Published in final edited form as:

Breast Cancer Res Treat. 2010 January ; 119(2): 305–314. doi:10.1007/s10549-009-0350-0.

Decreased TGF β signaling and increased COX2 expression in high risk women with increased mammographic breast density

Wei Tse Yang,

University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Michael T. Lewis,

Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA

Kenneth Hess,

University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Helen Wong,

Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA

Anna Tsimelzon,

Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA

Nese Karadag,

University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Micheline Cairo,

Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA

Caimao Wei,

University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Funda Meric-Bernstam,

University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Powel Brown,

Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA

Banu Arun,

University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Gabriel N. Hortobagyi,

University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Aysegul Sahin, and

University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Jenny C. Chang

Correspondence to: Wei Tse Yang.

Wei Tse Yang and Michael T. Lewis are the co-first authors.

Aysegul Sahin and Jenny C. Chang are the co-senior authors.

Electronic supplementary material: The online version of this article (doi:10.1007/s10549-009-0350-0) contains supplementary material, which is available to authorized users.

Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA

Breast Center at Baylor College of Medicine, 1 Baylor Plaza BCM 600, Houston, TX 77030, USA

Abstract

High mammographic density is associated with an increased risk of breast cancer. We hypothesized that specific pathways exist that are associated with increased mammographic density, and may therefore be used to identify potential targets for chemoprevention. Histologically confirmed normal breast tissue was collected from women undergoing breast surgery who had available demographic data and mammograms for review. Women with low versus high mammographic breast density were compared. Differentially expressed genes using Affymetrix HG U133Plus2 chips were identified in dense versus non-dense tissue. Immunohistochemical analysis (IHC) of estrogen receptor, progesterone receptor, Ki67, and COX2 expression was performed. About 66 women were identified, 28 (42%) had high, and 38 (58%) had low mammographic density. About 73 genes had differential expression between normal breast tissue with high and low mammographic density ($P < 0.001$, fold change > 1.5 with a low false discovery rate ($< 10\%$). Network and canonical pathway analysis indicated decreased TGF β signaling (TGFB2, SOS, SMAD3, CD44 and TNFRSF11B) in dense breast tissue relative to non-dense breast. By IHC, only COX2 expression in the stroma was statistically significant on multivariate analysis. TGF β ligands are currently the only growth factors known to prevent mammary epithelial cell proliferation. TGF β signaling has been reported to be inhibited by COX-2, and these molecules are highly differentially expressed in individuals at high risk of developing breast cancer. These results strongly suggest that COX2 inhibition should be investigated for breast cancer prevention despite possible increase in cardiovascular risk.

Keywords

Mammographic density; Chemoprevention; TGF-beta; COX2 Inhibitors

Introduction

Increased mammographic density is an independent, well-established, and robust risk factor for breast cancer. Women with dense breast tissue visible on mammography have an increased risk of breast cancer that is up to 6-fold higher compared to women with low mammographic breast density, matched for other variables, like age, menopausal status, weight, and parity [1–10]. In addition, mammographic breast density has been shown to improve the discriminatory power of the Gail model for prediction of absolute invasive breast cancer risk [11]. Yet, the genetic and perhaps environmental factors responsible for increased breast density remain unknown. Possible explanations for the strong association of increased breast density and greater breast cancer risk may include elevated intra-mammary estrogen production with associated increased aromatase activity, increased signaling through pro-proliferative growth factor pathways such as Wnt, Notch, or Hedgehog, or the loss of negative control of epithelial proliferation. Identification and understanding of molecular pathways altered in dense versus non-dense breasts could lead to early therapeutic interventions to prevent the development of breast cancer in these high-risk individuals.

Enhanced estrogen receptor signaling is a known risk factor. Tamoxifen and raloxifene, selective estrogen receptor modulators, have been shown in several large multi-center studies to reduce the risk of developing breast cancer by 49% in women with a Gail risk greater than 1.7%. [5, 12–15]. Likewise, aromatase inhibitors like letrozole, which lower plasma estradiol levels are being tested in clinical trials as chemopreventive agents [16]. These chemoprevention agents target the estrogen receptor (ER) pathway and are therefore only effective in preventing ER-positive breast cancers [5, 12–19]. Pathways that may drive the development of ER-negative breast cancers are less well established, and there are no effective therapies in preventing the development of this potentially more aggressive tumor phenotype. Importantly, factors responsible for increased breast density and breast cancer risk may be lowered by appropriate chemoprevention agents.

We hypothesized that specific molecular pathways exist that are associated with increased mammographic density and breast cancer risk. The goals of this study were to (1) identify women with high and low mammographic density, (2) discover differential gene expression patterns of normal breast tissue in women with breast cancer with high and low breast density using the Affymetrix U133Plus 2 platform, and (3) confirm relevant target pathways by immunohistochemical (IHC) analysis of the normal breast in women with high and low breast densities.

Methods

We identified suitable patient samples from a frozen tissue bank consisting of more than a thousand specimens of *normal breast* obtained by excisional biopsy from mastectomy specimens in women undergoing surgery for breast cancer who did not have neoadjuvant chemotherapy or prior radiation and had preoperative mammograms available for review. The biopsy specimens were obtained at a distant site, at least 5 cm from the primary tumor. Over 250 mammograms were reviewed, and breast parenchymal density was classified by two different methods; firstly, according to the American College of Radiology's Breast Imaging—Reporting and Data System (BIRADS) [20], and secondly, by quantitative assessment of percentage density of the mammographic parenchyma after digitizing mammograms using software from NIH Image and associated programs (<http://rsb.info.nih.gov/nihi-image/>). The mammograms were independently reviewed by two radiologists specializing in mammography with two and twelve years experience in breast imaging respectively. Quantitative classification of mammographic parenchyma was based on radiological assessment after digitizing mammograms and measuring percentage density [1, 2, 8]. For differential gene expression patterns by group comparisons, we selected women who had low breast density (BIRADS 1 and 2) versus high density (BIRADS 3 and 4). In addition, we also analyzed density as a continuous variable for differentially expressed genes in both groups.

Normal breast tissue samples obtained from the study populations were embedded in OCT and longitudinally cut into 8 μ m thick sections. A section of each specimen was examined by hematoxylinosin by a breast pathologist (AS), and only samples containing normal breast tissue with at least 60% epithelial content were selected. The methods for gene expression experiments have been described else-where [21]. Briefly, total RNA was isolated using

TRIzol reagent (Invitrogen Corporation, Carlsbad, CA). Samples were then subsequently passed over a Qiagen RNeasy column (Qiagen, Valencia, CA). After RNA recovery, using a T7 promoter primer cDNA was synthesized, then fluorescently labeled and the labeled cRNA samples hybridized onto an Affymetrix U133Plus 2 Gene Chip which contains ~34,000 genes, following the manufacturer's recommended procedures. Hybridized arrays were then scanned with a dual-laser based scanner. Pathway analysis of differentially expressed genes was performed using Ingenuity Pathway Analysis software. We also ran BioCarta pathway comparison using Functional class scoring method [22] with BRB Array Tools software (<http://linus.nci.nih.gov/BRB-ArrayTools.html>).

We then performed IHC analysis of relevant target signaling pathways, estrogen receptor (ER), progesterone receptor (PgR), proliferation index Ki67, and COX2, and scored the expression of these markers in the stroma and epithelium independently. The IHC analysis was performed on the corresponding mirror image of the specimen used for gene expression arrays. The IHC score for each biomarker was assessed by the pathologist (AS) who was blinded to the corresponding mammographic density of the specimens.

Statistical methods

For class comparison analyses (expression values versus density status, expression values versus menopausal status, and methylation values versus density status), we used t-tests on log values to rank the “genes”. To adjust for multiple testing, we used a beta-uniform mixture (BUM) model to estimate the false discovery rate (FDR). High and low density subjects were compared with respect to IHC variables using the Wilcoxon rank sum test. All IHC factors were included in a single multiple variable logistic regression model with density status (high versus low) as the response variable. The model performance was quantified with respect to discrimination and calibration. Discrimination (i.e. whether the relative ranking of individual predictions is in the correct order) was quantified with the area under the receiver operating characteristic curve or with the concordance index, which is similar to the area under the receiver operating characteristic curve (AUC).

Results

Patient and sample characteristics

A total of 74 specimens from 66 women were obtained from women with known mammographic density. The median age of these women was 48 years (range 29, 88), and 27 women were premenopausal, while 39 women were postmenopausal. Of the women with high breast density, 19 (68%) were premenopausal, 9 (33%) were postmenopausal of whom 3 (11%) were on hormone replacement therapy (HRT) for more than one year. In women with low breast density, 8 women (21%) were premenopausal while 30 women (79%) were postmenopausal, of whom 9 (24%) were on HRT for more than 1 year (Table 1).

Differential gene expression between high and low density normal breast tissue

The median RNA extracted was 25 µg (range 19–33 µg). The quality of the RNA was good, with high percent present calls between 40 and 50%, and with GAPDH ratios of less than 1. BUM analysis was then conducted linking FDR values to specific *P*-value cut-offs which

can then be compared to the observed P -values to compute the proportion of “genes” with lower P -values than the specified cut-off. The numbers of genes from the expression analysis on density (high/low density) that were significant with a FDR of 0.05 (5%) was 1,124, at 0.025: 423, at 0.010: 103, at 0.005: 29, and at 0.001: 4 genes.

Age and menopausal status were potential confounding variables, and by univariate analysis were associated with breast density. By logistic regression model with both age and menopausal status, breast density was only significantly associated with menopausal status (Supplementary Fig. 1). By further analysis, menopausal status alone yielded very few differentially expressed genes, there was only one gene from the expression analysis on menopausal status that was significant with an FDR of 5%.

We used an external cross-validation procedure (10-fold Monte Carlo Cross Validation, MCCV) on the top ranked 50-gene diagonal linear discriminant analysis (DLDA) model, with 2,500 iterations. The mean cross-validated Area Under the Receiver Operator Characteristic Curve (AUROCC) was 0.78, the standard deviation (SD) over the 2,500 iteration was 0.18, and the graph of the Area above the Curve (AAC = 1–AUC) is shown in Fig. 1. Using a permutation test on the 50-gene DLDA model using the same cross-validation (10-fold MCCV), and 3,500 permutations, the maximum AUROCC was 0.76, giving a predictive accuracy with a $P < 0.0003$. We also assessed how predictive accuracy changed with the number of top 5, 10, 20, 50, 75, 100, 150, 200, and 250 genes. The performance estimates were higher for the models with fewer genes, but this was not statistically significant with a standard deviation of 0.18.

We identified differentially expressed genes whose expression level differed by 1.5-fold at the $P = 0.001$ and $P = 0.01$ levels. At the $P = 0.001$ level, 26 uniquely named genes (30 probesets) were more highly expressed in dense versus non-dense breast (Table 2), with 47 unique genes (57 probesets) showing lower expression in dense versus non-dense breast (Table 3). At the $P = 0.01$ level, there were 84 probesets more highly expressed, and 150 probesets showing lower expression in dense versus nondense breasts.

Pathway analysis of genes differentially expressed in dense versus non-dense breast

For Ingenuity Pathway Analysis, we used the $P = 0.001$ dataset to afford a higher level of confidence in the analysis. Consistent with the difference in breast density due to increased connective tissue, the top six functions identified in the analysis (Fig. 2) were “tissue morphology” (12 molecules; $P = 1.43\text{E-}06 - 9.62\text{E-}03$), “connective tissue development and function” (12 molecules; $P = 7.12\text{E-}06 - 1.44\text{E-}02$), “connective tissue disorders” (8 molecules; $P = 2.28\text{E-}05 - 9.62\text{E-}03$), “developmental disorders” (13 molecules; $P = 2.28\text{E-}05 - 9.62\text{E-}03$), “skeletal and muscular disorders” (7 molecules; $2.28\text{E-}05 - 1.07\text{E-}02$) and “tumor morphology” (5 molecules; $2.28\text{E-}05 - 1.16\text{E-}02$).

The 12 genes represented in “tissue morphology” were identical with those related to “connective tissue development and function”. Most were involved in the regulation of adipose and connective tissue mass (DLK1, IGHG1, MARK2, PDPK1, SAT1, SMAD2, TNFSF11B) and/or in regulation of cell number (CD44, IL6ST, SMAD3, TNFRSF11B, ANGPT2).

The 13 genes were represented in “developmental disorders” (ANGPT2, ASPH, DCX, DLK1, IGHG1, IL6ST, NFIC, PDPK1, RELN S100A10, SMAD3, TGFBR2, TNFRSF11B) overlapped considerably with all genes in “connective tissue disorder” with the exception of MMP16 which was present only in the latter group. Likewise, there was considerable overlap with “skeletal and muscular disorders” given that many of these genes are involved in cartilage and bone development. These three disease-related categories thus also overlapped considerably with “tissue morphology” and “connective tissue development and function”.

With respect to known gene networks, many of the genes listed above were represented in a cell-cell signaling network regulating tissue morphology. In particular, both network and canonical pathway analysis suggested decreased TGF β signaling in dense breast relative to non-dense breast. These differences included reduced expression of pathway components TGFBR2, SOS, and SMAD3, as well as reduced expression of two known TGF β target genes, CD44 and TNFRSF11B. Likewise, similar results were obtained using a functional class scoring method for BioCarta Pathway comparison, with reduced TGF β signaling in dense versus non-dense breasts (permutation $P < 0.005$).

Confirmation of relevant signaling pathways by IHC analysis

We analyzed COX2 expression together with proliferation (Ki67) expression by IHC. The expression of steroid hormone receptors ER and PgR was also analyzed. The expression of each biomarker was scored according to the Allred scoring system by a breast pathologist (AS), who was blinded to the density status of the women. The epithelial and stromal marker expressions were scored independently.

By univariate analysis, COX2 expression in stroma and epithelium ($P < 0.001$ and $P < 0.02$, respectively), as well as Ki67 in the stroma, were significantly higher in dense versus non-dense breast ($P < 0.05$). Steroid receptor (ER and PgR) expression was not significantly increased in dense versus non-dense breasts. By a multiple variable logistic regression model, only COX2 expression in the stroma was statistically significant at $P < 0.01$.

Discussion

Mammographic density is associated with an established increased risk of breast cancer. Identification and analysis of molecular pathways responsible for this phenotype would be important for understanding the underlying causes of breast cancer, and could facilitate the design of effective therapeutic interventions targeted against these critical pathways that would ultimately prevent the development of the disease.

We have found that decreased TGF β signaling, and increased COX2 expression and proliferation in the stroma are highly differentially expressed in individuals at high risk of developing breast cancer.

TGF β ligands are currently the only growth factors known to prevent mammary epithelial cell proliferation. Reduced TGF β signaling, through reduced ligand or dominant negative TGFBR2 expression, was sufficient to promote tumor formation in mouse models [23].

TGF β has been shown to directly influence TNFRSF11B (osteoprotegerin) [24], both of which showed decreased expression in dense compared to non-dense breast tissue. In addition, CD44, another known TGF β target [25, 26], also showed reduced expression in dense versus non dense breast tissue. CD44 is regulated in conjunction with Wnt signaling, the second highest ranked canonical pathway identified by our gene expression analysis. We found elevated expression of MARK2 which regulates DVL, and SOX4—a TGF β 1 target that interacts with TCF4 downstream of Wnt expression, although no “canonical” Wnt signaling components were identified in our analysis. Finally, and perhaps most importantly, consistent with this analysis, TGF β has been reported to be inhibited by COX-2 [27, 28], and we found COX2 expression together with stromal Ki67 highly differentially expressed in normal breast tissue of dense breasts.

In a preclinical transgenic mouse MMTV-COX-2 model that overexpresses the COX-2 gene in the mammary glands, hyperplasia followed by frank transformation was observed in 100% of the mice within 7–14 months, supporting the role of COX2 as an oncogene [29]. MMTV-COX-2 induction of mammary hyperplasia is associated with increased prostaglandin E2 synthesis, increased aromatase activity, decreased BRCA1 levels, and induction of tumor-associated angiogenesis [30–32]. In humans, COX-2 expression has been observed in 41% of invasive breast cancers, and 80% of ductal carcinoma in situ (DCIS) [33], and now, we report a higher expression in normal high-risk breast tissue with increased mammographic density.

An earlier study by Schreinemacher and Everson suggested that non steroidal anti-inflammatory drugs (NSAIDs) may suppress the development of different cancers [34]. Later studies have focused on more specific inhibitors of the cyclooxygenase (COX) enzymes, COX-1 and COX-2. Of these, COX-2 inhibitors like celecoxib given for only 6 months have been shown to significantly reduce the number of colorectal polyps, a pre-cancerous condition, in individuals with familial polyposis [35, 36]. However, recently, in addition to the known rare side effects of gastrointestinal bleeding, increased cardiovascular events (thromboses, strokes, and myocardial infarction) were described in several large phase III studies (<http://www.fda.gov/drug/infopage/COX2/default.htm>). These studies demonstrating increased risk of cardiovascular events were primarily observed in individuals at risk of ischemic heart disease, who were administered COX-2 inhibitors at higher doses over several years. Based on these side effects and fears, recruitment to several COX-2 inhibitor chemopreventive studies in breast and other cancers have been curtailed [37, 38].

There were some potential pitfalls with this study. First, mammographic density patterns may reflect differences in the amounts of varied cell types, including epithelial, stromal, or adipose cells in normal human breast tissue. The samples of normal breast tissue in this study were collected in OCT and snap frozen. Before RNA was extracted, a frozen section of the mirror image sample was evaluated histopathologically to characterize the ratio of epithelial, stromal, and adipose elements, and stroma. Because we elected not to perform laser-capture microdissection in order to better understand the contribution of global stromal interactions, the samples were selected with approximately equal admixture of stromal and epithelial components. Second, we dichotomized density into low and high density, according to the BIRADS criteria, while breast density is a continuous spectrum. We,

however, found that the differentially expressed genes obtained by group comparison were highly statistically significant, with a low false discovery rate and accuracy, supporting the appropriate selection of specimens. We therefore analyzed breast density as a continuous variable, and found that the differentially expressed genes obtained by comparison were still statistically significant, with a low false discovery rate and accuracy, supporting the appropriate selection of samples. Third, these were normal samples in mastectomy specimens, and may not be truly representative of disease-free women. Although histologically normal areas were sampled, there may theoretically be field effect changes that effected outcome.

In summary, we have demonstrated that decreased $TGF\beta$ signaling, associated with increased stromal Ki67 expression and COX-2 expression is observed in mammographically dense breast, an established risk factor for breast cancer. These observations support the contention that COX2 inhibitors at different doses and schedules to minimize cardiovascular side effects should be further investigated for breast cancer prevention in high risk individuals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was funded by P50CA116199 grant from the National Cancer Institute (GNH, with a career development award to WTY). This study was supported in part by the Breast Cancer Research Foundation (JCC), the Emma Jacobs Clinical Breast Cancer Fund (JCC), the NCI Breast Cancer SPORE P50 CA50183 (JCC, MTL), R01 CA112305-01 from the National Cancer Institute (JCC), P30 CA125123 from the National Institute of Health (JCC).

References

1. Atkinson C, Warren R, Bingham SA, Day NE. Mammographic patterns as a predictive biomarker of breast cancer risk: effect of tamoxifen. *Cancer Epidemiol Biomarkers Prev*. 1999; 8(10):863–866. [PubMed: 10548313]
2. Boyd NF, Byng JW, Jong RA, Fishell EK, Little LE, Miller AB, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. *J Natl Cancer Inst*. 1995; 87(9):670–675. DOI: 10.1093/jnci/87.9.670 [PubMed: 7752271]
3. 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):886–894. DOI: 10.1056/NEJMoa013390 [PubMed: 12239257]
4. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med*. 2007; 356(3):227–236. DOI: 10.1056/NEJMoa062790 [PubMed: 17229950]
5. Boyd NF, Lockwood GA, Martin LJ, Byng JW, Yaffe MJ, Tritchler DL. Mammographic density as a marker of susceptibility to breast cancer: a hypothesis. *IARC Sci Publ*. 2001; 154:163–169. [PubMed: 11220655]
6. Boyd NF, Lockwood GA, Martin LJ, Knight JA, Jong RA, Fishell E, et al. Mammographic densities and risk of breast cancer among subjects with a family history of this disease. *J Natl Cancer Inst*. 1999; 91(16):1404–1408. DOI: 10.1093/jnci/91.16.1404 [PubMed: 10451446]
7. 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(5):961–966. DOI: 10.1158/1055-9965.EPI-05-0762 [PubMed: 16702377]
8. Byng JW, Yaffe MJ, Jong RA, Shumak RS, Lockwood GA, Tritchler DL, et al. Analysis of mammographic density and breast cancer risk from digitized mammograms. *Radiographics*. 1998; 18(6):1587–1598. [PubMed: 9821201]
 9. Harvey JA, Bovbjerg VE. Quantitative assessment of mammographic breast density: relationship with breast cancer risk. *Radiology*. 2004; 230(1):29–41. DOI: 10.1148/radiol.2301020870 [PubMed: 14617762]
 10. Warwick J, Pinney E, Warren RM, Duffy SW, Howell A, Wilson M, et al. Breast density and breast cancer risk factors in a high-risk population. *Breast*. 2003; 12(1):10–16. DOI: 10.1016/S0960-9776(02)00212-6 [PubMed: 14659350]
 11. Chen J, Pee D, Ayyagari R, Graubard B, Schairer C, Byrne C, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *J Natl Cancer Inst*. 2006; 98(17):1215–1226. [PubMed: 16954474]
 12. Cuzick J, Warwick J, Pinney E, Warren RM, Duffy SW. Tamoxifen and breast density in women at increased risk of breast cancer. *J Natl Cancer Inst*. 2004; 96(8):621–628. [PubMed: 15100340]
 13. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst*. 1998; 90(18):1371–1388. DOI: 10.1093/jnci/90.18.1371 [PubMed: 9747868]
 14. Freedman M, San Martin J, O’Gorman J, Eckert S, Lippman ME, Lo SC, et al. Digitized mammography: a clinical trial of postmenopausal women randomly assigned to receive raloxifene, estrogen, or placebo. *J Natl Cancer Inst*. 2001; 93(1):51–56. DOI: 10.1093/jnci/93.1.51 [PubMed: 11136842]
 15. Tiersten A, Ng YY, Pile-Spellman E, Nelsen C, Noguera-Irizarry W, Brafman R, et al. Relationship between mammographic breast density and tamoxifen in women with breast cancer. *Breast J*. 2004; 10(4):313–317. DOI: 10.1111/j.1075-122X.2004.21332.x [PubMed: 15239789]
 16. Goss PE, Strasser-Weippl K. Prevention strategies with aromatase inhibitors. *Clin Cancer Res*. 2004; 10(1 Pt 2):372S–379S. DOI: 10.1158/1078-0432.CCR-031210 [PubMed: 14734494]
 17. Chlebowski RT, McTiernan A. Biological significance of interventions that change breast density. *J Natl Cancer Inst*. 2003; 95(1):4–5. [PubMed: 12509389]
 18. van Gils CH, Hendriks JH, Holland R, Karssemeijer N, Otten JD, Straatman H, et al. Changes in mammographic breast density and concomitant changes in breast cancer risk. *Eur J Cancer Prev*. 1999; 8(6):509–515. DOI: 10.1097/00008469-199912000-00006 [PubMed: 10643940]
 19. van Gils CH, Otten JD, Hendriks JH, Holland R, Straatman H, Verbeek AL. High mammographic breast density and its implications for the early detection of breast cancer. *J Med Screen*. 1999; 6(4):200–204. [PubMed: 10693066]
 20. BI-RADS. BI-RADS. American College of Radiology; Reston, VA: 2003.
 21. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet*. 2003; 362(9381):362–369. DOI: 10.1016/S0140-6736(03)14023-8 [PubMed: 12907009]
 22. Pavlidis P, Qin J, Arango V, Mann JJ, Sibille E. Using the gene ontology for microarray data mining: a comparison of methods and application to age effects in human prefrontal cortex. *Neurochem Res*. 2004; 29(6):1213–1222. DOI: 10.1023/B:NERE.0000023608.29741.45 [PubMed: 15176478]
 23. Forrester E, Chytil A, Bieri B, Aakre M, Gorska AE, Sharif-Afshar AR, et al. Effect of conditional knockout of the type II TGF-beta receptor gene in mammary epithelia on mammary gland development and polyomavirus middle T antigen induced tumor formation and metastasis. *Cancer Res*. 2005; 65(6):2296–2302. DOI: 10.1158/0008-5472.CAN-04-3272 [PubMed: 15781643]
 24. Thirunavukkarasu K, Miles RR, Halladay DL, Yang X, Galvin RJ, Chandrasekhar S, et al. Stimulation of osteoprotegerin (OPG) gene expression by transforming growth factor-beta (TGF-

- beta). Mapping of the OPG promoter region that mediates TGF-beta effects. *J Biol Chem*. 2001; 276(39):36241–36250. DOI: 10.1074/jbc.M104319200 [PubMed: 11451955]
25. Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, et al. Molecular definition of breast tumor heterogeneity. *Cancer Cell*. 2007; 11(3):259–273. DOI: 10.1016/j.ccr.2007.01.013 [PubMed: 17349583]
 26. Yu Q, Stamenkovic I. Transforming growth factor-beta facilitates breast carcinoma metastasis by promoting tumor cell survival. *Clin Exp Metastasis*. 2004; 21(3):235–242. DOI: 10.1023/B:CLIN.0000037705.25256.d3 [PubMed: 15387373]
 27. Studer RK, Chu CR. p38 MAPK and COX2 inhibition modulate human chondrocyte response to TGF-beta. *J Orthop Res*. 2005; 23(2):454–461. DOI: 10.1016/j.orthres.2004.08.012 [PubMed: 15734262]
 28. Uttamsingh S, Bao X, Nguyen KT, Bhanot M, Gong J, Chan JL, et al. Synergistic effect between EGF and TGF-beta1 in inducing oncogenic properties of intestinal epithelial cells. *Oncogene*. 2008; 27(18):2626–2634. DOI: 10.1038/sj.onc.1210915 [PubMed: 17982486]
 29. Liu CH, Chang SH, Narko K, Trifan OC, Wu MT, Smith E, et al. Overexpression of cyclooxygenase-2 is sufficient to induce tumorigenesis in transgenic mice. *J Biol Chem*. 2001; 276(21):18563–18569. DOI: 10.1074/jbc.M010787200 [PubMed: 11278747]
 30. Subbaramaiah K, Hudis C, Chang SH, Hla T, Dannenberg AJ. EP2 and EP4 receptors regulate aromatase expression in human adipocytes and breast cancer cells. Evidence of a BRCA1 and p300 exchange. *J Biol Chem*. 2008; 283(6):3433–3444. DOI: 10.1074/jbc.M705409200 [PubMed: 18083712]
 31. Chang MC, Wu HL, Lee JJ, Lee PH, Chang HH, Hahn LJ, et al. The induction of prostaglandin E2 production, interleukin-6 production, cell cycle arrest, and cytotoxicity in primary oral keratinocytes and KB cancer cells by areca nut ingredients is differentially regulated by MEK/ERK activation. *J Biol Chem*. 2004; 279(49):50676–50683. DOI: 10.1074/jbc.M404465200 [PubMed: 15375172]
 32. Chang SH, Liu CH, Conway R, Han DK, Nithipatikom K, Trifan OC, et al. Role of prostaglandin E2-dependent angiogenic switch in cyclooxygenase 2-induced breast cancer progression. *Proc Natl Acad Sci USA*. 2004; 101(2):591–596. DOI: 10.1073/pnas.2535911100 [PubMed: 14688410]
 33. Soslow RA, Dannenberg AJ, Rush D, Woerner BM, Khan KN, Masferrer J, et al. COX-2 is expressed in human pulmonary, colonic, and mammary tumors. *Cancer*. 2000; 89(12):2637–2645. DOI: 10.1002/1097-0142(20001215)89:12B2637::AID-CNCR17C3.0.CO;2-B [PubMed: 11135226]
 34. Schreinemachers DM, Everson RB. Aspirin use and lung, colon, and breast cancer incidence in a prospective study. *Epidemiology*. 1994; 5(2):138–146. DOI: 10.1097/00001648-199403000-00003 [PubMed: 8172988]
 35. Huls G, Koornstra JJ, Kleibeuker JH. Non-steroidal anti-inflammatory drugs and molecular carcinogenesis of colorectal carcinomas. *Lancet*. 2003; 362(9379):230–232. DOI: 10.1016/S0140-6736(03)13915-3 [PubMed: 12885487]
 36. Steinbach G, Lynch PM, Phillips RK, Wallace MH, Hawk E, Gordon GB, et al. The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis. *N Engl J Med*. 2000; 342(26):1946–1952. DOI: 10.1056/NEJM200006293422603 [PubMed: 10874062]
 37. Senior K. COX-2 inhibitors: cancer prevention or cardio-vascular risk? *Lancet Oncol*. 2005; 6(2):68. doi: 10.1016/S1470-2045(05)01720-1 [PubMed: 15704289]
 38. Dannenberg AJ, Altorki NK, Boyle JO, Dang C, Howe LR, Weksler BB, et al. Cyclo-oxygenase 2: a pharmacological target for the prevention of cancer. *Lancet Oncol*. 2001; 2(9):544–551. DOI: 10.1016/S1470-2045(01)00488-0 [PubMed: 11905709]

Fig. 1.

Area above the curve (1-AUC) with Monte Carlo Cross Validation, MCCV) on the top ranked 50-gene diagonal linear discriminant analysis (DLDA) model, with 2,500 iterations. Monte-Carlo cross-validation involves repeated cross-validation on random partitions of the data. Each iteration yields a new random partition, thus the more iterations the better, but with diminishing returns. The 2,500 iterations represent a reasonable compromise between excessive computation time and adequate representation of all possible partitions

A Top biological functions

B Top canonical signaling pathways

C Top gene networks

D Canonical TGFβ signaling

Fig. 2. Ingenuity Pathway analysis. **a** Top biological functions associated with differentially expressed genes. **b** Top canonical signaling pathways associated with genes differentially expressed in dense versus non-dense breast. **c** Top gene network identified by analysis of genes differentially expressed in dense versus non-dense breast. Red fill indicates higher expression in dense breast; green fill indicates lower expression in dense breast. **d** Schematic

diagram of canonical TGF β signaling with differentially expressed genes highlighted. Green fill indicates genes showing lower expression in dense breast relative to nondense breast

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Clinical and mammographic data for 66 women with genetic analysis stratified by breast density

	High density (<i>n</i> = 28)	Low density (<i>n</i> = 38)	Total (<i>n</i> = 66)
Age (years)			
Median (range, years)	45.0 (32, 66)	56.0 (29, 88)	48.0 (29, 88)
Pre-menopausal	19 (68)	8 (21)	27 (41)
Post-menopausal			
No HRT	6 (22)	21 (55)	27 (41)
HRT > 1 year	3 (11)	9 (24)	13 (19)
BIRADS breast density			
Classification			
1	–	12 (32)	
2	–	26 (68)	
3	20 (71)	–	
4	8 (29)	–	

Numbers in parentheses represent percentages

Table 2

Uniquely named genes more highly expressed in dense versus non-dense breast tissue

Affy ID	UniGene ID	Gene title	Gene symbol	Fold change	P-density	t-test
209560_s_at	Hs.533717	Delta-like 1 homolog (Drosophila)	DLK1	3.771138	0.000226	
230378_at	Hs.62492	Secretoglobins, family 3A, member 1	SCGB3A1	2.770219	0.000981	
204851_s_at	Hs.34780	Doublecortin, lissencephaly, X-linked (doublecortin)	DCX	2.552658	0.000151	
235976_at	Hs.525105	SLIT and NTRK-like family, member 6	SLITRK6	1.860899	8.46E-05	
202965_s_at	Hs.496593	Calpain 6	CAPN6	1.860899	0.000908	
203355_s_at	Hs.434255	Pleckstrin and Sec7 domain containing 3	PSD3	1.836552	0.00013	
205967_at	Hs.46423	Histone cluster 1, H4c	HIST1H4C	1.770308	4.42E-08	
241450_at	Hs.135015	R-Spondin homolog (Xenopus laevis)	RSPO1	1.767855	0.000168	
220051_at	Hs.72026	Protease, serine, 21 (testisin)	PRSS21	1.756861	0.000772	
209982_s_at	Hs.372938	Neurexin 2	NRXN2	1.754427	8.28E-06	
202341_s_at	Hs.435711	Tripartite motif-containing 2	TRIM2	1.693491	8.35E-05	
205923_at	Hs.558371	Reelin	RELN	1.675974	0.000168	
211633_x_at	Hs.510635	Immunoglobulin heavy constant gamma 1	IGHG1	1.639209	0.000222	
213298_at	Hs.170131	Nuclear factor I/C (CCAAT-binding transcription factor)	NFIC	1.630145	1.29E-05	
203929_s_at	Hs.101174	Microtubule-associated protein tau	MAPT	1.617763	3.69E-05	
201416_at	Hs.643910	SRY (sex determining region Y)-box 4	SOX4	1.60103	0.000468	
207012_at	Hs.546267	Matrix metalloproteinase 16 (membrane-inserted)	MMP16	1.565909	6.11E-05	
211634_x_at	-	Immunoglobulin heavy constant mu	IGHM	1.561573	0.000555	
217627_at	Hs.531262	Zinc finger protein 573	ZNF573	1.560491	8.97E-05	
205381_at	Hs.567412	Leucine rich repeat containing 17	LRRCL17	1.558329	8.14E-05	
222593_s_at	Hs.146679	Spermatogenesis associated, serine-rich 2	SPATS2	1.547565	2.66E-08	
231192_at	Hs.527909	Endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7	EDG7	1.54328	0.000216	
203431_s_at	Hs.440379	Rho GTPase-activating protein	RICS	1.542211	0.000886	
203942_s_at	Hs.567261	MAP/Microtubule affinity-regulating kinase 2	MARK2	1.524145	9.85E-06	
204524_at	Hs.459691	3-Phosphoinositide dependent protein kinase	PDPK1	1.523089	1.87E-05	

Table 3

Uniquely named genes expressed at lower levels in dense versus non-dense breast tissue

Affy ID	UniGene ID	Gene title	Gene symbol	Fold change	P-density t-test
208607_s_at	Hs.632144	Serum amyloid A1///A2	SAA1///SAA2	-2.109645	0.0004318
224568_x_at	Hs.642877	Metastasis associated lung adenocarcinoma transcript 1 (non-coding RNA)	MALAT1	-2.082043	0.0008568
230333_at	Hs.28491	Spermidine/spermine N1-acetyltransferase 1	SAT1	-2.076279	0.00005668
220065_at	Hs.132957	Tenomodulin	TNMD	-2.039195	0.0006703
238320_at	-	Trophoblast-derived noncoding RNA	TncRNA	-2.029325	0.00001025
214456_x_at	Hs.632144	Serum amyloid A1	SAA1	-1.832737	0.0002159
235631_at	Hs.591469	Discoidin domain receptor family, member 2	DDR2	-1.816297	0.0008882
242736_at	Hs.38621	Sorbin and SH3 domain containing 1	SORBS1	-1.780151	0.00002815
1553243_at	Hs.498586	Inter-alpha (globulin) inhibitor H5	ITIH5	-1.74957	0.0005955
223578_x_at	-	PRO1073 protein	PRO1073	-1.741101	0.0001976
208116_s_at	Hs.102788	Mannosidase, alpha, class 1A, member 1	MAN1A1	-1.739895	0.00009267
238909_at	Hs.143873	S100 calcium binding protein A10	S100A10	-1.718322	0.000408
234989_at	-	Trophoblast-derived noncoding RNA	TncRNA	-1.715941	0.00005225
214967_at	Hs.98259	Sterile alpha motif domain containing 4A	SAMD4A	-1.713564	0.0001035
204933_s_at	Hs.81791	Tumor necrosis factor receptor superfamily, member 11b (osteoprotegerin)	TNFRSF11B	-1.712377	0.00006684
1555594_a_at	Hs.478000	Muscleblind-like (Drosophila)	MBNL1	-1.697016	0.0003952
204864_s_at	Hs.532082	Interleukin 6 signal transducer (gp130, oncostatin M receptor)	IL6ST	-1.69584	0.0005663
1569003_at	Hs.444569	Transmembrane protein 49	TMEM49	-1.680628	0.000712
205554_s_at	Hs.476453	Deoxyribonuclease I-like 3	DNASE1L3	-1.672493	0.0003356
210148_at	Hs.201918	Homeodomain interacting protein kinase 3	HIPK3	-1.66094	0.0004933
207334_s_at	Hs.82028	Transforming growth factor, beta receptor II (70/80 kDa)	TGFBR2	-1.656341	0.0001195
209209_s_at	Hs.509343	Pleckstrin homology domain containing, family C, member 1	PLEKHC1	-1.646041	0.0001743
212777_at	Hs.278733	Son of sevenless homolog 1 (Drosophila)	SOS1	-1.641483	0.0004977
203641_s_at	Hs.470457	COBL-like 1	COBLL1	-1.639209	0.0005501
242277_at	Hs.102471	Phosphatase and actin regulator 2	PHACTR2	-1.596597	0.0003963
242774_at	Hs.525392	Spectrin repeat containing, nuclear envelope 2	SYNE2	-1.594385	0.0005184
215236_s_at	Hs.163893	Phosphatidylinositol binding clathrin assembly protein	PICALM	-1.58447	0.0001136
231274_s_at	Hs.122514	Solute carrier family 25, member 37	SLC25A37	-1.583372	0.0003

Affy ID	UniGene ID	Gene title	Gene symbol	Fold change	P-density t-test
210896_s_at	Hs.332422	Aspartate beta-hydroxylase	ASPH	-1.578988	0.0002632
227260_at	Hs.525163	Ankyrin repeat domain 10	ANKRD10	-1.577894	0.0005416
205846_at	Hs.434375	Protein tyrosine phosphatase, receptor type, B	PTPRB	-1.5768	0.0003221
201325_s_at	Hs.436298	Epithelial membrane protein 1	EMP1	-1.55941	0.0001865
242853_at	Hs.529609	ATPase type 13A3	ATP13A3	-1.557249	0.00001849
201559_s_at	Hs.440544	Chloride intracellular channel 4	CLIC4	-1.55617	0.0004587
1564053_a_at	Hs.491861	YTH Domain family, member 3	YTHDF3	-1.554015	0.00006834
217523_at	Hs.502328	CD44 Molecule (Indian blood group)	CD44	-1.550786	0.00005697
210875_s_at	Hs.124503	Transcription factor 8 (represses interleukin 2 expression)	TCF8	-1.549712	0.0007159
202759_s_at	Hs.591908	A kinase (PRKA) anchor protein 2	AKAP2///PALM2-AKAP2	-1.549712	0.0002442
236907_at	Hs.387804	Poly(A) binding protein, cytoplasmic 1	PABPC1	-1.537941	0.0001446
1557241_a_at	Hs.306339	Sushi-repeat-containing protein, X-linked 2	SRPX2	-1.534746	0.0004343
240960_at	Hs.87752	Moesin	MSN	-1.534746	0.00007599
219552_at	Hs.522334	Sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	SVEP1	-1.530497	0.0008898
235421_at	Hs.432453	Mitogen-activated protein kinase kinase kinase 8	MAP3K8	-1.528377	0.0003225
1558111_at	Hs.478000	Muscleblind-like (Drosophila)	MBNL1	-1.523089	0.0001467
239448_at	Hs.36915	SMAD family member 3	SMAD3	-1.523089	0.0003767
211148_s_at	Hs.583870	Angiotensinogen 2	ANGPT2	-1.519925	0.0005004
1552656_s_at	Hs.127310	U2AF Homology motif (UHM) kinase 1	UHMK1	-1.518872	0.0003975