Full-term Pregnancy Induces a Specific Genomic Signature in the Human Breast

Jose Russo, Gabriela A. Balogh, and Irma H. Russo and the Fox Chase Cancer Center Hospital Network Participants

Breast Cancer Research Laboratory, Fox Chase Cancer Center, Philadelphia, Pennsylvania

Abstract

Breast cancer risk has traditionally been linked to nulliparity or late first full-term pregnancy, whereas young age at first childbirth, multiparity, and breast-feeding are associated with a reduced risk. Early pregnancy confers protection by inducing breast differentiation, which imprints a specific and permanent genomic signature in experimental rodent models. For testing whether the same phenomenon was detectable in the atrophic breast of postmenopausal parous women, we designed a case-control study for the analysis of the gene expression profile of RNA extracted from epithelial cells microdissected from normal breast tissues obtained from 18 parous and 7 nulliparous women free of breast pathology (controls), and 41 parous and 8 nulliparous women with history of breast cancer (cases). RNA was hybridized to cDNA glass microarrays containing 40,000 genes; arrays were scanned and the images were analyzed using ImaGene software version 4.2. Normalization and statistical analysis were carried out using Linear Models for Microarrays and GeneSight software for hierarchical clustering. The parous control group contained 2,541 gene sequences representing 18 biological processes that were differentially expressed in comparison with the other three groups. Hierarchical clustering of these genes revealed that the combined parity/absence of breast cancer data generated a distinct genomic profile that differed from those of the breast cancer groups, irrespective of parity history, and from the nulliparous cancer-free group, which has been traditionally identified as a high-risk group. The signature that identifies those women in whom parity has been protective will serve as a molecular biomarker of differentiation for evaluating the potential use of preventive agents. (Cancer Epidemiol Biomarkers Prev 2008;17(1):51–66)

Introduction

More than 300 years elapsed since a striking excess in breast cancer mortality was reported in nuns, in whom the increased risk was attributed to their childlessness (1), until MacMahon et al. (2), in a landmark case-control study, found an almost linear relationship between a woman's risk and the age at which she bore her first child. This work, which included areas of high, intermediate, and low breast cancer risk in seven parts of the world, confirmed that pregnancy had a protective effect that was evident from the early teen years and persisted until the middle twenties (2). Other studies have reported that additional pregnancies and breast-feeding confer greater protection to young women, including a statistically significantly reduced risk of breast cancer in women with deleterious *BRCA1* mutations who breast-fed for a cumulative total of >1 year (3, 4). Our studies, designed for unraveling what specific phenomena occurred in the breast during pregnancy for conferring a lifetime protection from developing cancer, led us to the discovery that endogenous endocrinological or environmental influences affecting breast development before the first full-term pregnancy were important modulators of the susceptibility of the breast to undergo neoplastic transformation (5, 6). The fact that exposure of the breast of young nulliparous females to environmental physical agents (7) or chemical toxicants (8, 9) results in a greater rate of cell transformation indicates that the immature breast possesses a greater number of susceptible cells that can become the site of origin of cancer, similarly to what has been reported in experimental animal models (5, 6, 10, 11). In these models, the initiation of cancer is prevented by the differentiation of the mammary gland induced by pregnancy (11, 12). The molecular changes involved in this phenomenon are just starting to be unraveled (13-15). In women, the protection conferred by pregnancy, however, is age specific because a delay in childbearing after age 24 progressively increases the risk of cancer development, which becomes greater than that of nulliparous women when the first full-term pregnancy occurs after 35 years of age (2, 16). The higher breast cancer risk that has been associated with early menarche (17) further emphasizes the importance of the length of the susceptibility ''window'' that encompasses the period of breast

Received 7/26/07; revised 10/3/07; accepted 11/1/07.

Grant support: National Cancer Institute grant RO1-CA093599.

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.

Note: Fox Chase Cancer Center Hospital Network Participants: Emily Penman and Nicholas J. Petrelli, Helen F. Graham Cancer Center, Christiana Care Health System, Newark, Delaware; Angela Lanfranchi, Somerset Medical Center, Somerville, New Jersey; Kathryn Evers, Diagnostic Imaging, and Monica Morrow, Department of Surgical Oncology, American Oncologic Hospital, Fox Chase Cancer Center, Philadelphia, Pennsylvania.

Requests for reprints: Jose Russo, Breast Cancer Research Laboratory, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, PA 19111. Phone: 215-728-4782; Fax: 215-728-2180. E-mail: J_Russo@fccc.edu

Copyright © 2008 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-07-0678

development occurring between menarche and the first pregnancy, when the organ is more susceptible to either undergo complete differentiation under physiologic hormonal stimuli, and hence to be protected from breast cancer, or to suffer genetic or epigenetic damage that might contribute to increasing the lifetime risk of developing breast cancer (9, 18). The damage caused by a single or a combination of putative cancer-causing agents might, in turn, be amplified by the genetic make up of the patient, such as the inheritance of the *BRCA1* or *BRCA2* susceptibility genes, which influences the pattern of breast development and differentiation and is responsible for at least 5% of all the breast cancer cases (19-21). This postulate is supported by our observations that the architectural pattern of lobular development in parous women with cancer differs from that of parous women without cancer, being similar to that of nulliparous women with or without cancer. Thus, the higher breast cancer risk in parous women might have resulted from either a failure of the breast to fully differentiate under the influence of the hormones of pregnancy (22, 23) and/or stimulation of the growth of foci of transformed cells initiated by early damage or genetic predisposition (9, 18, 20).

Numerous studies have been done for understanding how the dramatic modifications that occur during pregnancy in the pattern of lobular development and differentiation (22, 23), cell proliferation, and steroid hormone receptor content of the breast (24) influence cancer risk. Studies at molecular level using different platforms for global genome analysis have confirmed the universality of this phenomenon in various strains of rats and mice (13-15, 25). For testing the permanence of these changes, we have analyzed the pattern of gene expression occurring during and after pregnancy in rodents. Hierarchical cluster analysis of the genomic profile of rat mammary glands in the 15th and 21st days of pregnancy and at 21 and 42 days postpartum revealed four different patterns of expression in relation to the time of pregnancy (25). During pregnancy, genes related to the secretory properties of the mammary epithelium (Cluster A) become up-regulated, decreasing to control values after 21 and 42 days postpartum. Cluster B includes genes related to the apoptotic pathways, the fatty acid binding protein and cathecol-*O*-methyltransferase, among others, which become up-regulated from the end of pregnancy until the 21st day postpartum and decreasing thereafter. Cluster C represents differentiation-associated genes whose level of expression continuously and progressively increases with time of pregnancy, reaching their highest levels between 21 and 42 days postpartum, and cluster D comprises genes that are up-regulated around the 15th day of pregnancy and become progressively down-regulated from the end of pregnancy until the 42nd day postpartum (25). These observations confirm at genomic level our previous morphologic and physiologic findings indicating that temporal and sequential changes have to occur in the development of the mammary gland for accomplishing a protective degree of differentiation (11, 12, 25-28). The importance of identifying a specific signature by 42 days postpartum is highlighted by the observations that administration of the polycyclic hydrocarbon 7,12-dimethylbenz(*a*)anthracene to parous rats results in a markedly reduced tumorigenic response, supporting the concept that the differentiation induced by pregnancy shifts the susceptible ''intermediate cells'' that originate mammary cancer in the terminal end buds of the virginal gland (5, 10) to transformation-resistant cells (11, 12).

Studies in experimental animal models have been useful for uncovering the sequential genomic changes occurring in the mammary gland in response to the multiple hormonal stimuli of pregnancy that lead to the imprinting of a permanent genomic signature. Work reported here was designed with the purpose of testing whether a similar phenomenon occurs in the atrophic breast of postmenopausal parous women, specifically in the epithelium of lobule type 1 (Lob 1), the site of origin of breast carcinomas (5, 6). Our results support our hypothesis that parous women that had not developed breast cancer after menopause exhibit a genomic ''signature'' that differs from that present in the breast of parous postmenopausal women with cancer or in nulliparous women who traditionally represent a high breast cancer risk group (1-6).

Materials and Methods

Patients and Methods for Sample Collection. For this three-center hospital-based study, patients were enrolled from the American Oncologic Hospital of the Fox Chase Cancer Center in Philadelphia, Pennsylvania; Christiana Care Health System, Newark, Delaware; and Somerset Medical Center, Sommerville, New Jersey. The study protocol had been approved by the Institutional Review Board of each participating institution, and written informed consent was obtained from every participant. Patients eligible for the study were postmenopausal women ≥50 years of age and whose menses had naturally ceased 1 year before enrollment. Excluded from this study were women whose ovaries had been surgically removed; who had a history of cancer other than nonmelanoma skin cancer; who were taking medications that could interfere with the study protocol such as estrogens (including Tamoxifen and Raloxifene), progestins, androgens, prednisone, thyroid hormones, and insulin; and women with Alzheimer's disease or severe cognitive deficit and were unable to give informed consent.

Participant Identification. Potential participants were identified by a trained research nurse that carried out daily searches of surgical breast consultation visit summaries at the Breast Evaluation Clinic of the three participating hospitals. Those women that fulfilled the eligibility criteria listed above and who were recommended for a breast biopsy by their treating breast surgeon were selected for the study. Information included in visit summaries, such as age, menopausal status, history of cancer, and current medications, was used to determine if a woman was potentially eligible for this study. A letter was sent to each potential participant describing the study and informing them of their eligibility, which was confirmed in a telephone interview placed within 2 weeks of initial clinical evaluations when biopsies were recommended.

Data and Specimen Collection. Data were collected at preoperative clinic visits before biopsies and during breast biopsy procedures. At the preoperative visits,

informed consent was obtained, participants were asked to complete a study questionnaire, and height and weight were measured. Each one of the participating hospitals was provided specifically designed kits for breast tissue collection, which included tissue specimen containers partially filled with 70% ethanol, blood collecting tubes, copies of the eligibility criteria, patient data questionnaires, and labels with coded numbers for the biospecimens and questionnaires. All patients were accessed to a Fox Chase Cancer Center database using the originally assigned coded numbers. Patient names and medical record numbers were known only by the treating physician and authorized personnel at each participating hospital.

Breast tissue specimens were obtained by the operating surgeon following standard procedures for surgical breast biopsies at each site only after tissues were evaluated for presence of tumor, and if present, assessment of tumor size, margin identification, and adequacy of the tissue available for pathologic diagnosis. Normal-appearing tissues were taken from areas at a distance ≥2 cm from any grossly identifiable lesion and immediately fixed in 70% ethanol for 8 h, followed by dehydration, paraffin embedding, sectioning, and staining for histologic analysis and laser capture microdissection following previously described procedures (28). Histopathologic diagnosis of tumor type was made by pathologists at each site (Table 1). Only women diagnosed with invasive breast cancer (cases) or benign breast disease without hyperplasia or atypia (controls) were included in the study. Seventy-four postmenopausal women fulfilled the criteria of eligibility for this study; among them, 59 (80%) were parous and 15 (20%) nulliparous. Eighteen of the 59 parous women that had benign breast biopsies but were free of cancer served as controls and 41 women that had a diagnosis of breast cancer were selected as cases. Among the nulliparous women, 7 were free of cancer (control) and 8 had breast cancer (cases). Average ages at the time of diagnosis and at first birth in the parous cases and controls are shown in Table 2. The number of cases per group represents the distribution of cases at each one of the participating hospitals.

cDNA Human Microarray Analysis. RNA isolation and amplification from laser capture microdissection samples were done as previously described (28). Microarrays were prepared by the Fox Chase Cancer Center National Cancer Institute–supported Microarray Facility. Mirror glass slides were used for robotically spotting 40,000 cDNAs representing 28,000 distinct human transcripts, 10,000 identified by expressed sequence tags, and 2,000 controls and blank spots. Probe construction using direct labeling with random hexamer primer and purification using the QIA-quick PCR purification kit (Qiagen) were done as previously described (28). After the last centrifugation at 13,000 rpm for 1 min, the concentration of the eluted material was determined, then partially dried in a vacuum centrifuge and resuspended in 15 μL of hybridization buffer containing 20× SSC and 0.6 μL of 10% (w/v) SDS. Thereafter, the probes were denatured at 95°C, centrifuged for 3 min at 13,000 rpm, and the products were pipetted onto prehybridized arrays; the slides were coverslipped and placed in hybridization chambers (Gene Machine). Arrays were incubated in a 42°C water bath for 16 to 18 h and subsequently washed with 0.5× SSC, 0.01% (w/v) SDS, followed by 0.06× SSC, at room temperature for 10 min each. The slides were centrifuged for 8 min at 800 rpm (130 × *g*) at room temperature. The glass microarrays were hybridized, placing in the red channel (labeled with Cy5) the amplified RNA from the breast samples and in the green channel (labeled with Cy3) the human universal reference amplified RNA (Stratagene Technologies, Inc.). Each hybridization compared Cy5-labeled cDNA reverse transcribed from amplified RNA isolated from each patient with the Cy3-labeled cDNA reverse transcribed from a universal human reference amplified RNA sample. Equal amounts of fluorescent probes were used to hybridize the cDNA microarrays in triplicate, and after quality verification in the Nanodrop, replicates from the same sample were combined and redistributed into three separate tubes to have identical replicates. Arrays were read in an Affymetrix 428 fluorescent scanner (MWG) at 10-μm resolution, with variable voltage of the photomultiplier tube for obtaining the maximal signal intensities with <1% (w/v) probe saturation. The resulting images were analyzed using ImaGene software version 4.2 (Biodiscovery).

Data Analysis. Normalization and statistical analysis of the expression data were carried out using Linear Models for Microarray Data (29-31). For detecting the differential expression of genes that might not necessarily be highly expressed, background correction using the ''normexp'' method in Linear Models for Microarray Data was done for adjusting the local median background estimates, a correction strategy that avoids problems with background estimates that are greater than foreground values and ensures that there were no missing or negative corrected intensities. An offset of 100 was used for both channels to further damp down the variability of log ratios for low-intensity spots. The resulting log ratios were normalized by using the print-tip group Lowess method with span 0.4, as recommended by Smyth (31).

Moderated *t* statistic was used as the basic statistic for significance analysis; it was computed for each probe and for each contrast (31). False discovery rate was controlled using the BH adjustment of Benjamini and Hochberg (32, 33). All genes with *P* value below a threshold of 0.05 were selected as differentially expressed, maintaining the proportion of false discoveries in the selected group below the threshold value, in this case 5% (34). Hierarchical clustering was done using GeneSight software (version 2.4; BioDiscovery, Inc.).

Gene Validation by Reverse Transcription-PCR Amplification. Genes that were found to be up-regulated in the parous control breast were validated by real-time reverse transcription-PCR (RT-PCR) using nucleotide sequences that were found using the gene accession number obtained from the cDNA glass microarrays and searching the National Cancer Institute Blast website.[1] TaqMan primer and probe sets sequences are listed in Table 4. The sense and antisense primer sequences were

[1] http://www.ncbi.nlm.nih.gov/BLAST/
[2] http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi

Table 1. Profile of the four groups of patients and diagnosis of breast lesions from which normal Lob 1 epithelium was obtained by laser capture microdissection

	Case ID	Age at diagnosis (y)	Age at first birth (y)	Breast biopsy diagnosis	Parity status	RNA (ng)*	aaRNA (ng)	Ratio 260/280
1	81719	50	33	Fibrocystic changes	Parous control	28.90	997.80	1.99
2	84453	59	25	Ductal hyperplasia	Parous control	58.10	896.10	2.01
3	110857	55	22	Fibroadenoma	Parous control	54.60	938.20	2.02
4	119747	61	25	Fibrocystic changes	Parous control	25.90	961.10	2.04
5	131682	55	17	Ductal hyperplasia, mild	Parous control	25.80	725.10	1.99
6	134134	61	34	Fibroadenoma, adenosis	Parous control	80.90	6,072.00	2.00
7	135125	52	31	Adenosis, ductal ectasia	Parous control	224.70	1,483.30	2.03
8	135447	77	23	Apocrine metaplasia	Parous control	28.90	865.00	2.05
9	135990	64	27	Adenosis	Parous control	47.90	901.40	2.03
10	136383	71	24	Adenosis	Parous control	46.50	3,823.00	2.02
11	136880	59	20	Papilloma	Parous control	39.70	968.30	2.03
12	137340	63	18	Ductal hyperplasia, mild	Parous control	35.80	235.40	2.01
13	139641	61	21	Stromal fibrosis	Parous control	49.70	1,468.20	1.99
14	141007	77	21	Adenosis	Parous control	40.60	1,283.50	2.02
15	141300	78	24	Adenosis	Parous control	161.90	2,358.10	2.02
17	143793	72	27	Adenosis	Parous control	20.30	439.30	2.01
18	148115	60	20	Benign breast disease	Parous control	46.30	412.40	2.00
19	131453	71	17	Invasive Ductal carcinoma	Parous case	52.50	2,115.60	2.03
20	132370	61	26	Invasive Ductal Carcinoma	Parous case	20.10	65.00	2.05
21	132452	55	26	Invasive Ductal and lobular carcinoma	Parous case	51.10	1,225.80	2.07
22	132454	60	25	Invasive ductal carcinoma	Parous case	46.30	544.40	2.02
23	132456	72	19	Invasive ductal carcinoma	Parous case	100.50	53.20	2.03
24	133360	57	19	Invasive ductal carcinoma	Parous case	54.90	1,534.00	2.01
25	133931	74	26	Invasive ductal and lobular carcinoma	Parous case	22.20	1,682.90	2.00
26	134133	75	26	Invasive ductal carcinoma	Parous case	41.70	6,421.80	2.00
27	154855	75	25	Invasive ductal carcinoma	Parous case	32.50	5,432.00	2.00
28	135984	76	20	Mucinous adenocarcinoma	Parous case	21.90	1,443.80	2.00
29	137805	78	23	Mucinous adenocarcinoma	Parous case	401.40	1,235.00	2.02
30	138206	59	28	Invasive Ductal carcinoma and DCIS	Parous case	45.40	1,574.20	2.00
31	138993	76	26	Invasive Ductal carcinoma	Parous case	83.40	7,473.40	2.04
32	139128	84	31	Invasive ductal carcinoma	Parous case	79.30	1,310.20	2.00
33	140569	67	24	Invasive ductal carcinoma	Parous case	41.10	1,093.00	2.00
34	141008	55	29	Invasive ductal carcinoma and DCIS	Parous case	29.70	1,405.60	2.00
35	141299	75	27	Invasive ductal carcinoma and DCIS	Parous case	42.90	762.30	2.05
38	145563	65	23	Invasive ductal carcinoma	Parous case	17.90	764.00	2.00
39	145564	74	25	Invasive ductal carcinoma	Parous case	28.50	828.20	2.00
40	145565	62	28	Invasive ductal carcinoma and DCIS	Parous case	26.10	682.60	2.02
41	146980	65	26	Invasive ductal carcinoma	Parous case	17.10	411.60	2.02
42	147715	81	25	Invasive ductal carcinoma and DCIS	Parous case	106.30	416.70	2.02
43	149911	56	32	Invasive ductal carcinoma and DCIS	Parous case	107.30	1,425.76	2.02
44	153163	82	30	Invasive lobular carcinoma	Parous case	377.90	1,326.24	2.00
45	153556	65	30	Invasive ductal carcinoma	Parous case	309.10	1,427.00	2.00
46	154250	76	20	Invasive ductal carcinoma and DCIS	Parous case	310.10	1,428.24	2.11
47	155065	79	28	Invasive ductal carcinoma	Parous case	1,003.40	1,129.98	2.00
48	155844	75	26	Invasive ductal carcinoma	Parous case	1,537.70	1,430.24	2.06
49	155845	82	21	Invasive ductal carcinoma	Parous case	310.10	1,531.00	2.00
50	156062	58	26	Invasive ductal carcinoma	Parous case	311.10	1,432.24	2.08
51	156105	73	26	Invasive lobular carcinoma and LCIS	Parous case	305.80	1,233.24	2.00
52	157584	70	23	Invasive lobular carcinoma	Parous case	325.80	1,434.00	2.20
53	157678	92	19	Invasive lobular Carcinoma	Parous case	1,784.00	1,635.24	2.00
54	158532	70	25	Invasive ductal carcinoma	Parous case	1,655.80	1,146.94	2.00
55	158972	60	31	Invasive ductal carcinoma and DCIS	Parous case	966.90	1,437.50	2.01
56	158973	61	19	Invasive ductal carcinoma	Parous case	1,011.90	1,444.24	2.00
57	160038	60	16	Invasive ductal carcinoma	Parous case	429.80	1,439.24	2.00
58	160039	66	30	Invasive ductal carcinoma and DCIS	Parous case	1,783.50	1,440.00	2.01
59	160827	63	28	Invasive ductal carcinoma and DCIS	Parous case	355.70	1,441.24	2.00
60	15737	65	N/A	Adenosis	Nulliparous control	579.00	1,342.67	2.00
61	45853	62	N/A	Fibroadenoma	Nulliparous control	131.90	1,443.24	2.04
62	131161	58	N/A	Papilloma	Nulliparous control	51.10	2,005.80	2.00
63	132404	51	N/A	Fibroadenoma, papilloma	Nulliparous control	81.20	2,006.80	2.00
64	141009	53	N/A	Stromal fibrosis	Nulliparous control	108.80	977.40	2.07
65	143964	50	N/A	Apocrine metaplasia, stromal fibrosis	Nulliparous control	56.60	618.10	2.00
66	149204	58	N/A	Adenosis	Nulliparous control	31.10	401.60	2.00
67	132372	53	N/A	Invasive ductal carcinoma	Nulliparous case	20.90	557.50	2.05

(*Continued on the following page*)

Table 1. Profile of the four groups of patients and diagnosis of breast lesions from which normal Lob 1 epithelium was obtained by laser capture microdissection (Cont'd)

	Case ID	Age at diagnosis (y)	Age at first birth (y)	Breast biopsy diagnosis	Parity status	RNA (ng)*	aaRNA (ng)	Ratio 260/280
68	132382	68	N/A	Invasive ductal and lobular carcinoma	Nulliparous case	27.60	386.60	2.00
69	132402	77	N/A	Invasive ductal carcinoma	Nulliparous case	776.70	387.60	2.00
70	136596	74	N/A	Invasive ductal carcinoma	Nulliparous case	51.70	891.60	1.99
71	142667	87	N/A	Invasive ductal carcinoma	Nulliparous case	37.10	1,217.50	2.00
72	144166	57	N/A	Invasive ductal carcinoma	Nulliparous case	646.70	1,218.50	2.00
73	155958	57	N/A	Invasive lobular carcinoma	Nulliparous case	150.50	1,219.50	2.08
74	156622	55	N/A	Invasive ductal carcinoma	Nulliparous case	433.60	1,220.50	2.00

NOTE: Groups of patients: parous controls, women with benign breast biopsies; parous cases, women with breast cancer; nulliparous controls, childless women with benign breast biopsies; and nulliparous cases, childless women with breast cancer.
Abbreviations: aaRNA, amplified RNA; DCIS, ductal carcinomas *in situ*; LCIS, lobular carcinoma *in situ*.
*Total amount of RNA in nanograms obtained by laser capture microdissection from each sample.

designed using Primer3 software[2] and synthesized by the DNA Sequencing Facility at the Fox Chase Cancer Center. A β-actin primer was included as a control for gene expression. Primers were labeled with SyBro Green dye (Applied Biosystems); for avoiding competition in the multiplex PCR reaction, tube primer concentrations were limited and standardized. All RT-PCR reactions were done on the ABI Prism 7000 Sequence Detection System using the fluorescent SyBro Green methods (SYBRO Green RT-PCR Master Mix Reagents, all from Applied Biosystems). For each RT-PCR reaction, 100 ng of amplified RNA in a total volume of 50 μL were used. Primer and probe concentrations for target genes were optimized according to the manufacturer's recommended procedure. The following thermal cycling conditions were used: 30 min at 48°C, 10 min at 95°C, and 40 cycles of 15 s; denaturation at 95°C for 60 s; and annealing at 60°C. Each gene was analyzed in triplicate, normalized against β-actin, and expressed in relation to a calibrator sample. Results were expressed as relative gene expression using the ΔC_t method, as previously described (28).

Results

Identification of Differentially Expressed Genes in Breast Epithelium. For the analysis of the effect of parity on the genomic profile of epithelial cells from Lob 1, cDNA microarray expression profiling of the 74 breast tissue samples described in Table 1 was done. Genes whose expression changes differed by at least 1.2-fold and that were considered to be statistically significant between nulliparous and parous women with and without cancer using established algorithms were selected for further analysis (33). A total of 2,541 gene sequences were found to be differentially expressed (t test with false discovery rate $P < 0.05$) in the breast epithelium of the parous control group in comparison with nulliparous control and cases and parous cases. The parous control group had 126 genes up-regulated and 103 down-regulated (Table 3) with respect to the nulliparous control and case groups and to the parous group with breast cancer (cases).

Hierarchical Cluster Analysis. Unsupervised hierarchical clustering done using the expression profiles of 2,541 globally varying genes across the nulliparous and parous data sets representing the four groups revealed that samples clustered primarily based on parity status (Fig. 1). This suggested that the principal source of global variation in gene expression across these data sets was due to genetic differences between women due to reproductive history. This observation suggested that determining which parity-induced gene expression changes were conserved among these highly divergent groups could represent a powerful approach to defining a parity-related gene expression signature. Results of clustering set depicted in Fig. 2A and B indicate that the combined parity and absence of breast cancer data generate a distinct genomic profile that differs from the breast cancer groups, irrespective of parity history, and from the nulliparous cancer-free group, which has been traditionally identified as a high-risk group.

Gene Functional Category Analysis. We measured the relevance of Gene Ontology (GO) terms (35) belonging to the category of biological processes in the

Table 2.

Group	Age at diagnosis, mean ± SD (y)	Age at first birth (y)	RNA (ng)*	aaRNA (ng)	Ratio 260/280
Parous Control	63.23 ± 8.77 †	24.70 ± 4.88	59.79 ± 53.51 ‡,§	1,460.48 ± 1,454.13§	2.01 ± 0.01
Parous case	69.35 ± 9.21 †,‖	24.97 ± 4.06	365.35 ± 525.62 ‡	1,596.34 ± 1,502.99	2.02 ± 0.03
Nulliparous control	56.71 ± 5.60‖	NA	148.52 ± 192.97	1,256.51 ± 630.21	2.06 ± 0.14
Nulliparous case	66.00 ± 12.43	NA	268.10 ± 307.61§	887.41 ± 387.16	2.01 ± 0.03

*Total amount of RNA in nanograms obtained by laser capture microdissection from each sample.
† Parous controls vs parous cases, $t = 2.31$, $P < 0.02$.
‡ Parous controls vs parous cases, $t = 2.37$, $P < 0.02$.
§ Parous control vs nulliparous case, $t = 2.76$, $P = 0.01$.
‖ Parous case vs nulliparous control, $t = 3.94$, $P < 0.001$.

Table 3. Genes differentially expressed in the breast epithelium of parous control women

Gene name	Gene ID	Symbol	GO no.	Molecular function GO no.	$P_{adjusted}$	Fold increase/decrease
Apoptosis						
BCL2-associated X protein	AI565203	*BAX*	GO:0006915	GO:0005515	0.023	2.65
CASP2 and RIPK1	AA285065	*CRADD*	GO:0042981	GO:0005515	0.004	1.89
TNF receptor–associated factor 1	R71691	*TRAF1*	GO:0006915	GO:0006461	0.017	1.72
TIA1 cytotoxic granule-associated RNA binding protein	R82978	*TIA1*	GO:0006915	GO:0000166	0.017	1.56
TNFRSF1A-associated via death domain	AA916906	*TRADD*	GO:0006915	GO:0005515	0.027	1.42
Protein phosphatase 1F	AA806330	*PPM1F*	GO:0006915	GO:0016787	0.014	1.35
Mdm4	AI310969	*MDM4*	GO:0006915	GO:0004842	0.013	−1.25
Programmed cell death	AA416757	*PDCD5*	GO:0006915	GO:0005554	0.001	−2.15
Antiapoptosis						
Baculoviral inhibitor of apoptosis protein repeat–containing 6	H10434	*BIRC6*	GO:0006916	GO:0004840	0.013	−1.26
BCL2-associated athanogene 4	H22928	*BAG4*	GO:0006916	GO:0005057	0.026	−1.27
Cell adhesion						
Sema domain	AA436152	*SEMA5A*	GO:0007155	GO:0004872	0.050	1.81
Fibulin 5	H17615	*FBLN5*	GO:0007160	GO:0004888	0.010	1.79
Intercellular adhesion molecule 3	W95068	*ICAM3*	GO:0016337	GO:0005178	0.019	1.70
Formin binding protein 4	N49573	*FNBP4*	GO:0007155	GO:0005198	0.027	1.29
Sidekick homologue 1 (chicken)	N23940	*SDK1*	GO:0007155	GO:0005515	0.026	1.26
Epithelial V-like antigen 1	AA668897	*EVA1*	GO:0007155	GO:0005515	0.011	1.25
Neuropilin 1	AA098867	*NRP1*	GO:0007155	GO:0004872	0.017	1.25
Discs, large homologue 5 (*Drosophila*)	AA478949	*DLG5*	GO:0016337	GO:0005515	0.013	−1.80
Collagen, type XVI, α1	AA088202	*COL16A1*	GO:0007155	GO:0005198	0.010	−1.78
Down syndrome cell adhesion molecule	N53145	*DSCAM*	GO:0007155	GO:0005515	0.012	−2.10
Laminin, γ1 (formerly LAMB2)	AA599005	*LAMC1*	GO:0007155	GO:0005515	0.001	−2.91
Cell signaling-signal transduction						
Egf-like module containing	AI174266	*EMR2*	GO:0007165	GO:0004872	0.011	1.51
Low-density lipoprotein receptor–related protein 5	R83038	*LRP5*	GO:0016055	GO:0004872	0.014	1.40
G protein–coupled receptor kinase interactor 1	AI079118	*GIT1*	GO:0008277	GO:0005096	0.012	1.35
Insulin receptor substrate 1	AA456306	*IRS1*	GO:0007165	GO:0004871	0.022	1.29
Cornichon homologue 2 (*Drosophila*)	R42919	*CNIH2*	GO:0007242	GO:0005554	0.010	1.25
Ankyrin 2, neuronal	AI018106	*ANK2*	GO:0007165	GO:0005200	0.014	1.25
Galanin receptor 2	N75473	*GALR2*	GO:0007186	GO:0004966	0.030	1.20
Development and differentiation enhancing factor 2	AI054096	*DDEF2*	GO:0043087	GO:0005096	0.013	−1.30
BRCA2 and CDKN1A interacting protein	AI033172	*BCCIP*	GO:0000079	GO:0005554	0.007	−1.37
Rap guanine nucleotide exchange factor (GEF) 6	AA911005	*RAPGEF6*	GO:0007264	GO:0005085	0.002	−1.40
Endothelin receptor type A	AA909960	*EDNRA*	GO:0007186	GO:0001599	0.011	−1.45
Neuropeptide Y receptor Y1	R43817	*NPY1R*	GO:0007165	GO:0001584	0.019	−1.46
Neuropeptide S receptor 1	H91700	*NPSR1*	GO:0007165	GO:0004872	0.014	−1.46
GIPC PDZ domain containing family, member 1	AI094796	*GIPC1*	GO:0007186	GO:0005515	0.011	−1.55
RAB27A, member RAS oncogene family	AI309109	*RAB27A*	GO:0007264	GO:0000166	0.005	−1.85
Ankyrin repeat and death domain containing 1A	AI053438	*ANKDD1A*	GO:0007165	GO:0005515	0.009	−2.31
Small inducible cytokine subfamily E	H05323	*SCYE1*	GO:0007267	GO:0005125	0.002	−3.87
Coiled-coil domain containing 132	R49442	*CCDC132*	GO:0000160	GO:0000155	0.002	−3.89
Cell cycle and growth						
Homeodomain interacting protein kinase 2	N38891	*HIPK2*	GO:0000074	GO:0000166	0.007	1.87
Retinoblastoma binding protein 6	AA398302	*RBBP6*	GO:0000074	GO:0003676	0.012	1.58
DnaJ (Hsp40) homologue, subfamily A, member 2	AI273537	*DNAJA2*	GO:0000074	GO:0008270	0.002	1.56
Dynactin 1 (p150, glued homologue, *Drosophila*)	AA488168	*DCTN1*	GO:0007067	GO:0003774	0.011	1.48
Transmembrane and coiled-coil domains 7	AI057241	*TMCO7*	GO:0007076	GO:0005488	0.011	−1.60
Sestrin 3	AI190194	*SESN3*	GO:0007050	GO:0005554	0.002	−1.90
LATS, large tumor suppressor, homologue 1	AI023733	*LATS1*	GO:0000086	GO:0000166	0.009	−1.98
Transforming, acidic coiled-coil containing protein 1	AA598796	*TACC1*	GO:0007049	GO:0005515	0.007	−2.17
Protein phosphatase 2	H09640	*PPP2R1B*	GO:0000074	GO:0000158	0.019	−2.35
Katanin p60 subunit A 1	T47614	*KATNA1*	GO:0007049	GO:0000166	0.005	−2.66
G_1 to S phase transition 1	R62452	*GSPT1*	GO:0000082	GO:0000166	0.000	−3.50
Response to exogenous agents						
Chromosome 10 open reading frame 59	AI093491	*C10orf59*	GO:0006725	GO:0004497	0.013	1.93
Thioredoxin reductase 1	AA464849	*TXNRD1*	GO:0045454	GO:0015036	0.006	1.92
Epoxide hydrolase 1, microsomal (xenobiotic)	AA838691	*EPHX1*	GO:0006805	GO:0004301	0.012	1.78
Retinol dehydrogenase 11 (all-*trans*/9-*cis*/11-*cis*)	H82421	*RDH11*	GO:0008152	GO:0016491	0.016	1.64
N-Acetyltransferase 2 (arylamine *N*-acetyltransferase)	AI460128	*NAT2*	GO:0008152	GO:0004060	0.009	1.50
Immunoglobulin (CD79A) binding protein 1	AA463498	*IGBP1*	GO:0042113	GO:0008601	0.005	1.38
Calcium binding atopy-related autoantigen 1	AA992324	*CBARA1*	GO:0006952	GO:0005509	0.013	1.38
Toll-interleukin 1 receptor	AI279454	*TIRAP*	GO:0006954	GO:0004888	0.009	1.38

(*Continued on the following page*)

Table 3. Genes differentially expressed in the breast epithelium of parous control women (Cont'd)

Gene name	Gene ID	Symbol	GO no.	Molecular function GO no.	$P_{adjusted}$	Fold increase/ decrease
Epoxide hydrolase 1, microsomal (xenobiotic)	AA838691	*EPHX1*	GO:0006805	GO:0004301	0.012	1.25
Glutathione *S*-transferase θ1	T64869	*GSTT1*	GO:0006950	GO:0004364	0.013	1.24
Cell transport						
Armadillo repeat containing 1	AA490502	*ARMC1*	GO:0030001	GO:0046872	0.0122	1.65
Solute carrier family 19, member 3	AA707858	*SLC19A3*	GO:0006810	GO:0005386	0.0232	1.64
Translocation protein 1	T98628	*TLOC1*	GO:0015031	GO:0004872	0.0120	1.63
SH3KBP1 binding protein 1	AA457723	*SHKBP1*	GO:0006813	GO:0005216	0.0232	1.63
Tweety homologue 1 (*Drosophila*)	R56769	*TTYH1*	GO:0006826	GO:0005381	0.0232	1.60
Solute carrier family 22	AA705565	*SLC22A9*	GO:0006810	GO:0005215	0.0119	1.56
Translocated promoter region	AA064778	*TPR*	GO:0006810	GO:0005554	0.0069	1.53
UDP-*N*-acetyl-α-D-galactosamine	AA598949	*GALNT10*	GO:0005794	GO:0003779	0.0071	1.47
HIV-1 Rev binding protein	AA927604	*HRB*	GO:0006406	GO:0003677	0.0124	1.40
Chloride channel 6	H72322	*CLCN6*	GO:0006811	GO:0005247	0.0220	1.35
Transient receptor potential cation channel	AI167481	*TRPM1*	GO:0006812	GO:0005262	0.0144	1.24
Dysbindin domain containing 2	AA598970	*DBNDD2*	GO:0015031	GO:0005515	0.0247	−1.31
Frequenin homologue (*Drosophila*)	H16821	*FREQ*	GO:0005794	GO:0005509	0.0174	−1.42
Sorting nexin 11	H16467	*SNX11*	GO:0007242	GO:0005554	0.0247	−1.46
Acyl-CoA oxidase 1, palmitoyl	AI079148	*ACOX1*	GO:0006118	GO:0003995	0.0101	−1.53
Ficolin (collagen/fibrinogen domain containing) 1	AI349250	*FCN1*	GO:0006817	GO:0003823	0.0084	−1.66
Cytochrome b_5 reductase 4	AI053851	*CYB5R4*	GO:0006118	GO:0004128	0.0139	−1.70
Solute carrier family 20 (phosphate transporter)	AA933776	*SLC20A2*	GO:0006810	GO:0004872	0.0108	−1.71
Stonin 2	AA992626	*STON2*	GO:0006886	GO:0005515	0.0056	−1.72
RAP1B, member of RAS oncogene family	AA598864	*RAP1B*	GO:0006886	GO:0005525	0.0142	−1.81
Kelch-like 2, Mayven (*Drosophila*)	AI348818	*KLHL2*	GO:0006886	GO:0005515	0.0106	−2.05
γ-Aminobutyric acid A receptor	R43452	*GABRB3*	GO:0006811	GO:0004890	0.0042	−2.36
Ras-GTPase–activating protein SH3 domain binding	AA598628	*G3BP*	GO:0006810	GO:0000166	0.0020	−3.82
Chromatin modification						
SET domain containing 1A	AA459896	*SETD1A*	GO:0016568	GO:0003723	0.016	1.87
Histone cluster 1, H2ac	N50797	*HIST1H2AC*	GO:0007001	GO:0003677	0.027	1.27
Development-morphogenesis						
Dopey family member 2	W15495	*DOPEY2*	GO:0007275	GO:0005554	0.012	2.57
DiGeorge syndrome critical region gene 14	AI369125	*DGCR14*	GO:0007399	GO:0005554	0.004	2.32
Fibroblast growth factor 11	AA936128	*FGF11*	GO:0007399	GO:0008083	0.004	2.26
Dishevelled, dsh homologue 2 (*Drosophila*)	R38325	*DVL2*	GO:0007275	GO:0004871	0.007	2.07
Latent transforming growth factor β binding protein 4	R87406	*LTBP4*	GO:0007275	GO:0008083	0.007	1.99
Ephrin-B3	AA485665	*EFNB3*	GO:0030154	GO:0005005	0.000	1.64
Twist homologue 1	AI220198	*TWIST1*	GO:0009653	GO:0003677	0.003	−1.25
Bruno-like 4, RNA binding protein (*Drosophila*)	R52541	*BRUNOL4*	GO:0009790	GO:0003676	0.0143	−1.34
Cysteine-rich transmembrane BMP regulator 1	R78638	*CRIM1*	GO:0007399	GO:0004867	0.012	−1.35
Protein inhibitor of activated STAT, 2	AI151206	*PIAS2*	GO:0007275	GO:0003677	0.001	−1.50
Hepatic leukemia factor	R59192	*HLF*	GO:0007275	GO:0003690	0.012	−1.50
Dual specificity phosphatase 22	AA454636	*DUSP22*	GO:0000188	GO:0008138	0.006	−1.64
Split hand/foot malformation (ectrodactyly) type 1	R38516	*SHFM1*	GO:0030326	GO:0005515	0.012	−1.72
Tropomyosin 3	AA206591	*TPM3*	GO:0007517	GO:0003779	0.002	−2.23
Microtubule-associated protein 1B	H17493	*MAP1B*	GO:0007517	GO:0005198	0.003	−2.64
DNA repair and replication						
Translin	AA460927	*TSN*	GO:0006310	GO:0003677	0.065	1.94
RAD51-like 3 (*S. cerevisiae*)	N29765	*RAD51L3*	GO:0006284	GO:0005524	0.024	1.92
Nth endonuclease III-like 1 (*E. coli*)	AI369190	*NTHL1*	GO:0006284	GO:0003677	0.009	1.92
Ankyrin repeat domain 17	R37816	*ANKRD17*	GO:0006298	GO:0003676	0.055	1.78
Three prime repair exonuclease 1	AI352447	*TREX1*	GO:0006281	GO:0003697	0.011	1.54
Polymerase (DNA-directed)	AI017254	*POLD3*	GO:0000731	GO:0003891	0.007	1.50
Excision repair cross-complementing rodent repair deficiency	N49276	*ERCC8*	GO:0006281	GO:0003702	0.040	1.25
Ubiquitin-activating enzyme E1	AA598670	*UBE1*	GO:0006260	GO:0000166	0.044	−1.25
Structural maintenance of chromosomes 2	AA598549	*SMC2*	GO:0000067	GO:0000166	0.016	−1.61
Miscellaneous processes						
Diaphanous homologue 3 (*Drosophila*)	AI018026	*DIAPH3*	GO:0016043	GO:0003779	0.004	2.22
Thrombospondin, type I, domain containing 4	AA120866	*THSD4*	GO:0031012	GO:0008233	0.009	1.85
Carcinoembryonic antigen-related cell adhesion	AI242105	*CEACAM4*	GO:0005887	GO:0005554	0.010	1.65
Sarcospan (Kras oncogene-associated gene)	AA458998	*SSPN*	GO:0006936	GO:0005554	0.006	1.59
Annexin A5	AI269079	*ANXA5*	GO:0007596	GO:0004859	0.007	−1.39
Lactamase, β	AI273225	*LACTB*	GO:0046677	GO:0016787	0.002	−1.81
RNA processing						
DEAD (Asp-Glu-Ala-Asp) box polypeptide 17	H82870	*DDX17*	GO:0006396	GO:0008026	0.003	3.02
Tetratricopeptide repeat domain 8	W37689	*TTC8*	GO:0007600	GO:0005488	0.016	1.87

(*Continued on the following page*)

Table 3. Genes differentially expressed in the breast epithelium of parous control women (Cont'd)

Gene name	Gene ID	Symbol	GO no.	Molecular function GO no.	$P_{adjusted}$	Fold increase/ decrease
Survival of motor neuron protein interacting protein 1	N26026	*SIP1*	GO:0000245	GO:0031202	0.017	1.74
Eukaryotic translation initiation factor 4A, isoform 3	N79030	*EIF4A3*	GO:0006364	GO:0005524	0.021	1.71
Processing of precursor 7	H71218	*POP7*	GO:0008033	GO:0003676	0.010	1.54
Pseudouridylate synthase 7 homologue (*S. cerevisiae*)	AA434411	*PUS7*	GO:0008033	GO:0004730	0.017	1.44
BMS1-like, ribosome assembly protein (yeast)	AA915891	*BMS1L*	GO:0007046	GO:0000166	0.002	−1.69
Brix domain containing 2	AI025116	*BXDC2*	GO:0007046	GO:0005554	0.000	−2.10
Splicing factor, arginine/serine-rich 10	AI583623	*SFRS10*	GO:0000398	GO:0000166	0.002	−2.20
Metabolism						
Dihydrolipoamide branched chain transacylase E2	AI004719	*DBT*	GO:0008152	GO:0005515	0.002	2.98
Homer homologue 1 (*Drosophila*)	AA903860	*HOMER1*	GO:0007206	GO:0005515	0.005	2.02
Heparan sulfate (glucosamine) 3-*O*-sulfotransferase 4	AA973808	*HS3ST4*	GO:0030201	GO:0008467	0.046	1.98
Dehydrodolichyl diphosphate synthase	AA995913	*DHDDS*	GO:0008152	GO:0016740	0.009	1.96
Acyl-CoA synthetase short-chain family member 1	N67766	*ACSS1*	GO:0008152	GO:0003824	0.022	1.76
Fumarylacetoacetate hydrolase	AA010559	*FAH*	GO:0006559	GO:0000287	0.028	1.29
SID1 transmembrane family, member 2	T98941	*SIDT2*	GO:0016042	GO:0003847	0.018	1.27
Protein tyrosine phosphatase, receptor type, B	N66422	*PTPRB*	GO:0006796	GO:0005529	0.014	1.26
Acyl-CoA synthetase long-chain family member 3	AI206454	*ACSL3*	GO:000663	GO:0003824	0.012	−1.25
Amylase, α 1A; salivary	R64129	*AMY1A*	GO:0005975	GO:0004556	0.009	−125
Protein biosynthesis and metabolism						
Lysyl oxidase	H80737	*LOX*	GO:0006464	GO:0004720	0.002	3.67
Tubulin tyrosine ligase-like family, member 5	R34225	*TTLL5*	GO:0006464	GO:0004835	0.011	3.20
Vacuolar protein sorting 13 homologue C (*S. cerevisiae*)	AA663968	*VPS13C*	GO:0008104	GO:0005554	0.004	2.39
Protein tyrosine phosphatase, receptor type, C	AA703526	*PTPRC*	GO:0006470	GO:0004725	0.007	1.32
Ribosomal protein L9	AI199007	*RPL9*	GO:0006412	GO:0003723	0.011	1.26
GrpE-like 1, mitochondrial (*E. coli*)	AA449720	*GRPEL1*	GO:0006457	GO:0000774	0.011	1.25
Protein tyrosine phosphatase, nonreceptor type 21	W72293	*PTPN21*	GO:0006470	GO:0004725	0.028	1.25
β-site APP-cleaving enzyme 2	AA457119	*BACE2*	GO:0006464	GO:0004194	0.028	1.24
Hypothetical protein MGC42105	AA416627	*MGC42105*	GO:0006468	GO:0046872	0.022	1.24
Eukaryotic translation initiation factor 2B	AI174400	*EIF2B1*	GO:0006412	GO:0003743	0.001	−1.45
Par-3 partitioning defective 3 homologue (*C. elegans*)	AA902790	*PARD3*	GO:0006461	GO:0005515	0.002	−1.55
Transient receptor potential cation channel	AA598596	*TRPC4AP*	GO:0006461	GO:0005524	0.007	−1.57
GrpE-like 2, mitochondrial (*E. coli*)	AA598831	*GRPEL2*	GO:0006457	GO:0000774	0.019	−1.61
Mitochondrial ribosomal protein S16	AA887401	*MRPS16*	GO:0006412	GO:0003735	0.001	−1.67
Collagen, type IV	AA971606	*COL4A3BP*	GO:0006468	GO:0004674	0.013	−1.67
Mitochondrial ribosomal protein S11	AI032875	*MRPS11*	GO:0006412	GO:0003735	0.013	−1.67
Farnesyltransferase, CAAX box, α	AA283874	*FNTA*	GO:0018347	GO:0004660	0.024	−1.72
Tryptophanyl tRNA synthetase 2 (mitochondrial)	R43272	*WARS2*	GO:0006412	GO:0000166	0.017	−1.72
Capping protein (actin filament) muscle Z-line	W92769	*CAPZA1*	GO:0006461	GO:0003779	0.004	−1.78
Lipase, hepatic	AI054269	*LIPC*	GO:0006487	GO:0004806	0.0028	−2.13
Eukaryotic translation initiation factor 1A	AI214283	*EIF1AY*	GO:0006412	GO:0003723	0.013	−2.15
Serine/threonine/tyrosine interacting-like 1	AI205036	*STYXL1*	GO:0006470	GO:0008138	0.0066	−2.69
Proteolysis and ubiquitination						
Cathepsin B	AI091648	*CTSB*	GO:0006508	GO:0004213	0.004	2.16
E3 ubiquitin protein ligase	W86992	*EDD1*	GO:0006511	GO:0004840	0.016	1.66
Dipeptidyl-peptidase 3	AA430361	*DPP3*	GO:0006508	GO:0004177	0.007	1.54
Peptidase D	AA481543	*PEPD*	GO:0006508	GO:0004251	0.031	1.42
Ring finger protein 44	AI675516	*RNF44*	GO:0016567	GO:0004842	0.014	1.32
Gem (nuclear organelle) associated protein 4	AA041254	*GEMIN4*	GO:0000398	GO:0005515	0.028	1.30
Arginyltransferase 1	AI015417	*ATE1*	GO:0006512	GO:0004057	0.009	1.26
Heterogeneous nuclear ribonucleoprotein R	AA779191	*HNRPR*	GO:0006397	GO:0000166	0.013	1.25
SUMO-1 activating enzyme subunit 1	AA598486	*SAE1*	GO:0016567	GO:0004839	0.017	−1.58
Ubiquitin specific peptidase 30	AI055850	*USP30*	GO:0006511	GO:0004197	0.000	−1.77
TRIAD3 protein	AI051657	*TRIAD3*	GO:0006512	GO:0008270	0.005	−2.04
Ubiquitin-conjugating enzyme E2E 1	AA197307	*UBE2E1*	GO:0006511	GO:0004840	0.013	−2.07
Ariadne homologue	AI185068	*ARIH1*	GO:0006511	GO:0004842	0.003	−2.14
AFG3 ATPase family gene 3-like 2 (yeast)	AI219905	*AFG3L2*	GO:0006508	GO:0000166	0.001	−2.18
Transcription						
Suppressor of Ty 5 homologue (*S. cerevisiae*)	R21511	*SUPT5H*	GO:0000122	GO:0003711	0.043	2.15
Inhibitor of DNA binding 4	AA464856	*ID4*	GO:0006357	GO:0003714	0.001	2.10
Bromodomain PHD finger transcription factor	AA704421	*BPTF*	GO:0000122	GO:0005515	0.066	2.00
Zinc finger protein 498	W94267	*ZNF498*	GO:0006355	GO:0003676	0.005	2.00
SRY (sex determining region Y)-box 10	AA976578	*SOX10*	GO:0006350	GO:0003677	0.028	1.93
Zinc finger protein 710	AI025842	*ZNF710*	GO:0006355	GO:0003676	0.005	1.90

(*Continued on the following page*)

Table 3. Genes differentially expressed in the breast epithelium of parous control women (Cont'd)

Gene name	Gene ID	Symbol	GO no.	Molecular function GO no.	$P_{adjusted}$	Fold increase/ decrease
General transcription factor IIB	H23978	*GTF2B*	GO:0006355	GO:0016251	0.009	1.54
Zinc finger protein 26	R97944	*ZNF26*	GO:0006355	GO:0003676	0.017	1.53
Zinc finger protein 268	AI277336	*ZNF268*	GO:0006355	GO:0008270	0.014	1.50
Protein inhibitor of activated STAT, 1	N91175	*PIAS1*	GO:0006350	GO:0003677	0.0280	1.31
Kv channel interacting protein 3, calsenilin	H39123	*KCNIP3*	GO:0006350	GO:0003677	0.024	1.30
Zinc finger protein 275	AA406125	*ZNF275*	GO:0006355	GO:0003677	0.032	1.28
Homeobox D1	W68537	*HOXD1*	GO:0006355	GO:0003700	0.032	1.26
HIR histone cell cycle regulation defective homologue A	AA609365	*HIRA*	GO:0006357	GO:0003700	0.022	1.26
Forkhead box K2	AA136472	*FOXK2*	GO:0006350	GO:0003700	0.014	1.26
Transducin-like enhancer of split 3 (E(sp1) homologue)	AI216623	*TLE3*	GO:0006355	GO:0005554	0.013	1.25
p300/CBP-associated factor	N74637	*PCAF*	GO:0006350	0.001368735	0.050	1.25
Zinc finger protein 544	AA885065	*ZNF544*	GO:0006355	GO:0003676	0.010	1.25
Regulatory factor X-associated protein	AI365571	*RFXAP*	GO:0006366	GO:0003700	0.013	1.24
Bromodomain adjacent to zinc finger domain, 2A	AA699460	*BAZ2A*	GO:0006355	GO:0003677	0.015	1.24
Zinc finger protein 16	H17016	*ZNF16*	GO:0006350	GO:0003677	0.012	1.23
Ring finger protein 12	AA598809	*RNF12*	GO:0006350	GO:0003714	0.007	−1.25
POU domain, class 6, transcription factor 1	AI123130	*POU6F1*	GO:0006355	GO:0003700	0.008	−1.37
RAR-related orphan receptor A	AI022327	*RORA*	GO:0006350	GO:0003700	0.011	−1.40
Myeloid/lymphoid or mixed-lineage leukemia	AI197974	*MLLT6*	GO:0006355	GO:0005515	0.009	−1.41
Zinc finger protein 425	H20279	*ZNF425*	GO:0006355	GO:0003676	0.005	−1.62
PBX/knotted 1 homeobox 2	AI024125	*PKNOX2*	GO:0006355	GO:0003700	0.004	−1.64
D4, zinc and double PHD fingers family 2	AA496782	*DPF2*	GO:0006350	GO:0003676	0.016	−1.69
General transcription factor IIIC	AI184450	*GTF3C4*	GO:0006350	GO:0003677	0.002	−1.87
GATA zinc finger domain containing 2A	AA458840	*GATAD2A*	GO:0006306	GO:0030674	0.022	−1.97
SRY (sex determining region Y)-box 3	AI359981	*SOX3*	GO:0006355	GO:0003677	0.004	−2.11
HDAC8	AI053481	*HDAC8*	GO:0000122	GO:0004407	0.002	−2.20
Methyl-CpG binding domain protein 3	AI017865	*MBD3*	GO:0006350	GO:0003677	0.002	−3.17
Biological process unknown						
DEAH (Asp-Glu-Ala-Asp/His) box polypeptide 57	AI125363	*DHX57*	GO:0000004	GO:0003697	0.012	2.96
Frequenin homologue (*Drosophila*)	AA918755	*FREQ*	GO:0000004	GO:0005509	0.008	1.98
WD repeat domain 44	W80619	*WDR44*	GO:0000004	GO:0008270	0.016	1.86
Fibulin 2	AA452840	*FBLN2*	GO:0000004	GO:0005509	0.008	1.83
Ectonucleoside triphosphate diphosphohydrolase 3	AI247824	*ENTPD3*	GO:0000004	GO:0004050	0.023	1.26
Zinc finger, DHHC-type containing 9	AI346102	*ZDHHC9*	GO:0000004	GO:0000166	0.028	1.24
Docking protein 5	R39924	*DOK5*	GO:0000004	GO:0005158	0.026	−1.25
Progesterone receptor membrane component 2	AA456304	*PGRMC2*	GO:0000004	GO:0003707	0.011	−1.25
RB1-inducible coiled-coil 1	R38102	*RB1CC1*	GO:0000004	GO:0016301	0.026	−1.30
B-Cell CLL/lymphoma 7A	H90147	*BCL7A*	GO:0000004	GO:0003779	0.013	−1.30
Dedicator of cytokinesis 5	AA932511	*DOCK5*	GO:0000004	GO:0005085	0.004	−1.37
Zinc finger protein 320	AI025436	*ZNF320*	GO:0000004	GO:0003676	0.012	−1.50
Heterogeneous nuclear ribonucleoprotein M	AI220112	*HNRPM*	GO:0000004	GO:0003723	0.001	−1.73
Hypothetical protein MGC4562	AI184226	*MGC4562*	GO:0000004	GO:0003723	0.006	−1.85
Phosphatase and actin regulator 1	R99333	*PHACTR1*	GO:0000004	GO:0005096	0.005	−1.91
DEAD (Asp-Glu-Ala-Asp) box polypeptide 46	AI003503	*DDX46*	GO:0000004	GO:0000166	0.0020	−2.19
Biological process and molecular function unknown						
ORM1-like 1 (*S. cerevisiae*)	AA437132	*ORMDL1*	GO:0000004	GO:0005554	0.002	3.02
Ankyrin repeat domain 12	AA938440	*ANKRD12*	GO:0000004	GO:0005554	0.005	1.89
Transmembrane protein 27	AA999850	*TMEM27*	GO:0000004	GO:0005554	0.005	1.78
DKFZp434A0131 protein	AA032084	*DKFZP434A0131*	GO:0000004	GO:0005554	0.028	1.25
Vitamin K epoxide reductase complex	AI279477	*VKORC1L1*	GO:0000004	GO:0005554	0.010	−1.66
Zinc finger, RAN-binding domain containing 1	AI033098	*ZRANB1*	GO:0000004	GO:0005554	0.005	−1.85
Family with sequence similarity 57, member A	H23091	*FAM57A*	GO:0000004	GO:0005554	0.024	−3.86
Microcephaly, primary autosomal recessive 1	AA156424	*MCPH1*	GO:0000004	GO:0005554	0.010	−3.94
Transmembrane protein 32	AA251026	*TMEM32*	GO:0000004	GO:0005554	0.004	−4.43
Neurensin 2	AI199579	*NRSN2*	GO:0000004	GO:0005554	0.002	−4.85

Abbreviations: CLL, chronic lymphocytic leukemia; ORM1, *Homo sapiens* orosomucoid 1; RAR, retinoic acid receptor; TNFRSF1A, TNF receptor superfamily, member 1A.

breast epithelium of parous women and analyzed the biological significance of those terms that were found to be deregulated in response to an early reproductive event with high statistical significance (Tables 3 and 4). Among the 18 categories identified to contain deregulated genes, the most highly represented biological process was gene transcription, in which 21 (64%) genes were up-regulated and 12 (36%) genes were down-regulated. Higher gene expression was observed in 11 processes that included proteolysis and ubiquitination,

Table 4. RT-PCR validation of genes up-regulated in the breast epithelium of parous women

Gene name	Gene symbol	Primer sequence	Parous control	Nulliparous control	Parous case	Nulliparous case
TNFRSF1A-associated via death domain	*TRADD*	gatggccttagggttccttc	11488.00 ± 985.00	0.57 ± 0.33	2.18 ± 1.57	1.89 ± 1.85
Eukaryotic translation initiation factor 4A, isoform 3	*EIF4A3*	aagaaaggtggactggctga	3822.18 ± 764.10	1.09 ± 1.04	2.29 ± 2.72	12.97 ± 27.51
Suppressor of Ty 5 homologue (*S. cerevisiae*)	*SUPT5H*	ctttgaggggaaccgttaca	1517.76 ± 234.55	0.26 ± 0.12	16.84 ± 26.09	2.90 ± 3.15
SRY (sex determining region Y)-box 5	*SOX5*	agggactcccgagagcttag	267.61 ± 24.87	0.79 ± 0.71	2.73 ± 3.05	6.04 ± 5.71
Carcinoembryonic antigen–related cell adhesion molecule 1	*CEACAM1*	acccacctgcacagtactcc	12.58 ± 1.01	1.97 ± 0.16	0.64 ± 0.06	1.74 ± 0.14
Homeo box D1	*HOXD1*	ttcagcaccaagcaactgac	9.49 ± 3.15	2.57 ± 3.54	2.64 ± 2.34	1.11 ± 1.11
Ephrin B3	*EFNB3*	cttcccaagatctcccttcc	3.63 ± 3.23	1.11 ± 0.08	0.7 ± 0.59	1.17 ± 0.84
p300/CBP-associated factor	*PCAF*	acgttcacctgctggtccaa	98.36 ± 21.44	1.38 ± 0.33	9.87 ± 3.76	2.95 ± 5.34
Inhibitor of DNA binding 4	*ID4*	atgggatgaggaaatgcttg	830.28 ± 100.33	0.21 ± 0.23	33.80 ± 63.44	3.37 ± 3.83
Surfeit 5	*Surfeit*	cctgcctgcaggttagaaag	1.12 ± 1.13	1.75 ± 0.08	1.35 ± 0.67	1.53 ± 2.16

cell adhesion, response to exogenous agents, metabolism, DNA repair and replication, RNA processing, apoptosis, miscellaneous processes, antiapoptosis, and chromatin modification, in which the ratios of up-regulated to down-regulated genes ranged from 1.75 to 11 (Table 3). A greater number of genes with lower level of expression were observed in various processes that included: cell transport, protein biosynthesis and metabolism, cell signaling-signal transduction, biological process unknown, and biological process and molecular function unknown. The genes composing these categories are listed in Table 3.

A number of genes that in the arrays of the parous control breast epithelial cells were either significantly up-regulated or not modified by the reproductive process were confirmed by RT-PCR. They included tumor necrosis factor receptor superfamily, member 1A–associated via death domain (*TRADD*), eukaryotic translation initiation factor 4A, isoform 3 (*EIF4A3*), suppressor of Ty 5 homologue (*S. cerevisiae*) (*SUPT5H*; ref. 35), sex determining region Y (SRY)-box 5 (*SOX5*), carcinoembryonic antigen-related cell adhesion molecule 1 (*CEACAM1*), homeobox D1 (*HOXD1*), ephrin B3 (*EFNB3*), p300/CREB-binding protein (CBP)–associated factor (*PCAF*), inhibitor of DNA binding 4 (*ID4*), and *Surfeit* (Table 4). All genes detected as differentially expressed by the microarray platform were confirmed to be differentially expressed by RT-PCR ($P < 0.5$), whereas those that did not differ among parous and nulliparous control and cases, such as *Surfeit*, did not differ in the level of expression by RT-PCR (Table 4).

Discussion

The present work is the first demonstration that an early first full-term pregnancy imprints in the involuted breast lobules of postmenopausal parous women free of breast cancer a specific genomic signature that significantly differs from that of parous women with cancer and nulliparous women with or without the disease. The cDNA microarray analysis of epithelial RNA of completely involuted lobules, represented by Lob 1, obtained by laser capture microdissection, revealed that these cells express a genomic signature composed of 232 deregulated genes representing 18 functional categories.

The signature is composed of both up-regulated and down-regulated genes. Deregulated genes predominated in the category of transcription, in which 63% were up-regulated and 37% down-regulated. The fact that the number of down-regulated genes was slightly higher in the cell transport, protein biosynthesis metabolism, cell signaling-signal transduction, development and morphogenesis, cell cycle and growth, as well as in those categories in which the biological process and the molecular functions are unknown indicates that down-regulation and/or silencing of gene expression plays an important role in the differentiation of the breast induced by pregnancy, as shown in experimental models (11-15, 25).

Twenty-three genes were found to be significantly up-regulated in the parous breast epithelium in the categories of transcription and chromatin modification, an indication that modifications in transcriptional activity during pregnancy play an important role and become a permanent component of the genomic signature imprinted by this physiologic process in the postmenopausal breast epithelium. More than 2-fold significant increase ($P < 0.05$) over control values was observed in the bromodomain PHD finger transcription factor (*BPTF*); *SUPT5*, which has 50% similarity to yeast *SPT5* and is part of a protein complex involved in transcriptional repression by modulating chromatin structure (36); and zinc finger protein 498 (*ZNF498*), which is involved in the regulation of nucleobase, nucleoside, nucleotide, and nucleic acid metabolism. The expression of *BPTF* has been reported to be lost or significantly reduced in primary carcinomas and in cell lines established from different human carcinomas, supporting our postulate that this gene may play a role in suppression of tumors originating from epithelial tissue (37, 38). *ID4*, a member of the ID family of proteins (Id1–Id4) that function as dominant-negative regulators of basic helix-loop-helix transcription factors, was increased in the parous women epithelium, as confirmed by RT-PCR that detected significant increase in the levels of expression from 0.21 ± 0.23 in nulliparous controls to 830.28 ± 100.33 in parous controls (Table 4). *ID4* mRNA has been reported to be expressed in normal breast epithelium and myoepithelium but to be absent in estrogen receptor α (ER-α)–positive invasive carcinomas, sporadic breast cancers expressing both *ER-α* and *BRCA1* (39), ductal carcinomas *in situ*, and atypical ductal hyperplasias (40). Epigenetic inactivation of *ID4* has been reported in human leukemia (41), colorectal cancer (42), and gastric

adenocarcinoma (43). Its complete or partial epigenetic inactivation also occurs in both ER-α–positive and ER-α–negative cells (i.e., T47D, MCF-7, and HBL-100, BT20, BT549, and BR2, respectively; ref. 44), findings that support the role of this gene as a putative tumor-suppressor gene and as a key controller of cell differentiation.

The fact that SRY box 10 or *SOX10*, a gene that is methylated in the breast cancer cell line MCF7 (45), is significantly up-regulated in the breast of parous women indicates that it may play an integral role in the specification and transcription of the terminal differentiation that has been reported in other systems, such as astrocytes and oligodendrocytes (46). In contrast, SRY box 3 (*SOX3*), which is involved in the regulation of embryonic development and determination of cell fate (47) and is essential for the maintenance of spermatogonial stem cells (48), is down-regulated in the parous breast. These observations suggest that these genes might play in the breast a role similar to that described in neural and male reproductive organs, respectively. Nevertheless, the final molecular mechanisms by which these transcription factors regulate the differentiation of the parous breast epithelium need further investigation. Transcription factors also associated to coactivators and chromatin remodeling, such *PCAF*, which have previously found to be significantly up-regulated in breast epithelial cells of parous women (6, 25-27), seem to play

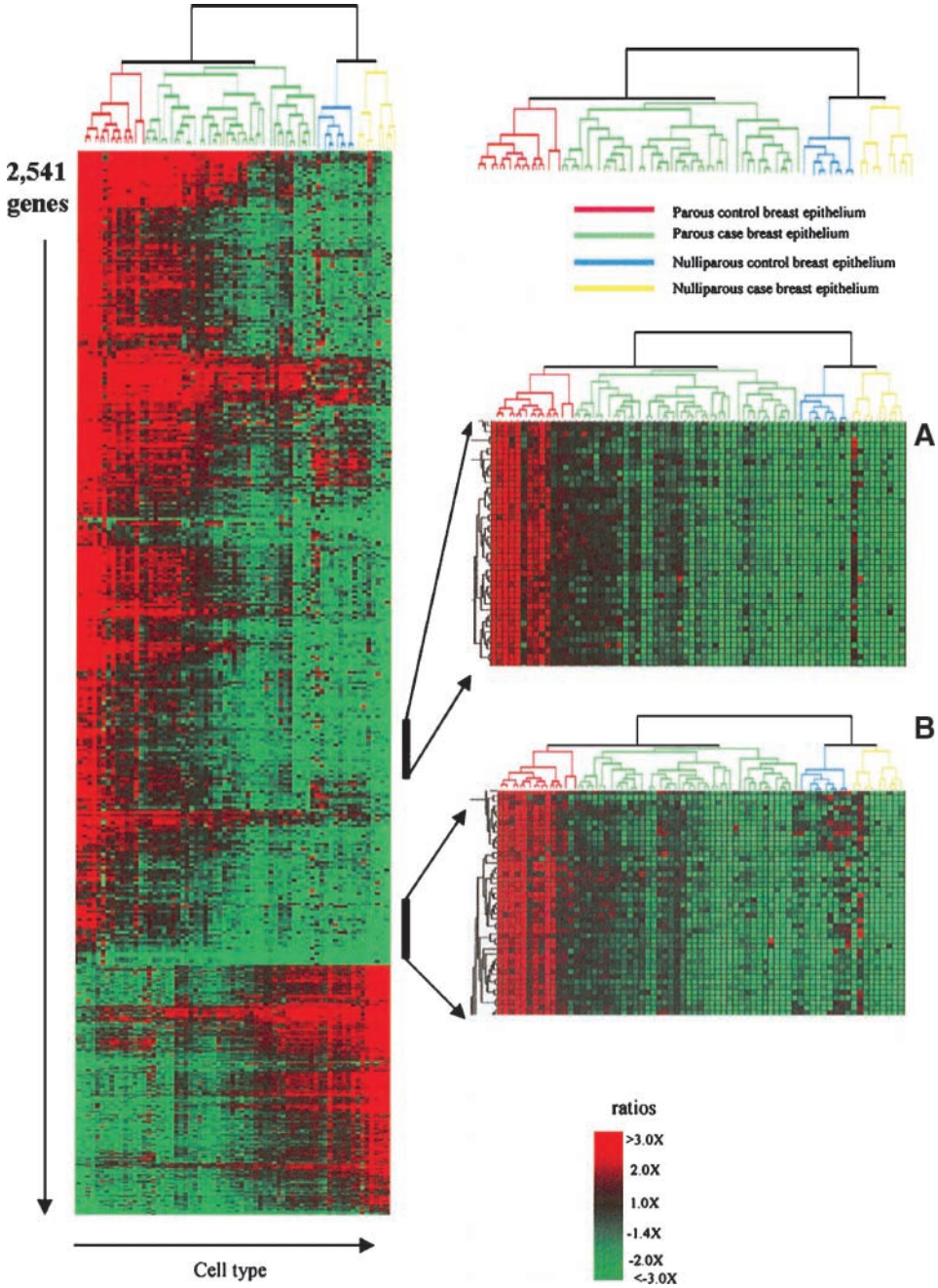

Figure 1. Unsupervised hierarchical clustering analysis using the expression profiles of 2,541 globally varying genes across the nulliparous and parous data sets representing parous controls (—), parous cases (—), nulliparous controls (—), and nulliparous cases (—). The clustering procedure used to derive the dendrogram is described in Materials and Methods.

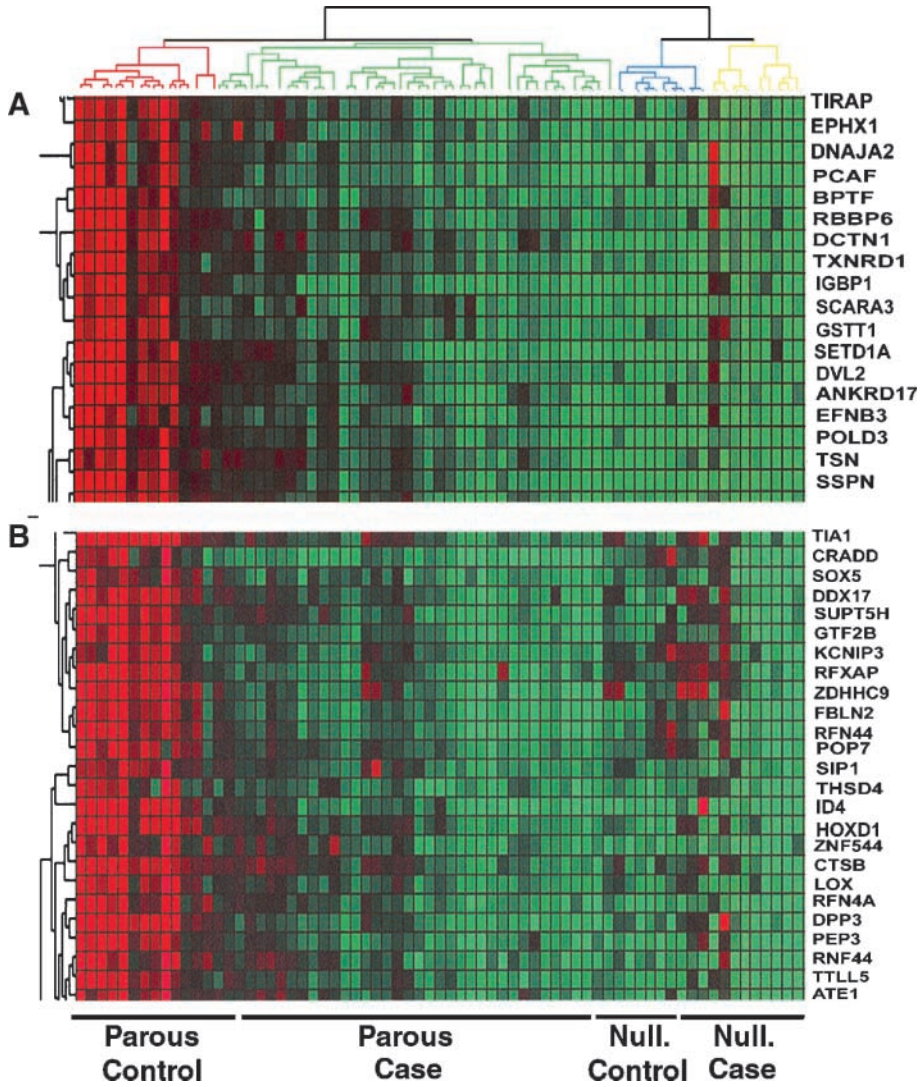

Figure 2. A and **B,** unsupervised hierarchical analysis of subsets of 18 matched breast epithelia from the parous control specimens shown in Fig. 1 that were microdissected and hybridized independently as biological replicates. The combined parity/absence of breast cancer data generated a distinct genomic profile that differed from those of the breast cancer groups, irrespective of parity history, and of the nulliparous cancer-free control group. Groups identified as for Fig. 1.

an important role in the genomic signature induced by pregnancy in breast epithelial cells. The p300/CBP family of coactivators can interact with the isolated A/B domain of the ER-α, enhancing its AF-1 activity and thus contributing to ligand-independent activity of the receptor under the stimulus of steroid receptor coactivator-1 (49). Interestingly, p300/CBP is recruited by steroid receptor coactivator-1 and cofactors such as transcription intermediary factor 2 and amplified in breast cancer 1, which interact with nuclear receptors in a ligand-dependent manner for enhancing transcriptional activation by the receptor via histone acetylation/methylation (50). PCAF is also a coactivator of the tumor suppressor p53 and participates in p53-mediated transactivation of target genes through acetylation of both bound p53 and histones within p53 target promoters (51). The up-regulation of *PCAF* in the differentiated breast epithelial cells of parous women might be associated with an increase in the protein levels of the histone acetyl transferases p300, whereas CBP suppresses the level of histone deacetylase (HDAC) and increases the level of acetylated histone H4, as it has been reported for metastatic breast cancer cells after treatment with all-*trans* retinoic acid, which also up-regulates the expression of *BAX* (52), a proapoptotic gene that is also up-regulated in the parous breast epithelial cells.

The general transcription factor IIB (*GTF2B*), which encodes one of the ubiquitous factors required for transcription initiation by RNA polymerase II and *HOXD1*, is also up-regulated in the parous breast. Of great interest is the fact that transcription factors encoded by the *HOX* genes, which play a crucial role in *Drosophila, Xenopus*, and mammalian embryonic differentiation and development, up-regulate *HOXC6, HOXD1*, and *HOXD8* expression in human neuroblastoma cells that are chemically induced to differentiate, an indication that *HOX* is associated with maturation toward a differentiated neuronal phenotype (53).

Two protein inhibitors of activated STAT (*PIAS*) were found to be deregulated in the breast epithelium of parous women; *PIAS1* was up-regulated and *PIAS2* (also called *PIASx*) was down-regulated. Members of the PIAS protein family have been identified as negative regulators of STAT signaling and of transcription factors such nuclear factor κB and p53 (54). PIAS members have small ubiquitin-like modifier (SUMO)

E3-ligase activity, PIAS1 exerting a direct inhibition of STAT1 DNA binding whereas PIAS2 recruiting HDAC3 for repressing STAT4-dependent transcription. Several PIAS can also cause STAT sumoylation, which is likely to inhibit STAT signaling (55). The down-regulation of *PIAS2* needs further analysis because the extent of *PIAS* and *SUMO* family expression in breast tissues remains unclear, although preliminary evidence suggests that dysregulation of *PIAS* expression does occur in human breast cancers.

Among the genes that are down-regulated in the involuted lobular epithelium of postmenopausal parous women are *HDAC8* and methyl-CpG binding domain protein 3 (*MBD3*). The importance of the down-regulation of these two genes is highlighted by the fact that HDACs interact with DNA methyltransferases and methyl CpG-binding domain (MBD) proteins, which are associated with CpG island methylation, another epigenetic modification involved in transcriptional repression and heterochromatin remodeling (56-59). The inhibition of HDAC by trichostatin A induces terminal differentiation of mouse erythroleukemia cells and apoptosis of lymphoid and colorectal cancer cells. In addition, trichostatin A treatment of cells expressing the PML zinc finger protein derepresses transcription and allows cells to differentiate normally. These findings have led to the development of HDAC inhibitors as potential agents for the treatment of certain forms of cancer (57). Interestingly, the deacetylase activity of HDAC8 is inhibited by protein kinase A–mediated phosphorylation, resulting in the hyperacetylation of histones H3 and H4, a phenomenon similar to that induced by human chorionic gonadotropin in the human breast (58) and which represents a novel mechanism of regulation of the activities of human class I HDAC by protein kinases (56). MBD3 is one of the five members of the MBD family that recruit various HDAC-containing repressor complexes leading to silencing by generating repressive chromatin structures at relevant binding sites. It plays an important role in mediating the HDAC-specific small-molecule inhibitor (*HDI*)–induced gene regulations associated with cancer-selective cell death, imparting HDI-induced selectivity in cancer cells via differential transcriptional regulation (59). Silencing of *MBD3* abrogates HDI-induced transcriptional reprogramming and growth inhibition in HDI-treated lung cancer cells but not in normal cells. In response to HDI treatment, *MBD3* relocalizes within cells in a different manner in cancer and normal cells, an indication that the relocation of *MBD3* to the nucleus may facilitate its recruitment to the genome and allow *MBD3* to function as a regulatory molecule (59). Our ongoing studies have been designed for clarifying whether intracellular relocation plays a role in differential transcriptional reprogramming in response to pregnancy-induced differentiation.

We found of great interest our findings that genes that are involved in the metabolism of xenobiotic substances and oxidative stress were significantly up-regulated in the breast epithelium of postmenopausal parous women. Among them are epoxide hydrolase or *EPHX1*, which plays an important role in both the activation and detoxification of exogenous chemicals such as polycyclic aromatic hydrocarbons found in cigarette smoke (60), and thioredoxin reductase 1 or *TXNRD1* (61), a member of the pyridine nucleotide family of oxidoreductases and one of the major antioxidant and redox regulators in mammals. TXNRD1 protein reduces thioredoxins and other substrates, playing a role in selenium metabolism, protecting against oxidative stress, and supporting the function of p53 and of other tumor suppressors. The up-regulation in the parous breast epithelium of glutathione *S*-transferase (GST) θ1 (*GSTT1*), which belongs to a family of important enzymes involved in the detoxification of a wide variety of known and suspected carcinogens, including potential mammary carcinogens identified in charred meats and tobacco smoke, is of importance because a substantial proportion of the Caucasian population has a homozygous deletion (null) of the *GSTM1* or *GSTT1* gene, which results in lack of production of these isoenzymes and a significantly elevated risk of breast cancer associated with cigarette smoking (62). *N*-Acetyltransferase 2-arylamine *N*-acetyltransferase (*NAT2*), which is involved in the metabolism of different xenobiotics including potential carcinogens (63), indicates that the lifetime sequel of the differentiation of the breast by an early pregnancy is the activation of a system of defense that makes the parous breast cells less vulnerable to genotoxic substances. This contention is supported by *in vitro* data showing that breast epithelial cells from parous women do not express phenotypes of cell transformation when treated with chemical carcinogens, whereas those from nulliparous women do (64).

Seven DNA repair controlling genes were found to be significantly up-regulated in the Lob 1 of the parous breast, an indication that an improved DNA repair system was involved in the protective effect induced by pregnancy, as we have previously shown in the rodent experimental model in which mammary epithelial cells of parous animals remove 7,12-dimethylbenz(*a*)anthraacene DNA adducts more efficiently than those of virgin animals (65). DNA repair is central to the integrity of the human genome and reduced DNA repair capacity has been linked to genetic susceptibility to cancer, including that of the breast (66). Among the genes that were up-regulated in the epithelial cells of the parous breast was RAD51-like 3 or *RAD51D*, which is one of the five RAD51 paralogues that are required in mammalian cells for normal levels of genetic recombination and damaging agents (67). We have previously reported that the X ray repair complementing defective repair I (*XRCC4*) gene is up-regulated in breast epithelial cells of parous women (6, 26). XRCC4 is a DNA repair factor that is essential for the resolution of DNA double-strand break during V(D)J recombination, acting as a caretaker of the mammalian genome in both normal development and suppression of tumors. In the present study, we found in the same cells the up-regulation of excision repair cross-complementing rodent repair deficiency, complementation group 8 (*ERCC8*), also known as *CSA* (68), which interacts with CSB, and when mutated the transcription-coupled repair, a DNA repair defect found in Cockayne syndrome, is impaired (68). Ankyrin repeat domain 17 (*ANKRD17*), translin or *TSN*, which encodes a DNA-binding protein that specifically recognizes conserved target sequences at the breakpoint junction of chromosomal translocations (69), and three prime repair exonuclease 1 (*TREX1*) are also up-regulated in the parous control group. The protein encoded by this

latter gene uses two different open reading frames from which the upstream open reading frame encodes proteins that interact with the ataxia telangiectasia and Rad3–related protein, a checkpoint kinase. The proteins encoded by this upstream open reading frame localize to intranuclear foci following DNA damage and are essential components of the DNA damage checkpoint (70, 71). These data indicate that the activation of genes involved in the DNA repair process is part of the signature induced in the mammary gland by pregnancy, confirming previous findings that, *in vivo*, the ability of the cells to repair carcinogen-induced damage by unscheduled DNA synthesis and adduct removal is more efficient in the post pregnancy mammary gland (65).

Among the genes that control apoptosis, eight were deregulated, six were up-regulated, and two down-regulated. The former included the BCL2-associated X protein (*BAX*), a proapoptotic gene that belongs to the BCL2 protein family whose transcription is stimulated by the active p53 and the proapoptotic and cell cycle regulator gene *p21* (72). To the same category belong the cytotoxic granule–associated RNA binding protein (*TIA1*), tumor necrosis factor (TNF) receptor–associated factor 1 (*TRAF1*), *TRADD*, CASP2 and RIPK1 domain containing adaptor with death domain (*CRADD*), and protein phosphatase 1F (*PPM1F*). *TIA1* possesses nucleolytic activity against CTL target cells inducing in them DNA fragmentation (73). *TNFR1* can initiate several cellular responses, including apoptosis that relies on caspases and necrotic cell death, which depends on receptor-interacting protein kinase 1 (RIPK1; 74, 75). TRADD protein has been suggested to be a crucial signal adaptor that mediates all intracellular responses from TNFR1 (76). Caspase-2 is one of the earliest identified caspases engaged in the mitochondria-dependent apoptotic pathway by inducing the release of cytochrome *c* and other mitochondrial apoptogenic factors into the cell cytoplasm (77). *PPM1F* encodes a protein that is a member of the protein phosphatase 2C family of Ser/Thr protein phosphatases; overexpression of this phosphatase has been shown to mediate caspase-dependent apoptosis (78). Two apoptotic and two antiapoptotic genes are down-regulated in the breast epithelium of parous women: the programmed cell death 5 or *PDCD5* and the transformed 3T3 cell double minute 4 (*MDM4*) in the former, and baculoviral inhibitor of apoptosis protein repeat–containing 6 (*BIRC6*) and BCL2-associated athanogene 4 (*BAG4*) in the latter. The *Mdm4* gene that encodes structurally related oncoproteins that bind to the p53 tumor suppressor protein and inhibit p53 activity is amplified and overexpressed in a variety of human cancers (79). The Split hand/foot malformation (ectrodactyly) type 1 (*SHFM1*) encodes a protein with a BIR (baculoviral) domain and UBCc (ubiquitin-conjugating enzyme E2, catalytic) domain (80). This protein inhibits apoptosis by facilitating the degradation of apoptotic proteins by ubiquitination. BAG4 is a member of the BAG1-related antiapoptotic protein family that functions through interactions with a variety of cell apoptosis– and growth-related proteins including BCL2, Raf protein kinase, steroid hormone receptors, growth factor receptors, and members of the heat shock protein 70 kDa family. This protein was found to be associated with the death domain of TNF receptor type 1 and death receptor 3, and thereby negatively regulates downstream cell death signaling (81). Altogether these clusters of genes seem to maintain the programmed cell death pathway very active in the parous breast epithelium when compared with the epithelium obtained from the breast of parous women with cancer and from nulliparous women with or without cancer. Supporting evidence for this statement comes from data obtained from experimental models (6, 25) and from normal breast tissues of parous women obtained from reduction mammoplasties (6, 26), in which genes involved in the pathway of apoptosis are significantly deregulated. Another cluster of genes that are up-regulated in the parous control group are those related to immunosurveillance. We have previously reported that breast epithelial cells from parous women significantly overexpressed genes related to the immune system (82); therefore, this category will not be further discussed here.

Altogether, our data indicate that the first full-term pregnancy induces in the breast epithelium a specific genomic profile that is still identifiable in parous women at postmenopause. Furthermore, this genomic signature is constituted by genes that cluster differently than those genes expressed in the epithelial cells of parous and nulliparous women with breast cancer as well as from nulliparous women without cancer. This genomic signature allowed us to evaluate the degree of mammary gland differentiation induced by pregnancy. Of importance is the fact that this signature serves for characterizing at molecular level the fully differentiated condition of the breast epithelium that is associated with a reduction in breast cancer risk, thus providing a useful molecular tool for predicting when pregnancy has been protective, for identifying women at risk irrespective of their pregnancy history, and for its use as an intermediate biomarker for evaluating cancer preventive agents.

Acknowledgments

We thank Joanne F. Dorgan, M.P.H., Ph.D. (Genetic Epidemiology, Population Science Division, Fox Chase Cancer Center) for supervising the collection of breast tissues and of patient data; Debra Riordan, M.S., Ryan Hopson, M.S., and Irene Shandruk, M.S. (Genetic Epidemiology, Population Science Division, Fox Chase Cancer Center), Gladwyn Downes (Christiana Care Health System, Newark, DE), and Barbara Carney (Somerset Medical Center, Somerville, NJ) for collecting the clinical data of all the patients entered in this study; Peter Morrison (Department of Biostatistics, Biodiscovery, Inc., San Diego, CA) for his contribution in the analysis of cDNA microarray data; and Daniel A. Mailo, Rebecca C. Heulings, Patricia A. Russo, and Fathima S. Sheriff for their technical and administrative contributions for the successful completion of this work.

References

1. Mustacchi P. Ramazzini and Rigoni-Stern on parity and breast cancer. Clinical impression and statistical corroboration. Arch Intern Med 1961;108:639–42.
2. MacMahon B, Cole P, Lin TM. Age at first birth and breast cancer risk. Bull World Health Organ 1970;43:209–21.
3. Tryggvadottir L, Tulinius H, Eyfjord JE, Sigurvinsson T. Breastfeeding and reduced risk of breast cancer in an Icelandic cohort study. Am J Epidemiol 2001;154:37–42.
4. Jernstrom H, Lubinski J, Lynch HT, et al. Breast-feeding and the risk of breast cancer in *BRCA1* and *BRCA2* mutation carriers. J Natl Cancer Inst 2004;96:1094–8.
5. Russo J, Tay LK, Russo IH. Differentiation of the mammary gland and susceptibility to carcinogenesis: a review. Breast Cancer Res Treat 1982;2:5–73.

6. Russo J, Balogh GA, Chen J, et al. The concept of stem cell in the mammary gland and its implication in morphogenesis, cancer and prevention. Front Biosci 2006;11:151–72.
7. Cutuli B, Borel C, Dhermain F, et al. Breast cancer occurred after treatment for Hodgkin's disease: analysis of 133 cases. Radiother Oncol 2001;59:247–55.
8. Key J, Hodgson S, Omar RZ, et al. Meta-analysis of studies of alcohol and breast cancer with consideration of the methodological issues. Cancer Causes Control 2006;17:759–70.
9. Band PR, Le ND, Fang R, Deschamps M. Carcinogenic and endocrine disrupting effects of cigarette smoke and risk of breast cancer. Lancet 2002;360:1044–9.
10. Russo J, Tait L, Russo IH. Susceptibility of the mammary gland to carcinogenesis. III. The cell of origin of rat mammary carcinoma. Am J Pathol 1983;113:50–66.
11. Russo IH, Russo J. Developmental stage of the rat mammary gland as determinant of its susceptibility to 7,12-dimethylbenz(*a*)anthracene. J Natl Cancer Inst 1978;61:1439–49.
12. Russo IH, Koszalka M, Russo J. Comparative study of the influence of pregnancy and hormonal treatment on mammary carcinogenesis. Br J Cancer 1991;64:481–4.
13. D'Cruz CM, Moody SE, Master SR, et al. Persistent parity-induced changes in growth factors, TGF-β3, and differentiation in the rodent mammary gland. Mol Endocrinol 2002;16:2034–51.
14. Ginger MR, Rosen JM. Pregnancy-induced changes in cell-fate in the mammary gland. Breast Cancer Res 2003;5:192–7.
15. Blakely CM, Stoddard AJ, Belka GK, et al. Hormone-induced protection against mammary tumorigenesis is conserved in multiple rat strains and identifies a core gene expression signature induced by pregnancy. Cancer Res 2006;66:6421–31.
16. Shantakumar S, Terry MB, Teitelbaum SL, et al. Reproductive factors and breast cancer risk among older women. Breast Cancer Res Treat 2007;102:365–74.
17. Iwasaki M, Otani T, Inoue M, et al. Role and impact of menstrual and reproductive factors on breast cancer risk in Japan. Eur J Cancer Prev 2007;16:116–23.
18. Ha M, Mabuchi K, Sigurdson AJ, et al. Smoking cigarettes before first childbirth and risk of breast cancer. Am J Epidemiol 2007;166:55–61.
19. Kote-Jarai Z, Matthews L, Osorio A, et al. Accurate prediction of *BRCA1* and *BRCA2* heterozygous genotype using expression profiling after induced DNA damage. Clin Cancer Res 2006;12:3896–901.
20. Epidemiological Study of BRCA1 and BRCA2 Mutation Carriers (EMBRACE); Gene Etude Prospective Sein Ovaire (GENEPSO); Gen en Omgeving studie van de werkgroep Hereditiair Borstkanker Onderzoek Nederland (GEO-HEBON); International BRCA1/2 Carrier Cohort Study (IBCCS) Collaborators' Group; Andrieu N, Easton DF, Chang-Claude J, et al. Effect of chest X-rays on the risk of breast cancer among *BRCA1/2* mutation carriers in the International *BRCA1/2* Carrier Cohort Study: a report from the EMBRACE, GENEPSO, GEO-HEBON, IBCCS Collaborators' Group. J Clin Oncol 2006;24:3361–6.
21. Russo J, Lynch H, Russo IH. Mammary gland architecture as a determining factor in the susceptibility of the human breast to cancer. Breast J 2001;7:278–91.
22. Russo J, Russo IH. Differentiation and breast cancer development. In: Heppner G, editor. Advances in oncobiology. Vol. 2. New Jersey: JAI Press, Inc.; 1998. p. 1–10.
23. Russo J, Russo IH. Toward a unified concept of mammary tumorigenesis. Prog Clin Biol Res 1997;396:1–16.
24. Russo J, Ao X, Grill C, Russo IH. Pattern of distribution of cells positive for estrogen receptor α and progesterone receptor in relation to proliferating cells in the mammary gland. Breast Cancer Res Treat 1999;53:217–27.
25. Russo J, Mailo D, Hu Y-F, et al. Breast differentiation and its implication in cancer prevention. Clin Cancer Res 2005;11:931–6s.
26. Balogh GA, Heulings R, Mailo DA, et al. Genomic signature induced by pregnancy in the human breast. Int J Oncol 2006;28:399–410.
27. Russo J, Balogh GA, Heulings R, et al. Molecular basis of pregnancy induced breast cancer protection. Eur J Cancer Prev 2006;15:306–42.
28. Balogh GA, Heulings R, Mailo D, et al. Methodological approach to study the genomic profile of the human breast. Int J Oncol 2007;31: 253–60.
29. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004;5:R80.
30. Wettenhall JM, Smyth GK. LimmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics 2004;20: 3705–6.
31. Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, editors. Bioinformatics and computational biology solutions using R and bioconductor. New York: Springer; 2005. p. 397–420.
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc B 1995;57:289–300.
33. Ritchie ME, Diyagama D, Neilson J, et al. Empirical array quality weights for microarray data. BMC Bioinformatics 2006;7:261.
34. Perelman E, Ploner A, Calza S, Pawitan Y. Detecting differential expression in microarray data: comparison of optimal procedures. BMC Bioinformatics 2007;8:28.
35. Maglietta R, Piepoli A, Catalano D, et al. Statistical assessment of functional categories of genes deregulated in pathological conditions by using microarray data. Bioinformatics 2007;23:2063–72.
36. Stachora AA, Schafer RE, Pohlmeier M, Maier G, Ponstingl H. Human Supt5h protein, a putative modulator of chromatin structure, is reversibly phosphorylated in mitosis. FEBS Lett 1997;409:74–8.
37. Kurochkin IV, Yonemitsu N, Funahashi SI, Nomura H. ALEX1, a novel human armadillo repeat protein that is expressed differentially in normal tissues and carcinomas. Biochem Biophys Res Commun 2001;280:340–7.
38. Hsia N, Cornwall GA. DNA microarray analysis of region-specific gene expression in the mouse epididymis. Biol Reprod 2004;70: 448–57.
39. Roldan G, Delgado L, Muse IM. Tumoral expression of BRCA1, estrogen receptor α and ID4 protein in patients with sporadic breast cancer. Cancer Biol Ther 2006;5:505–10.
40. de Candia P, Akram M, Benezra R, Brogi E. Id4 messenger RNA and estrogen receptor expression: inverse correlation in human normal breast epithelium and carcinoma. Hum Pathol 2006;37:1032–41.
41. Yu L, Liu C, Vandeusen J, et al. Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia. Nat Genet 2005;37:265–74.
42. Umetani N, Takeuchi H, Fujimoto A, et al. Epigenetic inactivation of ID4 in colorectal carcinomas correlates with poor differentiation and unfavorable prognosis. Clin Cancer Res 2004;10:7475–83.
43. Chan AS, Tsui WY, Chen X, et al. Down-regulation of ID4 by promoter hypermethylation in gastric adenocarcinoma. Oncogene 2003;22:6946–53.
44. Umetani N, Mori T, Koyanagi K. Aberrant hypermethylation of ID4 gene promoter region increases risk of lymph node metastasis in T1 breast cancer. Oncogene 2005;24:4721–7.
45. Kim JY, Tavaré S, Shibata D. Human hair genealogies and stem cell latency. BMC Biol 2006;4:2.
46. Ito Y, Wiese S, Funk N, et al. Sox10 regulates ciliary neurotrophic factor gene expression in Schwann cells. Proc Natl Acad Sci U S A 2006;103:7871–6.
47. Weiss J, Meeks JJ, Hurley L, Raverot G, Frassetto A, Jameson JL. Sox3 is required for gonadal function, but not sex determination, in males and females. Mol Cell Biol 2003;23:8084–91.
48. Raverot G, Weiss J, Park S, Hurley L, Jameson JL. Sox3 expression in undifferentiated spermatogonia is required for the progression of spermatogenesis. Dev Biol 2005;283:215–25.
49. Dutertre M, Smith CL. Ligand-independent interactions of p160/steroid receptor coactivators and CREB-binding protein (CBP) with estrogen receptor-α: regulation by phosphorylation sites in the A/B region depends on other receptor domains. Mol Endocrinol 2003;17: 1296–314.
50. Iwase H. Molecular action of the estrogen receptor and hormone dependency in breast cancer. Breast Cancer 2003;10:89–96.
51. Watts GS, Oshiro MM, Junk DJ, et al. The acetyltransferase p300/CBP-associated factor is a p53 target gene in breast tumor cells. Neoplasia 2004;6:187–94.
52. Hayashi K, Goodison S, Urquidi V, et al. Differential effects of retinoic acid on the growth of isogenic metastatic and non-metastatic breast cancer cell lines and their association with distinct expression of retinoic acid receptor β isoforms 2 and 4. Int J Oncol 2003;22:623–9.
53. Manohar CF, Salwen HR, Furtado MR, Cohn SL. Up-regulation of HOXC6, HOXD1, and HOXD8 homeobox gene expression in human neuroblastoma cells following chemical induction of differentiation. Tumour Biol 1996;17:34–47.
54. Khwaja A. The role of Janus kinases in haemopoiesis and haematological malignancy. Br J Haematol 2006;134:366–84.
55. Clevenger CV. Roles and regulation of Stat family transcription factors in human breast cancer. Am J Pathol 2004;165:1449–60.
56. Lee H, Rezai-Zadeh N, Seto E. Negative regulation of histone deacetylase 8 activity by cyclic AMP-dependent protein kinase A. Mol Cell Biol 2004;24:765–73.
57. Melnick A, Licht JD. Histone deacetylases as therapeutic targets in hematologic malignancies. Curr Opin Hematol 2002;9:322–32.

58. Jiang X, Russo IH, Russo J. Human chorionic gonadotropin and inhibin induce histone acetylation in human breast cancer cells. Int J Oncol 2002;20:77–9.
59. Noh EJ, Jang ER, Jeong G, Lee YM, Min CK, Lee J-S. Methyl CpG-binding domain protein 3 mediates cancer-selective cytotoxicity by histone deacetylase inhibitors via differential transcriptional reprogramming in lung cancer cells. Cancer Res 2005;65: 11400–10.
60. Huang W-Y, Chatterjee N, Chanock S, et al. Microsomal epoxide hydrolase polymorphisms and risk for advanced colorectal adenoma. Cancer Epidemiol Biomarkers Prev 2005;14:152–7.
61. Yoo MH, Xu XM, Carlson BA, Gladyshev VN, Hatfield DL. Thioredoxin reductase 1 deficiency reverses tumor phenotype and tumorigenicity of lung carcinoma cells. J Biol Chem 2006;281:13005–8.
62. Zheng W, Wen WQ, Gustafson DR, Gross M, Cerhan JR, Folsom AR. GSTM1 and GSTT1 polymorphisms and postmenopausal breast cancer risk. Breast Cancer Res Treat 2002;74:9–16.
63. Mucci LA, Wedren S, Tamimi RM, Trichopoulos D, Adami HO. The role of gene-environment interaction in the aetiology of human cancer: examples from cancers of the large bowel, lung and breast. J Int Med 2001;249:477–93.
64. Russo J, Calaf G, Sohi N, et al. Critical steps in breast carcinogenesis. Ann N Y Acad Sci 1993;698:1–20.
65. Tay LK, Russo J. Formation and removal of 7,12-dimethylbenz(*a*)anthracene-nucleic acid adducts in rat mammary epithelial cells with different susceptibility to carcinogenesis. Carcinog 1981;2:1327–33.
66. Kennedy DO, Agrawal M, Shen J, et al. DNA repair capacity of lymphoblastoid cell lines from sisters discordant for breast cancer. J Natl Cancer Inst 2005;97:127–32.
67. Wiese C, Hinz JM, Tebbs RS, et al. Disparate requirements for the Walker A and B ATPase motifs of human RAD51D in homologous recombination. Nucleic Acid Res 2006;34:2833–43.
68. Laine JP, Egly JM. When transcription and repair meet: a complex system. Trends Genet 2006;22:430–6.
69. Mellon SH, Bair SR, Depoix C, Vigne JL, Hecht NB, Brake P. Translin coactivates steroidogenic factor-1-stimulated transcription. Mol Endocrinol 2007;21:89–105.
70. Bomgarden RD, Yean D, Yee MC, Cimprich KA. A novel protein activity mediates DNA binding of an ATR-ATRIP complex. J Biol Chem 2004;279:13346–53.
71. Zou L, Elledge SJ. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science 2003;300:1542–8.
72. Eissing T, Waldherr S, Allgower F, Scheurich P, Bullinger E. Response to bistability in apoptosis: roles of bax, bcl-2, and mitochondrial permeability transition pores. Biophys J 2007;92:3332–4.
73. Mori N, Murakami YI, Shimada S, et al. TIA-1 expression in hairy cell leukemia. Mod Pathol 2004;17:840–6.
74. Xie P, Hostager BS, Munroe ME, Moore CR, Bishop GA. Cooperation between TNF receptor-associated factors 1 and 2 in CD40 signaling. J Immunol 2006;176:5388–400.
75. Bryce PJ, Oyoshi MK, Kawamoto S, Oettgen HC, Tsitsikov EN. TRAF1 regulates Th2 differentiation, allergic inflammation and nuclear localization of the Th2 transcription factor, NIP45. Int Immunol 2006;18:101–11.
76. Zheng L, Bidere N, Staudt D, et al. Competitive control of independent programs of tumor necrosis factor receptor-induced cell death by TRADD and RIP1. Mol Cell Biol 2006;26:3505–13.
77. Guo Y, Srinivasula SM, Druilhe A, Fernandes-Alnemri T, Alnemri ES. Caspase-2 induces apoptosis by releasing proapoptotic proteins from mitochondria. J Biol Chem 2002;277:13430–7.
78. Mi J, Guo C, Brautigan DL, Larner JM. Protein phosphatase-1α regulates centrosome splitting through Nek2. Cancer Res 2007;67: 1082–9.
79. Francoz S, Froment P, Bogaerts S, et al. Mdm4 and Mdm2 cooperate to inhibit p53 activity in proliferating and quiescent cells *in vivo*. Proc Natl Acad Sci U S A 2006;103:3232–7.
80. Colnaghi R, Connell CM, Barrett RM, Wheatley SP. Separating the anti-apoptotic and mitotic roles of survivin. J Biol Chem 2006;281: 33450–6.
81. Eichholtz-Wirth H, Fritz E, Wolz L. Overexpression of the ''silencer of death domain,'' SODD/BAG-4, modulates both TNFR1- and CD95-dependent cell death pathways. Cancer Lett 2003;194:81–9.
82. Balogh GA, Russo IH, Spittle C, Heulings R, Russo J. Immune surveillance and programmed cell death related genes are significantly overexpressed in the normal breast epithelium of postmenopausal parous women. Int J Oncol 2007;31:303–12.