Radiation-Induced Gene Signature Predicts Pathologic Complete Response to Neoadjuvant Chemotherapy in Breast Cancer Patients

Authors: Oh, Daniel S., Cheang, Maggie C. U., Fan, Cheng, and Perou, Charles M.

Radiation-Induced Gene Signature Predicts Pathologic Complete Response to Neoadjuvant Chemotherapy in Breast Cancer Patients

Daniel S. Oh,[a,1] Maggie C. U. Cheang,[b] Cheng Fan[b] and Charles M. Perou[b,c,d]

[a] Department of Radiation Oncology, Leo W. Jenkins Cancer Center, The Brody School of Medicine at East Carolina University, Greenville, North Carolina 27834; [b] Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina 27599; [c] Department of Genetics, University of North Carolina, Chapel Hill, North Carolina 27599; and [d] Department of Pathology and Laboratory Medicine, University of North Carolina, Chapel Hill, North Carolina 27599

Oh, D. S., Cheang, M. C. U., Fan, C. and Perou, C. M. Radiation-Induced Gene Signature Predicts Pathologic Complete Response to Neoadjuvant Chemotherapy in Breast Cancer Patients. *Radiat. Res.* **181**, 193–207 (2014).

The identification of biomarkers predictive of neoadjuvant chemotherapy response in breast cancer patients would be an important advancement in personalized cancer therapy. In this study, we hypothesized that due to similarities between radiation- and chemotherapy-induced cellular response mechanisms, radiation-responsive genes may be useful in predicting response to neoadjuvant chemotherapy. Murine p53 null breast cancer cell lines representative of the luminal, basal-like and claudin-low human breast cancer subtypes were irradiated to identify radiation-responsive genes across subtypes. These murine tumor radiation-induced genes were then converted to their human orthologs, and subsequently tested as a predictor of pathologic complete response (pCR), which was validated on two independent published neoadjuvant chemotherapy datasets of genomic data with chemotherapy response. A radiation-induced gene signature consisting of 30 genes was identified on a training set of 337 human primary breast cancer tumor samples that was prognostic for survival. Mean expression of this signature was calculated for individual samples on two independent published datasets and was found to be significantly predictive of pCR. Multivariate logistic regression analysis in both independent datasets showed that this 30 gene signature added significant predictive information independent of that provided by standard clinical predictors and other gene expression-based predictors of pCR. This study provides new information for radiation-induced biology, as well as information regarding response to neoadjuvant chemotherapy and a possible means of improving the prediction of pCR. © 2014 by Radiation Research Society

INTRODUCTION

Neoadjuvant chemotherapy has been widely used in recent years as part of the standard of care for locally advanced breast cancer patients. Although neoadjuvant and adjuvant chemotherapy have similar efficacy in terms of disease-free and overall survival rates (*1*), neoadjuvant chemotherapy has been shown to improve breast-conserving operability in locally advanced breast cancers (*2, 3*). Another advantage of neoadjuvant chemotherapy is that it allows for the direct and timely observation of tumor treatment response. Twenty to thirty percent of breast cancer patients who receive neoadjuvant chemotherapy achieve a pathologic complete response (pCR) (*4*), which has been correlated with improved long-term, disease-free and overall survival (*1–3, 5–9*) and thus is a valuable surrogate end point for survival.

The prediction of neoadjuvant chemotherapy response (i.e., pCR) is an active area of research. Since patients who achieve pCR after neoadjuvant chemotherapy are more likely to experience excellent cancer-free, long-term survival (*1–3, 5–9*), the accurate prediction of pCR would be of significant value. It would serve to identify those patients who could benefit most from neoadjuvant chemotherapy and identify those unlikely to benefit from therapy and thus be spared treatment-associated toxicities and be selected to receive alternative therapeutic approaches. Rouzier *et al.* (*10*) published in 2005, a nomogram consisting of clinical variables including tumor size, histologic grade and estrogen receptor (ER) status to predict pCR to neoadjuvant FAC (fluorouracil, doxorubicin and cyclophosphamide) or T/FAC (paclitaxel, FAC) chemotherapy. This same group reported in another study (*11*) that intrinsic breast cancer subtype as determined by gene expression profiling of tumors prior to treatment, was associated with pCR. In this study, basal-like and HER2-enriched (HER2-E) intrinsic subtypes were associated with the highest rates of pCR (45%) while luminal A and luminal B tumors showed the lowest rate of pCR (6%). Building on these studies, we sought to determine whether we could identify a novel biomarker that could add predictive

information independent of that provided by known clinical variables and other gene expression-based predictors of pCR.

Cellular responses to radiation and chemotherapy involve multiple overlapping pathways. Both therapies involve cytotoxic agents that can ultimately induce apoptosis in target cells (*12, 13*). Radiation therapy and many chemotherapeutics result in DNA damage, activating multiple overlapping proteins/pathways, such as the ATM and ATR protein kinases, TP53 and TP53-independent pathways including CHK2-mediated signaling (*14*). Other proteins/pathways induced by both radiation and chemotherapy includes the extracellular signal-regulated kinase (ERK) pathway, Jun-N-terminal kinase (JNK) pathway, p38 MAPK family, NK-kB, AKT, mTOR and the checkpoint proteins (*14–18*). We hypothesized that due to the similarities in the cellular response mechanisms to radiation and chemotherapy, radiation-responsive genes may be useful in predicting response to neoadjuvant chemotherapy for breast cancer.

MATERIALS AND METHODS

Cell Culture

To identify radiation-responsive genes, we used cell lines derived from the T11, 2225L and 2250L murine mammary tumors described by Herschkowitz *et al.* (*19*). Briefly, these tumors were produced by removal and transplantation of 6-week-old BALB/c p53$^{-/-}$ mammary tissues into 3-week-old wild-type BALB/c recipients. This was necessary to circumvent the appearance of other tumor types that occurred with short latency in mice homozygous for p53 loss. As described by Herschkowitz *et al.* (*19*), based on murine-intrinsic gene list expression analysis, the T11 tumor had characteristics of the claudin-low human breast carcinoma subtype; the 2225L tumor had characteristics of the basal-like human breast carcinoma subtype; the 2250L tumor had characteristics of the luminal human breast carcinoma subtype. Respective tumor samples stored in liquid nitrogen were thawed and placed into appropriate media to allow cell line formation as follows: for the T11 cell line, the media used was RPMI 1640 containing 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin; for the 2225L and 2250L cell lines, media used were HMEC media containing 5% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin and HMEC supplement (Gibco® Life Technologies, Grand Island, NY). Cells were incubated in a humidified incubator at 37°C in 95% air/5% CO$_2$. Genomic DNA was harvested from cell lines using QIAGEN DNeasy Kit (QIAGEN, Valencia, CA), and presence of the p53 null transgene was verified using polymerase chain reaction.

Cell Irradiation and Collection of RNA

Cells were plated in 150 mm dishes and grown until 50% confluence. Cells were then irradiated to a dose of 8 Gy using an RS 2000 irradiator (Rad Source Technologies Inc., Suwanee, GA) operating at a dose rate of 100 cGy/min. Cells were then immediately returned to the incubator and harvested at 4, 8, 12, 24 and 48 h after irradiation, at which point RNA was isolated using the QIAGEN RNeasy Mini Kit. A control nonirradiated RNA sample for each cell line was collected from cells harvested immediately after mock irradiation.

Microarray Experiments

Mouse whole-genome 4×180,000 features microarrays (Agilent Technologies Inc., Palo Alto, CA) were hybridized according to manufacturer's protocol with Cy3-CTP-labeled cRNA from mock-irradiated cells (2 ug/sample) and Cy5-CTP-labeled cRNA from irradiated cells (2 ug/sample), with replicates for a 24 h time point for each cell line. Microarrays were scanned and image files analyzed as described previously (*19*). All primary microarray data are available from the University of North Carolina (Chapel Hill, NC) Microarray Database (https://genome.unc.edu) and the Gene Expression Omnibus [GEO; National Center for Biotechnology Information (NCBI)] (http://www.ncbi.nlm.nih.gov/geo/) with series number GSE48073.

Analysis of Microarray Data to Identify Radiation-Responsive Genes

Data from microarray experiments were calculated as described (*19*), where we used the Lowess normalized log2R/G ratio. To identify radiation-responsive genes, we used a one-class Significance Analysis of Microarrays (SAM) to identify genes that changed in *all* time points for *all* three cell lines (as a single class) relative to the mock-irradiated cells (*20*). Using a false discovery rate (FDR) of 0%, SAM identified 4,278 radiation-induced genes and 2,689 radiation-repressed genes. Hierarchical cluster analysis was conducted using Cluster 3.0 (*21*) and results were visualized in Java TreeView (*22*).

To identify genes that were differentially expressed after irradiation across the cell lines, we used a multi-class SAM analysis to identify genes that changed in all time points after irradiation, but that were different between the cell lines (each cell line as its own class) relative to mock-irradiated cells. Using an FDR of 0%, SAM identified 1,289 genes differentially expressed after irradiation across the cell lines. Hierarchical cluster analysis was conducted using Cluster 3.0 (*21*) and results were visualized in Java TreeView (*22*).

Analysis of Human Primary Breast Tumor Microarray Data Using the Radiation-Induced Gene Set

The human primary breast tumor samples used as the training dataset are described in Prat *et al.* (*23*); this training dataset comprised a total of 337 tumor samples represented by microarray experiments from breast cancer patients (consisting of 320 breast tumor samples and 17 normal breast samples) heterogeneously treated in accordance with standard of care. To analyze this human training dataset with the radiation-induced murine gene set identified by the above described SAM, the list of 4,278 radiation-induced murine genes identified from the cell lines were converted to their human gene orthologs using the UCSC genome website (http://genome.ucsc.edu/). A total of 1,964 unique human genes present on our human microarrays were identified. These genes were then used to hierarchically cluster the Prat *et al.* 337 human breast sample dataset using Cluster 3.0 to identify gene sets/clusters composed of genes that are highly coordinately expressed across these patient samples. We hypothesized that the mean expression of one or more of these gene sets/clusters may be useful in predicting an individual tumor's response to neoadjuvant chemotherapy. To avoid spurious results that may occur with small gene clusters or poorly correlated gene clusters, we limited the analysis to the most correlated gene clusters (node correlation >=0.45) and that were greater than 20 genes in size. This identified 23 non-overlapping gene clusters, where we next calculated a mean gene expression value of each gene cluster, for each of the 337 samples from the Prat *et al.* dataset (*23*).

By matching Entrez gene identifiers (www.ncbi.nlm.nih.gov/gene), microarray data for as many genes as possible for each of the 23 gene clusters identified above was obtained on two independent test microarray datasets (*24, 25*). Briefly, the Hatzis *et al.* (*24*) dataset consisted of microarray data from 473 HER2-negative breast cancer tumors taken prior to sequential taxane and anthracycline-based neoadjuvant chemotherapy. This dataset had associated pCR and distant relapse free survival (DRFS) data. The Gluck *et al.* (*25*) dataset consisted of microarray data from 95 HER2-negative breast cancer tumors taken prior to neoadjuvant capecitabine and docetaxel

chemotherapy; this dataset had associated pCR data and TP53 mutation status data using the AmpliChip TP53 assay. As done for the Prat *et al.* dataset, a mean gene expression value of each gene cluster was calculated for each patient in the Hatzis *et al.* dataset, and we validated the most significant gene cluster on the Gluck *et al.* dataset.

Statistical and Survival Analysis

Kaplan-Meier survival plots were compared using the Cox-Mantel log-rank test in WinSTAT® (R. Fitch Software, Staufen, Germany) for Excel (Microsoft Corp., Redmond, WA). Univariable and multivariable logistic regression analysis were used to assess the significance of gene cluster mean values to pCR. All statistical tests were two-tailed and $P < 0.05$ were declared significant. Analyses were performed using JMP Pro 9.0 (SAS Institute, Cary, NC) and R version 2.13.1 (R Development Core Team, Vienna, Austria).

RESULTS

Identification of Radiation-Responsive Genes

To identify radiation-responsive genes, we used cell lines derived from three different mouse p53-deficient tumors, including the T11 (claudin-low), 2225L (basal-like) and 2250L (luminal) tumors described by Herschkowitz *et al.* (*19*). A one-class SAM with a false discovery rate (FDR) of 0% identified 4,278 radiation-induced and 2,689 radiation-repressed genes in microarray experiments on the combined T11, 2225L and 2250L cell lines at 4, 8, 12, 24 and 48 h after a single radiation dose of 8 Gy. Hierarchical clustering of the radiation-induced genes is shown in Fig. 1 and that of the radiation-repressed genes is shown in Fig. 2.

Many genes identified as radiation induced in our analysis were previously known to be radiation responsive in humans including CREM, BNIP3, FAS, TNFRSF11B, IFITM1, LGALS3PB, COX7B and SESN1 (*26–29*), indicating the conservation of the radiation responsiveness of many genes in this p53-deficient mouse model. Using the program DAVID [Database for Annotation, Visualization and Integrated Discovery (*30, 31*); http://david.abcc.ncifcrf.gov], the gene ontology categories that were significantly over-represented relative to chance in the set of all radiation-induced genes included: "protein catabolic process" ($P = 3.40 \times 10^{-9}$); "apoptosis" ($P = 3.45 \times 10^{-5}$); "cell death" ($P = 1.17 \times 10^{-4}$) and "macromolecule catabolic process" ($P = 8.24 \times 10^{-8}$). Kyoto Encyclopedia of Genes and Genomes [KEGG (*32*); http://www.genome.jp/kegg/] pathways (which are graphical diagrams representing knowledge on molecular interaction and reaction networks for a wide variety of cellular processes) that were significantly over-represented relative to chance included: "pyruvate metabolism" ($P = 3.65 \times 10^{-5}$); "ubiquitin mediated proteolysis" ($P = 0.003$); "VEGF signaling pathway" ($P = 0.003$), "mTOR signaling pathway" ($P = 0.004$); and interestingly, "p53 signaling pathway" ($P = 0.02$). Since the cell lines used in our experiments were p53 null, the enrichment for members of the KEGG pathway "p53 signaling pathway" in the radiation-induced gene set suggested that these members

known to be activated by p53 [such as Scotin (*33*) and the sestrins SESN1 and SESN2 (*34, 35*)] can also be activated independent of p53; Scotin has been confirmed independently by Terrinoni *et al.* to be able to be activated independent of p53 (*36*).

Subclusters of radiation-induced genes, as opposed to the complete list, showed enrichment of many distinct ontologies in the various gene clusters. For example, cluster A in Fig. 1, which consisted of genes whose expression was induced at a relatively constant level with respect to time across all cell lines, was significantly enriched for members of the gene ontology categories "mitochondrion" ($P = 6.24 \times 10^{-9}$) and "cellular protein catabolic process" ($P = 0.02$). In addition, cluster C in Fig. 1, which consisted of genes whose expression increased with time across all cell lines, was enriched for members of gene ontology categories "positive regulation of programmed cell death" ($P = 0.002$), "antigen processing and presentation" ($P = 6.40 \times 10^{-7}$), "lysosome" ($P = 3.71 \times 10^{-5}$) and "immune response" ($P = 2.44 \times 10^{-4}$). Other clusters of radiation induced genes and the respective gene ontologies enriched in those clusters are shown in Fig. 1.

Gene ontology categories significantly over-represented relative to chance in the set of all radiation-repressed genes included "spliceosome" ($P = 3.19 \times 10^{-14}$), "cell cycle" ($P = 2.03 \times 10^{-11}$), "cell division" ($P = 3.33 \times 10^{-7}$) and "mitosis" ($P = 3.98 \times 10^{-5}$). KEGG pathways significantly over-represented relative to chance in the set of radiation-repressed genes included "ribosome" ($P = 5.45 \times 10^{-22}$), "cell cycle" ($P = 5.24 \times 10^{-4}$) and "spliceosome" ($P = 3.78 \times 10^{-19}$). The enrichment of genes involved in apoptosis in the radiation-induced gene set, and enrichment of genes involved in mitosis in the radiation-repressed gene set, is consistent with prior studies examining radiation-responsive genes in humans (*26–29*) indicating conservation of the radiation responsiveness of these gene ontology categories/pathways in our mouse model. As for the radiation-induced genes, clusters of radiation-repressed genes showing highly correlated expression were observed (Fig. 2) and gene ontology analysis of these gene clusters showed enrichment. For example, cluster A in Fig. 2, which consisted of genes whose expression was more repressed in the 2250L cell line than in the other cell lines, was enriched for the gene ontology category "Zinc ion binding" ($P = 0.01$), while cluster G in Fig. 2 that consisted of genes whose repression increased with time across all cell lines, was enriched for the gene ontology categories "cell cycle" ($P = 2.93 \times 10^{-5}$), "RNA processing" ($P = 1.20 \times 10^{-21}$), "cell division" ($P = 0.003$), "mitosis" ($P = 0.02$) and "nuclear division" ($P = 0.02$); the repression of these gene clusters likely reflect reduced cell proliferation, which is a known cellular response to radiation. Other clusters of radiation-repressed genes and the respective gene ontologies enriched in those clusters are shown in Fig. 2.
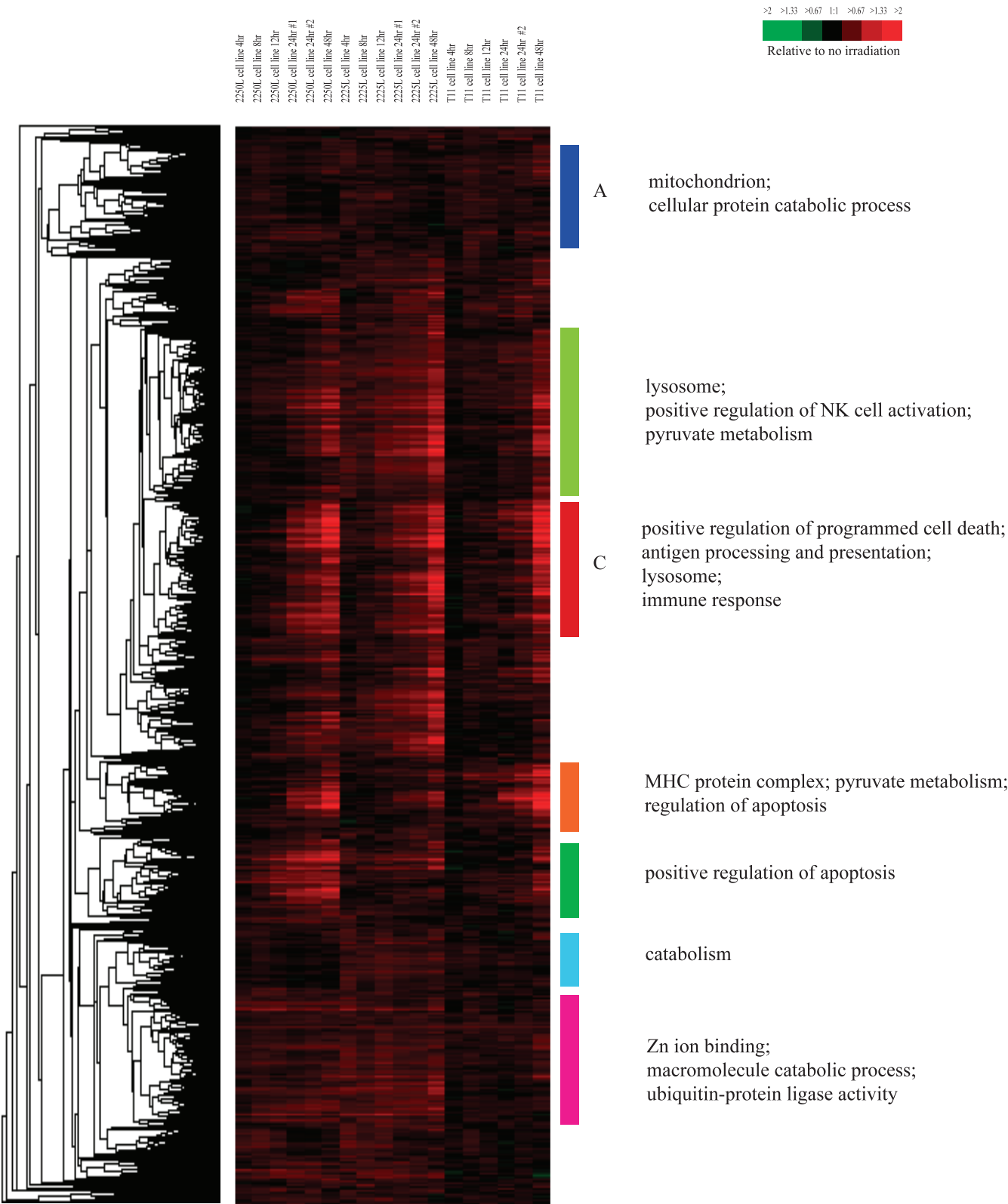
FIG. 1. Hierarchical clustering analysis of genes determined by one-class significance analysis of microarrays (SAM) to be radiation-induced (4,278 genes) when testing three different mouse mammary tumor cell lines. Colored bars and/or letters indicate various gene clusters as discussed in this article. To the left of the diagram are the gene ontology categories significantly enriched in the corresponding gene clusters indicated by the colored bars and/or letters.
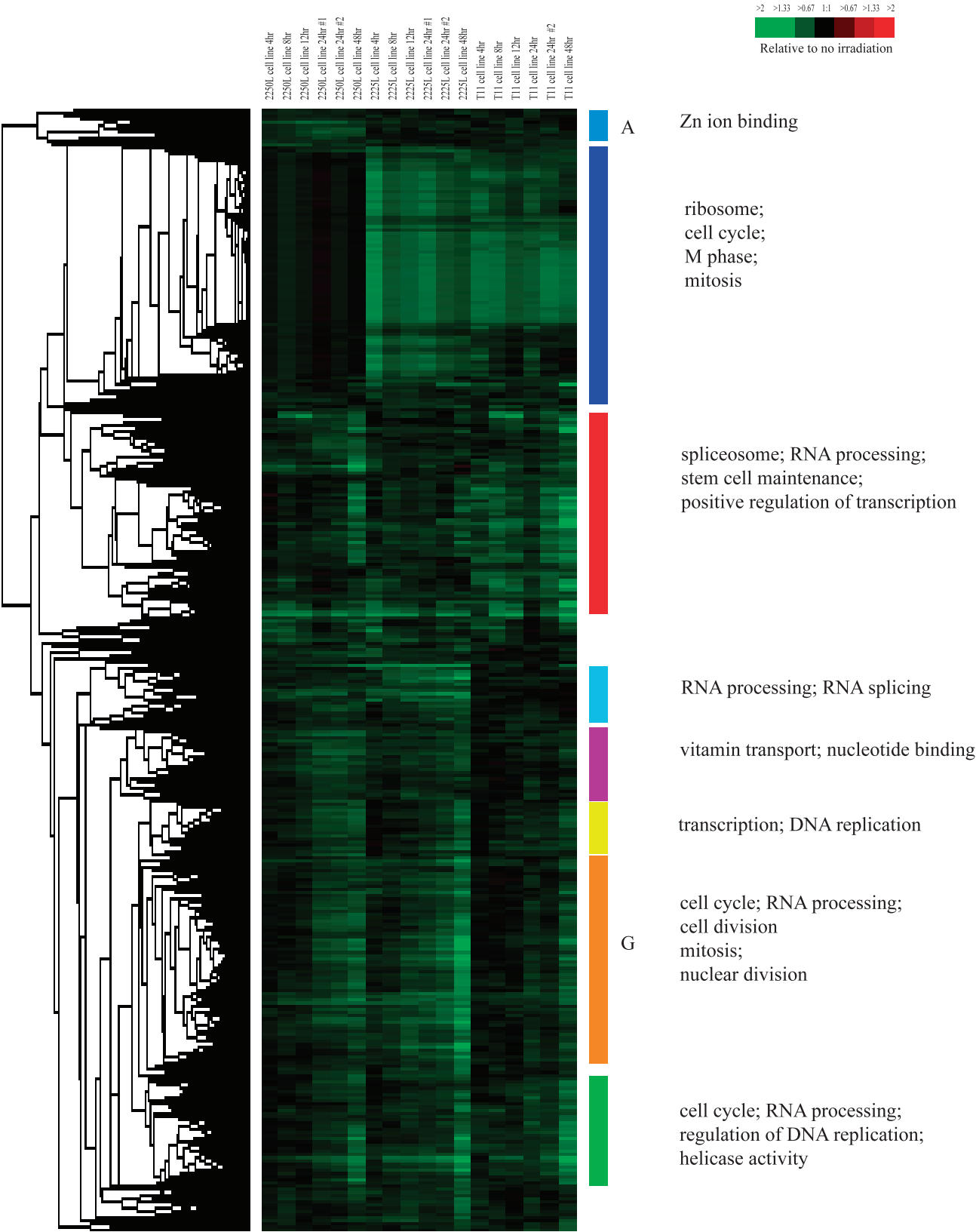
FIG. 2. Hierarchical clustering analysis of genes determined by one-class significance analysis of microarrays (SAM) to be radiation-repressed (2,689 genes) when testing three different mouse mammary tumor cell lines. Colored bars and/or letters indicate various gene clusters as discussed in the text. To the left of the diagram are the gene ontology categories significantly enriched in the corresponding gene clusters indicated by the colored bars and/or letters.

Differential Response to Radiation among the Cell Lines

To identify genes that were differentially expressed after irradiation between the T11, 2225L and 2250L cell lines, we used a multi-class SAM analysis with each cell line as its own class. Using an FDR of 0%, SAM identified 1,289 genes differentially expressed after irradiation between the cell lines (hierarchical clustering of these genes is shown in Fig. 3), indicating that while the cell lines do share common genes that respond similarly to radiation (as discussed above and shown in Figs. 1 and 2), there are also clear differences in the cell lines' radiation response. These data suggest that the different intrinsic subtypes of breast cancer may have differences in their response to radiation.

Figure 3 shows that the genes found by SAM to be differentially expressed after irradiation between the cell lines fall into distinct gene clusters that differ in terms of how they respond across the three cell lines. For example, cluster A in Fig. 3 (Supplementary Fig. S1; http://dx.doi.org/10.1667/RR13485.1.S1) consists of genes that are repressed in the 2225L basal-like cell line, but that are induced in the other 2 cell lines in response to radiation. Gene ontology categories significantly enriched in this cluster according to DAVID include ''response to wounding'' ($P = 0.005$) and ''inflammatory response'' ($P = 0.02$). This gene cluster was also significantly enriched for members of the KEGG pathway ''MAPK signaling'' ($P = 0.01$). Cluster B in Fig. 3 (Supplementary Fig. 2; http://dx.doi.org/10.1667/RR13485.1.S1) consists of genes observed to be repressed in the T11 claudin-low and 2225L basal-like cell lines, while less repressed in the 2250L luminal cell line in response to radiation. Gene ontology categories significantly enriched in this cluster include ''M phase'' ($P = 0.006$), ''mitotic cell cycle'' ($P = 0.01$) and ''ribosome'' ($P = 1.25 \times 10^{-4}$). Cluster C in Fig. 3 (Supplementary Fig. 3; http://dx.doi.org/10.1667/RR13485.1.S1) consists of genes repressed in the 2250L and 2225L cell lines while less repressed or slightly induced in the T11 cell line in response to radiation; gene ontology categories significantly enriched in this cluster include ''myofibril assembly'' ($P = 0.002$), ''muscle contraction'' ($P = 0.02$) and ''muscle cell development'' ($P = 0.02$). Cluster E in Fig. 3 (Supplementary Fig. 5; http://dx.doi.org/10.1667/RR13485.1.S1) consists of genes induced in the 2250L and 2225L cell lines while repressed or less induced in the T11 cell line in response to radiation; gene ontology categories significantly enriched in this cluster include ''zinc ion binding'' ($P = 5.76 \times 10^{-4}$) and ''DNA binding'' ($P = 0.006$).

Cluster D in Fig. 3 (Supplementary Fig. 4; http://dx.doi.org/10.1667/RR13485.1.S1) consists of genes induced in the 2225L basal-like cell line while repressed or less induced in the other 2 cell lines in response to radiation. Gene ontology categories significantly enriched in this cluster include ''blood vessel development'' ($P = 0.003$), ''positive regulation of mesenchymal cell proliferation'' ($P = 0.008$) and ''cell motion'' ($P = 0.02$). Interestingly this

gene cluster contained the genes SESN3 and TCF4 found to be regulated by the transcription factor YBX1 by Evdokimova *et al.* (*37*). In the experiments done by Evdokimova *et al.* the enforced expression of YBX1 in noninvasive breast epithelial cells was found to directly activate translation of mRNAs encoding Snail1 and other transcription factors such as SESN3 and TCF4 that are implicated in activation of mesenchymal genes, resulting in induction of an epithelial-mesenchymal transition (EMT) accompanied by enhanced metastatic potential. Other genes in cluster D in Fig. 3 such as DDR1, FLT1, NOTCH3 and PDZRN3 have been identified as candidate YBX1-regulated genes through chromatin immunoprecipitation (ChiP)-on-chip analysis performed by Finkbeiner *et al.* (*38*). Also included in cluster D is the gene STAT3, the prosurvival pathway of which has been shown to be engaged by YBX1 to protect cells from apoptosis (*39*). Another study has suggested that STAT3 may indirectly activate transcription of YBX1 by activating transcription of TWIST, which in turn activates transcription of YBX1 (*40*). In summary, our multiclass SAM analysis has identified genes differentially expressed after irradiation between the cell lines, suggesting that differences in radiation response exist within the various breast cancer intrinsic subtypes and further, suggests possible biological bases for these differences in radiation response.

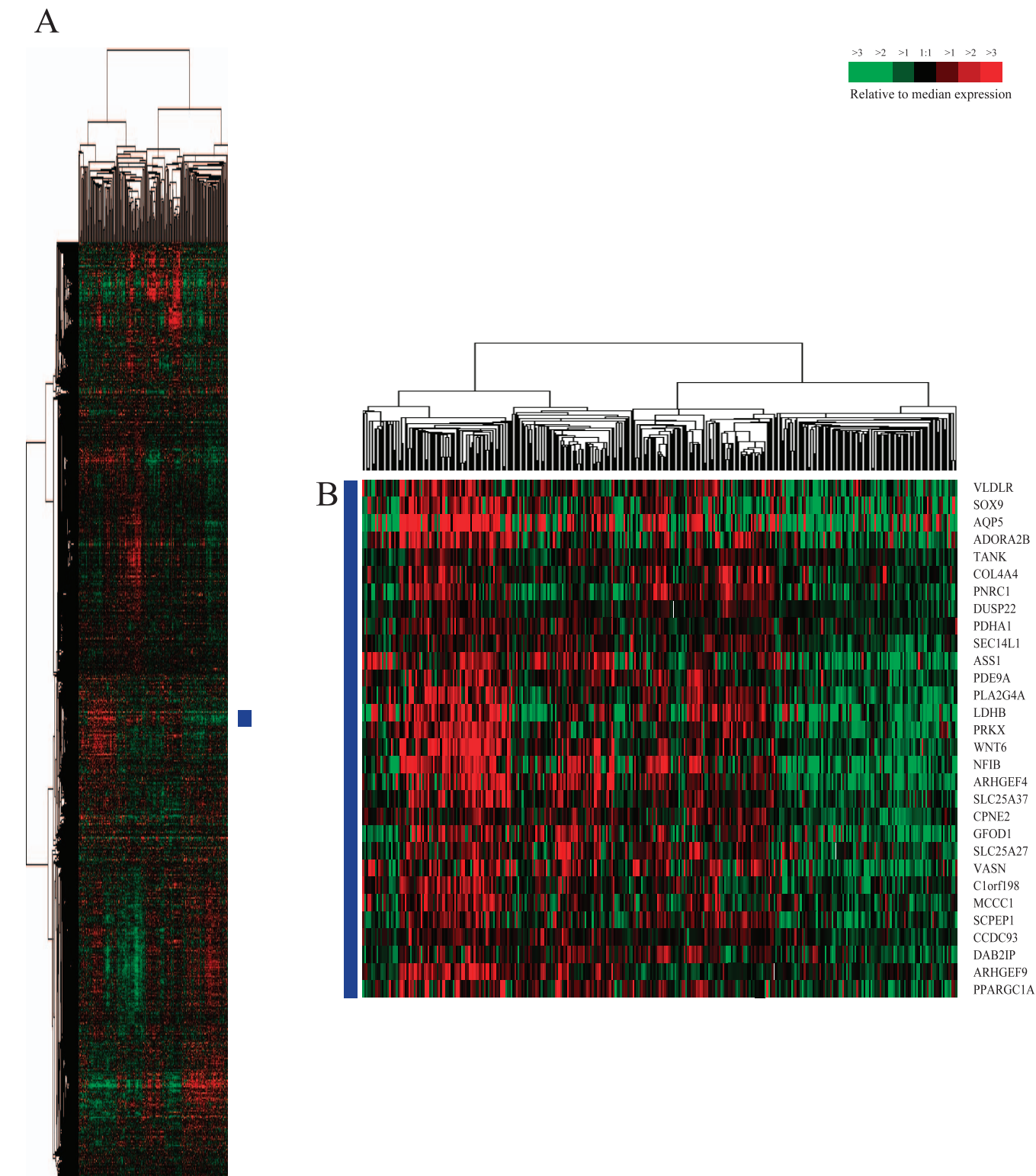*Analysis of Human Breast Tumors Using the Radiation-Induced Gene Set*

We hypothesized that due to similarities between the cellular response mechanisms to radiation and chemotherapy, expression differences of radiation-induced genes may be useful in predicting response to neoadjuvant chemotherapy. To test this hypothesis, first we hierarchically clustered a training set of 337 human primary breast samples (consisting of 320 breast tumor samples and 17 normal breast samples) published by Prat *et al.* (*23*), using the murine radiation-induced gene set of 4,278 genes (shown in Fig. 1) converted into their human orthologs (complete cluster diagram shown in Fig. 4A); this was done to identify gene sets/clusters composed of genes that are highly concordantly expressed within human breast samples.

We hypothesized that the mean expression of one or more of these gene sets/clusters may be useful in predicting an individual tumor's response to neoadjuvant chemotherapy. To avoid spurious results that may occur with small gene clusters, or poorly correlated gene clusters, we limited analysis to the most correlated gene clusters (node correlation $>=0.45$) and those nodes that were greater than 20 genes in size. This identified 23 non-overlapping gene clusters. We then tested the ability of the mean expression of each of these 23 gene sets to predict response to neoadjuvant chemotherapy by using a dataset published by Hatzis *et al.* (*24*), which consisted of microarray data from 473 HER2-negative breast cancer tumors taken prior to
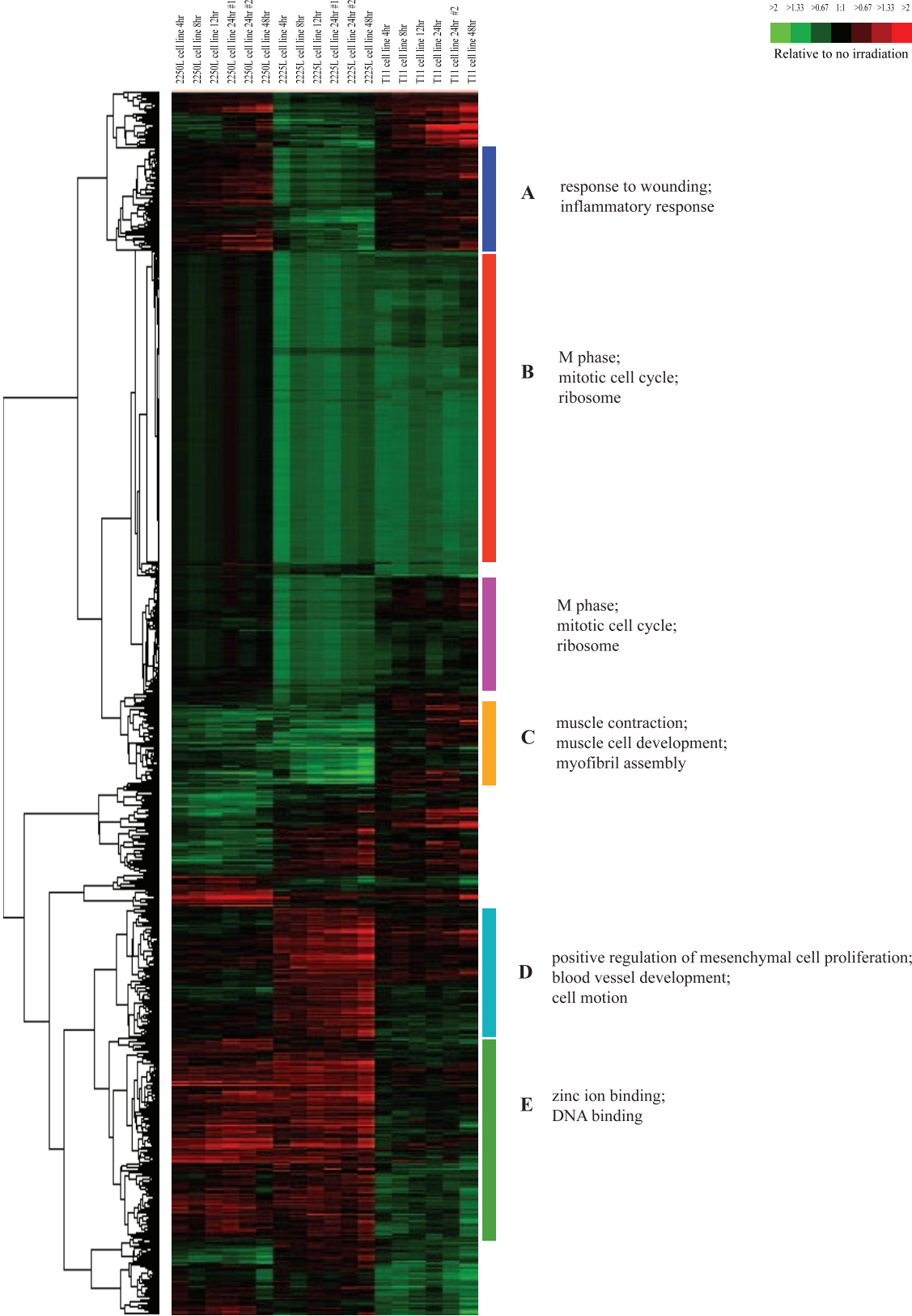
FIG. 3. Hierarchical clustering analysis of genes determined by multi-class significance analysis of microarrays (SAM) to be differentially expressed across the three murine-derived mammary tumor cell lines T11, 2225L and 2250L in response to radiation (1,289 genes). Colored bars and/or letters indicate various gene clusters as discussed in this article. To the left of the diagram are the gene ontology categories significantly enriched in the corresponding gene clusters indicated by the colored bars and/or letters.

FIG. 4. Hierarchical cluster analysis of the Prat *et al.* (*49*) 337 primary human breast sample dataset (consisting of 320 breast tumor samples and 17 normal breast samples) using the radiation-induced murine gene list converted into human orthologous genes. Panel A: Scaled down representation of the complete cluster diagram. Panel B: 30 gene set found to be significant in predicting pCR and survival.

TABLE 1
Radiation-Induced 30 Gene Signature List, Which Was Predictive of Pathologic Complete
Response (pCR) to Neoadjuvant Chemotherapy in Breast Cancer Patients

Gene symbol	Gene name	Entrez gene ID
ADORA2B	adenosine A2b receptor	136
AQP5	aquaporin 5	362
ARHGEF4	Rho guanine nucleotide exchange factor (GEF) 4	50649
ARHGEF9	Cdc42 guanine nucleotide exchange factor (GEF) 9	23229
ASS1	argininosuccinate synthase 1	445
C1orf198	chromosome 1 open reading frame 198	84886
CCDC93	coiled-coil domain containing 93	54520
COL4A4	collagen, type IV, alpha 4	1286
CPNE2	copine II	221184
DAB2IP	DAB2 interacting protein	153090
DUSP22	dual specificity phosphatase 22	56940
GFOD1	glucose-fructose oxidoreductase domain containing 1	54438
LDHB	lactate dehydrogenase B	3945
MCCC1	methylcrotonoyl-Coenzyme A carboxylase 1 (alpha)	56922
NFIB	nuclear factor I/B	4781
PDE9A	phosphodiesterase 9A	5152
PDHA1	pyruvate dehydrogenase (lipoamide) alpha 1	5160
PLA2G4A	phospholipase A2, group IVA (cytosolic, calcium-dependent)	5321
PNRC1	proline-rich nuclear receptor coactivator 1	10957
PPARGC1A	peroxisome proliferator-activated receptor gamma, coactivator 1 alpha	10891
PRKX	protein kinase, X-linked	5613
SCPEP1	serine carboxypeptidase 1	59342
SEC14L1	SEC14-like 1 (S. cerevisiae)	6397
SLC25A27	solute carrier family 25, member 27	9481
SLC25A37	solute carrier family 25, member 37	51312
SOX9	SRY (sex determining region Y)-box 9	6662
TANK	TRAF family member-associated NFKB activator	10010
VASN	vasorin	114990
VLDLR	very low density lipoprotein receptor	7436
WNT6	wingless-type MMTV integration site family, member 6	7475

sequential taxane and anthracycline-based neoadjuvant chemotherapy; this dataset had associated pCR and distant relapse-free survival (DRFS) data. Using univariate logistic regression modeling, only 4/23 of the tested gene clusters had a significant Bonferroni-corrected P value in predicting pCR. To determine which of these 4 gene clusters could be useful in predicting pCR, we then performed multivariate logistic regression modeling with each of these 4 gene sets in a model that included the standard clinical variables and other gene expression based predictors including those used by Hatzis et al. (24, 41). Only 1 gene cluster (consisting of 30 genes, Fig. 4B and Table 1) was significant in multivariate analysis (Table 3). As seen in Tables 2 and 3, a higher mean expression of this 30 gene set was associated with a higher likelihood of pCR. As shown in Table 2B, this 30 gene set was not only able to significantly predict pCR in univariate logistic regression modeling on the entire Hatzis et al. patient dataset, but was also able to significantly predict pCR within individual clinical subsets of patients, including the clinically relevant triple-negative subset and the more biologically relevant basal-like patient subset.

We next tested the ability of the 30 gene cluster to predict pCR on a second and completely independent test dataset published by Gluck et al. (25), which consisted of microarray data from 95 HER2-negative breast cancer tumors taken prior to neoadjuvant capecitabine and docetaxel chemotherapy. This dataset had associated pCR data and TP53 mutation status data using the AmpliChip TP53 assay. As shown in Table 4, the 30 gene set significantly predicted pCR in univariate logistic regression modeling on the entire dataset as well as on the biologically relevant basal-like patient subset. In multivariable logistic regression modeling (Table 5) that included the standard clinical parameters, intrinsic subtype and TP53 mutation status, the 30 gene set again remained a statistically significant predictor of pCR.

As both the Prat et al. and Hatzis et al. datasets had associated survival data, we also tested the prognostic value of the 30 gene set to predict survival. To achieve this goal, we grouped patients from each dataset into halves or tertiles based on rank order mean expression value of the 30 gene set. Kaplan-Meier analysis showed significant differences among patients when divided into halves or tertiles for distant relapse-free survival (DRFS) in the Hatzis et al. dataset ($P = 2.48 \times 10^{-7}$ and 0.0002, respectively), and for overall survival (OS; $P = 1.45 \times 10^{-6}$ and 1.18×10^{-7}, respectively) and relapse-free survival (RFS; $P = 2.08 \times 10^{-5}$ and 7.95×10^{-5}, respectively) in the Prat et al. dataset.

TABLE 2

Univariate Logistic Regression Analysis to Predict Pathologic Complete Response (pCR) Using the Radiation-Induced 30 Gene Signature for (Part A) All Patients and (Part B) Select Patient Subsets of the Hatzis *et al*. (*24*) Dataset

Part A	Number of patients	Number with pCR (%)	Odds ratio[a] (95% CI) of pCR vs. no pCR	*P*	AUC
Patient set					
All patients	473	96 (20%)	**8.313 (4.64–15.47)**	**<0.0001**	0.735
Part B					
Patient subset					
ER–	192	66 (34%)	**8.86 (2.98–29.78)**	**<0.0001**	0.687
ER+	280	30 (11%)	**3.47 (1.17–10.16)**	**0.0258**	0.593
Triple negative	167	56 (32%)	**8.16 (2.51–30.20)**	**0.0003**	0.684
Basal-like subtype[b]	133	54 (41%)	**7.51 (2.10–31.07)**	**0.0016**	0.672
Claudin-low subtype	73	17 (23%)	11.90 (0.83–271.22)	0.089	0.658

Notes. Bold numbers indicate statistical significance (*P* < 0.05). AUC = area under curve from receiver operating characteristic (ROC) curve for logistic regression model, pCR= pathological complete response, CI = confidence interval.

[a] Odds ratio defined for per unit increase in regressor variable (i.e., mean expression value of radiation-induced 30 gene set, defined as a continuous variable).

[b] Intrinsic subtype defined using PAM50 predictor (*43*).

In the Hatzis *et al*. dataset, patients with tumors of the basal-like subtype, when grouped into halves or tertiles based on rank order mean expression value of the 30 gene set, showed significant differences in DRFS (Fig. 5), with patients having higher expression of the 30 gene set having better DRFS.

Gene ontology analysis of the 30 gene set using DAVID revealed significant enrichment for the following categories:

TABLE 3

Multivariate Logistic Regression Analysis to Predict Pathologic Complete Response for the Hatzis *et al*. (*24*) Dataset Using Multiple Possible Predictive Factors Including the Radiation-Induced 30 Gene Signature, the Standard Clinical Parameters and Additional Genomic Biomarkers

Variable	Odds ratio (95% CI) of pCR vs. no pCR	*P*
Radiation-induced 30 gene set[a]	**4.52 (1.46–14.68)**	**0.0087**
SET index: High vs. Low [Symmans *et al*. (*41*)]	0.67 (0.07–4.47)	0.69
SET index: Int vs. Low [Symmans *et al*. (*41*)]	1.14 (0.21–4.84)	0.86
Chemosensitivity predictor (Hatzis *et al*. (*24*)): Rx sensitive vs. insensitive	**2.45 (1.23–5.01)**	**0.010**
Excellent pathologic response predictor [(Hatzis *et al*. (*24*)]: RCB–0/I vs. RCB-II/III	**6.85 (3.26–15.27)**	**<0.0001**
ER pos vs. neg	0.83 (0.33–2.03)	0.67
PR pos vs. neg	1.13 (0.5–2.61)	0.76
HER2 pos vs. neg	2.21 (0.40–12.59)	0.36
Histologic grade 2 vs. 1	0.80 (0.12–16.0)	0.85
Histologic grade 3 vs. 1	1.61 (0.23–32.7)	0.66
Clinical T stage 2 vs. 1	1.20 (0.10–27.9)	0.89
Clinical T stage 3 vs. 1	1.04 (0.08–24.3)	0.97
Clinical T stage 4 vs. 1	0.62 (0.05–14.9)	0.72
Basal-like vs. LumA intrinsic subtype[b]	**4.61 (1.01–23.7)**	**0.049**
LumB vs. LumA intrinsic subtype	**5.04 (1.35–22.1)**	**0.015**
Normal vs. LumA intrinsic subtype	3.16 (0.59–16.8)	0.17
Claudin vs. LumA intrinsic subtype	4.07 (0.88–20.8)	0.073
Her2 vs. LumA intrinsic subtype	2.80 (0.50–15.8)	0.23

Notes. Bold numbers indicate variables found to be significant (*P* < 0.05) in the logistic regression model. pCR = pathological complete response, CI = confidence interval, RCB–0/I = pCR or minimal residual cancer burden, defining excellent response. RCB–II/III = moderate or extensive residual cancer burden, defining lesser response.

[a] Odds ratio defined for per unit increase in regressor variable (i.e., mean expression value of radiation-induced 30 gene set, defined as a continuous variable).

[b] Intrinsic subtype defined using PAM50 predictor (*43*).

TABLE 4
Univariate Logistic Regression Analysis to Predict Pathologic Complete Response (pCR) Using the Radiation-Induced 30 Gene Signature for (Part A) All Patients, and (Part B) Basal-like Patient Subset of the Gluck et al. (22) Dataset

Part A	Number of patients	Number with pCR (%)	Odds ratio[a] (95% CI) of pCR vs. no pCR	P	AUC
Patient set					
All patients	95	10 (10%)	**14.82 (2.00–180.8)**	**0.0071**	0.738
Part B					
Patient subset					
Basal-like subtype[b]	28	6 (21%)	**25577 (7.82–2.8 × 10⁹)**	**0.01**	0.82

Notes. Bold numbers indicate statistical significance ($P < 0.05$). AUC = area under curve from receiver operating characteristic (ROC) curve for logistic regression model, pCR = pathological complete response, CI = confidence interval.

[a] Odds ratio defined for per unit increase in regressor variable (i.e., mean expression value of radiation-induced 30 gene set, defined as a continuous variable).

[b] Intrinsic subtype defined using PAM50 predictor (43).

''pyruvate metabolic process'' (genes LDHB, PPARGC1A, PDHA1; $P = 0.002$); ''response to extracellular stimulus'' (genes ADORA2B, PPARGC1A, PLA2G4A, VLDLR; $P = 0.007$); and ''generation of precursor metabolites and energy'' (genes LDHB, PPARGC1A, PDHA1, SLC25A27; $P = 0.017$). Interestingly, our 30 gene set was also significantly enriched for members of the KEGG pathway ''vascular smooth muscle contraction'' (genes PLA2G4A, PRKX, ADORA2B; $P = 0.027$). When it was compared with other gene expression-based predictors, only one common gene (NFIB) was found between our 30 gene predictor and the predictors published by Hatzis et al. (24). There was no overlap with either the 11-gene proliferation signature developed by Parker et al. (42), the PAM50 gene signature used by Gluck et al. or Prat et al. (23, 25, 43) or

the 21 gene recurrence score assay developed by Paik et al. (44), which is known as Oncotype DX.

DISCUSSION

The search for biomarkers that predict pCR to neoadjuvant chemotherapy for breast cancer has been an area of intense research in recent years. pCR is an important clinical end point, as patients who achieve pCR to neoadjuvant chemotherapy are more likely to have improved overall and disease-free survival (3). Gene expression-based analyses have contributed to this research, with the development of several predictive gene sets and assays based on selection of genes that directly correlate with patient/tumor outcomes (24, 45, 46). We took a different approach and selected genes using no knowledge of survival outcomes or response

TABLE 5
Multivariate Logistic Regression Analysis to Predict Pathologic Complete Response for the Gluck et al. (22) Dataset Using Multiple Possible Predictive Factors Including the Radiation-Induced 30 Gene Signature, the Standard Clinical Parameters and Other Possible Biomarkers

Variable	Odds ratio (95% CI) of pCR vs. no pCR	P
Radiation-induced 30 gene set[a]	**2167 (2.41–4.4 × 10⁷)**	**0.02**
ER pos vs. neg	0.07	0.33
PR pos vs. neg	**4.4 × 10⁻⁹**	**0.02**
Histologic grade 2 vs. 1	0.83	0.93
Histologic grade 3 vs. 1	1.27	0.90
Clinical T stage 2 vs. 1	0.10	0.26
Clinical T stage 3 vs. 1	**9 × 10⁻¹¹**	**0.004**
Basal-like vs. LumA intrinsic subtype[b]	0.0006	0.11
LumB vs. LumA intrinsic subtype	6.4 × 10⁻⁸	0.99
Normal vs. LumA intrinsic subtype	**1.4 × 10⁻¹¹**	**0.01**
Her2 vs. LumA intrinsic subtype	0.01	0.25
TP53 mutation by AmpliChip assay: Mutant vs. wild-type	0.28	0.44

Notes. Bold numbers indicate variables found to be significant ($P < 0.05$) in the logistic regression model.
[a] Odds ratio defined for per unit increase in regressor variable (i.e., mean expression value of radiation-induced 30 gene set, defined as a continuous variable). [b] Intrinsic subtype defined using PAM50 predictor (43).
Abbreviations: pCR = pathological complete response; CI = confidence interval.
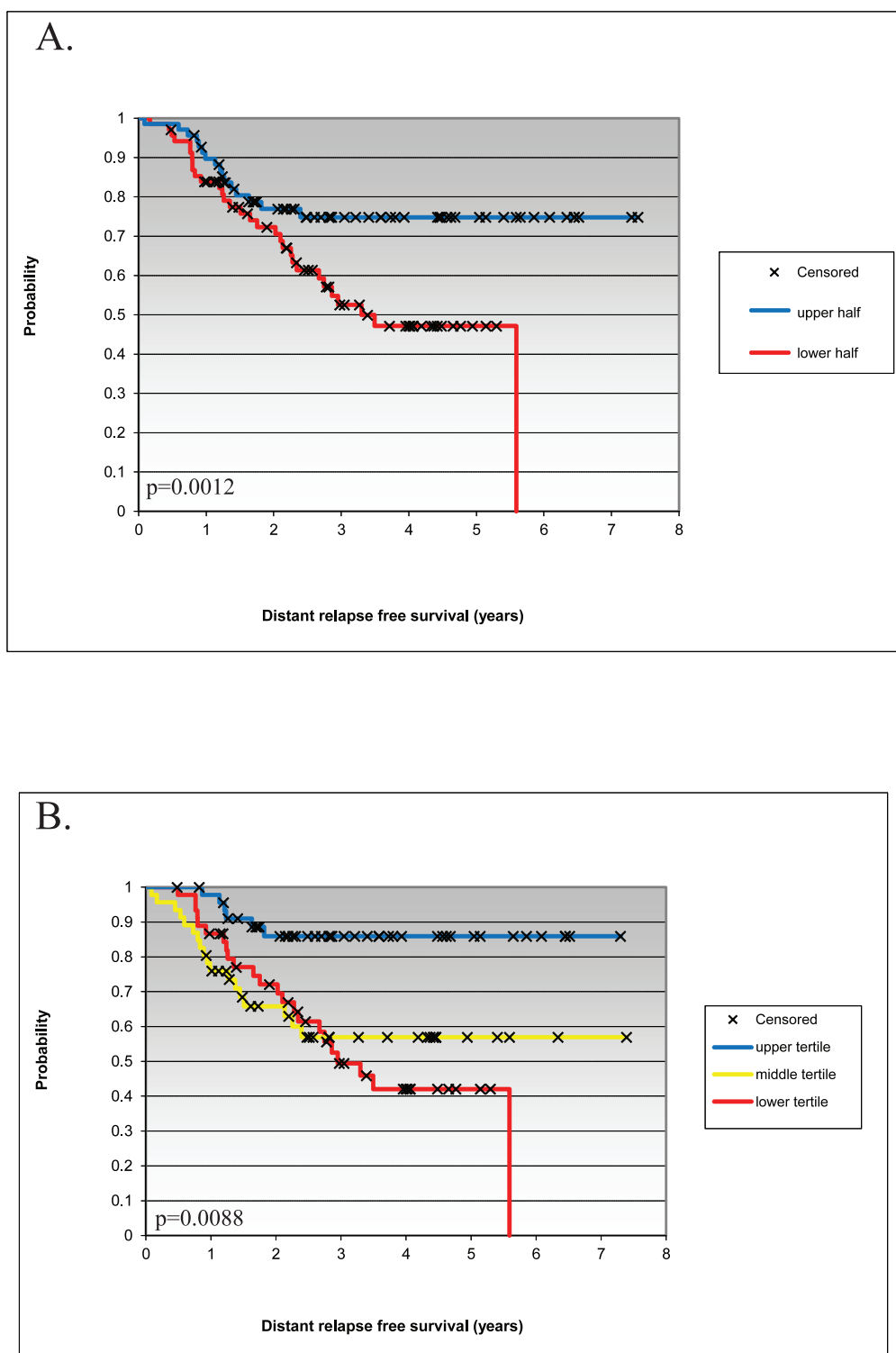
FIG. 5. Kaplan-Meier survival curves of patients with basal-like tumors from the Hatzis *et al.* dataset (*24*) divided into (panel A) halves or (panel B) tertiles according to rank order mean expression value of the radiation-induced 30 gene set. *P* values were calculated using the log-rank test.

to therapy; these were instead selected on the basis of radiation induction and their natural patterns of expression in primary breast tumors. We objectively identified gene sets/clusters composed of radiation-induced genes that were

highly concordantly expressed in human breast tumor samples. We then systematically tested the ability of each of these gene sets to predict pCR on an independent test dataset (*24*). Only one gene cluster consisting of 30 genes

was able to significantly predict pCR in both univariate analysis and multivariate analysis that included standard clinical variables and breast intrinsic subtype. In addition, this 30 gene signature was able to predict pCR in both univariate and multivariate analysis on a second independent test dataset (25), and thus it has shown utility on two different validation data sets.

As determined by gene ontology analysis, our radiation-induced 30 gene signature was significantly enriched for genes involved in pyruvate metabolism, generation of precursor metabolites and energy, suggesting that tumors most sensitive to chemotherapy are those with highest metabolic activity, consistent with the known mechanisms of action of several chemotherapeutics used in the datasets we analyzed, including doxorubicin and capecitabine (47–53). Genes composing our 30 gene signature include lactate dehydrogenase B (LDHB), which is a subunit of the lactate dehydrogenase enzyme, the key glycolytic enzyme catalyzing formation of lactic acid from pyruvate. LDHB has recently been shown to be a downstream target of mTOR critical for oncogenic mTOR-mediated tumorigenesis (54). It was also recently shown to be an essential gene in basal-like/triple-negative breast cancer metabolism (55). Proteomic analysis done by Cortesi et al. (56) showed that high expression of the LDHB protein in tumor interstitial fluid was associated with response to chemotherapy in breast cancer patients, consistent with our results. Interestingly, the downregulation of DAB2 interacting protein (DAB2IP), another member of our 30 gene set and a novel member of the Ras GTPase-activating protein family, was recently reported by Kong et al. (57) to result in resistance to ionizing radiation in prostate cancer cells. Our results suggest that the low expression of DAB2IP is also associated with resistance to chemotherapy in breast cancer.

Not only has our analysis of radiation-responsive genes yielded a useful tool for predicting pCR, it has also yielded information indicating that the different intrinsic subtypes of breast cancer may have distinct biological differences in their response to radiation. As shown in Fig. 3, the genes found to be differentially expressed after irradiation between the murine cell lines representative of the different intrinsic subtypes (claudin-low, basal-like and luminal) fall into distinct clusters. These clusters differ not only in how they respond in the cell lines to radiation but also their biological function, possibly reflecting differential activation or repression of various pathways in the different subtypes in response to radiation. To support this hypothesis, genome wide sequencing of the 2225L basal-like tumor has shown a possible YBX1 activating mutation in this tumor, but not in the T11 or 2250L tumors (unpublished results). Consistent with the possibility of an YBX1 activating mutation in the 2225L tumor, we observed induction of multiple YBX1 known and putative targets in the 2225L basal-like cell line (Fig. 3, gene cluster D; Supplementary Fig. S4: http://dx.doi.org/10.1667/RR13485.1.S1) in re-sponse to radiation, which was not seen in the other cell lines. Active/expressed YBX1 appears to be an important feature of the basal-like subtype (58–61). With our observations, this suggests an important role of this gene not only in defining the basal-like subtype but also in the radiation response of this subtype. Experiments are ongoing in our lab to test this hypothesis.

In conclusion, our radiation-induced 30 gene signature shows significant promise as a clinically relevant predictor of pCR in breast cancer patients. We have shown its efficacy in predicting pCR in two independent test datasets. Furthermore we have shown by multivariate analysis that it adds significant information in predicting pCR beyond what is provided by the standard clinical parameters, breast tumor intrinsic subtypes, p53 mutation status and other gene expression-based predictors including those published by Hatzis et al. (24). Our results are hypothesis generating, but based on the analysis presented here, we feel that our predictor warrants further investigation, including validation in prospectively designed clinical trials, to confirm the clinical validity and utility of our 30 gene signature.

SUPPLEMENTARY INFORMATION

Supplementary Fig. S1. http://dx.doi.org/10.1667/RR13485.1.S1; Gene cluster A shown Fig. 3 (upper left) significantly enriched for genes involved in wound and inflammatory response as determined by DAVID.

Supplementary Fig. S2. http://dx.doi.org/10.1667/RR13485.1.S1; Gene cluster B from Fig. 3 (upper left) significantly enriched for genes involved in M phase, mitotic cell cycle and ribosome as determined by DAVID.

Supplementary Fig. S3. http://dx.doi.org/10.1667/RR13485.1.S1; Gene cluster C from Fig. 3 (upper left) significantly enriched for genes involved in muscle contraction, muscle cell development, and myofibril assembly as determined by DAVID.

Supplementary Fig. S4. http://dx.doi.org/10.1667/RR13485.1.S1; Gene cluster D from Fig. 3 (upper left) significantly enriched for genes involved in positive regulation of mesenchymal cell proliferation, blood vessel development and cell motion as determined by DAVID.

Supplementary Fig. S5. http://dx.doi.org/10.1667/RR13485.1.S1; Gene cluster E from Fig. 3 (upper left) significantly enriched for genes involved in zinc ion and DNA binding as determined by DAVID.

REFERENCES

1. Fisher B, Bryant J, Wolmark N, Mamounas E, Brown A, Fisher ER, et al. Effect of preoperative chemotherapy on the outcome of women with operable breast cancer. J Clin Oncol 1998; 16:2672–85.

2. Kaufmann M, von Minckwitz G, Smith R, Valero V, Gianni L, Eiermann W, et al. International expert panel on the use of primary (preoperative) systemic treatment of operable breast cancer: review and recommendations. J Clin Oncol 2003; 21:2600–8.

3. Rastogi P, Anderson SJ, Bear HD, Geyer CE, Kahlenberg MS, Robidoux A, et al. Preoperative chemotherapy: updates of National Surgical Adjuvant Breast and Bowel Project Protocols B-18 and B-27. J Clin Oncol 2008; 26:778–85.

4. Cianfrocca M, Gradishar WJ. Controversies in the therapy of early stage breast cancer. Oncologist 2005; 10:766–79.

5. Kuerer HM, Newman LA, Smith TL, Ames FC, Hunt KK, Dhingra K, et al. Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy. J Clin Oncol 1999; 17:460–9.

6. Ferriere JP, Assier I, Cure H, Charrier S, Kwiatkowski F, Achard JL, et al. Primary chemotherapy in breast cancer: correlation between tumor response and patient outcome. Am J Clin Oncol 1998; 21:117–20.

7. Chang J, Powles TJ, Allred DC, Ashley SE, Clark GM, Makris A, et al. Biologic markers as predictors of clinical outcome from systemic therapy for primary operable breast cancer. J Clin Oncol 1999; 17:3058–63.

8. Pierga JY, Mouret E, Laurence V, Dieras V, Savigioni A, Beuzeboc P, et al. Prognostic factors for survival after neoadjuvant chemotherapy in operable breast cancer. the role of clinical response. Eur J Cancer 2003; 39:1089–96.

9. Cleator SJ, Makris A, Ashley SE, Lal R, Powles TJ. Good clinical response of breast cancers to neoadjuvant chemoendocrine therapy is associated with improved overall survival. Ann Oncol 2005; 16:267–72.

10. Rouzier R, Pusztai L, Delaloge S, Gonzalez-Angulo AM, Andre F, Hess KR, et al. Nomograms to predict pathologic complete response and metastasis-free survival after preoperative chemotherapy for breast cancer. J Clin Oncol 2005; 23:8331–9.

11. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 2005; 11:5678–85.

12. Luo Y, Leverson JD. New opportunities in chemosensitization and radiosensitization: modulating the DNA-damage response. Expert Rev Anticancer Ther 2005; 5:333–42.

13. Martin NM. DNA repair inhibition and cancer therapy. J Photochem Photobiol B 2001; 63:162–70.

14. Norbury CJ, Zhivotovsky B. DNA damage-induced apoptosis. Oncogene 2004; 23:2797–808.

15. Perona R, Sanchez-Perez I. Control of oncogenesis and cancer therapy resistance. Br J Cancer 2004; 90:573–7.

16. Bold RJ, Termuhlen PM, McConkey DJ. Apoptosis, cancer and cancer therapy. Surg Oncol 1997; 6:133–42.

17. Viktorsson K, Lewensohn R, Zhivotovsky B. Apoptotic pathways and therapy resistance in human malignancies. Advances Cancer Res 2005; 94:143–96.

18. Bensimon A, Aebersold R, Shiloh Y. Beyond ATM: the protein kinase landscape of the DNA damage response. FEBS Lett 2011; 585:1625–39.

19. Herschkowitz JI, Zhao W, Zhang M, Usary J, Murrow G, Edwards D, et al. Comparative oncogenomics identifies breast tumors enriched in functional tumor-initiating cells. Proc Natl Acad Sci U S A 2012; 109:2778–83.

20. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001; 98:5116–21.

21. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95:14863–8.

22. Saldanha AJ. Java Treeview–extensible visualization of microarray data. Bioinformatics 2004; 20:3246–8.

23. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 2010; 12:R68.

24. Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. JAMA 2011; 305:1873–81.

25. Gluck S, Ross JS, Royce M, McKenna EF, Jr., Perou CM, Avisar E, et al. TP53 genomics predict higher clinical and pathologic tumor response in operable early-stage breast cancer treated with docetaxel-capecitabine +/- trastuzumab. Breast Cancer Res Treat 2012; 132:781–91.

26. Piening BD, Wang P, Subramanian A, Paulovich AG. A radiation-derived gene expression signature predicts clinical outcome for breast cancer patients. Radiat Res 2009; 171:141–54.

27. Tsai MH, Cook JA, Chandramouli GV, DeGraff W, Yan H, Zhao S, et al. Gene expression profiling of breast, prostate, and glioma cells following single versus fractionated doses of radiation. Cancer Res 2007; 67:3845–52.

28. Amundson SA, Do KT, Vinikoor LC, Lee RA, Koch-Paiz CA, Ahn J, et al. Integrating global gene expression and radiation survival parameters across the 60 cell lines of the National Cancer Institute Anticancer Drug Screen. Cancer Res 2008; 68:415–24.

29. Kabacik S, Mackay A, Tamber N, Manning G, Finnon P, Paillier F, et al. Gene expression following ionising radiation: identification of biomarkers for dose estimation and prediction of individual response. Int J Radiat Biol 2011; 87:115–29.

30. 30. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4:44–57.

31. 31. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009; 37:1–13.

32. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 2012; 40 (Database issue):D109–14.

33. Bourdon JC, Renzing J, Robertson PL, Fernandes KN, Lane DP. Scotin, a novel p53-inducible proapoptotic protein located in the ER and the nuclear membrane. J Cell Biol 2002; 158:235–46.

34. Budanov AV, Lee JH, Karin M. Stressin' Sestrins take an aging fight. EMBO Mol Med 2010; 2:388–400.

35. Budanov AV, Sablina AA, Feinstein E, Koonin EV, Chumakov PM. Regeneration of peroxiredoxins by p53-regulated sestrins, homologs of bacterial AhpD. Science 2004; 304:596–600.

36. Terrinoni A, Ranalli M, Cadot B, Leta A, Bagetta G, Vousden KH, et al. p73-alpha is capable of inducing scotin and ER stress. Oncogene 2004; 23:3721–5.

37. Evdokimova V, Tognon C, Ng T, Ruzanov P, Melnyk N, Fink D, et al. Translational activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition. Cancer Cell 2009; 15:402–15.

38. Finkbeiner MR, Astanehe A, To K, Fotovati A, Davies AH, Zhao Y, et al. Profiling YB-1 target genes uncovers a new mechanism for MET receptor regulation in normal and malignant human mammary cells. Oncogene 2009; 28:1421–31.

39. Lee C, Dhillon J, Wang MY, Gao Y, Hu K, Park E, et al. Targeting YB-1 in HER-2 overexpressing breast cancer cells induces apoptosis via the mTOR/STAT3 pathway and suppresses tumor growth in mice. Cancer Res 2008; 68:8661–6.

40. Kalra J, Sutherland BW, Stratford AL, Dragowska W, Gelmon KA, Dedhar S, et al. Suppression of Her2/neu expression through ILK inhibition is regulated by a pathway involving TWIST and YB-1. Oncogene 2010; 29:6343–56.

41. Symmans WF, Hatzis C, Sotiriou C, Andre F, Peintinger F, Regitnig P, et al. Genomic index of sensitivity to endocrine therapy for breast cancer. J Clin Oncol 2010; 28:4111–9.

42. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res 2010; 16:5222–32.

43. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009; 27:1160–7.

44. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. The New Engl J Med 2004; 351:2817–26.

45. Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. J Clin Oncol 2006; 24:4236–44.

46. Lin Y, Lin S, Watson M, Trinkaus KM, Kuo S, Naughton MJ, et al. A gene expression signature that predicts the therapeutic response of the basal-like breast cancer to neoadjuvant chemotherapy. Breast Cancer Res Treat 2010; 123:691–9.

47. Momparler RL, Karon M, Siegel SE, Avila F. Effect of adriamycin on DNA, RNA, and protein synthesis in cell-free systems and intact cells. Cancer Res 1976; 36:2891–5.

48. Walko CM, Lindley C. Capecitabine: a review. Clin Ther 2005; 27:23–44.

49. Longley DB, Harkin DP, Johnston PG. 5-fluorouracil: mechanisms of action and clinical strategies. Nat Rev Cancer 2003; 3:330–8.

50. Fanciulli M, Bruno T, Cerboni C, Bonetto F, Iacobini C, Frati L, et al. Cloning of a novel human RNA polymerase II subunit downregulated by doxorubicin: new potential mechanisms of drug related toxicity. FEBS Lett 1996; 384:48–52.

51. Chuang RY, Chuang LF. Inhibition of chicken myeloblastosis RNA polymerase II activity by adriamycin. Biochemistry 1979; 18:2069–73.

52. Aubel-Sadron G, Londos-Gagliardi D. Daunorubicin and doxorubicin, anthracycline antibiotics, a physicochemical and biological review. Biochimie 1984; 66:333–52.

53. Barranco SC, Novak JK. Survival responses of dividing and nondividing mammalian cells after treatment with hydroxyurea, arabinosylcytosine, or adriamycin. Cancer Res 1974; 34:1616–8.

54. Zha X, Wang F, Wang Y, He S, Jing Y, Wu X, et al. Lactate dehydrogenase B is critical for hyperactive mTOR-mediated tumorigenesis. Cancer Res 2011; 71:13–8.

55. McCleland ML, Adler AS, Shang Y, Hunsaker T, Truong T, Peterson D, et al. An Integrated Genomic Screen Identifies LDHB as an Essential Gene for Triple-Negative Breast Cancer. Cancer Res 2012; 72:5812–23.

56. Cortesi L, Barchetti A, De Matteis E, Rossi E, Della Casa L, Marcheselli L, et al. Identification of protein clusters predictive of response to chemotherapy in breast cancer patients. J Proteome Res 2009; 8:4916–33.

57. Kong Z, Xie D, Boike T, Raghavan P, Burma S, Chen DJ, et al. Downregulation of human DAB2IP gene expression in prostate cancer cells results in resistance to ionizing radiation. Cancer Res 2010; 70:2829–39.

58. Habibi G, Leung S, Law JH, Gelmon K, Masoudi H, Turbin D, et al. Redefining prognostic factors for breast cancer: YB-1 is a stronger predictor of relapse and disease-specific survival than estrogen receptor or HER-2 across all tumor subtypes. Breast Cancer Res 2008; 10:R86.

59. Stratford AL, Habibi G, Astanehe A, Jiang H, Hu K, Park E, et al. Epidermal growth factor receptor (EGFR) is transcriptionally induced by the Y-box binding protein-1 (YB-1) and can be inhibited with Iressa in basal-like breast cancer, providing a potential target for therapy. Breast Cancer Res 2007; 9:R61.

60. Stratford AL, Fry CJ, Desilets C, Davies AH, Cho YY, Li Y, et al. Y-box binding protein-1 serine 102 is a downstream target of p90 ribosomal S6 kinase in basal-like breast cancer cells. Breast Cancer Res 2008; 10:R99.

61. Larsson O, Wahlestedt C, Timmons JA. Considerations when using the significance analysis of microarrays (SAM) algorithm. BMC Bioinformatics 2005; 6:129.