

# Cell-Type-Specific Responses to Chemotherapeutics in Breast Cancer

Melissa A. Troester,<sup>1</sup> Katherine A. Hoadley,<sup>2,3</sup> Therese Sørlie,<sup>6</sup> Brittney-Shea Herbert,<sup>7</sup> Anne-Lise Børresen-Dale,<sup>6</sup> Per Eystein Lønning,<sup>8</sup> Jerry W. Shay,<sup>9</sup> William K. Kaufmann,<sup>1,4,5</sup> and Charles M. Perou<sup>1,2,3,4</sup>

<sup>1</sup>Department of Pathology and Laboratory Medicine, <sup>2</sup>Curriculum in Genetics and Molecular Biology, <sup>3</sup>Department of Genetics, <sup>4</sup>Lineberger Comprehensive Cancer Center, and <sup>5</sup>Center for Environmental Health and Susceptibility, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; <sup>6</sup>Departments of Genetics, The Norwegian Radium Hospital, Montebello, Oslo, Norway; <sup>7</sup>Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana; <sup>8</sup>Department of Medicine, Section of Oncology, Haukeland University Hospital, Bergen, Norway; and <sup>9</sup>Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, Texas

## ABSTRACT

Recent microarray studies have identified distinct subtypes of breast tumors that arise from different cell types and that show statistically significant differences in patient outcome. To gain insight into these differences, we identified *in vitro* and *in vivo* changes in gene expression induced by chemotherapeutics. We treated two cell lines derived from basal epithelium (immortalized human mammary epithelial cells) and two lines derived from luminal epithelium (MCF-7 and ZR-75-1) with chemotherapeutics used in the treatment of breast cancer and assayed for changes in gene expression using DNA microarrays. Treatment doses for doxorubicin and 5-fluorouracil were selected to cause comparable cytotoxicity across all four cell lines. The dominant expression response in each of the cell lines was a general stress response; however, distinct expression patterns were observed. Both cell types induced DNA damage-response genes such as *p21<sup>waf1</sup>*, but the response in the luminal cells showed higher fold changes and included more p53-regulated genes. Luminal cell lines repressed a large number of cell cycle-regulated genes and other genes involved in cellular proliferation, whereas the basal cell lines did not. Instead, the basal cell lines repressed genes that were involved in differentiation. These *in vitro* responses were compared with expression responses in breast tumors sampled before and after treatment with doxorubicin or 5-fluorouracil/mitomycin C. The *in vivo* data corroborated the cell-type-specific responses to chemotherapeutics observed *in vitro*, including the induction of *p21<sup>waf1</sup>*. Similarities between *in vivo* and *in vitro* responses help to identify important response mechanisms to chemotherapeutics.

## INTRODUCTION

The response of breast tumors to cytotoxic chemotherapeutic agents such as doxorubicin (DOX) and 5-fluorouracil (5FU) varies significantly across individuals. Sensitivity to these compounds is associated with *HER2* overexpression (1), *p53* status (2), and *topoisomerase II $\alpha$*  amplification or deletion (3), but the mechanisms of chemoresistance are still poorly understood. To better understand variations in clinical responses to treatment, recent studies have used gene expression patterns to identify major biological subtypes of breast cancer. These studies identified a previously unrecognized tumor subtype with characteristics of breast basal epithelium (4–8); Basal-like tumors are estrogen receptor  $\alpha$  (ER $\alpha$ )-negative, do not overexpress *HER2*, and they have a poor prognosis compared with tumors derived from luminal epithelium (5, 7).

Basal and luminal breast tumors are often treated with the same

chemotherapeutic agents, but little is known about how each cell type responds to these drugs. To improve our understanding of how basal and luminal epithelium differ in their responses to chemotherapy, we selected two representative cell lines from each of these breast epithelial cell types to study; two human mammary epithelial (HME) cell lines immortalized by the overexpression of the catalytic subunit of telomerase (hTERT) represent the basal subtype, and two breast tumor-derived cell lines (MCF-7 and ZR-75-1) represent the luminal subtype (9). All four cell lines express wild-type p53 protein. True to their corresponding tumor subtypes, the HME lines are ER $\alpha$ -negative and the luminal cancer cell lines are ER $\alpha$ -positive. We treated all four cell lines with DOX and 5FU and performed expression profiling to identify patterns of response.

Transcriptional profiling is a powerful approach for investigating cellular responses to drugs. This approach has led to greater understanding of pathway inhibition and off-target drug effects (10), the response of yeast to genotoxic agents and environmental stresses (11), and the effects of different kinds of DNA-damaging agents in human cells (12). Our analyses showed that the transcriptional responses of the basal and luminal cell lines to chemotherapeutics are quite distinct. We also correlated our *in vitro* data with *in vivo* data on breast tumors sampled before and after treatment with DOX or 5FU/mitomycin C (4, 5, 7), and we identified commonalities. Taken together, these *in vitro* and *in vivo* data sets illustrate that cell type is an important determinant of response to commonly used chemotherapeutics.

## MATERIALS AND METHODS

**Cells and Culture Conditions.** HME31-hTERT no. 16C (ME16C) cells were derived from a clone of a finite life span HME cell culture isolated from the uninvolved tissue of a 53-year-old woman with unilateral breast cancer and no family history of breast cancer; HME31 postselection cells were infected with the retrovirus pBabe-puro-hTERT and an immortal population was established (ME16C). A second immortal HME clone, HME-CC, was a gift from Christopher Counter (Duke University, Durham, NC); to derive the HME-CC cells, a HME cell isolate (Clonetics) was infected with the retrovirus pBabe-hygro-hTERT and an immortal population was established. ME16C and HME-CC cells were maintained in mammary epithelium growth media (Cambrex Bio Science, Walkersville, MD). Karyotyping on the HME-CC and ME16C lines was conducted as described in Wang and Federoff (13) at the University of North Carolina at Chapel Hill Chromosome Imaging Core Facility. A single isolate of HME-CC was found to be trisomic for chromosome 20 in 65% of metaphase spreads and contained 9q+ and 18q+ in 25% of metaphase spreads. Two different isolates of ME16C were studied, and both were shown to be trisomic for chromosome 20 in 50% of metaphase spreads and marker chromosomes 3p- and iso10q were recognized in 25% of metaphase spreads. MCF-7 cells (a gift from F. Tamanoi, University of California-Los Angeles, Los Angeles, CA) and ZR-75-1 cells (American Type Culture Collection) were maintained in RPMI 1640 supplemented with L-glutamine (Life Technologies, Inc.), 10% fetal bovine serum (Sigma), and 50 units/ml penicillin/50 units/ml streptomycin. Before conducting these experiments and at regular intervals, thereafter, all of the cell lines were tested by the Lineberger Comprehensive Cancer Center Tissue Culture Facility and were found to be negative for *Mycoplasma* contamination. Cells were maintained at 37°C and 5% carbon dioxide.

Received 1/13/04; revised 3/10/04; accepted 4/8/04.

**Grant support:** This work was supported by National Institute of Environmental Health Sciences (NIEHS), NIH, Grants U19-ES11391, P30-ES10126, F32-ES012374 (M. Troester), and T32-ES07017 (M. Troester), and by NIH Grant T32-GM07092 (K. Hoadley).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org>); M. Troester and K. Hoadley contributed equally to the work.

**Requests for reprints:** Charles Perou, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Campus Box 7295, Chapel Hill, NC 27599-7295. Phone: (919) 843-5740; Fax: (919) 843-5718; E-mail: cperou@med.unc.edu.

**Cytotoxicity Assay.** A mitochondrial dye conversion assay (Cell Titer 96, Promega) was used to quantitate cell line responses to chemotherapeutics. Five thousand cells were seeded per well of a 96-well plate. The cells were allowed to adhere overnight and then the media was replaced with fresh media containing a range of drug doses (DOX, 0–10  $\mu\text{M}$ ; 5FU, 0–10  $\text{mM}$ ). After 36 h of drug treatment, 15  $\mu\text{l}$  of tetrazolium dye solution were added, and culture was incubated at 37°C for 1 h before adding Cell Titer 96 Stop Solution. Dye conversion products were allowed to solubilize in a humidified chamber overnight, and absorbance was measured at 570 nm (minus background absorbance at 650 nm).

**Estimating the  $IC_{50}$ .** The  $IC_{50}$  for 36 h of treatment for each drug in each cell line was estimated using nonlinear regression (SAS Statistical Software, Cary, NC) and the following relationship:

$$y = \frac{k}{1 + \left(\frac{x}{x_0}\right)^b}$$

where  $y$  is the absorbance value corrected for medium-only wells, and  $x$  is the dose (in  $\mu\text{M}$  for DOX and in  $\text{mM}$  for 5FU; Ref. 14). The parameter  $k$  represents the value of  $y$  (in absorbance units) when  $x$  is zero. The  $IC_{50}$  value is represented by  $x_0$ , and  $-b$  is a unitless scalar representing the slope of the line on logit-log scale. In our experiments, if  $b$  is greater than zero, the response is monotonically decreasing.

**Collection of mRNA for Microarray Experiments.** Cell lines were grown in 150-mm dishes to 70–80% confluence and then were treated for 3, 12, 24, or 36 h with DOX (doxorubicin hydrochloride) or 5FU (Sigma) at the  $IC_{50}$  concentration. Cells were harvested by scraping and mRNA was isolated using a Micro-FastTrack kit (Invitrogen). To generate feeding control (sham) mRNA samples for each cell line, cells were treated with medium only, in parallel with drug-treated samples. Individual harvests of treated or sham mRNA were not pooled before microarray analysis. However, a reference mRNA sample was generated for each of the four cell lines by harvesting untreated mRNA from each cell line at 80% confluence and then pooling four harvests together (*i.e.*, four MCF-7 harvests were pooled and served as the reference mRNA for all MCF-7 experiments), using each cell line as its own reference controlled for baseline differences between the cell lines.

**Microarray Experiments.** Syntheses of labeled cDNA were performed as described previously (4), with reference cDNAs labeled with Cy3-dUTP and treated and sham cDNAs labeled with Cy5-dUTP. Each cDNA sample mix was hybridized overnight to an oligonucleotide microarray created in the University of North Carolina at Chapel Hill Genomics Core Facility (<http://genomicscore.unc.edu/>). These microarrays were created by spotting the Compugen Human oligomers library representing 18,861 human genes (<http://www.labonweb.com/chips/libraries.html>) onto coated microarray slides (Corning no. 40016). All of the microarray raw data tables are available at the University of North Carolina Microarray Database (<https://genome.unc.edu/>) at the supporting website for this article,<sup>10</sup> and have been deposited in the Gene Expression Omnibus under the accession number of GSE763 (submitter C. Perou). The direction of gene expression change was verified by real-time reverse transcription PCR for a subset of samples using commercially available primers (Applied Biosystems) for *p21<sup>waf</sup>*, *ferredoxin reductase*, *prostate differentiation factor*, and *inhibitor of DNA binding 3*. To normalize the target sample variation, we used the average of three control genes: *splicing factor 3A subunit 1 (SF3A1)*, *pumilio homolog 1 (PUM1)*, and  $\beta$ -actin. *SF3A1* and *PUM1* were selected as control genes because they had the lowest variation across the tumor data set presented in Perou *et al.* (4). Sham-adjusted real-time-PCR values were regressed on the average of sham-adjusted  $\log_2(\text{red}/\text{green ratio})$  array values for each gene; the regression yielded a positive slope of 4.3, Pearson  $r = 0.75$  (data not shown).

**SAM Using Cell Line Data.** Genes that were significantly induced or repressed were identified using the significance analysis of microarrays (SAM) package Add-In for Microsoft Excel (15). Before conducting SAM, genes were excluded that did not have a mean signal intensity greater than twice the median background value for both the red and green channels in at least 70% of the experiments. For genes that passed these filtering criteria, the log-base-2 of median red intensity over median green intensity was calculated.

The gene expression changes in the 3 h time points were very modest (data not shown); therefore, this time point was excluded from all analyses. To identify genes whose steady-state expression was altered, we combined the 12-, 24-, and 36-h time points for each cell line and treatment group into a single class. This eliminated artifacts caused by random temporal variation in steady-state RNA levels. Two or three replicate arrays were used for each treatment condition for each cell line.

To identify a general stress response pattern, DOX- and 5FU-treated experiments were combined into a single class and compared against sham experiments for each cell line (*i.e.*, MCF-7 DOX- and 5FU-treated *versus* MCF-7 sham). Missing data were imputed using SAM with 100 permutations and 10 k-nearest neighbors. A two-class unpaired SAM analysis was conducted on the imputed data set. The SAM  $\delta$  values were adjusted to obtain the largest gene list that gave a false discovery rate of less than 5%.

**SAM Using Breast Tumor Data.** All of the tumor data were published previously (4, 5, 7) except for data from five new tumors samples collected after chemotherapy, which are now publicly available at the Stanford Microarray Database (<http://genome-www5.stanford.edu/>). This breast tumor dataset encompassed two different cohorts of breast cancer patients, one of which received neoadjuvant DOX and a second of which received neoadjuvant 5FU and mitomycin C (16, 17); in both cohorts, we obtained samples of the tumors before therapy and at the time of surgical resection (after therapy sample). All before- and after-samples were labeled with Cy5-dUTP, mixed with Cy3-dUTP-labeled Stanford common reference sample, and hybridized to cDNA microarrays produced at Stanford University (4). The gene expression patterns of all of the before-samples were compared with the gene expression patterns of all of the after-samples using a two-class, unpaired SAM analysis. A total of 81 before- and 50 after-samples were assessed, representing all of the tumor subtypes identified in Sørli *et al.* (5). Consistent with the *in vitro* data analyses, SAM  $\delta$  values were adjusted to obtain the largest gene list that gave a false discovery rate of less than 5%.

To study the genes differentially regulated in basal or luminal tumor subtypes separately, we also classified each tumor into one of two groups using the intrinsic list of Sørli *et al.* (7): one group contained those tumors that represented the luminal epithelium-derived tumors (both Luminal A and B for a total of 51 before- and 30 after-samples) and a second group representing the basal subtype (for a total of 11 before- and 10 after-samples). Each of these two groups was then analyzed using a two-class unpaired SAM analysis; gene expression patterns of before-samples were compared with gene expression patterns of after-samples, and false discovery rates were estimated.

**Hierarchical Clustering of Gene Expression Responses.** Average linkage hierarchical cluster analysis using Pearson correlation was conducted using the program Cluster, and the data were visualized in Treeview (18, 19). To visualize the gene expression patterns for the luminal cell lines, the data from the union of the genes identified by SAM for MCF-7 and ZR-75-1 were identified, combined into a nonredundant list, and clustered. These clusters illustrate the fold change relative to control levels for each gene. Following the same procedure, data from the union of the gene sets identified for ME16C and HME-CC were extracted, combined into a nonredundant list, and clustered. Cluster analysis was also performed using the top 100 genes identified by SAM for distinguishing between luminal and basal cell lines responses to DOX-treatment and 5FU-treatment. For all of the clusters, genes were excluded that did not have a mean intensity greater than twice the median background for both the red and green channel in at least 80% of the experiments. For the breast tumor data, similar gene filtering, SAM, and clustering analyses were performed.

**Western Blot Analysis.** Cells were treated for 36 h with DOX or 5FU at the 36-h  $IC_{50}$  concentration. Cells were rinsed with PBS and then harvested with M-PER Mammalian Protein Extraction reagent (Pierce) containing Halt Protease Inhibitor and 5 mM EDTA (Pierce). Protein concentrations were determined using Micro BCA Protein Assay Reagent kit (Pierce). Lysates were combined with 2 $\times$  Laemmli Sample Buffer (Bio-Rad) containing  $\beta$ -mercaptoethanol and were boiled for 5 min. Forty  $\mu\text{g}$  of protein were electrophoresed on a 4–20% Tris-HCl Criterion precast gel (Bio-Rad) and transferred to a Hybond-P membrane (Amersham Biosciences) by electroblotting. The blots were probed with antibodies against p21<sup>waf1</sup> (Neomarkers; Ab-11) and  $\beta$ -actin (Abcam, AC-15). Blots were washed three times with Tris-buffered saline supplemented with 0.1% TWEEN and then were probed with antimouse IgG horseradish peroxidase-linked whole antibody from sheep (Amersham). The

<sup>10</sup> All figures and text can be obtained at the supporting website for this article, <https://genome.unc.edu/pubsup/TOX/>.

blots were rewashed, and detection was by enhanced chemiluminescence (SuperSignal West Pico Chemiluminescent substrate; Pierce).

## RESULTS

**Cell-Type-Specific Transcriptional Responses *in Vitro*.** To investigate the response of four distinct cell lines to treatment with chemotherapeutics, we used a mitochondrial dye conversion assay [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT)] to measure the cytotoxicity of 5FU and DOX after 36 h of treatment. We then treated these four cell lines (MCF-7, ZR-75-1, ME16C, and HME-CC) with two mechanistically distinct chemotherapeutics (DOX and 5FU) at doses that produced similar levels of toxicity across all four lines ( $IC_{50}$ ). The  $IC_{50}$  concentrations and their 95% confidence intervals are shown in Table 1.

Our experimental design was aimed at identifying the steady-state and cell-type-specific transcriptional response of these cell lines and was not focused on defining chemotherapeutic-specific responses or temporal variation. By combining 12-, 24-, and 36-h-treated experiments into a single class for supervised analyses, we avoided temporal artifacts and identified only those gene expression changes that were consistent over time.

**Chemotherapeutic-Induced Gene Expression Patterns in Luminal Cell Lines.** Differences between basal and luminal cell lines responding to treatment were immediately evident given the absolute number of genes whose expression was altered when treated experiments were compared with sham experiments (Table 2). In each luminal cell line, ~10-fold more genes were altered in response to drug. To visualize these expression changes, we combined the SAM-supervised lists for the two luminal cell lines and performed a hierarchical clustering analysis (Fig. 1 and Supplemental Fig. 1 for the complete cluster diagram with all gene names). Each cell line had a unique expression response to chemotherapy that was distinct enough to cause the two treated luminal lines to cluster into different dendrogram branches (Fig. 1B). Clusters of genes that distinguish between MCF-7 and ZR-75-1 responses can be found in Supplemental Fig. 1.

Common features dominated the overall expression patterns in the two luminal cell lines despite some cell-line-specific responses. For example, a cluster of genes that reflect cell proliferation *in vitro* and *in vivo* (4, 5, 9, 20, 21) was identified (Fig. 1C). These genes had slightly increased expression in the sham experiments because of feeding but had greatly diminished expression during drug treatment. This cluster included well-characterized cell cycle regulators (20) such as *cyclin A2*, *cyclin B1*, *cell division cycle 2*, and many genes involved in specific phases of the cell cycle such as *Ki-67*, *ribonucleotide reductase M2*, *polo-like kinase*, and *topoisomerase IIA*. This cluster also included *pituitary tumor-transforming 1*, a gene that is overexpressed in many cancers, is tumorigenic *in vivo*, and has been shown to bind p53 (22). The gene product of *serine/threonine kinase*

Table 2 Number of oligonucleotides significantly altered by treatment with chemotherapeutic as determined by Significance Analysis of Microarrays

Sample	No. of oligonucleotides	No. of false significant <sup>a</sup>
MCF7 treated	998	48.5 (4.9)
ZR-75-1 treated	783	38.5 (4.9)
ME16C treated	84	3.3 (3.9)
HME-CC treated	84	3.0 (3.4)
All tumors	28	0.7 (2.5)
Luminal tumors	15	0.8 (5.3)
Basal tumors	10	2 (20)

<sup>a</sup> Percentage false significant indicated in parentheses.

6 (*STK6*) is also present and has cell-cycle-dependent expression, with maximum expression in G<sub>2</sub>-M (23); in addition, *STK6* has been shown to bind chromosome 20 open reading frame 1 (24), which is also repressed and in this cluster. *Squalene epoxidase* was down-regulated in the luminal cell lines and is a gene that was differentially expressed between luminal and basal tumors *in vivo* (5).

A large cluster of genes that include DNA-damage and stress-response genes was up-regulated in response to treatment in the luminal lines (Fig. 1D). *p21<sup>waf1</sup>* and the DNA-damage response gene *GADD45* were induced strongly in both lines. Also present in this cluster were a number of genes involved in xenobiotic metabolism including *carboxylesterase 2*, *epoxide hydrolase*, and *ferredoxin reductase*. The latter two of these genes, along with *p21<sup>waf1</sup>* and *GADD45*, are all known to be p53-regulated (25, 26). Induction of xenobiotic metabolism genes may represent a stereotyped adaptive response of the cell to DNA damage.

**Chemotherapeutic-Induced Gene Expression Patterns in Basal Cell Lines.** A much smaller list of genes showed significantly altered expression in the ME16C or HME-CC basal cell lines (Fig. 2 and Supplemental Fig. 2). Using the combined list of genes that were significantly altered in either basal cell line in a hierarchical clustering analysis showed that the basal lines did not cluster as distinctly as the luminal cell lines (Fig. 2B). Within the treated branch, some time points for the HME-CC line clustered on separate branches, but the drug-treated ME16C experiments all grouped together. This suggests that the changes induced in basal cells treated with chemotherapeutics were subject to more temporal variation. The changes also appeared more subtle; strong signatures like those observed in the luminal cell lines were not nearly as evident in these basal cell lines. We identified a small cluster of genes that was slightly induced in the sham experiments, but that was down-regulated in the treated experiments (Fig. 2C). Many of these genes are involved in cellular differentiation including *integrin-β4*, *collagen type XIIα1*, *COX2*, and *core promoter element-binding protein*. A proliferation signature (similar to Fig. 1C) was not identified in the treated basal lines. However, a set of genes involved in the DNA damage and/or stress response was identified (Fig. 2D) and similar to the luminal cell lines (Fig. 1D), *p21<sup>waf1</sup>* was induced, although less dramatically. Several xenobiotic metabolism genes were also up-regulated including the p53-regulated genes *ferredoxin reductase* and *quinone oxidoreductase homolog*, as well as *glutathione-S-transferase π (GST-π)*. *Inhibitor of DNA binding 3*, an inhibitor of differentiation (27), was also induced in both basal cell lines.

**Comparison of Basal versus Luminal Cell Lines.** In the analyses above, we identified genes that differed between shams and treated samples on a cell-line by cell-line basis. To assess differences between basal and luminal cell lines, we first compared the lists of chemotherapeutic-induced genes for the luminal (1000 genes) and the basal (100 genes) cell lines. There were 42 genes on both lists, but only two genes (*chitinase 3-like 1* and *p21<sup>waf1</sup>*) were up-regulated in all four lines (no genes down-regulated). We then used SAM to directly

Table 1 Estimated  $IC_{50}$  for 5-fluorouracil and doxorubicin based on mitochondrial dye conversion assay

Cell line	$IC_{50}$ <sup>a</sup>	Treatment dose
5-fluorouracil		
MCF-7	0.34 mM (0.13–0.55)	0.3 mM
ZR-75-1	3.3 mM (2.8–3.7)	3.0 mM
ME16C	0.064 mM (0.055–0.074)	0.06 mM
HME-CC	0.011 mM (0.009–0.013)	0.01 mM
Doxorubicin		
MCF-7	0.86 μM (0.74–0.97)	0.9 μM
ZR-75-1	0.43 μM (0.37–0.50)	0.4 μM
ME16C	0.52 μM (0.49–0.54)	0.5 μM
HME-CC	0.16 μM (0.14–0.18)	0.2 μM

<sup>a</sup> Values in parentheses represent 95% confidence intervals.



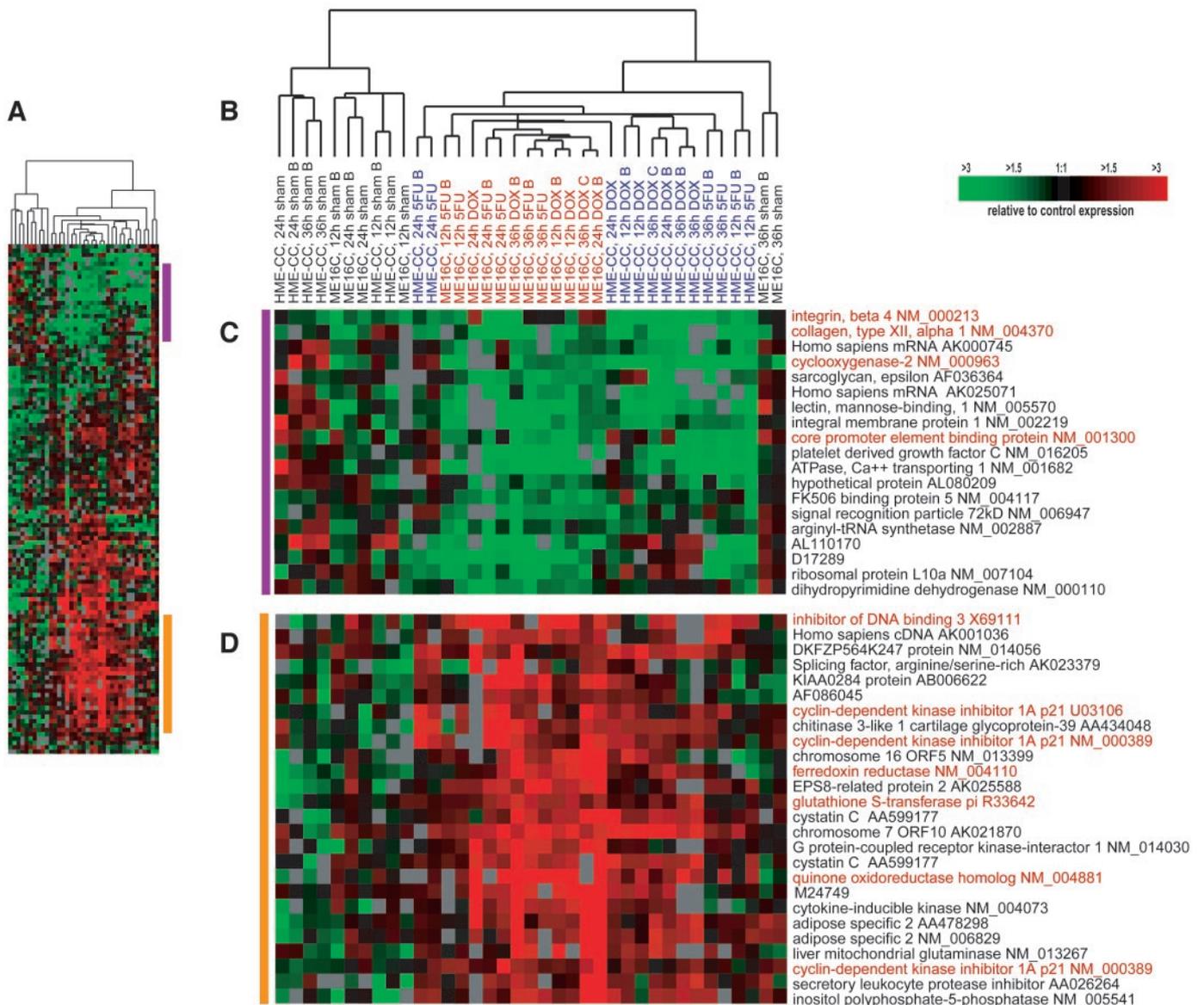


Fig. 2. Gene expression pattern for genes significantly altered by chemotherapeutics in ME16C and HME-CC cell lines. Cluster analysis was conducted using 25 treated and 12 sham experiments. Data from the union of the genes identified by Significance Analysis of Microarray (SAM) for ME16C or HME-CC were identified and combined into a nonredundant list; the scaled-down cluster diagram is shown in **A** (complete cluster available in Supplemental Fig. 2). Colored bars illustrate the location of clusters **C** and **D**. The dendrogram in **B** shows that the experiments were divided into two primary branches, one consisting primarily of treated experiments and one consisting exclusively of shams. A large cluster enriched for genes involved in differentiation and possibly correlated with proliferation (**C**) and a cluster enriched for genes involved in responding to stress or DNA damage (**D**) are shown. Highlighted in red, genes discussed in the text.

identify the set of genes that distinguished the treated luminal lines from the treated basal lines. With a 5% false discovery rate, 920 genes were statistically different. The top 100 distinguishing genes were used to cluster the experiments (Fig. 3 and Supplemental Fig. 3). The grouping of the cell lines identified two primary dendrogram branches (Fig. 3B), one representing only basal cell lines and one representing predominantly luminal cell lines. A few of the late time points for the basal cell lines fell within the luminal branch but remained distinct on their own secondary branches because of their unique expression profiles. This finding again illustrates the temporal variability across the basal cell lines time points.

Consistent with our previous analysis, the luminal cell lines showed a greatly reduced proliferation signature, which was relatively unchanged in the basal cell lines (Fig. 3, C and D). This gene set included *retinoblastoma 1*, *ribonucleotide reductase M2*, *MCM4*, *chromosome 20 open reading frame 1* and *pituitary tumor-transform-*

*ing 1*, all of which regulate cell proliferation or have cell cycle-dependent expression (20). A cluster of genes whose expression was induced in luminal lines and repressed in basal lines is shown in Fig. 3E, whereas the gene set in Fig. 3F was induced in both cell types, but was more highly induced in luminal cells versus basal cells. Among these genes was *X-box binding protein 1 (XBPI)*, a gene whose expression was previously shown to be highly expressed in luminal tumors *in vivo* (5). *XBPI* is a transcription factor involved in mediating the unfolded protein response (28), which may represent a stress response that is more prominent in secretory luminal cells. *HER2* also appeared to be induced more distinctly in luminal cells treated with chemotherapeutics (Fig. 3E). *HER2* has been extensively studied in breast cancer, and it has been shown that MCF-7 cells that overexpress *HER2* retained their proliferative advantage after DOX treatment in a human breast cancer xenograft model (29). Fig. 3F illustrates that the DNA damage response was much more dramatic in the

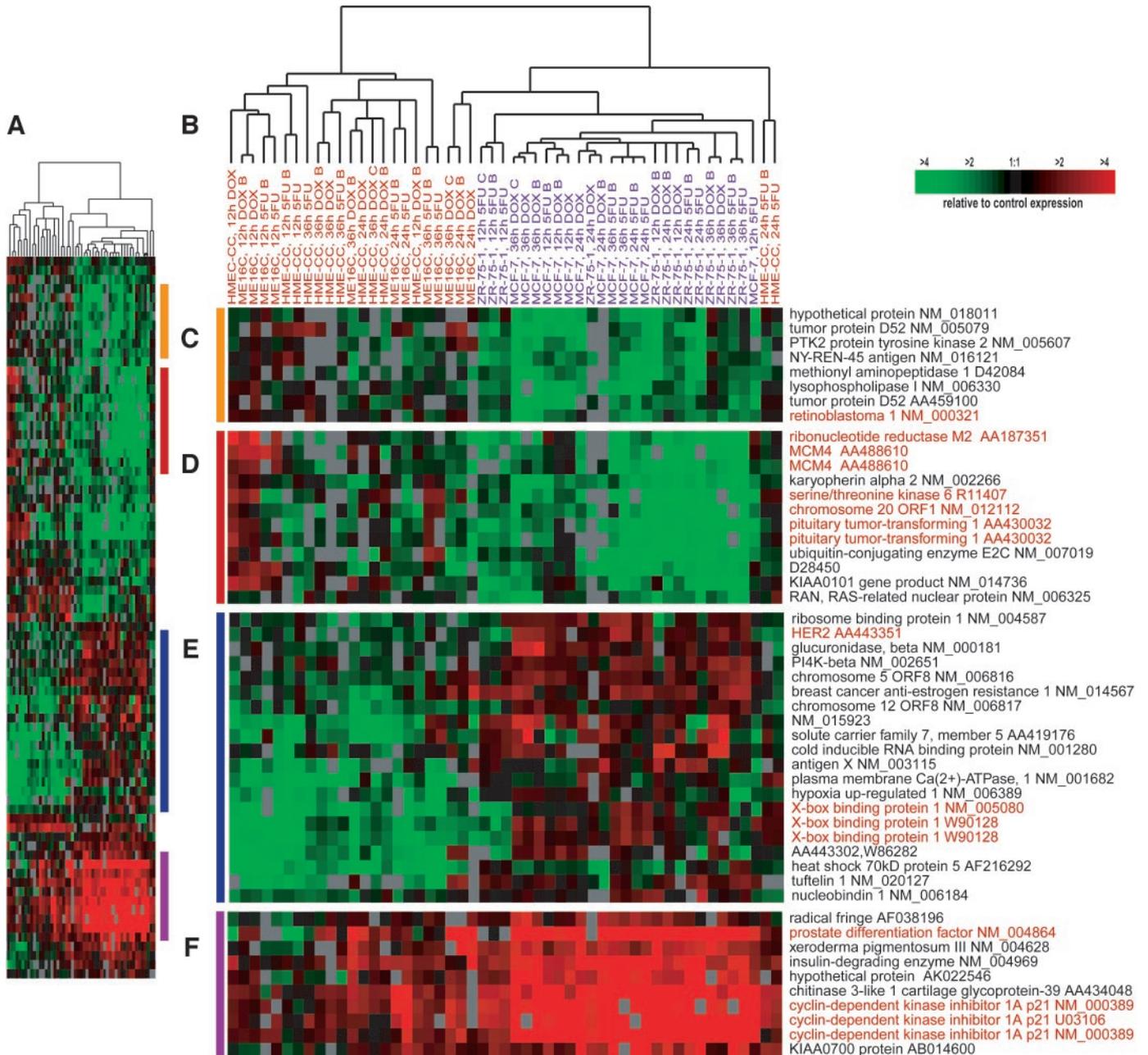


Fig. 3. Gene expression pattern of top 100 genes that distinguished between basal and luminal chemotherapeutic-treated cell lines. Cluster analysis was conducted on the 51 chemotherapeutic-treated MCF-7, ZR-75-1, HME-CC, and ME16C experiments. A, the scaled down cluster diagram (complete cluster diagram is available in Supplemental Fig. 3). Colored bars illustrate the location of clusters C, D, E, and F. B, the 51 experiments were divided into two dendrogram branches based on gene expression. Blue, the luminal cell lines; red, the basal cell lines. Clusters of genes are shown whose expression were more drastically down-regulated (C and D) or up-regulated (E and F) in luminal cell lines compared with basal cell lines. Highlighted in red, genes discussed in the text.

luminal cell lines, with expression of  $p21^{waf1}$  and *prostate differentiation factor* being highly up-regulated in luminal cells and less dramatically induced in basal cell lines. We also confirmed the cell-type-specific differences in basal *versus* luminal induction of  $p21^{waf1}$  on the protein level by Western blot (Fig. 4).

**In Vivo Responses to Chemotherapeutics.** We have previously profiled 115 breast tumors and have identified clinically distinct subtypes using patterns of gene expression (4, 5, 7). Tumor biopsies were sampled before chemotherapy, and for 46 of these tumors, tumor biopsies were also sampled after chemotherapy (16, 17). To allow comparisons with our *in vitro* work, we conducted a supervised analysis using SAM to identify gene expression differences between before- and after-chemotherapy samples. For the first analysis, we

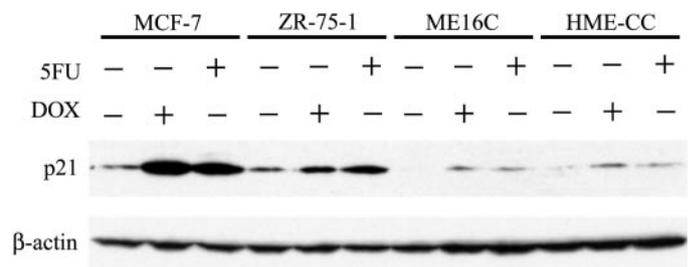


Fig. 4. Protein levels of  $p21^{waf1}$  in chemotherapeutic-treated basal and luminal cell lines. Cell lines were treated with an  $IC_{50}$  dose of doxorubicin (DOX) or 5-fluorouracil (5FU), and lysates were collected at 36 h.  $p21^{waf1}$  levels were induced in all chemotherapeutic-treated samples relative to sham samples.  $\beta$ -actin was assayed as a loading control.

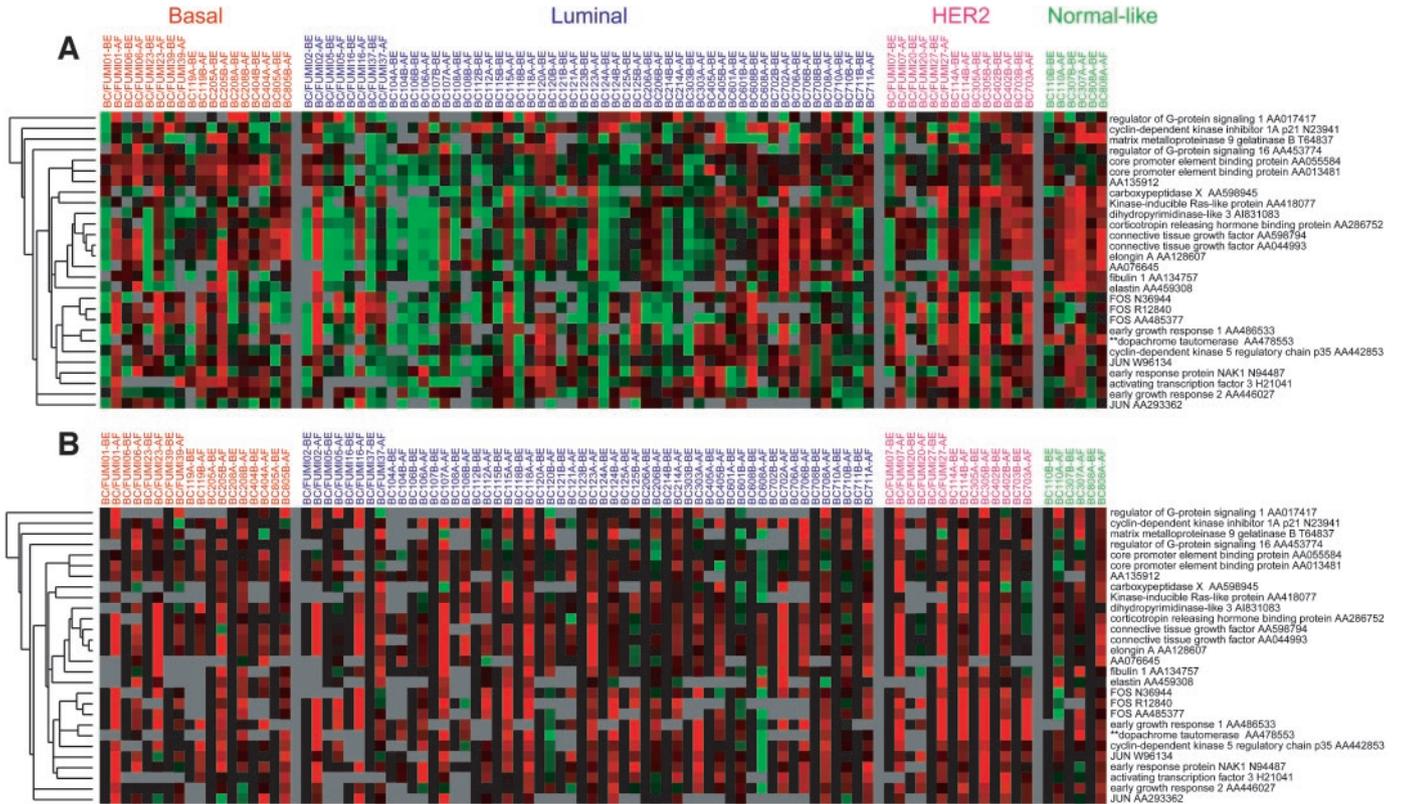


Fig. 5. Gene expression pattern for genes altered by chemotherapeutics in tumors. *A*, median-centered expression data using the gene set determined by SAM to be significantly changed in the before *versus* the after doxorubicin (DOX)- and 5-fluorouracil (5FU)-treated tumors. Tumor pairs from all of the tumor subtypes characterized in Sørlie *et al.* (5) are included; *tumor samples in red* are basal, *tumor samples in blue* are luminal (subtype A and B); *tumor samples in pink* are ERBB2/HER2 positive; *tumor samples in green* are normal-like. *B*, each after-sample in *A* was normalized to the expression value for its corresponding before-sample. A gray square was assigned to both samples if either the before- or the after-sample had missing data. *A* and *B* exclude all tumor samples for which a complete before and after set was unavailable.

disregarded tumor subtype differences and treatment differences (patients were treated with either DOX or 5FU/mitomycin) and looked for consistent differences between 81 before-samples and 50 after-samples. All of the breast tumor subtypes identified in Sørlie *et al.* (5) were represented. The list of genes that differed between the before- and after-samples is shown in a cluster diagram (Fig. 5A) in which the samples are arranged by tumor subtype as defined in Sørlie *et al.* (Ref. 5; not clustered); genes were clustered and all fold-changes are displayed relative to the median gene expression level. A total of 28 cDNA clones representing 23 genes were more highly expressed in the after-samples relative to the before-samples (no genes were significantly lower in the after-samples). These findings agree (at least 13 genes in common) with a similar analysis performed on a subset of these data (30). Among these 23 genes were the AP-1 coactivators *FOS* and *JUN*, *p21<sup>waf1</sup>*, and a number of other genes involved in wound healing including *connective tissue growth factor* and *matrix metalloproteinase 9*.

In Fig. 5B, we normalized the expression ratio in each after-sample to its paired before-sample (displayed in black) on a gene-by-gene basis. This allowed visualization of the changes caused by chemotherapy in each patient. These changes are difficult to discern in Fig. 5A because of the diversity of initial expression values in the before-samples. Relative to the paired before-samples, nearly all of the after-samples in all four of the tumor subtypes showed induced expression of these genes, despite their diverse expression ranges in Fig. 5A. A few of the tumors had anomalous behavior, underscoring the individuality of tumor responses even within a subtype.

Based on our cell line data, we hypothesized that there might also be tumor subtype-specific responses; therefore, we conducted analyses on the before-samples *versus* the after-samples for the basal and

luminal subtypes separately. Using 81 luminal tumor samples and SAM analysis, we identified 14 genes that were changed in expression after treatment. Using 21 basal tumor samples, we identified nine genes that were induced after treatment (Table 3). In this analysis, there was a five-gene overlap between the basal and luminal gene lists. A number of genes that were seen in the combined analysis were significantly altered in only one of the subtypes. For example, *p21<sup>waf1</sup>* was present only on the luminal list and *core promoter element-binding protein (COPEB)* was present only on the basal list. Next, we

Table 3 Genes altered by chemotherapy in luminal and basal breast tumor subtypes

Luminal tumors	Basal tumors
<i>Connective tissue growth factor</i> AA598794	<i>Connective tissue growth factor</i> AA598794
<i>Connective tissue growth factor</i> AA044993	<i>Connective tissue growth factor</i> AA044993
<i>Early growth response 1</i> AA486533	<i>Early growth response 1</i> AA486533
<i>Early response protein NAK1</i> N94487	<i>Early response protein NAK1</i> N94487
<i>Elongin A</i> AA128607 <sup>a</sup>	<i>Elongin A</i> AA128607
<i>FOS</i> N36944	<i>FOS</i> R12840
<i>Corticotropin releasing hormone binding protein</i> AA286752	<i>Core promoter element-binding protein</i> AA013481 <sup>a</sup>
<i>Cyclin-dependent kinase 5, regulatory subunit 1 (p35)</i> AA442853	<i>Dermatan sulfate proteoglycan 3</i> AA131238
<i>Cyclin-dependent kinase inhibitor 1A, p21<sup>waf1</sup></i> N23941 <sup>a</sup>	<i>Homo sapiens mRNA</i> AA135912
<i>Dihydropyrimidinase-like 3</i> AI831083	<i>RAB21</i> AA076645
<i>Dopachrome tautomerase</i> AA478553 <sup>b</sup>	
<i>Kinase-inducible Ras-like protein</i> AA418077	
<i>Prostate differentiation factor</i> N26311 <sup>a,b</sup>	
<i>Spondin 1</i> H09099	
<i>Thrombospondin 1</i> AA464532 <sup>a</sup>	

<sup>a</sup> This gene was also altered by treatment in the corresponding cell line experiments.

<sup>b</sup> A potentially chimeric cDNA clone that maps to two different Unigene entries.

compared each cell type's *in vivo* list with its corresponding *in vitro* list and identified four genes that were altered in luminal tumors and luminal cell lines (*p21<sup>waf1</sup>*, *elongin A*, *prostate differentiation factor*, and *thrombospondin 1*). *COPEB* was the only gene that was significantly altered in both the basal tumors and cell lines.

Finally, to identify additional similarities between the cell line and tumor data sets, we used the SAM-generated luminal and basal cell line-derived gene lists in a clustering analysis of the tumor samples. Basal and luminal gene expression signatures identified in the cell lines also appeared differentially expressed in tumor subtypes (Supplemental Figs. 4 and 5). For example, when the basal cell line list was used to cluster all of the tumor samples, *p21<sup>waf1</sup>* and *MDM2* clustered together and showed higher expression in the luminal tumors, whereas *COPEB* and *GST- $\pi$*  showed higher expression in the basal tumors (Supplemental Fig. 4). When the luminal cell line list was used to cluster all of the tumor samples, subtype-specific responses were also evident; for example, the basal tumors showed high expression of the proliferation signature both before and after chemotherapy. This is consistent with the *in vitro* findings because the proliferation signature was unchanged in the basal cell lines after chemotherapy treatment.

## DISCUSSION

The mammary gland contains a heterogeneous population of epithelial cells in different stages of differentiation. Basal and luminal epithelium represent two cell populations that are thought to arise from a common progenitor, but they each express unique markers and perform unique functions (31, 32). Luminal epithelia are widely believed to give rise to the majority of breast cancers, but there is evidence that up to 15% of breast cancers show some characteristics of basal epithelium (4, 5). MCF-7 cells and HME cell lines have been extensively studied as models of breast cancer; however, they represent different types of breast cancers. MCF-7 and ZR-75-1 cells (data not shown) have expression similarities with ER $\alpha$ -positive breast tumors, whereas HME lines (finite life span or immortalized) have expression similarities with basal breast tumors (33).

In our model of breast cancer, basal and luminal epithelial cells have unique transcriptional responses to the chemotherapeutics DOX and 5FU. The two luminal cell lines showed similar response patterns to one another including the strong induction of DNA damage/stress response genes, notably *p21<sup>waf1</sup>* (Figs. 1D and 3F). The basal cell lines showed a much less dramatic induction of *p21<sup>waf1</sup>* (Fig. 2D and 3F). All four of our cell lines are wild type for *p53* by sequence analysis and express p53 protein (data not shown); therefore, the differences in *p21<sup>waf1</sup>* expression cannot be attributed to differences in *p53* status. *p21<sup>waf1</sup>* is involved in the G<sub>1</sub> checkpoint response, and others have reported an impaired G<sub>1</sub> checkpoint in HME cell lines (34). Consistent with a strong cell cycle checkpoint in the luminal cell lines, MCF-7 and ZR-75-1 cells also repressed a large set of proliferation genes (Fig. 1C). This suggests that their G<sub>1</sub> checkpoint in response to DNA damage is intact (20). The two basal cell lines did not repress the proliferation signature, but they did down-regulate genes involved in differentiation (Fig. 2C).

The basal cell lines that we used for this study were hTERT-immortalized HME cells, whereas the luminal cell lines were derived from human tumors. Telomerase expression is a hallmark of breast cancer (35), but increased telomerase expression is one of many changes that are observed as cells progress toward a malignant state (36). Although breast tumor derived cell lines of luminal origin are widely studied, analogous lines of basal origin have not yet been identified. We acknowledge that comparison of breast cancer lines *versus* immortalized breast lines represents a starting point for investigations of these cell types. Future comparisons using additional cell

lines, and preferably cancer cell lines of basal origin, may yield more data of greater significance. We note that some of the expression differences observed between the basal and luminal cell lines could be due to differences in tumorigenicity. However, we found that our cell lines recapitulated some of the cell type differences seen *in vivo* in response to these same agents (Table 3; Supplemental Figs. 4 and 5). The overlap observed between the tumors and cell lines is significant, especially considering three differences between these data sets: (a) the tumor data were acquired using cDNA microarrays and a common reference sample whereas the cell lines were assayed using 60mer oligonucleotide arrays and a cell-line specific reference (untreated pooled reference); (b) the cell lines were all *p53* wild type, whereas ~40% of the tumors were *p53* mutant. Thus, the *in vivo* analysis is more likely to have excluded some p53-dependent responses to chemotherapy; and (c) the tumors represent a heterogeneous cell population and the cell lines represent only a single cell type.

A strength of tumor profiling studies is that they capture the heterogeneity of tumors in their natural environment. However, this heterogeneity makes it difficult to study the chemotherapy responses of specific cell types. The role of each cell type in a tumor can begin to be dissected using cell-line models, preferably with multiple cell lines representing each cell type. Cell lines are as unique as the tumors from which they were derived, but common response patterns can only become identifiable when looking at multiple cell lines in concert. This was illustrated in a recent study of 60 cell lines and 60,000 compounds (33, 37) in which relationships between sets of cell lines, sets of genes, and toxicant sensitivity were identified. In the work presented here, we used four cell lines with two cell lines representing each of two tumor subtypes. Characterizing common responses and interindividual variation in these cell lines will help to identify those responses that are stereotypical for each cell type.

Recent studies have demonstrated that DNA-damaging agents induce generic stress responses. In 2000, Gasch *et al.* (11) showed that yeast displayed a stereotypic pattern of gene expression when exposed to a wide range of stresses including heat shock, growth factor deprivation, and treatment with hydrogen peroxide. These authors termed the stereotypic response the "environmental stress response (ESR)." The environmental stress response included repression of growth-related genes and genes encoding ribosomal proteins and induction of genes involved in DNA damage response and metabolism. These results are in agreement with our finding that a major response to treatment included repression of genes involved in cell growth and induction of DNA damage response genes. Our work with breast cell lines corroborates other recent human cell line studies that have demonstrated common stress responses after DNA-damaging treatments (12, 26, 38–40). In this article, we have demonstrated that some of the changes seen *in vitro* were also observed *in vivo*.

Finally, we note that DOX and 5FU have distinct mechanisms of action (41, 42). For example, DOX is thought to target topoisomerase IIA blocking the G<sub>2</sub>-M transition and 5FU targets thymidylate synthase blocking S-phase progress. In our experiments with luminal cell lines, both drugs affected gene expression in all phases of the cell cycle (Fig. 1C). These cell cycle genes serve as proliferation markers and are not specific to a single mode of action. The specific mechanisms of action of DOX and 5FU may be evident in a subset of genes expressed in our experiments and subsequent analyses will attempt to identify this gene set. However, to fully validate toxicant-specific gene sets, it must also be demonstrated that the gene set predicts mode of action for independent data sets on mechanistically similar drugs. Our primary objective for this work was to understand how cell types differed in their stress response patterns, which are the dominant gene expression responses to DNA damage. The identification of cell-type

specific stress responses *in vitro* and *in vivo* has implications for understanding the biological response to therapy.

## ACKNOWLEDGMENTS

We thank Elizabeth Livanos of the University of North Carolina at Chapel Hill Chromosome Imaging Core Facility for karyotyping the HME cell lines.

## REFERENCES

1. Thor AD, Berry DA, Budman DR, et al. erbB-2, p53, and efficacy of adjuvant therapy in lymph node-positive breast cancer. *J Natl Cancer Inst (Bethesda)* 1998;90:1346–60.
2. Aas T, Borresen AL, Geisler S, et al. Specific P53 mutations are associated with de novo resistance to doxorubicin in breast cancer patients. *Nat Med* 1996;2:811–4.
3. Jarvinen TA, Tanner M, Rantanen V, et al. Amplification and deletion of topoisomerase IIalpha associate with ErbB-2 amplification and affect sensitivity to topoisomerase II inhibitor doxorubicin in breast cancer. *Am J Pathol* 2000;156:839–47.
4. Perou CM, Sørli T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature (Lond)* 2000;406:747–52.
5. Sørli T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001;98:10869–74.
6. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature (Lond)* 2002;415:530–6.
7. Sørli T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 2003;100:8418–23.
8. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci USA* 2003;100:10393–8.
9. Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 2000;24:227–35.
10. Hughes TR, Marton MJ, Jones AR, et al. Functional discovery via a compendium of expression profiles. *Cell* 2000;102:109–26.
11. Gasch AP, Spellman PT, Kao CM, et al. Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* 2000;11:4241–57.
12. Heinloth AN, Shackelford RE, Innes CL, et al. Identification of distinct and common gene expression changes after oxidative stress and gamma and ultraviolet radiation. *Mol Carcinog* 2003;37:65–82.
13. Wang HC, Federoff S. Banding in human chromosomes treated with trypsin. *Nat New Biol* 1972;235:52–4.
14. Van Ewijk PH, Hoekstra JA. Calculation of the EC50 and its confidence interval when subtoxic stimulus is present. *Ecotoxicol Environ Saf* 1993;25:25–32.
15. Tusher V, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001;98:5116–21.
16. Geisler S, Borresen-Dale AL, Johnsen H, et al. TP53 gene mutations predict the response to neoadjuvant treatment with 5-fluorouracil and mitomycin in locally advanced breast cancer. *Clin Cancer Res* 2003;9:558–8.
17. Geisler S, Lonning PE, Aas T, et al. Influence of TP53 gene alterations and c-erbB-2 expression on the response to treatment with doxorubicin in locally advanced breast cancer. *Cancer Res* 2001;61:2505–12.
18. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998;95:14863–8.
19. Eisen MB, Brown PO. DNA arrays for analysis of gene expression. *Methods Enzymol* 1999;303:179–205.
20. Whitfield ML, Sherlock G, Saldanha AJ, et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell* 2002;13:1977–2000.
21. Perou CM, Jeffrey SS, van de Rijn M, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci USA* 1999;96:9212–7.
22. Bernal JA, Luna R, Espina A, et al. Human securin interacts with p53 and modulates p53-mediated transcriptional activity and apoptosis. *Nat Genet* 2002;32:306–11.
23. Kimura M, Kotani S, Hattori T, et al. Cell cycle-dependent expression and spindle pole localization of a novel human protein kinase, Aik, related to Aurora of *Drosophila* and yeast Ipl1. *J Biol Chem* 1997;272:13766–71.
24. Kufer TA, Sillje HH, Korner R, Gruss OJ, Meraldi P, Nigg EA. Human TPX2 is required for targeting Aurora-A kinase to the spindle. *J Cell Biol* 2002;158:617–23.
25. Liu G, Chen X. The ferredoxin reductase gene is regulated by the p53 family and sensitizes cells to oxidative stress-induced apoptosis. *Oncogene* 2002;21:7195–204.
26. Park WY, Hwang CI, Im CN, et al. Identification of radiation-specific responses from gene expression profile. *Oncogene* 2002;21:8521–8.
27. Sikder HA, Devlin MK, Dunlap S, Ryu B, Alani RM. Id proteins in cell growth and tumorigenesis. *Cancer Cell* 2003;3:525–30.
28. Ma Y, Hendershot LM. The unfolding tale of the unfolded protein response. *Cell* 2001;107:827–830.
29. Pegram MD, Finn RS, Arzoo K, Beryt M, Pietras RJ, Slamon DJ. The effect of HER-2/neu overexpression on chemotherapeutic drug sensitivity in human breast and ovarian cancer cells. *Oncogene* 1997;15:537–47.
30. Korn EL, McShane LM, Troendle JF, Rosenwald A, Simon R. Identifying prepost chemotherapy differences in gene expression in breast tumours: a statistical method appropriate for this aim. *Br J Cancer* 2002;86:1093–6.
31. Smith GH, Chepko G. Mammary epithelial stem cells. *Microsc Res Tech* 2001;52:190–203.
32. Pechoux C, Gudjonsson T, Ronnov-Jessen L, Bissell MJ, Petersen OW. Human mammary luminal epithelial cells contain progenitors to myoepithelial cells. *Dev Biol* 1999;206:88–99.
33. Ross DT, Perou CM. A comparison of gene expression signatures from breast tumors and breast tissue derived cell lines. *Dis Markers* 2001;17:99–109.
34. Meyer KM, Hess SM, Tlsty TD, Leadon SA. Human mammary epithelial cells exhibit a differential p53-mediated response following exposure to ionizing radiation or UV light. *Oncogene* 1999;18:5795–805.
35. Herbert BS, Wright WE, Shay JW. Telomerase and breast cancer. *Breast Cancer Res* 2001;3:146–9.
36. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57–70.
37. Scherf U, Ross DT, Waltham M, et al. A gene expression database for the molecular pharmacology of cancer [see comments]. *Nat Genet* 2000;24:236–44.
38. Weigel AL, Handa JT, Hjelmeland LM. Microarray analysis of H2O2-, HNE-, or tBH-treated ARPE-19 cells. *Free Radic Biol Med* 2002;33:1419–32.
39. Morgan KT, Ni H, Brown HR, et al. Application of cDNA microarray technology to *in vitro* toxicology and the selection of genes for a real-time RT-PCR-based screen for oxidative stress in Hep-G2 cells. *Toxicol Pathol* 2002;30:435–51.
40. Sesto A, Navarro M, Burslem F, Jorcano JL. Analysis of the ultraviolet B response in primary human keratinocytes using oligonucleotide microarrays. *Proc Natl Acad Sci USA* 2002;99:2965–70.
41. Gewirtz DA. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem Pharmacol* 1999;57:727–41.
42. Longley DB, Harkin DP, Johnston PG. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* 2003;3:330–8.