Molecular Profiling of Laser-Microdissected Matched Tumor and Normal Breast Tissue Identifies Karyopherin α2 as a Potential Novel Prognostic Marker in Breast Cancer

Edgar Dahl,[1] Glen Kristiansen,[2] Kathrin Gottlob,[3] Irina Klaman,[3] Elke Ebner,[4] Bernd Hinzmann,[6] Klaus Hermann,[6] Christian Pilarsky,[7] Matthias Dürst,[8] Monika Klinkhammer-Schalke,[5] Hagen Blaszyk,[9] Ruth Knuechel,[1] Arndt Hartmann,[4] André Rosenthal,[6] and Peter J. Wild[10]

Abstract

Purpose: The aim of the present study was to identify human genes that might prove useful in the diagnosis and therapy of primary breast cancer.

Experimental Design: Twenty-four matched pairs of invasive ductal breast cancer and corresponding benign breast tissue were investigated by a combination of laser microdissection and gene expression profiling. Differential expression of candidate genes was validated by dot blot analysis of cDNA in 50 pairs of matching benign and malignant breast tissue. Cellular expression of candidate genes was further validated by RNA *in situ* hybridization, quantitative reverse transcription-PCR, and immunohistochemistry using tissue microarray analysis of 272 nonselected breast cancers. Multivariate analysis of factors on overall survival and recurrence-free survival was done.

Results: Fifty-four genes were found to be up-regulated and 78 genes were found to be down-regulated. Dot blot analysis reduced the number of up-regulated genes to 15 candidate genes that showed at least a 2-fold overexpression in >15 of 50 (30%) tumor/normal pairs. We selected phosphatidic acid phosphatase type 2 domain containing 1A (*PPAPDC1A*) and karyopherin α2 (*KPNA2*) for further validation. PPAPDC1A and KPNA2 RNA was up-regulated (fold change >2) in 84% and 32% of analyzed tumor/normal pairs, respectively. Nuclear protein expression of KPNA2 was significantly associated with shorter overall survival and recurrence-free survival. Testing various multivariate Cox regression models, KPNA2 expression remained a highly significant, independent and adverse risk factor for overall survival.

Conclusions: Gene expression profiling of laser-microdissected breast cancer tissue revealed novel genes that may represent potential molecular targets for breast cancer therapy and prediction of outcome.

Mammary carcinogenesis is facilitated by a sequential accumulation of genetic alterations, including the inactivation of tumor suppressor genes and the activation of oncogenes (1). This complex process is poorly understood and no disease-specific cytogenetic marker has been identified in breast cancer thus far. Loss of heterozygosity analyses revealed that all chromosome arms could be affected to varying degrees. Genetic amplifications have been found in various chromosomal regions that harbor established oncogenes such as *FGFR1*, *FGFR2*, *MYC*, *CCND1*, and *STK15* (reviewed in ref. 2). The large number of tumor-related genes underscores the complexity of breast cancer with variable effect on carcinogenesis.

Expression profiling by DNA microarray technology has extensively been applied to identify breast cancer–associated genes (3–7). Many of such studies, however, are limited by exclusive expression analysis of bulk tissue containing a mixture of tumor, stromal, endothelial, and inflammatory cells.

The aim of this study was to identify potential diagnostic marker genes and molecular targets for breast cancer therapy using a combination of oligonucleotide arrays, multi-tissue Northern blots, and tissue microarray technologies in laser-microdissected pure tissue samples.

Materials and Methods

Patients and samples. Twenty-four patients with primary invasive ductal carcinoma of the breast (stage I-IIIC) were studied using a combination of laser microdissection and gene expression profiling. Tissue samples of the primary tumor and matching normal tissue were

Authors' Affiliations: [1]Institute of Pathology, University Hospital of the RWTH Aachen, Aachen, Germany; [2]Institute of Pathology, Charité University Hospital; [3]metaGen Pharmaceuticals i.L., Berlin, Germany; [4]Institute of Pathology, University of Regensburg; [5]Central Tumor Registry, Regensburg, Germany; [6]Signature Diagnostics AG, Potsdam, Germany; [7]University Hospital Carl Gustav Carus, Department of Surgery, Dresden, Germany; [8]Department of Gynecology, Friedrich Schiller University, Jena, Germany; [9]Department of Pathology, University of Vermont College of Medicine, Burlington, Vermont; and [10]Institute of Pathology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Received 11/1/05; revised 3/23/06; accepted 4/27/06.

Grant support: German Ministry for Education and Research, BMBF grant 01KW0401, as part of the German Human Genome Project (E. Dahl).

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: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

P.J. Wild and A. Rosenthal share the senior authorship for this work.

Requests for reprints: Edgar Dahl, Institute of Pathology, University Hospital of the RWTH Aachen, Pauwelsstrasse 30, 52047 Aachen, Germany. Phone: 49-241-8088431; Fax: 49-241-8082439; E-mail: edahl@ukaachen.de.

doi:10.1158/1078-0432.CCR-05-2090

obtained between 1994 and 2002 from three different pathology laboratories (Charité University Hospital, Berlin, Germany; University of Regensburg, Regensburg, Germany; and Friedrich-Schiller University of Jena, Jena, Germany). Histopathologic data are summarized in Table 1. Informed consent was obtained from all patients and the Institutional Review Board of all participating centers approved the study. Specimens were immediately transferred from the operating room to the pathology laboratory. Samples were stored at −80°C after serial sectioning of tissue for routine pathology examination. Tumor type and stage were assigned according to UICC and WHO criteria. Tumors were graded according to Elston and Ellis (8).

Laser microdissection, RNA isolation, and chip hybridization. Pure populations of >90% benign or neoplastic breast cancer cells were obtained using the SL-Microtest UV laser microdissection system (Molecular Machines & Industries, Glattbrugg, Switzerland) or the PALM laser microdissection device (Wolfratshausen, Germany) as previously described (9). Total RNA was isolated from twenty 5-μm frozen sections using sterilized (3 hours, 250°C) glass slides and racks. To prevent RNase contamination, frozen slides (−80°C) were directly transferred into 100% ethanol (1 minute) and subsequently stained with 1% methylene blue, pretreated with 0.1% diethylpyrocarbonate.

Poly(A)$^+$ RNA was isolated from lysed tissue cells by magnetic separation (PolyATract System 1000; Promega, Heidelberg, Germany) according to the specifications of the manufacturer. cDNA synthesis was followed by three cycles of *in vitro* transcription as previously described (10). In brief, RNA was primed with the Affymetrix T7-oligo-dT promoter-primer combination (5′-GGCCAGTGAATTGTAATACGACT-CACTATAGGGAGGCGGT$_{24}$-3′). TaqMan PCR tested the cDNA of each round of amplification for its integrity and cDNAs of low quality were excluded from further processing. Biotinylated nucleotides were incorporated into the aRNA during the last *in vitro* transcription. Hybridization and detection of labeled aRNA used the metg001A Affymetrix GeneChip array according to the instructions of the manufacturer.

Chip design and bioinformatics analysis. The custom-designed Affymetrix oligonucleotide array (Gene Chip metg001A) consists of 6,117 probe sets representing 3,950 cDNAs based on the annotation of the probe sets with the GoldenPath assembly[11] and has previously been described (11). In short, we selected genes from signal transduction pathways known to be involved in the progression of human tumors (e.g., transforming growth factor β, RAS, and WNT pathways) and cDNA fragments derived from systematic screening of EST libraries for genes that are differentially expressed in normal and tumor tissues (12).

Preprocessing and chip-level corrections of the Affymetrix metg001A gene chip data has previously been described (13). In brief, Gene Chips were scanned using an Agilent GeneArray Scanner (Agilent Technologies, Palo Alto, CA). Raw intensity values were extracted from the CEL-files. Using the 75% percentile of the perfect match intensities for each probe set (PMQ value) generated a representative expression value. For each probe set, a nonparametric Wilcoxon test was calculated by comparing the intensities of the perfect match and mismatch probes to test the probe sets for presence or absence of an expression signal. To minimize data perturbation caused by experimental variation, a model-fitting algorithm was applied to the PMQ data creating a theoretical reference chip. Expression data of each individual chip were compared with the reference chip. For further analysis, data were transformed into log space (ln).

Probe sets were ranked according to three different variables as described by Kristiansen et al. (11): first, according to the ratio of the median of expression values in tumor samples divided by the median of expression values in normal samples (change fold median); second, according to the percentage of patients who had a tumor/normal fold change exceeding 2; and third, according to the Golub criteria (14). The cross section of the top 5% of genes for each method was selected for further analysis.

Table 1. Characteristics of 24 matched tumor and normal tissue pairs

Variable	Categorization	*n* analyzable	%
Clinicopathologic data			
Tumor stage			
	pT_1	7	29.2
	pT_2	16	66.6
	pT_3	1	4.2
Lymph node status			
	pN_0	12	50.0
	pN_1	8	33.3
	pN_2	4	16.7
Histologic grade			
	G1	1	4.2
	G2	11	45.8
	G3	12	50.0

Breast cancer profiling array. The breast cancer profiling array (BD Clontech, Heidelberg, Germany) contains 50 pairs of cDNAs generated from matching tumor and normal tissue samples of individual patients, spotted on a nylon membrane (http://www.clontech.com/techinfo/manuals/pdf/pt3424-1.pdf). Hybridizations using 25 ng of a gene-specific ^{32}P-labeled cDNA probe digested from Unigene cDNA clones were done according to the recommendations of the manufacturer. The tumor/normal intensity ratio was calculated using a STORM-860 phosphoimager (Molecular Dynamics, Eugene, OR) and normalized against the background. We defined a candidate gene as up-regulated in breast cancer by this technique if a >2-fold up-regulation was detectable in at least 30% of analyzed tumor/normal pairs (Table 2).

Quantitative reverse-transcription PCR. For TaqMan assays (Applied Biosystems, Weiterstadt, Germany), cDNA reverse transcribed from 1 ng of amplified RNA was used. Genes were amplified with the TaqMan Universal PCR Master Mix according to the conditions of the manufacturer using the ABI PRISM 5700 Sequence Detection System. Gene expression was quantified by the comparative C_T Method, normalizing C_T values to glyceraldehyde-3-phosphate dehydrogenase and calculating the relative expression values of tumor and normal tissue (15). Oligonucleotide primers and hybridization probes are listed in Supplementary Figure 1.

RNA* in situ *hybridization. IMAGE consortium plasmid clones were linearized with *Hin*dIII and *Eco*RI for sense and antisense probes, respectively. A 500-bp cDNA fragment (GenBank accession no. N25267) was used for PPAPDC1A. Riboprobes were generated and digoxigenin labeled using the Dig RNA labeling kit (Roche Applied Science, Mannheim, Germany). Sections of paraffin-embedded breast tissues were deparaffinized, rehydrated, and processed according to the instructions of the manufacturer (Roche Applied Science). Hybridized probes were detected using alkaline phosphatase–conjugated anti-DIG antibodies and BM Purple as substrate (Roche Applied Science).

Immunohistochemistry. Immunohistochemical studies for the expression of TP53 and HER2 used an avidin-biotin peroxidase method with a 3,3′-diaminobenzidine chromatogen. For karyopherin α2 (KPNA2), the ChemMate detection kit (DAKO, Glostrup, Denmark) was used. After antigen retrieval [microwave oven for 30 minutes (TP53, HER2) or 10 minutes (KPNA2) at 250 W], immunohistochemistry was carried out in a NEXES immunostainer (Ventana, Tucson, AZ) following the instructions of the manufacturer. The following primary antibodies were used: anti-TP53 (mouse monoclonal Bp53-12,

[11] http://genome.ucsc.edu.

Santa Cruz Biotechnology, Inc., Santa Cruz, CA; dilution 1:1,000), anti-KPNA2 (goat polyclonal SC6917, Santa Cruz Biotechnology; dilution 1:200), and anti-HER2 (rabbit polyclonal A0485, DAKO; dilution 1:400). Normal testicular parenchyma was chosen as internal positive control for KPNA2 immunohistochemistry. One surgical pathologist (A.H.) did a blinded evaluation of the slides without knowledge of clinical data. Causes of noninterpretable results included lack of tumor tissue and presence of necrosis or crush artifact. TP53 and KNPA2 positivity was defined as strong nuclear staining in at least 10% cells. HER2 expression was scored according to the DAKO HercepTest (16).

Breast cancer tissue microarray. A tissue microarray was constructed as previously described (17) and contained 289 nonselected formalin-fixed, paraffin-embedded primary breast cancers (stage I -IIIC) together with matched normal breast tissue of each patient. An experienced surgical pathologist (A.H.) evaluated H&E-stained slides of all specimens before construction of the tissue microarray to identify representative tumor areas. Clinical follow-up, provided by the Central Tumor Registry Regensburg, Germany was available for 272 breast cancer patients with a median follow-up period of 79 months (0-148 months). Clinicopathologic variables of breast cancer cases included in the tissue microarray are summarized in Table 3.

Statistical analysis of tissue microarray data. Statistical analyses were completed with SPSS version 10.0 (SPSS, Chicago, IL). Differences were considered statistically significant when $P < 0.05$. Contingency table analysis and two-sided Fisher's exact tests were used to study the statistical association between clinicopathologic and immunohistochemical variables. The Wilcoxon test (two-sided) for dependent variables was used to compare tumors and matched normal samples on the tissue microarray. Recurrence-free and overall survival curves comparing patients with or without any of the factors were calculated using the Kaplan-Meier method, with significance evaluated by two-sided log-rank statistics. Overall survival and recurrence-free survival were measured from time of surgery. For the analysis of recurrence, patients were censored at the time of their last tumor-free clinical follow-up appointment. For overall survival analysis, patients were censored at the time of their last tumor-free clinical follow-up appointment or at their date of death not related to the tumor. A stepwise multivariable Cox regression model was adjusted, testing the independent prognostic relevance of nuclear KPNA2 immunoreactivity. The limit for reverse selection procedures was $P = 0.01$. The proportionality assumption for all variables was assessed with log-negative-log survival distribution functions.

Table 2. Genes significantly up-regulated in primary breast cancer compared with matched normal breast tissue

No.	Gene name, description	HUGO name	Gene annotation*	Map NCBI	Chip (%),† *n* = 24	BCPA (%),†‡ *n* = 50	RT-PCR (%),† *n* = 6
1	Phosphatidic acid phosphatase type 2 domain containing 1A	**PPAPDC1A**§	Lipid phosphatase	10q26.13	92	84	100
2	Osteopontin	SPP1	Cell adhesion, cell structure	4q21-q25	83	52	83
3	Normal mucosa of esophagus specific 1	**NMES1**	Electron transport	15q21.1	75	49	NA
4	Sulfatase 1	**SULF1**	Esterase	8q13	67	52	100
5	Karyopherin α2 (RAG cohort 1, importin α1)	**KPNA2**	Nuclear transport, protein targeting	17q23	67	32	NA
6	Calgranulin B	S100A9	Cell communication, immunity, and defense	1q12-q22	64	46	50
7	G protein-coupled receptor (RAI3/RAIG1)	**GPCR5A**	G-protein coupled receptor	12p13-p12.3	64	64	67
8	CDC28 protein kinase regulatory subunit 2	**CKS2**	Cell cycle control	9q22	63	54	NA
9	Peroxiredoxin 4	PRDX4	Antioxidation and free radical removal	Xp22.11	58	42	17
10	Matrix metalloproteinase 3/ stromelysin 1	MMP3	Metalloprotease, proteolysis	11q22.3	52	NA	NA
11	IFN-α inducible protein 27	IFI27	Immunity and defense	14q32	46	42	50
12	Aurora kinase A (STK15)	AURKA	Intracellular signaling, chromosome segregation	20q13	43	38	NA
13	AE binding protein 1	**AEBP1**	Metalloprotease, mRNA transcription regulation	7p13	35	56	NA
14	Periostin, osteoblast specific factor (OSF2)	POSTN	Cell adhesion, skeletal development	13q13.3	32	43	83
15	Trophoblast glycoprotein (5T4 antigen)	TPBG	Not classified	6q14-q15	30	50	50

Abbreviations: RT-PCR, reverse transcription-PCR; NA, data not available.
*Molecular function or biological process according to the PANTHER classification system (https://panther.appliedbiosystems.com).
†Percentage of matched tumor and normal tissue pairs with change folds >2.
‡BCPA, breast cancer profiling array (BD Clontech).
§Boldface represents gene previously not found to be differentially expressed using bulk tissue.

Table 3. Clinicopathologic and immunohistochemical variables in relation to nuclear KPNA2 immunoreactivity, and univariate analysis of factors on tumor-specific survival and recurrence-free survival

Variable	Categorization	KPNA2 immunoreactivity				Tumor-related death			Tumor recurrence		
		n analyzable	<10%	≥10%	*P**	*n*	Events	*P*†	*n*	Events	*P*†
Clinicopathologic data											
Age at diagnosis											
	<50 y	55	19	36	0.109	91	17	**0.0188**	91	30	0.3088
	≥50 y	136	65	71		181	61		172	65	
Tumor stage											
	pT_1	54	34	20	**0.011**	99	14	**<0.0001**	96	19	**<0.0001**
	pT_2	97	35	62		123	37		121	48	
	pT_3	15	6	9		18	8		17	9	
	pT_4	25	9	16		32	19		29	19	
Lymph node status											
	pN_0	86	44	42	0.101	133	16	**<0.0001**	131	21	**<0.0001**
	pN_{1-3}	97	37	60		129	54		125	67	
Histologic grade											
	G1	21	17	4	**<0.001**	39	6	**<0.0001**	38	8	**<0.0001**
	G2	89	49	40		129	27		123	32	
	G3	80	17	63		100	44		98	53	
Multifocality											
	Unifocal tumor	165	73	92	1,000	233	62	0.984	226	78	0.2275
	Multifocal tumor	26	11	15		38	15		36	16	
Histologic type											
	Ductal	150	62	88	0.541	218	62	0.6302	213	79	0.4282
	Lobular	17	9	8		25	9		23	8	
	Other	20	10	10		23	5		21	5	
Immunohistochemistry											
Estrogen receptor status											
	Negative	51	15	36	0.083	65	26	0.0069	65	31	0.0193
	Positive	109	49	60		139	32		134	40	
Progesterone receptor status											
	Negative	115	36	79	**<0.001**	141	55	**0.0001**	133	60	**0.0006**
	Positive	58	36	22		76	10		76	16	
TP53 immunohistochemistry											
	<10%	102	50	52	**0.010**	135	29	0.0526	129	43	0.0642
	≥10%	53	14	39		60	24		58	27	
HER2 immunohistochemistry											
	Negative (0-1+)	127	61	66	**0.002**	170	40	**0.0003**	161	52	**0.0001**
	Positive (2+-3+)	40	8	32		45	23		45	26	

*Fisher's exact test (two-sided); boldface represents significant data.
†Log-rank test (two-sided); boldface represents significant data.

Results

Identification of up-regulated and down-regulated novel putative marker genes in breast cancer. We analyzed the expression profiles of 24 invasive ductal breast carcinomas and matching normal breast tissues. Gene expression profiles were examined using custom-designed Affymetrix oligonucleotide microarrays. All specimens were laser-microdissected to obtain pure populations of at least 2,000 benign and malignant cells. RNA integrity was tested during the first two rounds of linear amplifications.

Up-regulated and down-regulated genes were ranked according to our recently described stringent bioinformatics method (11), which requires that a candidate gene is selected among the top 5% of genes in any of three methods that are regularly used for gene expression analysis. Specifically, the Golub approach, the fold change 2 approach, and the fold change median approach identified 54 up-regulated and 78 down-regulated genes in our samples (Supplementary Figures 2 and 3). Differential expression of all up-regulated candidate genes and two of the down-regulated candidate genes was validated by dot blot analysis on a nylon array containing spotted cDNAs derived from 50 matching pairs of normal and tumor breast tissue (Fig. 1). The specificity of each gene probe was controlled on a multi-tissue Northern blot (BD Clontech) containing poly(A)$^+$ RNA samples from 16 human tissues before hybridization to

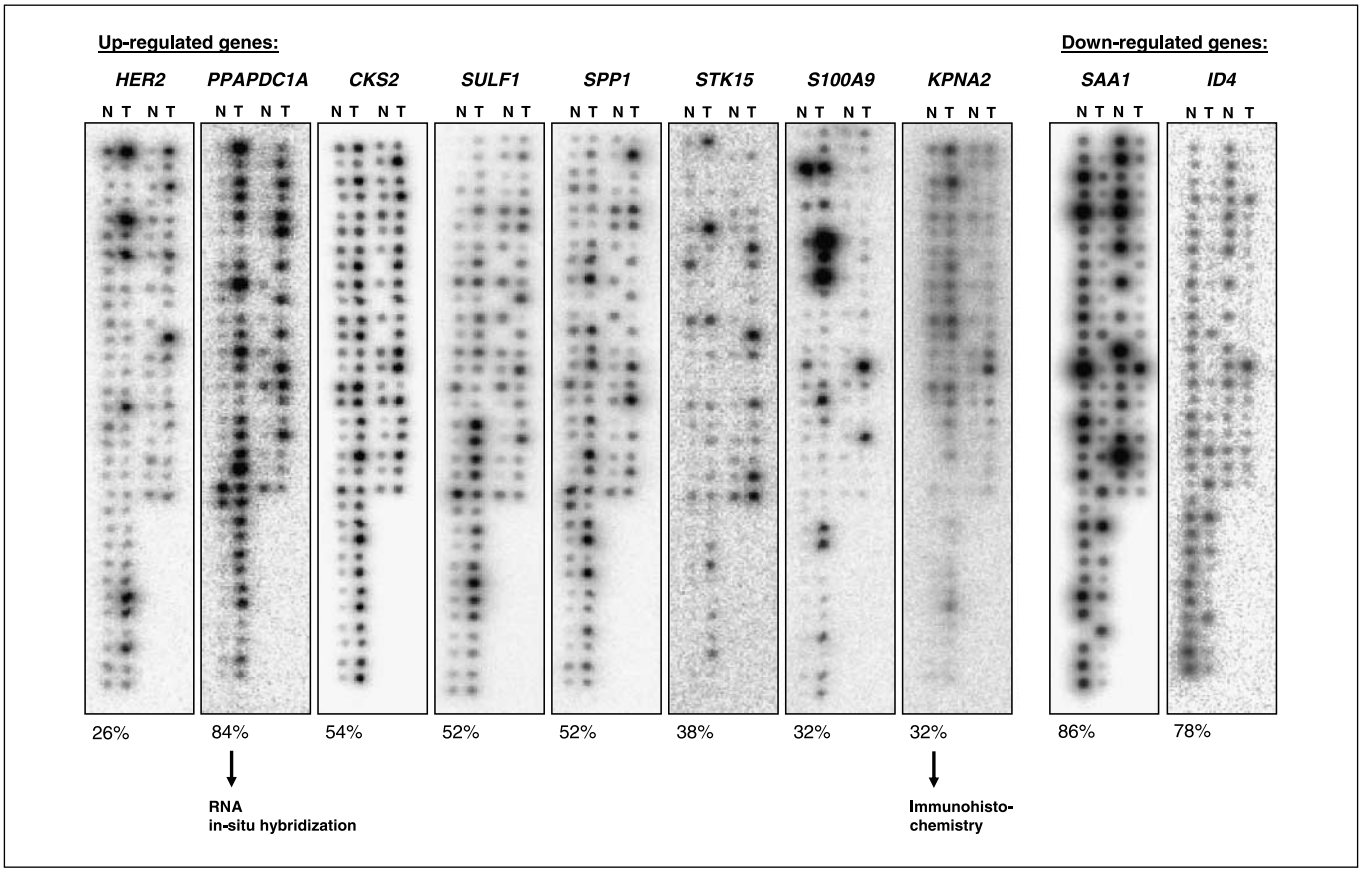

Fig. 1. Representative examples of up-regulated and down-regulated genes in breast cancer. Expression profiles were determined using breast cancer profiling arrays (BD Clontech) containing cDNA pairs derived from 50 matched tumor/normal samples. N, normal; T, tumor. Percentage of up-regulated and down-regulated tumor/normal pairs is given at the bottom.

the breast cancer profiling array (Fig. 2A-C). Candidate genes needed to show an at least 2-fold up-regulation in >30% of analyzed normal/tumor pairs to be at least as differentially expressed as HER2 (Fig. 1). Table 2 summarizes the 15 up-regulated genes fulfilling these criteria and compares their grade of differential expression in DNA array analysis (24 matched pairs), dot blot analysis (50 matched pairs), and reverse transcription-PCR analysis (6 matched pairs). The molecular function of these 15 highly overexpressed genes was annotated according to the Panther classification system (Applied Biosystems). We found that these genes are involved in important biological processes, such as cell cycle control, intracellular signaling, and cell adhesion (Table 2). Our 15 genes reached overexpression values ranging from 32% to 84%

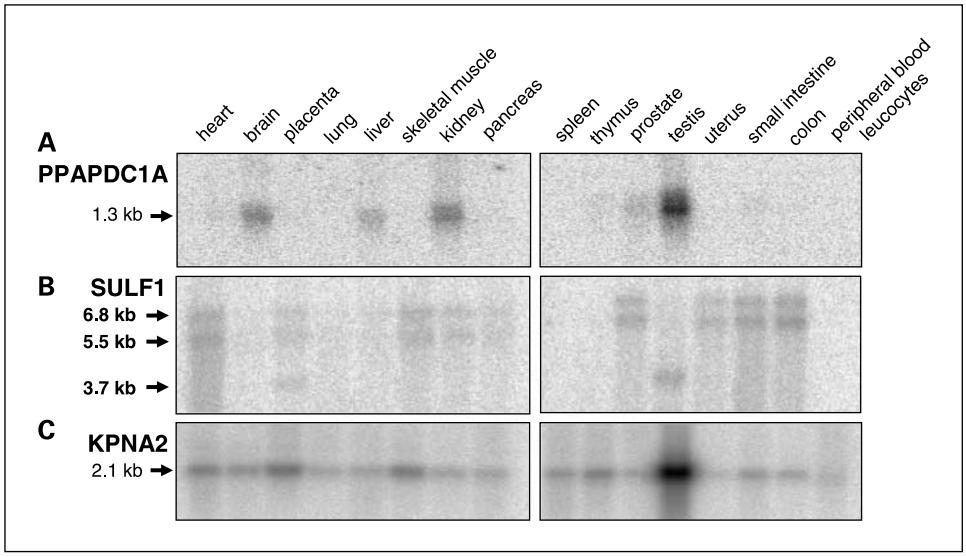

Fig. 2. Expression profiles for PPAPDC1A (*A*), SULF1 (*B*), and KPNA2 (*C*) using multi-tissue Northern blots (BD Clontech) containing cDNA derived from various organs.

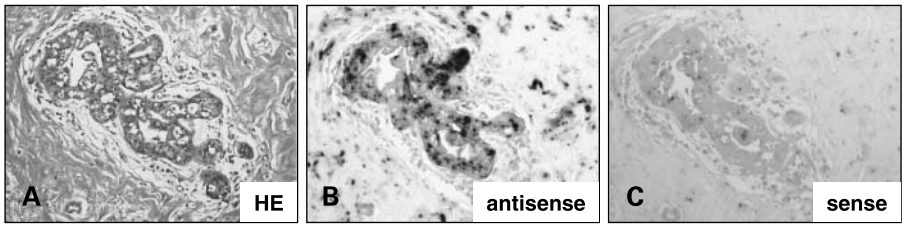

Fig. 3. RNA *in situ* hybridization of PPAPDC1A. Representative area of breast cancer (original magnification, ×100). *A,* H&E; *B,* strongly positive staining in the tumor cells; *C,* negative control.

whereas HER2 was overexpressed in only 26% of analyzed tissue pairs (Fig. 1). Seven of the 15 genes have never been associated with breast tumorigenesis before.

PPAPDC1A, SULF1, and KPNA2 are novel genes strongly overexpressed in breast cancer. Three of the seven novel breast cancer–associated genes are of special interest because of their presumed biological functions and will be presented in more detail. Phosphatidic acid phosphatase type 2 domain containing 1A (*PPAPDC1A*) encodes an integral membrane phosphatase of 271 amino acids that was up-regulated in at least 84% of tissue pairs as analyzed by DNA microarray, cDNA dot blot analysis, and reverse transcription-PCR techniques (Table 2; Fig. 1). According to BLAST analysis, PPAPDC1A represents a novel member of the family of lipid phosphate phosphatases also known as phosphatic acid phosphatases (18, 19). These lipid phosphatases catalyze the conversion of phosphatidic acid to diacylglycerol and anorganic phosphate. Because diacylglycerol is an important molecule in cell signaling, phosphatidic acid phosphatases may play an important role in signal transduction (20) and cancer development. PPAPDC1A was strongly expressed in human testis (Fig. 2A) and less abundant in kidney, brain, and liver. We analyzed PPAPDC1A RNA expression in 34 matched tissue samples using nonradioactive RNA *in situ* hybridization and found a significant up-regulation in tumor tissue in 15 of the 34 samples ($P = 0.001$; Fig. 3A-C).

The human sulfatase 1 gene (*SULF1*) is the human orthologue of the recently characterized quail *SULF1* gene (21). The human SULF1 mRNA exhibited two major transcripts of 6.8 and 5.5 kb and was weakly expressed in most human tissues (Fig. 2B). Expression was most abundant in colon, small intestine, prostate, and heart. A smaller transcript of 3.7 kb was present in placenta and represented the predominant SULF1 mRNA in testis (Fig. 2B). SULF1 was found to be overexpressed in 52% of matched tissue pairs according to breast cancer profiling array data (Fig. 1) and in 67% according to our DNA array data (Table 2). Expression data provided by Su et al. (22) also indicated strong SULF1 expression in breast, ovarian, lung, and pancreatic cancers. We also found very high SULF1 expression in the breast cancer cell line HS578T (data not shown). In contrast to the potential tumor-promoting function of SULF1 in breast cancer, Lai et al. (23) have recently described a cell growth–suppressing role of SULF1 in head and neck squamous cell carcinoma cell lines, possibly through inhibition of the mitogen-activated protein kinase signaling pathway.

KPNA2 is a nuclear and cytoplasmic protein also known as importin α1 or Rch1. KPNA2 is thought to connect karyophilic proteins to the nuclear import machinery (24). We found uniform and moderate expression of KPNA2 mRNA in most human normal tissues (Fig. 2C) and prominent expression in testis. The human KPNA2 transcript was determined to be 2.1 kb in size. KPNA overexpression in matching tissues was 67% in the DNA array experiment and 32% on the breast cancer profiling array (Table 2). Because human normal breast tissue was not present on the multi-tissue Northern blot from BD Clontech, we further analyzed KPNA2 expression by real-time PCR in eight microdissected breast normal tissues in comparison with other human normal tissues also present on the multi-tissue Northern blot (Fig. 4). This analysis showed variable KPNA2 expression in human normal tissue and a mean expression level that was comparable to that found in thymus.

Validation of KPNA2 expression by immunohistochemistry. KPNA2 was selected from the 15-gene set for further validation on the protein level because a primary antibody was commercially available. Investigation of KPNA2 protein expression in a large series of breast cancers by tissue microarray technology was informative in 191 of 272 (70.2%) of tumor specimens and in 171 of 272 (62.9%) of normal samples. KPNA2 expression in at least 10% of nuclei was detected in 107 of 191 (56.0%) of breast cancers whereas only 3 of 171 (1.8%) of matched normal tissues were KPNA2 positive ($P < 0.001$). Representative KPNA2 immunostaining patterns in malignant and normal breast tissue are shown in Fig. 5A-H.

Clinicopathologic and immunohistochemical characteristics were associated with KPNA2 immunohistochemistry for descriptive data analysis (Table 3). Nuclear KPNA2 staining in at least 10% of tumor cells was significantly associated with higher tumor stage ($P < 0.011$) and higher tumor grade ($P < 0.001$), but not with age at diagnosis, lymph node status,

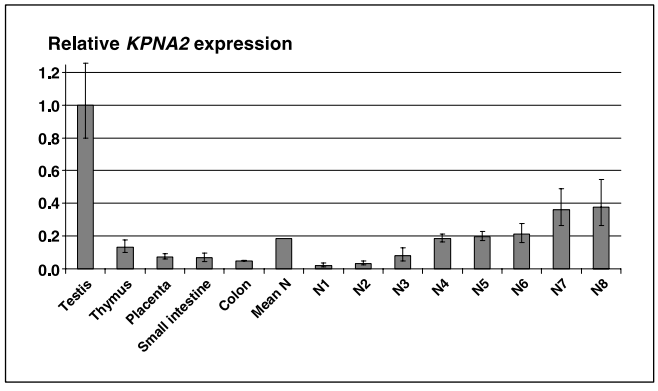

Fig. 4. Real-time PCR analysis of *KPNA2* expression in normal breast tissue (*N1-N8*) in comparison with other normal human tissues (normalization to testis). As found by the multi-tissue Northern blot (Fig. 2C), *KPNA2* is most abundantly expressed in normal testis. *KPNA2* expression in normal breast tissue is similar to that found in normal thymus, reaching ~20% of the expression found in testis. Bars, SD.

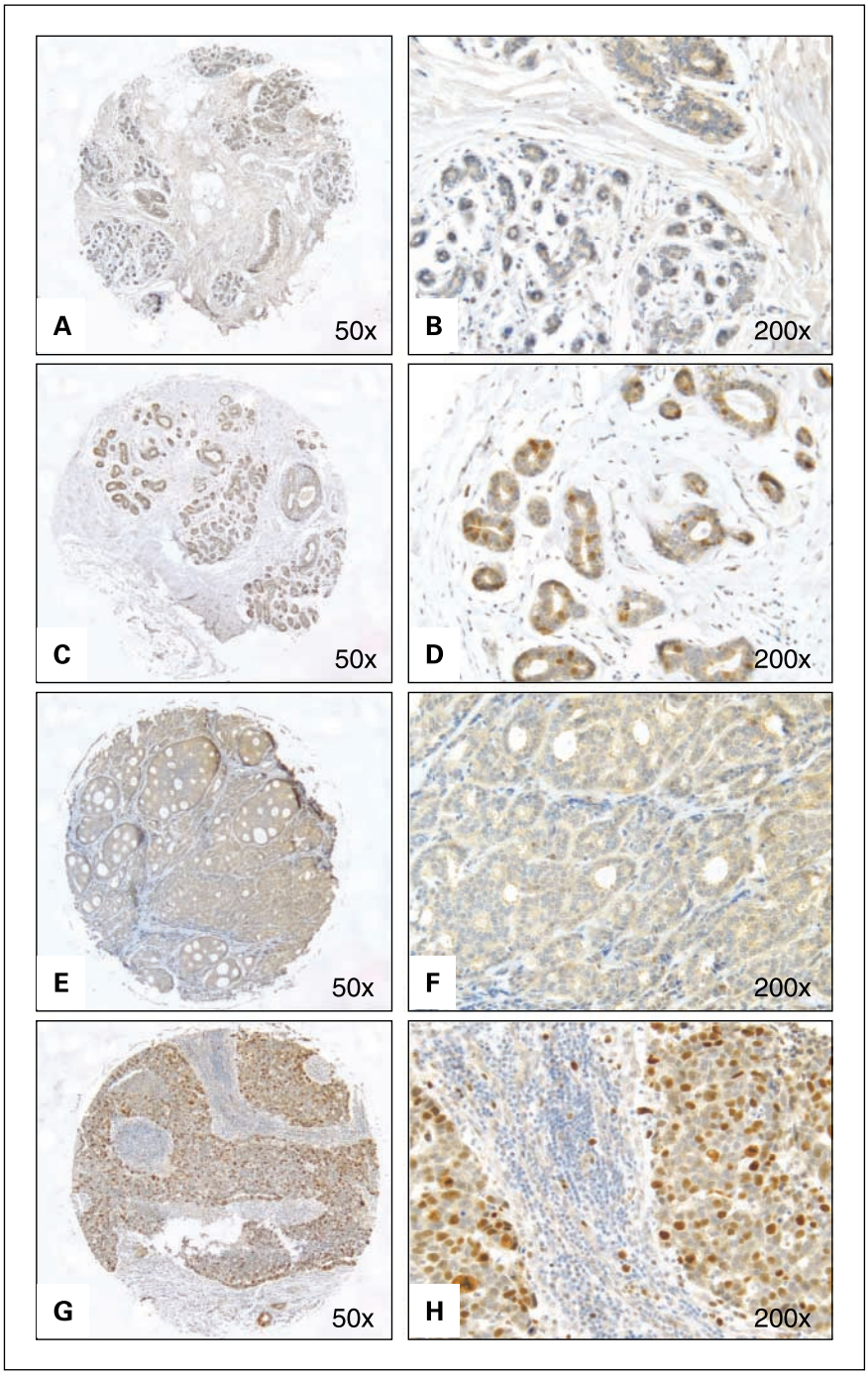

Fig. 5. Nuclear KPNA2 staining of normal breast tissue and invasive breast cancer on the tissue microarray. Original magnification is given in the bottom right corner. *A* to *D*, representative examples of normal breast tissue with negative (*A* and *B*) and positive (*C* and *D*) immunoreactivity of few cells for KPNA2. *E* to *H*, representative sections of invasive breast cancer showing negative (*E* and *F*) and strongly positive (*G* and *H*) KPNA2 immunohistochemistry.

multifocality, estrogen receptor status, and histologic tumor type. When analyzing tumors of all stages, negative progesterone receptor status ($P < 0.001$), TP53 positivity ≥10% ($P = 0.010$), and positive HER2 status ($P = 0.002$) were significantly associated with KPNA2 expression.

Overall survival and recurrence-free survival were compared between KPNA2-negative and KPNA2-positive cases by univariate log-rank statistics (Table 3). Patients with KPNA2-positive tumors (≥10%) had an estimated mean overall survival time of 101 months (95% confidence interval, 90-112) compared with 120 months (95% confidence interval, 110-129) in patients with negative KPNA2 staining ($P = 0.0047$; Fig. 6A). KPNA2 expression was also associated with shorter recurrence-free survival ($P = 0.0013$).

Two different Cox regression models were developed for the assessment of overall survival. Only nuclear KPNA2 immunohistochemistry, age at diagnosis, tumor stage, grade, lymph node status, estrogen receptor status, TP53 immunohistochemistry, and HER2 immunohistochemistry were considered (Table 4). In the global model, age ≥50 years ($P = 0.041$), positive lymph node status ($P = 0.015$), and negative estrogen receptor status ($P = 0.021$) were significant (Table 4). After reverse selection, the same three variables and KPNA2 immunohistochemistry remained in the model. Nuclear KPNA2

expression was an independent negative risk factor for overall survival (hazard ratio, 2.420; 95% confidence interval 1.200-4.882; $P = 0.014$). Because of the assumption of proportional hazards, the probability of death was consistently valid during the entire observation period. Finally, multiplicative terms of interaction (Inter 1-7) were considered, representing interactions between nuclear KPNA2 expression and dichotomous covariables. In the global model, only the interaction between lymph node status and KPNA2 immunohistochemistry was significant ($P = 0.031$). After reverse selection, a model containing age $\geq$50 years ($P = 0.023$), negative estrogen receptor status ($P = 0.003$), and the interaction between lymph node status and KPNA2 immunohistochemistry (Inter 3; hazard ratio, 4.558; 95% confidence interval, 2.397-8.664; $P < 0.001$) was found.

We did an exploratory subgroup analysis using Kaplan-Meier plots (Fig. 6B and C) because the estimated hazard ratio for Inter 3 was significantly below 1.0. KPNA2 protein expression was associated with shorter overall survival in the subgroup of node-positive patients ($P = 0.0062$) whereas KPNA2 immunohistochemistry had no prognostic effect in node-negative patients.

Discussion

This is the first gene expression study investigating laser-microdissected matched benign and malignant tissue pairs ($n = 24$) in breast cancer. A custom-designed Affymetrix gene chip with only 6,117 probe sets ($\sim$3,950 cDNA fragments) found 15 genes (Table 2) that are more up-regulated than *HER2*. Differential expression was evaluated by hybridization of gene-specific probes to a dot blot array containing 50 cDNA pairs derived from breast tumor and matching normal breast tissue (Breast Cancer Profiling Array, BD Clontech). Whereas HER2 RNA was up-regulated at least 2-fold in 26% of matching tumor/normal pairs, the 15 genes reached scores between 32% (*KPNA2*) and 84% (*PPAPDC1A*). Several recently discovered breast cancer–associated genes were included in the 15 gene list [e.g., aurora kinase A (*AURKA*; ref. 25) and calgranulin B (*S100A9*; ref. 26)]. However, seven genes have not been implied in mammary carcinogenesis thus far and may increase the pool of potentially useful molecules for breast cancer diagnosis and therapy.

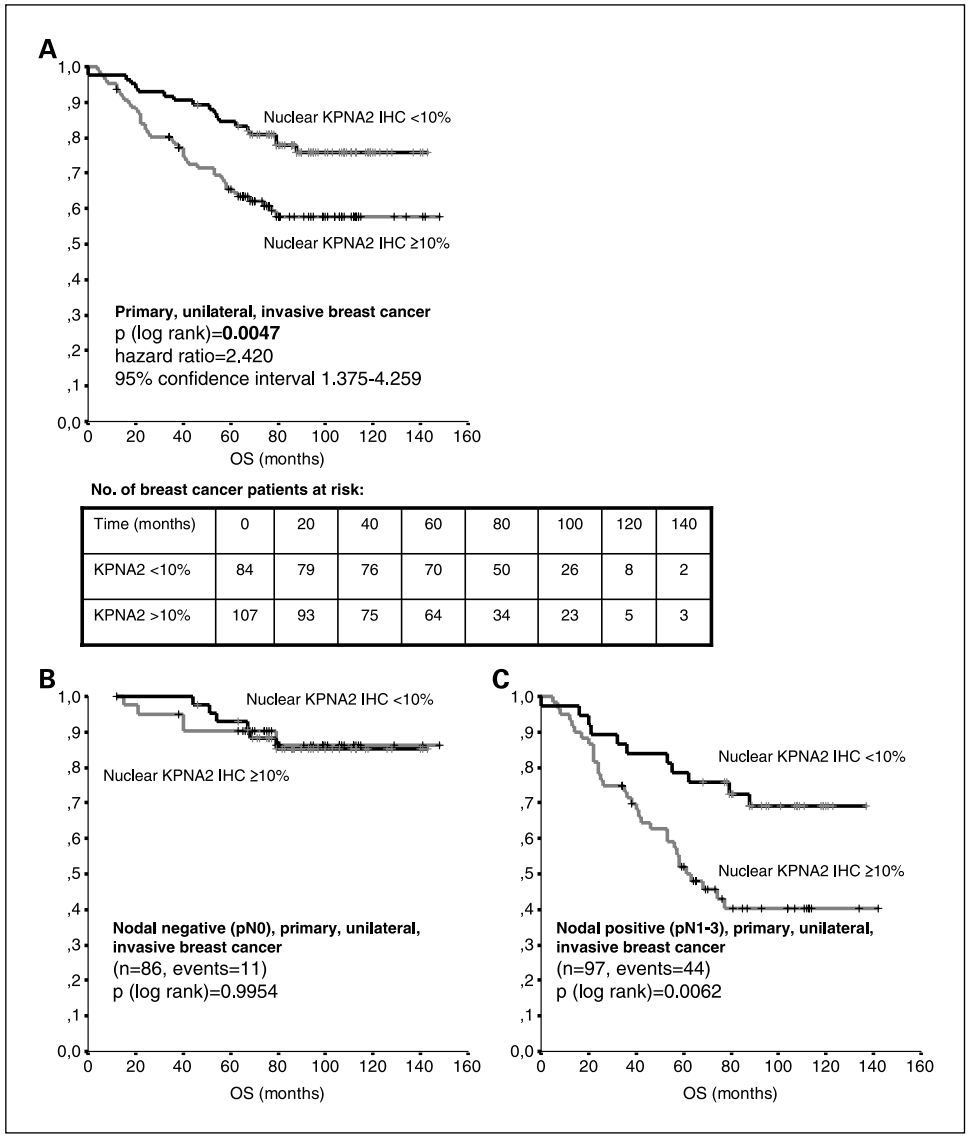

Fig. 6. *A,* distribution of time (months) to tumor-related death among 272 breast cancer patients with negative (<10%) and positive (≥10%) KPNA2 immunoreactivity as estimated by the method of Kaplan and Meier. *B* and *C,* distribution of time to tumor-related death among nodal negative (*B*) and nodal positive (*C*) breast cancer patients on KPNA2 immunoreactivity as estimated by the method of Kaplan and Meier.

Table 4. Multivariate Cox regression analysis of factors possibly influencing tumor-specific survival

Name	Variables*	Categorization		Global *P*	Stepwise reverse selection†	
					Hazard ratio (95% confidence interval)	*P*
(a) Raw model						
AGE	Age at diagnosis	0	<50 y	**0.041**	2,646 (1.095-6.394)	**0.031**
		1	≥50 y			
T	Tumor stage	0	pT_{1-2}	0.378	—	
		1	pT_{3-4}			
N	Lymph node status	0	pN_0	**0.015**	3,254 (1.582-6.694)	**0.001**
		1	pN_{1-3}			
G	Histologic grade	0	G_{1-2}	0.113	—	
		1	G_3			
ER	Estrogen receptor status	0	Negative	**0.021**	0.422 (0.224-0.759)	**0.008**
		1	Positive			
TP53	TP53 IHC	0	<10%	0.443	—	
		1	≥10%			
HER2	HER2 IHC	0	Negative (0-1+)	0.268	—	
		1	Positive (2+-3+)			
KPNA	KPNA2 IHC	0	<10%	0.142	2,420 (1.200-4.882)	**0.014**
		1	≥10%			
(b) Model with multiplicative terms of interaction (INTER)						
AGE	Age at diagnosis	0	<50 y	0.386	2,770 (1.148-6.684)	**0.023**
		1	≥50 y			
T	Tumor stage	0	pT_{1-2}	0.090	—	
		1	pT_{3-4}			
N	Lymph node status	0	pN_0	0.591	—	
		1	pN_{1-3}			
G	Histologic grade	0	G_{1-2}	0.271	—	
		1	G_3			
ER	Estrogen receptor status	0	Negative	0.175	0.385 (0.203-0.728)	**0.003**
		1	Positive			
TP53	TP53 IHC	0	<10%	0.731	—	
		1	≥10%			
HER2	HER2 IHC	0	Negative (0-1+)	0.098	—	
		1	Positive (2+-3+)			
KPNA	KPNA2 IHC	0	<10%	0.990	—	
		1	≥10%			
Inter 1	Interaction between KPNA and AGE		KPNA × AGE	0.945	—	
Inter 2	Interaction between KPNA and T		KPNA × T	0.155	—	
Inter 3	Interaction between KPNA and N		KPNA × N	**0.031**	4,558 (2.397-8.664)	**<0.001**
Inter 4	Interaction between KPNA and G		KPNA × G	0.685	—	
Inter 5	Interaction between KPNA and ER		KPNA × ER	0.954	—	
Inter 6	Interaction between KPNA and TP53		KPNA × TP53	0.883	—	
Inter 7	Interaction between KPNA and HER2		KPNA × HER2	0.207	—	

Abbreviation: IHC, immunohistochemistry.
*Limit for stepwise reverse selection procedures, $P = 0.1$.
†Boldface represents significant data.

In 1999, Blobel and colleagues showed that proteins of the karyopherin α and β families play a central role in nucleocytoplasmic transport (reviewed in ref. 27). KPNA2 is an adaptor protein within the classic nuclear protein import machinery which mediates the import of signaling factors in the nucleus and the export of response molecules to the cytoplasm (24). The classic nuclear protein import requires a nuclear localization signal within the cargo protein to be imported. This signal is recognized by the adaptor protein karyopherin/importin α. In blood lymphocytes, activation of cellular signaling was associated with strongly increased KPNA2 expression and redistribution of the KPNA2 protein from the cytoplasm to both the nucleus and the plasma membrane (28). Functional studies of BRCA1 by Thakur et al. have provided conclusive evidence that BRCA1 is a nuclear protein transported into the nucleus by the karyopherin

pathway (29), possibly linking KPNA2 expression with breast carcinogenesis. The present study observed strong nuclear staining of KPNA2 in breast tumor cells compared with a weak or absent staining in normal breast tissue, consistent with the hypothesis that increased nuclear signaling is present in a high percentage of human breast tumors. Lack of KPNA2 expression in normal breast tissue and overexpression in the majority of carcinomas exposes KPNA2 immunohistochemistry as a potential diagnostic marker. Investigation of KPNA2 expression in preneoplastic lesions of the breast is mandatory to establish the time point of KPNA2 up-regulation in the multistep process of mammary carcinogenesis.

Many studies have focused on the potential of gene expression profiles to predict the clinical outcome of breast cancer (4–6, 30–32). However, results of former studies are limited by exclusive analysis of bulk tissue containing a mixture of tumor, stromal, endothelial, and inflammatory cells. Some of the genes in the signatures presented to date seem to be derived from nonepithelial components of the tumor (31), suggesting that stromal elements represent an important contributing factor to the metastatic phenotype.

Survival differences were also noted between the different subtypes of breast tumors as defined by expression patterns (32, 33). Five distinct gene expression patterns were distinguished (32–34), including one basal-like, one HER2-overexpressing, two luminal-like, and one normal breast tissue–like subgroup. Cluster analysis of two published independent data sets, representing different patient cohorts from different laboratories, uncovered the same breast cancer subtypes (32). The patients with basal-like and HER2-subtypes were associated with the shortest survival. This strongly supported the idea that many of these breast tumor subtypes represent biologically distinct disease entities with different clinical outcome.

Supervised clustering with patient survival as the supervising variable by Sorlie et al. (33) resulted in a final list of 264 cDNA clones. This 264-clone set was then used for a hierarchical clustering analysis. The resulting diagram showed that almost all of the 264 cDNA clones selected fell into the three main gene expression clusters: the luminal/estrogen receptor positive cluster, the basal epithelial cluster, and the proliferation cluster. Interestingly, *KPNA2* was strongly expressed in tumors of the proliferation cluster (i.e., in the group of tumors with the shortest overall and recurrence-free survival). Notably, Dressman et al. (35) defined molecular signatures that correlate with response to neoadjuvant chemotherapy. Karyopherin α6 (*KPNA6*) was among the genes that predicted clinical response of stage IIB/III breast cancer patients (35). In our study, KPNA2 expression (≥10%) was significantly associated with shorter recurrence free survival ($P = 0.0013$). Due to missing retrospective data on chemotherapy and radiation therapy, no conclusions about therapy stratification of breast cancer patients based on KPNA2 expression can be drawn. A prospective clinical trial is currently planned together with the Department of Gynecology, University of Regensburg, Regensburg, Germany, to address this issue.

A remarkable feature of the expression signatures identified in previous studies is that they usually involve fewer than 100 genes (5, 30), in one instance even only 17 genes (31). However, the incomplete overlap between the different sets of defined genes is confusing (6, 32). Comparison of our 54 up-regulated and 78 down-regulated genes with the gene signatures of van de Vijver et al. (4) and Wang et al. (7) is difficult because of differences in patients and techniques. No overlap with the study of van de Vijver et al. (ref. 4; 70 genes) and only a single match (*KPNA2*) to the study by Wang et al. (ref. 7; 76 genes) were found. However, these studies compared breast tumors from long-term and short-term breast cancer survivors whereas our study compared normal breast tissue and invasive breast cancer. Therefore, a large overlap, which was even not detectable between the two closely related studies mentioned above, was not expected. Of note is that *KPNA2*, the gene we have focused on in our study, was also found by Wang et al. (7).

Sortiriou et al. (36) have examined whether histologic grade was associated with gene expression profiles of breast cancer and whether such profiles could be used to improve histologic grading; expression profiles between histologic grade 1 and histologic grade 3 tumors were compared. Interestingly, *KPNA2* was among the top overexpressed genes that were associated with histologic grade (36). Sortiriou et al. concluded that grade based on gene expression may reclassify grade 2 tumors into two groups with high and low risk of recurrence, improving the accuracy of histologic tumor grading and thus its prognostic value. In our tissue microarray study, KPNA2 expression and high tumor grade coassociated significantly ($P < 0.001$). However, in the subgroup of grade 2 tumors, no difference in tumor-related survival ($P = 0.1807$) and tumor recurrence ($P = 0.3156$) on KPNA2 expression could be observed. Previous studies have shown that mutations of the *TP53* gene predict poor prognosis (reviewed in ref. 37) and are associated with poor response to systemic chemotherapy (38). Overexpression of the HER2 protein is a well-known prognostic factor associated with poor survival in breast cancer, which also was found for the HER2-positive group defined by Sorlie et al. (33). In our study, high rates of KNPA2 expression were significantly associated with positive TP53 and HER2 immunohistochemistry and a high proliferation index (Table 3). KPNA2 seems to be characteristic of the basal-like subtype of breast cancers, possibly representing a different clinical entity of breast tumors, which is associated with shorter survival times and a high frequency of *TP53* mutations.

In summary, novel genes with clinical utility to select patients more likely to develop aggressive disease have been identified using a combination of laser microdissection of matched tumor/normal pairs and array expression analysis. We illustrated an independent negative correlation between KPNA2 expression in the primary tumor and overall survival in node-positive breast cancer patients. KPNA2 expression may represent a potential diagnostic marker to predict differential clinical responses to treatment and disease outcome. Prospective studies are currently conducted to validate our findings.

Acknowledgments

We thank Beate Petschke, Nicole Creutzburg, Monika Kerscher, Nina Nieβl, and Rudolf Jung for excellent technical assistance, Armin Pauer (Central Tumor Registry, Regensburg) for help in obtaining the clinical data, and metaGen Pharmaceuticals i.L. (Berlin, Germany) for providing oligonucleotide arrays (Affymetrix) and arrays for Northern blot analysis (BD Clontech).

References

1. Vogelstein B, Kinzler KW. The multistep nature of cancer. Trends Genet 1993;9:138–41.
2. Ethier SP. Identifying and validating causal genetic alterations in human breast cancer. Breast Cancer Res Treat 2003;78:285–7.
3. Ma XJ, Wang Z, Ryan PD, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell 2004;5:607–16.
4. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999–2009.
5. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6.
6. Huang E, Cheng SH, Dressman H, et al. Gene expression predictors of breast cancer outcomes. Lancet 2003;361:1590–6.
7. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365:671–9.
8. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;19:403–10.
9. Wild P, Knuechel R, Dietmaier W, Hofstaedter F, Hartmann A. Laser microdissection and microsatellite analyses of breast cancer reveal a high degree of tumor heterogeneity. Pathobiology 2000;68:180–90.
10. Kristiansen G, Pilarsky C, Wissmann C, et al. ALCAM/CD166 is up-regulated in low-grade prostate cancer and progressively lost in high-grade lesions. Prostate 2003;54:34–43.
11. Kristiansen G, Pilarsky C, Wissmann C, et al. Expression profiling of microdissected matched prostate cancer samples reveals CD166/MEMD and CD24 as new prognostic markers for patient survival. J Pathol 2005;205:359–76.
12. Schmitt AO, Specht T, Beckmann G, et al. Exhaustive mining of EST libraries for genes differentially expressed in normal and tumour tissues. Nucleic Acids Res 1999;27:4251–60.
13. Grutzmann R, Foerder M, Alldinger I, et al. Gene expression profiles of microdissected pancreatic ductal adenocarcinoma. Virchows Arch 2003;443:508–17.
14. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531–7.
15. Fink L, Seeger W, Ermert L, et al. Real-time quantitative RT-PCR after laser-assisted cell picking. Nat Med 1998;4:1329–33.
16. Graziano C. HER-2 breast assay, linked to Herceptin, wins FDA's okay. CAP Today 1998;12:14–6.
17. Kasper G, Weiser AA, Rump A, et al. Expression levels of the putative zinc transporter LIV-1 are associated with a better outcome of breast cancer patients. Int J Cancer 2005;117:961–73.
18. Waggoner DW, Xu J, Singh I, Jasinska R, Zhang QX, Brindley DN. Structural organization of mammalian lipid phosphate phosphatases: implications for signal transduction. Biochim Biophys Acta 1999;1439:299–316.
19. Kanoh H, Kai M, Wada I. Phosphatidic acid phosphatase from mammalian tissues: discovery of channel-like proteins with unexpected functions. Biochim Biophys Acta 1997;1348:56–62.
20. Sciorra VA, Morris AJ. Roles for lipid phosphate phosphatases in regulation of cellular signaling. Biochim Biophys Acta 2002;1582:45–51.
21. Dhoot GK, Gustafsson MK, Ai X, Sun W, Standiford DM, Emerson CP, Jr. Regulation of Wnt signaling and embryo patterning by an extracellular sulfatase. Science 2001;293:1663–6.
22. Su AI, Welsh JB, Sapinoso LM, et al. Molecular classification of human carcinomas by use of gene expression signatures. Cancer Res 2001;61:7388–93.
23. Lai JP, Chien J, Strome SE, et al. HSulf-1 modulates HGF-mediated tumor cell invasion and signaling in head and neck squamous carcinoma. Oncogene 2004;23:1439–47.
24. Leung SW, Harreman MT, Hodel MR, Hodel AE, Corbett AH. Dissection of the karyopherin α nuclear localization signal (NLS)-binding groove: functional requirements for NLS binding. J Biol Chem 2003;278:41947–53.
25. Ewart-Toland A, Dai Q, Gao YT, et al. Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. Carcinogenesis 2005;26:1368–73.
26. Arai K, Teratani T, Kuruto-Niwa R, Yamada T, Nozawa R. S100A9 expression in invasive ductal carcinoma of the breast: S100A9 expression in adenocarcinoma is closely associated with poor tumour differentiation. Eur J Cancer 2004;40:1179–87.
27. Chook YM, Blobel G. Karyopherins and nuclear import. Curr Opin Struct Biol 2001;11:703–15.
28. Andrade R, Alonso R, Pena R, Arlucea J, Arechaga J. Localization of importin α (Rch1) at the plasma membrane and subcellular redistribution during lymphocyte activation. Chromosoma 2003;112:87–95.
29. Thakur S, Zhang HB, Peng Y, et al. Localization of BRCA1 and a splice variant identifies the nuclear localization signal. Mol Cell Biol 1997;17:444–52.
30. Bertucci F, Nasser V, Granjeaud S, et al. Gene expression profiles of poor-prognosis primary breast cancer correlate with survival. Hum Mol Genet 2002;11:863–72.
31. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. Nat Genet 2003;33:49–54.
32. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003;100:8418–23.
33. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98:10869–74.
34. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406:747–52.
35. Dressman HK, Hans C, Bild A, et al. Gene expression profiles of multiple breast cancer phenotypes and response to neoadjuvant chemotherapy. Clin Cancer Res 2006;12:819–26.
36. Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 2006;98:262–72.
37. Hartmann A, Blaszyk H, Kovach JS, Sommer SS. The molecular epidemiology of p53 gene mutations in human breast cancer. Trends Genet 1997;13:27–33.
38. Berns EM, Foekens JA, Vossen R, et al. Complete sequencing of TP53 predicts poor response to systemic therapy of advanced breast cancer. Cancer Res 2000;60:2155–62.