

## Research Article

## Lipid Metabolism Genes in Contralateral Unaffected Breast and Estrogen Receptor Status of Breast Cancer

Jun Wang<sup>1</sup>, Denise Scholtens<sup>2,5</sup>, Michelle Holko<sup>2</sup>, David Ivancic<sup>1</sup>, Oukseub Lee<sup>1</sup>, Hong Hu<sup>1</sup>, Robert T. Chatterton, Jr.<sup>3,5</sup>, Megan E. Sullivan<sup>4</sup>, Nora Hansen<sup>1</sup>, Kevin Bethke<sup>1</sup>, Carola M. Zalles<sup>6</sup>, and Seema A. Khan<sup>1,5</sup>

## Abstract

Risk biomarkers that are specific to estrogen receptor (ER) subtypes of breast cancer would aid the development and implementation of distinct prevention strategies. The contralateral unaffected breast of women with unilateral breast cancer (cases) is a good model for defining subtype-specific risk because women with ER-negative (ER<sup>-</sup>) index primaries are at high risk for subsequent ER-negative primary cancers. We conducted random fine needle aspiration of the unaffected breasts of cases. Samples from 30 subjects [15 ER-positive (ER<sup>+</sup>) and 15 ER<sup>-</sup> cases matched for age, race and menopausal status] were used for Illumina expression array analysis. Findings were confirmed using quantitative real-time PCR (qRT-PCR) in the same samples. A validation set consisting of 36 subjects (12 ER<sup>+</sup>, 12 ER<sup>-</sup> and 12 standard-risk healthy controls) was used to compare gene expression across groups. ER<sup>-</sup> case samples displayed significantly higher expression of 18 genes/transcripts, 8 of which were associated with lipid metabolism on gene ontology analysis (GO: 0006629). This pattern was confirmed by qRT-PCR in the same samples, and in the 24 cases of the validation set. When compared to the healthy controls in the validation set, significant overexpression of 4 genes (*DHRS2*, *HMGCS2*, *HPGD* and *ACSL3*) was observed in ER<sup>-</sup> cases, with significantly lower expression of *UGT2B11* and *APOD* in ER<sup>+</sup> cases, and decreased expression of *UGT2B7* in both subtypes. These data suggest that differential expression of lipid metabolism genes may be involved in the risk for subtypes of breast cancer, and are potential biomarkers of ER-specific breast cancer risk. *Cancer Prev Res*; 6(4); 321–30. ©2013 AACR.

## Introduction

Breast cancer prevention requires accurate identification of high-risk women, not only those who are at risk of developing breast cancer, but also by risk for estrogen-receptor (ER) subtype, because effective pharmacologic prevention agents are currently only effective against ER positive (ER<sup>+</sup>) breast cancer. A wide array of established risk factors for breast cancer relates to endogenous life-time estrogen exposure (1–4), but the relative risk (RR) estimates for each of these factors are modest (1.5–2.0), and associations with the receptor subtype of breast cancer (5) are not strong enough to guide clinical decision making. Statistical models in clinical use have only modest discriminatory ability on hormone receptor subtype (6, 7). Currently, the only validated breast-based biomarkers of sporadic breast

cancer risk consist of epithelial atypia (8, 9) and mammographic density (10, 11), with RR estimation of 4 or greater, but again without specificity for ER subtypes of subsequent cancer (12). Molecular changes in benign breast lesions have potential value for the improved identification of high-risk women (13–15), but so far no hormone receptor (HR) subtype specificity has been showed.

Endocrine interventions that prevent ER<sup>+</sup> breast cancer (16–18) have adverse effects that have prevented wide adoption by high-risk women (19). Identification of women at risk specifically for ER<sup>+</sup> breast cancer, therefore likely to benefit from endocrine prevention, may improve acceptance of these drugs. Similarly, the identification of women at risk for ER-negative (ER<sup>-</sup>) breast cancer would spare women the toxicity of endocrine agents from which they would not benefit, and allow their recruitment to studies of new interventions with potential efficacy against ER<sup>-</sup> disease.

The contralateral unaffected breast of women undergoing surgical therapy for newly diagnosed unilateral breast cancer is an efficient model for discovery of subtype-specific risk markers, because studies of metachronous contralateral breast cancer show that second primary tumors in the contralateral breast resemble the ER status of the index primary. In a pooled analysis of NSABP studies, the OR for the association between ER status of the index primary and metachronous contralateral cancer was 14.8 [95% confidence interval (CI), 3.8–74.3%;  $P < 0.001$ ; ref. 20]. In

**Authors' Affiliations:** Departments of <sup>1</sup>Surgery, <sup>2</sup>Preventive Medicine, <sup>3</sup>Obstetrics & Gynecology, <sup>4</sup>Pathology, and <sup>5</sup>The Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois; and <sup>6</sup>Mercy Hospital, Miami Florida

**Note:** Supplementary data for this article are available at Cancer Prevention Research Online (<http://cancerprevres.aacrjournals.org>).

**Corresponding Author:** Seema A. Khan, Department of Surgery, Northwestern University, 303 E. Superior St, Lurie 4-111, Chicago, IL 60611. Phone: 312-503-4236; Fax: 312-503-2555; E-mail: [skhan@nmh.org](mailto:skhan@nmh.org)

doi: 10.1158/1940-6207.CAPR-12-0304

©2013 American Association for Cancer Research.

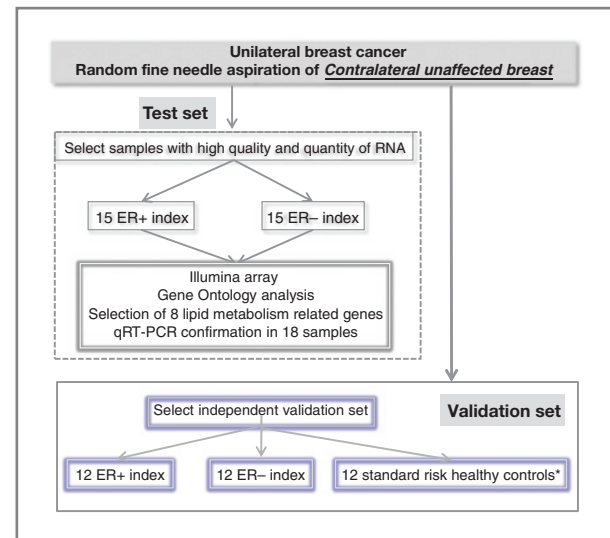
recent analyses of SEER data and the Geneva tumor registry, women whose first breast tumors were ER– were much more likely to develop an ER– second tumor. Women with HR-negative first tumors have nearly a 10-fold elevated risk (standardized incidence ratio = 9.81; 95% CI, 9.0–10.7%) of developing HR-negative second tumors (21, 22). The mechanism for HR status similarity of bilateral tumors is not clear, besides an association with younger age onset and possible hereditary mutations in *BRCA1*, *BRCA2* and *TP53* (21). Although the results of these studies need to be considered in the context of their methodological limitations (some studies lack detailed treatment or recurrence data, and are subject to a survival bias), the unaffected contralateral breast provides a window of opportunity for investigation of the molecular environment before the development of a second tumor. On the one hand, the contralateral breast may display a local physiological environment for specific subtype tumorigenesis, similar to the affected breast before the first tumor occurs. On the other hand, the contralateral breast may be affected by the first tumor through systemic factors (such as growth factors and paracrine factors) or affected by treatment for the first tumor, and modulated to a microenvironment favorable to similar subtype tumorigenesis. In either situation, studies on gene expression profiles in contralateral breast would provide us clues on tumor subtype preference and potential ER-specific biomarkers.

In this study, we have examined gene expression profiles in random fine needle aspirate (rFNA) samples from the contralateral breasts of women with new unilateral breast cancer (cases) to seek candidate panels of ER-specific risk biomarkers. On a discovery set of 30 women, we have identified gene expression differences in the contralateral breast that associate with ER+ or ER– index primary tumors. These findings were confirmed using qRT-PCR, validated in an independent set of 24 case samples, and compared to rFNA samples from 12 healthy controls. This report on a total of 66 women is larger than other studies of molecular features of normal breast epithelium, where total sample size has generally been under 50 women (23). Thus, our present findings based on 54 cases and 12 healthy controls constitute one of the largest of these challenging studies.

## Materials and Methods

### Study population

Cases diagnosed with unilateral breast cancer proceeding to primary surgical therapy were recruited under an Institutional Review Board (IRB) approved protocol (NU-835-040), with exclusions for neoadjuvant treatment, prior endocrine therapy, or pregnancy/lactation during the prior 2 years. rFNA was conducted in the operating room, on the contralateral unaffected breast of 85 women using the method of Fabian and colleagues (24). Samples were collected in a PBS-heparin solution and were kept on ice until processing. Samples from 54 subjects yielded sufficient high-quality RNA and could be matched by age, race and menopausal status. Of these, 30 subjects were included in



**Figure 1.** Diagram of the experimental design. Among 85 women who underwent rFNA of the CUB during surgery, 54 samples yielded sufficient high-quality RNA and could be matched by age, race, and menopausal status. Of these, 30 subjects were included in the test set for microarray analysis and 24 were used in the validation set. Twelve healthy control subjects were recruited to a different protocol where rFNA of both breasts was conducted for development of breast cancer risk biomarkers.

the test set for microarray analysis and 24 were used in the validation set (Fig. 1). In addition, a standard risk control group consisted of 12 healthy women recruited to another IRB approved protocol (NU08B2), with rFNA being conducted in the clinic. Thus, the total number of women included in the study (test, validation, control) was 66. By design, age, menopausal status and race were balanced between ER+ and ER– groups, and between controls and cases. The menopausal status was determined using last menstrual period date and hormonal criteria with true postmenopausal status classification were serum estradiol <30 pg/mL and follicle-stimulating hormone >30 IU/L. Cytological evaluation was done on one tenth of the sample; liquid-based cytological processing (Cytyc Corporation) was used to prepare a Papanicolaou-stained slide. Cytomorphology was assessed (by CZ) and classified by Masood semiquantitative index (25), and National Institute Consensus Panel Criteria (26). The projected probability of developing cancer at 5 years was calculated for each subject according to the Gail model (7) at the time of aspiration (for cases the calculation was meant only as a summary description of breast cancer risk factors, for comparison purposes).

### Total RNA extraction and gene expression profiling

Total RNA was extracted from rFNA samples using Trizol reagent (Invitrogen) and purified using RNeasyPlus Micro Kit (Qiagen). RNA was then treated with DNase and checked for integrity using Agilent 2100 Bioanalyzer. Samples with adequate RNA quality (RNA integrity number >8) were selected from 30 patients (15 with ER+ and 15 with ER– contralateral primaries) for Illumina array hybridization. Patients in the 2 groups were matched by age, race,

menopause status, and menstrual phase from premenopausal women as far as possible. Total RNA sample reverse transcription, labeling, and hybridization were conducted by Northwestern University Genomics Core Facility using Illumina Human WG6 BeadArray which contains 48,804 50-mers probes to detect >25,400 unique curated genes. The microarray data were deposited on Gene Expression Omnibus (GEO) with accession number GSE41400.

### Quantitative RT-PCR

Total RNA (10 ng) was reverse-transcribed using high-capacity RNA-to-cDNA Master Mix (Applied Biosystems) for qRT-PCR in 9 ER+/ER- pairs of test set, additional 12 ER+/ER- pairs in an independent validation set, with 12 rFNA samples from healthy individuals at standard cancer risk as controls. Eight selected genes and 2 housekeeping genes (*GAPDH* and *HPRT1*) for normalization were pre-loaded in 384 well microfluidic cards (each gene in triplicate) for Taqman low-density gene expression assays from Applied Biosystems. Assays were designed with small amplicons (<100 bp) to enhance detection sensitivity. RT-PCR reactions were then carried out in an Applied Biosystems 7900HT machine. Both test and validation set samples were tested in triplicate for all 8 target genes and 2 housekeeping genes. For each sample, the average of Ct values of each gene was calculated and used for data analysis. For each target gene, expression level was normalized against the average expression of 2 housekeeping genes.

### Statistical analysis

The raw data were normalized using the robust spline (RSN) algorithm available in the lumi Bioconductor package for R statistical software. Nonspecific filtering was done by limiting analyses only to those with interquartile range

above the median. Gene expression levels were compared between ER+ and ER- groups using paired *t* tests. Acknowledging the potential for false discovery, but aiming toward an encompassing list of candidates, genes with a nominal *P*-value less than 0.05 and at least 2-fold difference between two groups were considered differentially expressed seed candidate genes. The differentially expressed genes were analyzed for gene function annotation enrichment and gene ontology terms using DAVID (<http://david.abcc.ncifcrf.gov/home.jsp>). The annotation categories with at least 3 involved genes and a modified Fisher exact *P*-value less than 0.05 were considered strongly enriched in the gene function. Average linkage hierarchical clustering analysis was done using CIMminer from NCI Genomics and Bioinformatics Group (<http://discover.nci.nih.gov/>) to generate color-coded Clustered Image Maps (CIMS). All *P*-values are 2-tailed.

### Results

All results presented pertain to the contralateral unaffected breasts of breast cancer cases, which are referred to as ER+ or ER- depending on the ER status of the index lesion.

### Study subjects and rFNA cytology assessment

By design, age, menopausal status, and race were balanced between ER+ and ER- groups as well as across test and validation sets. The detailed patient clinical information and index tumor features are listed in Supplementary Tables S1 (test set) and S2 (validation set). There is no significant difference in Her2 status between ER+ and ER- groups in either the test or the validation sets ( $P = 0.328$  and  $0.871$ , respectively using  $\chi^2$  test). Tumor grade in the ER- group was significantly higher than in ER+ group in test set ( $P = 0.0028$ ) and in validation set ( $P = 0.043$ ) using  $\chi^2$  test. As summarized in Table 1, there were also no significant

**Table 1.** Summary of clinical information of patients for rFNA samples studies

	Healthy controls	Patients with ER+ cancer	Patients with ER- cancer	<i>P</i> -value
Test set in microarray				
N		15	15	
Age		50.3 ± 8.2	51.5 ± 9.3	
Premenopausal vs. postmenopausal		5/10	7/8	
European vs. other race		10/5	10/5	
BMI		27.5 ± 6.8	27.0 ± 4.0	0.98
Gail 5-year risk		1.1 ± 0.7	1.6 ± 1.2	0.64
Independent validation set in qRT-PCR				
N	12	12	12	
Age	53.1 ± 6.2	52.5 ± 8.6	54.4 ± 8.9	
Premenopausal vs. postmenopausal	4/8	4/8	4/8	
European vs. other race	10/2	10/2	10/2	
BMI	29.6 ± 9.3	27.0 ± 8.0	28.3 ± 6.3	0.74
Gail 5-year risk	1.2 ± 0.3	1.6 ± 1.2	1.6 ± 0.9	0.45

The data is mean and *SD*. Wilcoxon rank-sum test was used to compare the difference on age, BMI, and Gail 5-year risk between ER+ and ER- groups within testing set. One-way ANOVA test was used to compare the difference on age, BMI, and Gail 5-year risk among healthy control group, ER+ group and ER- group within validation set.

differences between ER+ and ER- groups with regard to BMI and 5-year Gail risk. For the validation set, there were no significant differences in these parameters between healthy controls and breast cancer cases. All cases have been followed for a minimum of 3 years, and one contralateral cancer has emerged, at an interval of 2 years.

In the test set there was no significant difference between ER+ and ER- groups in cytological category ( $P = 0.72$ ) and Masood score ( $P = 0.14$ ) of rFNA samples (Fig. 2). In the validation set, Masood score was similar between healthy controls and breast cancer cases ( $P = 0.19$ ). However, the cytological category differed, with the ER- cases showing a significantly higher frequency of contralateral breast atypia (9/15) than the healthy controls ( $P = 0.015$ ) and ER+ cases ( $P = 0.026$ ). There was no difference between healthy controls and ER+ cases ( $P = 0.75$ ). The median fraction of atypical cells in the atypical samples was 5%.

### Identification of differentially expressed genes

From 35,964 genes and 12,838 undesigned transcripts, 18 transcripts (with 5 duplicated genes) were found to be differentially expressed between ER+ and ER- groups by using 2-tailed paired  $t$  test (difference > 2-fold,  $P$ -value < 0.05). The expression of these genes was higher in the ER- group (Supplementary Table S3). As the  $P$ -values were not adjusted for multiple testing, we considered these to be *most likely* differentially expressed genes, compared to other genes that did not meet these criteria. After eliminating 5 duplicated genes, gene set enrichment analysis was done on 13 unique genes (Supplementary Table S4) using DAVID. Among the annotation categories that contained at least 3 target genes, the lipid metabolic process (GO:0006629) showed the most significant gene enrichment ( $P = 1.56 \times 10^{-8}$ ), which indicated that 8 genes from the gene set were involved in lipid metabolism and enriched in other

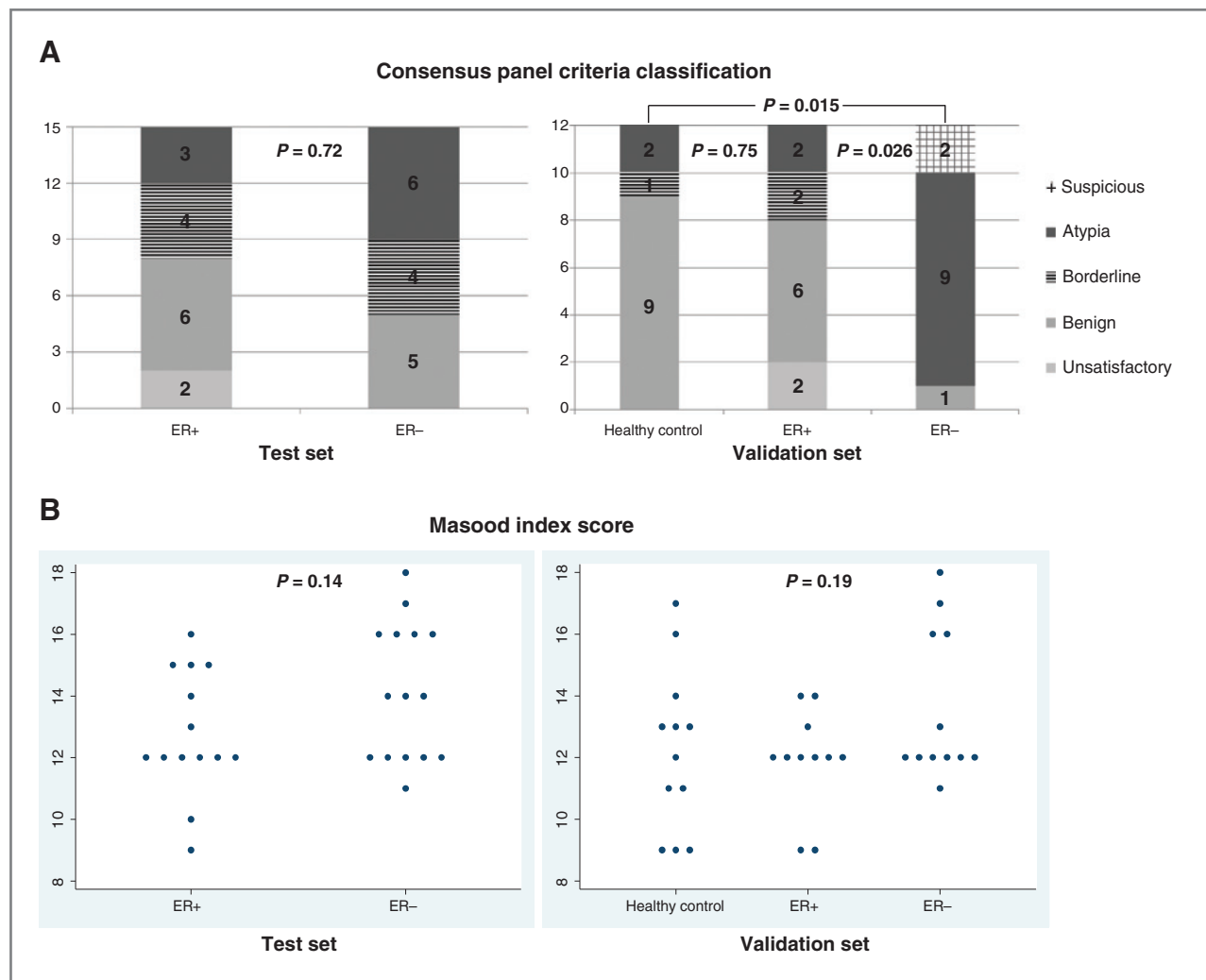


Figure 2. rFNA cytormorphology assessment using (A) cytology consensus panel criteria and (B) Masood index score.  $\chi^2$  test was used to compare difference on cytology category within test and validation sets. Wilcoxon rank-sum test was used to compare difference on Masood index score between ER+ and ER- groups within test sets. One-way ANOVA test was used to compare the difference on Masood index score among healthy control group, ER+ group and ER- group within validation set.

metabolic processes, such as fatty acid metabolism, steroid metabolism, and lipid biosynthesis (Supplementary Table S5). These were *DHRS2*, *HMGCS2*, *HPGD*, *ACSL3*, *UGT2B7*, *UGT2B11*, *APOD*, and *ALOX15B*.

There were 9 pairs of rFNA samples out of 15 pairs of the test set with sufficient residual RNA to conduct confirmatory qRT-PCR assay of these 8 genes. Paired *t* tests showed that the expression of 7 genes were significantly higher in samples from ER- cases than those from ER+ cases (Supplementary Fig. S1), confirming the microarray results. The expression of *UGT2B7* was also consistent with microarray results (higher in samples from ER- cases than in ER+ cases), but was not statistically significant.

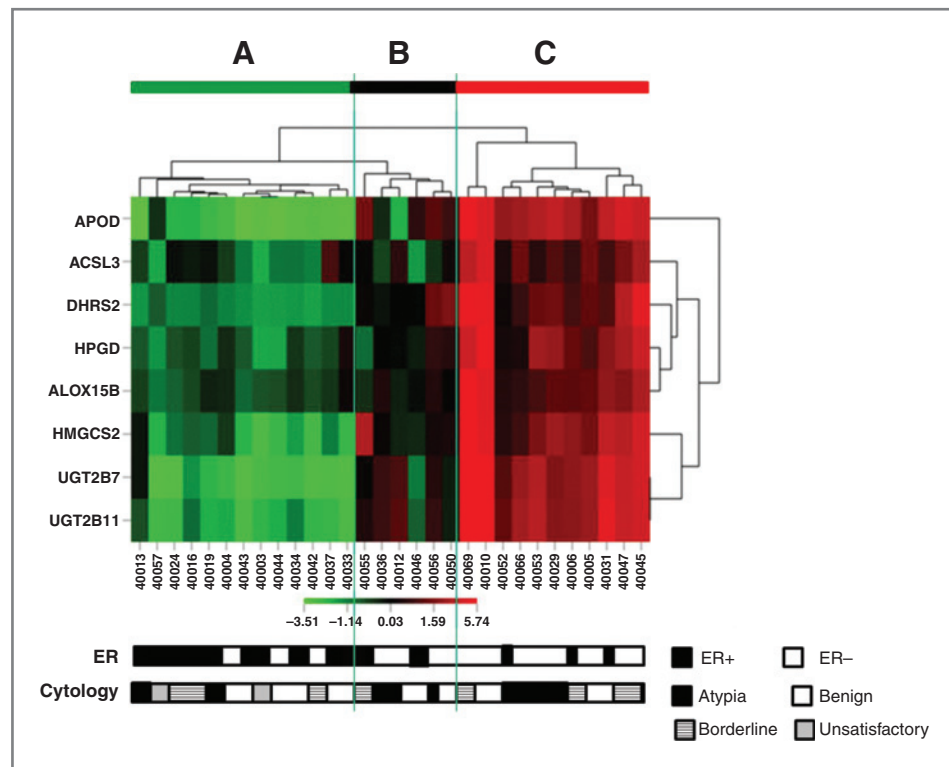
Hierarchical clustering analysis on 30 rFNA samples using 8 lipid metabolism genes (normalized to the median abundance of each gene) discriminated the samples into 3 major groups on the heatmap (Fig. 3). In group A, 10 of 13 samples (77%) from ER+ cases showed low expression. In group B, 2 of 6 (33%) samples (gene expression in the middle range) came from ER+ cases. In group C (highest expression level), 8 of 11 samples (73%) were associated with ER- cases. The association between the 3 main clusters and the ER status of the cases was significant (Fisher exact  $P = 0.035$ ). There was no clear dissection on the cytology categories.

#### Expression of 8 lipid metabolism-related genes in an independent validation set

To further validate the expression pattern of this 8 gene-set and to examine whether these genes were differentially expressed between high- and standard-risk breasts, we test-

ed their expression in an independent validation set of rFNA samples from 12 healthy controls and 24 cases (12 ER+ and 12 ER-), matched for age, race and menopausal status (Table 1). The expression of 8 genes in samples from ER- cases was significantly higher than in ER+ cases (*UGT2B11* being the only gene with a marginal significance), thus confirming the microarray and qRT-PCR results from the test set. Figure 4 shows one-way ANOVA results for the 8 genes, expression of which is significantly different among the 3 groups of subjects as reflected by the Bonferroni-adjusted  $P$  values. In addition, when comparing gene expression in the ER- cases to the control group, 4 genes were expressed at significantly higher levels: *DHRS2* (6.3-fold,  $P = 0.009$ ), *HMGCS2* (4.4-fold,  $P = 0.029$ ), *HPGD* (4.3-fold,  $P = 0.002$ ), and *ACSL3* (2.2-fold,  $P = 0.031$ ), whereas there was no significant difference in expression of these 4 genes between ER+ cases and the control group. In contrast, expression of *UGT2B11* (-5.6-fold,  $P = 0.0005$ ) and *APOD* (-3.0-fold,  $P = 0.0004$ ) was significantly lower in the ER+ cases compared to controls. The expression of *UGT2B7* was significantly decreased in both ER+ cases (-15.5,  $P = 0.0005$ ) and ER- cases (-11.5,  $P = 0.0005$ ) compared to controls, and was also different between ER- and ER+ groups (3.9,  $P = 0.020$ ). Overall, among the 8 genes that were tested in the validation set, expression of 4 genes (*DHRS2*, *HMGCS2*, *HPGD*, *ACSL3*) was increased in ER- cases, expression of 2 genes (*UGT2B11* and *APOD*) was decreased in ER+ cases, and expression of 1 gene (*UGT2B7*) was decreased in both subtypes. *ALOX15B* was not associated with either subtype.

**Figure 3.** Clustering analysis of 30 contralateral rFNA samples using 8 lipid metabolism-related genes. The expression level of the genes was normalized with the median abundance of each gene with red color indicating increased level and green color showing decreased level. The ER status was of the unilateral tumor.



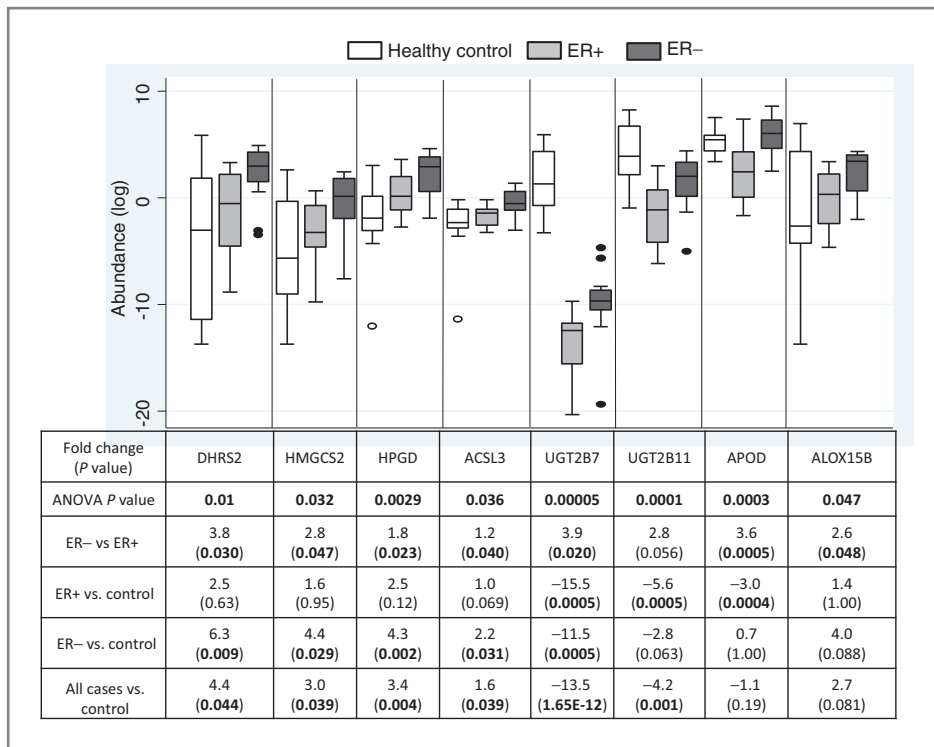


Figure 4. qRT-PCR validation of 8 lipid metabolism-related genes in independent rFNA samples. One-way ANOVA test was used to compare the difference (fold change) on log-transformed gene expression abundance among healthy control, ER+ and ER- groups. The Bonferroni-adjusted P-value was used for multiple comparisons between groups.

Hierarchical clustering analysis on 36 rFNA samples using 8 lipid metabolism genes (normalized to the median abundance of each gene) not only discriminated 24 case

samples into 3 major groups, but also dissected 12 control samples from 24 case samples (Fig. 5). Similar to the test set, 7 of 10 samples (70%) in group A with lower expression

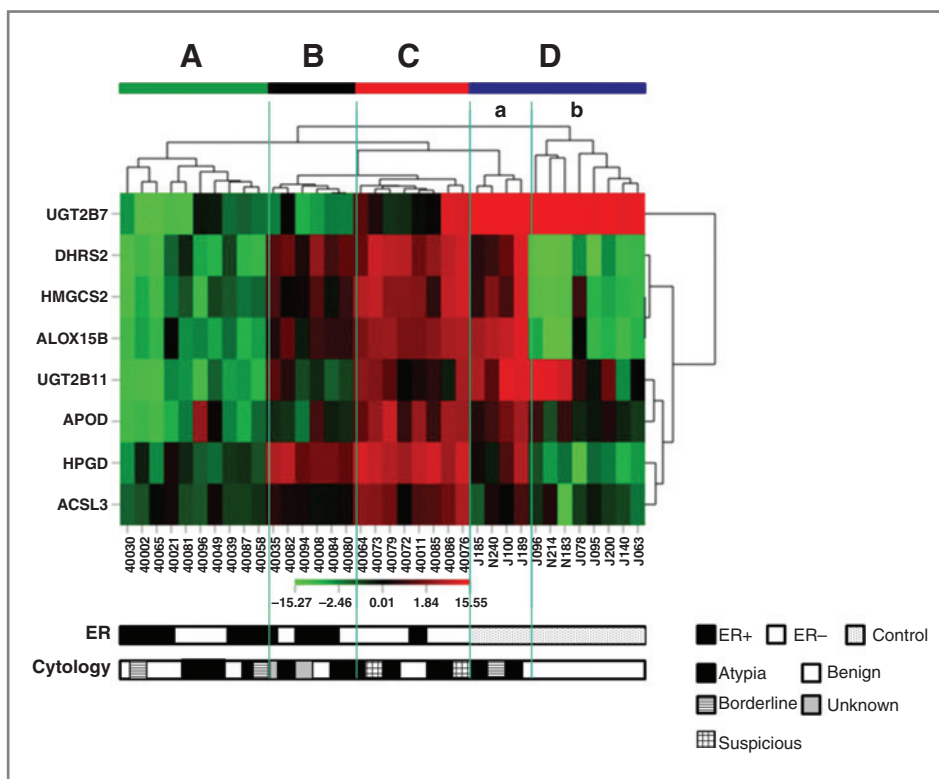


Figure 5. Clustering analysis of 24 contralateral rFNA samples and 12 healthy control samples using 8 lipid metabolism-related genes. The expression level of the genes was normalized with the median abundance of each gene with red color indicating increased level and green color showing decreased level. The ER status was of the unilateral tumor.

levels were ER+ cases; 4 of 6 (67%) samples in group B with gene expression in the middle range came from ER+ cases. In group C with highest expression level, 7 of 8 samples (88%) were associated with ER- cases. The association between the 3 main clusters and the ER status of the cases was significant (Fisher exact  $P = 0.034$ ). There was no clear dissection on the cytology categories on 24 cases. The control samples in group D were dissected into 2 clusters. The 4 samples in the first cluster (D-a) showed high gene expression levels, similar to group C samples (mainly ER- cases). Their cytology assessment was 2 atypia, 1 borderline, and 1 benign. In contrast, the 8 samples in the second cluster (D-b) had lower expression levels (except *UGT2B7*), similar to group A samples (mainly ER+ cases). All the cytology assessments of the 8 samples were benign.

## Discussion

Using the clinically normal contralateral breasts of breast cancer cases, we have shown that analysis of gene expression is feasible in rFNA samples, generating the first report of extensive gene expression analysis with such samples. Until now rFNA samples have been used for immunohistochemistry (27, 28) and DNA-based analyses (29–34), although feasibility of gene expression measurement has been explored (35). Considering the topographic heterogeneity in the breast, rFNA sampling is a useful tool for the discovery of breast cancer risk biomarkers; it provides a predominantly epithelial sample from a large area of the breast and allows the evaluation of women with clinically normal breasts, therefore greatly expanding the population for study, and eventual application of results. Finally, the use of the contralateral unaffected breast as a discovery platform allows the avoidance of the potentially initiated field around the tumor, which has been the subject of previous studies (36–39).

We have identified a candidate group of genes that are differentially expressed in the unaffected breasts of ER+ and ER- breast cancer cases. The risk of ER- metachronous second primaries is increased up to 22-fold in women with an index ER- primary (21). The presence of distinct gene expression patterns in the contralateral breast suggests that distinct gene expression patterns precede the development of different tumor subtypes. To our knowledge, this is the first report of a gene signature defined in the unaffected breasts of women with new unilateral ER+/ER- breast cancer, matched for age, race, and menopausal status.

The candidate genes were identified by comparison of gene array results between 2 high-risk groups; we selected 8 lipid metabolism-related genes from a total of 18 genes on the basis of gene ontology analysis. The differential expression of these genes was confirmed by qRT-PCR in the test set and verified in the validation set. Upon comparison with healthy control samples in the validation set, we observed that although the ER+/ER- differences were preserved, there were ER-specific differences in expression relative to healthy controls. The expression of 4 genes (*DHRS2*, *HMGCS2*, *HPGD*, and *ACSL3*) was significantly higher only

in ER- cases (compared to ER+ cases and to healthy controls), suggesting a specific association of high expression of these 4 genes with risk for ER- breast cancer. In contrast, expression of 2 genes (*UGT2B11* and *APOD*) was significantly decreased in ER+ cases compared with either control or ER- groups, indicating a possible protective association with ER+ cancer. *UGT2B7* on the other hand, showed significant decrease in both ER+ group and ER- group compared with controls, suggesting a protective function of *UGT2B7* against both subtypes of breast cancer.

The potential involvement of lipid metabolism-related genes to ER- breast cancer is unexpected, though evidence pointing to a link of lipid/steroid metabolism with ER- breast cancer risk and outcomes exists. ER-/PR- tumors are more common in obese premenopausal women (5), and large hip circumference has a particularly strong association with premenopausal ER-/PR- breast cancer (40). Specific estradiol metabolites may affect risk of ER- breast cancer (41), and studies of low-fat dietary interventions suggest a bigger impact on the outcome of ER- than ER+ tumors (42). Our novel findings that lipid metabolism-related genes are expressed at higher levels in the contralateral breasts of ER- cases point to potential interactions of the lipid metabolism-sex steroid axis and the metabolic environment of high-risk breasts.

The functions of these genes in relation to lipid modification and elimination, and to transportation and detoxification of distinct lipid compounds, suggest that their expression results in a specific microenvironment of steroid hormone metabolites, which may determine whether initiated cells progress to ER+ or ER- tumors. For example, UDP-glucuronosyl-transferase (UGT) 2B subfamily enzymes (*UGT2B7*, *UGT2B11*, *UGT2B28*) catalyzes conjugation glucuronic acid and facilitate elimination of potentially toxic compounds, particularly steroids, from target cells. *UGT2B7* has unique specificity for 3,4-catechol estrogens and estriol, and may protect breast tissues from carcinogenic estrogen metabolites locally (43, 44). Expression of *UGT2B7* and *2B11* genes was found to be decreased in breast cancer and in benign biopsies from healthy women with high mammographic density, a strong breast cancer risk factor (45). *DHRS2* (Hep27), a member of short-chain alcohol dehydrogenase/reductase (SDR) family of enzymes, has roles in ER+ breast cancer and is more frequently expressed in ER+ tumors (46). Its presence at higher levels in ER- cases suggests other roles in normal breast tissue. The *HPGD* gene encoding an alcohol dehydrogenase regulated by progesterone is a poor prognostic marker in breast cancer (47), and is overexpressed in ER- cell lines and ER- tumors (48). The specific roles and interaction of these genes in the genesis of ER- tumors clearly needs further work, but just as clearly, point to a heretofore unsuspected role of lipid metabolism abnormalities in the genesis of ER- breast cancer.

Weaknesses of our study include the fact that  $P$ -values for the initial selection of genes were not adjusted for multiple comparisons because of small sample size. There is possibility that the differences in the test set are chance

observations. Although the possibility of false discovery is mitigated by the verification of higher expression using qRT-PCR in the same samples, and in an independent validation set of similar samples with striking and distinct differences between ER+ and ER- case samples and healthy controls, further confirmation in a large sample set is needed.

The subjects in this study did not undergo neoadjuvant treatment, prior endocrine therapy, or pregnancy/lactation during the prior 2 years. Thus, we have minimized the potential effects on these factors on our study results. Although premenopausal samples were not matched for menstrual phase, we find in a parallel study that the expression of these genes is unrelated to serum estradiol and progesterone, and does not fluctuate during the menstrual cycle (manuscript in preparation). Another major exposure of interest that may modulate these genes, in our opinion, is obesity. Even though we did not find correlations between BMI and the 8 lipid metabolism gene set, we do not have the power to fully assess the impact of BMI because of the small sample size in this study. In addition, as we do not have detailed data on diet, physical activity variables, and other systemic lipid markers other than BMI, we cannot explore those relationships. We plan to include exploration of association with body mass index and other systemic lipid markers in future studies.

An additional question relates to the "normality" of the case samples. The cytomorphology of rFNA samples was similar in the case groups of the test set. In the validation set, the more quantitative Masood scores were not significantly different, but the categorical assessment of cytomorphology showed significantly more atypia in the ER- cases. However, atypical cells comprised a median of 5% of the epithelial sample, and are therefore unlikely to be driving the observed gene expression patterns. We will pursue this in future studies, because a relationship of lipid metabolism-related genes with cytological atypia, if it exists, would be of significant interest. Finally, with more than 3 years minimum follow-up on these subjects, only 1 of 54 cases developed contralateral breast cancer, providing some confidence that our samples came from truly unaffected high-risk breasts.

Prior examinations of molecular signatures of morphologically benign epithelial samples adjacent to breast tumors (36, 37, 39) have shown marked similarities to the molecular features of the adjacent cancer and is useful in defining the molecular "field" around the tumor (49), which may include paracrine effects of tumor cells on adjacent tissue. The use of the contralateral breast avoids these potential problems, and retains the advantage of analyzing tissue where the ER status of future cancers is likely to be similar to the index lesion (21, 22).

In this study, we observed the association between differential expression of lipid metabolism genes in the unaffected contralateral breast and tumor ER status. As the rFNA samples from contralateral breast were taken at the same time of tumor removal, the differential gene expression in contralateral breast could reflect a systemic alteration induced by tumor itself; if so, it seems that systemic changes

differ between women with ER- and ER+ breast cancer. Future studies of tissue acquired before tumor development would clarify the question. However, comparing our results with previous study by Graham and colleagues (38) on gene expression in tumor-adjacent epithelium point to the possibility of important differences in the tumor-adjacent epithelium and the contralateral high-risk tissue. The recent study of nonproliferative tumor-adjacent epithelium described a 198-gene signature differentiating the field surrounding 15 ER+ tumors from that surrounding 15 ER- tumors (38). Although 5 of our 13 lipid metabolism genes (Supplementary Table S3) were on that list (*UGT2B28*, *HPGD*, *SERHL*, *HMGCS2*, and *ALOX15B*), the difference was in the opposite direction (*higher* expression in samples adjacent to ER+ tumors), and concordant with expression patterns in the tumors themselves. Thus, differentially expressed genes in the surrounding field may not parallel those in the contralateral breast. With a similar sample size and high-quality RNA, although a different array platform, we found a much smaller number of differentially expressed genes in ER- and ER+ cases, implying that gene expression in the contralateral breasts of ER+ and ER- cases displays much smaller differences. Thus, continued investigations of unaffected breast tissue from the breast cancer population are vital; molecular changes that support the initial development of cancer may not be detectable in tumors and their surrounding tissue because they are swamped by the abundance of changes present in fully developed tumors. In contrast, the molecular features of unaffected high-risk tissue may reflect predisposing physiological conditions in the breast microenvironment, such as inflammatory states, oxidative stress, and deranged metabolism that promote tumor formation, which are no longer detectable in fully developed tumors, but are targets for preventive intervention.

In future studies, we will compare the profiles of lipid metabolism genes in normal epithelial cells adjacent to the tumor with those located in different quadrants and contralateral breasts. However, the eventual clinical utility of a validated signature may well be in rFNA samples, to offer risk stratification to women with clinically normal breasts.

#### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

#### Authors' Contributions

**Conception and design:** M.L. Holko, H. Hu, R.T. Chatterton, S.A. Khan  
**Development of methodology:** J. Wang, D. Ivancic, O. Lee, H. Hu, R.T. Chatterton, S.A. Khan

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** M.L. Holko, D. Ivancic, O. Lee, H. Hu, R.T. Chatterton, M.E. Sullivan, K. Bethke, C.M. Zalles, S.A. Khan

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** J. Wang, D. Scholtens, M.L. Holko, O. Lee, H. Hu, R.T. Chatterton, S.A. Khan

**Writing, review, and/or revision of the manuscript:** J. Wang, O. Lee, H. Hu, R.T. Chatterton, M.E. Sullivan, N. Hansen, K. Bethke, C.M. Zalles, S.A. Khan

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** H. Hu, S.A. Khan

**Study supervision:** S.A. Khan



## Grant Support

This work was supported by the Lynn Sage Cancer Research Foundation, Avon Foundation, and a generous donation from Ms. Dariel Eklund.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked

*advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 12, 2012; revised January 17, 2013; accepted January 18, 2013; published OnlineFirst March 19, 2013.

## References

- Hankinson SE, Colditz GA, Willett WC. Towards an integrated model for breast cancer etiology: the lifelong interplay of genes, lifestyle, and hormones. *Breast Cancer Res* 2004;6:213–8.
- Vona-Davis L, Howard-McNatt M, Rose DP. Adiposity, type 2 diabetes and the metabolic syndrome in breast cancer. *Obes Rev* 2007;8:395–408.
- Friedenreich CM, Orenstein MR. Physical activity and cancer prevention: etiologic evidence and biological mechanisms. *J Nutr* 2002;132:345S–64S.
- Kendall A, Folkert EJ, Dowsett M. Influences on circulating oestrogens in postmenopausal women: relationship with breast cancer. *J Steroid Biochem Mol Biol* 2007;103:99–109.
- 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.
- Chlebowski RT, Anderson GL, Lane DS, Aragaki AK, Rohan T, Yassmeen S, et al. Predicting risk of breast cancer in postmenopausal women by hormone receptor status. *J Natl Cancer Inst* 2007;99:1695–705.
- Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989;81:1879–86.
- Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. *N Engl J Med* 1985;312:146–51.
- Hartmann LC, Sellers TA, Frost MH, Lingle WL, Degnim AC, Ghosh K, et al. Benign breast disease and the risk of breast cancer. *N Engl J Med* 2005;353:229–37.
- McCormack VA, dos SS I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;15:1159–69.
- Kerlikowske K, Ichikawa L, Miglioretti DL, Buist DS, Vacek PM, Smith-Bindman R, et al. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. *J Natl Cancer Inst* 2007;99:386–95.
- Becker S, Kaaks R. Exogenous and endogenous hormones, mammographic density and breast cancer risk: can mammographic density be considered an intermediate marker of risk? *Recent Results Cancer Res* 2009;181:135–57.
- Stark A, Hulka BS, Joens S, Novotny D, Thor AD, Wold LE, et al. HER-2/neu amplification in benign breast disease and the risk of subsequent breast cancer. *J Clin Oncol* 2000;18:267–74.
- Gobbi H, Dupont WD, Simpson JF, Plummer WD Jr, Schuyler PA, Olson SJ, et al. Transforming growth factor-beta and breast cancer risk in women with mammary epithelial hyperplasia. *J Natl Cancer Inst* 1999;91:2096–101.
- Visscher DW, Pankratz VS, Santisteban M, Reynolds C, Ristimaki A, Vierkant RA, et al. Association between cyclooxygenase-2 expression in atypical hyperplasia and risk of breast cancer. *J Natl Cancer Inst* 2008;100:421–7.
- 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;90:1371–88.
- Vogel VG, Costantino JP, Wickerham DL, Cronin WM, Cecchini RS, Atkins JN, et al. Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes: the NSABP Study of Tamoxifen and Raloxifene (STAR) P-2 trial. *JAMA* 2006;295:2727–41.
- Goss PE, Ingle JN, Ales-Martinez JE, Cheung AM, Chlebowski RT, Wactawski-Wende J, et al. Exemestane for breast-cancer prevention in postmenopausal women. *N Engl J Med* 2011;364:2381–91.
- Freedman AN, Graubard BI, Rao SR, McCaskill-Stevens W, Ballard-Barbash R, Gail MH. Estimates of the number of US women who could benefit from tamoxifen for breast cancer chemoprevention. *J Natl Cancer Inst* 2003;95:526–32.
- Swain SM, Wilson JW, Mamounas EP, Bryant J, Wickerham DL, Fisher B, et al. Estrogen receptor status of primary breast cancer is predictive of estrogen receptor status of contralateral breast cancer. *J Natl Cancer Inst* 2004;96:516–23.
- Kurian AW, McClure LA, John EM, Horn-Ross PL, Ford JM, Clarke CA. Second primary breast cancer occurrence according to hormone receptor status. *J Natl Cancer Inst* 2009;101:1058–65.
- Bouchardy C, Benhamou S, Fioretta G, Verkooijen HM, Chappuis PO, Neyroud-Caspar I, et al. Risk of second breast cancer according to estrogen receptor status and family history. *Breast Cancer Res Treat* 2011;127:233–41.
- Euhus DM, Cler L, Shivapurkar N, Milchgrub S, Peters GN, Leitch AM, et al. Loss of heterozygosity in benign breast epithelium in relation to breast cancer risk. *J Natl Cancer Inst* 2002;94:858–60.
- Fabian CJ, Kimler BF, Zalles CM, Klomp JR, Kamel S, Zeiger S, et al. Short-term breast cancer prediction by random periareolar fine-needle aspiration cytology and the Gail risk model. *J Natl Cancer Inst* 2000;92:1217–27.
- Masood S, Frykberg ER, McLellan GL, Scalapino MC, Mitchum DG, Bullard JB. Prospective evaluation of radiologically directed fine-needle aspiration biopsy of nonpalpable breast lesions. *Cancer* 1990;66:1480–7.
- The uniform approach to breast fine-needle aspiration biopsy. National Cancer Institute fine-needle aspiration of breast workshop subcommittees. *Diagn Cytopathol* 1997;16:295–311.
- Khan QJ, Kimler BF, Clark J, Metheny T, Zalles CM, Fabian CJ. Ki-67 expression in benign breast ductal cells obtained by random periareolar fine needle aspiration. *Cancer Epidemiol Biomarkers Prev* 2005;14:786–9.
- Sharma P, Kimler BF, Warner C, Metheny T, Xue Q, Zalles CM, et al. Estrogen receptor expression in benign breast ductal cells obtained from random periareolar fine needle aspiration correlates with menopausal status and cytomorphology index score. *Breast Cancer Res Treat* 2006;100:71–6.
- Bean GR, Bryson AD, Piliie PG, Goldenberg V, Baker JC Jr, Ibarra C, et al. Morphologically normal-appearing mammary epithelial cells obtained from high-risk women exhibit methylation silencing of INK4a/ARF. *Clin Cancer Res* 2007;13:6834–41.
- Bean GR, Ibarra DC, Goldenberg VK, Baker JC Jr, Troch MM, Paisie C, et al. Hypermethylation of the breast cancer-associated gene 1 promoter does not predict cytologic atypia or correlate with surrogate end points of breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2007;16:50–6.
- Baker JC Jr, Ostrander JH, Lem S, Broadwater G, Bean GR, D'Amato NC, et al. ESR1 promoter hypermethylation does not predict atypia in RPFNA nor persistent atypia after 12 months tamoxifen chemoprevention. *Cancer Epidemiol Biomarkers Prev* 2008;17:1884–90.
- Vasilatos SN, Broadwater G, Barry WT, Baker JC Jr, Lem S, Dietze EC, et al. CpG island tumor suppressor promoter methylation in non-BRCA-associated early mammary carcinogenesis. *Cancer Epidemiol Biomarkers Prev* 2009;18:901–14.
- Euhus DM, Bu D, Milchgrub S, Xie XJ, Bian A, Leitch AM, et al. DNA methylation in benign breast epithelium in relation to age and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2008;17:1051–9.
- Lewis CM, Cler LR, Bu DW, Zochbauer-Muller S, Milchgrub S, Naftalis EZ, et al. Promoter hypermethylation in benign breast epithelium in relation to predicted breast cancer risk. *Clin Cancer Res* 2005;11:166–72.

35. Petroff BK, Phillips TA, Kimler BF, Fabian CJ. Detection of biomarker gene expression by real-time polymerase chain reaction using amplified ribonucleic acids from formalin-fixed random periareolar fine needle aspirates of human breast tissue. *Anal Quant Cytol Histol* 2006;28:297–302.
36. 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.
37. Graham K, de Las MA, Tripathi A, King C, Kavanah M, Mendez J, et al. Gene expression in histologically normal epithelium from breast cancer patients and from cancer-free prophylactic mastectomy patients shares a similar profile. *Br J Cancer* 2010;102:1284–93.
38. Graham K, Ge X, de Las MA, 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;17:236–46.
39. Tripathi A, King C, de la Morenas A, Perry VK, Burke B, Antoine GA, et al. Gene expression abnormalities in histologically normal breast epithelium of breast cancer patients. *Int J Cancer* 2008;122:1557–66.
40. Fagherazzi G, Chabbert-Buffet N, Fabre A, Guillas G, Boutron-Ruault MC, Mesrine S, et al. Hip circumference is associated with the risk of premenopausal ER–/PR– breast cancer. *Int J Obes (Lond)* 2012;36:431–9.
41. Eliassen AH, Missmer SA, Tworoger SS, Hankinson SE. Circulating 2-hydroxy- and 16alpha-hydroxy estrone levels and risk of breast cancer among postmenopausal women. *Cancer Epidemiol Biomarkers Prev* 2008;17:2029–35.
42. Chlebowski RT, Blackburn GL, Thomson CA, Nixon DW, Shapiro A, Hoy MK, et al. Dietary fat reduction and breast cancer outcome: interim efficacy results from the Women's Intervention Nutrition Study. *J Natl Cancer Inst* 2006;98:1767–76.
43. Blevins-Primeau AS, Sun D, Chen G, Sharma AK, Gallagher CJ, Amin S, et al. Functional significance of UDP-glucuronosyltransferase variants in the metabolism of active tamoxifen metabolites. *Cancer Res* 2009;69:1892–900.
44. Mitra PS, Basu NK, Owens IS. Src supports UDP-glucuronosyltransferase-2B7 detoxification of catechol estrogens associated with breast cancer. *Biochem Biophys Res Commun* 2009;382:651–6.
45. 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:R65.
46. Deisenroth C, Thormer AR, Enomoto T, Perou CM, Zhang Y. Mitochondrial Hep27 is a c-Myb target gene that inhibits Mdm2 and stabilizes p53. *Mol Cell Biol* 2010;30:3981–93.
47. Greenland KJ, Jantke I, Jenatschke S, Bracken KE, Vinson C, Gellersen B. The human NAD+-dependent 15-hydroxyprostaglandin dehydrogenase gene promoter is controlled by Ets and activating protein-1 transcription factors and progesterone. *Endocrinology* 2000;141:581–97.
48. Lehtinen L, Vainio P, Wikman H, Reemts J, Hilvo M, Issa R, et al. 15-hydroxyprostaglandin dehydrogenase associates with poor prognosis in breast cancer, induces epithelial-mesenchymal transition and promotes cell migration in cultured breast cancer cells. *J Pathol* 2012;226:674–86.
49. Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. *Science* 1996;274:2057–9.

# Cancer Prevention Research

## Lipid Metabolism Genes in Contralateral Unaffected Breast and Estrogen Receptor Status of Breast Cancer

Jun Wang, Denise Scholtens, Michelle Holko, et al.

*Cancer Prev Res* 2013;6:321-330. Published OnlineFirst March 19, 2013.

<b>Updated version</b>	Access the most recent version of this article at: <a href="https://doi.org/10.1158/1940-6207.CAPR-12-0304">doi:10.1158/1940-6207.CAPR-12-0304</a>
<b>Supplementary Material</b>	Access the most recent supplemental material at: <a href="http://cancerpreventionresearch.aacrjournals.org/content/suppl/2021/03/17/1940-6207.CAPR-12-0304.DC1">http://cancerpreventionresearch.aacrjournals.org/content/suppl/2021/03/17/1940-6207.CAPR-12-0304.DC1</a>

<b>Cited articles</b>	This article cites 49 articles, 15 of which you can access for free at: <a href="http://cancerpreventionresearch.aacrjournals.org/content/6/4/321.full#ref-list-1">http://cancerpreventionresearch.aacrjournals.org/content/6/4/321.full#ref-list-1</a>
-----------------------	--

<b>Citing articles</b>	This article has been cited by 2 HighWire-hosted articles. Access the articles at: <a href="http://cancerpreventionresearch.aacrjournals.org/content/6/4/321.full#related-urls">http://cancerpreventionresearch.aacrjournals.org/content/6/4/321.full#related-urls</a>
------------------------	---

<b>E-mail alerts</b>	<a href="#">Sign up to receive free email-alerts</a> related to this article or journal.
----------------------	--

<b>Reprints and Subscriptions</b>	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at <a href="mailto:pubs@aacr.org">pubs@aacr.org</a> .
-----------------------------------	--

<b>Permissions</b>	To request permission to re-use all or part of this article, use this link <a href="http://cancerpreventionresearch.aacrjournals.org/content/6/4/321">http://cancerpreventionresearch.aacrjournals.org/content/6/4/321</a> . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.
--------------------	--