Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 326 (2005) 218–227
BBRC
www.elsevier.com/locate/ybbrc

Loss of annexin A1 expression in human breast cancer detected by multiple high-throughput analyses☆

Dejun Shen[a], Helena R. Chang[a,*], Zugen Chen[b], Jianbo He[a], Victor Lonsberry[c], Yahya Elshimali[c], David Chia[c], David Seligson[c], Lee Goodglick[c], Stanley F. Nelson[b], Jeffrey A. Gornbein[d]

[a] *Gonda/UCLA Breast Cancer Research Laboratory, Department of Surgery, Revlon/UCLA Breast Center, University of California at Los Angeles, David Geffen School of Medicine, Los Angeles, CA 90095, USA*
[b] *Department of Human Genetics, UCLA, David Geffen School of Medicine, Los Angeles, CA 90095, USA*
[c] *UCLA Department of Pathology and Laboratory Medicine and Jonsson Comprehensive Cancer Center, University of California at Los Angeles, David Geffen School of Medicine, Los Angeles, CA 90095, USA*
[d] *Department of Biomathematics, University of California at Los Angeles, David Geffen School of Medicine, Los Angeles, CA 90095, USA*

Received 23 October 2004

Abstract

To test the efficacy of combined high-throughput analyses (HTA) in target gene identification, screening criteria were set using >fivefold difference by microarray and statistically significant changes ($p < 0.01$) in SAGE and EST. Microarray analysis of two normal and seven breast cancer samples found 129 genes with >fivefold changes. Further SAGE and EST analyses of these genes identified four qualified genes, ERBB2, GATA3, AGR2, and ANXA1. Their expression pattern was validated by RT-PCR in both breast cell lines and tissue samples. Loss of ANXA1 in breast cancer was further confirmed at mRNA level by Human Breast Cancer Tissue Profiling Array and at protein level by immunohistochemical staining. This study demonstrated that combined HTA effectively narrowed the number of genes for further study, while retaining the sensitivity in identifying biologically important genes such as ERBB2 and ANXA1. A distinctive loss of ANXA1 in breast cancer suggests its involvement in maintaining normal breast biology.

Keywords: Annexin A1; Breast cancer-related genes; Microarray; Multiple high-throughput analysis; SAGE; EST; CGAP

High-throughput gene expression analyses are important in characterizing global gene expression and have emerged as major tools for studying functional genomics of cancer [1]. Multiple bioinformatic infrastructures have been established to compile data from HTA. For example, the Cancer Genome Anatomy Project (CGAP) is a collaborative network to decipher genetic changes that occur during cancer formation and progression [2,3]. CGAP generates and disseminates genomic data on normal, pre-cancerous, and malignant processes. Since the results collected in CGAP come from established laboratories and represent a pool of similar studies, analysis of these accumulated data is similar to the meta-analysis. Multiple data query tools are available in the CGAP database, categorized in *Genes*, *Tissues*, *Pathways*, *RNAi*, *Chromosomes*, *SAGE Genie*, and *Tools*. For example, through *Gene Finder* under *Gene*, information

☆ *Abbreviations:* ANXA1, annexin A1; BCRG, breast cancer-related gene; CGAP, cancer genome anatomy project; DCIS, ductal carcinoma in situ; EST, expressed sequence tag; HTA, high-throughput analysis; IDC, invasive ductal carcinoma; RT-PCR, reverse transcriptase polymerase chain reaction; SAGE, series analysis of gene expression; vNorthern, virtual Northern.

* Corresponding author. Fax: +1 310 206 2982.
E-mail address: hchang@mednet.ucla.edu (H.R. Chang).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.10.214

about a specific gene can be browsed using a gene symbol, accession number, or keyword. Tissue-specific gene expression patterns can be found using *Virtual Northern* (*vNorthern*), which includes the data from both serial analyses of gene expression (SAGE) and expressed sequence tag (EST) analyses. Previous applications of data mining using CGAP resource have led to the identification of several novel or known cancer-related genes [4–7]. The CGAP gene index was also used to build the cDNA microarray for cancer classification [8].

A major challenge for high-throughput technologies is data analysis rather than data production [9,10]. Each of the three different high-throughput technologies, SAGE, EST, and microarray, possesses a different sensitivity and/or specificity in detecting specific expression variations. This often becomes more complicated considering the heterogeneity in samples, pathologies, and experimental methodologies. These factors may affect the outcome of individual gene analysis by a single method, although the genes consistently detected by multiple reliable methodologies are likely to reflect the underlying pathology. The accuracy of detecting distinct patterns of gene expression in cancer will be improved as more data are compiled in CGAP. It is possible that the findings of multiple HTA are complementary; thus, a combination of these methods may enhance the likelihood of identifying biologically important genes.

In this study, we explored the possibility of adding SAGE and EST database searches to simplify microarray analysis for identifying and validating important genes. Microarray was performed to screen for differentially expressed genes in normal and breast cancer specimens from patients. The identified differentially expressed genes were analyzed in conjunction with the bioinformatics of SAGE and EST. The candidate genes were selected by the presence of a minimal fivefold change in microarray analysis and statistically significant changes ($p < 0.01$) in bioinformatic SAGE and EST analyses. Only the markedly differentially expressed genes identified by all three methodologies were further analyzed. Of the four genes identified, annexin A1 (ANXA1) expression was found to be consistently lower or absent in breast cancer tissue. The loss of ANXA1 was further validated by real-time RT-PCR in both breast cancer cell lines and breast cancer specimens, breast cancer tissue profiling arrays consisting of 30 paired breast cancer specimens and their normal counterparts, and immunohistochemical staining of 10 samples each of normal and different stages of breast lesions, including hyperplasia, and in situ and invasive tumors.

Materials and methods

Sample collection and RNA extraction. Human breast tissues were collected according to NIH guidelines, and by protocols approved by the UCLA Institutional Review Board. Fresh surgical samples of breast cancer and normal breast tissues were grossly dissected to remove as much fat tissues as possible and then stored in liquid nitrogen. TriZol (Invitrogen, Carlsbad, CA) was used for total RNA extraction according to the manufacturer's instructions. RNA quality was examined by formaldehyde-denatured agarose gel electrophoresis and OD 260/280 ratio, and then quantified based on its OD260 reading by a DU 640B spectrophotometry (Beckman Coulter, Fullerton, CA).

Microarray analysis of human breast cancer. Seven breast cancer (two invasive ductal carcinomas, three ductal carcinoma in situ, and two invasive lobular carcinomas) and two normal breast tissue samples were used for microarray analysis. Ten micrograms of total RNA was used to synthesize double-stranded cDNA using the Superscript Choice System (Invitrogen, Carlsbad, CA). The double-stranded cDNA was purified using Phase Lock Gel (Pharmacia) and employed for synthesis of biotin-labeled anti-sense cRNA using the BioArray High Yield RNA Transcript Labeling Kit (Enzo Diagnostics, Farmingdale, NY). Labeled cRNA was fragmented, hybridized, and scanned as described in the Affymetrix GeneChip protocol (Affymetrix, Santa Clara, CA). The Affymetrix U95A array containing about 12,500 known human genes was used. Expression profiles were analyzed using GeneChip Analysis Suite 5.0 (Affymetrix, Santa Clara, CA). Only relative changes in gene expression equal to or greater than fivefold were further screened by bioinformatic SAGE and EST analyses. The hierarchical clustering of these genes was performed using the Cluster program and visualized with the TreeView program [11].

Virtual Northern analysis of gene expression. Virtual Northern analysis enables researchers to view information about the expression of a specific gene in all EST cDNA and SAGE libraries in the CGAP database [2,3,12]. Through *Gene Finder* at http://cgap.nci.nih.gov/Genes/GeneFinder, a multitude of organized information about a particular gene can be found by querying either the unique gene identifier or keyword. Included in the available information is the EST and SAGE *vNorthern* expression pattern across all the available libraries classified according to their tissue origins. The p values representing the statistical significance for both bioinformatic SAGE and EST are in the CGAP database, and their calculations have been previously described [13].

Real-time quantitative RT-PCR analysis. An independent panel of nine paired breast cancer tissue samples, different from those used in the microarray, was subjected to real-time RT-PCR analysis. The samples included five invasive ductal carcinomas, two invasive lobular carcinomas, and one each of DCIS and invasive mucinous carcinoma. Eight breast-derived cell lines including six breast cancer cell lines BT474, MCF-7, MDA-MB231, SKBR3, T47D, and ZR75 and two benign breast cell lines HBL100 and MCF-10A were also used in the study. Reverse transcription was performed using SuperScript II RNase H^- reverse transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. One microliter of cDNA was subjected to real-time RT-PCR amplification, using the QuantiTect SYBR Green PCR kit (Qiagen, Valencia, CA) in an iCycler IQ real-time PCR detection system (Bio-Rad Laboratory, Hercules, CA). All the primers for PCR were designed based on GenBank sequences with a web-based program, Primer3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). Primer sequence information is listed in Table 1. A-20 μl reaction contained 0.3 μM of each of the primers, 0.5 μl cDNA, and 1× QuantiTect SYBR Green PCR mix. The following PCR amplification program was used: initial DNA denaturation and HotStarTaq activation at 95 °C for 15 min; 45 PCR cycles at 94 °C for 20 s, 65 °C for 30 s, and 72 °C for 20 s; data acquisition was set at the chain extension step; and melting curve analysis was performed between 55 and 95 °C, with 0.5 °C-increments. All data were analyzed using the iCycler IQ optical system software version 3.0a (Bio-Rad Laboratory).

Validation of differential ANXA1 expression in human breast cancer using the human breast cancer tissue profiling array. The human

Table 1
Primer sequences used in real-time RT-PCR

Primers	Sequences (5′ → 3′)
ANXA1 forward	TTGAGGAGGTTGTTTTAGCTCTG
ANXA1 reverse	AGTTCTTGATGCCAAAATCTCAA
ERBB2 forward	CTCTACAGCGGTACAGTGAGGAC
ERBB2 reverse	AACATCTGGCTGGTTCACATATT
GATA3 forward	AGGGTCTCTAGTGCTGTGAAAA
GATA3 reverse	CATCCTTCATGCCTTACAGCTAC
AGR2 forward	AAGCAACAAACCCTTGATGATTA
AGR2 reverse	TAAACCAGATTGAGGAGGACAAA
β-Actin forward	TCAAGATCATTGCTCCTCCTGA
β-Actin reverse	CTCGTCATACTCCTGCTTGCTG

breast cancer tissue profiling array was obtained from BD Bioscience Clontech (Palo Alto, CA). The array contained normalized cDNA from 30 pairs of breast tumors, including 22 invasive ductal carcinomas, five invasive lobular carcinomas, and one each of DCIS, medullary, and tubular carcinomas, along with their corresponding normal tissues derived from the same patients. The array also contained negative and positive controls, and cDNA from three breast cancer cell lines, MCF-7, MDA-MB231, and MDA-MB435S. A plasmid containing the full-length cDNA of ANXA1 gene was obtained from Open Biosystems (Huntsville, AL). The clone was re-sequenced in the laboratory, and the sequence of the ANXA1 insert was verified. A 1.2-kb ANXA1 probe was prepared by restrictive digestion and purified using a Qiagen column (Valencia, CA). The probe was labeled with [^{32}P]dCTP, using random primer labeling (Stratagene, La Jolla, CA). The hybridization of the ANXA1 probe to the array was performed according to the manufacturer's instructions. The same array was then stripped and re-hybridized with a control ubiquitin probe provided by Clontech. The differential expression of ANXA1 in breast cancers and their corresponding normal breast tissues was compared.

Immunohistochemical staining of annexin A1 in human breast cancer tissue samples. The monoclonal anti-ANXA1 antibody (IgG1, clone 29) was purchased from BD Transduction Laboratories (Lexington, KY). Ten samples each of normal breast tissues, ductal hyperplasia, carcinoma in situ, and invasive carcinoma were obtained from the archived pathological tissue bank in the UCLA Department of Pathology and Laboratory Medicine. Paraffin-embedded tissues were sectioned, warmed to 56 °C for 30 min, and then deparaffinized using xylenes. Following rehydration, endogenous peroxidase activity was quenched with 3% H_2O_2 for 5 min (Dako Envision+ kit), as

Table 2
Gene expression of 30 breast cancer-related genes determined by multiple HTA

Gene symbol	Gene name	Differential gene expression (C/N)		
		Microarray	EST	SAGE
(1) Oncogenes				
BCL2	B-cell CLL/lymphoma 2	4.89	1.48	1.08
CCND1	Cyclin D1	3.15	1.28	7.55*
CTSD	Cathepsin D	3.60	3.77*	3.30*
EGFR	Epidermal growth factor receptor	0.13	0.07*	0.32
ERBB2	Her2/neu	6.71	25.20*	32.11*
IGF1	Insulin-like growth factor 1	0.36	1.48	1.62
IGF1R	Insulin-like growth factor 1 receptor	2.15	2.97	0.65
MYC	c-myc	0.34	0.49	0.34*
(2) Tumor suppressor genes				
ATM	Ataxia telangiectasia mutated gene	1.14	0.12**	1.15
BRCA1	Breast cancer 1, early onset	1.36	0.49	0.32
BRCA2	Breast cancer 2, early onset	0.87	1.00	0.69*
CDKN1A	p21/WAF1/CIP1	0.48	0.40**	0.41*
CDKN1B	p27/Kip1	0.76	0.40	0.70
CDKN2A	p16/INK4a	1.11	0.25	0.43
FHIT	Fragile histidine triad gene	1.66	1.00	0.32
IGFBP3	Insulin-like growth factor binding protein 3	0.53	0.14*	0.16*
TP53	Tumor protein p53	1.07	0.59	1.13
(3) Other breast cancer-related genes				
AR	Androgen receptor	1.55	0.49	0.57
CD44	CD44 antigen	1.64	1.03	0.45
CDH1	E-cadherin	0.89	1.14	1.74
ESR1	Estrogen receptor 1	1.47	1.98	12.30*
ESR2	Estrogen receptor 2	0.71	0.49	1.08
FGF2	Basic fibroblast growth factor	0.30	0.74	0.14*
KLK3	Prostate-specific antigen	3.56	30.00*	1.00
KRT19	keratin 19	32.39	1.11	2.06*
MGB1	Mammaglobin	58.43	9.39*	3.27**
MKI67	Ki-67antigen	3.33	6.92**	6.15
MUC1	Mucin 1	6.54	0.74	2.94*
PGR	Progesterone receptor	0.56	1.00	0.00
VEGF	Vascular endothelial growth factor	0.78	0.49	1.01

* $p < 0.01$.
** $p < 0.05$, but >0.01.

described by the manufacturer. Samples were incubated in sodium citrate buffer (pH 6.0) for 30 min at 95 °C to obtain maximal antigen retrieval. Non-specific binding was blocked by pre-incubation with 5% normal goat serum in PBS for 20 min. Anti-ANXA1 monoclonal antibody was added to samples with a final concentration of 0.06 μg/ml diluted in 1% bovine serum albumin and incubated for 40 min at room temperature. Following extensive washing, samples were incubated for 40 min at room temperature with secondary antibody-peroxidase labeled polymer (DAKO Envision+ kit). Samples were extensively washed, treated with DAB substrate-chromagen solution for 5 min at room temperature, as described by the manufacturer (DAKO Envision+ kit), and counterstained with hematoxylin. Controls included incubation with primary isotype antibody. The immunochemical staining was evaluated and scored by a board-certified pathologist (Y.E.), based on relative staining intensity (scale of 0, 1, 2, and 3) and the percentage of cells of a given histology/pathology staining positively. The staining score (SS) was calculated by combining both staining intensity and percentage of cells stained based on the following formula: $SS = 3a\% + 2b\% + 1c\%$. The staining difference among different groups of samples was compared using Student's t test.

Results

Differential gene expression detected by multiple HTA

CGAP employed several approaches, including EST and SAGE, to catalog all genes expressed during cancer development [3]. Through May 2004, the CGAP database has collected about four million EST and 12 million SAGE tags. To explore the feasibility of using this resource for measuring gene expression levels, *Virtual Northern* was used to study the expression of a panel of 30 known breast cancer-related genes in cancerous and benign breast tissues (Table 2). Microarray, SAGE, and EST *vNorthern* each identified the expressional changes of a set of breast cancer-related genes. Microarray detected four differentially expressed genes, and SAGE and EST, 11 and seven, respectively (Table 2). Only a few differentially expressed genes detected are not consistent with the current knowledge, such as the down-regulation of epidermal growth factor receptor (EGFR) identified by EST *vNorthern* in human breast cancer. Some differences were also noticed among the three different high-throughput approaches. This is not surprising, due to the heterogeneity in pathologies, samples, and methodologies. Nonetheless, our study showed a clear reduction of tumor suppressor genes and upregulation of oncogenes in cancers (Table 2). Consistent differential expression of ERBB2, CTSD, MYC, CDKN1A, CDKN2A, IGFBP3, MKI67, and MGB1 was identified by all three HTA. If the screening criteria were set as microarray > twofold, and SAGE and EST $p < 0.05$, ERBB2, CDKN1A, CDKN2A, IGFBP3, and MGB1 are selected. However, if more stringent criteria were used, with microarray > fivefold, and SAGE and EST $p < 0.01$, only ERBB2 would be selected. These results suggest that by increasing the screening stringency, one may narrow down the important genes for further investigation.

Analysis of human breast cancer by microarray combined with bioinformatic SAGE and EST

Seven fresh breast cancer and two normal breast tissue samples were analyzed using Affymetrix U95A arrays. Among 12,627 genes analyzed, 730 genes with more than twofold and 129 genes with more than fivefold differential expression were identified (Figs. 1 and 2). These included several known important breast cancer-related genes, such as c-erb B2, MUC1, keratin 19, and mammaglobin. The 129 genes with more than fivefold differential expression were further analyzed against the data from bioinformatic SAGE and EST. Four genes, c-erb B2 (ERBB2), T lymphocyte transcriptional factor GATA binding protein 3 (GATA3), *Xenopus laevis* homolog gene of anterior gradient 2 (AGR2), and annexin A1 (ANXA1), met the screening criteria of fivefold by microarray and $p < 0.01$ by both SAGE and EST (Fig. 3). Therefore, they were selected for further analysis.

Loss of ANXA1 expression in breast cancer validated by real-time RT-PCR and human breast cancer tissue profiling array

Consistent detection of four differentially expressed genes, ERBB2, GATA3, AGR2, and ANXA1, by all three HTA implies a role of these genes in breast cancer. The expression of these genes was further verified using real-time RT-PCR quantification in six breast cancer

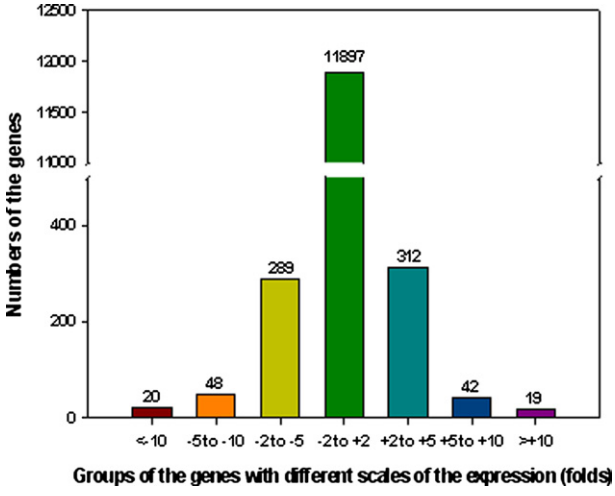

Fig. 1. Distribution of genes in microarray analysis of human breast cancer. Seven breast cancer and two normal breast tissue samples were analyzed using Affymetrix U95A chips. The distribution of genes was presented in differential expression scales. The "+" and "−" indicate the up-regulation and down-regulation of genes, respectively, in breast cancer versus normal breast tissues.

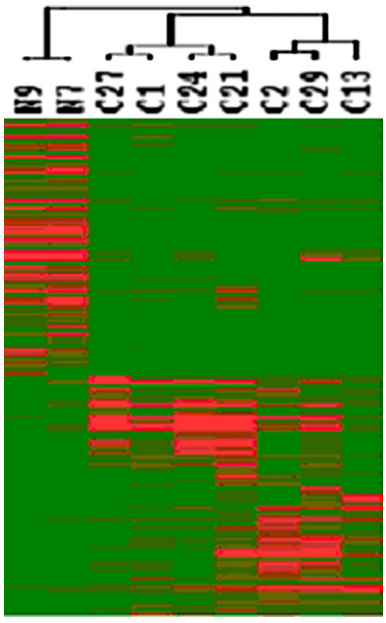

Fig. 2. Hierarchical clustering of genes identified by microarray analysis breast cancer. The hierarchical clustering of 129 genes with more than fivefold differential expression was performed using the program Cluster and visualized with TreeView. "N" and "C" represent normal and cancer, respectively.

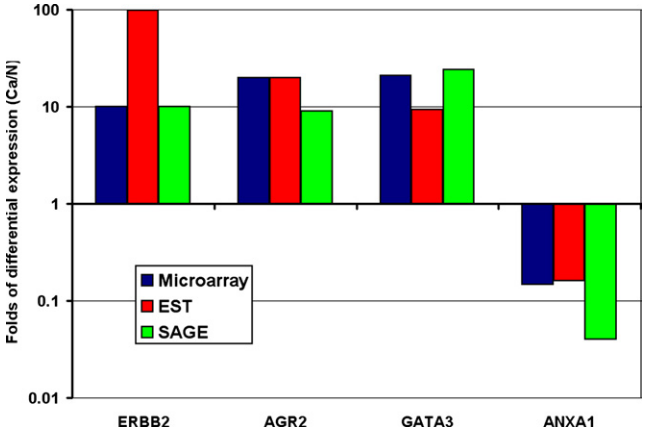

Fig. 3. Expression of breast cancer-related genes in human breast cancer determined by multiple HTA. One hundred and twenty-nine genes with more than fivefold differential expression identified by microarray analysis were further analyzed by SAGE and EST *vNorthern*. Four genes, ERBB2, AGR2, GATA3, and ANXA1, met the screening criteria of expressional difference of >fivefold by microarray and $p < 0.01$ by both SAGE and EST.

cell lines, two benign breast epithelial cell lines (Fig. 4), and nine paired human breast cancer samples and their matched normal tissues (Fig. 5). The expression of ERBB2, GATA3, and AGR2 was not detectable in the two immortalized benign breast cell lines, HBL100 and MCF10A, compared to various levels of expression of these genes in breast cancer cell lines. ANXA1 expression was minimal in both malignant and immortalized benign cell lines (Fig. 4). In human tissues, eight of nine

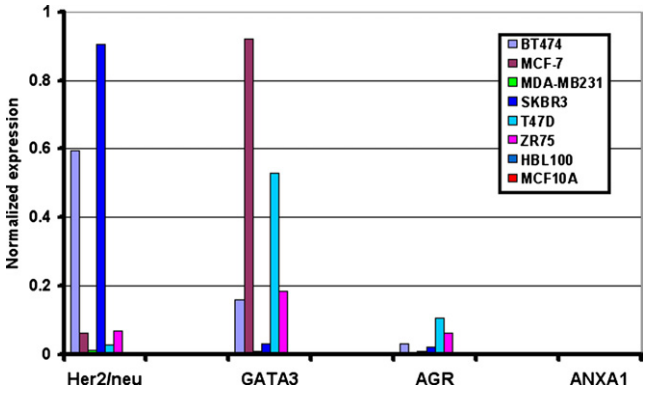

Fig. 4. Real-time RT-PCR analysis of breast cancer-related genes in breast derived cell lines. Six breast cancer and two immortalized benign breast cell lines were analyzed for the expression of four genes identified by multiple HTA. The expression was presented as normalized expression by β-actin.

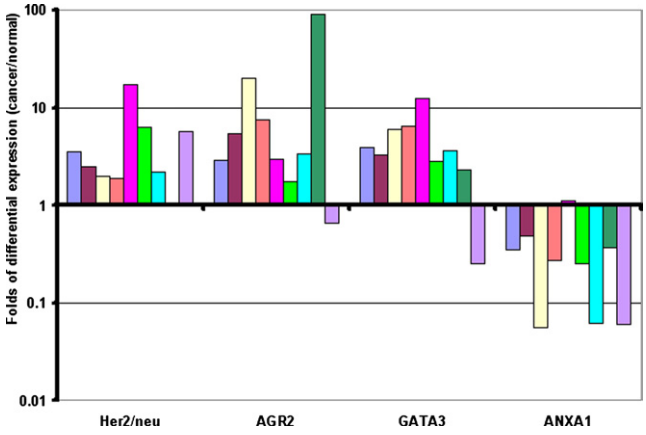

Fig. 5. Real-time RT-PCR analysis of breast cancer-related gene in paired breast cancer tissue samples. The expression changes in paired breast cancer tissues were presented as folds of normalized expression by β-actin in cancers versus in their normal counterparts. Each color bar represents a separate pair of tissue samples.

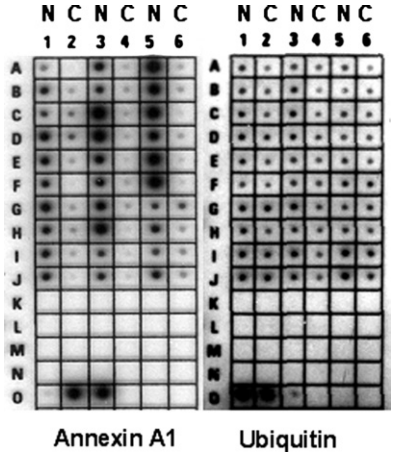

Fig. 6. ANXA1 expression in human breast cancer tissue profiling array. The human breast cancer tissue profiling array contains 30 pairs of cDNA from human breast cancer and their normal counterparts. "N" and "C" represent normal and cancer, respectively. O1, O2, and O3 are three breast cancer cell lines, MCF-7, MDA-MB-231, and MDA-MB435S as experimental controls.

samples had elevated expression of ERBB2, GATA3, and AGR2 in breast cancer, while ANXA1 expression was reduced or lost in eight out of nine cancer tissue samples (Fig. 5).

We further investigated the expression of ANXA1 in a commercially available human breast cancer tissue profiling array (BD Clontech). This array contains normalized cDNA from 30 paired breast cancer and corresponding normal breast tissue samples. The results confirmed the loss of ANXA1 expression in breast cancer tissues. ANXA1 expression was significantly reduced or lost in 29 of 30 breast cancer samples, as compared to their normal counterparts (Fig. 6).

Decreased ANXA1 expression in human breast cancer epithelial cells detected by immunohistochemical staining

Controversy about ANXA1 expression in breast cancer was reported in studies using immunohistochemical staining [14–16]. Therefore, we further examined the

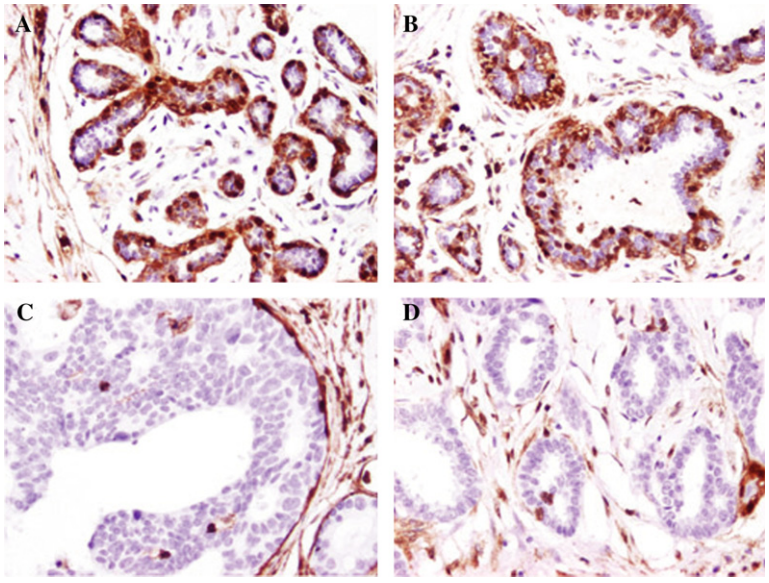

Fig. 7. Immunohistochemical staining of ANXA1 in benign and malignant breast tissues. (A) Normal; (B) ductal hyperplasia; (C) in situ carcinoma; and (D) invasive carcinoma.

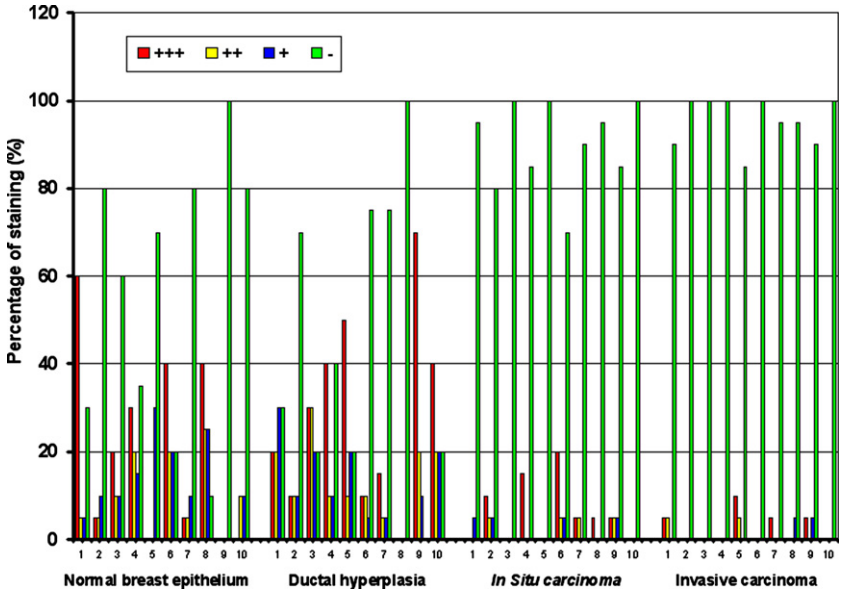

Fig. 8. Scoring of immunohistochemical staining of ANXA1 in different types of breast lesions. The immunohistochemical staining of ANXA1 was scored according to a combination of both intensity as +++, ++, +, –, and percentage.

expression pattern of ANXA1 in normal and various human breast lesions by immunohistochemical staining. As shown in Figs. 7 and 8, breast glandular epithelial cells in normal and hyperplastic breast tissue expressed constitutive levels of ANXA1. Moreover, relatively abundant levels of ANXA1 were observed in myoepithelial cells of both malignant lesions and non-malignant glands (Fig. 7). However, there was a significantly decreased epithelial staining of ANXA1 in in situ and invasive carcinomas when compared to benign breast tissues ($p < 0.0001$). The average score of annexin A1 staining was 43.5 ± 30.6 in normal breast tissue, 52.0 ± 32.7 in hyperplasia, and only 10.0 ± 10.0 and 4.5 ± 5.5 in in situ and invasive carcinomas, respectively (Fig. 8). There was no significant difference for annexin A1 staining between normal and ductal hyperplasia ($p = 0.27$) and between in situ and invasive carcinoma ($p = 0.60$). A significant loss of annexin A1 staining was observed both in in situ tumors versus normal breast epithelium ($p = 0.0025$) and ductal hyperplasia ($p = 0.0001$), as well as in invasive carcinoma versus normal breast epithelium ($p = 0.0006$) and ductal hyperplasia ($p < 0.0001$).

Discussion

Multiple HTA

HTA has proven to be a powerful tool in novel gene discovery, classification, and prognosis prediction in cancers. Recent studies focused on identification of disease-related genes [17,18], expressional profiling to classify cancer into various subtypes [19–21], and to predict prognosis [22–24]. Since different HTA have different sensitivities and/or specificities in target gene discovery, we hypothesized that a properly combined use of these technologies might improve the identification of differentially expressed genes. While analytical variations may occur due to the heterogeneity in samples, pathology, and methodologies, the pattern of differential gene expression should be consistently detected by reliable analyses to reflect the underlying pathology, despite the multiple possible variations. Based on these rationales, we designed an approach to combine multiple HTA, including cDNA microarray and bioinformatic SAGE and EST, to search for a subset of candidate genes potentially important in the development and/or progression of cancer. We first compared the gene expression levels of a panel of 30 known breast cancer-related genes between normal and malignant breast tumors, and the results were compared to those of microarray analysis. As shown in Table 2, the expression of ERBB2, CTSD, MYC, CDKN1A, CDKN2A, IGFBP3, MKI67, and MGB1 was consistent in all three HTA. ERBB2, CDKN1A, CDKN2A, IGFBP3, and MGB1 were found to be significant when the search criteria were set at the level of microarray > twofold, and SAGE and EST $p < 0.05$. However, only ERBB2 would be detected if more stringent criteria of microarray > fivefold plus SAGE and EST $p < 0.01$ were applied. This suggests that multiple HTA could be complementary when proper criteria are used. With increasing stringency of the screening criteria applied to combined HTA analysis, the number of candidate genes would be reduced, and hence, the important differentially expressed genes could be identified more efficiently.

This strategy was then applied to our microarray data analysis. Seven primary human breast cancer and two normal breast tissue samples were analyzed using Affymetix U95A chips. Among the 12,627 genes analyzed, 730 genes were found to have more than twofold differential expression and 129 genes if the selection was limited to those genes with a minimum differentiation of fivefold (Figs. 1 and 2). After cross-analysis by bioinformatic SAGE and EST analyses with a criteria of $p < 0.01$, only four of the 129 genes were significant, including three overexpressed genes, ERBB2, GATA3, and AGR2, and a suppressed gene, ANXA1, in breast cancer (Fig. 3). This suggested that, at the given screening criteria, combined microarray with SAGE and EST analyses effectively narrowed the number of genes for further analysis, yet retaining the sensitivity to identify the biologically important genes, such as ERBB2 and ANXA1.

Studying differential expression between breast cancer and normal breast can be technically challenging, given that normal tissues are poor in epithelium as compared to adipocytes, whereas tumors may be epithelial and endothelial cell-enriched. One approach to address this concern is to incorporate laser capture microdissection or antibody purification to enrich the epithelial cells from normal or malignant tissues. However, these processes may not be feasible when only small amounts of tissues are available. Further, the time-consuming procedures may also compromise the quality of RNA preparation. To address the problems in tissue sampling, the following steps were taken: (1) all visible fat was removed from normal breast tissues for RNA extraction; (2) the results of microarray analysis was cross-validated with bioinformatic SAGE and EST; and (3) the findings were then confirmed in breast cell lines and tissue samples at both RNA and protein levels. Although this approach may miss the genes with less significant changes, it allows us to select and to focus on those genes with the most striking differences. In our study, several known breast cancer genes such as ERBB2 and ANXA1 are successfully identified by the described method using a very stringent screening criterion and the grossly dissected tissue samples.

Characteristics of the genes identified by multiple HTA

ERBB2 is a member of an epidermal growth factor family with tyrosine kinase activity, and is a well-known breast cancer oncogene with both gene amplification and overexpression in up to 25% of breast cancer patients. The overexpression of ERBB2 has been associated with poor prognosis of patients with breast cancer [25] and its overexpression detected by microarray analysis was also previously reported [20]. Increased expression of two other genes identified in this study, GATA3 and AGR2, was observed by others as well [26,27]. They were found to co-express with estrogen receptor in human breast cancer and were also detected by HTA [20,28,29]. GATA3 was also shown to be involved in the switch of host immunity from cellular to humoral [30,31]. The pattern of ANXA1 expression in human cancer is not well defined. While up-regulation of ANXA1 was previously observed in a stomach cancer cell line [32], glioblastoma [4], and mammary adenocarcinoma [15,16,33], many investigators reported that ANXA1 expression was reduced or lost in cancer tissues. Down-regulation of ANXA1 was reported in many types of cancers, including head and neck [34], esophageal [35], gastric [36], prostate [35,37], and breast [14]. In this study, we clearly demonstrated that ANXA1 expression was diminished or lost in breast cancer tissues by all three of the high-throughput technologies, microarray, and bioinformatic SAGE and EST. This was further validated both at the mRNA level by real-time RT-PCR and human breast cancer tissue profiling array, and at the protein level by immunochemical staining of human breast cancer tissues.

Loss of ANXA1 expression in human breast cancer

Annexins are a group of structurally related calcium-binding proteins, which share a domain that binds to phospholipids and an amino terminal domain that determines specificity [38]. The functions of the annexin family proteins are associated with regulation of membrane trafficking and cellular adhesion. Several members of the annexin family, including annexins 1, 2, 5, 6, and 7, were reported to be involved in the oncogenic process [39–44]. These annexins are often linked with negative regulation of cell proliferation [38]. Although the role of ANXA1 in cancer development has not been well characterized, the down-regulation of ANXA1 expression in multiple types of cancers suggests that ANXA1 may be a negative regulator of cancer cell growth. ANXA1, also known as lipocortin 1(LC-1), is a 37-kDa calcium- and phospholipid-binding protein of the annexin superfamily found in a wide range of organisms, including vertebrates, invertebrates, and plants [38]. ANXA1 is an important mediator in glucocorticoid-regulated inflammatory response [45,46]. It was also found to be involved in the regulation of cytokine production [47] and the MAPK/ERK signal transduction pathway [48,49]. Forced expression of ANXA1 activated the MAPK/ERK pathway and reduced cell proliferation by disruption of the actin skeleton and ablation of cyclin D1 expression [49]. When treated with an anticancer agent, perillyl alcohol, ANXA1 expression in breast cancer was upregulated, along with other apoptosis-related genes, p21 (Cip1/WAF1), Bax, and Bad [50]. In our study, ANXA1 expression is decreased or lost at the mRNA level, as measured by multiple HTA, real-time RT-PCR, and human breast cancer tissue profiling array, and at the protein level as determined by immunohistochemical staining in most breast cancer tissue samples examined, including both in situ and invasive carcinomas. ANXA1 expression at very low levels was observed in both cancerous and immortalized normal breast epithelial cell lines, suggesting that ANXA1 loss might occur at an early stage of malignant transformation. Immunohistochemical staining of various human breast lesions showed that significant loss of annexin A1 staining occurred between ductal hyperplasia, and in situ and invasive carcinomas.

Conclusion

In summary, four differentially expressed genes in human breast cancer, ERBB2, GATA3, AGR2, and ANXA1, were detected by multiple HTA consisting of microarray and bioinformatic SAGE and EST. The loss of ANXA1 was further confirmed both at the mRNA level by real-time RT-PCR and human breast cancer profiling array, and at the protein level by immunohistochemical staining. Our study suggests that multiple HTA is a useful strategy for discovering biomarkers with clinical relevance in cancer detection and disease classification. Furthermore, the consistent loss of ANXA1 in breast cancer suggests a regulatory role in human breast cancer development.

Acknowledgments

We thank the UCLA DNA microarray core for technical assistance in microarray analysis and Rebecca Radbod for statistical analysis. This study is supported in part by grants from the National Cancer Institute (5R01 CA093736), the Gonda Foundation, the UCLA Human Gene Medicine Program, and the Early Detection Research Network (NCI CA986366).

References

[1] W. Zhang, P.M. Laborde, K.R. Coombes, D.A. Berry, S.R. Hamilton, Cancer genomics: promises and complexities, Clin. Cancer Res. 7 (2001) 2159–2167.

[2] K. Boon, E.C. Osorio, S.F. Greenhut, C.F. Schaefer, J. Shoemaker, K. Polyak, P. Morin, K.H. Buetow, R.L. Strausberg, S.J. De Souza, G.J. Riggins, An anatomy of normal and malignant gene expression, Proc. Natl. Acad. Sci. USA 99 (2002) 11287–11292.

[3] R.L. Strausberg, A.A. Camargo, G.J. Riggins, C.F. Schaefer, S.J. de Souza, L.H. Grouse, A. Lal, K.H. Buetow, K. Boon, S.F. Greenhut, A.J. Simpson, An international database and integrated analysis tools for the study of cancer gene expression, Pharmacogenomics J. 2 (2002) 156–164.

[4] W.T. Loging, A. Lal, I.M. Siu, T.L. Loney, C.J. Wikstrand, M.A. Marra, C. Prange, D.D. Bigner, R.L. Strausberg, G.J. Riggins, Identifying potential tumor markers and antigens by database mining and rapid expression screening, Genome Res. 10 (2000) 1393–1402.

[5] I.E. Krop, D. Sgroi, D.A. Porter, K.L. Lunetta, R. LeVangie, P. Seth, C.M. Kaelin, E. Rhei, M. Bosenberg, S. Schnitt, J.R. Marks, Z. Pagon, D. Belina, J. Razumovic, K. Polyak, HIN-1, a putative cytokine highly expressed in normal but not cancerous mammary epithelial cells, Proc. Natl. Acad. Sci. USA 98 (2001) 9796–9801.

[6] B. Ryu, J. Jones, M.A. Hollingsworth, R.H. Hruban, S.E. Kern, Invasion-specific genes in malignancy: serial analysis of gene expression comparisons of primary and passaged cancers, Cancer Res. 61 (2001) 1833–1838.

[7] P. Buckhaults, Z. Zhang, Y.C. Chen, T.L. Wang, B. St Croix, S. Saha, A. Bardelli, P.J. Morin, K. Polyak, R.H. Hruban, V.E. Velculescu, I.M. Shih, Identifying tumor origin using a gene expression-based classification map, Cancer Res. 63 (2003) 4144–4149.

[8] A.A. Alizadeh, M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, J.C. Boldrick, H. Sabet, T. Tran, X. Yu, J.I. Powell, L. Yang, G.E. Marti, T. Moore, J. Hudson Jr., L. Lu, D.B. Lewis, R. Tibshirani, G. Sherlock, W.C. Chan, T.C. Greiner, D.D. Weisenburger, J.O. Armitage, R. Warnke, R. Levy, W. Wilson, M.R. Grever, J.C. Byrd, D. Botstein, P.O. Brown, L.M. Staudt, Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling, Nature 403 (2000) 503–511.

[9] R. Simon, M.D. Radmacher, K. Dobbin, L.M. McShane, Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification, J. Natl. Cancer Inst. 95 (2003) 14–18.

[10] V. Detours, J.E. Dumont, H. Bersini, C. Maenhaut, Integration and cross-validation of high-throughput gene expression data: comparing heterogeneous data sets, FEBS Lett. 546 (2003) 98–102.

[11] M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein, Cluster analysis and display of genome-wide expression patterns, Proc. Natl. Acad. Sci. USA 95 (1998) 14863–14868.

[12] D.A. Porter, I.E. Krop, S. Nasser, D. Sgroi, C.M. Kaelin, J.R. Marks, G. Riggins, K. Polyak, A SAGE (serial analysis of gene expression) view of breast tumor progression, Cancer Res. 61 (2001) 5697–5702.

[13] A. Lal, A.E. Lash, S.F. Altschul, V. Velculescu, L. Zhang, R.E. McLendon, M.A. Marra, C. Prange, P.J. Morin, K. Polyak, N. Papadopoulos, B. Vogelstein, K.W. Kinzler, R.L. Strausberg, G.J. Riggins, A public database for gene expression in human cancers, Cancer Res. 59 (1999) 5403–5407.

[14] R. Schwartz-Albiez, K. Koretz, P. Moller, G. Wirl, Differential expression of annexins I and II in normal and malignant human mammary epithelial cells, Differentiation 52 (1993) 229–237.

[15] S.H. Ahn, H. Sawada, J.Y. Ro, G.L. Nicolson, Differential expression of annexin I in human mammary ductal epithelial cells in normal and benign and malignant breast tissues, Clin. Exp. Metastasis 15 (1997) 151–156.

[16] S.D. Pencil, M. Toth, Elevated levels of annexin I protein in vitro and in vivo in rat and human mammary adenocarcinoma, Clin. Exp. Metastasis 16 (1998) 113–121.

[17] T. Bouras, M.C. Southey, A.C. Chang, R.R. Reddel, D. Willhite, R. Glynne, M.A. Henderson, J.E. Armes, D.J. Venter, Stanniocalcin 2 is an estrogen-responsive gene coexpressed with the estrogen receptor in human breast cancer, Cancer Res. 62 (2002) 1289–1295.

[18] D. Porter, S. Weremowicz, K. Chin, P. Seth, A. Keshaviah, J. Lahti-Domenici, Y.K. Bae, C.L. Monitto, A. Merlos-Suarez, J. Chan, C.M. Hulette, A. Richardson, C.C. Morton, J. Marks, M. Duyao, R. Hruban, E. Gabrielson, R. Gelman, K. Polyak, A neural survival factor is a candidate oncogene in breast cancer, Proc. Natl. Acad. Sci. USA 100 (2003) 10931–10936.

[19] K.J. Martin, B.M. Kritzman, L.M. Price, B. Koh, C.P. Kwan, X. Zhang, A. Mackay, M.J. O'Hare, C.M. Kaelin, G.L. Mutter, A.B. Pardee, R. Sager, Linking gene expression patterns to therapeutic groups in breast cancer, Cancer Res. 60 (2000) 2232–2238.

[20] C.M. Perou, T. Sorlie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, C.A. Rees, J.R. Pollack, D.T. Ross, H. Johnsen, L.A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S.X. Zhu, P.E. Lonning, A.L. Borresen-Dale, P.O. Brown, D. Botstein, Molecular portraits of human breast tumours, Nature 406 (2000) 747–752.

[21] M. West, C. Blanchette, H. Dressman, E. Huang, S. Ishida, R. Spang, H. Zuzan, J.A. Olson Jr., J.R. Marks, J.R. Nevins, Predicting the clinical status of human breast cancer by using gene expression profiles, Proc. Natl. Acad. Sci. USA 98 (2001) 11462–11467.

[22] D.A. Zajchowski, M.F. Bartholdi, Y. Gong, L. Webster, H.L. Liu, A. Munishkin, C. Beauheim, S. Harvey, S.P. Ethier, P.H. Johnson, Identification of gene expression profiles that predict the aggressive behavior of breast cancer cells, Cancer Res. 61 (2001) 5168–5178.

[23] M.J. van de Vijver, Y.D. He, L.J. van't Veer, H. Dai, A.A. Hart, D.W. Voskuil, G.J. Schreiber, J.L. Peterse, C. Roberts, M.J. Marton, M. Parrish, D. Atsma, A. Witteveen, A. Glas, L. Delahaye, T. van der Velde, H. Bartelink, S. Rodenhuis, E.T. Rutgers, S.H. Friend, R. Bernards, A gene-expression signature as a predictor of survival in breast cancer, N. Engl. J. Med. 347 (2002) 1999–2009.

[24] L.J. van 't Veer, H. Dai, M.J. van de Vijver, Y.D. He, A.A. Hart, M. Mao, H.L. Peterse, K. van der Kooy, M.J. Marton, A.T. Witteveen, G.J. Schreiber, R.M. Kerkhoven, C. Roberts, P.S. Linsley, R. Bernards, S.H. Friend, Gene expression profiling predicts clinical outcome of breast cancer, Nature 415 (2002) 530–536.

[25] M.F. Press, L. Bernstein, P.A. Thomas, L.F. Meisner, J.Y. Zhou, Y. Ma, G. Hung, R.A. Robinson, C. Harris, A. El-Naggar, D.J. Slamon, R.N. Phillips, J.S. Ross, S.R. Wolman, K.J. Flom, HER-2/neu gene amplification characterized by fluorescence in situ hybridization: poor prognosis in node-negative breast carcinomas, J. Clin. Oncol. 15 (1997) 2894–2904.

[26] D.A. Thompson, R.J. Weigel, hAG-2, the human homologue of the *Xenopus laevis* cement gland gene XAG-2, is coexpressed with estrogen receptor in breast cancer cell lines, Biochem. Biophys. Res. Commun. 251 (1998) 111–116.

[27] R.V. Hoch, D.A. Thompson, R.J. Baker, R.J. Weigel, GATA-3 is expressed in association with estrogen receptor in breast cancer, Int. J. Cancer 84 (1999) 122–128.

[28] M. Nacht, A.T. Ferguson, W. Zhang, J.M. Petroziello, B.P. Cook, Y.H. Gao, S. Maguire, D. Riley, G. Coppola, G.M. Landes, S.L. Madden, S.D.S. Sukumar, Combining serial analysis of gene expression and array technologies to identify genes differentially expressed in breast cancer, Cancer Res. 59 (1999) 5464–5470.

[29] F. Bertucci, R. Houlgatte, A. Benziane, S. Granjeaud, J. Adelaide, R. Tagett, B. Loriod, J. Jacquemier, P. Viens, B. Jordan, D. Birnbaum, C. Nguyen, Gene expression profiling of primary breast carcinomas using arrays of candidate genes, Hum. Mol. Genet. 9 (2000) 2981–2991.

[30] M.C. Nawijn, G.M. Dingjan, R. Ferreira, B.N. Lambrecht, A. Karis, F. Grosveld, H. Savelkoul, R.W. Hendriks, Enforced expression of GATA-3 in transgenic mice inhibits Th1 differentiation and induces the formation of a T1/ST2-expressing Th2-committed T cell compartment in vivo, J. Immunol. 167 (2001) 724–732.
[31] M. Li-Weber, P.H. Krammer, Regulation of IL4 gene expression by T cells and therapeutic perspectives, Nat. Rev. Immunol. 3 (2003) 534–543.
[32] P. Sinha, G. Hutter, E. Kottgen, M. Dietel, D. Schadendorf, H. Lage, Increased expression of annexin I and thioredoxin detected by two-dimensional gel electrophoresis of drug resistant human stomach cancer cells, J. Biochem. Biophys. Methods 37 (1998) 105–116.
[33] D.C. Sgroi, S. Teng, G. Robinson, R. LeVangie, J.R. Hudson Jr., A.G. Elkahloun, In vivo gene expression profile analysis of human breast cancer progression, Cancer Res. 59 (1999) 5656–5661.
[34] J.M. Garcia Pedrero, M.P. Fernandez, R.O. Morgan, A. Herrero Zapatero, M.V. Gonzalez, C. Suarez Nieto, J.P. Rodrigo, Annexin A1 down-regulation in head and neck cancer is associated with epithelial differentiation status, Am. J. Pathol. 164 (2004) 73–79.
[35] C.P. Paweletz, D.K. Ornstein, M.J. Roth, V.E. Bichsel, J.W. Gillespie, V.S. Calvert, C.D. Vocke, S.M. Hewitt, P.H. Duray, J. Herring, Q.H. Wang, N. Hu, W.M. Linehan, P.R. Taylor, L.A. Liotta, M.R. Emmert-Buck, E.F. Petricoin III, Loss of annexin 1 correlates with early onset of tumorigenesis in esophageal and prostate carcinoma, Cancer Res. 60 (2000) 6293–6297.
[36] Y. Hippo, M. Yashiro, M. Ishii, H. Taniguchi, S. Tsutsumi, K. Hirakawa, T. Kodama, H. Aburatani, Differential gene expression profiles of scirrhous gastric cancer cells with high metastatic potential to peritoneum or lymph nodes, Cancer Res. 61 (2001) 889–895.
[37] W. Xin, D.R. Rhodes, C. Ingold, A.M. Chinnaiyan, M.A. Rubin, Dysregulation of the annexin family protein family is associated with prostate cancer progression, Am. J. Pathol. 162 (2003) 255–261.
[38] V. Gerke, S.E. Moss, Annexins: from structure to function, Physiol. Rev. 82 (2002) 331–371.
[39] J. Theobald, A. Hanby, K. Patel, S.E. Moss, Annexin VI has tumour-suppressor activity in human A431 squamous epithelial carcinoma cells, Br. J. Cancer 71 (1995) 786–788.
[40] B.C. Bastian, Annexins in cancer and autoimmune diseases, Cell. Mol. Life Sci. 53 (1997) 554–556.
[41] T.R. Kataoka, A. Ito, H. Asada, K. Watabe, K. Nishiyama, K. Nakamoto, S. Itami, K. Yoshikawa, M. Ito, H. Nojima, Y. Kitamura, Annexin VII as a novel marker for invasive phenotype of malignant melanoma, Jpn. J. Cancer Res. 91 (2000) 75–83.
[42] H. Sato, H. Ogata, L.M. De Luca, Annexin V inhibits the 12-*O*-tetradecanoylphorbol-13-acetate-induced activation of Ras/extracellular signal-regulated kinase (ERK) signaling pathway upstream of Shc in MCF-7 cells, Oncogene 19 (2000) 2904–2912.
[43] A. Chetcuti, S.H. Margan, P. Russell, S. Mann, D.S. Millar, S.J. Clark, J. Rogers, D.J. Handelsman, Q. Dong, Loss of annexin II heavy and light chains in prostate cancer and its precursors, Cancer Res. 61 (2001) 6331–6334.
[44] M. Srivastava, L. Bubendorf, L. Nolan, M. Glasman, X. Leighton, G. Miller, W. Fehrle, M. Raffeld, O. Eidelman, O.P. Kallioniemi, S. Srivastava, H.B. Pollard, ANX7 as a bio-marker in prostate and breast cancer progression, Dis. Markers 17 (2001) 115–120.
[45] J.C. Buckingham, R.J. Flower, Lipocortin 1: a second messenger of glucocorticoid action in the hypothalamo–pituitary–adrenocortical axis, Mol. Med. Today 3 (1997) 296–302.
[46] F. Roviezzo, S.J. Getting, M.J. Paul-Clark, S. Yona, F.N. Gavins, M. Perretti, R. Hannon, J.D. Croxtall, J.C. Buckingham, R.J. Flower, The annexin-1 knockout mouse: what it tells us about the inflammatory response, J. Physiol. Pharmacol. 53 (2002) 541–553.
[47] A.M. Kamal, S.F. Smith, M. De Silva Wijayasinghe, E. Solito, C.J. Corrigan, An annexin 1 (ANXA1)-derived peptide inhibits prototype antigen-driven human T cell Th1 and Th2 responses in vitro, Clin. Exp. Allergy 31 (2001) 1116–1125.
[48] L.C. Alldridge, H.J. Harris, R. Plevin, R. Hannon, C.E. Bryant, The annexin protein lipocortin 1 regulates the MAPK/ERK pathway, J. Biol. Chem. 274 (1999) 37620–37628.
[49] L.C. Alldridge, C.E. Bryant, Annexin 1 regulates cell proliferation by disruption of cell morphology and inhibition of cyclin D1 expression through sustained activation of the ERK1/2 MAPK signal, Exp. Cell Res. 290 (2003) 93–107.
[50] E.A. Ariazi, Y. Satomi, M.J. Ellis, J.D. Haag, W. Shi, C.A. Sattler, M.N. Gould, Activation of the transforming growth factor beta signaling pathway and induction of cytostasis and apoptosis in mammary carcinomas treated with the anticancer agent perillyl alcohol, Cancer Res. 59 (1999) 1917–1928.