Tumorigenesis and Neoplastic Progression

Early Dysregulation of Cell Adhesion and Extracellular Matrix Pathways in Breast Cancer Progression

Lyndsey A. Emery,* Anusri Tripathi,* Chialin King,*
Maureen Kavanah,[†] Jane Mendez,[†]
Michael D. Stone,[†] Antonio de las Morenas,[‡]
Paola Sebastiani,[§] and Carol L. Rosenberg^{**}

From the Departments of Medicine,* Surgery,[†] Pathology,[‡]
and Biostatistics,[§] Boston University Medical Center,
Boston, Massachusetts

Proliferative breast lesions, such as simple ductal hyperplasia (SH) and atypical ductal hyperplasia (ADH), are candidate precursors to ductal carcinoma *in situ* (DCIS) and invasive cancer. To better understand the relationship of breast lesions to more advanced disease, we used microdissection and DNA microarrays to profile the gene expression of patient-matched histologically normal (HN), ADH, and DCIS from 12 patients with estrogen receptor positive sporadic breast cancer. SH were profiled from a subset of cases. We found 837 differentially expressed genes between DCIS-HN and 447 between ADH-HN, with >90% of the ADH-HN genes also present among the DCIS-HN genes. Only 61 genes were identified between ADH-DCIS. Expression differences were reproduced in an independent cohort of patient-matched lesions by quantitative real-time PCR. Many breast cancer-related genes and pathways were dysregulated in ADH and maintained in DCIS. Particularly, cell adhesion and extracellular matrix interactions were overrepresented. Focal adhesion was the top pathway in each gene set. We conclude that ADH and DCIS share highly similar gene expression and are distinct from HN. In contrast, SH appear more similar to HN. These data provide genetic evidence that ADH (but not SH) are often precursors to cancer and suggest cancer-related genetic changes, particularly adhesion and extracellular matrix pathways, are dysregulated before invasion and even before malignancy is apparent. These findings could lead to novel risk stratification, prevention, and treatment approaches. (Am J Pathol 2009, 175:1292–1302; DOI: 10.2353/ajpath.2009.090115)

Terminal ductal lobular units are the functional unit of the breast. Each lobule is composed of a layer of luminal epithelial cells surrounded by a basal layer of myoepithelial cells (MECs), and supported by a heterogeneous stromal microenvironment. One widely accepted model of breast cancer progression proposes that normal epithelial cells of the terminal ductal lobular unit move through a series of distinct precursors that acquire increasingly aberrant histological and molecular abnormalities, resulting in invasive cancer.^{1,2}

In an attempt to understand these molecular abnormalities, extensive gene expression studies have been done in invasive breast cancers. These studies have revealed that invasive breast cancer is a heterogeneous disease that can be subclassified based on distinct molecular features.³ These classifications have been repeatedly correlated with clinically relevant tumor phenotypes, including hormone receptor status and histological grade.³ Understanding the molecular changes that occur before the development of invasive cancer is important and is the focus of our studies.

Ductal carcinoma *in situ* (DCIS) is a nonobligate precursor to invasive breast cancer.⁴ Approximately 50% of DCIS will progress to invasive cancer if left untreated, and 12% to 20% will recur within 10 years despite appropriate treatment.⁵ Gene expression studies have revealed similarities between DCIS and invasive ductal carcinoma (IDC).^{6,7} In fact, DCIS have been found to be

Supported by DAMD17-01-1-0159 and PHS CA115434 to C.L.R. and W81XWH-09-1-0019 to L.A.E.

Accepted for publication June 16, 2009.

This work was presented in part at the 98th Annual Meeting of the United States and Canadian Academy of Pathology, Boston, Massachusetts, March 7, 2009.

Supplemental material for this article can be found on <http://ajp.amjpathol.org>.

Current address of A.T.: Department of Biochemistry and Medical Biotechnology, Calcutta School of Tropical Medical, Kolkata, West Bengal, India.

Address reprint requests to Carol L. Rosenberg, Boston University Medical Center, 650 Albany Street, EBRC 4th Floor, Boston, MA 02118. E-mail: crosenbe@bu.edu.

more similar to their IDC component than to other DCIS, suggesting that DCIS are both direct precursors to invasive cancer and are highly heterogeneous.^{2,6–9} Although what determines the transition from DCIS to IDC remains uncertain, MECs, stromal microenvironment, and extracellular matrix (ECM) may play an important role.¹⁰

Despite advances in understanding the progression of DCIS to invasive cancer, current knowledge of earlier events in breast tumorigenesis is limited. Proliferative breast lesions, such as simple ductal hyperplasia (SH) and atypical ductal hyperplasia (ADH), may be important precursors to DCIS and IDC. ADH are morphologically heterogeneous and have some, but not all, of the cytologic and architectural features of DCIS.¹¹ An ADH diagnosis confers up to a fourfold increased risk of breast cancer, but only 15% of women with ADH develop invasive disease.^{12,13} Little is known about ADH's genetic abnormalities. Loss of heterozygosity studies reveal that ADH are genetically related, although not identical, to DCIS.^{11,14–16} Gene expression analyses of ADH are severely limited. We are aware of only two studies: one including nine patient-matched ADH (whose data are not publicly available), and another including eight samples limited to only ADH.^{6,17}

SH represent a large, heterogeneous group of proliferative breast lesions. While very common, only 5% of women diagnosed with SH eventually develop carcinomas.¹² There are relatively few studies of these lesions' genetic alterations, and interpretation of their results is complicated because some report few chromosomal abnormalities and others report many.¹⁸ Little is known about the genetic relationship of these lesions with cancer. While a small subset of SH may be clonal precursors to invasive cancer, supporting evidence remains weak.¹⁸

Taken together, there is a paucity of information regarding the gene expression abnormalities associated with breast cancer initiation and progression. Using microdissection to isolate co-existing, or patient-matched, histologically normal (HN), SH (when available), ADH, and DCIS epithelium from estrogen receptor positive (ER+) sporadic breast cancers, we characterized each lesion's gene expression profile. This design accounts for many patient-specific variables such as age, medication use, and individual genomic variation. Our goal is to reveal the genetic relationships between normal epithelium, proliferative lesions, and preinvasive breast cancer (and by extension invasive disease). Ultimately, we aim to better understand breast cancer progression, improve precursor risk stratification, and identify novel prevention and treatment strategies.

Materials and Methods

Tissue Acquisition

Primary breast tissue not needed for diagnosis was obtained with appropriate institutional review board approval from surgical oncology patients at Boston University Medical Center. Once surgical specimen were procured, tissue was snap frozen in liquid nitrogen and

immediately embedded in frozen tissue blocks by using TissueTek OCT media (VWR Scientific Products Corporation, San Diego, CA) to ensure RNA integrity.

Sectioning and Pathological Diagnosis

Twelve cases were chosen based on a diagnosis of ER+ DCIS or IDC and the absence of chemo- or radiation therapy before surgery. Frozen tissue blocks were cryosectioned into serial 10 μ m sections. H&E-stained sections were reviewed by a breast pathologist who identified lesions of interest. Each case contained HN, ADH, and DCIS (some cases also contained SH). DCIS were assigned a grade and histology according to the Van Nuys classification scheme.¹⁹ When present, concurrent IDC was classified according to the Nottingham combined histological grade scheme.²⁰ Estrogen Receptor and Progesterone Receptor status was determined by immunohistochemistry and human epidermal growth factor receptor 2 (Her2) status was determined by a combination of immunohistochemistry and fluorescence *in situ* hybridization. Without an IDC component present, DCIS Her2 status was unavailable.

Laser Capture Microdissection, RNA Isolation, and RNA Amplification

HN, SH (when available), ADH, and DCIS lesions were isolated via laser capture microdissection (LCM) from consecutive sections by using the PixCell IIe LCM system (Arcturus Engineering, Mountain View, CA) according to our standard procedures.^{21,22} Briefly, serial sections were mounted on uncoated glass slides and immediately stored at -80°C . Slides were lightly stained with diluted H&E to enable visualization and to prevent nucleotide alterations introduced from H&E. Lesions were isolated via LCM from 30 to 50 consecutive slides, depending on the size of the lesion and its continuity throughout the block. In one case, (380) HN microdissected from different blocks was pooled. The presence of any IDC, lymphocyte infiltration, and/or necrosis was avoided, and extreme care was taken to keep every sample separate to avoid contamination.

Following the manufacturer's protocol, total RNA was extracted and purified by using the Picopure RNA Isolation Kit (Arcturus Engineering). RNA concentration was quantitated by using the NanoDrop ND-100 Spectrophotometer (NanoDrop Technologies, Wilmington, DE) and its quality was assessed by using the RNA Pico Chips and the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). To convert the RNA to cDNA, the purified total RNA was linearly amplified for two rounds by using the MessageAMP aRNA Kit (Ambion, Austin, TX), in which the first round is based on an oligo-dT primer and T7 RNA polymerase. Our lab has successfully used amplified RNA from microdissected samples for microarray analysis from samples obtained by LCM with a high degree of reliability and reproducibility.^{21,22}

Oligonucleotide Microarray Hybridization and Data Acquisition

For second round amplification, cDNA for each sample was *in vitro* transcribed to biotin-labeled cRNA with the *in vitro* transcribed labeling kit (Affymetrix, Santa Clara, CA). Gene expression was determined by using the U133A GeneChip microarray (Affymetrix) using standard procedures for quality control, hybridization, staining, and scanning.²² The CEL files for each sample (40 total) were processed with MAS5.0. The unnormalized data were then rescaled to a target of 100 to remove bias. The microarray data from these samples are available from the National Center for Biotechnology Information Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) under accession GSE16873.

Statistical Analysis

The data were analyzed in the following three comparisons: DCIS-HN; ADH-HN; and ADH-DCIS. An initial filter of the microarray output removed all probesets with <15% present call rates across samples. Remaining probesets were analyzed by using a two-step procedure. The data were first analyzed by using Bayesian Analysis of Differential Gene Expression (BADGE) to identify a smaller set of probesets to be analyzed with a more sophisticated, but computationally intensive, procedure.²³ BADGE uses an efficient model-averaging approach to identify probesets with different expression in two biological conditions. The model-averaging approach is used to calculate the posterior probability of a fold-change >1, so that probesets in which this probability is either very small (<0.025) or very large (>0.975) can be taken as differentially expressed. This method has a very large sensitivity but lower specificity because it employs large sample approximation. We have shown that with 10 samples per group the sensitivity to detect a fold-change of two or larger can be 100%, but the specificity can be <70%.²³ Therefore, from the BADGE analysis, we selected the set of probesets in which the probability of a fold-change >1 was either >0.90 or <0.10. These relaxed thresholds were chosen to also take into account the fact that patient-matching was not included in the analysis and could impact the evidence for differential expression. These probesets were further analyzed by using a regression model with log-normal errors, covariates to represent the biological conditions, and a random effect per patient to take into account patient-matching across all probesets. The model was fitted by using a stochastic algorithm known as the Gibbs sampling that is implemented in WinBugs 1.4. The Gibbs sampling was run for 10,000 iterations that were sufficient to reach convergence. The values generated in the 10,000 iterations were used to estimate the fold-change (in the real scale), the 95% credible intervals (CI), and the SE evidence of differential expression was based on 95% CI nonoverlapping one. Heatmaps were generated by using the HeatPlus software package (Bioconductor, Se-

attle, WA). Simple hierarchical clustering was used to cluster samples based on their expression profiles.

Gene Annotation and Pathway Analysis

Gene annotation and pathway analyses were conducted by using several approaches including the following: Gene Ontology (www.geneontology.org); the Database for Annotation, Visualization and Integrated Discovery (DAVID; <http://david.abcc.ncifcrf.gov>); and the Ingenuity Pathway Analysis software, version 5.5 (Ingenuity Systems, Redwood City, CA). *P* values reported from DAVID were calculated by using a modified Fisher Exact test (EASE Score) to measure the gene enrichment in annotation terms.^{24,25}

Candidate Gene and Endogenous Control Gene Selection

Genes were selected for validation based on the degree, consistency, and direction of differential gene expression between HN, ADH, and/or DCIS; whether there was overlap between several comparisons; potential pathway involvement; and potential role in tumorigenesis. Using these criteria, we chose six genes for validation: fibronectin 1 (*FN1*), G-protein coupled receptor 5A (*GPCR5A*), calpain 6 (*CAPN6*), membrane metallo-endopeptidase (*MME*), oxytocin receptor (*OXTR*), and cyclin-dependent kinase inhibitor 1C/p57^{KIP2} (*CDKN1C*). 5-methyltetrahydrofolate-homocysteine methyltransferase (*MTR*) was used as the endogenous control gene, because the microarray indicated stable expression in all samples.²⁶

Quantitative Real-Time PCR Validation

For initial validation, we used the original, unamplified RNA from 8 of the 12 cases analyzed by microarray. For prospective validation, 12 independent breast cancer cases were chosen based on the same criteria used for initial case selection. Following the manufacturer's protocol, total RNA was extracted and purified from these independent cases by using a different kit (RNAqueous-Micro Kit, Ambion).

Reverse transcription (RT) was conducted in 50 μ l volumes by using the appropriate amount of unamplified total RNA required for a given reaction plate (2 ng cDNA/reaction well for quantitative real-time PCR). RT was performed by using random hexamer primers from the TaqMan Reverse Transcription Reagent Kit (Applied Biosystems, Foster City, CA) and standard thermocycler conditions.²² Each real-time PCR reaction was performed in 20 μ l volumes by using PCR Universal Master Mix and TaqMan Gene Expression Assays (Applied Biosystems) for the gene to be validated. Using the criteria described, we selected the following primers: *FN1* (Hs00415008_m1); *GPCR5A* (Hs01551896_m1); *CAPN6* (Hs01020899_m1); *MME* (Hs01115449_g1); *OXTR* (Hs00168573_m1); *CDKN1C* (Hs00908986_g1); and *MTR* (Hs01090032_g1). TaqMan Gene Expression Assays (Applied Biosystems)

Table 1. Clinical and Histopathological Features of Patient Samples Used in Microarray Studies

Case number	Patient age	Samples	ER/PR/Her2 status	DCIS histology	DCIS grade	IDC grade
444	48	HN*, SH, ADH*, DCIS*	+ / + / NA	Solid	2 and 3	Not present
248	49	HN*, ADH*, DCIS	+ / + / -	Micropapillary	2	2
274	49	HN*, ADH*, DCIS*	+ / + / -	Cribriform and micropapillary	2	1
380	53	HN*, SH, ADH*, DCIS*	+ / + / +	Micropapillary	2	2
237	55	HN, ADH, DCIS	+ / - / -	Comedo	2	2
232	59	HN, ADH, DCIS	+ / + / NA	Micropapillary	2	Not present
226	61	HN*, ADH*, DCIS	- / - / NA	Micropapillary	2	3
258	65	HN, ADH*, DCIS*	+ / + / -	Cribriform	2	2
405	67	HN*, SH, ADH*, DCIS*	+ / + / -	Comedo and cribriform	2 and 3	2
275	85	HN, SH, ADH, DCIS	+ / + / -	NA	2	2
228	92	HN*, ADH*, DCIS*	+ / - / -	Solid	3	3
323	92	HN, ADH, DCIS	+ / - / -	Micropapillary	1	NA

*Indicates samples with sufficient RNA remaining and inclusion in initial validation.

HN, histologically normal; SH, simple ductal hyperplasia; ADH, atypical ductal hyperplasia; DCIS, ductal carcinoma *in situ*.

used for the validation studies are indicated in parentheses. The Prism 7500 Sequence Detector System (Applied Biosystems) was used to measure fluorescence under standard cycling parameters.²² Each condition was analyzed in duplicate with appropriate controls (-RT, -RNA, and no template) to conserve RNA.

Relative quantification by quantitative real-time PCR (qRT-PCR) was assessed by using the delta delta Ct (ddCt) method, and relative gene expression was calculated after normalization to *MTR*.²⁷ First, for both the technical and prospective validation studies, we compared changes in gene expression among the lesions as a group by using the same regression model of the ddCt data with a random effect per patient (to take into account patient-matching) as in the analysis of the microarray data. The Gibbs sampling was used to estimate the ddCt, the fold-changes as $2^{(-ddCt)}$, 95% CI, and SEs. Gene expression differences were considered to validate if the direction of change was the same and if the fold-change estimated from the microarray data were within two SEs from that estimated by qRT-PCR.

Next, we compared the expression of each gene in each patient-matched pair used for validation by qRT-PCR. Gene expression differences were considered to validate if the direction of change was the same (as estimated by microarray) and if the fold-change was ≥ 1.5 (as measured by the ddCt method).

Results

Clinical and Histopathological Features of Cases Used for Microarray

Summary patient information for the 12 cases used for microarray studies can be seen in Table 1. Patient age ranged from 48 to 92 years. A total of 40 samples (12 HN, 4 SH, 12 ADH, and 12 DCIS) were identified and analyzed by microarray. All cancers were ER+ and most were also progesterone receptor positive (PR+) and Her2 negative (Her2-). The only exception is case 226, which was believed to be ER+ PR+ at the time of microarray, but later determined to be ER- PR-. DCIS

histology and grade reflects the specific lesion microdissected. Representative lesions are depicted in Figure 1.

Gene Expression Profiles

Microarray Data

The following three primary comparisons were made: DCIS-HN; ADH-HN; and ADH-DCIS. BADGE identified a subset of 1850 differentially expressed probesets. Further analysis, accounting for patient-matching, identified 1048 probesets (representing 883 genes) that changed in at least one comparison and had a fold-change ≥ 1.5 in either direction. A comprehensive probeset list can be seen in Supplemental Table 1 at <http://ajp.amjpathol.org>. Gene expression changes are related to the lesion listed first (ie, "up" in ADH-HN indicates overexpression in ADH relative to HN) and all reported fold-changes are in the real scale.

Figure 1. Representative lesions from four cases used for microarray and independent validation studies. Case numbers correspond to an individual patient. H&E stained 10 μ m sections; original magnification, $\times 200$.

Figure 2. Gene expression profiles by hierarchical clustering. **A:** Heatmap with HN, ADH, and DCIS lesions. **B:** Heatmap with HN, SH, ADH, and DCIS lesions. Patient samples are labeled along the **top** of the heatmap by case number and lesion type. Each of the 1048 probesets differentially expressed in at least one comparison is presented as a separate row. The green color represents relative gene overexpression, while the red color represents relative gene underexpression.

ADH and DCIS grouped mostly together by hierarchical clustering, while the HN expression profile was distinct from both (Figure 2A). Within the ADH-DCIS grouping, ADH and DCIS lesions from the same patient clustered together. These results indicate that ADH and DCIS share highly similar expression profiles regardless of whether they are derived within or between patients. Patient age did not influence gene expression similarities between ADH-DCIS (see Supplemental Figure 1A at <http://ajp.amjpathol.org>). Of interest, three ADH (cases 405, 275, and 228) clustered with HN. No technical or clinical variables (eg, patient age, PR/Her2 status, DCIS grade or histology, presence of IDC) accounting for this were identified. While these ADH were from older patients, there was one ADH sample from another 92 year-old patient that clustered with DCIS. The next youngest ADH sample was from a 65 year-old patient and also clustered with DCIS. For these reasons, we favor that these ADH did not resemble HN from any age effect, and may be biologically different. Of note is that case 226 shares a similar expression profile to the other ER+ cases.

Among the 883 genes identified, there was differential expression of 837 genes between DCIS-HN, 447 genes between ADH-HN, and 61 genes between ADH-DCIS. There was a tendency toward underexpression in DCIS-HN and ADH-HN, with approximately 2/3 of the genes underexpressed in the more histologically ab-

normal lesion. Over 90% (407 of 447) of the ADH-HN genes are also in the DCIS-HN gene set. Thus, the great majority of ADH expression abnormalities appear to be cancer-related, since they are also present in DCIS.

Initial Validation of Microarray Data using qRT-PCR

To confirm the microarray, qRT-PCR was used to measure the relative expression of six genes (overexpressed: *FN1*, *GPRC5A*; underexpressed: *CAPN6*, *MME*, *OXTR*, *CDKN1C*) on unamplified RNA from the original microarray cases (Table 1). These genes are involved in a wide range of cancer-related functions including cancer marker (*MME*), hormone receptor (*OXTR*), focal adhesion (*FN1*, *CAPN6*), cell signaling (*GPRC5A*), and cell cycle regulation (*CDKN1C*). qRT-PCR results are summarized in Supplemental Figure 2 at <http://ajp.amjpathol.org>. When we compared the gene expression differences between each patient-matched pair used in the technical validation, a total of 58 of 65 (89%) patient-matched lesions validated the differential expression predicted by microarray. These results verify the accuracy of the microarray data and indicate that amplification did not misrepresent true expression differences.

Table 2. Most Differentially Expressed Genes and Overrepresented Pathways for Each Comparison

Gene set	Underexpressed genes (Fold change)	Overexpressed genes (Fold change)	Overrepresented KEGG pathways	No. of genes	% genes	P value (EASE score)
DCIS-HN	<i>MME</i> * (0.08)	<i>TOP2A</i> (6.5)	Focal adhesion (hsa04510)	33	3.94	1.2E-06
	<i>IGJ</i> (0.12)	<i>NEK2</i> (5.8)	Regulation of actin cytoskeleton (hsa04810)	22	2.63	3.9E-02
	<i>SCN3A</i> (0.13)	<i>CRIP1</i> (5.7)	Cell communication (hsa01430)	17	2.03	1.5E-02
	<i>OXTR</i> * (0.13)	<i>SLC27A2</i> (5.3)	Cell cycle (hsa04110)	16	1.91	6.8E-03
	<i>HOXA9</i> (0.15)	<i>GPRC5A</i> * (4.7)	Small cell lung cancer (hsa05222)	15	1.79	1.3E-03
ADH-HN	<i>MME</i> * (0.12)	<i>COL10A1</i> (3.9)	Focal adhesion	26	5.84	1.4E-09
	<i>OXTR</i> * (0.24)	<i>GPRC5A</i> * (3.8)	Regulation of actin cytoskeleton	15	3.37	7.0E-03
	<i>SCN3A</i> (0.24)	<i>CRIP1</i> (3.3)	ECM-receptor interaction (hsa04512)	11	2.47	3.9E-04
ADH-DCIS	<i>CHL1</i> (0.24)	<i>SLC27A2</i> (3.2)	Prostate cancer (hsa05215)	11	2.47	3.5E-04
	<i>LMOD1</i> (0.26)	<i>LYZ</i> /// <i>LILRB1</i> (3.2)	Cell communication	10	2.25	2.7E-02
	<i>TOP2A</i> (0.35)	<i>IGJ</i> (4.0)	Focal adhesion	5	8.06	1.5E-02
	<i>NEK2</i> (0.43)	<i>MME</i> * (3.6)	p53 signaling (hsa04115)	3	4.84	4.4E-02
	<i>NCAPG</i> (0.45)	<i>EREG</i> (3.1)				
	<i>RRM2</i> (0.46)	<i>SAA1</i> /// <i>SAA2</i> (3.0)				
	<i>MELK</i> (0.49)	<i>ADH1B</i> (2.9)				
Nonoverlapping genes from DCIS-HN	<i>WIF1</i> (0.13)	<i>TOP2A</i> (6.5)	Cell cycle	12	2.75	3.0E-03
	<i>IGHA1</i> /// <i>IGHA2</i> (0.21)	<i>NEK2</i> (5.8)	p53 signaling pathway	9	2.06	3.3E-03
	<i>OLFM4</i> (0.23)	<i>E2F8</i> (4.4)	Small cell lung cancer	8	1.83	4.1E-02
	<i>ADH1B</i> (0.23)	<i>RRM2</i> (4.1)				
	<i>KIT</i> (0.26)	<i>CEP55</i> (3.9)				

Note: In a given gene set, the lesion on the right is the reference. For example, *TOP2A* is overexpressed in DCIS relative to HN, and underexpressed in ADH relative to DCIS. The average fold-change is reported for genes with multiple probesets. Each probeset for these genes was differentially expressed in the same direction and to the same degree.

*Indicates genes used in validation studies.

Gene Annotation and Pathway Analysis

Gene annotation and pathway analysis identified the dysregulation of several biologically important genes and pathways. Results were similar among various pathway analysis tools. Here we report the results from DAVID, because this tool provided the most succinct summary of overrepresented pathways (Table 2). We found several cancer-related pathways including cell cycle, ECM-receptor interaction, and p53 signaling (Table 2). Pathways involving cell adhesion and ECM were implicated repeatedly. These included focal adhesion, regulation of actin cytoskeleton, cell communication, and ECM-receptor interaction. Many of these pathways were dysregulated in ADH-HN and sustained in DCIS (Table 2). Focal adhesion was identified as the most overrepresented Kyoto Encyclopedia of Genes and Genomes pathway in each gene set. Table 3 presents selected differentially expressed genes involved in cell adhesion and ECM interactions.

In contrast, overrepresentation of adhesion and ECM-related genes and pathways was not found among the 437 DCIS-HN genes that do not overlap with the ADH-HN genes (Table 2). These nonoverlapping genes also had different Gene Ontology molecular function annotations than the other gene sets (data not shown). In addition, some pathways of known cancer relevance were not prominent among our gene sets (eg, maintenance of genomic stability, apoptosis, and damage repair). Together, these results suggest that adhesion and ECM interaction pathways are dysregulated early

during tumorigenesis of ER+ cancers, even in the transition to ADH.

Consistent with what we would expect to find in ER+ samples, markers of basal breast cancers (*KRT5*, *KRT14*, and *KRT17*) were underexpressed in DCIS.²⁸ Conversely, a marker of ER+ luminal cancers (*KRT19*) was overexpressed in DCIS.²⁸ Additionally, many genes of known breast cancer relevance were identified, including the following: *BRCA1*; *CCND1*; *EZH2*; *microRNA21* (all modestly overexpressed); and *FGFR2* (modestly underexpressed). Genes related to mammary "stemness" were absent (*CD44*, *ALDH1*), as were many genes that induce or characterize epithelial-mesenchymal transition (*Twist1*, *SNAI1*, *CDH1*, *VIM*).^{29,30} *SNAI2* was modestly underexpressed and *FN1* was overexpressed.

Prospective Validation of Microarray using qRT-PCR in Independent Cases

To further demonstrate that expression differences between HN, ADH, and DCIS were biologically meaningful, validation via qRT-PCR was performed by using unamplified total RNA from 29 microdissected lesions (12 HN, 8 ADH, and 9 DCIS) derived from 12 independent breast cancer cases. Representative lesions are shown in Figure 1. For a summary of patient information see Supplemental Table 2 at <http://ajp.amjpathol.org>.

We used the same six primers and criteria as in the initial validation. When comparing each of the six gene's expression in HN, ADH, and DCIS as groups, we found

Table 3. Selected Differentially Expressed Genes in Adhesion and ECM-Related Pathways

Gene	Fold change			Pathway*/Function
	DCIS-HN	ADH-HN	ADH-DCIS	
<i>ACTG2</i>	0.24	0.36	—	RAC
<i>ADAMTS1</i>	0.47	0.67	—	ECM protease
<i>CAPN6</i>	0.16	0.31	2.5	FA
<i>CAV1</i>	0.48	0.58	1.6	FA
<i>CAV2</i>	0.45	0.62	—	FA
<i>CCND2</i>	0.52	0.59	—	FA
<i>COL14A1</i>	0.37	—	2.2	Cell adhesion; ECM Structural component
<i>COL15A1</i>	0.48	0.51	—	Cell adhesion; ECM Structural component
<i>EGF</i>	0.50	0.51	—	FA, RAC
<i>EGFR</i>	0.34	0.47	—	FA, RAC
<i>FGF1</i>	0.39	0.55	—	RAC
<i>FGF2</i>	0.38	—	2.1**	RAC
<i>FGFR2</i>	0.52	0.45	—	RAC
<i>FIGF</i>	0.23	0.50	2.2	FA
<i>FN1</i>	3.8	2.8	—	FA, CC, ECM, RAC
<i>FYN</i>	0.57	0.54	—	FA
<i>IGF1</i>	0.35	0.44	—	FA
<i>ITGA10</i>	0.51	0.60	—	FA, ECM, RAC
<i>LAMA2</i>	0.47	—	1.7	FA, CC, ECM
<i>LAMA3</i>	0.53	0.60	—	FA, CC, ECM
<i>LAMB1</i>	—	0.51	—	FA, CC, ECM
<i>LAMB3</i>	0.32	0.38	—	FA, CC, ECM
<i>LAMC2</i>	0.37	0.50	—	FA, CC, ECM
<i>MME</i>	0.08	0.21	3.6	Cell communication; Cancer marker
<i>MYLK</i>	0.22	0.28	—	FA, RAC
<i>NCAM1</i>	0.40	0.51	—	Cell adhesion; Transmembrane molecule
<i>PAK3</i>	0.30	0.51	2.4	FA, RAC
<i>PDGFA</i>	0.46	0.49	—	FA, RA
<i>PDGFD</i>	0.44	0.48	—	FA
<i>PDGFRA</i>	0.31	—	2.6	FA, RAC
<i>PIK3R1</i>	0.52	0.58	—	FA, RAC
<i>PIK3R3</i>	1.7	1.7	—	FA, RAC
<i>PIP5K1B</i>	0.50	0.59	—	RAC
<i>PPP1R12B</i>	0.61	0.61	—	RAC
<i>RELN</i>	0.44	0.52	—	FA, CC, ECM
<i>SPP1</i>	—	2.0	—	FA, CC, ECM
<i>THBS1</i>	2.0	2.0	—	FA, CC, ECM
<i>TIAM1</i>	1.7	1.7	—	RAC
<i>TNN</i>	0.33	0.44	—	FA, CC, ECM
<i>TNXB</i>	0.44	0.47	—	FA, CC, ECM
<i>VCAM1</i>	0.52	0.49	—	Cell adhesion; Transmembrane molecule
<i>VEGFA</i>	1.5	1.7	—	FA

Note: The average fold-change is reported for genes with multiple probesets. Each probeset for these genes was differentially expressed in the same direction and to the same degree.

*DAVID pathways: FA = focal adhesion; CC = cell communication; and RAC = regulation of actin cytoskeleton.

***FGF2* is differentially expressed in ADH-DCIS; however, the RAC pathway was not identified as overrepresented for this gene set.

that each gene was significantly different in both DCIS and ADH compared with HN. There were no differences in gene expression between ADH and DCIS, but we noted trending of the fold-change as lesions become more histologically abnormal (Figure 3).

When we compared the gene expression differences between each patient-matched pair used in the prospective validation, a total of 93 of 106 (88%) patient-matched lesions validated the microarray. For DCIS-HN, 51 of 54 (94%) pairs were validated; for ADH-HN, 36 of 40 (90%) pairs were validated; and for ADH-DCIS, 6 of 12 (50%) pairs were validated. For detailed results see Supplemental Table 3 at <http://ajp.amjpathol.org>. For both DCIS-HN and ADH-HN, only case 345 did not validate. Review of the guide slides revealed many inflammatory cells surrounding the lesions; and, while the case had ductal features, lobular carcinoma *in situ* was microdis-

sected (see Supplemental Table 2 at <http://ajp.amjpathol.org>). These features may have confounded the validation of this case. In contrast, the weaker validation between ADH-DCIS could not be explained by outlier cases or other obvious reasons.

SH: Microarray Analysis and qRT-PCR Validation

RNA from SH samples was isolated and hybridized alongside their matched lesions in four microarray cases (Table 1). Because of the limited SH sample size, expression data were analyzed by using only the 1048 probesets identified in the HN-ADH-DCIS analyses. Hierarchical clustering including these SH suggested that three of the four SH samples more closely resembled HN than

Figure 3. Prospective validation of differential gene expression by qRT-PCR in independent cases. Fold-changes and direction of change were in accordance with the microarray (see inset). All genes were significantly differentially expressed between DCIS-HN, ADH-HN, and DCIS-SH. All genes, except *FN1*, were significantly differentially expressed between ADH-SH. No genes were significantly differentially expressed between ADH-DCIS or SH-HN. Error bars represent the SE estimated with Gibbs sampling. HN ($n = 12$), SH ($n = 6$), ADH ($n = 8$), and DCIS ($n = 9$).

ADH or DCIS; the fourth SH sample (case 444) did not closely resemble any of the other lesions studied (Figure 2B). Patient age did not influence gene expression similarities between SH-HN (see Supplemental Figure 1B at <http://ajp.amjpathol.org>).

For qRT-PCR validation, we identified SH in 6 of the 12 independent cases (see Supplemental Table 2 at <http://ajp.amjpathol.org>). When we compared gene expression in lesions as groups, we found that none of the six genes differed between SH-HN (Figure 3). In contrast, gene expression differed in four of the five genes (all except *FN1*) between ADH-SH; all six genes differed between DCIS-SH. When considering each patient-matched pair used in the prospective validation study, a total of 30 of 36 (83%) patient-matched SH-HN pairs showed no differential gene expression. These data provide evidence that SH expression resembles that of normal breast epithelium and has not acquired the changes characteristic of higher-risk precursors.

Discussion

There are gaps in our current understanding of early stages of breast tumorigenesis. To address them, we characterized gene expression profiles of patient-matched candidate precursors from patients with ER+ sporadic breast cancers to identify progression-specific changes. Our results show that ADH and DCIS share highly similar expression profiles and that most of the ADH-associated abnormalities are also present in DCIS. In contrast, SH resemble HN epithelium. Many cancer-related pathways are already dysregulated in ADH, before malignancy is apparent, with an overrepresentation of adhesion and ECM-related pathways. The early dysregulation of these pathways is, to our knowledge, a novel finding.

One gap in knowledge regarding early breast tumorigenesis is the genetic relationship between ADH and DCIS. In the only previous progression study we are aware of, just seven of the nine ADH were associated with the full spectrum of lesions.⁶ Also, of the ADH studied, most patient-matched DCIS were grade 1. This study reported few expression differences between ADH and DCIS, but the data are not available for comparison. Our study includes more ADH (plus other candidate precursors when available), only cases with the full spectrum of lesions (HN, ADH, DCIS), and more heterogeneous and higher grade DCIS (mostly grade 2). We expected our design would reveal robust progression-specific HN-ADH-DCIS expression differences. We find ADH and DCIS to be almost indistinguishable from one another. This is supported by the ~90% validation of expression differences between both DCIS-HN and ADH-HN, and only 50% validation between ADH-DCIS. Over 90% of the expression changes observed in ADH overlap those in DCIS. Despite a larger DCIS-HN gene set (~400 additional genes characterize DCIS-HN compared with ADH-HN), only 61 genes are different between ADH-DCIS. One possible explanation for these results is that ADH-DCIS expression differences are subtle and may require a larger sample size to detect. Another possibility is that these ADH are not distinct from concurrent DCIS, are direct cancer precursors, and already contain cancer-associated expression abnormalities.

However, epidemiological evidence suggests ADH are far from committed cancer precursors. In support of this, we also identified a small ADH subset whose expression more closely resembles normal epithelium. It is possible these ADH are at an earlier stage in progression or at a lower cancer risk than the other ADH. Consistent with this are results from a comparison of ADH from four patients

with cancer and four patients without cancer, identifying ~540 differentially expressed genes.¹⁷ This implies that factors other than ADH epithelial cell expression (eg, microenvironment, ~400 DCIS-specific genes) may influence ADH progression, and that these factors remain variable at this stage. It will be of interest to compare larger sets of ADH from patients with and without cancer, to determine whether their expression profiles can better stratify ADH and their risk for progression.

Despite heterogeneity within ADH, by including DCIS of varying grade and histology, our study design may have identified an ER+ DCIS-specific profile. The genes identified may not only be progression-specific, but may be dysregulated regardless of ER+ DCIS subgroup. This is demonstrated by a single outlier from the independent validation containing lobular carcinoma *in situ*, and supports data indicating lobular and ductal tumors may be genetically distinct.⁴

A second gap in knowledge we address is the potential role that cell adhesion and ECM play during breast cancer progression.⁵ ECM-neoplastic epithelial cell interactions are known to be important in later stages of breast tumorigenesis, specifically during invasion and metastasis.^{10,31,32} Recent studies have identified stromal gene expression signatures that are prognostic in invasive breast cancer.^{33–36} Other studies suggest that the stromal microenvironment plays an important role in the transition from DCIS to invasive disease.³⁷ Our data suggest that ECM-epithelial cell interactions become dysregulated even earlier, before visible evidence of invasion. In fact, we identify adhesion and ECM-related genes and pathways before evidence of malignancy (ie, not only in DCIS, but also in ADH). Also, we note a high degree of overlap between the stromal expression signatures that can predict prognosis in invasive breast cancer and the ECM-related genes identified within our study. Furthermore, when considering those genes dysregulated only in DCIS (ie, the ~400 DCIS-HN genes that do not overlap with ADH-HN), adhesion and ECM-related pathways are not overrepresented. Thus, cancer-associated dysregulation of adhesion and ECM pathways may already have occurred by the time ADH is apparent. Early involvement of ECM-related pathways may help explain why *FGFR2* (underexpressed in both ADH and DCIS) could be a breast cancer susceptibility gene.³⁸

Among the adhesion and ECM-related pathways we identified, focal adhesion is most overrepresented. Focal adhesions serve as mechanical links to the ECM, signaling centers by which cells communicate with the surrounding ECM, and have been implicated in breast cancer invasion.^{39–41} Focal adhesions may be conduits by which abnormal luminal epithelial cells communicate with the microenvironment and vice versa (an abnormal microenvironment communicating with epithelial cells).

One of these focal adhesion genes is caveolin 1 (*CAV1*). We found *CAV1* to be dysregulated in all of our comparisons, indicating that *CAV1* expression progressively decreases as these ER+ lesions become more histologically abnormal. These data are consistent with a recent study showing that *CAV1* null mice develop DCIS-like lesions when treated with estrogen.⁴² Furthermore,

the mammary stroma in these mice also promoted the growth of transplanted tumors.⁴³ Together, these results suggest that *CAV1* may be an important tumor suppressor in the mammary gland and that a loss of *CAV1* function may contribute to the development of ER+ breast tumors.

Another important component of ECM-epithelial cell interactions are MECs that surround terminal ductal lobular units and help to maintain the basement membrane.^{44,45} The lesions we profiled undoubtedly include MECs. A recent study identified phenotypic differences in protein expression between DCIS-associated MECs and those from normal breast.⁴⁴ We found decreased expression of many of these genes, not only in DCIS, but also in ADH (eg, *MME*, *MYH11*, *CNN1*). Taken together, there is evidence supporting the hypothesis that the appearance of higher risk lesions (like ADH) may not be due only to epithelial cell abnormalities. Instead, it may require specific dysregulation of and interaction with the surrounding microenvironment. ECM-MEC interactions, in particular, may dysregulate important luminal epithelial cell genes, thereby influencing precancerous lesions to progress toward invasive disease.

A final gap in knowledge addressed here is the role of SH in breast cancer progression. Simple hyperplasias have been greatly understudied at the gene expression level. While our SH sample size is small, we find SH consistently resembles HN epithelium. Therefore, SH may not yet have acquired abnormalities, or been exposed to stimuli from the microenvironment, that place them on the precursor pathway. It will be important to expand the study of these lesions and compare them to other proliferative lesions without atypia, some of which do appear to contain abnormalities.⁴⁶

Our study has limitations, most notably its small sample size. We have taken pains to minimize the potential biases that could arise. We used microdissection of fresh frozen primary human tissues, patient-matching, a statistical analysis accounting for subtle gene expression changes in small sample sizes, an independent set of samples for validation, and we have made the results publicly available. Together, these make our data unique and valuable.

In summary, we used patient-matched samples to define specific expression profiles of potential breast cancer precursors, assess genetic relationships between these precursors, and illuminate molecular events that occur during the early stages of tumorigenesis. Recently, mathematical modeling suggests a “parallel” pathway model, by which DCIS and IDC may progress through different grades in parallel and may arise from a common progenitor, corresponds better with clinical observations than the traditional linear model of breast cancer progression.⁴⁷ This parallel pathway may imply that, for tumors containing both DCIS and IDC (like most samples in our study), essential genetic changes occur before histological evidence of DCIS.⁴ Our findings not only support this hypothesis, but suggest that precursors may diverge and progress along this parallel pathway even earlier, in the transition from normal (or SH) to ADH. Ultimately, these observations could foster the development of novel

breast cancer treatments, prevention strategies, and risk assessment tools. Targeting cell adhesion and ECM pathways is a logical candidate for further exploration.

Acknowledgments

We are grateful for the contributions to this project from our late collaborator, Dr. Erwin F. Hirsch, Department of Surgery, Boston University Medical Center. We also thank Dr. Hongying Huang, Department of Pathology, Boston University Medical Center, for assistance with the histology images.

References

- Wellings SR, Jensen HM, Marcum RG: An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. *J Natl Cancer Inst* 1975, 55:231–273
- Allred DC, Wu Y, Mao S, Nagtegaal ID, Lee S, Perou CM, Mohsin SK, O'Connell P, Tsimelzon A, Medina D: Ductal carcinoma in situ and the emergence of diversity during breast cancer evolution. *Clin Cancer Res* 2008, 14:370–378
- Sotiriou C, Pusztai L: Gene-expression signatures in breast cancer. *N Engl J Med* 2009, 360:790–800
- Kuerer HM, Albarracin CT, Yang WT, Cardiff RD, Brewster AM, Symmans WF, Hylton NM, Middleton LP, Krishnamurthy S, Perkins GH, Babiera G, Edgerton ME, Czerniecki BJ, Arun BK, Hortobagyi GN: Ductal carcinoma in situ: state of the science and roadmap to advance the field. *J Clin Oncol* 2009, 27:279–288
- Thompson A, Brennan K, Cox A, Gee J, Harcourt D, Harris A, Harvie M, Holen I, Howell A, Nicholson R, Steel M, Streuli C: Evaluation of the current knowledge limitations in breast cancer research: a gap analysis. *Breast Cancer Res* 2008, 10:R26
- Ma XJ, Salunga R, Tuggle JT, Gaudet J, Enright E, McQuary P, Payette T, Pistone M, Stecker K, Zhang BM, Zhou YX, Varnholt H, Smith B, Gadd M, Chatfield E, Kessler J, Baer TM, Erlander MG, Sgroi DC: Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci USA* 2003, 100:5974–5979
- Vincent-Salomon A, Lucchesi C, Gruel N, Raynal V, Pierron G, Goudefroye R, Reyat F, Radvanyi F, Salmon R, Thierry JP, Sastre-Garau X, Sigal-Zafrani B, Fourquet A, Delattre O: Integrated genomic and transcriptomic analysis of ductal carcinoma in situ of the breast. *Clin Cancer Res* 2008, 14:1956–1965
- Tamimi RM, Baer HJ, Marotti J, Galan M, Galaburda L, Fu Y, Deitz AC, Connolly JL, Schnitt SJ, Colditz GA, Collins LC: Comparison of molecular phenotypes of ductal carcinoma in situ and invasive breast cancer. *Breast Cancer Res* 2008, 10:R67
- Iakovlev VV, Arneson NC, Wong V, Wang C, Leung S, Iakovleva G, Warren K, Pintilie M, Done SJ: Genomic differences between pure ductal carcinoma in situ of the breast and that associated with invasive disease: a calibrated aCGH study. *Clin Cancer Res* 2008, 14:4446–4454
- Hu M, Yao J, Carroll DK, Weremowicz S, Chen H, Carrasco D, Richardson A, Violette S, Nikolskaya T, Nikolsky Y, Bauerlein EL, Hahn WC, Gelman RS, Allred C, Bissell MJ, Schnitt S, Polyak K: Regulation of in situ to invasive breast carcinoma transition. *Cancer Cell* 2008, 13:394–406
- Moulis S, Sgroi DC: Re-evaluating early breast neoplasia. *Breast Cancer Res* 2008, 10:302
- Hartmann LC, Sellers TA, Frost MH, Lingle WL, Degnim AC, Ghosh K, Vierkant RA, Maloney SD, Pankratz VS, Hillman DW, Suman VJ, Johnson J, Blake C, Tlsty T, Vachon CM, Melton LJ, 3rd, Visscher DW: Benign breast disease and the risk of breast cancer. *N Engl J Med* 2005, 353:229–237
- Degnim AC, Visscher DW, Berman HK, Frost MH, Sellers TA, Vierkant RA, Maloney SD, Pankratz VS, de Groen PC, Lingle WL, Ghosh K, Penhater L, Tlsty T, Melton LJ, 3rd, Reynolds CA, Hartmann LC: Stratification of breast cancer risk in women with atypia: a Mayo cohort study. *J Clin Oncol* 2007, 25:2671–2677
- Larson PS, de las Morenas A, Cerda SR, Bennett SR, Cupples LA, Rosenberg CL: Quantitative analysis of allele imbalance supports atypical ductal hyperplasia lesions as direct breast cancer precursors. *J Pathol* 2006, 209:307–316
- Chuaqui RF, Zhuang Z, Emmert-Buck MR, Liotta LA, Merino MJ: Analysis of loss of heterozygosity on chromosome 11q13 in atypical ductal hyperplasia and in situ carcinoma of the breast. *Am J Pathol* 1997, 150:297–303
- O'Connell P, Pekkel V, Fuqua SA, Osborne CK, Clark GM, Allred DC: Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. *J Natl Cancer Inst* 1998, 90:697–703
- Poola I, DeWitty RL, Marshalleck JJ, Bhatnagar R, Abraham J, Leffall LD: Identification of MMP-1 as a putative breast cancer predictive marker by global gene expression analysis. *Nat Med* 2005, 11:481–483
- Reis-Filho JS, Lakhani SR: The diagnosis and management of pre-invasive breast disease: genetic alterations in pre-invasive lesions. *Breast Cancer Res* 2003, 5:313–319
- Silverstein MJ, Poller DN, Waisman JR, Colburn WJ, Barth A, Gierson ED, Lewinsky B, Gamagami P, Slamon DJ: Prognostic classification of breast ductal carcinoma-in-situ. *Lancet* 1995, 345:1154–1157
- Dalton LW, Pinder SE, Elston CE, Ellis IO, Page DL, Dupont WD, Blamey RW: Histologic grading of breast cancer: linkage of patient outcome with level of pathologist agreement. *Mod Pathol* 2000, 13:730–735
- King C, Guo N, Frampton GM, Gerry NP, Lenburg ME, Rosenberg CL: Reliability and reproducibility of gene expression measurements using amplified RNA from laser-microdissected primary breast tissue with oligonucleotide arrays. *J Mol Diagn* 2005, 7:57–64
- Tripathi A, King C, de la Morenas A, Perry VK, Burke B, Antoine GA, Hirsch EF, Kavanah M, Mendez J, Stone M, Gerry NP, Lenburg ME, Rosenberg CL: Gene expression abnormalities in histologically normal breast epithelium of breast cancer patients. *Int J Cancer* 2008, 122:1557–1566
- Sebastiani P, Xie H, Ramoni MF: Bayesian analysis of comparative microarray experiments by model averaging. *Bayesian Anal* 2006, 1:707–732
- Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA: DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003, 4:P3
- Huang da W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009, 4:44–57
- Dydensborg AB, Herring E, Auclair J, Tremblay E, Beaulieu JF: Normalizing genes for quantitative RT-PCR in differentiating human intestinal epithelial cells and adenocarcinomas of the colon. *Am J Physiol Gastrointest Liver Physiol* 2006, 290:G1067–G1074
- Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001, 25:402–408
- Gudjonsson T, Adriance MC, Sternlicht MD, Petersen OW, Bissell MJ: Myoepithelial cells: their origin and function in breast morphogenesis and neoplasia. *J Mammary Gland Biol Neoplasia* 2005, 10:261–272
- Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S, Schott A, Hayes D, Birnbaum D, Wicha MS, Dontu G: ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* 2007, 1:555–567
- Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, Campbell LL, Polyak K, Briskin C, Yang J, Weinberg RA: The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008, 133:704–715
- Kaspar M, Zardi L, Neri D: Fibronectin as target for tumor therapy. *Int J Cancer* 2006, 118:1331–1339
- Allinen M, Beroukhim R, Cai L, Brennan C, Lahti-Domenici J, Huang H, Porter D, Hu M, Chin L, Richardson A, Schnitt S, Sellers WR, Polyak K: Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* 2004, 6:17–32
- Helleman J, Jansen MP, Ruigrok-Ritstier K, van Staveren IL, Look MP, Meijer-van Gelder ME, Sieuwerts AM, Klijn JG, Sleijfer S, Foekens JA, Berns EM: Association of an extracellular matrix gene cluster with breast cancer prognosis and endocrine therapy response. *Clin Cancer Res* 2008, 14:5555–5564
- Bergamaschi A, Tagliabue E, Sorlie T, Naume B, Triulzi T, Orlandi R, Russnes HG, Nesland JM, Tammi R, Auvinen P, Kosma VM, Menard S, Borresen-Dale AL: Extracellular matrix signature identifies breast

- cancer subgroups with different clinical outcome. *J Pathol* 2008, 214:357–367
35. Lukes L, Crawford NP, Walker R, Hunter KW: The origins of breast cancer prognostic gene expression profiles. *Cancer Res* 2009, 69:310–318
36. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, Chen H, Omeroglu G, Meterissian S, Omeroglu A, Hallett M, Park M: Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med* 2008, 14:518–527
37. Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC: Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res* 2009, 11:R7
38. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struwing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odefrey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, Fletcher O, Johnson N, Seal S, Stratton MR, Rahman N, Chenevix-Trench G, Bojesen SE, Nordestgaard BG, Axelsson CK, Garcia-Closas M, Brinton L, Chanock S, Lissowska J, Peplonska B, Nevanlinna H, Fagerholm R, Eerola H, Kang D, Yoo KY, Noh DY, Ahn SH, Hunter DJ, Hankinson SE, Cox DG, Hall P, Wedren S, Liu J, Low YL, Bogdanova N, Schurmann P, Dork T, Tollenaar RA, Jacobi CE, Devilee P, Klijn JG, Sigurdson AJ, Doody MM, Alexander BH, Zhang J, Cox A, Brock IW, MacPherson G, Reed MW, Couch FJ, Goode EL, Olson JE, Meijers-Heijboer H, van den Ouweland A, Uitterlinden A, Rivadeneira F, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Hopper JL, McCredie M, Southey M, Giles GG, Schroen C, Justenhoven C, Brauch H, Hamann U, Ko YD, Spurdle AB, Beesley J, Chen X, Mannermaa A, Kosma VM, Kataja V, Hartikainen J, Day NE, Cox DR, Ponder BA: Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007, 447:1087–1093
39. Chen CS, Alonso JL, Ostuni E, Whitesides GM, Ingber DE: Cell shape provides global control of focal adhesion assembly. *Biochem Biophys Res Commun* 2003, 307:355–361
40. Behmoaram E, Bijian K, Jie S, Xu Y, Darnel A, Bismar TA, Alaoui-Jamali MA: Focal adhesion kinase-related proline-rich tyrosine kinase 2 and focal adhesion kinase are co-overexpressed in early-stage and invasive ErbB-2-positive breast cancer and cooperate for breast cancer cell tumorigenesis and invasiveness. *Am J Pathol* 2008, 173:1540–1550
41. Provenzano PP, Inman DR, Eliceiri KW, Beggs HE, Keely PJ: Mammary epithelial-specific disruption of focal adhesion kinase retards tumor formation and metastasis in a transgenic mouse model of human breast cancer. *Am J Pathol* 2008, 173:1551–1565
42. Mercier I, Casimiro MC, Zhou J, Wang C, Plymire C, Bryant KG, Daumer KM, Sotgia F, Bonuccelli G, Witkiewicz AK, Lin J, Tran TH, Milliman J, Frank PG, Jasmin JF, Rui H, Pestell RG, Lisanti MP: Genetic ablation of caveolin-1 drives estrogen-hypersensitivity and the development of DCIS-like mammary lesions. *Am J Pathol* 2009, 174:1172–1190
43. Sotgia F, Del Galdo F, Casimiro MC, Bonuccelli G, Mercier I, Whitaker-Menezes D, Daumer KM, Zhou J, Wang C, Katiyar S, Xu H, Bosco E, Quong AA, Aronow B, Witkiewicz AK, Minetti C, Frank PG, Jimenez SA, Knudsen ES, Pestell RG, Lisanti MP: Caveolin-1/- null mammary stromal fibroblasts share characteristics with human breast cancer-associated fibroblasts. *Am J Pathol* 2009, 174:746–761
44. Hilson JB, Schnitt SJ, Collins LC: Phenotypic alterations in ductal carcinoma in situ-associated myoepithelial cells: biologic and diagnostic implications. *Am J Surg Pathol* 2009, 33:227–332
45. Adriance MC, Inman JL, Petersen OW, Bissell MJ: Myoepithelial cells: good fences make good neighbors. *Breast Cancer Res* 2005, 7:190–197
46. Lee S, Medina D, Tsimelzon A, Mohsin SK, Mao S, Wu Y, Allred DC: Alterations of gene expression in the development of early hyperplastic precursors of breast cancer. *Am J Pathol* 2007, 171:252–262
47. Sontag L, Axelrod DE: Evaluation of pathways for progression of heterogeneous breast tumors. *J Theor Biol* 2005, 232:179–189