Oncogene (2005) 24, 2705–2714

www.nature.com/onc

Analysis of methylation-sensitive transcriptome identifies GADD45a as a frequently methylated gene in breast cancer

Wei Wang[1,3], Gudrun Huper[1], Yueqi Guo[1], Susan K Murphy[2], John A Olson Jr[1] and Jeffrey R Marks*[,1]

[1]*Department of Surgery, Duke University Medical Center, Durham, NC 27710, USA;* [2]*Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC 27710, USA*

Treatment of the breast cancer cell line, MDAMB468 with the DNA methylation inhibitor, 5-azacytidine (5-AzaC) results in growth arrest, whereas the growth of the normal breast epithelial line DU99 (telomerase immortalized) is relatively unaffected. Comparing gene expression profiles of these two lines after 5-AzaC treatment, we identified 36 genes that had relatively low basal levels in MDAMB468 cells compared to the DU99 line and were induced in the cancer cell line but not in the normal breast epithelial line. Of these genes, 33 have associated CpG islands greater than 300 bp in length but only three have been previously described as targets for aberrant methylation in human cancer. Northern blotting for five of these genes (α-Catenin, DTR, FYN, GADD45a, and Zyxin) verified the array results. Further analysis of one of these genes, GADD45a, showed that 5-AzaC induced expression in five additional breast cancer cell lines with little or no induction in three additional lines derived from normal breast epithelial cells. The CpG island associated with GADD45a was analysed by bisulfite sequencing, sampling over 100 CpG dinucleotides. We found that four CpG's, located approximately 700 bp upstream of the transcriptional start site are methylated in the majority of breast cancer cell lines and primary tumors but not in DNA from normal breast epithelia or matched lymphocytes from cancer patients. Therefore, this simple method of dynamic transcriptional profiling yielded a series of novel methylation-sensitive genes in breast cancer including the BRCA1 and p53 responsive gene, GADD45a.
Oncogene (2005) **24,** 2705–2714. doi:10.1038/sj.onc.1208464
Published online 14 February 2005

Keywords: GADD45a; methylation; breast cancer; expression array

Introduction

Aberrant gene expression is a characteristic of all cancers. Gene amplifications, deletions, translocations, mutations, alternative splicing, and methylation changes underlie much of this variation in expression. Secondary or indirect effects induced by primary genetic/epigenetic events that influence the level of transcription factors add another level of complexity. It is widely assumed that many of these changes in gene expression are etiologic with respect to neoplastic progression.

Changes in DNA methylation status can have profound effects on the expression of genes; however, these changes can be difficult to identify and in many cases, even more difficult to determine whether they are causal in altering gene expression. CpG islands are genomic regions enriched in CG dinucleotides and commonly associated with the promoter regions of many genes. Most strategies designed to detect methylation changes do not provide comprehensive sampling of CpG islands working under the assumption that methylation status is uniform throughout a given island. Critical or common methylation changes cannot be predicted *a priori* based upon sequence or transcription factor-binding sites, therefore, the subset of residues sampled may leave important methylated regions undetected. Further, inappropriate tissue or cell type comparisons may produce misleading results since aberrant methylation patterns can only be determined by ultimately knowing what constitutes the normal condition. For example, comparing the methylation status of an epithelial cancer to normal lymphocytes or fibroblasts can produce a high rate of false positives and false negatives.

A number of genes are commonly hypermethylated resulting in reduced expression in breast cancer including p16, Cyclin D2, 14-3-3σ, BRCA1, HIN1, Maspin, RARβ, and RASSF1A (Dobrovic and Simpfendorfer, 1997; Woodcock *et al*., 1999; Domann *et al*., 2000; Ferguson *et al*., 2000; Krop *et al*., 2001; Evron *et al*., 2001b; Fackler *et al*., 2003). Since hypermethylated genes may be involved in the etiology of the disease and may be early events in neoplastic progression, these genes make attractive targets for therapeutics and disease detection/monitoring (Bovenzi *et al*., 1999; Evron *et al*., 2001a; Umbricht *et al*., 2001; Muller *et al*.,

*Correspondence: JR Marks, Department of Surgery, Duke University Medical Center, Box 3873, Medical Sciences Research Building, Durham, NC 27710, USA; E-mail: marks003@mc.duke.edu
[3]Current address: Department of Medicine, University of Washington, Seattle, WA 98195, USA
Received 2 September 2004; revised 17 December 2004; accepted 17 December 2004; published online 14 February 2005

2003; Oshiro *et al*., 2003; Parker *et al*., 2003). Agents that reactivate gene expression silenced by methylation have favorable therapeutic indices and are now entering clinical trials for a variety of cancers (Baylin, 2004; Dowell and Minna, 2004; Egger *et al*., 2004; Zelent *et al*., 2004).

Breast cancer is a heterogeneous disease at the microscopic and molecular levels. In addition to the primary divisions of invasive breast cancer by histologic subtype, for example, ductal, lobular, medullary, and mucinous, there are now a number of molecular markers that discriminate subsets of the disease. The most prominent of these continue to be treatment related, that is, estrogen receptor and HER2 status. However, with the advent of gene expression arrays, further functional categorization is possible. One of these categories, the basal subtype, is characterized by lack of hormone receptor expression, absence of HER2 amplification, frequent p53 mutations, and expression of a constellation of intermediate filament and basement membrane proteins commonly found in normal basal cells (Dairkee *et al*., 1988; Perou *et al*., 2000). These cancers are more likely to recur and lead to disease-related mortality (Sorlie *et al*., 2001).

In an effort to identify new genes that are functionally affected by altered methylation status in breast cancer, we performed a discovery gene expression array experiment on a prototypical basal cancer cell line after treating with the DNA methylation inhibitor, 5-azacytidine (5-AzaC). In this report, we describe a series of transcripts that are induced after 24 h of treatment with this agent when compared to normal mammary epithelial cells, also of the basal phenotype. One of these genes, GADD45a, was further analysed for aberrant methylation in cancer. A contiguous series of CpG residues located approximately 700 bp upstream of the start of GADD45a transcription are hypermethylated in breast cancer but not in normal breast epithelial cells. This simple but effective approach may have significant utility for discovery of hypermethylated genes in cancer.

Results

Gene expression patterns after 5-AzaC treatment

We chose two established breast cell lines to compare methylation-sensitive transcription, one derived from normal mammary epithelial cells and immortalized by telomerase expression (DU99) and the other a well-studied cancer line (MDAMB468). The choice of these cells was dictated by the following considerations: The guiding principle was to ensure a relevant normal versus cancer comparison with cells that are likely to be of the same lineage. The DU99 cell line began as primary mammary epithelial cells (HMEC) cultured from a reduction mammoplasty and maintained in DFC1 medium (Band and Sager, 1989). Primary HMEC isolated and grown in this manner exhibit the characteristics of basal mammary epithelial cells as defined by expression of specific cytokeratins, vimentin, basement membrane proteins, and smooth muscle actin (for a few cell divisions) (Pechoux *et al*., 1999). After retroviral transfer of the gene encoding the enzymatic subunit of telomerase, an immortalized clone was derived and has been in continuous culture in our laboratory for over 3 years. These cells have been extensively characterized and were found to have the following properties: (1) they retain expression of enzymatically active telomerase; (2) they have similar growth kinetics to the parental cells and remain contact inhibited at confluence; (3) the cells are nontumorigenic in immunocompromised mice; (4) the DU99 line is karyotypically identical to the parental HMEC culture and contains a diploid content of DNA; (5) these cell respond to ionizing radiation by inducing p53 protein, p21 mRNA, and cell cycle arrest; (6) their growth in three-dimensional culture mimics the parental cell culture in forming regular acinar structures surrounded by a basement membrane. Under these culture conditions, prolactin was capable of stimulating milk protein production; and (7) unlike many other telomerase immortalized lines, DU99 cells do retain expression of the CDK inhibitor, p16 (Flowers *et al.,* in preparation).

We chose to use these cells instead of primary normal HMEC since they are more homogeneous (primary cultures have both basal and luminal lineages represented, particularly at early passages) and have all of the advantages inherent in immortalized cell lines. We chose the MDA468 line among the many available breast cancer cell lines on the basis of the following three criteria: (1) the gene expression pattern of these cells closely resembles the basal phenotype, including high level expression of the epidermal growth factor receptor (JRM, unpublished data); (2) they contain near the median level of CpG island hypermethylation ($\sim$30%) and total 5-methylcytosine ($\sim$4%) from a series of 70 cancer cell lines analysed (Paz *et al*., 2003); (3) treatment with 5-AzaC is nontoxic but still has a demonstrable effect on the cells indicating that physiologically relevant genes are affected.

Figure 1 shows a growth curve of cells (both DU99 and MDA468) treated with a single dose of 5-AzaC. Viability of the cell lines was unaffected by treatment at both doses (1 and 5 μM), and growth of the DU99 normal breast line was slightly growth inhibited only at the higher dose. In contrast, the growth of MDA468 cells was almost entirely suppressed in the presence of either dose of 5-AzaC. A parallel flow cytometric cell cycle analysis revealed that the growth arrest did not occur in a particular phase, nor was there a significant increase in the subdiploid (apoptotic) fraction (data not shown). Therefore, comparing transcriptional patterns after 5-AzaC treatment in these two lines should provide a relevant and interesting data set to explore.

For microarray expression analysis, each cell line was treated with 5 μM 5-AzaC (or vehicle) for 24 h and total RNA was extracted. cRNA probe was prepared, and then hybridized to Affymetrix U133A gene expression arrays ($>$22 000 probe sets). We used several criteria to identify genes that may be targets for hypermethylation

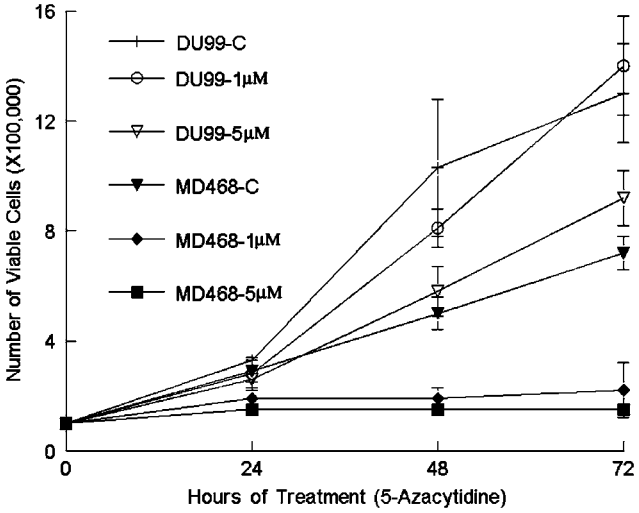

Figure 1 Growth suppression by 5-azaC of the breast cancer cell line, MDAMB468 but not the normal DU99 line. Cells (1×10^5) were treated at time 0 with the vehicle control, 1, or 5 μM 5-AzaC. Viable cells (trypan blue excluding) from triplicate wells were counted at the indicated time points and the resulting data were averaged and plotted. For the MDA468 line, both adherent and nonadherent cells were counted. No increase in the number of dead cells was observed over the course of the treatment in either line

in cancer but not in normal cells. First, expression had to be induced at least twofold in MDA468 cells versus control and show little or no induction in the normal DU99 line. Second, the basal level of expression in the cancer line had to be lower than that observed in the normal breast line, indicative of possible repression in the cancer. Finally, the level of expression after treatment had to be scored as 'present' by the Affymetrix analysis software. This excluded many genes where both basal and induced levels did not reach what is considered a detectable level of hybridization. In total, 36 genes met these criteria and are listed in alphabetical order in Table 1. Many of these genes were represented on the arrays by more than one probe set, and in these cases the average induction is shown across all relevant probes. All but three of these genes have CpG islands that span at least 300 bp over their 5′ ends. Further, only three of the genes have published information indicating aberrant methylation in cancer.

Northern blot validation of five identified genes

We chose five genes for further validation of the expression array data. Each of the genes contains a well-defined CpG island but none of them have any previously published information regarding transcriptional regulation via promoter methylation. MDA468 and the DU99 cell lines were treated for 24 h with 5-AzaC (1 and 10 μM), and total RNA was probed by Northern blotting for steady-state levels of α-Catenin (CTNNA1), Diptheria toxin receptor (DTR), the FYN oncogene, GADD45a, and Zyxin (ZYX) (Figure 2). For each gene, we observed induction in the MDA468 line, but either very little or no increase in the normal DU99 line, confirming the results from the expression arrays. Induction of steady-state transcript levels ranged from twofold (DTR) to greater than sixfold (FYN and GADD45). These values are in good agreement with the level of induction measured by the expression arrays (Table 1).

GADD45a is commonly regulated by DNA methylation in breast cancer cell lines

From this initial set of five genes, we chose GADD45a for more detailed characterization. GADD45 is transcriptionally activated by two important breast cancer tumor suppressor genes, p53 and BRCA1, and therefore may itself be an important target for downregulation. To broaden the results obtained above, we next treated a series of breast epithelial lines with 5-AzaC and the histone deacetylation inhibitor, trichostatin A (TSA). We examined three additional normal breast and five more cancer-derived cell lines for GADD45a mRNA levels (Figure 3). Of the normal cell lines, only the MCF10A spontaneously immortalized line showed induction of GADD45a after 5-AzaC treatment. No change in the steady-state level of the GADD45a transcript was observed with TSA treatment in these noncancerous cell lines. In contrast, 5-AzaC induced GADD45 expression in each of the five additional cancer lines. HCC1937, a line harboring a hemizygous BRCA1 mutation, showed the highest level of induction. GADD45a transcription responded to TSA treatment only in the T47D and MCF7 lines. Little or no induction by TSA was observed in HCC1937, ZR75-1, BT474, and MDA468 cells. Thus, GADD45 is responsive to the DNA methylation inhibitor in a number of breast cancer cell lines but not normal breast epithelial cells. Aberrant histone acetylation may also be a factor in regulating expression of this gene in some cancer lines.

GADD45a gene is methylated in breast cancers

To determine whether GADD45a is hypermethylated in breast cancer, we performed bisulfite sequencing through the CpG island that covers the 5′ end of the gene. This analysis included over 1.4 kb of sequence starting from the first cluster of CpG residues about 700 bases upstream of the transcription start site through most of exon 2 (Figure 4a). Since our initial results were on the MDA468 breast cancer line and GADD45a was specifically and highly induced by 5-AzaC but not TSA in these cells, we chose DNA from these cells to screen for methylation in this region. Counter to expectations, we found that among the 109 CpG dinucleotides covered in our analysis, only four were methylated (Figure 4b). These residues are sequential and located approximately 700 bp upstream of the transcriptional start site. There was no detectable methylation at any other CpG throughout this region. We next analysed the same region in a series of normal and cancer-derived cell lines to determine whether the methylation was cancer specific. DNA from eight normal breast epithelial cultures (five primary HMEC and three immortalized

Table 1 Genes induced by 5-AzaC treatment

Gene symbol	*Gene name*	*MDA468 induction*[a]	*DU99 induction*[a]	*MDA468*[b] *DU99*	*CpG island (#)*[c]
CAV1	Caveolin 1	3.5	1.1	0.2	Y (148)
CDC25A		4.6	1.2	0.8	Y (110)
CLC	Cardiotrophin-like cytokine	2.5	1.3	1	Y (105)
COPEB	Core promoter element binding protein	6.4	1	0.9	Y (132)
CTNNA1	α-Catenin	3.2	1	0.7	Y (104)
DKK1	Dickkopf homolog 1	10.2	1	0.9	Y (64)
DTR	Diptheria toxin receptor	22	1.1	1.9	Y (64)
DUSP1	Dual specificity phosphatase 1	5.6	1	1.1	Y (194)
EPHA2	Ephrin receptor	3.3	1.4	0.6	Y (118)
ETV5	Ets variant gene 5	5	1	1	Y (139)
F3	Coagulation factor III	4.3	1.2	0.3	Y (102)
FOSL1	Fos-like antigen 1	15.4	1	1.3	Y (154)
FYN	Fyn oncogene	12.7	0.9	1.1	Y (95)
GADD45A	Growth arrest and DNA damage 45A	6	0.9	3.8	Y (150)
GJB3	Connexin 31	3.3	1.2	0.6	Y (26)
INHBA	Inhibin beta A	7.3	1.2	2	Y (26)
IRS2	Insulin receptor substrate 2	4.2	0.8	2.3	Y (548)
KRT17	Keratin 17	3.1	1.1	0.3	N
KRTHB1	Keratin 1	3.3	1.1	1	N (34)
MAFF	v-maf	4.5	0.8	2.6	Y (139,62)
MYO1B	Myosin 1B	3.1	1	0.6	N
PLAB	Prostate differentiation factor	46	1.3	0.4	Y (20,74)
PTPRG	Protein tyrosine phosphatase, receptor type G	5.7	0.6	1.5	Y (266)
PTRF	Polymerase 1 and transcript release factor	3.9	1	0.6	Y (156,97)
SERPINB8	Serine protease inhibitor B8	4.9	1.1	1.1	Y (48)
SLC3A2	Solute carrier family 3	3.7	1.1	1.3	Y (47,81)
SLC7A5	Solute carrier family 5	3	1.2	0.4	Y (105)
SLC9A1	Solute carrier family 9	5.4	1	0.5	Y (40)
SMURF2	E3 ubiquitin ligase	3.4	1	1.3	Y (114)
SPHK1	Sphingosine kinase 1	3.4	1	0.5	Y (376)
STK17A	Serine/threonine kinase 17A	2.6	1	0.5	Y (80)
THBS1	Thrombospondin 1	3.0	1.2	0.2	Y (109)
TNFRSF10B	TNF receptor 10B	2.7	1	0.3	Y (109)
TUBB	Tubulin beta	3.6	0.8	2.2	Y (108)
UPP1	Uridine phosphorylase	11	1.1	1.9	Y (31,104)
ZYX	Zyxin	3.4	1.2	0.9	Y (156)

[a]Fold induction of expression, averaged between all probe sets for a given gene on the U133 arrays. [b]Ratio of expression in the MDA468 line compared to DU99 after treatment. [c]Existence of a CpG island spanning an area of at least 300 bp over the 5′ end of the gene and number of CpG's within this region

normal cell lines, DU99, MCF10A, and 26NC) showed no evidence of methylation within this amplicon. Of the nine additional breast cancer cell lines, all but the MDA361, BT549, and HCC1937 lines had detectable methylation in this region.

We next analysed a series of 20 primary breast cancer specimens for evidence of methylation. DNA from 12 of these cancer specimens demonstrated detectable levels of these four C residues after bisulfite modification indicative of methylation (Figure 4c). Matching normal DNA (isolated from peripheral blood lymphocytes) from eight of these patients was also analysed and showed no sign of methylation at these residues. Therefore, the methylation pattern appears to be cancer specific.

To determine the pattern of methylation of individual DNA molecules, bisulfite-treated and PCR-amplified DNA from four primary tumor and two normal breast epithelial cell cultures were cloned. From each ligation, 10 insert containing plasmids were sequenced across the commonly methylated region as defined above. Neither of the normal epithelial specimens contained any protected C residues within this region. Of the 40 sequences generated from the cancers, a relatively complex picture emerged. Of the 40 clones, 21 exhibited methylation at one or more of the four CpG's; however, 11 of these clones contained one, two, or three protected C residues while 10 were methylated at all four. In the clones that were less than fully methylated, residues more distal to the 5′ end of the GADD45 gene were more likely to be methylated. These results are concordant with the direct sequencing (Figure 4) that most commonly indicated a gradient of intensity of the methyl CpG that decreased across this region from 5′ to 3′ with respect to the gene.

Expression of Gadd45 mRNA in breast epithelia

Methylation status is often associated with mRNA expression, particularly for genes where the CpG island is broadly methylated. We discovered methylation of GADD45 by virtue of derepressed expression differences between a normal and cancer cell line. To broaden this analysis, we examined steady-state mRNA levels in a series of breast epithelial cultures and primary tumors (Figure 5). A panel of cultured cells, both primary and

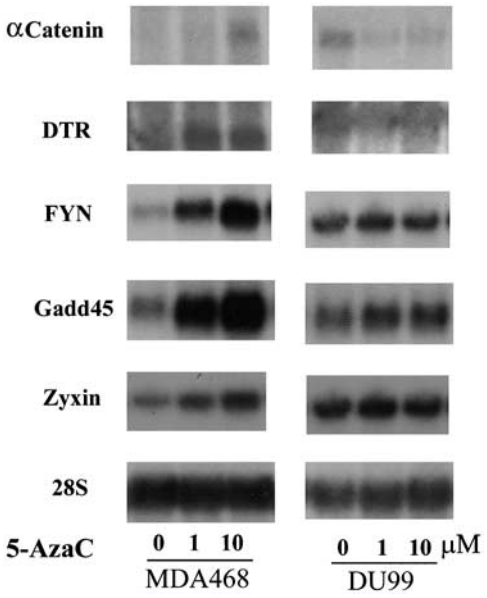

Figure 2 Expression of five genes identified by the expression array experiment analysed by Northern blotting. The cells were treated for 24 h with the indicated doses of 5-AzaC before total RNA was harvested. The 28S rRNA band from each gel is shown as a loading control

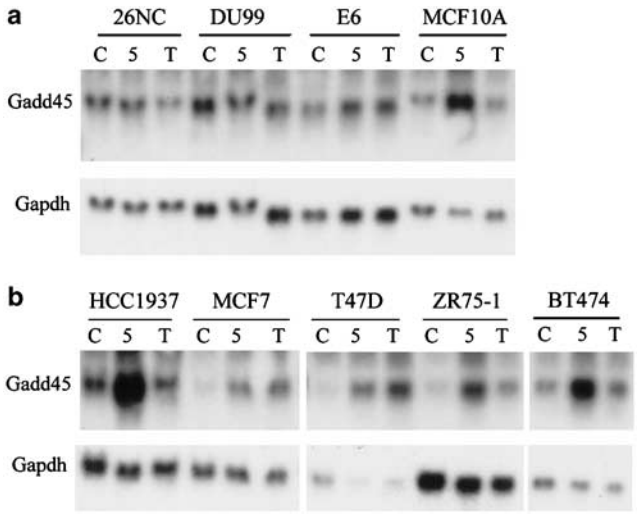

Figure 3 Northern blotting for GADD45a transcript in normal and breast cancer cell lines after treatment with either 5-AzaC (5) or Trichostatin A (T). Cells treated with vehicle (a combination of 1 : 10 000 ethanol and 1 : 33 000 acetic acid) are labeled 'C'. After hybridization to detect GADD45, the blots were stripped and reprobed for GAPDH. Equal amounts of total RNA (10 μg) were loaded for each cell line, with GAPDH expression varying greatly between different lines

immortalized normal breast cells and breast cancer cell lines, exhibited widely varying levels of the Gadd45 transcript (Figure 5a). Three separate primary HMEC cultures and the telomerase-immortalized DU99 line (lanes 1–4) all expressed high levels, while the normal epithelial-derived MCF10A and HPVE6 immortalized lines had relatively low levels (lanes 5–6). None of these cultures demonstrated detectable levels of methylation. Of the cancer cell lines, levels comparable to HMEC were observed only in the HCC1937 line (lane 14). The HCC1937 and two other lines had no detectable methylation in the region (BT549 and MDA361, lanes 8–9). Methylation at these residues does not seem to be the sole determinant of steady-state expression of Gadd45 as some of the methylated lines express equivalent or higher levels than unmethylated cells.

Primary cancer specimens were also examined in this manner (Figure 5b). The first three lanes are cell lines represented on the upper blot (lanes b1, 2, 3 = a4, 11, 12), while the remaining lanes contain RNA from the same cancers assayed for methylation status. As for the cell lines, widely varying levels of Gadd45 transcript are expressed in primary breast cancers. The cancer with the highest level of expression (lane 8) had no detectable methylation, however as with the cell lines, there does not appear to be a simple and direct correlation between basal steady-state mRNA levels and methylation status.

Discussion

Epigenetic silencing of gene expression is associated with the onset and progression of human cancer. Silencing may occur via promoter methylation and/or modification of histones. These two types of events may also be causally related, that is, the presence of one inducing the other (e.g. CpG methylation inducing methylation on H3 lysine 9) (Coombes *et al*., 2003). It is likely that the known complexity of these interactions will increase and aberrant forms of these epigenetic transcriptional control processes will be further implicated in the etiology of human cancer. In this study, we describe a series of genes that are induced by treatment with 5-AzaC in a breast cancer cell line but exhibit little or no change in expression with the same treatment in a comparable normal breast epithelial line. Of the 36 genes listed in Table 1, all but three have clearly defined CpG islands near their 5′ end. The presence of CpG-rich regions does not immediately implicate them as regulated by methylation in general or specifically in breast epithelial cells since more than 50% of all human genes may have associated CpG clusters (Murphy and Jirtle, 2001). Only three of these genes have been shown to be hypermethylated in human cancer, CAV1, THBS1, and TNFRSF10B (Li *et al*., 1999; Cui *et al*., 2001; Wiechen *et al*., 2001; Oue *et al*., 2003; van Noesel *et al*., 2003). Of these, THBS1 and CAV1 methylation have been linked specifically to breast cancer (Engelman *et al*., 1999; Li *et al*., 1999). Conversely, a number of genes that have been shown to be hypermethylated in breast cancer were not scored in this assay including p16, RARβ, 14-3-3σ, and Cyclin D2 (Foster *et al*., 1998; Ferguson *et al*., 2000; Sirchia *et al*., 2000; Evron *et al*., 2001b). A number of explanations could account for this including: (1) our initial assay that utilized a single breast cancer cell line, (2) 5-AzaC does not reactivate all genes that are methylated, and (3) the absence of high-quality probe sets for measuring transcript levels of the genes on the

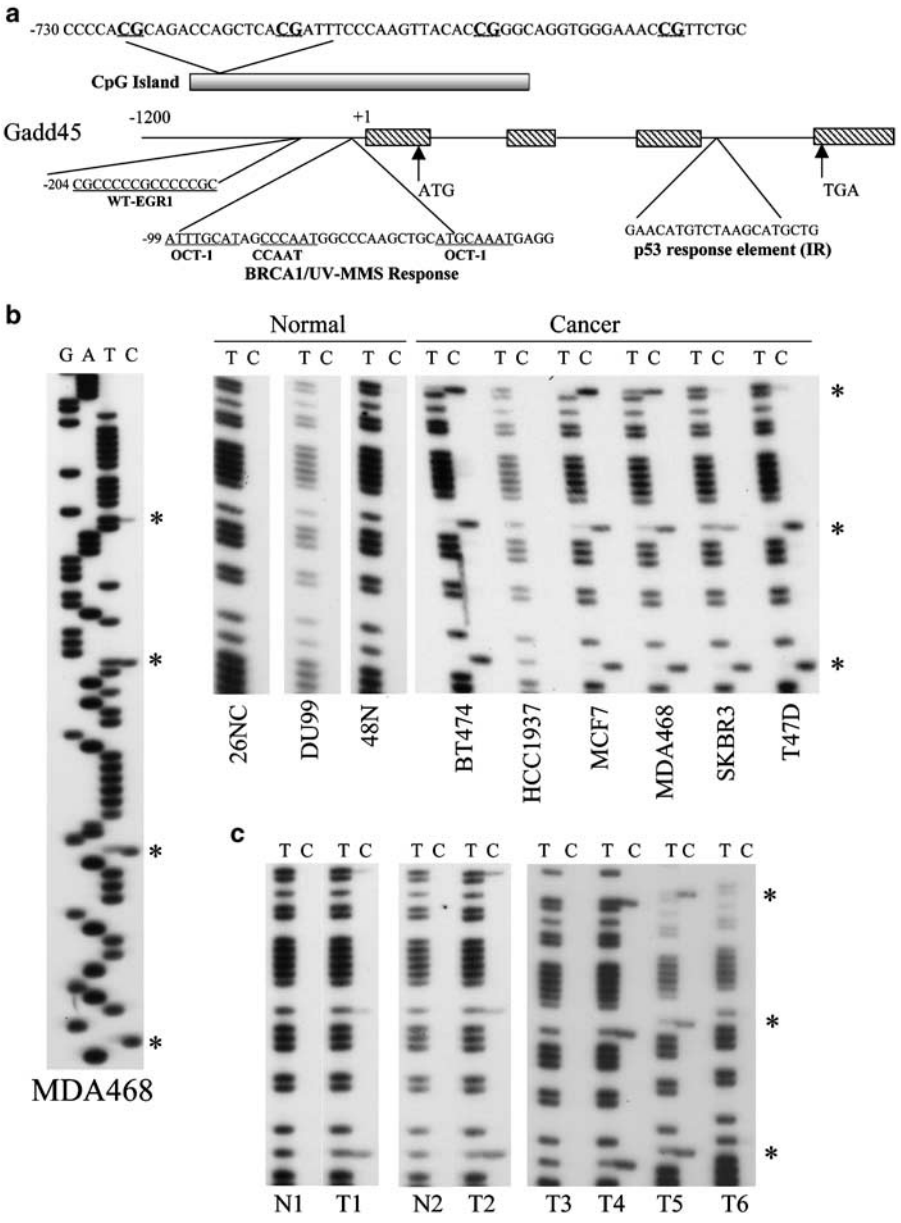

Figure 4 Methylation of the GADD45a promoter region in breast cancer. (**a**) Diagram of the GADD45a gene (Genbank accession #L24498) with the known transcriptional control regions indicated including the upstream WT-EGR1 site, the combination OCT-1/BRCA1-binding site, and the p53 response element in intron 3. The CpG island is indicated by a solid bar above the sequence and the specific region that is commonly methylated in breast cancer is shown on top with the methylated residues in bold. (**b**) Bisulfite sequence read from primer 1F (Table 2) for a series of breast cancer and normal breast epithelial lines showing methylated C residues in the cancer but not normal cells. The 48N cells are primary, nonimmortalized breast epithelial cells. (**c**) Bisulfite sequencing of two normal (lymphocyte) : tumor pairs and an additional four primary breast cancers

Affymetrix arrays. Nonetheless, the simple assay approach was effective in identifying a series of genes that are regulated by the DNA methylation inhibitor.

Of the five genes that were chosen for additional analysis (α-Catenin, DTR, FYN, GADD45, and Zyxin), all demonstrated induction by Northern blotting in the MDA468 line but not in the DU99 normal breast epithelial line. Further, each of these five genes was also induced by 5-AzaC in at least one other breast cancer cell line and not in additional normal breast cells (data not shown).

Finally, we chose to focus on the GADD45a gene for more detailed characterization. From a systematic examination of the entire CpG island, we found four consecutive residues that were methylated in breast cancers, both cell lines and primary tissue specimens. Methylation at these residues was observed in 7/10 cancer cell lines, 12/20 primary breast cancers, but none of the eight normal breast epithelial cultures or eight lymphocyte DNAs (from a subset of the breast cancer patients whose tumors were analysed). Therefore, methylation at these sites appears to be a common

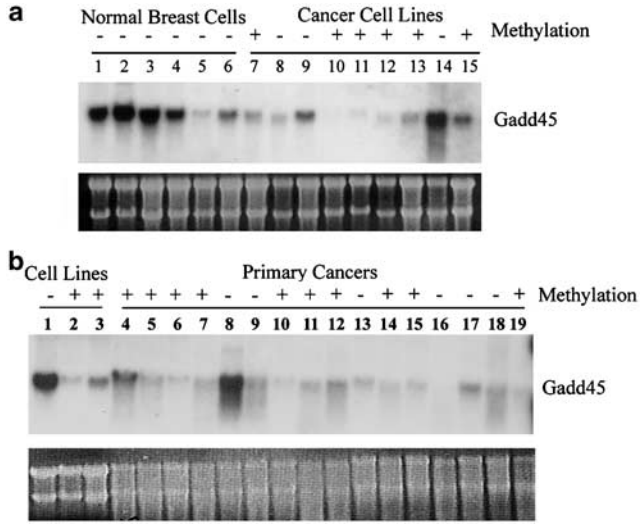

Figure 5 Expression of Gadd45 mRNA in breast epithelia with known methylation status. (**a**) Northern blot of a series of cultured breast epithelial cells. Lanes 1–3, primary breast epithelial cells; lane 4, DU99; lane 5, MCF10A; lane 6, E6-immortalized HMEC; lane 7, MDAMB435; lane 8, MDAMB361; lane 9, BT549; lane 10, T47D; lane 11, MCF7; lane 12, SKBR3; lane 13, BT20; lane 14, HCC1937; lane 15, MDAMB231. Methylation status of the GADD45 gene is indicated above the blot with any degree of methylation indicated by '+'. (**b**) Northern blot of a series of primary breast cancer. Lane 1, DU99; lane 2, MCF7; lane 3, SKBR3, lanes 4–19, primary invasive ductal breast cancers. Below each blot is the ethidium stained gel showing the loading. The 28 and 18 S bands are less distinct in the primary tumor specimens (**b**, 4–19) compared to the cell lines (lanes 1–3) due to slight degradation also reflected in the increased smear of the hybridization signal in these samples

feature of breast cancer and does not occur in normal breast epithelial cells or lymphocytes. This makes it a possible candidate for disease detection or monitoring with a profile similar to Cyclin D2, Twist, and RAR β (Evron *et al*., 2001a).

That GADD45a is hypermethylated in breast cancer does not come as a surprise. What is surprising is that this is the first description of the phenomenon. GADD45a is involved in the G2 arrest after DNA damage, inducible by both of the commonly mutated tumor suppressors, p53 and BRCA1, and therefore is a logical target for inactivation in its own right (Harkin *et al*., 1999). Our discovery of hypermethylation of this gene was accomplished through two avenues – transcriptional induction observed using the expression arrays and bisulfite sequencing of the entire CpG island. If we did not have the information from the expression arrays, we would have been less prone to fully examine the CpG island. The commonly used alternative approach, methylation-specific PCR, samples only a few CpG dinucleotides per reaction (specifically, those included in the PCR primer) and given the limited nature of methylation in this promoter region, we would not have observed these methylation events relatively far upstream of the transcription start site.

The GADD45a methylated residues, over 700 bp upstream of the transcription start site, are not contained with any recognized response element. BRCA1 acts via several proximal elements that lack CpG's including two OCT-1 sites (Jin *et al*., 2000; Fan *et al*., 2002). The p53 response element resides within intron 3, entirely outside of the CpG island. (Zhan *et al*., 1994). A WT1/EGR1-binding site that mediates Myc repression (Amundson *et al*., 1998) is located at −204 and does contain two CpG dinucleotides that show no evidence of methylation in our study.

The significance of methylation of the GADD45a gene must be evaluated at several levels. First, it does occur preferentially in breast cancers and not in normal breast epithelium; hence, it can be considered a marker of disease. The transcript is highly inducible by 5-AzaC in a number of breast cancer cell lines but again, not in normal breast cells indicating that methylation is involved in the regulation of the transcript. Further, steady-state levels of the GADD45a mRNA are typically higher in normal versus cancer cell lines (see Figures 3 and 5). Many primary breast cancers also have relatively low transcript levels (Figure 5). However, there does not appear to be a one-to-one correlation between methylation and induction by 5-AzaC. In particular, the HCC1937 cancer cell line and the MCF10A normal line harbor no detectable methylation at these sites, yet GADD45a is highly induced by 5-AzaC. Basal expression in the HCC1937 line is approximately the same as the levels in normal breast epithelial cells, suggesting that 5-AzaC-induced expression of GADD45a transcription may have several components, that is, an intrinsic one related to methylation of the promoter and an extrinsic one related to methylation of a trans-acting factor. Neither does there seem to be a simple direct correlation between methylation status and basal steady-state expression. Given the limited region of methylation relatively far upstream of the transcriptional initiation, it is not surprising that this signal does not control basal gene expression. We consider it likely that these residues may control the induction of GADD45 in response to some physiologic stimulus with methylation blunting the response.

Of additional interest in this regard is that the HCC1937 line harbors a hemizygous BRCA1 inactivating mutation (Tomlinson *et al*., 1998). Exogenous BRCA1 induces the expression of GADD45a and this relationship may be an important component of the mechanism through which BRCA1 functions in response to DNA damage (Harkin *et al*., 1999; Mullan *et al*., 2001). The prediction would be that the HCC1937 line (also expressing mutant p53) might have low basal GADD45a levels, yet the opposite is observed. This could be indicative of a feedback regulatory loop involving BRCA1 and GADD45a transcription.

The role of GADD45a and other inducible genes in the dramatic 5-AzaC-mediated growth arrest observed in the MDA468 line is of interest. Overexpression of GADD45a results in a G2 arrest in breast cancer cells (Wang X *et al*., 1999; Wang XW *et al*., 2004); therefore, increased transcription of this gene may be at least partially responsible for the phenotypic effect of 5-AzaC

on the breast cancer cell line. The full list of genes specifically induced in the cancer line is impressive for its diversity. Some of the more notable candidates that could be involved in growth suppression include SMURF2 (E3 ubiquitin ligase), STK17A (also known as a death-associated protein kinase), and Inhibin β-A (McPherson *et al*., 1997; Sanjo *et al*., 1998; Lin *et al*., 2000). Other genes identified using this approach offer intriguing possibilities for further study including genes involved in the differentiated phenotype (Keratin 1, Keratin 17, Myosin 1b, and Prostate Differentiation Factor), transcriptional activation of other genes (ETV5, α-Catenin, and v-MAF), and mitosis (Zyxin and CDC25a). This simple approach has identified a series of new target genes that may be involved in the etiology of breast cancer through aberrant methylation. It remains to be determined whether DNA methylation controls these genes directly or indirectly.

Materials and methods

Cell culture and treatment

The human breast cancer cell lines MCF7, MDAMB468, SKBR3, T47D, ZR75-1, MDAMB435, MDAMB361, BT539, BT20, and BT474 were obtained from the American Type Culture Collection and maintained in RPMI 1640 medium supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Life Technologies). Normal breast epithelial cultures and lines were also used in these studies. All of these cells were cultured in DFCI-1 medium (Band *et al*., 1990). Primary human mammary epithelial cells such as the 48N culture were obtained from reduction mammoplasties. The 26NC cell line is a chemically immortalized (dimethylbenzanthracene) derivative of the 26N primary culture and has been maintained in our laboratory for over 8 years (Davis *et al*., 1996). The BE20E6 (E6) line was immortalized by stable transfection of a plasmid expressing the human papillomavirus (HPV) E6 gene (provided by Ray White, University of Utah) and MCF10A are spontaneously immortalized adherent mammary epithelial cells obtained from the Michigan Cancer Foundation. DU99 cells are telomerase-immortalized normal human mammary epithelial cells (Flowers *et al*., in preparation).

Cells were treated with 5-AzaC (Sigma, St Louis, MO, USA) from a 200 mM stock dissolved in 50% acetic acid. Cells were also treated with 100 ng/ml of Trichostatin A from a stock solution of 1 mg/ml in ethanol (Calbiochem, La Jolla, CA, USA). All treatments were carried out in complete medium with cells in logarithmic growth phase. Control cultures were treated with the appropriate vehicle in parallel.

Tissue specimens

Tissues were obtained under an IRB approved protocol from patients undergoing breast surgery at Duke University Medical Center. These specimens were flash frozen and maintained at -135°C. Before extraction, all specimens were sectioned and stained with hematoxylin and eosin to evaluate the percentage of cancer and normal epithelium. Only tissues that were predominantly of the appropriate cell type were used for this study. DNA was extracted using the Purgene DNA Isolation Kit (Gentra Systems, Minneapolis, MN, USA). Total RNA was extracted from flash frozen tumor specimens by the following method: between 50 and 100 mg of tissue was placed in 700 μl RLT (RNeasy Mini Kit, Qiagen, Valencia, CA, USA) and then homogenized using a Mini Beadbeater with Lysing Matrix A (Q-BIOgene, Irvine, CA, USA) for 20 s at room temperature. The homogenate was centrifuged briefly to pellet the beads and the supernatant was drawn through a 21-gauze needle five times before following the remaining steps of the Rneasy Mini Kit. RNA concentration and integrity were assessed by spectrophotometry and agarose gel electrophoresis.

Expression array analysis and Northern blotting

Total RNA was extracted from treated and control cells using the Qiagen RNeasy reagents. Extracted RNA was then treated with DNase (Qiagen, Valencia, CA, USA). The integrity and

Table 2 Primers used for bisulfite sequencing of GADD45[a]

Primer	*Primer positions*[a]	*Primer sequences*[b]	*PCR product*
1F-1R	1462–1487	TTTGAGAAAGGAGAATTTGGGTTGTT	391nt
	1852–1825	CTATTTCCAAAATTAATACACTAAAATC	
2F-2R	1489–1518	GGGATTTTTATATGTGGTTATTAGTTTTTT	320nt
	1808–1778	ACAAATTACATTAAAAAAATATCTCCAAACC	
3F-3R	1934–1950	GGTAGTTTGTTTTAGTTTAAGTTGAGG	337nt
	2270–2244	CCTACCAACCACTAAAAAACAAAAAAC	
4F-4R1	2207–2232	TTGGTTGAGGGTTGGTAGGATAATTT	272nt
	2478–2454	CCTACTTTCTACACTCACTCACAAA	
–4R2	2564–2537	CCTCCAAAATCATATTACAAACTACAAA	358nt
5F-5R	2539–2563	TGTAGTTTGTAATATGATTTTGGAG	284nt
	2822–2800	CCACTACTACCCTTTACAAAAAC	
6F1-6R	2798–2821	GGGTTTTTGTAAAGGGTAGTAGTG	328nt
	3125–3101	CTTTACTAAACACTTCCTCCAAAAC	
6F2	2851–2875	TGGTAGTTTGTGGTAGGGGTATTTT	275nt

[a]Positions derived from Genbank sequence L24498. [b]In some cases, multiple reverse primers were used both for amplification and sequencing to resolve difficult areas

concentration of RNA was assessed using the Agilent Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The RNA was then used for cRNA probe synthesis and hybridization to the Affymetrix (Santa Clara, CA, USA) human U133A array as previously described (West *et al*., 2001). Northern blotting, using total RNA and probed with ^{32}P-labeled DNA probes, was carried out as previously described (Marks *et al*., 1988). All probes were derived from RT–PCR products using the GenBank deposited coding sequence for each gene (primer sequences available upon request).

Bisulfite modification and sequencing

Sodium bisulfite modification of DNA was performed based upon a previously published method (Grunau *et al*., 2001). Briefly, 1 μg of genomic DNA was denatured with 3 M NaOH for 20 min at 42°C followed by deamination in saturated sodium bisulfite/10 mM hydroquinone (Sigma, St Louis, MO, USA) solution pH 5.0 for 4 h at 55°C. The DNA was desalted using the Wizard DNA Clean-up System (Promega, Madison, WI, USA), then desulfonated in 3 M NaOH (20 min at 37°C) and ethanol precipitated. Samples were resuspended in 25 μl water and stored at 4°C for no more than 2 weeks before amplification.

For sequencing, bisulfite-treated DNA was PCR amplified with primers specific for the bisulfite-converted sequences (Table 2). The PCR products were resolved on agarose gels, purified using QiaQuick Gel Extraction systems (Qiagen, Valencia, CA, USA) followed by cycle sequencing (Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit; Amersham Biosciences, Piscataway, NJ, USA) with both amplification primers. Reactions were then resolved on 5% acrylamide/urea gels and imaged by autoradiography.

Acknowledgements
We are indebted to Nancy Glover for her ongoing technical assistance. This work was supported by the NCI funded Early Detection Research Network Grant, CA84955 awarded to JRM.

References

Amundson SA, Zhan Q, Penn LZ and Fornace Jr AJ. (1998). *Oncogene*, **17,** 2149–2154.
Band V and Sager R. (1989). *Proc. Natl. Acad. Sci. USA*, **86,** 1249–1253.
Band V, Zajchowski D, Kulesa V and Sager R. (1990). *Proc. Natl. Acad. Sci. USA*, **87,** 463–467.
Baylin SB. (2004). *Novart. Found. Symp.*, **259,** 226–233.
Bovenzi V, Le NL, Cote S, Sinnett D, Momparler LF and Momparler RL. (1999). *Anticancer Drugs*, **10,** 471–476.
Coombes MM, Briggs KL, Bone JR, Clayman GL, El-Naggar AK and Dent SY. (2003). *Oncogene*, **22,** 8902–8911.
Cui J, Rohr LR, Swanson G, Speights VO, Maxwell T and Brothman AR. (2001). *Prostate*, **46,** 249–256.
Dairkee SH, Puett L and Hackett AJ. (1988). *J. Natl. Cancer Inst.*, **80,** 691–695.
Davis P, Bazar K, Huper G, Lozano G, Marks J and Iglehart JD. (1996). *Oncogene*, **13,** 1315–1322.
Dobrovic A and Simpfendorfer D. (1997). *Cancer Res.*, **57,** 3347–3350.
Domann FE, Rice JC, Hendrix MJ and Futscher BW. (2000). *Int. J. Cancer*, **85,** 805–810.
Dowell JE and Minna JD. (2004). *J. Clin. Oncol.*, **22,** 1353–1355.
Egger G, Liang G, Aparicio A and Jones PA. (2004). *Nature*, **429,** 457–463.
Engelman JA, Zhang XL and Lisanti MP. (1999). *FEBS Lett.*, **448,** 221–230.
Evron E, Dooley WC, Umbricht CB, Rosenthal D, Sacchi N, Gabrielson E, Soito AB, Hung DT, Ljung B, Davidson NE and Sukumar S. (2001a). *Lancet*, **357,** 1335–1336.
Evron E, Umbricht CB, Korz D, Raman V, Loeb DM, Niranjan B, Buluwela L, Weitzman SA, Marks J and Sukumar S. (2001b). *Cancer Res.*, **61,** 2782–2787.
Fackler MJ, McVeigh M, Evron E, Garrett E, Mehrotra J, Polyak K, Sukumar S and Argani P. (2003). *Int. J. Cancer*, **107,** 970–975.
Fan W, Jin S, Tong T, Zhao H, Fan F, Antinore MJ, Rajasekaran B, Wu M and Zhan Q. (2002). *J. Biol. Chem.*, **277,** 8061–8067.
Ferguson AT, Evron E, Umbricht CB, Pandita TK, Chan TA, Hermeking H, Marks JR, Lambers AR, Futreal PA, Stampfer MR and Sukumar S. (2000). *Proc. Natl. Acad. Sci. USA*, **97,** 6049–6054.
Foster SA, Wong DJ, Barrett MT and Galloway DA. (1998). *Mol. Cell. Biol.*, **18,** 1793–1801.
Grunau C, Clark SJ and Rosenthal A. (2001). *Nucleic Acids Res.*, **29,** E65.
Harkin DP, Bean JM, Miklos D, Song YH, Truong VB, Englert C, Christians FC, Ellisen LW, Maheswaran S, Oliner JD and Haber DA. (1999). *Cell*, **97,** 575–586.
Jin S, Zhao H, Fan F, Blanck P, Fan W, Colchagie AB, Fornace Jr AJ and Zhan Q. (2000). *Oncogene*, **19,** 4050–4057.
Krop IE, Sgroi D, Porter DA, Lunetta KL, LeVangie R, Seth P, Kaelin CM, Rhei E, Bosenberg M, Schnitt S, Marks JR, Pagon Z, Belina D, Razumovic J and Polyak K. (2001). *Proc. Natl. Acad. Sci. USA*, **98,** 9796–9801.
Li Q, Ahuja N, Burger PC and Issa JP. (1999). *Oncogene*, **18,** 3284–3289.
Lin X, Liang M and Feng XH. (2000). *J. Biol. Chem.*, **275,** 36818–36822.
Marks J, Lin J, Miller D, Lozano G, Herbert J and Levine AJ. (1988). *Prog. Clin. Biol. Res.*, **284,** 163–186.
McPherson SJ, Thomas TZ, Wang H, Gurusinghe CJ and Risbridger GP. (1997). *J. Endocrinol.*, **154,** 535–545.
Mullan PB, Quinn JE, Gilmore PM, McWilliams S, Andrews H, Gervin C, McCabe N, McKenna S, White P, Song YH, Maheswaran S, Liu E, Haber DA, Johnston PG and Harkin DP. (2001). *Oncogene*, **20,** 6123–6131.
Muller HM, Widschwendter A, Fiegl H, Ivarsson L, Goebel G, Perkmann E, Marth C and Widschwendter M. (2003). *Cancer Res.*, **63,** 7641–7645.
Murphy SK and Jirtle RL. (2001). *Cancer Handbook* Alison MR, Gooderham NG (eds). Nature Publishing Reference: London, pp. 317–333.
Oshiro MM, Watts GS, Wozniak RJ, Junk DJ, Munoz-Rodriguez JL, Domann FE and Futscher BW. (2003). *Oncogene*, **22,** 3624–3634.
Oue N, Matsumura S, Nakayama H, Kitadai Y, Taniyama K, Matsusaki K and Yasui W. (2003). *Oncology*, **64,** 423–429.
Parker BS, Cutts SM, Nudelman A, Rephaeli A, Phillips DR and Sukumar S. (2003). *Cancer Biol. Ther.*, **2,** 259–263.
Paz MF, Fraga MF, Avila S, Guo M, Pollan M, Herman JG and Esteller M. (2003). *Cancer Res.*, **63,** 1114–1121.
Pechoux C, Gudjonsson T, Ronnov-Jessen L, Bissell MJ and Petersen OW. (1999). *Dev. Biol.*, **206,** 88–99.

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Aksien LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO and Botstein D. (2000). *Nature*, **406,** 747–752.

Sanjo H, Kawai T and Akira S. (1998). *J. Biol. Chem.*, **273,** 29066–29071.

Sirchia SM, Ferguson AT, Sironi E, Subramanyan S, Orlandi R, Sukumar S and Sacchi N. (2000). *Oncogene*, **19,** 1556–1563.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de RM, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein LP and Borresen-Dale AL. (2001). *Proc. Natl. Acad. Sci. USA*, **98,** 10869–10874.

Tomlinson GE, Chen TT, Stastny VA, Virmani AK, Spillman MA, Tonk V, Blum JL, Schneider NR, Wistuba II, Shay JW, Minna JD and Gazdar AF. (1998). *Cancer Res.*, **58,** 3237–3242.

Umbricht CB, Evron E, Gabrielson E, Ferguson A, Marks J and Sukumar S. (2001). *Oncogene*, **20,** 3348–3353.

van Noesel MM, van BS, Voute PA, Herman JG, Pieters R and Versteeg R. (2003). *Genes Chromosomes Cancer*, **38,** 226–233.

Wang W, Nahta R, Huper G and Marks JR. (2004). *Mol. Cancer Res.*, **2,** 442–452.

Wang XW, Zhan Q, Coursen JD, Khan MA, Kontny HU, Yu L, Hollander MC, O'Connor PM, Fornace Jr AJ and Harris CC. (1999). *Proc. Natl. Acad. Sci. USA*, **96,** 3706–3711.

West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson Jr JA, Marks JR and Nevins JR. (2001). *Proc. Natl. Acad. Sci. USA*, **98,** 11462–11467.

Wiechen K, Diatchenko L, Agoulnik A, Scharff KM, Schober H, Arlt K, Zhumabayeva B, Siebert PD, Dietel M, Schafer R and Sers C. (2001). *Am. J. Pathol.*, **159,** 1635–1643.

Woodcock DM, Linsenmeyer ME, Doherty JP and Warren WD. (1999). *Br. J. Cancer*, **79,** 251–256.

Zelent A, Waxman S, Carducci M, Wright J, Zweibel J and Gore SD. (2004). *Clin. Cancer Res.*, **10,** 4622–4629.

Zhan Q, Lord KA, Alamo Jr I, Hollander MC, Carrier F, Ron D, Kohn, KW, Hoffman B, Liebermann DA and Fornace Jr AJ. (1994). *Mol. Cell. Biol.*, **14,** 2361–2371.