Published in final edited form as:

Cancer Res. 2004 November 1; 64(21): 7748–7755. doi:10.1158/0008-5472.CAN-04-1827.

From Mice to Humans: Identification of Commonly Deregulated Genes in Mammary Cancer *via* Comparative SAGE Studies

Yuhui Hu¹, Hongxia Sun¹, Jeffrey Drake¹, Frances Kittrell², Martin C. Abba¹, Li Deng³, Sally Gaddis¹, Aysegul Sahin⁴, Keith Baggerly⁵, Daniel Medina², and C. Marcelo Aldaz¹

¹Department of Carcinogenesis, The University of Texas M.D. Anderson Cancer Center, Science Park-Research Division, Smithville

²Baylor College of Medicine, Department of Molecular and Cellular Biology, Houston

³Rice University, Houston

⁴Department of Pathology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas

⁵Department of Biostatistics, The University of Texas M.D. Anderson Cancer Center, Houston, Texas

Abstract

Genetically engineered mouse mammary cancer models have been used over the years as systems to study human breast cancer. However, much controversy exists on the utility of such models as valid equivalents to the human cancer condition. To perform an interspecies gene expression comparative study in breast cancer we used a mouse model that most closely resembles human breast carcinogenesis. This system relies on the transplant of p53 null mouse mammary epithelial cells into the cleared mammary fat pads of syngeneic hosts. Serial analysis of gene expression (SAGE) was used to obtain gene expression profiles of normal and tumor samples from this mouse mammary cancer model (>300,000 mouse mammary-specific tags). The resulting mouse data were compared with 25 of our human breast cancer SAGE libraries (>2.5 million human breast-specific tags). We observed significant similarities in the deregulation of specific genes and gene families when comparing mouse with human breast cancer SAGE data. A total of 72 transcripts were identified as commonly deregulated in both species. We observed a systematic and significant down-regulation in all of the tumors from both species of various cytokines, including *CXCL1* (*GRO1*), *LIF*, *interleukin 6*, and *CCL2*. All of the mouse and most human mammary tumors also displayed decreased expression of genes known to inhibit cell proliferation, including *NFKBIA* (*IKBα*), *GADD45B*, and *CDKN1A* (*p21*); transcription-related genes such as *CEBP*, *JUN*, *JUNB*, and *ELF1*; and apoptosis-related transcripts such as *IER3* and *GADD34/PPP1R15A*. Examples of overexpressed transcripts in tumors from both species include proliferation-related genes such as *CCND1*, *CKS1B*, and *STMN1* (*oncoprotein 18*); and genes related to other functions such as *SEPWI*, *SDFR1*, *DNCI2*, and *SP110*. Importantly, abnormal

©2004 American Association for Cancer Research.

Requests for reprints: C. Marcelo Aldaz, The University of Texas M.D. Anderson Cancer Center, Department of Carcinogenesis, Science Park-Research Division, Smithville, TX 78957. Phone: 512-237-2403; Fax: 512-237-2475; maldaz@odin.mdacc.tmc.edu.

Supplementary data for this article can be found at http://sciencepark.mdanderson.org/ggeg/sage_proj_8.htm.

expression of several of these genes has not been associated previously with breast cancer. The consistency of these observations was validated in independent mouse and human mammary cancer sets.

This is the first interspecies comparison of mammary cancer gene expression profiles. The comparative analysis of mouse and human SAGE mammary cancer data validates this p53 null mouse tumor model as a useful system closely resembling human breast cancer development and progression. More importantly, these studies are allowing us to identify relevant biomarkers of potential use in human studies while leading to a better understanding of specific mechanisms of human breast carcinogenesis.

INTRODUCTION

The utility of genetically engineered mouse models of mammary cancer as relevant tools for the study of human breast cancer is a topic of constant debate in the breast cancer research community. Some of the major criticisms refer to the fact that most transgenic models based on the overexpression of dominantly acting oncogenes (*e.g.*, *Wnt*, *HRas*, and *ErbB2*) generate tumors that are estrogen receptor-negative and that regress once the oncogenic stimuli are suppressed or withdrawn (1, 2). Phenotypically, these tumors differ considerably from their human counterparts histopathologically and in their progression from early to more advanced lesions (3). Furthermore, tumors produced by these models are mostly diploid and rarely metastasize. These observations differ considerably from the human condition in which most breast cancers are aneuploid, intraductal in origin, estrogen receptor-positive, and regional, and distant metastases is the major complicating factor during tumor progression even from very early stages. Recently, however, a few models have been developed that closely mimic various characteristics of human breast carcinogenesis. The p53 null model of mammary carcinogenesis (4, 5) reproduces many of the critical features of human breast cancer. In this model, epithelial cells from p53 null mice are transplanted into cleared mammary fat pads of syngeneic hosts. More than 60% of these transplants develop invasive mammary adenocarcinomas, and upon hormonal stimulation 100% of the grafts are tumorigenic. Most tumors are intraductal in origin, and premalignant lesions are detectable, such as atypical hyperplasias and *in situ* lesions. Most of these early lesions are aneuploid, express estrogen and progesterone receptors, and ~20% of the invasive adenocarcinomas that finally develop are estrogen receptor-positive (4, 5).

In a previous gene expression study we focused on the early effects that lack of p53 function exerts in normal mouse mammary epithelium *in vivo* (6). In the present study, we chose this p53 null model to obtain a comprehensive gene expression profile of spontaneous mammary tumors by serial analysis of gene expression (SAGE) and more importantly to perform an interspecies comparison with human breast cancer SAGE data generated in our laboratory. The final goal is the identification of commonly deregulated transcripts and pathways in both species that could lead to better understanding of the mechanisms of breast carcinogenesis and to identification of relevant biomarkers.

MATERIALS AND METHODS

Mouse Mammary Samples

Mouse mammary tumors developed spontaneously from intramammary fat pad transplanted p53 null mammary epithelium (4). Three p53 null mammary tumors (bulk samples) were used for generating three independent SAGE libraries (MT1, MT2, and MT3). Thirteen additional p53 null mammary tumors were dissected and snap frozen for RNA isolation and validation studies. As normal control for SAGE and Northern analyses, enriched mammary epithelium (>90% epithelial cells) from p53 wild-type and p53 null transplants were used as described previously (5, 6). To decrease the chances potential artifacts due to sample heterogeneity, each normal sample (MN1 and MN2) represents a pool of mammary epithelial samples from five age-matched separate mice (6).

Human Breast Samples

Snap-frozen samples were obtained from the M.D. Anderson breast cancer tumor bank for total RNA isolation. A total of 25 stage I to stage II breast carcinomas (bulk frozen samples) were used to generate the SAGE libraries. A set of 12 additional human breast tumors was used for real-time quantitative reverse transcription-PCR (RT-PCR) validation studies. Normal mammary epithelial organoids were isolated from four different reduction mammoplasty specimens and used as normal controls for validation studies.

Serial Analysis of Gene Expression Analysis

The 25 human and 5 mouse SAGE libraries were generated following standard procedures as described previously (7, 8) and using commercially available reagents (I-SAGE Kit, Invitrogen, Carlsbad, CA). Sequencing was performed using an ABI 3700 DNA Analyzer (Applied Biosystems, Foster City, CA). Human SAGE libraries were generated at an approximate resolution of 100,000 SAGE tags per library, and mouse SAGE libraries reached ~60,000 tags per library. The additional SAGE data of 4 normal human breast tissue samples used in the final comparisons were downloaded from the Cancer Genome Anatomy Project (CGAP)-SAGE Genie database⁶ (libraries generated at the laboratory of Dr. Kornelia Polyak, Dana-Farber Cancer Institute, Boston, MA). The libraries identified as Human Normal #1 and #2 were generated from human luminal mammary epithelial samples BerEp4 antibody-purified cells, whereas libraries #3 and #4 were generated from human breast epithelial organoids (9, 10).

Northern Blot Analyses

Total RNA from mouse samples was isolated, gene probes generated, and hybridization performed as described previously using standard procedures (8).

Real-Time Quantitative RT-PCR Analyses

Total RNA from human samples was isolated, and cDNA was synthesized following standard procedures. Primers and probes were either designed using Primer Express 1.5

⁶Internet address: <http://cgap.nci.nih.gov/SAGE>.

software (Applied Biosystems) or directly obtained from Applied Biosystems (Assays-on-Demand, Gene Expression Products). All of the real time RT-PCR reactions were performed using the TaqMan PCR Core Reagents kit and the ABI Prism 7700 Sequence Detection system (Applied Biosystems). Experiments were performed in triplicate for each data point. Amplification of 18S rRNA was used as the control for normalization.

Bioinformatics and Statistical Analyses

SAGE tag extraction from sequencing files was performed by using the SAGE2000 software version 4.0 (a kind gift of Dr. Kenneth Kinzler, John Hopkins University, Baltimore, MD). SAGE data management, tag to gene matching, as well as additional gene annotations and links to publicly available resources such as Gene Ontology (GO), UniGene, and LocusLink, were performed using a suite of web-based SAGE library tools developed by us.⁷ The large number of human libraries allowed us to use robust rank-based methods (Mann-Whitney test) to assess differential expression (4 normals *versus* 25 breast cancers). Due to the smaller number of libraries involved, this was not possible with the mouse libraries (2 normals *versus* 3 tumors). Here, a parametric model, overdispersed logistic regression, was used to assess differential expression between the groups. This model was chosen to capture the two types of variability that we have noticed with SAGE libraries: Poisson sampling variability within a library, and natural variation in the true proportion of a tag between different individuals in the same group. The logistic model captures the first, overdispersion the second, and using both groups provides a pooled estimate of the SD associated with the difference.⁸ This was implemented in R using the glm function with a quasi-binomial link. All of the proportions (tag count/library size) were slightly modified (tag count + 1)/(library size + 1) to avoid convergence problems with the logistic fits. *Ps* for the logistic coefficients were computed using likelihood ratio tests. For both statistical methods used, significant differences were considered at $P < 0.05$.

To enable visualization and illustration of the commonly deregulated genes in mouse and human breast carcinomas, we used the TIGR MultiExperiment Viewer (MeV 2.2) software. This tool was used for normalization and average clustering of the SAGE based on the fold change of tag counts for each transcript comparing tumor to normal in the mouse and human libraries, respectively. To this end we calculated the ratio of tag counts for each transcript in each tumor over the average representation of the same individual tags in the corresponding normal libraries. The calculated ratios were subjected to a log2 transformation before the clustering analysis. Two-way (by gene and by tumor sample) hierarchical clustering was used to examine the relationships among the human breast libraries. Mouse tumor clustering was based on one-way hierarchical clustering (by tumor sample), maintaining the same order of sorted genes as in the human breast cancer analysis.

For automated functional annotation and classification of genes of interest based on GO terms, we used the EASE software (11) available at the Database for Annotation, Visualization and Integrated Discovery (DAVID) (12).⁹ The EASE software calculates

⁷Internet address: http://spi.mdacc.tmc.edu/bitools/about/sage_lib_tool.html.

⁸Keith Baggerly, unpublished observations.

⁹Internet address: <http://david.niaid.nih.gov/david>.

over-representation of specific GO terms with respect to the total number of genes assayed and annotated. Statistical measures of specific enrichment of GO terms are determined by an EASE score. The EASE score is a conservative adjustment of the Fisher exact probability that weights significance in favor of themes supported by more genes and is calculated using the Gaussian hypergeometric probability distribution that describes sampling without replacement from a finite population (11). This allows one to identify biological themes within a specific list of EASE analyzed genes.

All of the raw SAGE data reported as Supplementary Tables in this article are publicly available.

RESULTS AND DISCUSSION

Primary Mouse SAGE Analyses

We generated five mouse SAGE libraries from three p53 null mammary tumors (MT1, MT2, and MT3) and two samples from p53 wild-type and p53 null mammary transplants representing normal mouse mammary epithelium (MN1, MN2 respectively). This resulted in the sequencing of ~300,000 tags (60,000 tags per sample), thus monitoring the behavior of a total of 21,850 transcript tags.

In a previous study we observed that the p53 wild-type and null normal mammary epithelial samples were very comparable having the expression of only very few genes (0.2%) significantly different between the referred two normal samples (6).

Using the overdispersed logistic regression test, we identified 499 differentially expressed genes ($P < 0.05$) when comparing SAGE data from the p53 null mammary tumors *versus* normal mammary epithelium (see Supplementary Table 1).

To validate some of the most significant observations, we performed Northern analyses of various relevant transcripts. As shown in Fig. 1A and as predicted by SAGE, we confirmed the consistent increase in expression of several genes in most mouse mammary tumors, including cell proliferation related transcripts such as *CDC28 protein kinase 1 (Cks1)*, *stathmin 1 (Stmn1 or Op18, oncoprotein 18)*, and *CDK2 (cyclin-dependent kinase 2)-associated protein 1 (Cdkap1)*. Other validated up-regulated genes include *ring-box 1 (Rbx1)*, *protein phosphatase 1, catalytic subunit, γ isoform (Ppp1cc)*, *stromal cell derived factor receptor 1 (Sdfr1)*, and *selenoprotein W 1 (Sepw1)*.

Examples of SAGE predicted transcripts with decrease expression in tumors are shown in Fig. 1B. As can be observed, a significant loss of expression of several transcripts was confirmed in all of the p53 null mammary tumors, including transcription factors *Fos (FBJ osteosarcoma oncogene)*, *c-Jun*, and *Elf1 (E74-like factor 1)*; genes inhibiting cell proliferation such as *nuclear factor of κ light chain gene enhancer in B-cells inhibitor, α [*Nfkbia/inhibitor of nuclear factor- κ B α (I κ B α)*, *nuclear factor κ B (NF κ B) inhibitor*]; proapoptosis regulator *immediate early response gene 3 (Ier3 or Iex-1, immediate early response gene X-1)*; various cytokines such as *leukemia inhibitory factor (Lif)* and *chemokine (C-X-C motif) ligand 1 (Cxcl1 or Gro1 oncogene)*; *chemokine (C-C motif) ligand**

7 (*Ccl7*) and tumor necrosis factor, α -induced protein 2 (*Tnfaip2*); and RIKEN cDNA 2310004N11 gene (homologue of *MGC4796*, *ser/thr like kinase*).

Comparative Mouse-Human Mammary Cancer SAGE Studies

We performed a SAGE interspecies (mouse-human) comparative analysis of genes commonly deregulated in mammary cancer. To this end, every statistically significant, differentially expressed gene found in the mouse SAGE experiments (Supplementary Table 1) was investigated for its status of expression in a set of 25 human SAGE libraries, which, in turn, were compared with four normal human breast epithelial libraries.

These 25 stage I to II human breast cancer SAGE libraries were generated at a resolution of 100,000 transcript tags per library, representing >2.5 million tags. This allowed us to monitor the pattern of expression of ~35,000 human transcript tags.

To compare the SAGE data from the mouse and human mammary cancers, we used National Center for Biotechnology Information LocusLink and HomologGene databases to identify the human homo-logue genes for the commonly deregulated mouse transcripts (Supplementary Table 1). Human serial analysis of gene expression tags to gene matches were determined by the CGAP SAGE Genie database. Only reliable (unambiguous) SAGE tags matching mouse/human genes were considered in our analyses.

Among the 499 differentially expressed mouse transcripts (Supplementary Table 1), 43 genes have no known human homologues (8.6%), and 12 transcripts have no reliable human tag matches. Thus, a total of 444 genes were reliably represented in the SAGE libraries from both species. The comparative behavior of these 444 transcripts was analyzed in our 25 human breast cancer SAGE libraries relative to the expression observed in 4 normal human breast epithelial SAGE libraries.

Interestingly, we determined that 72 genes were commonly deregulated in tumors from both species; in other words, 72 transcripts that behaved in the same fashion both in mouse mammary tumors and human breast cancers. Fig. 2 depicts the expression level fold changes of these 72 genes when compared with the corresponding normal mammary epithelial samples in both species. The raw tag values for each of these transcripts in both species and for each normal and tumor sample is available as Supplementary Table 2.

Nineteen of the 72 transcripts were overexpressed in mouse and human tumors when compared with their normal counterparts, whereas 53 genes were underexpressed in tumors from both species relative to normal mammary epithelium.

Table 1 shows the 72 genes classified by their main functional roles based on literature searches and information obtained from LocusLink. Nevertheless, these 72 genes were additionally analyzed for over-representation of functional categories by the EASE software available at DAVID on-line resource (see Materials and Methods; refs. 11, 12). The EASE software performs the automated annotation and classification based on GO terms of the genes of interest (genes shown in Table 1, in our case) to biological themes and calculates their representation relative to the whole human breast cancer SAGE data set (see

Supplementary Table 3). This analysis largely agreed with the classification categories shown in Table 1. Among the statistically significant over-represented categories under “Molecular Function,” we found cytochrome oxidases, genes with cytokine, growth factor activity, and transcriptional regulator activity. Categories of genes found in the extracellular space are also highly represented (EASE score 0.003). Among the gene categories falling under Biological Processes, genes related to cell proliferation, protein kinase cascades, and cell death received highly statistically significant EASE scores (0.002, 0.002, and 0.009, respectively; see Supplementary Table 3).

To validate these SAGE comparative observations in additional human breast carcinomas, we performed real-time quantitative RT-PCR analyses of representative transcripts in an independent set of 12 invasive ductal breast carcinomas compared with normal human breast epithelial organoids representing four different donors. We analyzed expression levels for a total of 9 transcripts from the list of 72. As representative of up-regulated transcripts in breast carcinomas, we validated the overexpression of genes: *CKS1B** (*CDC28 protein kinase 1*), *STMN1* (*Stathmin 1, oncoprotein 18*), *DNCI2** (*Dynein, cytoplasmic, intermediate chain 2*), *SDFR1** (*Stromal cell derived factor receptor 1*), *SEPWI** (*Selenoprotein W1*), and *SP110** (*SP110 nuclear body protein*). The down-regulated transcripts validated were those representing genes *IER3* (*immediate early response 3*), *TNFAIP2** (*tumor necrosis factor, α -induced protein 2*), and *MGC4796** (*Ser/Thr-like kinase*). All of the transcripts analyzed behaved as predicted by the SAGE observations (Fig. 3). It is worth mentioning that to the best of our knowledge, of the validated 9 transcripts, those marked with an asterisk (*) have not been associated previously with breast cancer.

Specific Transcripts Commonly Deregulated in Human and Mouse Mammary Cancers

Cytokine Down-Regulation in Mammary Carcinomas—Among transcripts with decreased expression in mammary tumors from both species, we identified several cytokines. This is in agreement with similar observations in other, more limited human breast cancer SAGE studies (9, 10). Here we not only observed the significant down-regulation of such cytokines in each of the human breast carcinomas analyzed, but strikingly, the exact same phenomenon was observed in every mouse mammary tumor analyzed as compared with normal epithelium (Fig. 1B and Fig. 2)

The cytokine transcripts more consistently down-regulated were those encoding for *LIF* (*leukemia inhibitory factor*), *interleukin 6 (IL-6)*, *small inducible cytokine A2 (CCL2)*, and *CXCL1 (GRO1)*; the latter being several hundred-fold down-regulated. All of these transcripts are highly expressed in normal mammary epithelia from both mouse and humans. Intriguingly, the exact functional role for these abundantly secreted cytokines in normal mammary epithelium is unclear at this point.

LIF is a glycoprotein highly conserved between species and originally identified as an IL-6-like cytokine that induces macrophage differentiation (13). Numerous additional functional roles have been assigned to this molecule. A variety of transcriptional pathways have been reported to be activated by LIF binding to its receptor, thus affecting differentiation and proliferation in various cell types. A progressive increase in LIF expression in mouse mammary epithelium after weaning has been reported recently, suggesting that LIF plays an

important role in apoptosis induction during the first stage of mammary gland involution, an event at least partially mediated by Stat3 activation, a known downstream signal transduction target of LIF (14, 15). Interestingly, it was also suggested that LIF exerts growth-inhibitory effects on normal and malignant breast epithelial cells (16). However, the opposite effect in breast cancer cell lines was also reported by other investigators (17).

In our SAGE studies, *LIF* expression was significantly decreased in both human and mouse mammary tumors. As seen in Fig. 1B, these observations were validated by Northern analysis of 15 mouse mammary tumors that showed an almost complete lack of *LIF* expression comparing with normal mammary epithelial cells. On the basis of these observations and the aforementioned physiologic roles of LIF in apoptosis induction and potential growth-controlling functions in mammary epithelial cells, we speculate that the significant down-regulation observed could be associated to an escape of tumor cells from the constraints of such growth regulatory mechanisms.

IL-6 is also a multifunctional cytokine very similar to LIF, known to affect the proliferation of epithelial cells, and its effect depends on the cell type. The cytostatic effects of IL-6 on breast carcinoma cells, its normal production and secretion by normal mammary epithelial cells, and its loss of expression in breast carcinomas has been studied by various groups (18, 19). IL-6 in particular was shown to inhibit the growth of estrogen receptor-positive carcinoma cells, and its expression correlates with estrogen receptor status (20, 21). Nevertheless, the true role of this cytokine in breast carcinogenesis is still unclear and a matter of controversy.

CXCL1 is also known as GRO1 oncogene or melanoma growth stimulating activity, α (MGSA). GRO1 was originally identified as a mitogenic substance secreted by melanoma cells (22). CXC chemokines are multifunctional molecules that, other than their role in immunomodulation, have been associated with invasion, tumor cell migration, and neovascularization. However, other than the transformation-related role for CXCL1/GRO1 in melanoma (23), there is a paucity of clear examples or studies demonstrating a protumorigenic role for CXCL1/GRO1 in carcinoma development and none in breast cancer. Furthermore, in early studies, it was observed that *GRO1* was highly expressed by normal cultured human mammary epithelial cells, whereas most carcinoma cell lines analyzed did not express this cytokine (24).

In our studies, we observed a significant decrease in *CXCL1/GRO1* expression in breast carcinomas when compared with normal mammary epithelium in both species and in all of the carcinomas (Fig. 1B and Fig. 2). Interestingly, we also observed that *CXCL1/GRO1* expression increases considerably in normal mouse mammary epithelium during the first stage of mammary gland involution after weaning, shadowing the previously mentioned expression behavior of *LIF* (data not shown). On the basis of these and previously discussed observations, we can speculate that the role of CXCL1/GRO1 in breast epithelium appears to be related to a transformation-protective role, perhaps related to regulation of apoptosis and cell proliferation.

CCL2, or monocyte chemoattractant protein 1 (MCP-1), is also structurally related to the CXC subfamily of cytokines. The role of this cytokine in breast cancer development is also controversial. It has been shown recently that estrogen decreases the levels of this chemokine in mouse mammary tissues, an event apparently contributing to mammary tumor development (25). Similar to our observations in breast cancer and in contrast to other neoplasias, it was observed that MCP-1 is not expressed in the carcinomatous tissue of prostate (26). Furthermore, it was concluded recently that CCL2/MCP1 could be a relevant negative regulator of pancreatic cancer progression (27). These observations allow us to speculate that CCL2 could play a similar transformation protective role in mammary epithelium.

In summary, the down-regulation of specific cytokines in 100% of mouse and human tumors (*LIF*, *IL6*, *CXCL1*, and *CCL2*) appears to point to a very relevant and evolutionary conserved event related to mammary carcinogenesis.

Transcription Factors—In this study, we also detected 13 transcription factors deregulated both in mouse and human mammary tumors (Table 1).

ELF1 (*E74-like factor 1*) was found highly expressed in normal mammary epithelium both from mouse and humans (Fig. 1B and Supplementary Table 2). This observation is in agreement with a very recent report that analyzed expression of various Ets transcription factors in mammary tissues (28). *ELF1* is an Ets-related transcription factor and is known to be a key component in the transcriptional program during hematopoietic stem cell development (29). We observed a very significant down-regulation in expression of this transcription factor in all three of the mouse tumor SAGE libraries, in most mouse tumors analyzed by Northern (Fig. 1B), and in 13 of 25 human breast carcinoma serial analysis of gene expression libraries (Table 1, Supplementary Table 2, and Fig. 2). The decreased expression of *ELF1* observed in breast cancer appears contrary to most Ets transcription factors of which the overexpression is usually associated with malignant processes.

It has been shown that impairment of AP-1 transcription complex activity leads to failure of involution in the mouse mammary gland implying that at least postlactational programmed cell death depends on functional AP-1 (30). It was also shown that breast cancer cells have lower AP-1 activity (c-Fos and c-Jun levels) than normal mammary epithelial cells *in vitro* (31). Although many reports have been written on AP-1 family members in breast cancer biology, therapeutic response, and prognosis, the picture on the true role of each of these factors is far from clear (32). Intriguingly, our SAGE study with p53 null mammary tumors revealed a significant decrease of *Fos* expression in all of the tumors. Whereas the normal mammary epithelial samples expressed on average 70 SAGE tags representing Fos, the three mammary tumors expressed 5, 2, and 3 tags (per 60,000 total tags, MT1, MT2, and MT3 respectively, see Supplementary Table 1). Furthermore, we also observed that other AP-1 members (*e.g.*, *Jun* and *Junb*) were systematically down-regulated in all three of the p53 null mammary tumors. These observations were additionally validated in 15 additional mouse tumors by Northern analysis (Fig. 1B). It is known that Fos and Junb transcription can be regulated by p53. Therefore, the systematic down-regulation of AP-1 family members may be due to p53 loss and could be a characteristic intrinsic to this mouse model. However,

when comparing with human breast cancer SAGE data, we detected that Jun and Junb expression was also decreased in the majority of stage I and II breast tumors (18 of 25 and 17 of 25, respectively) when comparing with normal epithelium (Table 1 and Supplementary Table 2). On the other hand, we did not observe the same phenomenon affecting *FOS* in these human breast carcinomas. Therefore, we can conclude that at least some members that could form active AP-1 transcription factor complexes are down-regulated in breast carcinomas from both species and that such decrease in expression may be important in the carcinogenesis process.

NF κ B Pathway-Associated Transcripts—It has been postulated by various authors that activation of the NF κ B pathway is a very important event in breast cancer development (reviewed in ref. 33). In support of such a hypothesis and previous observations, in our SAGE studies we observed that several genes that appear to play inhibitory NF κ B roles were greatly down-regulated not only in mouse mammary tumors but also in human breast cancers. For instance, *NFKBIA/I κ B α* , the major inhibitor of NF κ B, was observed significantly down-regulated in 100% of mouse and human breast carcinomas (Fig. 1B, Table 1, and Supplementary Tables 1 and 2). Secretory leukocyte protease inhibitor (SLP1) was shown to prevent NF κ B activation by inhibiting the degradation of I κ B α and I κ B β , and we observed decreased expression of this gene in all of the mouse tumors and in 21 of 25 human breast carcinomas (Table 1; ref. 34). Human breast cancer data also showed that all of the tumors displayed lower expression of the diphtheria toxin receptor/heparin-binding epidermal growth factor-like growth factor (*DTR*; Table 1). This protein was demonstrated to suppress NF κ B activation by inhibiting I κ B kinase activation and I κ B degradation (35).

One of the roles of clusterin (CLU) is to inhibit NF κ B signaling through stabilization of I κ Bs (36). We observed decreased expression of the transcript encoding clusterin in all of the mouse tumors and 44% of human breast carcinomas. We also found down-regulation in every tumor of the transcript for *serine/threonine kinase MGC4796*, a protein reported to inhibit tumor necrosis factor-induced NF κ B activation and p53-induced transcription (37).

The significant decrease in expression of the aforementioned transcripts, all of which are related in different ways to preventing or decreasing NF κ B activity, strongly supports the hypothesis that high NF κ B activity is of critical relevance in human breast cancer (reviewed in ref. 33).

Cell Proliferation and Apoptosis—Not surprisingly our SAGE comparative study detected the common overexpression of positive regulators of cell proliferation such as *CCND1*, whereas negative regulators of proliferation-like *growth arrest and DNA-damage-inducible β* (*GADD45B*) and *cyclin-dependent kinase inhibitor 1A* (*CDKN1A/p21*) were found underexpressed in all three of the mouse tumors and most of human breast tumors (Table 1 and Fig. 2). Two genes regulating apoptosis, *IER3* and *protein phosphatase 1, regulatory (inhibitor) subunit 15A* (*PPP1R15A/GADD34*), also showed decreased expression in mouse and human tumors.

Interestingly, we detected the overexpression of *CDC28 protein kinase regulatory subunit 1B* (*CKS1B*) in human and mouse mammary tumors. These observations were validated

both with mouse and human carcinoma samples (Fig. 1A and Fig. 3). CKS1B functions as an important adaptor of SCF^{Skp2} ubiquitin ligase and facilitates SCF^{Skp2} targeting of the cell proliferation inhibitor p27 (Kip1) for ubiquitination and subsequent degradation (38). It was also suggested that CKS1B may be involved in p21 degradation in a similar fashion (39). Recently, overexpression of CKS1B has been observed in human non-small lung cancer and in gastric and colorectal carcinomas (40–42). Here we provide evidence for the first time that increased CKS1B expression is frequent in mammary carcinomas from both species, human and mouse.

The IER3 (immediate early response 3) or IEX-1 (immediate early response gene X-1) was suggested to play a role in regulation of cell growth and apoptosis. Its expression can be regulated by multiple transcription factors, including p53, NFκB/rel, c-Myc and Sp-1 (reviewed in ref. 43). We observed a significant down-regulation on the expression of IER3 in human and mouse mammary tumors (Fig. 1B, Fig. 3, Table 1, and Supplementary Tables). This is consistent with previous studies reporting that IER3 inhibited proliferation of breast cancer cells (43). Our finding additionally suggests that IER3 may be an important and common player in mammary carcinogenesis.

Various other interesting genes involved in signaling transduction and with other functional roles were found commonly deregulated in mouse and human breast tumors (Table 1). Interestingly, as mentioned previously, we identified and validated several transcripts of which the deregulation was never reported previously to be associated to human breast cancer such as *CKS1B*, *DNCI2*, *SDRF1*, *SEPW1*, *SP110*, *TNFAIP2*, and *Ser/Thr-like kinase MGC4796*.

In previous work, Desai *et al.* (44), using cDNA microarrays, compared gene expression profiles of mammary tumors generated in various transgenic mammary cancer models. They observed that the transgene-generated tumors shared similarities in their expression profiles as compared with normal mammary gland. However, major differences were observed that depended on the distinct oncogene-specific patterns of gene expression of each model. Very few similarities of commonly deregulated genes across all of the models were observed that were comparable with similar changes observed in human breast cancers (44).

On the basis of the results here reported and those previous observations we concluded that the p53 null model more closely resembles human breast cancer than the commonly used oncogene-driven transgenic models of mammary cancer.

In summary, to the best of our knowledge, this is the first mouse-human SAGE comparative study of mammary cancer gene expression profiles, providing information on evolutionary conserved pathway abnormalities involved in this malignancy. We observed that deregulation of genes related to the control of cell proliferation, differentiation, apoptosis, and mammary gland development is frequent in mammary carcinomas from both species. In addition, this comparative analysis validates the p53 mouse tumor model as a useful system that quite closely resembles human breast cancer development and progression. The significant similarities here described at the gene expression level add to the earlier mentioned biological and pathological commonalities. More importantly, these studies allow

us to identify relevant biomarkers of potential use in human studies while eventually leading to a better understanding of specific mechanisms of human breast carcinogenesis. Future studies are needed to focus on the functional characterization and mechanistic aspects of the most promising targets from the list of commonly deregulated genes across mammary tumors in this interspecies comparison.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support: NIH National Cancer Institute Grants U19 CA84978 (C. M. Aldaz), U01 CA84243 (D. Medina), and center grant ES-07784; L. Deng is supported by a fellowship from the W.M. Keck Foundation to the Gulf Coast Consortia.

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.

REFERENCES

1. Moody SE, Sarkisian CJ, Hahn KT, et al. Conditional activation of Neu in the mammary epithelium of transgenic mice results in reversible pulmonary metastasis. *Cancer Cell*. 2002; 2:451–61. [PubMed: 12498714]
2. Gunther EJ, Moody SE, Belka GK, et al. Impact of p53 loss on reversal and recurrence of conditional Wnt-induced tumorigenesis. *Genes Dev*. 2003; 17:488–501. [PubMed: 12600942]
3. Cardiff RD, Anver MR, Gusterson BA, et al. The mammary pathology of genetically engineered mice: the consensus report and recommendations from the Annapolis meeting. *Oncogene*. 2000; 19:968–88. [PubMed: 10713680]
4. Jerry DJ, Kittrell FS, Kuperwasser C, et al. A mammary-specific model demonstrates the role of the p53 tumor suppressor gene in tumor development. *Oncogene*. 2000; 19:1052–8. [PubMed: 10713689]
5. Medina D, Kittrell FS, Shepard A, Contreras A, Rosen JM, Lydon J. Hormone dependence in premalignant mammary progression. *Cancer Res*. 2003; 63:1067–72. [PubMed: 12615724]
6. Aldaz CM, Hu Y, Daniel R, Gaddis S, Kittrell F, Medina D. Serial analysis of gene expression in normal p53 null mammary epithelium. *Oncogene*. 2002; 21:6366–76. [PubMed: 12214277]
7. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. *Science*. 1995; 270:484–7. [PubMed: 7570003]
8. Charpentier AH, Bednarek AK, Daniel RL, et al. Effects of estrogen on global gene expression: identification of novel targets of estrogen action. *Cancer Res*. 2000; 60:5977–83. [PubMed: 11085516]
9. Porter DA, Krop IE, Nasser S, et al. A SAGE (serial analysis of gene expression) view of breast tumor progression. *Cancer Res*. 2001; 61:5697–702. [PubMed: 11479200]
10. Porter D, Lahti-Domenici J, Keshaviah A, et al. Molecular markers in ductal carcinoma in situ of the breast. *Mol Cancer Res*. 2003; 1:362–75. [PubMed: 12651909]
11. Hosack DA, Dennis G, Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. *Genome Biol*. 2003; 4:R70. [PubMed: 14519205]
12. Dennis G, Sherman BT, Hosack DA, et al. DAVID: database for annotation, visualization and integrated discovery. *Genome Biol*. 2003; 4:P3. [PubMed: 12734009]
13. Moreau JF, Donaldson DD, Bennett F, Witek-Giannotti J, Clark SC, Wong GG. Leukaemia inhibitory factor is identical to the myeloid growth factor human inter-leukin for DA cells. *Nature*. 1988; 336:690–2. [PubMed: 3143918]

14. Schere-Levy C, Buggiano V, Quaglini A, et al. Leukemia inhibitory factor induces apoptosis of the mammary epithelial cells and participates in mouse mammary gland involution. *Exp Cell Res.* 2003; 282:35–47. [PubMed: 12490192]
15. Kritikou EA, Sharkey A, Abell K, et al. A dual, non-redundant, role for LIF as a regulator of development and STAT3-mediated cell death in mammary gland. *Development.* 2003; 130:3459–68. [PubMed: 12810593]
16. Grant SL, Douglas AM, Goss GA, Begley CG. Oncostatin M and leukemia inhibitory factor regulate the growth of normal human breast epithelial cells. *Growth Factors.* 2001; 19:153–62. [PubMed: 11811789]
17. Kellokumpu-Lehtinen P, Talpaz M, Harris D, Van Q, Kurzrock R, Estrov Z. Leukemia-inhibitory factor stimulates breast, kidney and prostate cancer cell proliferation by paracrine and autocrine pathways. *Int J Cancer.* 1996; 66:515–9. [PubMed: 8635867]
18. Krueger J, Ray A, Tamm I, Sehgal PB. Expression and function of interleukin-6 in epithelial cells. *J Cell Biochem.* 1991; 45:327–34. [PubMed: 2045425]
19. Basolo F, Conaldi PG, Fiore L, Calvo S, Toniolo A. Normal breast epithelial cells produce interleukins 6 and 8 together with tumor-necrosis factor: defective IL6 expression in mammary carcinoma. *Int J Cancer.* 1993; 55:926–30. [PubMed: 8253529]
20. Chiu JJ, Sgagias MK, Cowan KH. Interleukin 6 acts as a paracrine growth factor in human mammary carcinoma cell lines. *Clin Cancer Res.* 1996; 2:215–21. [PubMed: 9816109]
21. Fontanini G, Campani D, Roncella M, et al. Expression of interleukin 6 (IL-6) correlates with oestrogen receptor in human breast carcinoma. *Br J Cancer.* 1999; 80:579–84. [PubMed: 10408869]
22. Richmond A, Balentien E, Thomas HG, et al. Molecular characterization and chromosomal mapping of melanoma growth stimulatory activity, a growth factor structurally related to beta-thromboglobulin. *EMBO J.* 1988; 7:2025–33. [PubMed: 2970963]
23. Dhawan P, Richmond A. Role of CXCL1 in tumorigenesis of melanoma. *J Leukoc Biol.* 2002; 72:9–18. [PubMed: 12101257]
24. Anisowicz A, Zajchowski D, Stenman G, Sager R. Functional diversity of gro gene expression in human fibroblasts and mammary epithelial cells. *Proc Natl Sci USA.* 1988; 85:9645–9.
25. Fanti P, Nazareth M, Bucelli R, et al. Estrogen decreases chemokine levels in murine mammary tissue: implications for the regulatory role of MIP-1 alpha and MCP-1/JE in mammary tumor formation. *Endocrine.* 2003; 22:161–8. [PubMed: 14665721]
26. Mazzucchelli L, Loetscher P, Kappeler A, et al. Monocyte chemoattractant protein-1 gene expression in prostatic hyperplasia and prostate adenocarcinoma. *Am J Pathol.* 1996; 149:501–9. [PubMed: 8701989]
27. Monti P, Leone BE, Marchesi F, et al. The CC chemokine MCP-1/CCL2 in pancreatic cancer progression: regulation of expression and potential mechanisms of antitumorigenic activity. *Cancer Res.* 2003; 63:7451–61. [PubMed: 14612545]
28. Galang CK, Muller WJ, Foos G, Oshima RG, Hauser CA. Changes in the expression of many Ets family transcription factors and of potential target genes in normal mammary tissue and tumors. *J Biol Chem.* 2004; 279:11281–92. [PubMed: 14662758]
29. Gottgens B, Nastos A, Kinston S, et al. Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *EMBO J.* 2002; 21:3039–50. [PubMed: 12065417]
30. Feng Z, Marti A, Jehn B, Altermatt HJ, Chicaiza G, Jaggi R. Glucocorticoid and progesterone inhibit involution and programmed cell death in the mouse mammary gland. *J Cell Biol.* 1995; 131:1095–103. [PubMed: 7490285]
31. Smith LM, Birrer MJ, Stampfer MR, Brown PH. Breast cancer cells have lower activating protein 1 transcription factor activity than normal mammary epithelial cells. *Cancer Res.* 1997; 57:3046–54. [PubMed: 9230221]
32. Gee JM, Barroso AF, Ellis IO, Robertson JF, Nicholson RI. Biological and clinical associations of c-jun activation in human breast cancer. *Int J Cancer.* 2000; 89:177–86. [PubMed: 10754497]
33. Cao Y, Karin M. NF-kappaB in mammary gland development and breast cancer. *J Mammary Gland Biol Neoplasia.* 2003; 8:215–23. [PubMed: 14635796]

34. Taggart CC, Greene CM, McElvaney NG, O'Neill S. Secretory leucoprotease inhibitor prevents lipopolysaccharide-induced IkappaBalpha degradation without affecting phosphorylation or ubiquitination. *J Biol Chem.* 2002; 277:33648–53. [PubMed: 12084717]
35. Mehta VB, Besner GE. Inhibition of NF-kappa B activation and its target genes by heparin-binding epidermal growth factor-like growth factor. *J Immunol.* 2003; 171:6014–22. [PubMed: 14634113]
36. Santilli G, Aronow BJ, Sala A. Essential requirement of apolipoprotein J (clusterin) signaling for IkappaB expression and regulation of NF-kappaB activity. *J Biol Chem.* 2003; 278:38214–9. [PubMed: 12882985]
37. Huang J, Teng L, Liu T, et al. Identification of a novel serine/threonine kinase that inhibits TNF-induced NF-kappaB activation and p53-induced transcription. *Biochem Biophys Res Commun.* 2003; 309:774–8. [PubMed: 13679039]
38. Bartek J, Lukas J. p27 destruction: Cks1 pulls the trigger. *Nat Cell Biol.* 2001; 3:E95–8. [PubMed: 11283628]
39. Bornstein G, Bloom J, Sitry-Shevah D, Nakayama K, Pagano M, Hershko A. Role of the SCFSkp2 ubiquitin ligase in the degradation of p21Cip1 in S phase. *J Biol Chem.* 2003; 278:25752–7. [PubMed: 12730199]
40. Inui N, Kitagawa K, Miwa S, et al. High expression of Cks1 in human non-small cell lung carcinomas. *Biochem Biophys Res Commun.* 2003; 303:978–84. [PubMed: 12670508]
41. Masuda TA, Inoue H, Nishida K, et al. Cyclin-dependent kinase 1 gene expression is associated with poor prognosis in gastric carcinoma. *Clin Cancer Res.* 2003; 9:5693–8. [PubMed: 14654553]
42. Shapira M, Ben-Izhak O, Bishara B, et al. Alterations in the expression of the cell cycle regulatory protein cyclin kinase subunit 1 in colorectal carcinoma. *Cancer.* 2004; 100:1615–21. [PubMed: 15073847]
43. Wu MX. Roles of the stress-induced gene IEX-1 in regulation of cell death and oncogenesis. *Apoptosis.* 2003; 8:11–8. [PubMed: 12510147]
44. Desai KV, Xiao N, Wang W, et al. Initiating oncogenic event determines gene-expression patterns of human breast cancer models. *Proc Natl Acad Sci USA.* 2002; 99:6967–72. [PubMed: 12011455]

Fig. 1.

Representative Northern blot analysis of commonly deregulated genes in p53 null mouse mammary tumors. *A*, transcripts with increased expression in p53 null mammary tumors as compared with normal mouse mammary epithelium. *B*, significantly down-regulated transcripts in p53 null mammary tumors relative to normal mouse mammary epithelium.

*2310004N11Rik, homologue of human *MGC4796* gene.

Fig. 2.

Heat map of commonly deregulated genes in mouse and human tumors displayed in fold change relative to normal tissue counterparts. *Color scale* at the *bottom* depicts the approximate fold change in expression for each transcript and library relative to normal mammary epithelium: negative fold change (transcripts with decreased expression in tumors) is represented in *green*, and positive fold change (transcripts with over-expression in tumors) is represented in *red*.

Fig. 3.

Real-time PCR validation of 9 deregulated genes in 12 additional human breast carcinomas. 18S rRNA was used as control. *Bar charts* were plotted based on the mean values $\pm$ 2 SEM of the ratio between target and control genes in normal human breast epithelium and invasive carcinomas. Values shown in the *top left corner* of each plot represent the equivalent fold change between the mean of tumors and normal tissue values. *, expression not detected in normal breast tissue.

Table 1

Commonly deregulated genes in mouse and human mammary tumors

Human gene name		Mouse tag	Mouse LocusID	Human tag	Human LocusID	Frequency in human tumors [*]
Regulation of cell proliferation						
<i>CCND1</i> (Cyclin D1)	↑	GTCCAGGAAA	12443	AAAGTCTAGA	595	20/25 (80%)
<i>CKS1B</i> (CKS1)	↑	TTACAAGCCT	54124	TTAAAAGCCT	1163	18/25 (72%)
<i>STMN1</i> (oncoprotein 18)	↑	AGAAGGAGGT	16765	AAGCTGAGGT	3925	9/25 (36%)
<i>NFKB1A</i> (IKBα)	↓	TGCTTAAAAG	18035	TAACAGCCAG	4792	25/25 (100%)
<i>GADD45B</i> (MYD118)	↓	CAGAGGCTGG	17873	AACTCCCAGT	4616	25/25 (100%)
<i>CDKN1A</i> (p21)	↓	TATTGTGGCT	12575	TGTCCTGGTT	1026	21/25 (84%)
Apoptosis related						
<i>IER3</i> (IEX-1)	↓	GATTGTCAGA	15937	ACCATCCTGC	8870	25/25 (100%)
<i>PPP1R15A</i> (GADD34)	↓	GGCACGCCTT	17872	ATCCGGACCC	23645	25/25 (100%)
<i>CLU</i> (Clusterin)	↓	TCTCCAGGCG	12759	CAACTAATTC	1191	11/25 (44%)
Regulation of transcription related						
<i>MSRB</i> (Pilin-like transcription factor)	↑	TGAATTGAGT	76467	GTACGTCTGG	22921	12/25 (48%)
<i>SP110</i> (nuclear body protein)	↑	AATACTAGAC	109032	GCATCTTCAA [†]	3431	10/25 (40%)
<i>KLF5</i> (Kruppel-like f5)	↓	AAGCGCCACC	12224	AAAAGCAGAA	688	22/25 (88%)
<i>CEBPB</i> (C/EBP-β)	↓	GCGGCCGTT	12608	GCTGAACGCG	1051	22/25 (88%)
<i>JUN</i> (c-Jun)	↓	CTAACGCAGC	16476	CCTTTGTAAG	3725	18/25 (72%)
<i>JUNB</i>	↓	GCCCCCTTCC	16477	ACCCACGTCA	3726	17/25 (72%)
<i>CSDA</i> (DBPA)	↓	GGGATTGCCC	56449	ATTAAAAAAA [†]	8531	16/25 (64%)
<i>STAT3</i>	↓	GCATCCTGTT	20848	TGAGGAGCTG	6774	15/25 (60%)
<i>ELF1</i> (E74-like factor 1)	↓	TAAAAGTTCT	13709	AAAAATTGGA	1997	13/25 (52%)
<i>SAP18</i> (Sin3 associated p. 18)	↓	GCAGTTCACA	20220	AGACCATATT	10284	12/25 (48%)
Signal transduction related						
<i>SARA2</i> (SAR1B)	↑	CAATAAAACA	66397	CAATAAAACA	51128	17/25 (68%)
<i>MGC4796</i> (Ser/Thr-like kinase)	↓	TCAGCAATAA	74178	TCAGCAATAA	83931	25/25 (100%)
<i>DTR</i> (HB-EGF)	↓	TCTGAACTGT	15200	GTCCTTTCTG	1839	25/25 (100%)
<i>DSCR1</i> (Calcipressin 1)	↓	CTTTGGGGAC	54720	CTTTGGAAAT	1827	25/25 (100%)
<i>ITPKC</i>	↓	GACTACGTGG	233011	ACAACACCCC	80271	25/25 (100%)
<i>MAP2K3</i> (MKK3)	↓	GTTTGGAGCT	26397	GTTTGGAGCT	5606	24/25 (96%)
<i>TACSTD2</i>	↓	AAGCGCCTCA	56753	GCCTACCCGA	4070	23/25 (92%)
<i>MT2A</i> (Metallothionein 2)	↓	TAAGTACAA	17750	GATCCCAACT	4502	21/25 (84%)
<i>DUSP8</i> (Dual S-phosphatase 8)	↓	AGACGGATGT	18218	GTGGAGGGGC	1850	20/25 (80%)
Cytokines						
<i>CXCL1</i> (GRO1)	↓	TGTGGGAGGC	14825	TTGAAACTTT	2919	25/25 (100%)
<i>IL-6</i>	↓	AAGAACAACCT	16193	GGCACCTCAG	3569	25/25 (100%)
<i>LIF</i>	↓	GTAGCGGCTT	16878	GCCTTGGGTG	3976	25/25 (100%)
<i>CCL2</i> (MCP-1)	↓	AATACTAAAA	20296	GTACTAGTGT	6347	25/25 (100%)

Human gene name		Mouse tag	Mouse LocusID	Human tag	Human LocusID	Frequency in human tumors*
ECM, cell adhesion-communication, secreted proteins						
<i>COL4A1 (Procollagen IV A1)</i>	↑	GCTTTCCTGT	12826	GACCGCAGGA	1282	11/25 (44%)
<i>SDFR1</i>	↑	TACTTGTTGTT	20320	TACTTGTTGTTG	27020	9/25 (36%)
<i>EMP1 (Epithelial membrane P1)</i>	↓	AGACGTAAAT	13730	TAATTTGCAT	2012	24/25 (96%)
<i>TNFAIP2</i>	↓	ATTCGAGGCT	21928	ACTCAGCCCG	7127	24/25 (96%)
<i>FST (Follistatin)</i>	↓	TAAATGTGCA	14313	AAGGAAGCTG	10468	22/25 (88%)
<i>MFGE8</i>	↓	TTCCATTCCG	17304	GGTTGGCAGG	4240	21/25 (84%)
<i>SLPI (Antileukoproteinase)</i>	↓	GCTCCCGGCT	20568	TGTGGGAAAT	6590	21/25 (84%)
<i>CYR61 (IGFBP10)</i>	↓	TTTACGGATG	16007	AGTGTCTGTG	3491	20/25 (80%)
<i>JUP (Plakoglobin)</i>	↓	GGTTTGGGGG	16480	GTGTGGGGGG	3728	9/25 (36%)
Chaperones						
<i>FKBP9</i>	↑	CAAATGCTGT	27055	GAATAAATGT	360132	8/25 (32%)
<i>HSPA1A (HSP70-1)</i>	↓	TGTTCAAGTTT	193740	CAGAGATGAA	3303	25/25 (100%)
<i>CRYAB (Crystallin α2)</i>	↓	TTCGTCCTGC	12955	GTTTCATCTC	1410	19/25 (76%)
Cytoskeleton related						
<i>DNC12 (dynein, interm.chain 2)</i>	↑	TCGGACAAAA	13427	TGGTAGTTAC	1781	11/25 (44%)
<i>CAPZA1 (Capping prot. α1)</i>	↑	CCACACTGTC	12340	TCAAAAAAAG	829	8/25 (32%)
<i>TUBA6 (Tubulin α6)</i>	↓	TGCTGCCATT	22146	GCTGCCCTTG	84790	23/25 (92%)
<i>KRT14 (Keratin 14)</i>	↓	CTGCTCAGGC	16664	GATGTGCACG	3861	23/25 (92%)
<i>MYO5B (Myosin Vb)</i>	↓	ACGTAAAAAA	17919	TTTCCAGCAT	4645	21/25 (84%)
<i>LMNA (Lamin A/C)</i>	↓	CCAAAGTCTT	16905	CGCAAGCTGG	4000	20/25 (80%)
<i>KIAA0992 (Palladin)</i>	↓	GGCCTATCTC	72333	GACCTATCTC	23022	11/25 (44%)
DNA repair related						
<i>RPA3 (Replication protein A3)</i>	↑	AATGAGCTTC	68240	ATTGATCTTG	6119	9/25 (36%)
<i>TOP1 (Topoisomerase I)</i>	↓	AATTCCCTTT	21969	AATTCCCTTT	7150	18/25 (72%)
RNA and protein biosynthesis and metabolism						
<i>RPL13 (Ribosomal protein L13)</i>	↓	ATACTGAAGC	270106	CCCGTCCGGA	6137	22/25 (88%)
<i>HNRPA0 (HnRNA binding protein)</i>	↓	CCTGGAATTT	77134	CCTGAAATTT	10949	24/25 (96%)
<i>SBDS</i>	↓	CTACCAAAAC	66711	CTGCCATAAC	51119	16/25 (64%)
<i>SUI1 (Putative transl.initiation f.)</i>	↓	CAATAAACTG	20918	CAATAAACTG	10209	16/25 (64%)
<i>EEF1G (Elongation factor 1G)</i>	↓	TGGGCAAAGC	67160	TGGGCAAAGC	1937	11/25 (44%)
<i>EEF2 (Elongation factor 2)</i>	↓	TGCTCGCAAA	13629	AGCACCTCCA	1938	10/25 (40%)
<i>PABPC1 (PolyA binding prot.1)</i>	↓	CTCGAGTCTC	18458	AAAAGAAACT	26986	9/25 (36%)
Metabolism, mitochondrion, miscellaneous						
<i>SEPW1 (Selenoprotein W, 1)</i>	↑	TTTCCAGGTG	20364	TCTTCCCCAG	6415	20/25 (80%)
<i>COX7B (Cytochrome C ox.VIIB)</i>	↑	GAAACTAGGT	66142	GACATATGTA	1349	19/25 (76%)
<i>SURF1 (Surfeit 1)</i>	↑	TGAAGTAAGG	20930	TGCTACTGGT	6834	18/25 (72%)
<i>COX6C (Cytochrome Cox.Vic)</i>	↑	AATATGTGTG	12864	AATATGTGGG	1345	17/25 (68%)
<i>ADH5 (Alcohol dehydrogen. 5)</i>	↑	AACTAGGCAA	11532	AACCTGTTTT	128	16/25 (64%)

Human gene name		Mouse tag	Mouse LocusID	Human tag	Human LocusID	Frequency in human tumors [*]
<i>TIPARP</i>	↓	TTACCATTGC	99929	AAATGGCCAA	25976	24/25 (96%)
<i>HMOX1 (Hemoxygenase)</i>	↓	GGCACTGGGC	15368	CGTGGGTGGG	3162	23/25 (92%)
Function unknown						
<i>HSPC148</i>	↑	GAGAAATATA	66070	GAGAAATATA	51503	16/25 (64%)
<i>BM-009</i>	↑	TTTAAGAATG	223601	TCTAAAGAAT	51571	8/25 (32%)
<i>FLJ23231</i>	↓	TATGCTGTGT [†]	230738	TAAATTTTAA [†]	80149	25/25 (100%)
<i>KIAA0346</i>	↓	CTGGGCGCGG	216850	GTGCTCAAAC	23135	21/25 (84%)
<i>HSHIN1</i>	↓	TAGGGCATTG	73945	GTGTGATGCT	54726	10/25 (40%)

NOTE. 72 genes commonly deregulated in mouse and human breast carcinoma SAGE libraries (normal VS tumor $P < 0.05$). 19 genes were found over expressed (↑) and 53 genes with decreased expression (↓) in breast carcinomas compared to normal mammary epithelium.

^{*} Transcript tags changing >3-fold when compared with normal in at least 8 of 25 (32%) human breast cancer SAGE libraries.

[†] Tag matches >1 gene.