



Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2014 December ; 23(12): 2810–2818. doi:
10.1158/1055-9965.EPI-14-0507.

Benign Breast Tissue Composition in Breast Cancer Patients: Association with Risk Factors, Clinical Variables, and Gene Expression

Xuezheng Sun¹, Rupninder Sandhu², Jonine D. Figueroa³, Gretchen L. Gierach³, Mark E. Sherman^{3,4}, and Melissa A. Troester^{1,2,*}

¹Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

²Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

³Hormonal and Reproductive Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA

⁴Breast and Gynecologic Cancer Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, MD, USA

Abstract

Introduction—Breast tissue composition (epithelium, non-fatty stroma, and adipose) changes qualitatively and quantitatively throughout the lifespan, and may mediate relationships between risk factors and breast cancer initiation. We sought to identify relationships between tissue composition, risk factors, tumor characteristics, and gene expression.

Methods—Participants were 146 patients from the Polish Breast Cancer Study, with data on risk factor and clinicopathological characteristics. Benign breast tissue composition was evaluated using digital image analysis of histological sections. Whole genome microarrays were performed on the same tissue blocks.

Results—Mean epithelial, non-fatty stromal, and adipose proportions were 8.4% (SD=4.9%), 27.7% (SD=24.0%), and 64.0% (SD=24.0%), respectively. Among women < 50 years old, stroma proportion decreased and adipose proportion increased with age, with approximately 2% difference per year ($p < 0.01$). The variation in epithelial proportion with age was modest (0.1% per year). Higher epithelial proportion was associated with obesity (7.6% in non-obese vs 10.1% in obese; $p=0.02$) and with poorly differentiated tumors (7.8% in well/moderate vs 9.9% in poor; $p=0.05$). Gene expression signatures associated with epithelial and stromal proportion were identified and validated. Stroma-associated genes were in metabolism and stem cell maintenance pathways, while epithelial genes were enriched for cytokine and immune response pathways.

Conclusions—Breast tissue composition was associated with age, BMI, and tumor grade, with consequences for breast gene expression.

*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

Impact—Breast tissue morphologic factors may influence breast cancer etiology. Composition and gene expression may act as biomarkers of breast cancer risk and progression.

Keywords

breast tissue composition; breast cancer risk factor; tumor characteristics; gene expression

Introduction

Benign breast tissue is comprised of epithelium, stroma, and adipose tissue, each of which is heterogeneous and complex in composition. Previous papers have suggested that during critical developmental windows, such as embryogenesis, pregnancy, and postmenopausal involution, important shifts occur in normal breast tissue composition (1, 2). Epithelial proportions increase with hormonal exposure, and stromal composition may also increase (3–5). As malignant transformation evolves, composition of benign breast tissues may change in parallel (6–9), exerting selective forces that influence tumor biology and behavior (10, 11). It is important to understand how the balance of epithelium and stroma is modulated by normal events throughout the lifespan, and how this balance contributes to carcinogenesis. The importance of understanding factors that influence benign breast tissue composition is underscored by the strong direct association between mammographic density (a radiological measure of fibroglandular breast tissue content) and breast cancer risk and by growing experimental evidence that the tumor microenvironment may influence breast cancer biology, prognosis, and treatment response (12–15).

The epidemiologic and clinical factors that affect benign breast tissue can be partly understood by examining mammographic density. Specifically, increasing age, menopause, elevated body mass index, and having a live birth are associated with lower mammographic density, suggesting less fibroglandular tissue relative to adipose tissue. Family history of breast cancer and use of menopausal hormones are related to higher mammographic density, and therefore greater fibroglandular content. In benign breast tissue of younger women, stroma accounts for the majority of mammographically dense tissue, but epithelial content may be an important driver of mammographic density-risk associations. Few studies have directly parsed density into stroma vs. epithelium, and even fewer have evaluated all three components (epithelium, fat, and stroma) in association with risk factors (16–18). Such studies could lead to improved understanding of mammographic density as a risk factor and may identify novel risk biomarkers.

To determine factors related to breast histology and to identify plausible biological pathways affected by changing breast tissue composition, we used breast tissues and gene expression data from the Polish Breast Cancer Study. Our aims were to determine the relationships of breast tissue composition to breast cancer risk factors and tumor characteristics, and to relate different tissue components (epithelial, non-fatty stroma and adipose) to gene expression profiles.

Materials and Methods

Patient samples

The study population is a subset of patients from the Polish Breast Cancer Study (PBCS), a population-based case-control study conducted in two large cities in Poland (Warsaw and Łódź) during 2000–2003 (19). PBCS cases were women aged 20–74 years with an incident, 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. Based on *in vitro* evidence that basal-like and luminal tumors are associated with distinctive microenvironments (20), we oversampled these subtypes. 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.

We limited our analysis to 146 invasive breast cancer cases in Warsaw with available data on breast composition and gene expression of extratumoral breast tissues. Our interest was to evaluate how normal tissue without evidence of pathologic change may vary in association with risk factors. Considering that epithelial tissue is a minority component (typically <10% by area) in normal breast (16), 11 women with outlying values suggestive of possible hyperplasia (epithelial proportion greater than 20%) were excluded as potentially non-normal. We further excluded women with incomplete data on breast risk factors and clinicopathological factors, which left 96 women in our main analysis. The 50 women with incomplete risk factor data were included as a validation data set.

Data collection

Information on demographic and anthropometric factors was collected from medical records and in-person interviews (19). Mammographic density of pre-treatment mammograms of the unaffected breast was assessed using a quantitative, reliable computer-assisted thresholding method (21). One expert reader measured absolute dense area (cm²) and total breast area (cm²); percentage mammographic density was calculated by dividing the dense breast area by the total breast area and multiplying by 100. Frozen non-neoplastic breast specimens of approximately 100 mg were cut over dry ice and then used to prepare 20 μm frozen sections from each end. The central portion was used for RNA extraction. Microarrays on the central section were performed using two-color 4X44K Agilent whole genome arrays. More information on RNA isolation, microarray, and data preprocessing has also been described previously (18). Microarray data used in this analysis are publicly available through the Gene Expression Omnibus (GSE49175).

After H&E staining, the sections were scanned into high-resolution (20X) digital images using the Aperio Scan-Scope XT Slide Scanner (Aperio Technologies, Vista, CA, USA) in the Translational Pathology Laboratory of University of North Carolina at Chapel Hill. The details of composition measurement and measurement validation have been described previously (18). Briefly, 15 representative digital slides were firstly manually annotated for epithelial area, non-fatty stromal area, and total area (mm²) using Aperio ImageScope

software. These annotations were used to train the composition estimator in Aperio's Genie Classifier. The trained classifier was positively and strongly correlated with manual annotation, and showed better performance in quantifying small percentages over visual assessment (by human eye on regular H&E slides). After training and validation, Aperio's Genie Classifier was applied to the remaining slides partitioning epithelium, adipose tissue, non-fatty stroma, and glass into percentages.

Statistical analysis

The distribution of adipose, non-fatty stroma, and epithelial composition and the Pearson's correlation between components were evaluated using R (version 3.0.0). To assess the influence of age on breast composition, we evaluated the distribution of composition by five age groups at diagnosis (<40 y, 40–49 y, 50–59 y, 60–69 y, and 70 y) and employed a cubic spline interpolation algorithm (knots=3) to model and visualize these associations (22). As stromal and adipose proportions were strongly inversely correlated ($\gamma=-0.98$, $p<0.001$), we only studied epithelial and stromal compositions in subsequent analyses. To study how composition changed in one compartment (e.g. epithelium) relative to another (e.g. stroma), the ratios of epithelial-to-stromal proportions and stromal-to-adipose proportions were also evaluated. Log-transformed ratios were used in the analysis to account for substantial skew.

Besides age, other breast cancer risk factors included percentage density, dense area, BMI, use of oral menopausal hormone therapy, menopausal status, family history of breast cancer in a first degree relative, previous benign breast disease history (defined as ever having any biopsy or partial removal procedure, with no cancer detected), age at menarche, and parity. Age at first birth and duration of lactation were assessed among parous women. Time since last birth could not be evaluated because approximately 95% patients had their most recent birth more than 15 years before breast cancer diagnosis. Tumor clinicopathological factors evaluated included tumor size, histological type, differentiation stage, number of positive axillary lymph nodes, and ER status. The age-adjusted differences in epithelial and stromal compositions according to breast cancer risk factors and tumor characteristics were compared using Wald test.

Linear regression with Linear Models for Microarray Data (LIMMA) (23) was used to identify genes associated with stromal and epithelial composition. Both epithelial and stromal proportions were included in the model as independent variables. The false discovery rate (FDR) approach was used to control for multiple comparisons, with 0.05 as the cutoff for statistical significance. Previous analyses showed distinct microenvironments associated with breast tumor subtypes (20, 24), so ER-stratified analysis were also performed.

To validate the signatures identified, a test data set ($n=50$, described above) was used. Principle component analysis was used to assess global gene expression in the test and training data sets, and showed considerable overlap (Supplementary Figure 1). To evaluate relative expression of a signature in a test set sample, the $\log_2(R/G)$ expression for each gene was multiplied by corresponding direction index (1 for up-regulated genes, and -1 for down-regulated gene) and results were summed across all genes. This score was also

evaluated for association with cellular composition of test samples. Although training and test datasets showed considerable similarity in demographic, anthropometric factors, and tumor characteristics (except that mammographic density of women in test dataset was slightly higher compared with women in training dataset), we performed a sensitivity analysis including the 50 subjects, to evaluate the impact of exclusion from the training dataset due to missing values. The magnitude of associations observed in the training dataset was not meaningfully changed by including these patients.

Results

Associations between histology and risk factors/clinicopathological characteristics

The analysis used histology data from 96 patients with complete data to assess risk factor associations. In this sample, mean epithelial proportion was 8.4% (SD=4.9%; range=0.5–19.4%), mean stromal proportion was 27.7% (SD=24.0%; range=0.02–85.0%), and mean adipose proportion was 64.0% (SD=24.0%; range=12.5–98.5%). While epithelial proportion was independent of the other two components (non-fatty stroma: $\gamma=-0.09$, $p=0.40$; adiposity: $\gamma=-0.12$, $p=0.26$), stromal and adipose proportions were strongly inversely correlated ($\gamma=-0.98$, $p<0.001$).

Among the risk factors examined, age showed the strongest association with breast composition. Interestingly, the association was not strictly monotonic linear, but had a turning point around age 50 (Figure 1). Adipose proportion significantly increased with age until age 50, with a 2.0% increment per year ($p=0.004$). After age 50, the rate of change in adipose proportion decreased to 0.1% per year ($p=0.800$). Conversely, stromal proportion decreased with age at a rate of -1.9% per year ($p=0.006$), with an inflection after 50 years of age, and a rate of -0.06% per year ($p=0.864$) thereafter. The variation in epithelial composition with age was relatively small, both before and after age 50, but a slight U-shape curve was observed, with declines in epithelial proportion prior to age 50, followed by a modest increase in proportion thereafter (< 50y: 0.11% decrement per year, $p=0.426$; 50y: 0.14% increment per year, $p=0.042$). The ratios of epithelium to non-fatty stroma and non-fatty stroma to adipose were also assessed (Supplementary Figure 2). For the ratio of epithelium to non-fatty stroma, we observed a slight but insignificant increase with age. The ratio of stromal-to-adipose proportion had an association with age that was very similar to that of stromal proportion, decreasing till age 50 years and then leveling off.

Given the high inverse correlation between stromal and adipose proportions, subsequent analyses focused on epithelium and non-fatty stroma, and did not separately assess associations with adipose proportion. However, the adipose associations are presented in Supplementary Table 1. Factors associated with breast composition after adjustment for age, included BMI and mammographic density measures (Table 1). We evaluated potential confounding of risk factor-composition relationships by BMI, and found no substantial confounding except for in models with mammographic density, and therefore reported the parsimonious model without BMI for all other associations. Women with higher stromal proportion were more likely to have higher percentage density and dense area, but were less likely to be obese. Interestingly, epithelial composition showed a positive association with BMI, but was suggested to be inversely associated with breast density. Although not

statistically significant, stromal but not epithelial content, was lower in parous compared with nulliparous women ($p=0.06$). We did not detect any significant associations of breast composition with tumor characteristics, except that higher epithelial proportion was associated with poorly differentiated tumors (Table 2).

Association between histology and gene expression

To identify gene expression that varied with breast composition on 96 women in the training sample with complete risk factor data, supervised analyses were performed. We detected an epithelial-content-associated signature ($n=716$ genes, $FDR<0.05$ and $\beta>0.02$), with approximately half of the genes up-regulated ($n=332$) (Supplementary Table 2). Pathway analysis showed that up-regulated genes included cytokine/chemokine activity, EIF2 signaling, and mTOR signaling, and suggested biological functions related to epithelial regulation and immune-related processes (Supplementary Table 3). No significant pathways (defined as Benjamini-Hochberg adjusted p -value <0.05) were detected among the down-regulated genes.

Stromal composition was associated with a broad biological profile, with the corresponding signature including 1065 genes (484 down-regulated and 581 up-regulated genes) defined by $FDR<0.05$ and $\beta>0.02$ (Supplementary Table 2). The up-regulated pathways were involved in regulating vascular permeability, epithelial to mesenchymal transition, and maintenance of breast cancer stem cells in self renewal state (Supplementary Table 3). The down-regulated pathways included metabolic and signaling pathways. The stromal and epithelial signatures were relatively independent, with less than one third of genes being identified in both signatures and only cell-sertoli cell junction signaling pathways being identified in both gene sets.

Given previous evidence of expression differences (20, 24–26) and our own findings suggesting epithelial content differences in cancer-adjacent normal by tumor ER status (27), we also performed analysis of gene expression stratified by ER status. In ER-positive patients, the stromal and epithelial signatures were similar to the signatures detected in unstratified analysis, with considerable overlap (stromal signature, 97%; epithelial signature, 65%). In ER-negative patients, no epithelial signature was detected. However, ER-negative stroma had a unique signature, with 425 genes identified (up=190; down=235) (Supplementary Table 2) that were not captured in unstratified/overall analysis. Pathway analysis showed these genes are involved in carcinogenesis and tumorigenesis, for example, actin cytoskeletal signaling, Wnt/beta-catenin signaling and leukocyte extravasation signaling (Supplementary Table 3).

We validated the composition-associated signatures in an independent subsample data ($n=50$) from the PBCS. As shown in Figure 2A, the average gene expression level of our stromal content signature increased with actual stromal proportion (p -value for trend <0.001 , age-adjusted). The high inverse correlation between non-fatty stroma and adipose compositions was also reflected at the gene expression level, where the expression of the stromal signature decreased with actual adipose proportion (Figure 2B, p -value for trend <0.001 , age-adjusted). Similarly, expression level of epithelial signature showed positive associations with epithelial proportion (Figure 2C, p -value for trend <0.001 , age-adjusted).

Thus the signatures identified in the training dataset showed strong relationships with composition in the independent, validation data.

Discussion

Benign breast tissue composition is responsive to exogenous exposures, endogenous physiological effects and pathological events. Understanding the factors that influence composition is important for understanding breast cancer etiology, risk, and progression. Our results showed that epithelial and stromal proportions were associated with well-recognized breast cancer risk factors (such as age, BMI, and mammographic density) and with tumor differentiation. The strongest associations were between age and breast composition. While the effect of age on breast epithelial content is well-recognized (16, 17, 28, 29), we also observed significant differences in stromal content. The association between age and tissue composition mirrors associations with percentage density (30), with neither appearing strictly monotonic, but both suggesting a turning point at age 50.

The observation that stroma is slowly displaced by fat even prior to menopause has potential implications for breast tumorigenesis broadly and for epidemiologic studies of breast cancer risk. It is increasingly recognized that stroma plays an important role in breast cancer progression (31–33). It has also been shown that expression of adipokines may be altered with obesity and in the peritumoral tissue of advanced breast tumors, suggesting a role for breast fat and adipokines in tumorigenesis (34). However, the impact of changing premenopausal breast tissue composition on risk is largely unexplored. Given the degree of changes in the breast microenvironment that occur prior to menopause, future studies should consider stromal composition as a possible modifier of risk among premenopausal women. Measures of microenvironment composition are not widely available, however mammographic density is increasingly available in epidemiologic studies and may be not only a surrogate marker (35), but also a risk modifier for some carcinogenic exposures.

It is interesting that the premenopausal stromal changes were much more pronounced in our study than the age-related changes in epithelium. In contrast to previous findings (16), we did not observe statistically significant age-associated differences in epithelial proportion among younger (age < 50 years) women. Differences in the age-epithelial content association between our study and Nurse's Health Study's (NHS) benign breast disease cohort (16) may be due to the older age of women in our dataset (percentage of population < 50 years: 22% vs. 68% in NHS). That is, given that our study included fewer younger women (where the effects of age are most pronounced), we were underpowered to detect age-related declines in epithelial content. However, our data including more older women yielded interesting novel insights, including suggestions of a modest increase in epithelial proportion after age 50. As women age, terminal duct lobular units (TDLUs) are well known to regress, likely due to hormonal changes; however, large ducts are believed to change less under the influence of age (27, 36). This could account for non-linearity in the age-epithelial content association. These associations should be further investigated in other cohorts, particularly if epithelial content serves as a marker of risk as suggested previously (28).

Composition changes with age may also reflect changes in breast volume. As non-fatty stroma declines, overall breast tissue volume declines and there is a concomitant increase in fat proportion. This is reflected in the ratio of non-fatty stroma to adipose: the ratio increased slightly in the age group of 60–69 years, decreasing after age 70. Thus this represents an alternative explanation for the increase in epithelial content in older women; the increasing proportion of epithelium may reflect a smaller breast rather than an increase in epithelium. However, even after normalizing epithelial proportion to fat percentage, the U-shaped curve morphology persisted (data not shown). Strong dichotomies in breast composition between young women (<50) and older women (≥ 50) are important to consider in light of observations that tumor biology in young women differs from that of old women (29).

BMI was strongly and inversely associated with non-fatty stroma content in PBCS. This result was consistent with previous studies (16, 17), perhaps reflecting the coordination of whole body composition and local breast composition. However, obese women also had a higher proportion of epithelium. This association has not been reported previously, but is in line with the positive relationship of obesity and breast cancer risk observed in postmenopausal women. Adipose cells are believed to express high levels of aromatase, and therefore local effects on epithelial cell histology are biologically plausible (37). Because 76% of our study population was postmenopausal, our data were well powered to capture the association of epithelial variation with obesity.

We did not observe strong associations between epithelial content and reproductive factors, in contrast to previous findings in the Nurse's Health Study (NHS) (16). The vast majority of our patients had experienced birth more than 15 years prior to sample collection, whereas the younger NHS study participants (mean age 45 compared to 56 in PBCS) were more recently pregnant. The differences in findings and study population suggest that effects of reproductive exposures, such as parity, on breast tissue composition may not be persistent. In fact, the prevailing hypothesis for long-term parity-associated protection emphasizes change in differentiation state of epithelium rather than change in cellular content (38). Other morphometric measures of epithelial size, shape, and or involution status may be more closely linked with reproductive factors.

Because patients in this study had cancer, it was also possible to evaluate breast composition in association with tumor characteristics. We found that greater epithelial composition was associated with poorer tumor differentiation and weakly with ER-negative status. While the effects suggesting relationships between cancer-adjacent composition and tumor characteristics are weak, future studies should consider other measures of benign morphometry. A previous study in PBCS (27) has suggested that TDLU involution may be inversely correlated with basal-like breast cancer and more careful measures of epithelial morphometry may be more sensitive than 'epithelial area' used in this study. Future studies should also evaluate inflammatory infiltrates as these cells may play an important role in gene expression patterns of ER-positive vs. negative tumors. However, an interesting observation of our study is that the associations of epithelium and non-fatty stroma with tumor characteristics were always in opposite directions. For example, large tumor size was associated with decreased stromal content and increased epithelial content. This could

reflect a repressive stroma and/or a common factor that increases epithelial cell growth for both benign and tumor tissue. Whatever the causes, these disparate findings for stroma and epithelium have important implications. Because epithelial and stromal content together are associated with increasing mammographic density, relationships between cellular content and tumor characteristics may be obscured (that is, changes in epithelium and stroma may effectively cancel each other out) in studies that rely on mammographic density as a surrogate measure of breast tissue composition. Nonetheless, while associations between mammographic density and intrinsic subtype may be largely null (39, 40), associations between tumor characteristics and particular cellular components may be important for progression.

By linking our risk factor exposures with composition, and composition with gene expression, our findings suggest biological pathways that are altered by risk factor exposure. For example, risk factors that alter stromal content may have effects on vascular permeability, tight junction signaling, metabolic signaling, loss of contact inhibition and cell polarity. Indeed, our previous experimental work has shown that stromal-epithelial interactions altered metabolism and glucose uptake of breast cancers (41). Thus, these gene expression changes may lead to novel hypotheses about breast cancer microenvironments. Furthermore, these independently validated composition signatures may be used in future studies where histological data is unavailable, to estimate composition.

Our study should be considered in light of limitations and strengths. 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 (24, 42), and therefore associations should be interpreted with cognizance of these possible differences. Future histological research on normal breast tissues from healthy women is needed. In particular, studies like the Komen Tissue Bank (43), in which needle biopsy sampling is essentially 'blind' to tissue structure, are important. In studies using whole tissue collected during pathologic evaluation, a bias toward sampling tissue which appears to contain parenchymal cells is well known. This bias was also present in the PBCS samples. Thus, our findings on composition are generalizable only as relative measures of composition and not as absolute measures. It is likely that our study, and moreover, most studies using surgical tissues, over-estimate stromal and epithelial percentages. In spite of these limitations, these data show compelling evidence that breast tissue composition varies according to epidemiologic risk factors, with important consequences for gene expression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support: This project was supported by grants from the National Cancer Institute to M. A. Troester (U01-ES019472 and R01-CA138255), and a Breast SPORE Project (P50CA058223, H.S. Earp PI), and a grant from the Avon Foundation to M. A. Troester. This work was supported (in part) by the Intramural Research Program of the National Institutes of Health, National Cancer Institute.

The authors thank Norman F. Boyd for contributing density measurements used in this analysis and for thoughtful comments on the manuscript. Also we thank Drs. Montserrat Garcia-Closas, Jolanta Lissowska, and Louise Brinton for their contributions in data and biospecimen collection.

References

1. Wiseman BS, Werb Z. Stromal effects on mammary gland development and breast cancer. *Science*. 2002; 296:1046–9. [PubMed: 12004111]
2. Robinson GW, Karpf AB, Kratochwil K. Regulation of mammary gland development by tissue interaction. *J Mammary Gland Biol Neoplasia*. 1999; 4:9–19. [PubMed: 10219903]
3. Russo J, Romero AL, Russo IH. Architectural pattern of the normal and cancerous breast under the influence of parity. *Cancer Epidemiol Biomarkers Prev*. 1994; 3:219–24. [PubMed: 8019370]
4. Ghosh K, Brandt KR, Reynolds C, Scott CG, Pankratz VS, Riehle DL, et al. Tissue composition of mammographically dense and non-dense breast tissue. *Breast Cancer Res Treat*. 2012; 131:267–75. [PubMed: 21877142]
5. Couto E, Qureshi SA, Hofvind S, Hilsen M, Aase H, Skaane P, et al. Hormone therapy use and mammographic density in postmenopausal norwegian women. *Breast Cancer Res Treat*. 2012; 132:297–305. [PubMed: 22052325]
6. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med*. 2011; 17:320–9. [PubMed: 21383745]
7. Polyak K, Kalluri R. The role of the microenvironment in mammary gland development and cancer. *Cold Spring Harb Perspect Biol*. 2010; 2:a003244. [PubMed: 20591988]
8. Kass L, Erler JT, Dembo M, Weaver VM. Mammary epithelial cell: Influence of extracellular matrix composition and organization during development and tumorigenesis. *Int J Biochem Cell Biol*. 2007; 39:1987–94. [PubMed: 17719831]
9. Barcellos-Hoff MH. Stromal mediation of radiation carcinogenesis. *J Mammary Gland Biol Neoplasia*. 2010; 15:381–7. [PubMed: 21181431]
10. Campbell LL, Polyak K. Breast tumor heterogeneity: Cancer stem cells or clonal evolution? *Cell Cycle*. 2007; 6:2332–8. [PubMed: 17786053]
11. Polyak K. Breast cancer: Origins and evolution. *J Clin Invest*. 2007; 117:3155–63. [PubMed: 17975657]
12. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2006; 15:1159–69. [PubMed: 16775176]
13. Jezierska-Drutel A, Rosenzweig SA, Neumann CA. Role of oxidative stress and the microenvironment in breast cancer development and progression. *Adv Cancer Res*. 2013; 119:107–25. [PubMed: 23870510]
14. Reisfeld RA. The tumor microenvironment: A target for combination therapy of breast cancer. *Crit Rev Oncog*. 2013; 18:115–33. [PubMed: 23237555]
15. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: Biology and implications for treatment. *Breast Cancer Res*. 2011; 13:227. [PubMed: 22078026]
16. Gertig DM, Stillman IE, Byrne C, Spiegelman D, Schnitt SJ, Connolly JL, et al. Association of age and reproductive factors with benign breast tissue composition. *Cancer Epidemiol Biomarkers Prev*. 1999; 8:873–9. [PubMed: 10548315]
17. 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; 14:343–9. [PubMed: 15734956]
18. Sun X, Gierach GL, Sandhu R, Williams T, Midkiff BR, Lissowska J, et al. Relationship of mammographic density and gene expression: Analysis of normal breast tissue surrounding breast cancer. *Clin Cancer Res*. 2013; 19:4972–82. [PubMed: 23918601]
19. 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; 95:123–9. [PubMed: 16755295]

20. 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; 9:3–13. [PubMed: 21131600]
21. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. *Phys Med Biol*. 1994; 39:1629–38. [PubMed: 15551535]
22. Hastie, TJTR. Generalized additive models. 1. New York: Chapman and Hall; 1990.
23. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004; 3:Article 3.
24. Casbas-Hernandez P, Arcy DM, Roman-Perez E, Brauer HA, McNaughton K, Miller SM, et al. Role of HGF in epithelial-stromal cell interactions during progression from benign breast disease to ductal carcinoma in situ. *Breast Cancer Res*. 2013; 15:R82. [PubMed: 24025166]
25. Finak G, Sadekova S, Pepin F, Hallett M, Meterissian S, Halwani F, et al. Gene expression signatures of morphologically normal breast tissue identify basal-like tumors. *Breast Cancer Res*. 2006; 8:R58. [PubMed: 17054791]
26. Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL, et al. Associations of breast cancer risk factors with tumor subtypes: A pooled analysis from the breast cancer association consortium studies. *J Natl Cancer Inst*. 2011; 103:250–63. [PubMed: 21191117]
27. Yang XR, Figueroa JD, Falk RT, Zhang H, Pfeiffer RM, Hewitt SM, et al. Analysis of terminal duct lobular unit involution in luminal A and basal breast cancers. *Breast Cancer Res*. 2012; 14:R64. [PubMed: 22513288]
28. Ghosh K, Hartmann LC, Reynolds C, Visscher DW, Brandt KR, Vierkant RA, et al. Association between mammographic density and age-related lobular involution of the breast. *J Clin Oncol*. 2010; 28:2207–12. [PubMed: 20351335]
29. 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; 21:1735–44. [PubMed: 22859400]
30. Ginsburg OM, Martin LJ, Boyd NF. Mammographic density, lobular involution, and risk of breast cancer. *Br J Cancer*. 2008; 99:1369–74. [PubMed: 18781174]
31. Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, van de Vijver MJ, et al. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci Transl Med*. 2011; 3:108ra113.
32. Dekker TJ, van de Velde CJ, van Pelt GW, Kroep JR, Julien JP, Smit VT, et al. Prognostic significance of the tumor-stroma ratio: Validation study in node-negative premenopausal breast cancer patients from the EORTC perioperative chemotherapy (POP) trial (10854). *Breast Cancer Res Treat*. 2013; 139:371–9. [PubMed: 23709090]
33. Downey CL, Simpkins SA, White J, Holliday DL, Jones JL, Jordan LB, et al. The prognostic significance of tumour-stroma ratio in oestrogen receptor-positive breast cancer. *Br J Cancer*. 2014; 110:1744–7. [PubMed: 24548861]
34. Gnerlich JL, Yao KA, Fitch PS, Goldschmidt RA, Bond MC, Cornwell M, et al. Peritumoral expression of adipokines and fatty acids in breast cancer. *Ann Surg Oncol*. 2013; 20 (Suppl 3):S731–8. [PubMed: 24052317]
35. Boyd NF, Martin LJ, Li Q, Sun L, Chiarelli AM, Hislop G, et al. Mammographic density as a surrogate marker for the effects of hormone therapy on risk of breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2006; 15:961–6. [PubMed: 16702377]
36. Walker RA, Martin CV. The aged breast. *J Pathol*. 2007; 211:232–40. [PubMed: 17200937]
37. Bulun SE, Mahendroo MS, Simpson ER. Aromatase gene expression in adipose tissue: Relationship to breast cancer. *J Steroid Biochem Mol Biol*. 1994; 49:319–26. [PubMed: 8043495]
38. Hawes D, Downey S, Pearce CL, Bartow S, Wan P, Pike MC, et al. Dense breast stromal tissue shows greatly increased concentration of breast epithelium but no increase in its proliferative activity. *Breast Cancer Res*. 2006; 8:R24. [PubMed: 16646977]
39. Morishita M, Ohtsuru A, Hayashi T, Isomoto I, Itoyanagi N, Maeda S, et al. Clinical significance of categorisation of mammographic density for breast cancer prognosis. *Int J Oncol*. 2005; 26:1307–12. [PubMed: 15809722]

40. Razzaghi H, Troester MA, Gierach GL, Olshan AF, Yankaskas BC, Millikan RC. Association between mammographic density and basal-like and luminal A breast cancer subtypes. *Breast Cancer Res.* 2013; 15:R76. [PubMed: 24008056]
41. Brauer HA, Makowski L, Hoadley KA, Casbas-Hernandez P, Lang LJ, Roman-Perez E, et al. Impact of tumor microenvironment and epithelial phenotypes on metabolism in breast cancer. *Clin Cancer Res.* 2013; 19:571–85. [PubMed: 23236214]
42. 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; 15:7020–8. [PubMed: 19887484]
43. Sherman ME, Figueroa JD, Henry JE, Clare SE, Rufenbarger C, Storniolo AM. The susan G. komen for the cure tissue bank at the IU simon cancer center: A unique resource for defining the “molecular histology” of the breast. *Cancer Prev Res (Phila).* 2012; 5:528–35. [PubMed: 22345117]

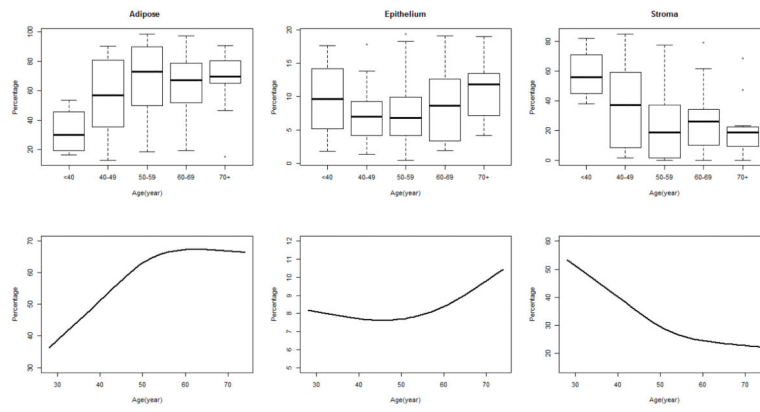


Figure 1. Boxplots and spline lines of breast tissue composition by age

Boxplots show the distribution of composition by five age groups at diagnosis (<40 y, 40–49 y, 50–59 y, 60–69 y, and 70 y). A cubic spline interpolation algorithm (knots=3) was used to smooth the curve and visualize these associations.

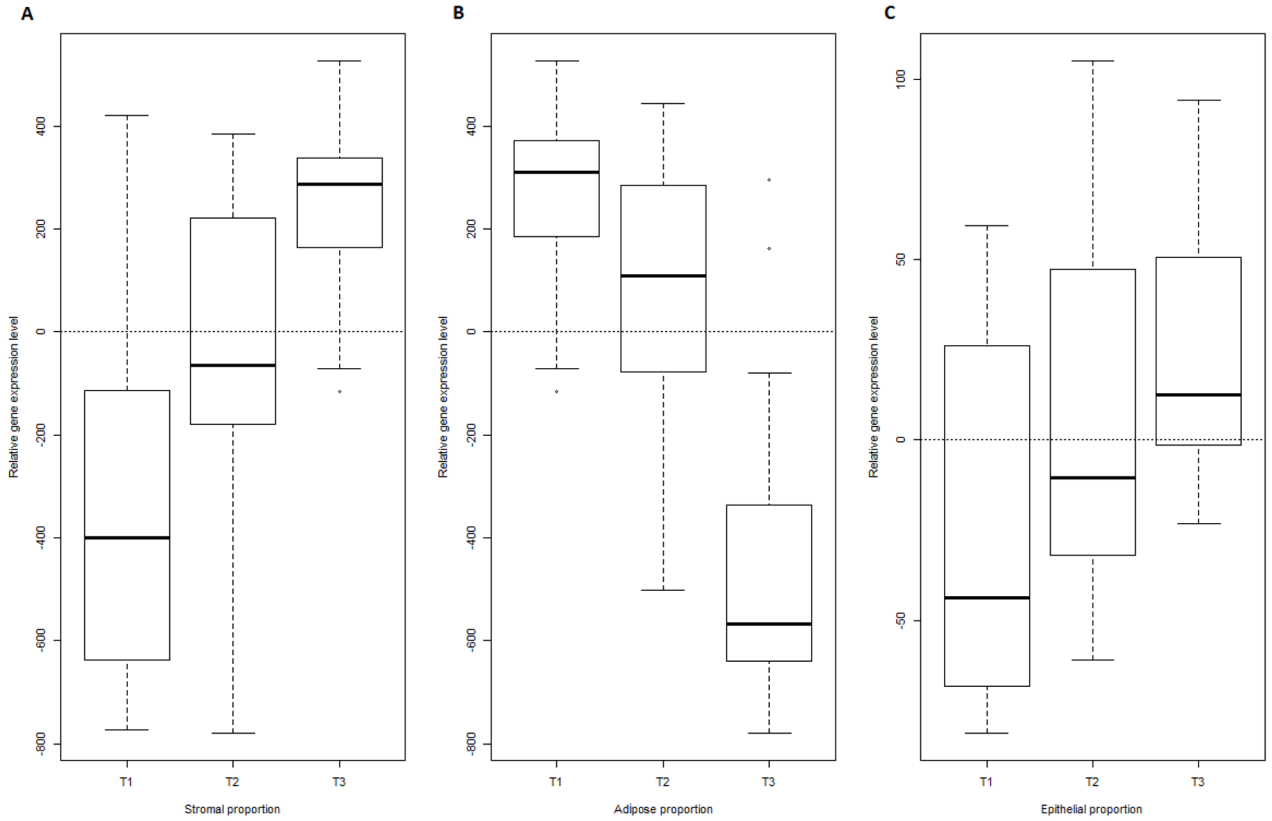


Figure 2. A–C. The average gene expression level of epithelial and stromal composition signatures across breast tissue composition groups in test dataset (n=50)

The association of relative expression of epithelial and stromal composition signatures with breast tissue composition was validated in a test set sample (n=50). The $\log_2(R/G)$ expression for each gene of each composition signature was multiplied by corresponding direction index (1 for up-regulated genes, and -1 for down-regulated gene), and results were summed across all genes as relative gene expression level. This relative gene expression level was evaluated for association with tissue composition by tertiles (cutoff points: epithelial proportion, 7.20% and 16.05%; stromal proportion, 10.78% and 34.31%; adipose proportion, 47.11% and 79.62%). **A** and **C**. The average gene expression level of the stromal and epithelial composition signatures increased with actual stromal and epithelial proportion (age-adjusted p-value for trend <0.001). **B**. The expression level of the stromal signature decreased with actual adipose proportion (age-adjusted p-value for trend <0.001).

Table 1

Age-adjusted means of epithelial and stromal proportion, by breast cancer risk factors^a

Characteristics	N (%)	Epithelial percentage			Stromal percentage		
		Mean (%)	SE	P	Mean (%)	SE	P
Percentage density							
<25%	50 (52)	8.69	0.69	0.50	19.27	3.11	<.01
25%	46 (48)	7.99	0.72		36.86	3.26	
Dense area							
<median ^b	52 (54)	8.25	0.64	0.81	22.98	3.00	0.02
median	44 (46)	8.48	0.70		33.28	3.27	
BMI							
<30 kg/m ²	67 (70)	7.59	0.57	0.02	31.31	2.74	0.02
30 kg/m ²	29 (30)	10.12	0.88		19.36	4.22	
Menopausal Status							
Premenopausal	23 (24)	9.22	1.15	0.41	28.07	5.52	0.94
Postmenopausal	73 (76)	8.08	0.59		27.58	2.81	
Age at menarche							
12 y	38 (40)	8.36	0.79	0.98	31.28	3.65	0.16
>12 y	58 (60)	8.38	0.64		24.57	2.97	
Parity							
0	20 (21)	7.62	1.07	0.44	36.18	5.04	0.06
1	76 (79)	8.55	0.55		25.47	2.57	
Age at 1 st birth							
25 y	45 (59)	9.15	0.72	0.23	27.06	3.35	0.59
>25 y	31 (41)	7.78	0.87		24.24	4.04	
Duration of lactation among parous women							
6 m	32(53)	9.43	0.90	0.57	20.70	3.68	0.37
> 6 m	28 (47)	8.67	0.97		25.56	3.94	
MHT							
Never	64 (69)	7.90	0.60	0.25	26.29	2.86	0.32
Current/ever	29 (31)	9.17	0.90		31.46	4.28	

Characteristics	N (%)	Epithelial percentage			Stromal percentage		
		Mean (%)	SE	P	Mean (%)	SE	P
Previous benign breast disease							
No	73 (78)	8.07	0.56	0.29	27.18	2.67	0.35
Yes	20 (22)	9.38	1.08		32.66	5.15	
Family history ^c							
No	77 (81)	8.12	0.54	0.24	27.75	2.59	0.84
Yes	18 (19)	9.60	1.12		28.93	5.38	

^a Age was adjusted as linear spline with the knot of 50 years. For percentage density and dense area, BMI was additionally adjusted. P-values were calculated using Wald test.

^b Median=36.52 cm².

^c Family history of breast cancer in first degree relative.

Abbreviation: BMI, body mass index; SE, standard error; MHT, menopausal hormone therapy.

Table 2
Age-adjusted means of epithelial and stromal proportion by tumor characteristics^a

Characteristics	N (%)	Epithelial percentage			Stromal percentage		
		Mean (%)	SE	P	Mean (%)	SE	P
Tumor size							
2 cm	51 (53)	8.11	0.67	0.59	30.38	3.17	0.22
>2cm	45 (47)	8.64	0.71		24.66	3.38	
Histological type							
Ductal	51 (53)	8.05	0.67	0.51	27.81	3.19	0.96
Other	45 (47)	8.70	0.71		27.58	3.40	
Affected lymph node							
0	70 (73)	8.33	0.66	0.96	28.62	3.17	0.67
1	26 (27)	8.38	0.72		26.61	3.44	
Differentiation							
Well/moderate	52 (54)	7.78	0.56	0.05	29.58	2.70	0.18
Poor	44 (46)	9.91	0.92		22.64	4.44	
ER status							
Negative	24 (25)	9.76	1.00	0.11	22.25	4.78	0.19
Positive	72 (75)	7.89	0.56		29.52	2.69	

^a Age was adjusted as linear spline with the knot of 50 years. P-values were calculated using Wald test.