, Dumont N, Rabban JT, Chen YY, Fontenay GV, et al. CD36 repression activates a multicellular stromal program shared by high mammographic density and tumor tissues. *Cancer Discov.* 2012; 2(9):826–39. [PubMed: 2277768]
44. Levental KR, Yu H, Kass L, Lakins JN, Egeblad M, Erler JT, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell.* 2009; 139(5):891–906. [PubMed: 19931152]
45. Paszek MJ, Weaver VM. The tension mounts: Mechanics meets morphogenesis and malignancy. *J Mammary Gland Biol Neoplasia.* 2004; 9(4):325–42. [PubMed: 15838603]
46. Butcher DT, Alliston T, Weaver VM. A tense situation: Forcing tumour progression. *Nat Rev Cancer.* 2009; 9(2):108–22. [PubMed: 19165226]
47. Eriksson L, Czene K, Rosenberg L, Humphreys K, Hall P. The influence of mammographic density on breast tumor characteristics. *Breast Cancer Res Treat.* 2012
48. Ding J, Warren R, Girling A, Thompson D, Easton D. Mammographic density, estrogen receptor status and other breast cancer tumor characteristics. *Breast J.* 2010; 16(3):279–89. [PubMed: 20408823]
49. Aiello EJ, Buist DS, White E, Porter PL. Association between mammographic breast density and breast cancer tumor characteristics. *Cancer Epidemiol Biomarkers Prev.* 2005; 14(3):662–8. [PubMed: 15767347]
50. Gierach GL, Ichikawa L, Kerlikowske K, Brinton LA, Farhat GN, Vacek PM, et al. Relationship between mammographic density and breast cancer death in the breast cancer surveillance consortium. *J Natl Cancer Inst.* 2012; 104(16):1218–27. [PubMed: 22911616]
51. Hackstadt AJ, Hess AM. Filtering for increased power for microarray data analysis. *BMC Bioinformatics.* 2009; 10:11. [PubMed: 19133141]
52. Bourgon R, Gentleman R, Huber W. Independent filtering increases detection power for high-throughput experiments. *Proc Natl Acad Sci U S A.* 2010; 107(21):9546–51. [PubMed: 20460310]
53. Graham K, Ge X, de Las Morenas A, Tripathi A, Rosenberg CL. Gene expression profiles of estrogen receptor-positive and estrogen receptor-negative breast cancers are detectable in histologically normal breast epithelium. *Clin Cancer Res.* 2011; 17(2):236–46. [PubMed: 21059815]

54. Boyd NF, Dite GS, Stone J, Gunasekara A, English DR, McCredie MR, et al. Heritability of mammographic density, a risk factor for breast cancer. *N Engl J Med.* 2002; 347(12):886–94. [PubMed: 12239257]
55. Butler LM, Gold EB, Greendale GA, Crandall CJ, Modugno F, Oestreicher N, et al. Menstrual and reproductive factors in relation to mammographic density: The study of women's health across the nation (SWAN). *Breast Cancer Res Treat.* 2008; 112(1):165–74. [PubMed: 18066689]
56. Assi V, Warwick J, Cuzick J, Duffy SW. Clinical and epidemiological issues in mammographic density. *Nat Rev Clin Oncol.* 2011; 9(1):33–40. [PubMed: 22143145]

### Translational Relevance

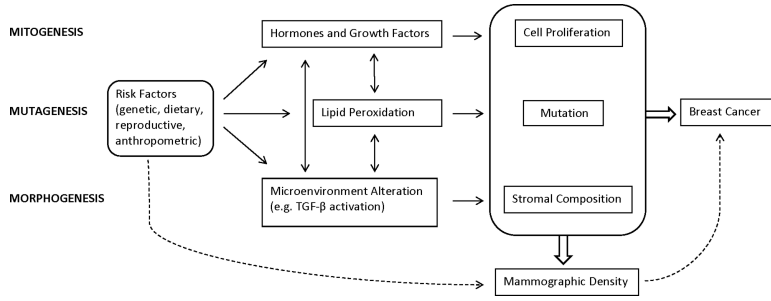
Mammographic density is the strongest risk factor for non-familial breast cancer among women apart from older age, but its mechanistic underpinnings are poorly understood. We recently reported two distinct molecular subtypes of normal tissue in breast cancer patients. We hypothesized that mammographic density would be associated with these subtypes based on their defining molecular pathways (e.g. fibrosis and cell-adhesion). Our results show that these well-defined molecular subtypes of normal tissue are strongly associated with both mammographic density and breast tissue composition, establishing novel molecular correlates of mammographic density. Many of the pathways enriched in patients with higher mammographic density are targetable, raising the possibility of developing prevention strategies for mitigating density-associated breast cancer risk.



**Figure 1. Breast tissue composition distribution by the Active/Inactive subtype and mammographic density**

Using quantitative analysis of digital histology images, each sample from the PWBCS for which gene expression was measured also was assessed for composition. Non-fatty stromal percentage was higher in Inactive subtype samples (A) and women with higher mammographic density (percentage, B; dense area, C). In contrast, adipose concentration was lower in the Inactive/higher density patients. Epithelial content did not vary significantly by any of these variables.





**Figure 2. Conceptual model for potential mechanisms of the association between mammographic density and breast cancer risk with a focus on the role of stroma**

This figure is adapted from the biological model proposed by Martin et al. (24), showing the underlying biological processes linking risk factors to breast cancer and suggesting the surrogate role of mammographic density in these settings. The genetic and environmental factors not only influence epithelial cell proliferation and induce genetic damage, but also regulate the surrounding microenvironment and stromal composition. The abnormalities in mitogenesis, mutagenesis, and morphogenesis result in breast cancer development. By comprehensively summarizing the variation in these three aspects, mammographic density works as an intermediate phenotype strongly indicating breast cancer risk.

**Table 1**

Participant characteristics by the Active/Inactive extratumoral microenvironment subtype

Characteristics	Total (n=121) n (%)	Active (n=59) n (%)	Inactive (n=62) n (%)	P <sup>*</sup>
Age at diagnosis				
Mean (SD)	55.26 (9.78)	56.03 (8.72)	54.52 (10.70)	0.39
<50 y	33 (27)	12 (20)	21 (34)	0.095
50 y	88 (73)	47 (80)	41 (66)	
BMI				
Mean (SD)	27.69 (5.30)	28.49 (5.48)	26.93 (5.06)	0.10
<30 kg/m <sup>2</sup>	87 (72)	39 (66)	48 (77)	0.17
30 kg/m <sup>2</sup>	34 (28)	20 (34)	14 (23)	
Age at menarche				
<12y	40 (33)	18 (31)	22 (36)	0.50
12 y	80 (67)	41 (69)	39 (64)	
Parity				
0	26 (21)	11 (19)	15 (24)	0.46
1	95 (79)	48 (81)	47 (76)	
Age at 1 <sup>st</sup> full term birth in parous women				
<25 y	52 (55)	25 (52)	27 (57)	0.60
25 y	43 (45)	23 (48)	20 (43)	
Use of oral MHT				
Never	80 (68)	37 (65)	43 (70)	0.52
Ever	38 (32)	20 (35)	18 (30)	
Menopausal status				
Premenopausal	34 (28)	14 (24)	20 (32)	0.30
Postmenopausal	87 (72)	45 (76)	42 (68)	
Family history of breast cancer <sup>†</sup>				
No	101 (84)	48 (83)	53 (85)	0.68
Yes	19 (16)	10 (17)	9 (15)	
Previous benign breast disease				
No	89 (75)	45 (79)	44 (72)	0.39
Yes	29 (25)	12 (21)	17 (28)	
Percentage density				
Mean (SD)	27.25 (16.00)	22.56 (13.90)	31.70 (16.69)	0.0013
<25%	57 (47)	35 (59)	22 (35)	0.0086
25%	64 (53)	24 (41)	40 (65)	
Dense area (cm <sup>2</sup> )				
Mean (SD)	37.70 (24.59)	33.21 (23.08)	41.97 (25.40)	0.049
<median <sup>‡</sup>	60 (50)	33 (56)	27 (44)	0.17
median	61 (50)	26 (44)	35 (56)	

Abbreviation: BMI, body mass index; SD, standard deviation; MHT, menopausal hormone therapy.

\* *P* values for the comparisons between Active and Inactive subtypes were calculated by t-test for continuous variables and Chi-square test for categorical variables. Missing values were excluded from percentage calculations.

† Family history of breast cancer in first degree relative.

‡ Median=36.52 cm<sup>2</sup>.

**Table 2**

Multivariable analysis of the association between mammographic density measures and the Active/Inactive extratumoral microenvironment subtype

Model	Mammographic density measures	Active (n=59) mean (SE)	Inactive (n=62) mean (SE)	<i>P</i> <sup>*</sup>
Crude mean	Dense area (cm <sup>2</sup> )	33.21 (3.16)	41.97 (3.09)	0.049
	Percentage density (%)	22.56(2.00)	31.70(1.95)	0.0014
Age-adjusted mean	Dense area (cm <sup>2</sup> )	33.84 (3.01)	41.37 (2.94)	0.076
	Percentage density (%)	23.20 (1.71)	31.09 (1.67)	0.0013
BMI-adjusted mean	Dense area (cm <sup>2</sup> )	33.17 (3.19)	42.01 (3.12)	0.051
	Percentage density (%)	23.60 (1.81)	30.71 (1.77)	0.0060
Age and BMI-adjusted mean	Dense area (cm <sup>2</sup> )	33.41 (3.01)	41.78 (2.93)	0.049
	Percentage density (%)	23.77 (1.63)	30.56 (1.59)	0.0038

Abbreviation: BMI, body mass index; SE, standard error

\* *P* values were calculated by t-test for univariate analysis and analysis of covariance (ANCOVA) for multivariable analysis. Age and BMI were used as continuous variables.

Table 3

Relationship of biologically relevant signatures with the Active/Inactive extratumoral microenvironment subtype, mammographic density, and breast composition

Signature (# genes used/# in signature)	Signature (# genes used/# in signature)	Microenvironment Subtype* n (%)		Percentage Density* n (%)		Breast Composition <sup>†</sup> mean (SE)		
		Active (n=59)	Inactive (n=62)	<25% (n=57)	25% (n=64)	Adipose (%)	Epithelium (%)	Stroma (%)
Age (779/802)	Old	56(95)	9 (15)	36(63)	29 (45)	74.52 (2.64)	8.74 (0.87)	16.74 (2.56)
	Young	3 (5)	53 (85)	21 (37)	35 (55)	47.99 (2.77)	10.85 (0.82)	41.15 (2.69)
	<i>P</i>	6.2E-21			0.049	3.6E-10	0.079	3.1E-09
BMI (677/760)	Obese	56 (95)	4 (6)	35 (61)	25 (39)	78.07 (2.52)	8.55 (0.86)	13.38 (2.46)
	Non-obese	3 (5)	58 (94)	22 (39)	39 (61)	46.86 (2.44)	10.85 (0.83)	42.29 (2.37)
	<i>P</i>	1.0E-26			0.014	7.3E-15	0.055	8.4E-14
DTF (581/758)	Positive	31 (53)	27 (44)	36 (63)	22 (34)	65.37 (3.27)	10.22 (0.88)	24.41 (3.11)
	Negative	28 (47)	35 (56)	21 (37)	42 (66)	58.83 (3.11)	9.31 (0.83)	31.87 (2.96)
	<i>P</i>	0.32			0.0016	0.14	0.45	0.079
TGF-β (210/234)	Positive	53 (90)	12 (19)	37 (65)	28 (44)	74.34 (2.61)	9.10 (0.82)	16.55 (2.50)
	Negative	6 (10)	50 (81)	20 (35)	36 (56)	47.72 (2.80)	10.47 (0.88)	41.81 (2.68)
	<i>P</i>	7.8E-15			0.020	2.7E-10	0.26	8.0E-10
E2 (711/754)	Positive	10 (17)	49 (79)	21 (37)	38 (59)	49.36 (2.80)	10.63 (0.85)	40.00 (2.70)
	Negative	49 (83)	13 (21)	36 (63)	26 (41)	74.08 (2.75)	8.88 (0.84)	17.04 (2.65)
	<i>P</i>	8.5E-12			0.013	5.3E-09	0.15	1.7E-08

Abbreviation: BMI, body mass index; SE, standard error; DTF, desmoid-type fibrosis; TGF, transforming growth factor; E2, 17β-estradiol.

\* *P* values were calculated by t-test for continuous variables and Chi-square test for categorical variables except that when expected cell count was less than 5, they were calculated by Fisher exact test.

<sup>†</sup>The sample size was 118 due to the missing of breast composition information.